PN-ABF-128
ISN 67403

B R I D G E S

BASIC RESEARCH AND IMPLEMENTATION
IN DEVELOPING EDUCATION SYSTEMS

RESEARCH TRIANGLE INSTITUTE

A USER'S GUIDE TO THE PRIMARY EDUCATION TRACKING SYSTEM

(PETS)

• JANUARY 1987 •

DEVELOPED BY THE RESEARCH TRIANGLE INSTITUTE
WITH
HARVARD UNIVERSITY
UNDER THE
BASIC RESEARCH AND IMPLEMENTATION
IN DEVELOPING EDUCATION SYSTEMS PROJECT
(BRIDGES)

Sponsored by the Office of Education, Bureau for Science and Technology,
and
Office of Development Resources, Bureau for Latin America and the Caribbean
United States Agency for International Development

POST OFFICE BOX 12194 RESEARCH TRIANGLE PARK, NORTH CAROLINA 27709

A USER'S GUIDE TO THE PRIMARY EDUCATION TRACKING SYSTEM

(PETS)

• JANUARY 1987 •

Developed by:
Luis Crouch and Scott Moreland
Research Triangle Institute

with:
Ernesto Cuadra and Noel McGinn
Harvard University

Sponsored by the Office of Education, Bureau for Science and Technology,
and
Office of Development Resources, Bureau for Latin America and the Caribbean
United States Agency for International Development

Table of Contents

CHAPTER 1. INTRODUCTION TO THE PETS SYSTEM 1
Introduction to the PETS System 1
PETS Modules 1
Using PETS on the Microcomputer 1
Hardware Environment 1
Software Environment 2
Organization of this Manual 2
CHAPTER 2. USING PETS 3
File Management Under PETS 3
PETS Modules: ESTIMATE 4
Description of the ESTIMATE Module 4
Use of the ESTIMATE Module 6
PETS Modules: PROJRATE 7
Description of the PROJRATE Module 7
Use of the PROJRATE Module 10
PETS Modules: EDPROJ 10
Description of the EDPROJ Module 10
Use of the EDPROJ Module 11
PETS Modules: COHORT 11
Description of the COHORT Module 11
Use of the COHORT Module 12
CHAPTER 3. USING HOST 14
Introduction 14
How Host is Used with PETS 14
What Host is and What it Does 14
Getting Started 15
Entering Host 15
How Host Uses Cursors 15
Using the Keyboard 16

Using the Menus 16
User Versus Manager Access 17
Using the Query Field Facility 17
Using the Toggle Field Facility 18
Editing Data (E) 18
Selecting a Variable 18
The Data Entry Cursor 18
If Data are not Found in the Data File 19
Changing Editor Supplied Default Values 19
Saving Editor Supplied Default Values 19
Variables With One or More Dimensions 19
Changing the Levels of the Row and Column Dimensions 20
Changing the Levels of the By-dimensions (Dimension) 20
Using Function Keys to Change the Period 20
The Edit Menu 21
Running a Module (R) 22
Specifying the Data File 23
Specifying the Starting Period 23
Specifying the Number of Periods 23
Executing the Application Module 23
Using Results: Generating Tables (T) 24
Specifying the Data File 24
Specifying the Comparison File 24
Specifying the Output File 25
Specifying the Period List 25
Managing Files (M) 26
Managing Data Files 26
Using the Management Utilities to Obtain PETS Documentation 27
Appendix A to Chapter 3: Installation 28
Required and Recommended Hardware 28
Copying Host to a Hard Disk or Cartridge Disk 29
Creating or Modifying a CONFIG.SYS File 30
Telling DOS How to Find Host 31
Appendix B to Chapter 3: Configuring Host (C) 32

Default Files 32
Default Output Devices 33
Default Directories 33
Default Printer Settings 34
Default Color Setting 35
Saving the Configuration Settings 35
Returning to the Main Menu 35
CHAPTER 4. TECHNICAL DESCRIPTION OF THE MODEL METHODOLOGY 36
The Grade-transition Enrolment Projection Method 36
Projections with Fixed Parameters 36
Projections with Variable Parameters 37
Linear Trend 37
Logistic Curve Fitting 37
Target Setting Method 38
Efficiency Indicators Based on Cohort Analysis 38
Basic Concept 38
Efficiency and Other Educational Indicators 39
Completion or Graduation Rate 39
"On Time" Graduates 39
Student/Years per graduate 39
Input-output ratio 39
Years Spent in School 40
Indicators Based on Enrolment Data 40
Gross Enrolment Ratio 40
Net Enrolment Ratio 40
Estimation of Transition Rates 40
Standard Method 40
Schiefelbein Method - Description 41
Schiefelbein Method - Example of Calculations 45
Appendix to Chapter 4 - Glossary of variables and symbols 47
CHAPTER 5. A STEP-BY-STEP SUMMARY 50
Installing Host and PETS 50
Configuring Host to run PETS 52

C

Entering or Editing Data 54
Running the ESTIMATE module 55
Running the PROJRATE module 57
Examining the projected rates with Tabulate 57
Re-checking recently entered data, re-running ESTIMATE. 58
Re-running PROJRATE 59
Editing the projected rates 59
Running the EDPROJ module 61
Running the COHORT module 61
Tabulating results of EDPROJ and COHORT 61
Labeling the results 62
Managing your files 62
What to do about problems 63

CHAPTER 1. INTRODUCTION TO THE PETS SYSTEM

Introduction to the PETS System

The Primary Education Tracking System (PETS) is designed to allow education planners and policy makers to keep track of and to project a set of basic educational indicators for the primary level of an education system. With it, based on a time series of enrolment data, and a population projection, the user can:

- estimate the transition rates within a primary school system;
- project those rates into the future using various approaches;
- project student enrolment into the future;
- calculate various educational efficiency measures;

The projection methodologies used in PETS have been used for several years in developing country settings. Those methodologies are described in more detail under the Technical Description section of this document. In this section we describe the implementation of the system on the microcomputer and how to use it.

PETS Modules

PETS uses a modular approach. This gives the user the flexibility to run modules independently and/or to concentrate on those modules that are of particular interest. Each of the four modules of PETS corresponds to a set of related tasks. The module ESTIMATE estimates the historical education transition rates for entry into the first grade of primary school, the repetition rates and the drop out rates; the module PROJRATE projects those rates forward for a specified time period; the module EDPROJ projects school enrolment; and the module COHORT estimates various educational efficiency measures.

Using PETS on the Microcomputer

Hardware Environment

Because of the portability, access and ease of use offered by today's microcomputers, and because IBM or compatible equipment has become the de facto standard for social science analysis and simulation modeling, PETS has been implemented for an IBM/PC or IBM/PC compatible microcomputer. Because of the large amount

of data used by PETS and because of the software under which it is implemented the following configuration of equipment is necessary:

- IBM PC or compatible;
- 256K of RAM;
- 3 MB of hard disk space;
- numeric coprocessor;
- printer (optional);

Software Environment

PETS uses and produces a large amount of information and data. Two different software approaches were tried during the initial stages of development of PETS. The first one used the spreadsheet program Lotus 1-2-3. This has the advantage of being familiar to many microcomputer users and is fairly easy to use. It has the disadvantage of becoming awkward and even difficult to use when the application becomes complex, large, modular, or menu-driven.

The second approach, and the one that was chosen, was development of a set of Turbo PASCAL programs that runs under a software "shell" called Host. Host was specifically developed for modeling applications such as PETS. It was developed by the Research Triangle Institute with funds provided by USAID through the Integrated Population Development Planning (INPLAN) project. It is as easy to use as Lotus 1-2-3, can manipulate large amounts of data, and can project faster than Lotus 1-2-3. It also has several features that make properly implemented models extremely flexible and adaptable to a wide variety of settings. An introduction to the facilities of Host, that it is necessary to know to use PETS, is given in a separate section.

Organization of this Manual

The rest of this manual is organized in three sections. Chapter 2 describes the use of the PETS modules. In Chapter 3, the Host modeling system is described. Finally, Chapter 4 presents a complete technical description of the methodology underlying PETS.

CHAPTER 2. USING PETS

In running PETS there are three fundamental tasks that the user performs with the help of Host: data entry and editing, running modules, and viewing results. Each of these tasks can be done by different facilities available under Host. Thus the Host EDIT facility allows the user to enter data values into the data base and/or to change existing values. The Host RUN facility allows the user to execute a PETS module for a specified number of periods. The Host TABULATE facility allows one to view variables in the system in predefined tabular formats and to display those tables on the screen, the printer or to a disk file. Each of these facilities is described in more detail in the section on using Host.

PETS has some thirty variables. Some of these variables or parameters must be supplied by the user, and others may be calculated by a PETS module as an output. Some outputs of one module may be used as inputs by another module. For example, repetition rates (REPEATRT) are calculated in ESTIMATE, projected forward in PROJRATE, and used as input to project enrolment in modules EDPROJ and COHORT.

File Management Under PETS

While the Host software described in the next chapter provides a facility for backing up and deleting data files, the PETS user might find it helpful to use DOS commands for file management. It is likely that the user may accumulate a number of data files and that he/she may wish to store these on floppy disks rather than on the hard disk. In this case the user should use the DOS "COPY" command to copy the files to a floppy disk and then after they are safely copied, use the DOS commands "DEL" or "ERASE" to delete the files from the hard disk.

Host data files are actually inter-linked DOS files with the same name but different extensions. Therefore the use of the DOS "wild-card" character is recommended. In the following example a set of Host data files with the name FILE1, corresponding to a projection database called FILE1, is transferred to a diskette and then erased. We assume that the user is in the data subdirectory and that the target diskette is in drive A. To copy the files to the floppy disk, use

```
C>copy file1.* a:[Enter]
```

or for added safety, use this command with the "verify" option:

```
C>copy file1.* a: /v[Enter]
```

To erase from the hard disk, use

```
C>del file1.*[Enter]
```

Make sure that you have indeed copied to the floppy disk before you erase from the hard disk.

Eventually, you may wish to copy the files back from the floppy disk to the hard disk. Assuming that you are in the sub-directory which contains the data files, and that the diskette containing the archived files is in drive A, use

C>copy a:file1.*[Enter]

There are two other sets of files that should be backed up occasionally, but not necessarily as frequently as the data files. These are the LABEL files and the TABLE files. To manipulate these, follow the same sets of commands as above, substituting "table" and "label" for "file1". Note that data files such as FILE1, FILE2, etc., share the same LABEL and TABLE files, so if you intend to continue working with the system, but with new data files (such as FILE2), do not delete your TABLE and LABEL files.

All PETS variables (input and output) are stored on disk in a file, called a projection file under the Host system. One of the most powerful features of the Host implemented PETS system is that it is possible to save a set of assumptions in a file and the results of a projection based on those assumptions. For example, the user could project the various education transition rates using a logistic curve in PROJRATE and then project school enrolment using EDPROJ. These rates and the resulting enrolment projections might be stored in a file called, for example, FILE1. Then the user might wish to project the rates using a linear trend in module PROJRATE and subsequently re-run module EDPROJ. In this case the user could store the second set of rates and the results in a second file called, for example, FILE2. The assumptions and results that were stored in FILE1 would still exist and it is even possible to produce tables which compare the two files.

The four modules that make up PETS are briefly described below and information is provided on their use. A description of the methodologies underlying these modules can be found in the Technical Description section. Their use is described in this section.

PETS Modules: ESTIMATE

Description of the ESTIMATE Module

The ESTIMATE module estimates repetition, dropout, and enrolment rates, using either the standard method based on reported repeater data or the Schiefelbein method. For this module the INPUT data are:

INPUT data for ESTIMATE module	
HISTORIC	indicates whether a given year is historic (represented by an "h") or projected (represented by a "p")
POP	population by single-year age groups for the historic period (and for as many periods in the future as the user may wish to project)
ENTAGE	age of entrance to school system
	The following information is necessary only when using the "Schiefelbein" method, or if age-specific enrolment rates for the historic period are desired:
AGGENROL	enrolment in the aggregate (last) age group by grade
ENROLAGE	enrolment by grade and single-year age groups
	The following information is necessary only when using the "standard" method:
NREPEAT	number of repeaters by grade
ENROL	enrolment by grade - can be aggregated automatically from the variable ENROLAGE if data have already been entered there

When the user chooses the Schiefelbein method, some additional information is required about the repetition rate in the last (usually 6th) grade, the number of dropouts, and the age distribution of enrolment. See below.

The primary OUTPUTS of this module are:

OUTPUT data from ESTIMATE module	
REPEATRT	repetition rate by grade
DROPRT	dropout rate by grade
ENTRT	enrolment rate of new students in first grade
ASER	age-specific enrolment rates
ENROL	enrolment by grade and sex

Note that the outputs of this module are calculated only for the historic period. When using the Schiefelbein method, the estimated historical repetition rate for the sixth grade is simply the one assumed by the reader and entered into the computer at run-time. Both with the Schiefelbein and the standard methods, the estimated sixth grade dropout rate will be zero by assumption. Note that with the standard method, the estimated dropout rates will be meaningless for any year followed by a year with no enrolment data.

Use of the ESTIMATE Module

While executing ESTIMATE from the RUN facility of Host, the user must specify, as with any module, the number of periods to execute the module. In ESTIMATE, the user should specify a number that is equal to the number of historic periods. If you cannot recall the length of the historic period, specify one that is AT LEAST as large as the number of historic periods. That is, if there are five historic periods, the user should specify at least FIVE. Since no actual "projections" take place in this module, any periods beyond the historic period are ignored.

The last year of historic enrolment data should be the last year for which the indicator variable HISTORIC is set to "h", both with the Schiefelbein and the standard method.

When the ESTIMATE module is executed, as with all modules, the user is provided with an opening screen and with a screen that briefly describes the module. Next, the user is asked whether he/she wishes to use "Standard method" which uses official data, or whether he/she wishes to use the method developed by Schiefelbein.

For the standard method, the user must have entered data on the number of repeaters through the Host EDIT facility. Such data are entered into the variable NREPEAT.

For the Schiefelbein method, other information must be entered:

First, the user is prompted to enter an *estimate of the 6th grade repetition rate*. Here, the user would enter his/her best guess of this rate. All other repetition and dropout rates are calculated based on that rate.

Second, the user is prompted for an *assumption about the percentage of non-promoted students who are dropouts*. In other words, if 90 out of 100 students in grade 1 pass to grade 2, the user must supply a percentage of the 10 students who do not pass who dropout. If the user puts that percentage at 50% then, in this example, 5 students would drop out and the other 5 would repeat first grade. In this case the user would enter 50.

Third, some additional information is required about the disaggregation of enrolment data. In many countries it is possible that there are some students, particularly in the later grades, who are "over-age". For example it is not unusual to have a number of 17 or 18 year olds in the sixth grade. Enrolment data by age and grade that are published in official sources often report the number of such over-age students, but the breakdown of their ages is often not sufficiently detailed for the Schiefelbein method to work. For example, the last reported age group may be 14, rather than 18. *The aggregated data for the last age group of the ENROLAGE matrix must be entered initially into the variable AGGENROL*. When these aggregated data are present the user must tell the computer to disaggregate the data when prompted.

If disaggregation is necessary, the user is prompted to *specify the last age group*. It is important to enter the correct value here, otherwise the disaggregation will be incorrect,

as will all the estimated rates.

For the standard method, the computer requests information on whether the enrolment data were entered in disaggregated form (through the variable ENROLAGE) or directly in aggregated form (through the variable ENROL). If the data were entered in disaggregated form, the computer asks whether the last age group is actually an aggregated one, as in the Schiefelbein method.

Finally, after the estimation of the transition rates is completed, the computer asks the user whether age-specific enrolment rates are desired. These rates have meaning only if age-specific enrolment data were entered. Note that there is no need to re-do these calculations unless new data on population or enrolment by grade and age have been entered.

PETS Modules: PROJRATE

Description of the PROJRATE Module

Future dropout, repetition, and first grade enrolment rates are necessary inputs for the calculation of enrolment projections. The user has the choice of entering his/her own scenario about the evolution of these rates, or he/she can let the computer do so by running PROJRATE. The module PROJRATE projects repetition, dropout, and enrolment rates for a specified number of periods, using one of five approaches that can be chosen by the user:

- future rates fixed at the level of the last historic value
- future rates fixed at the level of the historic mean
- varying future rates that follow a logistic curve
- varying future rates that follow a linear trend
- varying future rates that attain target levels at a target date, where the targets and the target year are given by the user

The INPUT data for this module are the historic values of:

INPUT data for PROJRATE module	
REPEATRT	repeater rates by grade
DROPRT	dropout rates by grade
ENTRT	entrance rate for the first grade
TARREP	target repetition rate - necessary only for target setting projections
TARDROP	target dropout rate - necessary only for target setting projections

Before running PROJRATE, the historic values of these rates must have already been estimated using the module ESTIMATE, or they could have been entered manually from

some other source.

The module also requires as INPUT information about which periods to use as the basis for the projections:

Indicator variables for PROJRATE	
HISTORIC	signals which periods are historic, which are to be projected
INCLUDE	signals which of the historic periods to be included as the basis for the projections - setting INCLUDE equal to "n" (no) will result in the data for that year being ignored in estimating projection trends

The variable INCLUDE is useful because it allows the user to ignore periods which may be considered atypical and so may not want to be included in the calculation of future trends. Note that any period designated as *not* included will be subsequently ignored in the module COHORT.

The variable HISTORIC should remain set as it was for the ESTIMATE module. If the Schiefelbein method was used for estimating historical rates, the variable INCLUDE should be set so that the last year for which there are enrolment data IS NOT included. If the standard method was used, then the variable INCLUDE should be set so that the last year with data on enrolment and numbers of repeaters IS NOT included. This is due to the fact that, with the Schiefelbein method, to estimate the rates for any given year, enrolment data for that year *and the next* are needed. With the standard method, enrolment data for the given year and enrolment and repeater data for the next are needed. For example, if the last year with enrolment data is 1984, the estimated rates for 1984 itself are meaningless. Therefore, 1984 must not be included in estimating the trends - in fact, the rates for 1984 must themselves be projected. NOTE: this applies also to any gaps in the data. Thus, for example, if there are enrolment data for 1974-1976, and 1981-1984, then even though 1974-1984 are designated "h" with the HISTORIC variable, both 1976 and 1984 must be designated "n" (for "no") with the INCLUDE variable: since there were no enrolment data in 1980 or 1985, the estimated rates for 1979 and 1984 are meaningless. Of course, 1977-1980 and 1985 and onward must also be designated "n" with the INCLUDE variable. Here is the rule: if you have n historic years WITH DATA, but with m gaps in the data (regardless of their length), then the number of included years can be AT MOST n-m-1, making sure that the last historic year is not included. All this should be checked carefully. Study the table below to make sure you understand it.

Example of use of HISTORIC and INCLUDE			
Year	Have historic data?	HISTORIC	INCLUDE
1974	Yes.	h	y
1975	Yes.	h	y
1976	Yes.	h	n
1977	No.	h	n
1978	No.	h	n
1979	No.	h	n
1980	No.	h	n
1981	Yes.	h	y
1982	Yes.	h	y
1983	Yes.	h	y
1984	Yes.	h	n
1985	No.	p	n
1986	No.	p	n
1987	No.	p	n
1988	No.	p	n
1989	No.	p	n
1990	No.	p	n

Note: there are 7 historic years with data, there is 1 gap in the data, therefore there are 7-1-1=5 years with INCLUDE set to "y".

Of course, all of the above applies also to years in which the estimated rates are considered to be "atypical". In these cases, the year with the "atypical" rate should not be included. If it is the enrolment data that are considered atypical, then the year with atypical enrolment data AND THE PREVIOUS ONE should not be included.

The OUTPUTS of this module are the projected values of:

OUTPUT data from PROJRATE module	
REPEATRT	repeater rates by grade
DROPRT	dropout rates by grade
ENTRT	enrolment rate for the first grade
TREP	t-statistic for the slope of the repetition regression
TDROP	t-statistic for the slope of the dropout regression

Use of the PROJRATE Module

Again the user will execute PROJRATE from the RUN facility of Host and, after selecting a "projection" or data file, will then be prompted to specify the number of periods to project. In PROJRATE, as with EDPROJ, the number of periods that are given here include the number of historic periods. Thus if the user has an historic period of data corresponding to 1979 - 1984, and wants to project until the year 1994, he/she would specify a projection period of 15 (1994-1979=15). If the user attempts to project for too short a period the PROJRATE module will detect this and will so inform the user.

Once the projection period has been specified and the user executes the module, he/she will subsequently be asked to choose the method wanted for projecting the rates. An initial choice is made between fixed and varying rates. Then a choice is made among either the two fixed methods or the three available varying methods, depending on whether the user chose fixed or varying rates.

After projecting the rates, the user should check the t-statistics for the trend estimates. Projections based on weak trends are very risky, and checking this statistic is the minimum that should be done by the careful user.

For all choices except the target method, there are no further decisions required of the user and the program proceeds.

If the user has chosen to project using target rates, he/she is reminded that these targets, that are in the variables TARREP and TARDROP, must be up to date. Targets can be updated through the Host EDIT facility. If the target variables are up to date, the user then specifies the date at which the targets are to be achieved. Obviously, this date must be within the projection period.

PETS Modules: EDPROJ

Description of the EDPROJ Module

The module EDPROJ projects enrolment in primary school for a specified number of periods using the standard grade transition method of projecting education enrolment. For this module the INPUT data are:

INPUT data for EDPROJ module	
POP	population by age
REPEATRT	repeater rates by grade
DROPRT	dropout rates by grade
ENTRT	entrance rate for the first grade
ENDAGE	oldest "legal" school age

These rates must have already been supplied by running the other modules in the

system, i.e. ESTIMATE and PROJRATES, or they may have been entered manually.

The OUTPUTS of this module are:

OUTPUT data from EDPROJ module	
ENROL	enrolment by grade
GER	gross enrolment rate
TOTENROL	total (aggregated across grades) enrolment in primary school
SCHLPOP	total (aggregated across ages) population of school age

Use of the EDPROJ Module

As with other modules, the user of EDPROJ will have to tell Host the number of periods to project. Here, there must be input data for each of the projection years. Hence if the user has an historic period of 1979 - 1984 and wants to project until 1994 he would specify a run of 15 periods, 10 of which would be projections. Here, the module PROJRATE would have had to have been run for 15 periods. In addition, population data for the projection period as well as the historic period would have to be present.

Once this module is executed, no further input is required on the part of the user.

PETS Modules: COHORT

Description of the COHORT Module

The module COHORT projects the history of a cohort of 1000 entrants in primary school according to given repeater and drop out rates. The module allows the user to use two alternative methods:

- fixed drop out and repeater rates
- varying dropout and repeater rates

The required INPUTS of the COHORT module are:

INPUT data for COHORT module	
REPEATRT	repeater rates by grade
DROPRT	drop out rates by grade

As with the EDPROJ module, these rates must have already been estimated by running other modules in the PETS system, or may have been entered manually.

The OUTPUTS of the COHORT module are:

OUTPUT data from COHORT module	
IORATE	the input/output ratio
GRADRATE	number who graduate out of 1000 entrants
YRSSCHL	average years spent in school
SYGRAD	student/years per graduate not counting student/years "wasted" by drop-outs
SYGRAD2	student/years per graduate, counting student/years "wasted" by drop-outs
ONTIME	number of graduates who graduate "on time", i.e. have not repeated

Use of the COHORT Module

Use of the COHORT module is similar to that of module EDPROJ. An appropriate number of periods must be specified and inputs present for the length of the projection period. Once the model is run the user is asked to choose between two alternative methods.

In the first method, all transition rates are assumed constant and equal to those during the entry year. For this method there must be as many periods of data for the transition rates as the current number of periods for which COHORT is being executed. This method is applied to the historic period for periods which are "included" by the INCLUDE variable being set to "y". It is also applied to the whole projection period.

In the second method, the transition rates that are applied change according to the actual calendar year in the future school history of the cohort of entrants. Experience has shown that it may take as long as 15 years for a cohort of entrants to finally graduate or dropout. In this case, a ten year projection would require AT LEAST twenty five years of future rates. For this reason, the program reminds the user who has chosen method 2 that a sufficient amount of data for future repeater and dropout rates must be in the data base. The user has the option of exiting the program and re-running PROJRATE if he/she determines that there are not enough periods of transition rate data.

Because this second method breaks down if there are gaps in the data, and since gaps in the data can only occur in the historic period, it cannot be applied in any periods designated "h" with the HISTORIC variable. The experienced user can trick the system by designating even the historic periods as "p", but this is fairly dangerous, since one might forget that this was done, and run into trouble later, when re-estimating or re-projecting the transition rates. Also, this should be done only if one is absolutely sure of the validity of the historical data, and if there are no gaps in them.

Note that the module COHORT can be used for analyzing the impact of programs oriented to the first few grades. For example, one can calculate the "fourth grade completion rate" by forcing the computer program to think there are only 4 grades right before running the module COHORT. This requires EXTREME caution, because it is easy to forget one has done this, and it will bias any subsequent exercises with the ESTIMATE, PROJRATE, and EDPROJ modules. The way to do this is as follows. Use the D)efine function in the Main Menu, then the D)imension function in the Define Menu, and then the E)dit function in the Dimension Menu. Enter the name of the GRADE dimension. Then use the cursor to go to the cell for the Current No. of Levels, and input the number 4, and press RETURN. Then press [s] for Save, [Enter] to save the dimension under its own name, and press [ESC] several times until you get back to the main menu. After running the module COHORT, go back and change the current number of levels of the grade dimension to what it was.

CHAPTER 3. USING HOST

Introduction

How Host is Used with PETS

As you have seen, the PETS system consists of four separate, but related modules which can be run in any of twenty different modes. In managing and running any model, three basic tasks are necessary: data entering and editing, model execution, and display of results. The Host system provides an easy-to-use and flexible environment in which PETS can be maintained and run. This chapter describes how Host can be used for carrying out the three fundamental modeling tasks necessary for PETS. A more complete Host manual is available from RTI.

What Host is and What it Does

Host is a specialized database management system designed to facilitate the implementation and utilization of all types of models. Host provides an integrated framework supporting all stages of model implementation from program design through result presentation.

Host's generalized facilities allow it to be used equally well for education, demographic, economic, ecological, epidemiological or any other type of model. Further, Host has specialized model support features that go beyond standard database operations. Data manipulation facilities included in Host provide basic input and output access routines for use by models, as well as interactive data entry, editing and reporting facilities to model users.

Host includes many of the characteristics common to most database management system (DBMS) software, including facilities for data definition, data labeling, full screen data editing, table generation and application program development. Unlike most DBMS software, Host's application development programming language is Turbo Pascal. Since Turbo Pascal is a widely used, general purpose, high-level language, this allows Host users to take immediate advantage of an immense body of existing Turbo Pascal literature and enhancement products.

Models implemented under Host are typically divided into logical sub-systems called "projection" modules. These modules pass data to one another through a common data file called the Host projection data base. Thus the output of one module may be the input to another. This modular approach allows greater flexibility and reliability than a single, large program. All Host modules are written in Turbo Pascal.

Host uses a multidimensional spreadsheet-like editor to allow users to access projection data. Data displayed in the editor can be clearly and meaningfully labeled using Host's labeling facility. Investigation, and comparisons of alternative modeling scenarios is made possible by the ability to set values for a base or reference run, and then use

that as the basis for comparison of alternative sets of values. Tables produced by Host include built-in comparison calculations so that differences between runs may be readily identified.

Any task which may be expressed as a Pascal language program can be developed into a Host module. Host places no restrictions upon what a projection module may do. Possibilities include large scale simulations, interactive training exercises, or computer graphics presentations. Programs may be highly interactive with values and decisions made by the operator as the program executes, or programs may take data prepared with the Host editor and run without any manual intervention.

Getting Started

Entering Host

Host is entered by typing "HOST" at the operating system prompt as follows:

```
C>host[Enter]
```

If the Host program is found in a subdirectory called, for example, HOST, then the proper sequence is

```
C>cd \host[Enter]
C>host[Enter]
```

Naturally, if you start from the A drive, the command

```
A>c:[Enter]
```

must precede all others.

Next, press any key to proceed, then select the appropriate operator status when requested by the Access Menu. You should choose the status "User" by pressing [Enter] unless you are developing projection modules or defining data structures.

After selecting the User status, you will be presented with Host's Main Menu.

How Host Uses Cursors

A cursor is a special symbol, color or shape used to show position on the display screen. Host uses several types of cursors. When you select a command from a horizontal menu or an item from a vertical list, Host indicates the current selection by placing it within a highlighted block. This type of cursor is also used to show position when editing data values in the Host Data Editor.

When entering data a character at a time, Host uses a blinking rectangle or square to indicate the character currently being entered. A blinking rectangular cursor indicates overwrite mode, while a blinking square cursor indicates insert mode.

Host occasionally combines the above cursor types to indicate both line and character position on the screen.

Using the Keyboard

No attempt is made in this Guide to familiarize you with the standard IBM PC keyboard arrangement. This information is included in the *Guide to Operations* manual provided with all IBM microcomputers and most IBM compatible microcomputers.

Host has been designed to make effective use of the standard IBM PC keyboard. Most special cursor movement and editing keys operate in standard ways. In addition, Host understands many Wordstar commands for cursor and for editing commands.

Using the Menus

Most of the functions in Host are selected from "Lotus-style" horizontal menus. These menus are displayed at the top of the screen (called here the Menu space). Whenever the Menu Space is used to display other information, a menu of available commands can be displayed by pressing [/] or [Esc]. The top line of the Menu Space is used to display a list of available commands. One of these commands is within a highlighted block called the menu cursor. The bottom line of the Menu Space is used to display a short explanation of the highlighted command. Often this explanation will include a list of commands which become accessible if the highlighted command is selected. There are two methods of selecting commands from these menus: *cursor* selection and *key* selection.

Cursor Selection. To select a command from the menu using the menu cursor, use the cursor movement keys to move the cursor to the desired selection. As the cursor moves, the explanation below the command list changes to correspond with the highlighted command. When the menu cursor is on the desired command, press [Enter].

Key Selection. Normally, each command in the command list will begin with a unique letter. To select any command on the menu, just press the key which corresponds to the first letter of the command. Using this method, it is not necessary to move the menu cursor to the desired selection. Key Selection is faster than cursor selection. It is designed to improve the efficiency of users familiar with the program. Some menu command lists, however, are composed of file names over which the program has little control. Such lists often contain commands which begin with the same letter. In this case, Key Selection will scan from left to right to select the first command which begins with the letter entered. Cursor Selection is the safest method in this situation.

When there are more commands in a menu than can be displayed across the Menu Space, an arrowhead will be displayed at one or both ends of the command list. These arrowheads indicate that additional commands are available in the indicated direction, and can be accessed by moving the cursor beyond the last command currently displayed

at that end of the Menu Space.

User Versus Manager Access

Host is designed to support two different categories of operators: those who use only the presently available projection modules, and those who develop and install new projection modules. Host labels the first category User and the second category Manager.

User or Manager access must be selected whenever Host is executed from the operating system. Once within Host, the access level may be changed at any time by selecting Access from the Main Menu. To set your access status to Manager, you must enter a password. This password may be changed once Manager access is granted. Operator access may be changed from Manager to User without a password.

User. Operators in this category are permitted access to all commands and data values except those which are closely interconnected and which could disrupt operation of Host if incorrectly modified. The essential structural core of a Data File and associated program modules are protected from this type of accidental modification. This allows users the maximum level of flexibility while reducing their concerns about any possible unintentional damage. Users are also prevented from changing any values which restrict data entry.

Manager. Operators in this category are permitted access to all Host commands and all data values which are not reserved for use by Host. Managers are permitted to define dimensions and variables, delete existing dimensions and variables, and establish data entry restrictions for Users. After completing restricted operations, a Manager may change status back to User to provide a greater level of system protection.

Using the Query Field Facility

There are many places in Host where you are asked to enter the name of some item. The particular item may be a label format, dimension, variable, table or file. In any of these cases, it may be difficult to remember the desired name. Pressing [?] in this situation will display a list containing the names and descriptions, if available, of all possible entries. Host prompts with this feature are called Query Fields and are always followed by a blank, or field, to be filled in, preceded by an inverse video "?". When the item to be selected is a file, the list is presented as a horizontal "Lotus- style" menu in the Menu Space. Otherwise, the list is displayed vertically in the Work Space. A cursor is provided and all the normal cursor, scrolling and paging functions are implemented to allow positioning of the cursor on any item in the list.

A special search feature is available to aid in finding a particular name from any vertically displayed list. The search feature is invoked by pressing [Alt]+[S]. A complete or partial name can then be entered and the page of the list containing that name will be displayed. Pressing [Enter] will select the item indicated by the cursor. Pressing

[/] or [Esc] will not select any item, but will return you to the previous step.

Using the Toggle Field Facility

Often you will be prompted to make decisions in Host that require a limited number of possible responses. In these situations, Host uses a feature called Toggle Fields. Toggle fields are easily identified by the inverse video "±" displayed next to the prompt. Some toggle fields have only two possible responses, such as YES or NO, while others present several choices such as selecting the Console, the Printer, or a Plotter as the destination for a graph. In all toggle field prompts, you will be shown the current setting, which you may leave unchanged or change to any of the other possible settings. If you wish to change a toggle field, several methods are available. You may use the [+] key or the [Space] bar to display the next choice, you may use the [-] key to display the previous choice, or you may type in any selection appropriate to the current toggle field. The toggle field will "wrap" around to the next choice whether you are scanning in a forward or reverse direction. Therefore, you can easily review all the choices, several times if desired, by repeatedly pressing the same toggle key. If you manually type in the selection and make an invalid entry, you will receive a message at the bottom of the screen,listing all the keystrokes or entries available to you for that field.

Editing Data (E)

Host's Data Editor can be used to enter new data values into a Data File, to edit existing data or to simply view Application Module input and output data. You can access the Data Editor by choosing the Edit selection on Host's Main Menu. The Editor begins by prompting for the name of the Application by presenting a horizontal menu of available data files. You can confirm the default entry by pressing [Enter], or you can enter the name of any Data File in your system. If you press [Esc] at this point, you will return to the Main Menu.

Selecting a Variable

Next, you are prompted for which variable in the selected Data File you wish to work with. Since this is a Query Field, a list of available variables can be displayed by pressing [?]. You may also select a different Data File or return to the Main Menu from this point by pressing [/] or [Esc] to raise a menu with the selections Resume, File and Quit.

The Data Entry Cursor

The Editor uses an inverse or colored video cursor to indicate your current row and column position within the data matrix. This is combined with a small rectangular or

square blinking cursor to indicate the current character position within the current entry.

If Data are not Found in the Data File

If data for the selected variable or current period of the selected variable cannot be found in the projection file, the Editor provides appropriate default values and displays the following message at the top of the screen:

Data not found in file; default values have been supplied.

For numeric variables, the default value is zero. For character and character string variables, the default value is the ASCII NUL character (^@ or decimal 0) which is displayed on the screen as a blank space. In any case, you may change these values at any time using the Editor.

Changing Editor Supplied Default Values

When you first change one of these default values for a variable (or period of a periodic variable), you will notice that the above informational message will disappear from the top of the Editor display. This indicates that the Editor has sensed the change and will therefore save the data to the Data File if so directed.

Saving Editor Supplied Default Values

For nonperiodic variables, Editor supplied default values will always be saved to the Data File when Save is selected from the Edit Menu. For periodic variables, the Editor will not save a period of data to the Data File unless at least one of these default values has been changed in that period, or unless that period of data is the one currently displayed by the Editor. This means that if you wish to save an entire period of Editor supplied default values to the Data File, that period must be the one currently displayed when you issue a Save command to the Editor. This feature is designed to minimize the amount of disk space occupied by unnecessary zero or NULL values.

Variables With One or More Dimensions

Since a screen display is a two dimensional space, only a two dimensional section of a multidimensional variable's data can be displayed. The names of these two dimensions are displayed on the left side. All other dimensions are displayed on the line directly above the data.

When a multidimensional variable is selected for editing, the Editor chooses the dimension having the largest current number of levels for the Row Dimension and the one having the next largest number of dimensions for the Column Dimension. This ensures that the largest two-dimensional slice of data will be the one displayed in the Work Space. All other dimensions are By- Dimensions and will appear on the Dimension Line. If the variable is periodic, the time dimension PERIOD will be the first By-

dimension on the Dimension Line.

Changing the Levels of the Row and Column Dimensions

The current level of the Row and Column Dimensions can be changed simply by using the cursor movement commands to change the row and column position of the cursor. Commands for moving through the levels of the row and column positions of the data matrix are identical to those of Lotus 1-2-3.

If there is more data in the two dimensional slice than can be displayed in the Work Space, you may display other portions of the slice using the paging and scrolling commands. End Mode is implemented as in Lotus 1-2-3 and is provided as a fast, convenient way to move around within the two dimensional slice of data. Scroll-Lock is also implemented as in Lotus 1-2-3 should you want to use it.

Changing the Levels of the By-dimensions (Dimension)

To change the level of any of the By-dimensions, display the Edit Menu by pressing [/], then to choose the selection Dimension. The cursor will immediately be placed on the current level of the first By-dimension in the Dimension Space. If the current variable is periodic, this will be the implied time dimension PERIOD. You can move the cursor to other By-dimension fields using [Right], [Left], [Home], [End], [Ctrl]+[Right] or [Ctrl]+[Left] the same way you would in a horizontal Lotus-style menu.

Since the current level of each By-dimension is a Toggle Field, there are several ways to change this value: you can use [+], [Space] or [-] to "toggle" through the list of available levels, or you can directly enter the desired level. For example, if the current level of a By-dimension is "Urban", you can change it to Rural by simply entering "rural" (upper or lower case is not important) or by toggling the entry until it is "Rural". You may change the current level of more than one By-dimension at a time.

When you are satisfied with the level settings you have chosen, simply press [Enter] to display the data slice corresponding to the new level settings and to return the cursor to the Work Space. If you do not wish to retain the changes you have made in the By-dimension levels, press [Esc] or [/] for an opportunity to cancel or retain the changes.

Using Function Keys to Change the Period

For periodic variables, you will most likely want to change the current level of the time dimension PERIOD. To make this as easy as possible, function keys have been assigned to display the next period, previous period, first period and last period of data.

Function keys used in the Data Editor for periodic variables.	
IBM PC Command	Function Key in Data Editor
[F3]	Display next period of data.
[F4]	Display previous period of data.
[F5]	Display first period of data.
[F6]	Display last period of data.

The "Last Period" of data is the last period of data in the current Data File which contains data for any variable, not just the current variable. Pressing any of these keys immediately displays the data for the indicated period. When editing a periodic variable, a function key status line will be displayed at the bottom of the Work Space to remind you that these commands are available.

The Edit Menu

If the Edit Menu is not displayed at the top of the screen, it can be displayed by pressing [/] or [Esc]. The Edit Menu offers the selections Resume, Area, Period, Interpolate, Save, Dimension, Variable, File and Quit. The selections Resume, Save, Dimension, Variable, File and Quit will be discussed in the following sections. The advanced functions Area, Period and Interpolate are not yet completed and will therefore not be discussed.

Resume. The Resume selection on the Edit Menu returns to the EDIT DATA mode, placing the Data Entry Cursor in its previous row-column position.

Area, Period, Interpolate. The Area, Period, Interpolate selections of the Edit Menu have not been completed. Attempting to choose any of these selections in this version of Host will result in the following message:

Sorry, that selection is not active on this menu.

Save. To save any changes you have made to the data of the current variable, select Save from the Edit Menu. You will be prompted as follows:

Save Variable as: (current variable name)

If you have been granted Manager Access, you may request that the data be saved to the Data File under a new name. If you choose to do this, several things can happen. If a variable of the specified name already exists, but has a different structure than the current variable, the save will not take place and you will be given an opportunity to specify a different name. If a variable of the specified name exists having the same structure as the current variable, you will be asked whether you want to proceed to overwrite any data which may exist for that variable with the data of the current variable. If you answer "No", the save will not take place and you will be given an opportunity to specify

a different name. Otherwise, the data will be saved under the specified name, replacing any data which might have existed for that variable. If no variable of the specified name exists, a definition matching that of the current variable will be added to the Data File under the new variable name and the data will be saved under that name. If you have been granted User access, you may not save the data under a new name, and should simply press [Enter] to save the data to the Data File, or [Esc] to return to editing the data.

Dimension. The purpose of the Dimension selection is discussed above in the section on Changing the Levels of the By-dimensions.

Variable. To edit a different variable, choose the Variable selection on the Edit Menu. If you have made any changes to the data of the current variable, you will be asked whether or not you wish to save those changes before continuing. You will be prompted to enter a variable name. This field is a Query Field, so enter [?] to select from a list of available variables. This is a "no return" selection, i.e. you cannot return to your previous state within the editor except by again selecting the same variable you were editing. This is because some of the memory space used when editing a variable must reallocated when attempting to select a new variable.

File. To select a different Data File to edit, choose the File selection from the Edit Menu. If you have made any changes to the data of the current variable, you will be asked whether or not you wish to save those changes before continuing. In the current version of Host, you are presented with a menu of Data Files.

Quit. To end the editing session and return to Host's Main Menu, select Quit from The Edit Menu. If you have made any changes to the data of the current variable, you will be asked whether or not you wish to save those changes before continuing.

Running a Module (R)

To execute any Application Module from within Host, select Run from Host's Main Menu.

After selecting Run, you will be prompted for the name of the Application Module you want to execute. You may enter a module name or, using the query feature, you may select from a list of modules.

Once a module is selected, the default settings for executing the module are displayed. You may choose to accept these settings and execute the module immediately (Go), or to Edit the settings before executing the module. The settings are: Data File, Starting Period, and Number of Periods. These settings are discussed in the following

sections.

Specifying the Data File

The default Data File, but this may be changed by the operator to any Data File in the user's system. The variables and dimensions required by the selected module must match existing variables and dimensions in the chosen Data File or the module will abort.

Specifying the Starting Period

This sets the starting time period for the projection. The default value of this setting is 1. Some modules may not function correctly if you choose a starting period other than 1; this depends upon the function and implementation of the module. You should consult any documentation provided with the module you intend to execute before changing this value. The maximum value of this field is 1023, but starting period and number of periods to project are interrelated such that starting period + number of periods must not exceed 1024.

Specifying the Number of Periods

This is the number of periods for which you wish the module to project. The default value of this field is 1. Normally you will want to change this setting. Some modules may yield nonsensical results if asked to project a large number of periods; this depends upon the purpose and implementation of the module. You should consult any documentation provided with the module you intend to execute before entering a very large value into this field. The maximum value of this field is 1023, but starting period and number of periods to project are interrelated such that starting period + number of periods must not exceed 1024.

Executing the Application Module

Once you are satisfied with these settings, you can display the Run Menu by pressing [/] or [Esc]. You may then select Go to execute the module with the current settings, Edit to further edit the settings or Quit to return to Host's Main Menu.

If you execute the selected module by choosing Go, Host will proceed to pass the current values of the displayed settings, then control of the microcomputer over to the Application Module.

Once the Application Module is in control, it may do whatever it has been programmed to do, including prompting you for information, manipulation of data in the selected Data File and displaying information. A module may even take over the entire screen to display graphics images. While the module is in control of the microcomputer, you will see the Mode Indicator displaying the name of the module.

When the Application Module has completed its task, it should return control to Host. Host will then ask if you wish to execute another Application Module. If you respond "Yes", you will see the same projection settings that were established in the prior run displayed. You may edit these settings to re-execute the same module, or to begin execution of other Application Modules.

Using Results: Generating Tables (T)

The results of running a module may be viewed by generating pre-defined tables with Host. Definition of these tables is possible under Host, and is explained in the more complete Host User's Guide, available from RTI. To generate tables from within Host, select the Tabulate command from the Host Main Menu or from the Table Menu.

After selecting Tabulate, you will be prompted to enter a table name. You may use the query feature to display a list of all currently defined tables. Move the cursor to the desired table and press [Enter]. Once you have selected a table, the default values for the TABLE parameters will be displayed, and you may choose to accept them and generate the table immediately (Go), or to Edit the parameters before producing the table. The parameters are: Data File, Comparison File, Output File and Period List. Each will be explained in the following sections.

Specifying the Data File

The default Data File may be changed by the operator to any Data File in the user's system. The variables and dimensions stored in the table definition must match existing variables and dimensions in the chosen Data File or the table will not be produced. The Data File is assumed to be in the Data Directory specified in the Configuration File unless an explicit path is specified.

Specifying the Comparison File

The default Comparison File is the file specified in the Configuration File, but this may be changed by the operator to any Data File in the user's system. The variables and dimensions stored in the table definition must match existing variables and dimensions in the chosen Data File or the table will not be produced. If there is a entry in the table definition that does not have a corresponding entry in the selected Data File, an error message will appear on the screen, and the Table generation will abort. If a Comparison File is specified but there are no comparison variables in the tables, the file is

ignored.

Specifying the Output File

The default output file is the default device specified in the Configuration File, and thus is either "CON" (console or display screen) or "PRN" (printer). If PRN is chosen and a printer is NOT connected to the computer, the system will "lock up" and the computer may have to be re-booted. The operator may change the output device or may specify output to a Lotus- Import File or to a print-ready file. To specify a Lotus-Import File, simply choose any valid file name and use the extension "PRN". Any other valid file name and extension specifies a print-ready file. Output is directed to the Data Directory specified in the Configuration File unless an explicit path is given with the Output File.

Specifying the Period List

The period list specifies the periods of data to be included in the table. This only makes sense for tables made up of periodic variables. A period list is the word "ALL", or a list of integer period values and period intervals separated by commas. A period value is an integer from 1 to 1024 inclusive. A period interval has the form

first-last

where first and last are period levels, and specifies the use of all periods from first to last inclusive. If first is omitted, all periods up to and including last are included; if last is omitted, all periods from first on are included. If first and last are omitted all periods are included, thus "-" is a synonym for ALL. Overlaps in groups are ignored, so each period specified is included in the table only once. The following table shows some valid period list entries and, for each entry, the periods which would appear in the output table.

Period List Entry	Periods included by the List
1,2-5,3-7,10	1,2,3,4,5,6,7,10
1,1,1,3,4	1,3,4
-	All periods
-5, 8-10	1,2,3,4,5,8,9,10
1-3,5,-	All periods except 4

Generating the Table

Once you are satisfied with the settings described in the above sections, you can display the Tabulate Menu by pressing [/] or [Esc]. You may then select Go to tell Host to generate the table using the current settings, Edit to further edit the settings to Quit to return to Host's Main Menu. If you choose to generate the table, by choosing Go, Host will proceed to pass the current values of the displayed settings, and then generate the table. At the top of the screen, in the Menu Space, Host will display the Host Instruction which would have achieved the same purpose as this interactive procedure had it been encountered in an Instruction File or entered at the DOS prompt.

When the table is completed, Host will ask if you wish to generate another table. If you respond "Yes", you will see the same tabulation settings that were established in the prior run displayed in the Work Space. You may edit these settings to re-execute the same module, or to begin execution of other Application Modules.

Managing Files (M)

Host provides a wide variety of file management capabilities and utilities for working with files. These facilities are accessed by selecting the Manage (M) option on the Host Main Menu. From the Manage Menu you can then choose to manage Data Files (Data), manage Library Files (Library), or to access the utilities (Utility). As with all the Host menus, Quit or [Esc] returns you to the previous menu. In this document we only describe the most common Manage functions which you may need in using PETS. For further information, please see the more complete Host User's Guide available from RTI.

Managing Data Files

File management functions available for Data Files include:

Create
Rename
Backup
Delete
Verify
Pack

We shall discuss only Rename, Backup, and Delete. All three of these functions can be performed also from DOS, but it is easier to do so from Host.

Renaming a Data File. Data Files can have the name modified to better fit your projection naming conventions. Simply choose the Rename option on the Manage Data Menu and select a Data File from the file menu. You will then be prompted for the new Data File name. If a Data File with the specified name already exists, you will be warned by Host and asked if you wish to overwrite the existing file.

Backing Up a Data File. It is important to create and keep backup copies of your Data Files. The Backup option on the Manage Data Menu provides a quick means of making copies of your important Data Files. After selecting a Data File from the file menu, you are prompted for the name of the target file. If a Data File by the specified name already exists in the target directory you will be warned by Host and asked if you wish to overwrite the existing file. You may not copy a file to itself. Host verifies that sufficient disk space remains for the file copy before attempting the backup. If there is not room enough you will get an error message stating the problem and you will be re-prompted for the name of the file copy. In this case, simply specify a file on another drive where more space is available, or insert a new diskette or cartridge with sufficient space for the file and retry the copy.

Deleting a Data File. Old Data Files can be removed when no longer needed. This is done by choosing the Delete option on the Manage Data Menu and selecting a Data File from the file menu. To prevent accidental erasure of files, you will be given a chance to confirm your delete selection before the actual deletion takes place.

Using the Management Utilities to Obtain PETS Documentation

The Host based PETS system comes with part of the manual which you are presently reading. However, there may be times when the manual is not next to your computer and you cannot remember the use of a certain PETS module. The Host Read facility allows you to view the contents of a file which contains information about the four PETS modules. Host has a "Read" function that allows you to display and scan forwards and backwards through any text file on your disk drives. To use the Read feature, select the Manage Utility option from the menu, and then choose the Read option. You are then prompted for the name of the file to be viewed. The file name can be preceded by a drive and/or subdirectory specification. You may scroll through the file in either direction with the [Pg Up] and [Pg Dn] keys. When you have finished examining the file and its contents, enter [Esc] to return to the menu. The file name relevant for Host is PETS.

Appendix A to Chapter 3: Installation

Installing the Host system on your microcomputer requires that you do the following:

- Verify that your computer system has the recommended hardware (see below).
- Copy the Host program disks to your hard disk or cartridge.
- Create or modify the file CONFIG.SYS.
- Tell DOS how to find Host.

Required and Recommended Hardware

Host is designed to run on IBM PC or compatible microcomputers running version 2.0 or later of PC-DOS or MS-DOS operating system. Although these two operating systems are not identical, none of their differences effect the operation of Host. Both operating systems are referred to throughout the remainder of this document as PC/MS-DOS or simply DOS. Host has been verified to function correctly with the following microcomputers and operating systems:

Microcomputer	Operating System	Version Number
IBM Personal Computer	IBM PC DOS	2.0 - 3.1
IBM Personal Computer	XT IBM PC DOS	2.0 - 3.1
IBM Personal Computer	AT IBM PC DOS	2.0 - 3.1
Compaq Portable	MS DOS	2.11
	IBM PC DOS	2.0 - 3.1
Compaq Deskpro	MS DOS	2.11
	IBM PC DOS	2.0 - 3.1
Wyse PC	MS DOS	2.11
Kaypro 2000	MS DOS	2.11
Kaypro 16	MS DOS	2.11
Sperry IT	MS DOS	3.1
Wang PC	IBM PC DOS	2.0 - 3.1

Note: when using Host with a Wang PC, an IBM emulation board must be present. Some minor differences between the Wang PC keyboard and the IBM PC keyboard are listed in the table below.

Wang PC Keyboard	IBM PC Keyboard
[Back tab]	[Esc]
[Prev]	[PgUp]
[Next]	[PgDn]
[Cancel]	[ScrollLock]
[GoTo]	[NumLock]
[Insert]	[Ins]
[Exec] or [Return]	[Enter]

Host itself will operate in 256K of Random Access Memory with all of the machines and operating systems listed above. The memory demands of projection modules running within host, however, depend upon their purpose and implementation. Documentation provided with such projection modules should list their specific memory requirements. Memory demanding utilities (print spoolers, RAM disks, etc.) and memory resident programs (spelling checkers, keyboard macro utilities, etc.) will decrease the amount of memory available to Host. In practice, we recommend at least 384K, and ideally 512K or more of Random Access Memory.

It is possible, but not practical, to run Host using only two 360K floppy disk drives. We strongly recommend that Host be installed on a hard disk, RAM disk, or high-speed cartridge disk system. A numeric coprocessor (Intel 8087 or 80287) is also required.

Copying Host to a Hard Disk or Cartridge Disk

Host is normally distributed on 5, 360K double-sided diskettes (Host is also available on a 10Mb IOMEGA Bernoulli Box cartridge upon request). All of the files on these diskettes must be copied to your hard disk or cartridge disk. First, verify that sufficient disk space is available. Approximately 1.3 megabytes is required to store all of the files on the distribution diskettes. The files actually required to run Host require about 1.0 megabytes (1,048,576 bytes).

To avoid conflicts with other files already on your system, it is best to place all Host system files together in a new subdirectory. In the following examples, the hard disk is assumed to be drive C. If this is not the case, simply replace the C in the following examples with the appropriate drive letter. Begin by creating a new directory as follows:

```
C>md host[Enter]
```

The files from each Host distribution diskette can now be copied into the new directory by first placing the diskette in drive A, then issuing the following command:

```
C>copy a:*.* \host[Enter]
```

Creating or Modifying a CONFIG.SYS File

Host requires that certain operating system parameters be set through the use of the CONFIG.SYS file. The purpose and usage of the operating system configuration file CONFIG.SYS is explained in detail in your PC DOS or MS DOS manual. Basically, CONFIG.SYS may contain certain commands which modify the way DOS functions or make it recognize certain foreign hardware devices. DOS searches the root directory of the boot disk for this file upon start-up. If found, its contents are interpreted by the operating system, otherwise DOS operates using default configuration values.

Host requires that the maximum number of concurrently open files permitted by DOS be at least 16. If a CONFIG.SYS file is already present on your DOS boot disk, you need to add the following statement to the file before executing Host:

```
files = 16
```

Using a number larger than 16 will not affect Host, but the memory required by the resident portion of DOS increases by 48 bytes for each additional file above the default value of 8.

The performance of Host (and many other programs as well) will be greatly improved if the following statement is also added to CONFIG.SYS:

```
buffers=24
```

This sets the number of sector buffers to 24, thereby providing some additional temporary storage space to DOS, and reducing the number of disk accesses required for many operations. You may request DOS to use between 1 and 99 buffers, but each buffer above the default value of 2 (3 for the IBM Personal Computer AT) requires 528 additional bytes.

If the file CONFIG.SYS is not on your DOS boot disk, you can easily create it by instructing the DOS COPY command to accept input from the console. The following example illustrates this technique, assuming your DOS boot disk is Drive C:

```
C>copy con config.sys[Enter]
files=16[Enter]
buffers=24[Enter]
^Z[Enter]

   1 File(s) copied
```

The symbol "^Z" is the DOS screen representation of the key combination [Ctrl]+[Z].

Telling DOS How to Find Host

In order to execute Host from a directory other than the one which contains it, you must tell DOS the path to the file HOST.COM. The simplest method is to enter the following command directly from the DOS prompt as shown below:

```
C>path c:\host[Enter]
```

This will work for the current session, but will not be in effect the next time you boot your microcomputer. It is better to include this line in your AUTOEXEC.BAT file. The purpose and usage of the AUTOEXEC.BAT file is explained in detail in your PC/MS DOS manual. Basically, the AUTOEXEC.BAT file may contain any standard DOS commands which you wish to be executed when DOS is started. DOS searches the root directory of the boot disk for this file upon start-up. If found, the commands in the file are executed in sequence. If an AUTOEXEC.BAT file is already present on your boot disk, you must add the following line to the file:

```
path c:\host
```

If a path command is already in the file, the path to the HOST.COM file may be added to the paths already listed. The resulting path command might appear as follows:

```
path c:\dos;c:\util\norton;c:\host
```

If the file AUTOEXEC.BAT is not already on your boot disk, you may easily create it by instructing the DOS COPY command to accept input from the console. This is the same technique used above to create the CONFIG.SYS file. In the following example, your DOS boot disk is assumed to be drive C:

```
C>copy con autoexec.bat[Enter]
path c:\host[Enter]
^Z[Enter]

   1 File(s) copied
```

Appendix B to Chapter 3: Configuring Host (C)

Host can be configured to best meet your equipment and modeling needs. You can use the default configuration or tailor any of the settings to your requirements. The system-wide configuration settings are stored in a Configuration File on disk named SYSTEM.ENV. This file includes the names of certain files, output devices for tables and graphs, printer settings, and screen display type. These settings are kept in the SYSTEM.ENV file so that you do not need to re-configure Host every time it is used. These files and other settings will be used by Host whenever they are required for an operation, unless a different file, subdirectory or other setting is explicitly specified.

When you select the Configure from the Main Menu, you are presented with a work screen that displays all the current settings. If the file containing the configuration settings is missing, Host creates one containing the standard initial settings. If there are settings that you wish to change, you can select that choice on the Configure Menu and make the modification. When you make a selection from the Configuration Menu, the cursor will move to the corresponding field in the screen work space. Each of the possible configuration choices will be discussed below.

When running Host from different subdirectories, it is necessary to have a SYSTEM.ENV file in each subdirectory from which you intend to execute Host.

Default Files

Host uses several types of files for storing different kinds of information (see The Host File System, p. 3-9 of the complete Host User's Guide, available from RTI). Three of these file types, Data, Comparison, and Instruction Files, are assigned names by a Host user when they are created. File extensions for these files are supplied by Host. Host requires the name of a specific file before it can perform any operations with files of these types. To prevent manual entry of file names with each Host operation, the Configuration setup allows default names for these files to be used throughout Host until changed, or until other files are explicitly entered for use in a given execution of an operation.

Data File. This is a file containing input data for performing projections and output data resulting from projections. When Host requires a Data File for an operation, the file selected in the Configuration setup will be the one used unless another Data File is explicitly specified.

Comparison File. This is a Data File which is designed to be used as a "base-run" for comparison purposes in certain tables and graphs. When data from a Comparison File is called for in a table or graph, data from this file will be used unless a different Comparison File is explicitly specified.

Instruction File. This is a file containing instructions to Host for executing projection modules, tables and graphs in a certain sequence. When Host requires the name of a specific Instruction File to perform some operation, this is the file which will be used

unless a different Instruction File is explicitly specified.

Default Output Devices

This selection allows you to set the destination for Tables and Graphs. The Configuration settings allow you to specify printer, console, or plotter output (when applicable). Table and graph output may also be sent to a file, although this cannot be predetermined in the Configuration File. If file output is desired the name file must be specified when the table or graph is generated.

Table Output Device. Initially the cursor will move to the Table setting, and you may choose output to the Printer or to the Console (screen). You can toggle between these two choices using the [+] and [-] keys.

Graph Output Device. You may also determine the output device for Graphs. Graphs may be displayed on three types of devices: the Printer, the Console, or a Plotter. The device type can be toggled between these choices with the [+] and [-] keys.

Default Directories

You can inform Host of the directory or directories where you will keep various Host-related files. This allows you the flexibility of structuring your subdirectories in the manner which is most convenient and useful for your purposes. Initially the current subdirectory is set as the default for all Host files. You may, however, wish to keep certain file types on different disk drives, or isolate specific groups of files in special subdirectories.

A Directory name can be configured for Data files, Module files and Library files. If you wish to modify the default settings for any of these file types simply move the cursor to the desired file type and enter the drive letter (if it will differ from the drive which is the Host default drive), and the hierarchical subdirectory name. If you are not certain how to specify subdirectory names, please refer to the IBM User's Manual, or to the PC-DOS or MS-DOS manual chapters on File Management and Tree-Structured Directories. The three directories which may be specified are presented below with the files which belong in each category.

Data Directory. A directory location may be specified for all Host data files. These include Data Files and Instruction Files. Unless another directory is explicitly entered for a given operation, Host will try to find necessary data files in this directory.

Module Directory. A directory location may be specified for all Host Application Modules. Unless another directory is explicitly entered for a given projection, Host will try to find run locate the necessary module in this directory.

Library Directory. A directory location may be specified for all Host Library files. These include Label Format and Table Definition files. Unless another directory is

explicitly entered for a given operation, Host will try to find necessary Table and Label library file necessary data files in this directory.

Default Printer Settings

You may select the printer settings which are most appropriate for your printer and for the types of reports that you wish to produce. The configuration settings allow you to specify margins and page lengths for reports. The configure file also allows you to specify general printer options for the use of Page ejects (or Form Feeds) and a Setup string. The page layout can be arranged to suit your desired report formats.

Each of the printer settings will be discussed briefly in the following section. The initial default settings assume standard (8 1/2 X 11 inch) paper, 10 characters per inch, and 6 lines per inch.

Left Margin. The left margin setting establishes the amount of blank space that is left on the left side of each page. The default setting of 10 spaces will result in a one inch left margin if your printer is set for 10 characters per inch. This can be set to any reasonable value to provide a more pleasing page layout.

Top Margin. This sets the number of blank lines above the first printed material on a page. The default setting is 6 lines.

Bottom Margin. The bottom margin sets the number of blank lines following the last line of text on a page. The default setting is 6 lines.

Page Length. The page length setting establishes the paper size that you use. Standard paper, (8 1/2 by 11 inches), uses the default setting of 66 lines. If you are using legal size, or any other non-standard size paper, then this should be adjusted accordingly.

Line Length. This value sets the length of the printed portion of a line. For a standard page with a left margin of 10, the default setting of 65 characters should be appropriate. This value will need to be adjusted if you use wide paper, or if you will be using a font type or print pitch that will affect the number of characters that will fit on a line. For example, if you will normally use a compressed print for your printouts you will be typically be able to fit 100 to 120 characters on a line.

Page Eject. Setting the page eject setting to "Yes" causes Host to format printouts to fit on a printed page. If Page Eject is not desired, then the printouts will not be formatted, and will print continuously across page boundaries, and without page numbers or headers. This may be useful to save paper, and to decrease printing time for draft printouts.

Setup String. This is a "String" of characters that is sent to the printer when Host first attempts to print. For most printers, the setup string is optional for normal printing. Establishing a setup string can be useful, however, for setting the printer to use special features, such as particular print sizes or font settings that you wish to use as the

standard mode for Host printouts.

Each printer uses its own printer control codes, so if you use this feature, you will need to refer to your printer manual to determine the appropriate characters to put into Host's Setup String. Host's method indicating characters and Control Characters for the setup string is the same as the method used by Lotus-123. If you have installed a setup string in Lotus for your printer, you can use the exact same string in Host.

The string, if used, should contain the letters or numbers in your printer manual indicated for the printer feature desired. The "ASCII" codes from 0 to 32 are called *Control Characters* and must be entered in a special way. These characters must be entered into the Setup String by entering a [\] (backslash) followed by the three digit code for the character. For example, the [Ctrl]+[G] character, sometimes written "^G" has the ASCII value seven, and would be entered:

\007

Characters with ASCII values above 32 may be entered in the same manner, or may be entered directly as the character associated with the value. For example, the capital letter "Q" has the value 81, so it may be entered as "Q", or as " 81". When entering letters, the case is significant, so follow the printer manual exactly. When forming the Setup String, simply enter each of the characters or backslash-digit combinations without intervening spaces or commas unless separators are indicated in the printer manual.

Default Color Setting

You may select a monitor setting for the type of screen display that you use. The two possible settings are presented as a Yes/No toggle field. If you have a monochrome Monitor, then set the Monochrome setting to Yes. If you have a color display, then set the Monochrome setting to No.

Saving the Configuration Settings

When you have adjusted these options to meet your requirements, select Save from the Configuration Menu, and the settings will be saved in the Configuration File (SYSTEM.ENV) in the current directory and it will they will be used as the new default settings for Host.

Returning to the Main Menu

If the configuration displayed is appropriate for your equipment and modeling plans, simply enter the Quit option and return to the Main Menu. If you have made changes which have not been saved to the Configuration File, you will be asked if you want to save them before returning to Host's Main Menu.

CHAPTER 4. TECHNICAL DESCRIPTION OF THE MODEL METHODOLOGY

The Grade-transition Enrolment Projection Method

Of the many available methods for projecting enrolment, the grade-transition method, while perhaps not the most sophisticated, is a good compromise for developing countries since it does not require as much data as the more difficult methods, yet is useful for policy analysis.

The method, at the elementary school level, is straightforward, and consists of two basic equations. The first equation applies to grade 1 only:

$$E_{1,t+1}=r_{1,t}E_{1,t}+n_{t+1}POP_{entage,t+1}$$

Here $E_{1,t}$ is enrolment in grade 1 at time t, $r_{1,t}$ is the repetition rate in grade 1 at time t, n_{t+1} is the proportion of the population of age entage ("entrance age") that enrols in grade 1 for the first time at time t+1, and $POP_{entage,t+1}$ is the population of age entage at time t+1. All these rates from now on will be collectively referred to as transition rates or parameters.

This equation means that the number of students in grade 1, in any given year t+1, comes from two sources: a proportion $r_{1,t}$ of last year's enrolment in grade 1, $E_{1,t}$, which repeats the grade, and a proportion n_{t+1} of the population $POP_{entage,t+1}$ that enters school for the first time at age entage, in year t+1.

For grades 2-6, the basic equation is:

$$E_{g,t+1}=r_{g,t}E_{1,t}+(1-r_{g-1,t}-d_{g-1,t})E_{g-1,t}$$

where the symbols are analogous to those above, and $d_{g,t}$ is the dropout rate from grade g at time t.

That is, enrolment in year t+1 in any grade g comes from two sources: a proportion $r_{g,t}$ of last year's enrolment $E_{g,t}$, which repeats the same course, and a proportion $(1-r_{g-1,t}-d_{g-1,t})$ of last year's enrolment in the previous grade, $E_{g-1,t}$, that is promoted forward. Note that $(1-r_{g-1,t}-d_{g-1,t})$ is the same thing as the promotion rate.

This methodology can be used with either fixed or varying transition parameters. In either case, it is necessary to project the parameters themselves *before projecting enrolment*.

Projections with Fixed Parameters

Projecting with fixed parameters simply means that $r_{g,t}=r_{g,t-1}$ and $d_{g,t}=d_{g,t-1}$ for all grades g, and for all periods t. When projecting with fixed parameters, two intuitively appealing ways to fix the parameters are:

- at the historical mean, and
- at the last historical level

The computer program allows the user to choose between these. Simple as it may seem, projecting with fixed parameters is often the more correct option (see below).

Projections with Variable Parameters

In projecting with variable parameters, there are several common options:

- projecting with a linear trend
- projecting with a logistic trend
- projecting according to optimistic, medium, and pessimistic targets

Linear Trend. If one uses a linear trend to project the transition parameters, the equation is (using the repetition rate for any grade g as an example, but the method applies to any rate and any grade):

$$r_{g,t} = \begin{cases} a+bt & \text{if } 0 \le a+bt \le 1 \\ 0 & \text{if } a+bt < 0 \\ 1 & \text{if } a+bt > 1 \end{cases}$$

where t is the time period, and the parameters a and b can be estimated using simple linear regression. Of course, if the statistical significance of these parameters, especially b, is questionable (one determines this by at least looking at its t-statistic - ideally the Durbin-Watson statistic should also be checked if available) it would be preferable to use fixed rates.

Logistic Curve Fitting. The problem with linear projections is a familiar one: there is nothing inherent in the method itself to prevent the projected rate from being greater than 1 or less than 0. Thus, it is necessary to impose those (or any other) bounds on it, as we have done above. Since this seems somewhat artificial (specially if the b coefficient is large and the rate is approaching either 1 or 0 fast), fitting a logistic curve, which automatically keeps the rate between 0 and 1, without restraining it artificially as it approaches those bounds, is usually a preferred method. Logistic curves are commonly used to model phenomena that are naturally bound from above and below, like the spread of disease or technology: the proportion of a population with a certain disease or using a certain technology cannot ever be below 0 or above 1. Furthermore, a logistic curve is somewhat "S"-shaped ("sigmoid"): at first, as it departs from the lower (or upper, if negatively sloped) bound, the slope is gentle, then it gets steeper, and as it approaches its upper (or lower, if negatively sloped) bound, the slope gets gentle again. This seems like a good approximation to the time trend of a variable like a repetition rate: at first, it may be hard to reduce it because of lack of experience and ingrained habits. But with success comes learning and more success, so it gets easier. At the end,

it gets difficult again, because the few remaining repeaters are "hard-core". The projecting equation would be:

$$r_{g,t} = \frac{1}{1+e^{a+bt}}$$

where e refers to the mathematical constant e. This implies that in estimating the parameters a and b, simple linear regression can be used, but only after first transforming the rate $r_{g,t}$ as follows:

$$\dot{r}_{g,t} = \ln\left(\frac{1-r_{g,t}}{r_{g,t}}\right)$$

It is this rate which is then regressed on t to obtain a and b. The same cautions about the statistical significance of a, and specially b, apply here.

Target Setting Method. Since these two methods are somewhat mechanical, in that they simply extrapolate current trends, a method based on target- setting is often used. In this case, and assuming a linear approach to the target, the projection equation is

$$r_{g,t} = r_{g,t-1} + \frac{\bar{r}_{g,tt} - r_{g,lht}}{tt - lht}$$

where the bar in $\bar{r}_{g,tt}$ indicates that this is a rate fixed by the user of the method, tt refers to the target period, and lht refers to the last historical period. With this method, the user has to determine both the target rate and the target year. If the statistical projection equations discussed above are untrustworthy, because of the low significance of the slope coefficients b, then the target-setting method might be the more forthright. This method can be used to set optimistic, pessimistic, and medium targets.

Efficiency Indicators Based on Cohort Analysis

Basic Concept

The internal efficiency measures of a school system which are described in the next section are derived from calculating the school history of an artificial cohort of 1000 entrants to the school system and tracking them as they progress (or drop out, or remain) in the school system. The method consists of starting with a matrix of enrolment by grade and sex with 1000 students in grade 1 and nobody in grades 2 - 6. Here we assume a six year primary cycle for exposition purposes. The grade transition model is then repeatedly applied to this enrolment matrix until there is nobody left in the school system, i.e. the initial 1000 children have either dropped out or have graduated.

As the grade transition model is iteratively applied, the number of students remaining in the system is tracked as well as the number of iterations. The number of iterations is conceptually equivalent to the amount of time the students remaining in the system have spent in school. Hence, if 100 of the 1000 entrants graduate "on time" after

six years (i.e. they did not repeat), then they would be counted as graduates after six iterations of the method. Those members of the cohort who repeated a grade will obviously remain in the system for longer than six years and it will require more than six iterations for them to graduate, if at all.

In doing these calculations the number of student/years spent in the system for various categories of students is also computed. Thus the 100 "on time" graduates spent 600 student/ years in the system. If 100 more students took seven years to graduate and no one else graduates, then the second group spent 700 years in the system and the total of 200 graduates spent 1300 student/years or an average of 6.5 years in the system per graduate. Similar calculations can be applied to students who leave the system without graduating.

Note that the method depends only on the drop out and repeater rates. For this reason the assumptions underlying the future transition rates during the future school history of the cohort are important. There are two basic methods which can be used here.

In the first method, all transition rates are assumed constant and equal to those during the entry year. In other words, the method projects the future history of a cohort entering school in 1986 according to the transition rates in 1986, and those rates are assumed to remain at 1986 levels.

In the second method, the transition rates change according to the actual "year" in the future projection of the cohort of entrants. In other words a cohort of 1000 entrants in 1986 would be subject to the transition rates in 1986 during the first iteration or "year", to the 1987 rates during the second iteration, to the 1988 rates during the third iteration, etc. If drop out rates are expected to fall over time, the second method would produce completion rates which are consistently higher and growing faster than those produced by the first approach.

Efficiency and Other Educational Indicators

Completion or Graduation Rate. Perhaps the most used, and certainly the simplest measure, this is the number of graduates per 1000 entrants.

"On Time" Graduates This is the number of graduates per 1000 entrants who complete the prescribed cycle within the prescribed number of years.

Student/Years per graduate This is average amount of time spent in a school system per graduate. There are two methods for calculating this. In the first method, the number of student/years spent in the system by all graduates is divided by the number of graduates. In the second method, the number of student/years spent both by graduates and by drop-outs is divided by the number of graduates. In a six year system with repetition and drop out rates equal to zero these would be equal to six.

Input-output ratio. A problem with the above measure is that it does not take into account the length of a cycle. A six year school system will require ideally six student/years per graduate which, from an efficiency standpoint should be no less efficient than an eight year system requiring ideally eight years. The input-output ratio standardizes for the length of the cycle by dividing the calculation for student/years per graduate by the length of the cycle. This ratio is ideally equal to one in any system.

Years Spent in School This is the average number of years spent in school by all of the 1000 entrants before leaving the system. It is the average of the student/years per graduate and and student/years per drop out, weighted respectively by the number of graduates and the number of drop outs. Note that the number of drop outs and the number of graduates must sum to 1000.

Indicators Based on Enrolment Data

Gross Enrolment Ratio This is the ratio of enrolment in primary school, in all age groups, divided by the population in the "legal schooling" age groups, usually 6-14 or 7-15, or thereabouts. Since the ratio is of actual or projected enrolment to actual or projected population, it has to be derived either from historical data or from an enrolment projection. It is not based on the analysis of an idealized cohort.

Net Enrolment Ratio. This is a similar concept to the one above, except that the numerator is not all students, but only those in the same age groups as the population. Thus, when compared with the Gross Enrolment Ratio, it is an indicator of the proportion of "overage" students in the system. It is important to note that this "overage" might be only in a legal sense: if a perfectly efficient system, with little repetition or dropouts, was unable to absorb the population right at the legal starting age, then the difference between the Net Enrolment Ratio and the Gross Enrolment Ratio would simply be a reflection of this fact, rather than a measure of inefficiency or wastage. In this sense, it differs from measures of "overage" derived from cohort analysis, such as the number of graduates not on time.

Estimation of Transition Rates

Standard Method

The simplest standard method of estimating the repetition and dropout rates involves using the official data on numbers of repeaters and enrolment to get the repetition rates, and solving the grade-transition equations for the dropout rates, using historical enrolment data and the previously estimated repetition rates.

The repetition rates are estimated as follows:

$$r_{g,t} = \frac{R_{g,t+1}}{E_{g,t}}$$

where $R_{g,t}$ is the number of repeaters in grade g, period t. Evidently, what we have here is just the ratio of repeaters to enrolment in the same grade the previous year. From this, we can go on to solve the grade-transition equations.

As we saw above, the standard equation of the grade-transition method for grades higher than 1 is:

$$E_{g,t+1}=r_{g,t}E_{1,t}+(1-r_{g-1,t}-d_{g-1,t})E_{g-1,t}$$

Solving $d_{g-1,t}$ for, we get

$$d_{g-1,t}=1-\frac{(E_{g,t+1}-r_{g,t}E_{g,t})}{E_{g-1,t}}-r_{g-1,t}$$

which is just 1 minus the promotion rate minus the repetition rate. Using this method we can solve for all the repetition and dropout rates. Note that the method implies that it will be easier to make the calculations if all repetition rates are calculated before the dropout rates, since to calculate the dropout rate for grade g we need to know the repetition rate not only for grade g, but also for grade g+1. Also note that the dropout rate implied in this method is the beginning-of-year to beginning-of-next-year dropout rate, i.e., it includes all those who abandon the school system between grades. Finally, it includes all those who die. This makes it unnecessary to refer to a death rate in making the enrolment projections.

Schiefelbein Method - Description

The motivation for the Schiefelbein* method is the suspicion, confirmed in some cases, that the numbers of repeaters in most developing countries are under-reported. The reasons given for this are many:

- repetition is a social stigma, so children, when asked by a new teacher whether they are repeaters, are not likely to be 100% truthful
- some children who appear as new in a classroom are actually repeaters who have migrated from another region or school
- children may repeat a course even if they have not officially failed the previous one

It is easy to see from the equations for the standard method that, under these circumstances, the repetition rates will tend to be under-estimated, and the dropout rates consequently over- estimated.

* See Schiefelbein, Ernesto, "Statistical Report on Repetition in Latin America," report prepared for UNESCO, Division of Statistics on Education, Office of Statistics, Paris, 1980.

The basic idea behind the Schiefelbein method is that there has to be a certain mathematical structure that determines the differences between two consecutive tables of enrolment by age and grade. This underlying mathematical structure must have age-specific repetition and dropout rates as its parameters. From the data on enrolment by age and grade for two consecutive years it may not be possible to uniquely identify a set of age-specific repetition and dropout rates, but, with certain unrestrictive assumptions, one can at least obtain a range within which the overall (i.e., not age-specific) rates MUST lie, given the age-grade enrolment data and the underlying mathematical structure.

The method consists of the following. We assume a primary school system consisting of six grades, an entrance age of 7, and a maximum school age of 23. These assumptions are for ease of exposition, but the method is easily generalizable. Also, the method can be used to calculate rates disaggregated by sex, if the enrolment data are available by sex. In the discussion below we abstract from this to keep things a little simpler.

We start with age-grade enrolment data for two consecutive years. Thus, we have $E_{g,a,t}$ and $E_{g,a,t-1}$, where g stands for grade, a stands for age, and t stands for time.

The calculations then proceed grade by grade, for each pair of consecutive ages. For any age group a, therefore, we first make the calculations for grade 6. We start by calculating the number of repeaters of age a, in grade 6 at time t:

$$R_{6,a,t}=r_6E_{6,a-1,t-1}$$

where r_6 is an assumed value. From here we can then go on to calculate the following. The total number of students promoted from grade 5 to grade 6 has to be:

$$P_{6,a,t}=\begin{cases}E_{6,a,t}-R_{6,a,t} & \text{if } E_{5,a-1,t-1}+R_{6,a,t}>E_{6,a,t}\\ E_{5,a-1,t-1} & \text{otherwise}\end{cases}$$

In words: if the sum of enrolment in grade 5 last year, plus those repeating grade 6 this year, is greater than enrolment in grade 6 this year, then we have to put a cap on how many students could have been promoted from grade 5 to grade 6. This cap is enrolment in grade 6 this year, minus repeaters in grade 6 this year. But if the sum of enrolment in grade 5 last year, plus those repeating grade 6 this year, is less than enrolment in grade 6 this year, then we have to promote everyone who was in grade 5 last year, unless we are ready to admit that there were a lot of students new to grade 6, which would not be logical.

The number of new students in grade 6* will then be:

$$N_{6,a,t}=\begin{cases}E_{6,a,t}-P_{6,a,t}-R_{6,a,t} & \text{if } E_{5,a-1,t-1}+R_{6,a,t}\leq E_{6,a,t}\\ 0 & \text{otherwise}\end{cases}$$

* Note: for national-level projections the only new students in grade 6 can be international

which follows from the determination of the number of promoted: if there were not enough people in grade 5 last year, such that the sum of enrolment in grade 5 last year, plus repeaters in grade 6 this year is less than enrolment in grade 6 this year, then there must have been some new students in grade 6 to make up the difference.

Having done the calculations for grade 6, we can then proceed with the earlier grades, still working with the same pair of consecutive age groups. The number of non-promoted students, that is, students who will either repeat any given grade g or drop out, will be:

$$NP_{g,a,t} = E_{g,a-1,t-1} - P_{g+1,a,t}$$

That is, the number of non-promoted from grade g is equal to the enrolment in that grade, minus those who were promoted to a higher grade. (Note that this implies that the calculation has to proceed from the higher grades to the lower ones.)

The number of repeaters is calculated according to:

$$R_{g,a,t} = \begin{cases} (1-\text{dropprop})NP_{g,a,t} & \text{if } NP_{g,a,t} \leq E_{g,a,t} \\ E_{g,a,t} & \text{otherwise} \end{cases}$$

where the proportion dropprop determines the proportion of non-promoted students who drop out. The number of repeaters in grade g is equal to a proportion of the non-promoted if the pool of non-promoted is less than enrolment, but if there are more non-promoted students from the previous grade than enrolment in grade g, then the maximum number of repeaters is set equal to enrolment. Note that the proportion dropprop plays a role, but only under certain conditions (the first condition in the equation above). The smaller the number of cases where those conditions are met, the less dependent on the particular assumed value of dropprop will be the results.

The number of students promoted to grade g will be:

$$P_{g,a,t} = \begin{cases} E_{g,a,t} - R_{g,a,t} & \text{if } E_{g-1,a-1,t-1} + R_{g,a,t} \geq E_{g,a,t} \\ E_{g-1,a-1,t-1} & \text{otherwise} \end{cases}$$

The number of new students in grade g will be:

$$N_{g,a,t} = \begin{cases} E_{g,a,t} - P_{g,a,t} - R_{g,a,t} & \text{if } E_{g-1,a-1,t-1} + R_{g,a,t} \leq E_{g,a,t} \\ 0 & \text{otherwise} \end{cases}$$

That is, if the pool of enrolment in the previous grade last year, plus repeaters of grade g, is larger than enrolment in grade g this year, then the number promoted has to be limited, and the limit is set to equal enrolment in grade g minus repeaters of grade g. But if the the number of repeaters of grade g plus enrolment in grade g-1 last year is smaller than enrolment in grade g this year, then we set the number promoted to equal

migrants. In regional projections, they may be inter-regional migrants.

the number enroled in g-1 last year.

And, finally, the number of dropouts can be calculated as:

$$D_{g,a,t}=NP_{g,a,t}-R_{g,a,t}$$

or the difference between the total that does not go on to grade g+1 and those who repeat grade g.

The equations for grade 1 are slightly different only because there is no previous grade. We list them for completeness only.

$$NP_{1,a,t}=E_{1,a-1,t-1}-P_{2,a,t}$$

$$R_{1,a,t}=\begin{cases}(1-\text{dropprop})NP_{1,a,t} & \text{if } NP_{1,a,t}\leq E_{1,a,t}\\ E_{1,a,t} & \text{otherwise}\end{cases}$$

$$P_{1,a,t}=0$$

(since there cannot be anyone promoted to grade 1)

$$N_{1,a,t}=\begin{cases}E_{1,a,t}-R_{1,a,t} & \text{if } R_{1,a,t}\leq E_{1,a,t}\\ 0 & \text{otherwise}\end{cases}$$

$$D_{1,a,t}=NP_{1,a,t}-R_{1,a,t}$$

After having calculated all these flows, we finally get the transition parameters as follows:

$$r_{g,t}=\frac{\sum_{a=5}^{23}R_{g,a,t+1}}{\sum_{a=5}^{23}E_{g,a-1,t}}$$

$$d_{g,t}=\frac{\sum_{a=5}^{23}D_{g,a,t+1}}{\sum_{a=5}^{23}E_{g,a-1,t}}$$

$$n_t = \frac{\sum_{a=5}^{23} N_{1,a,t}}{POP_{entage,t}}$$

Thus, the grade-specific transition rates are in fact weighted averages of the age-and-grade-specific transition rates. The presumption is that this averaging process will smooth out some of the irregularities inherent to the methodology, which arise because the data are not enough to uniquely identify all the age-specific transition rates.

Schiefelbein Method - Example of Calculations

Since the method is somewhat involved, a numerical example might help to understand the calculations required.

The following example uses data from Guatemala, from the years 1982 and 1983, for the age groups 11 (in 1982) and 12 (in 1983).

Grade	$E_{11,1982}$	P	NP	R	D	N	$E_{12,1983}$
First	17737		6477	6348	130	8299	14647
Second	21605	11260	6202	6078	124	0	17338
Third	24863	15403	5613	5500	112	0	20903
Fourth	23167	19250	2224	2180	44	0	21430
Fifth	19964	20493	457	448	9	0	21391
Sixth	6044	19507		302	0	0	19809

We start with the number of repeaters in grade 6, 302, which is 6044 multiplied by the assumed grade 6 repetition rate (0.05 here). This appears in the R column. Since we have 302 students repeating grade 6, and since 19964+302>19809, we cannot promote everyone that was in grade 5 in 1982. We therefore only promote 19809-302=19507. They appear in the P column. If grade 5 enrolment in 1982 had been, say, 19000, then, since 19000+302<19809, we would have promoted all 19000, and the difference between 19809 and 19302, which is 506, would have had to be new students. By promoting everyone when we have a situation like 19000+302<19809 we minimize the need to suppose there are new students entering grade 6.

Since 19507 of the 19964 students in grade 5 in 1982 were promoted, 457 (=19964-19507) must not have been promoted, so they appear in the NP column. These 457 students are split into repeaters and dropouts according to the assumed ratio dropprop, mentioned above, but only if the number of non-promoted is smaller then enrolment in grade 6 in 1983. Since 457 is indeed smaller than 21391, then the ratio is used in this case. Had we had, say, 23000 non-promoted (from grade 5) students, then only 21391 could have been repeaters, and the rest would have been dropouts. Now we can calculate the number promoted to grade 5. Since the enrolment pool in grade 4 in 1982, plus the number repeating grade 5 in 1983, is sufficient to supply the 22229 in grade 5 in

1983 (23167+448>21391), we need to cap the number that were promoted. Only 20493 (=21391-448) could have been promoted. The number of new students is 0, for grade 5 in 1983, because enrolment in grade 4 in 1982 (23167), plus those repeating grade 5 (448), was sufficient to supply the 21391 that were enroled in grade 5 in 1983 (since 23167+448>21391).

The process continues upward until we complete this table. Similar tables have to be completed for all consecutive age groups. This gives us total numbers of repeaters and dropouts by age. We aggregate these, and aggregate the enrolment as well, divide the repeaters and dropouts by enrolment, and get the repetition and dropout rates.

Appendix to Chapter 4 - Glossary of variables and symbols

The following table shows, for each variable, its symbol in this technical description, its name in the user's manual or in the computer program, and a brief description of its meaning. The variables are listed in the alphabetical order of their symbols.

Symbol	Name	Meaning
$d_{g,t}$	DROPRT	dropout rate grade g, time t
dropprop	DROPPROP	assumed share of dropouts in non-promoted students
$E_{g,a,t}$	ENROLAGE	enrolment in grade g, age a, at time t
$E_{g,t}$	ENROL	enrolment grade g, period t
endage	ENDAGE	oldest "legal" school age
entage	ENTAGE	age of entry into schooling
n_t	ENTRT	rate of entry from pop. to 1st grade
$POP_{a,t}$	POP	population of age a at time t
$\bar{r}_{g,tt}$	TARREP	target repetition rate
$R_{g,t}$	NREPEAT	repeaters, grade g, time t
$r_{g,t}$	REPEATRT	repeater rate grade g, time t
tt	TARYEAR	target time period
None	AGGENROL	enrolment in the aggregated age group
None	ASER	age-specific enrolment rate
None	GER	gross enrolment ratio
None	GRADRATE	number who graduate out of 1000 entrants
None	HISTORIC	indicator var.: historic=h, projected=p
None	INCLUDE	indicator var.: if include this year, include=y, if not, include=n.
None	IORATIO	input-output ratio: student years per graduate divided by the nominal school cycle length
None	ONTIME	number of "on-time" graduates,i.e., have not repeated
None	SCHLPOP	total (aggregated across ages) population of school age
None	SYGRAD	student/years per graduate not counting student/years "wasted" by drop-outs
None	SYGRAD2	student/years per graduate, counting student/years "wasted" by drop-outs
None	TARDROP	target dropout rate
None	TDROP	t-statistic for drop. rate regr. slope
None	TOTENROL	total (aggregated across grades) enrolment in primary school
None	TREP	t-statistic for rep. rate regr. slope
None	YRSSCHL	average no. of yrs. spent in school

The following variables are used only for intermediate calculations, and so the user need not be concerned about them. Within the program, some of these variables are

simply various columns of the matrix "flowmatrix".

Symbol	Name	Meaning
a	various	regression intercepts
b	various	regression slopes
$D_{g,a,t}$	flowmatrix	dropouts from grade g, of age a, at time t
lht	lasthist	last historical period
$N_{g,a,t}$	flowmatrix	new students in grade g, of age a, at time t
$NP_{g,a,t}$	flowmatrix	students not promoted from grade g, of age a, at time t
$P_{g,a,t}$	flowmatrix	students promoted to grade g, of age a, at time t
$R_{g,a,t}$	flowmatrix	repeaters of grade g, of age a, at time t
t	t	time period

CHAPTER 5. A STEP-BY-STEP SUMMARY

This is the chapter that must be read even if nothing else in this manual is read. It provides a step-by-step, or "recipe" guide to running PETS for a set of limited scenarios. Since PETS offers so many options, and since some of them involve some complex processes, a "recipe" approach is both dangerous and limiting. It is dangerous because it makes it possible to use the system without really understanding it. This can lead the user to unwarranted conclusions and shaky projections. And it is limiting because, by its nature, a "recipe" approach cannot cover all the options. So the reader is urged either to read the rest of the manual at some point, even if at first he/she reads only this part. A reasonable approach for the impatient or for those with some computer background would be to read this chapter first, experiment extensively with the system, and then read the rest of the manual before making serious use of the system.

The order in which this chapter is written is the recommended order for running the system, but it is not mandatory. As you learn about it, you can experiment and run it in different ways. Initially, however, follow the recommendations here.

Installing Host and PETS

This section describes how to install Host and PETS on a machine with a hard disk. It is fairly self-sufficient, but if you are a complete beginner in the use of PC DOS in machines with a hard disk, it may be wise to seek help from someone with some knowledge in this area. All this need be done only once, if the system is to be left installed in the hard disk. Of course, if you remove it from the hard disk (in order to save space - and this may be desirable, specially if you plan to use the system only for a week or so every year), then you will have to re-install it next time you want to use it. In this section it is assumed that the hard disk is drive C, but this is for illustrative purposes only.

Before you try to install Host and PETS, be sure that you have the following equipment. If you are not sure, one way to find out is to ask your in-office expert. Alternatively, you can run the first part of the routine included in the diagnostic diskette that came with your computer.

- IBM PC, XT, AT or compatible with 512 Kbytes of RAM
- Hard disk with at least 2 Mbytes free space
- 8087 or equivalent numeric co-processor
- PC DOS 2.0 or above or equivalent

To check on whether your version of PC DOS is the correct one, issue the command

C>ver[Enter]

On the diskette that your machine will boot from, or in the root directory of the hard disk, there has to be a file called CONFIG.SYS that needs to have the lines

```
files=16
buffers=24
```

To check if this file already exists, use the DOS command

```
C>dir config.sys[Enter]
```

If it does not exist (make sure! - the following process will erase any that already existed), the easiest way to create it is to issue the following commands:

```
C>copy con config.sys[Enter]
files=16[Enter]
buffers=24[Enter]
^Z[Enter]
```

where ^Z means pressing the keys [Ctrl] and [z] together. Note: the machine will not repond until the line with ^Z is entered. These instructions assume that the commands are issued from drive C, and that the file is being created in the root directory of drive C. If the file is to be created while the user is in drive C, but is to be put in a diskette in drive A, then the first line of the commands above would read:

```
C>copy con a:config.sys[Enter]
```

If the file config.sys exists, check its contents with the command

```
C>type config.sys[Enter]
```

If it does not contain the lines noted above, or it does not allow for at least as many files (16) and buffers (24), you can modify it with any editor (such as Wordstar, in the non-document mode) so that it contains those lines. Instructions on how to use an editor to modify this file is beyond the scope of this manual. Check the manual for whatever editor you choose to use, or have someone make these changes for you.

After you have created this file, or modified its contents, you have to re-boot the computer with a "warm-boot" procedure. For the great majority of IBM-compatible machines, this involves pressing the keys [Ctrl], [Alt], and [Del] at the same time. The easiest way is to depress, and keep depressed, the [Ctrl] and [Alt] keys, and then just tap the [Del] key.

Next, put all the Host and PETS files in a directory. We will assume you will call the directory "pets", but this is for illustration only. To create the directory, issue the command

```
C>md pets[Enter]
```

If for some reason the directory already exists, the computer will not be able to create it again. In that case, you may skip this step.

Now, make "pets" the default directory by entering:

C>cd pets[Enter]

The "cd" command will from now on be referred to as "change directory". Now put the first Host diskette in drive A, and enter

C>copy a:*.* /v[Enter]

Then do the same with all the other Host diskettes, and all the PETS diskettes.

Note: if some previous versions of Host and/or PETS were already installed in your machine, you might wish to erase all their files, to make room for the new version. To erase all the files in a given directory (called "oldpets" for illustration), first you must make that directory the default directory. If you are still within the "pets" directory, you can go to the "oldpets" directory by issuing the command

C>cd \oldpets[Enter]

The first command above takes you to the root directory, while the second makes "oldpets" the default directory.

Since the next command will erase all files in a directory, and you wouldn't want to do that in the wrong directory, make sure you are indeed in the "oldpets" directory with the command

c>dir /w[Enter]

The second line the computer puts on the screen after you hit [Enter] will tell you what directory you are in. If there are a lot of files in this directory the name of the directory will fly by somewhat quickly, so look carefully.

To erase all files in the "oldpets" directory, issue the command

C>erase *.*[Enter]

Now go back to the "pets" directory by issuing the command

C>cd \pets[Enter]

Configuring Host to run PETS

This section explains how to get into Host in order to configure the system to run PETS. The configuration procedure should be necessary only once. All the configuration information will be saved under a file called "system.env". Thus, if you remove the system from your hard disk, but make sure to back up the "system.env" file to a floppy (ideally one of the floppies containing the PETS system), then you should be able to copy this file back next time you install the system.

First, some general pointers about Host. Host is a menu-oriented data-base manager for modeling uses. Its key functions allow the programmer to implement rather complicated inter-linked models with relative ease. In Host, most keyboard keys work as in other popular software. For example, the [/] key calls the menu, as in Lotus

1-2-3. The [Esc] key takes you one level up in the menu structure. In almost all cases, pressing the [Esc] key repeatedly will take one back to the Main Menu. This is a trick worth remembering! Note that pressing [Esc] in certain situations is not the best procedure. For example, if you are entering or changing the value of a variable or parameter, and you press [Esc] without first pressing [Enter], you will go back to the Menu, but the machine may not accept the value you have just input. In these cases, you should instead press [/] to go back to the menu. The best way to develop a feel for how these things work is simply to experiment and practice. When editing, the [Backspace] key erases one character to the left, whereas the [Del] key deletes the character where the blinking cursor is. The [Ins] key turns the insert mode on, and the overwrite mode off, and vice-versa. Finally, the cursor (arrow) keys work just as one would expect: you can use them to choose items in the menu by moving to the right or to the left, and you can use them to go to different areas of the editor screen by moving right, left, up, or down. Finally, you can leave Host by pressing [q] for Quit, from the Main Menu.

To enter the system, make "pets" the default directory, as explained above, unless you are already in it. The easiest way to make sure is just to issue the command

C>cd \pets[Enter]

Once in the "pets" directory, go into Host simply by entering

C>host[Enter]

After an initial introductory screen, and after asking you to press any key, the machine will ask you whether you wish to enter as a regular user or as a manager. To enter as a user simply press the [u] key.

You will now be in the Main Menu. Choose the Configure option by pressing [c]. You will now be in the Configure Menu. Choose [f] to tell Host what the default files will be. The only important type is the default data file. To find out what your choices are, press [Shift] and, while holding [Shift], hit [?]. Point to the desired file using the cursor and press [Enter]. Initially, there will be only one file, so you will not have any problem choosing. After you have added files to the system, you will be responsible for remembering which files contain what, and for re-setting the default file accordingly. Press the [/] key to go back to the Configure Menu. Now press [o] to select the output devices. The important device in this case is the one where the tables will go. Generally, you will want to send the tables to the console. Thus, press the space bar until you get the word Console as the Tables Output Device. Then press [/] again to get the Configure Menu. The next step is to press [d] to set the directories. In the example we are working with, they should all be set to \pets. This can be accomplished simply by pressing the [\] key and entering the word pets in the first line. Go down one line with the cursor, and enter \pets again. And so forth until all four lines are completed. Note that even though you enter \pets, the machine will display \pets\. When you are finished, enter [/] to go back to the configure menu. You are now finished with the

essentials. NOTE: if the Configure screen already displays all the options listed here, then you need not re-enter them.

If you have a color monitor and prefer to work in color, press [m] to set the monitor options. Press the space bar until you set the Monochrome option to No. Then press [/] to go back to the Configure Menu.

At this point you should save your configuration by pressing [s]. Then press [q] to Quit. You will then be in the Main Menu again. All these options will be saved in a file called "system.env". Therefore, unless you want to change some of them, it should be necessary to go through this whole process only once.

Entering or Editing Data.

In this section we will go through the steps most commonly required to add a new year of data, or modify existing data. We will assume that a new year of data of enrolment by age and grade, and possibly of numbers of repeaters, has become available, or data for and old year have been changed, and we want to add it to the forecast base.

First, if you have not done so already, get into Host by issuing the command

C>host[Enter]

Enter as a user by responding with a [u] when asked.

To use the editor, press [e] when you get to the Main Menu. You will be asked to select a file to edit. The choice of files will be put up on the screen. Point to the one you wish (it should be the same one that was entered in the Configuration process above) with the cursor (arrow) keys, and press [Enter].

Now you will be prompted to enter a variable name. Enter the word "enrolage" for enrolment by age and grade, and press [Enter]. If you make a mistake entering this word, simply re-enter it; or press [Shift] and [?], and then point to the desired variable with the cursor (arrow) key and pick it by pressing the [Enter] key. The first year of data for the variable "enrolage" will now appear on the screen. Use the [F4] key to move forward in time until you get to the year you want. If you go too far, use [F3] to go back. Now enter the new data by typing the appropriate numbers in the appropriate places. You may overwrite any existing data. If the computer will not overwrite, try hitting the [Ins] key once and keep trying to enter the data. You can move around by using the cursor keys. After entering each number, for example, move down to the next one by pressing the down arrow. If you make a mistake while entering a number, and before moving to a new one, you can erase by using the Backspace key. If you notice an error in a number after having moved to another one, you can correct it by moving back to it and simply re-entering it over the old one. After finishing with the new year of data, enter [/] to go back to the Edit Menu. Then press [s] to save the data. The computer will suggest saving the variable under its own name. To say yes, simply press [Enter].

Then press [v] to select a new variable. Choose the variable "aggenrol" (aggregate group enrolment). If you make a mistake, re-enter the name, or use the [Shift] and [?] to get a list of variables from which you can pick "aggenrol" with the cursor and [Enter]. Use the [F4] and [f3] keys to move back and forth in time as needed. When you get to the year you want, enter here the data for the last, or aggregated, age group. Make sure that this is for the same age group as in previous years. That is, if in previous years the aggregated age group is 14 and higher, for the new year the aggregated age group cannot be 13 and higher, or 15 and higher, but must be 14 and higher as before. After entering these data, press [/] to get the Edit Menu, press [s] and [Enter] to save the variable under its own name, and [v] to get a new variable.

If you do not want to input new data on numbers of repeaters, then skip this paragraph. If you wish to input data on numbers of repeaters, at this point respond "nrepeat". Again, if you make a mistake, re-enter the name, or press [Shift] and [?], and pick the variable name with the cursor and [Enter]. Go through the same steps as for "enrolage" above, including saving with [s] and choosing a new variable with [v].

Either after entering new data on enrolment by age and grade ("enrolage") and aggregate age group enrolment ("aggenrol") or after entering data on numbers of repeaters ("nrepeat"), you will have asked for a new variable. If all you have done up to now is simply to fill in gaps in the data, or changed old values, you can skip this paragraph. But if you have actually added a year (or more) of data at the end of the series, you will need to tell this to the machine. Enter "historic" to choose the variable that tells the machine that there is a new year's worth of data. Then use the cursor keys to move to the recently entered years, and change the value of historic for each entered year from "p" to "h". Then press [/] to go to the Edit Menu, [s], and [Enter] to save the variable "historic" under its own name. Press [v] to select a new variable.

Whether or not you edited the variable "historic", you will have asked for a new variable. Enter "include" to choose the new variable. Go to the year BEFORE the last of the recently entered or changed year (or years) and, if "include" was set to "n", change the value of "include" from "n" to "y" for that year. For example, if you just entered enrolment data for 1985 and 1986, the value of "include" for 1984 and 1985 should be changed from "n" to "y". NOTE: if by entering new data you have closed gaps in the historical data, then all historical periods, except for the last, should have the the "include" variable set to "y". This is explained in full detail in Chapter 2. Then press [/], [s], and [Enter] to save the variable "include".

Finally, press [q] to quit the editor and go back to the main menu.

Running the ESTIMATE module.

The next step is to re-estimate the historical transition rates according to the new information. From the main menu, choose [r] for Run. Enter the word "estimate", or press [Shift] and [?], and pick the ESTIMATE module with the cursor and [Enter].

You will be presented with a series of parameters that control how ESTIMATE is to be run. The only one that will regularly need editing is the number of periods. Press [e] for Edit, and use the cursor (arrow) key to go down to the slot for the number of periods. Enter a number sufficiently large to cover all the existing historical periods. For example, if the first period is 1974 and the last historical period is 1985 enter at least 12, just to be sure (since 1985-1974=11). After you have entered the number, press [/] to get the menu again. Choose [g] for Go.

Since your data will now be altered in a significant way, the machine will ask you whether you wish to make a backup. If you have made a backup after your most recent session of significant data entry respond [n] for No. Otherwise, respond [y] for Yes. If you respond No, skip the rest of this paragraph; otherwise, read on. You will now be prompted for a name for your backup. You can respond, for example, "bu" for "back-up". Enter this name and press [Enter]. If the file already exists, the machine will ask you whether you wish to overwrite it. If the old backup file contains some important information, respond [n] for No. You will then be prompted for a new name, which you can give at this point. An appropriate name would be, for example, "buy2". Note that multiple backups take up a lot of disk space, so do this only sparingly. If the old backup file contains nothing you wish to save, then when asked if you wish to overwrite it, respond [y].

After this point your computer will be under the control of the ESTIMATE module. There will be three opening screens reminding you of where you are, what data the module needs, and what results you can expect. Press [Enter] after reading each one. Then you will be asked if all the data are present. Since you have just entered new data, and since the data for the variable "pop" (population) and "entage" (age of entrance to the school system) should already be in the system from past runs, you can respond with a 1 when asked if all data are present. If for some reason you want to change some of the population data, or have some other doubts, respond with a 2, and go back to the Main Menu, from which you can choose the Edit option, and proceed with the "pop" variable as for the "enrolage" variable as explained above. If the module goes ahead and runs without the required data, usually no permanent damage is done. Simply re-run it after entering the correct data.

VERY IMPORTANT: the module ESTIMATE expects to find some data, even if they are all zeroes, in the variables it uses, such as "enrolage", even in periods for which there are no data. If you get an error message while trying to run ESTIMATE saying that data could not be found in a certain period, that is a likely cause. Go back to the editor, as explained in the previous section, choose the name of the variable causing the problem, and for every period (in the historical set of periods) for which the machine says "Data could not be found...", perform the following steps: press [s], [Enter], [r] for resume, and [F4] to move forward one period, until either you are out of the set of historical periods, or all the historical periods have real data or zero data.

From this point on, the ESTIMATE module is self-explanatory. To the extent that it is not, it is too complicated for a "recipe" explanation. For example, this is not the place to explain the difference between the Schiefelbein vs. the standard procedure for estimating transition rates. If you want more information on the ESTIMATE module, go to Chapters 2 and 4. At the end of ESTIMATE, you will be asked if you want to calculate age-specific enrolment rates. Since you have just added new data on enrolment by age and grade, respond "yes". After the module has finished, it will present you with a screen saying it has successfully executed. Press [Enter]. The machine will ask you whether you wish to run another module. Respond [y] for Yes.

Running the PROJRATE module.

The next step is to re-project the transition rates to the future, based on the new historical information on transition rates, as estimated with ESTIMATE. The process is very similar, so the explanation of many of the steps is exactly the same.

Since at the end of ESTIMATE you chose to run another module, you will be prompted for the name of the module. Enter "projrate", or use [Shift] and [?] to pick PROJRATE as you should now know. You will be presented with a series of parameters that control how PROJRATE is to be run. The only one that will regularly need editing is the number of periods. Press [e] for Edit, and use the cursor key to go down to the space of the number of periods. Enter a number sufficiently large to cover all the future periods you want. For example, if the first historical period is 1974 and the last future period you want is 1990, enter at least 18 (since 1990-1974=16), just to be sure. After you have entered the number, press [/] to get the menu again. Choose [g] for Go. Since your data will now be altered in a significant way, the machine will ask you whether you wish to make a backup. Since you have just made a backup before running ESTIMATE, you probably do not want one again. After all, re-running ESTIMATE is not very time-consuming. There will be three opening screens reminding you of where you are, what data the module needs, and what results you can expect. Press [Enter] after reading each one. Then you will be asked if all the data are complete. Since the machine has just calculated the historical values of the necessary rates, and since you have entered new data for the variables "historic" and "include", you can respond with a 1 when asked if all data are present. From this point on, the PROJRATE module is self-explanatory. To the extent that it is not, it is too complicated for a "recipe" explanation. For example, this is not the place to explain the difference between the logistic vs. linear methods for projecting the rates. If you want more information on the PROJRATE module, go to Chapters 2 and 4. After the module has finished, it will present you with as screen saying it has successfully executed. Press [Enter]. The machine will ask you whether you wish to run another application module. Respond [n] for No.

Examining the projected rates with Tabulate.

At this point, you should examine the rates that were projected with PROJRATE. This is most easily done with the Host Tabulate facility. You should be in the Main Menu at this point. Press [t] for Tabulate. The machine will ask for a table name. Enter "rates", or pick it out by pressing [Shift] and [?], and pointing to it with the cursor and pressing [Enter]. The parameters that will control the tabulation will be displayed. To change them, press [e] for Edit. There are only two that might need changing. 1.) If you do not want the table to go to the monitor ("console"), but want it to go to the printer instead, move the cursor to the appropriate position and change "con" to "prn" (for "printer"). Note that if there is no printer connected to the machine, or if the printer is turned off, you may get locked out, and the only way back will be to warm boot the computer. 2.) The other parameter that needs to be set is the number of periods to put in the table. If your base year is 1974, for example, and you have projected to the year 2000, move to the slot for "periods" and enter "1-17" (for 1974-2000). You are now set to go. Press [/] to go to the Menu, and then press [g] to Go. If you sent the table to the printer, the whole table will be printed. If it is going to the monitor (con for "console"), the computer will put one page at a time on the screen. To get a new page, press [PgDn] (the key for 3 on your numeric keypad). You can go back up by pressing [PgUp] (the key for 9 on the numeric keypad). If these keys do not respond, it may be because the [NumLock] key (above the 8 in the numeric keypad) is set wrong. Press [NumLock] once and try again.

By moving up and down you should be able to check all the projected rates. You should examine these carefully and make sure the projections seem reasonable. (Note: the rates may be discouraging but still reasonable. By an unreasonable projection we mean, for example, one where a repetition rate is projected to go from 15% to 60% in ten years. If a repetition rate goes from 20% to 30% in ten years that may be discouraging, but it is not unreasonable.) If you do find unreasonable projections, take careful note of the following. 1. Are all the rates being projected in an unreasonable way? 2. If all the rates are projected unreasonably, are there one or two particular historical years that may be throwing the projection off? For example, if the repetition rates for all grades oscillate around certain values but suddenly become much higher for the last historical year, then it is clear that it is this last year that is throwing the projection off. Note which year it is. 3. If it is only some of the rates that are being projected badly, note exactly which. For example, note that it is the 5th grade repetition rate and the 3rd grade dropout rate. After you are finished with your examination, press [Esc]. The machine will ask you if you want to look at another table. Respond [n] for No. You will then be in the Main Menu.

Re-checking recently entered data, re-running ESTIMATE.

A likely cause of questionable estimation of historical rates is a mistake in entering data. If there are problems with some of the estimated rates in the historical periods, use the editor to make sure you have correctly entered the data necessary for the

ESTIMATE module (see section on Entering or Editing Data in this chapter). If you find some mistakes, correct them by re-entering the numbers, and re-run the ESTIMATE module as described above.

Re-running PROJRATE

If in examining the projected rates you discovered that there were one or two specific years that were throwing the projections off, and you have noted which were those years, and there were no data entry mistakes, the logical thing to do is to edit the variable "include", as described above in the section on "Entering Data", and change the questionable years from "y" to "n", so that the projection process does not consider those years part of the data base. So, review the section on editing data, change the "include" variable, and then re-run the projrate module. Of course, you do not need to re-enter all the "enrolage" numbers again - just go to the part of the "Entering Data" section that applies to the variable "include", noting that you still have to get into the editor, and so forth, as explained toward the beginning of the "Entering Data" section. After re-projecting the rates, examine them again, with the Tabulate facility, as described above. If there are problems still, it may be that you should be projecting using fixed, rather than variable, future rates. After you have produced a set of projected rates that appear reasonable (even though they may be discouraging), the next step would be to run the modules EDPROJ and COHORT.

Editing the projected rates

If it is only a few of the projected rates that are being projected badly, then the easiest way out is to directly edit the projected rates. If the problem is that one of the projected rates is going up too fast, for example, the following procedure is appropriate. Check (visually should be enough) to see whether there is a clear upward trend in the historical data for that rate, excluding strange years. If, ignoring the strange years, there is a clear historical trend upward, then use the editor (as described below) to put the highest reasonable historical rate in place of those projected by PROJRATE. This is an honest and conservative approach only if there is a clear trend, and only if the years you have singled out as strange years really are strange, rather than part of the trend. If the historical trend is downward, excluding any strange years, then you would be justified in using the lowest reasonable historical rate in place of the projected rates. The following example should make it clear. Suppose we have the values in the table below for the 3rd grade repetition rates. It is clear that there is a gradual upward trend in the historical values, but it is also clear that the values for 1980 and 1981 are off that trend. Therefore, the projection, as done automatically, is much too steep: the last "reasonable" (i.e., ignoring the off-trend values of 27% and 25%) historical value is 15%, and the projected values go up to 53% within 10 years. The suggestion, then, is to overwrite the automatically produced projections by putting in the highest historical value of 15% in all the projected periods.

Year	3rd grade repetition rate	Historic or projected	Value after editing
1974	.12	h	.12
1975	.12	h	.12
1976	.13	h	.13
1977	.14	h	.14
1978	.13	h	.13
1979	.15	h	.15
1980	.27	h	.27
1981	.25	h	.25
1982	.26	p	.15
1983	.29	p	.15
1984	.32	p	.15
1985	.35	p	.15
1986	.39	p	.15
1987	.42	p	.15
1988	.46	p	.15
1989	.49	p	.15
1990	.53	p	.15

To make these changes, and assuming you are in the Main Menu (if you are not, pressing [Esc] repeatedly - but slowly - should get you there), press [e] for Edit. Choose the appropriate file to edit by pointing to it and pressing [Enter]. Choose the appropriate variable ("repeatrt" in this case, but it could be some other). This can be done by pressing [Shift] and [?], pointing to a variable name with the cursor, and picking it out by pressing [Enter]. The machine will display the first year of data for that variable. Note that the display will not look anything like the table above, since the table above has only one rate in it, and has other extraneous information. Press [F4] until you get to the first projected year. If you overshoot, use [F3] to go back. When you are at the first projected year (1982 in the example above), enter .15 in the slot for 3rd grade, press [F4] to move forward, enter .15, and so forth for all projected years. When you are finished, press [/] to get the menu, and [s], [Enter] to save the variable under its own name. Then press [q] to go back to the main menu.

Note that this procedure of over-writing the projected values with assumed ones is unorthodox, to say the least. It should be used only in cases where there is no alternative, and the recipient of any projections made this way should be informed that assumed rates were used, and why.

At this point you do not want to re-run PROJRATE! In fact, if you re-run it, it will simply write in the old projected values on top of what you have just entered. The

next step is to run the modules EDPROJ and COHORT.

Running the EDPROJ module.

From the Main Menu, (if you are not there, pressing [Esc] repeatedly - but slowly - should get you there) pick [r] for Run. Point to the EDPROJ module and press [Enter]. A set of run parameters will be displayed. The only one that should need editing will be the number of periods. Select [e] for Edit. Go down to that slot with the cursor, and enter a number sufficiently large to cover both the historical periods and the projected periods. For example, if the base year is 1974 and we want to project to 2000, entering 28 periods would be sure to cover it. After entering the number of periods, press [/] and [g] for Go. The machine will then ask you whether you want to make a backup of the file. This will not normally be necessary, unless you have entered a lot of numbers by hand and you want to protect this work. If you decide to make a backup, it would be reasonable to overwrite the backup file made recently while running ESTIMATE. You may wish to review this procedure in the section on ESTIMATE, above, although the computer should lead you through it fairly well. After you take care of the backup question, there will be three initial screens. Press [Enter] after reading each one. Then the computer will ask whether all the necessary input data were entered. It is very likely that they were, since most of them will have been produced by the PROJRATE module, or were there from previous runs. If you are sure all the data are in the machine, then press 1. After the module is finished, you will get a message saying so. Press [Enter], and then respond [y] to the question whether you wish to run another application.

Running the COHORT module.

The procedure for running COHORT is precisely the same as for EDPROJ, so there is nothing new here, in terms of running the module. At one point the computer will ask whether the projection of the efficiency measures should be done with fixed or varying transition rates. Note that if you respond "varying", you need to have previously run the PROJRATE module for, typically, 15 or 20 periods more than you intend to run COHORT. A fuller explanation of these issues can be found in Chapters 2 and 4. After you run COHORT, the machine will ask whether you wish to run any other modules. Since COHORT is the last module, the answer should be [n] for No. That will put you in the Main Menu, where you can proceed to view the results of EDPROJ and COHORT.

Tabulating results of EDPROJ and COHORT.

There are many tables with results from EDPROJ and COHORT that one might wish to see, but here we will concentrate on only two of them: GER and COHORT. Starting from the Main Menu, press [t] for Tabulate. When asked for the name of the table, respond "ger" or use [Shift] and [?], with the cursor keys, to pick out the name "ger". Then the machine will display the table parameters. The parameters that will

control the tabulation will be displayed. To change them, press [e] for Edit. There are only two that might need changing. If you do not want the table to go to the monitor, but want it to go to the printer instead, move the cursor to the appropriate position and enter "prn" for "printer". Note that if there is no printer connected to the machine, or if the printer is turned off, you may get locked out, and the only way back will be to warm boot the computer. The other parameter that needs to be set is the number of periods to put in the table. If your base year is 1974, for example, and you have projected to the year 2000, move to the slot for "periods" and enter "1-17" (for 1974-2000). Press [/] to go to the Menu, and then press [g] to Go. If you sent the table to the printer, the whole table will be printed. If it is going to the console, the computer will put one page at a time on the screen. To get a new page, press [PgDn] (the key for 3 on your numeric keypad). You can go back up by pressing [PgUp] (the key for 9 on the numeric keypad). If these keys do not respond, it may be because the [NumLock] key (above the 8 in the numeric keypad) is set wrong. Press [NumLock] once and try again. When you have inspected the table, or when it has printed, the machine will ask you whether you wish to see another table. (If the table went to the screen you will need to press [Esc] first.) Respond [y], select "cohort", and go through the process again. To read these tables, refer to the list of variables at the end of Ch. 4 of this manual.

If the results of these projections are odd, then re-examine the projected rates (see the appropriate section above), re-project the rates using PROJRATE (and perhaps not using logistic or linear trends, but using fixed future rates, or targets), or re-edit the rates, and then re-project EDPROJ and COHORT. After you are satisfied with the projections, you might want to print the tables, rather than just sending them to the monitor.

Labeling the results

It is very important, for the reader of the results you produce, or for yourself at a later date, to clearly label any tables you produce with the assumptions and methods used to produce them. For example, if the historical rates were estimated using the Schiefelbein method, assuming 5% 6th grade dropout rate and 2% of non-promoted as dropouts, and if the rates were projected using a logistic curve, and the cohort analysis was done using fixed rates, then you should clearly write all this information on the tables produced before you file them or send them anywhere.

Managing your files

Different projections can be kept in different files, so that you need not erase past projections every time you want to make a new projection under different assumptions. The recommended procedure is as follows. Suppose we have just made a set of projections and these are in the default data file which we have called, for example, "guat" (for Guatemala). Let us suppose that these projections correspond to historical rates estimated using the Schiefelbein method. These can be backed up to a file which we

could call "guatsch". From the Main Menu, press [m] for Manage. Then choose [d] for Data, and [b] for Backup. The machine will ask which file to back up, assuming you wish to back up the one you are working with. Since this is the case, press [Enter]. It will ask you for the name of the file you want to back your data onto. At this point you would respond "guatsch", to remind yourself that this file will contain data for Guatemala, estimated using the Schiefelbein method. If this file already exists, it would ask whether you wish to over-write it. If it contains information you want to save, do not over-write it, and choose another name. Otherwise, overwrite it. If it does not exist, then there is no problem. At this point, you will have exactly the same information in "guat" and in "guatsch". But now let us say you produce estimates using the standard method, and projections based on those estimates. Then "guat" contains the standard method estimates and "guatsch" contains the Schiefelbein method estimates, and projections based on those estimates. You could then go on to back up "guat" onto a file called, for example, "guatstd" for Guatemala-standard. Naturally, this process can fill up your hard disk very quickly, so you may wish to save many of your files on floppy disk. This is easily done. Instead of giving the machine the name "guatsch" when backing up, give it the name "a:guatsch", after having put an empty, formatted diskette in drive A, and assuming the data file fits in a diskette. You can also use DOS, from outside Host, to accomplish many of these tasks, but this is beyond the scope of this chapter. See a short description of this process at the beginning of Ch. 2 of this manual, or refer to a DOS manual.

What to do about problems.

Most problems are not as bad as they seem. Some errors dump you out of Host and into DOS. No problem - just go back into Host and don't worry about the message you find as you go back into Host. If you can't get any reaction from the machine, press [Esc] several times. If that does not work, it may be necessary to warm boot, or even cold boot the machine. The problem is not the error itself, but trying to find out what is causing it, so you can go ahead with the projections.

If you do not have the right hardware or software you will certainly have problems. For example, if you do not have an 8087 numeric co-processor, or the equivalent, you will run into serious problems very fast. If you do not have sufficient space in your hard disk, you will also run into problems. If you have not copied all the files Host and PETS you need, again, there will be parts of the system (perhaps essential parts) that will not run.

Many problems are caused because the "config.sys" file was not properly created, or the machine was booted from a disk that does not have the proper "config.sys" file. See section above on installing Host.

Many problems are caused when Host cannot find files it has been told are there. Make sure you have indeed put the files you have said in the directories you have said.

This is particularly important before you run the Configure option in the Main M

Many errors are caused by trying to do things that are methodologically though the system will catch some of these. For example, projecting a trend few historical years of observations is likely to cause unpredictable errors. Al problems are created in attempting to run a particular module without having the necessary data. It is also important to run the modules in order - you shoul EDPROJ without having, at some point, run PROJRATE.

If you try to send a table to the printer and there is no printer, or it is not to the computer, you might get "locked out", and you might have to warm machine.

Make sure that modules have the data they need before you run them. Re opening screens before you run the module, and if you are not sure whether entered the necessary information, or whether you have run the previous mo automatically enter the information), then you might want to take advantag opportunity to leave the module presented after the third informative screen module, and correct the situation.

Note that it is not the same thing to have a variable with all zero values in period as to have no data for that period. If there really are no data, then when the editor for that variable, the machine will display a sign at the top that says "] found; default values supplied". To change this non-data period into one with z data, and assuming you are already in the editor (if you are not, press [Esc] u get to the Main Menu, choose [e] for Editor, point to a particular file and [E choose it, and enter the name of a variable), you have to save each period indi by hitting the sequence of keys [s], [Enter], [r], [F4], until you run into peri have data already, or are not relevant.

FILE

PETSPETSPETSPETSPETSPETSPETSPETSPETS

PRIMARY
EDUCATION
TRACKING
SYSTEM

PETSPETSPETSPETSPETSPETSPETSPETSPETS

1987

Developed by

RTI RESEARCH TRIANGLE INSTITUTE
Research Triangle Park, NC 27709 USA

with

HARVARD UNIVERSITY
Cambridge, MA 02138 USA

Under

B·R·I·D·G·E·S

Basic Research and Implementation in DevelopinG Education Systems

61

Sponsored by

Office of Education
Bureau for Science and Technology

and

Office of Development Resources
Bureau for Latin America and the Caribbean

United States Agency for
International Development

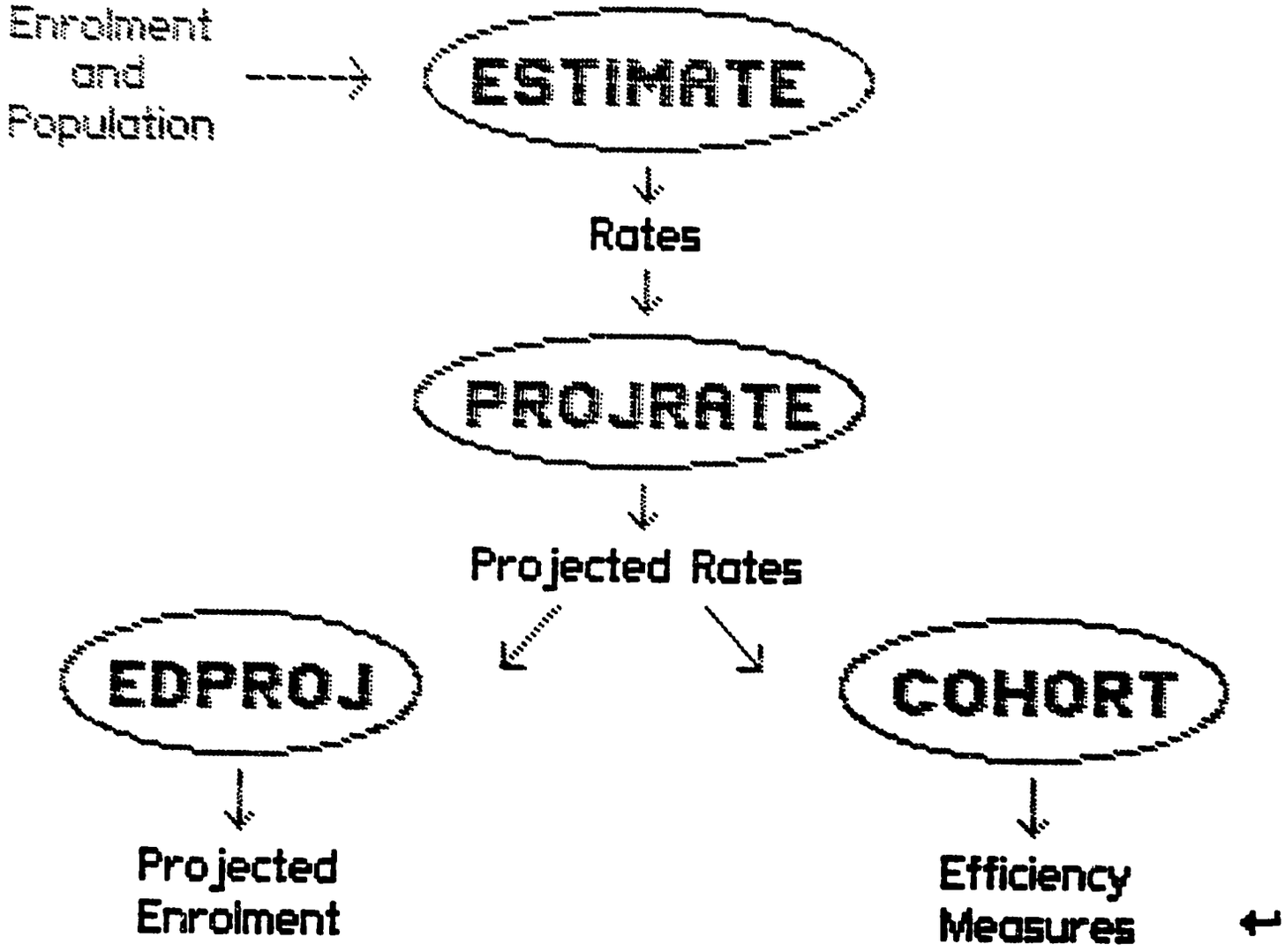
18

Primary
Education
Tracking
System

- Estimate transition rates
- Project rates into the future
- Project student enrolment
- Evaluate efficiency

PETS is made up of 4 modules

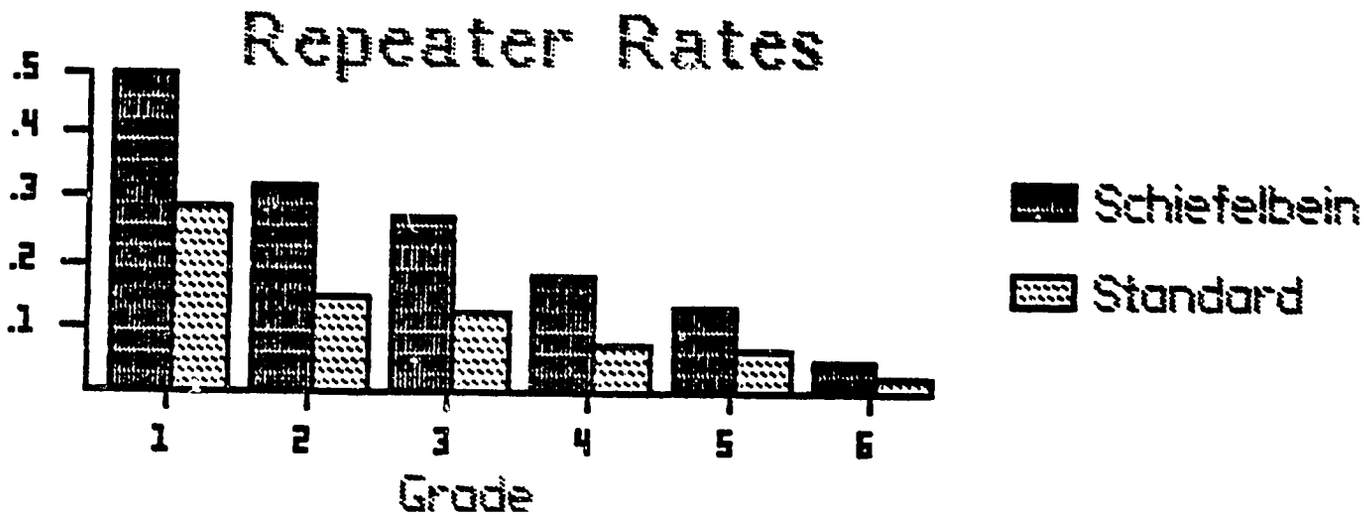
ESTIMATE	estimates past repetition, dropout and enrolment rates
PROJRATE	projects these rates forward in time
EDPROJ	projects school enrolment
COHORT	estimates educational efficiency measures

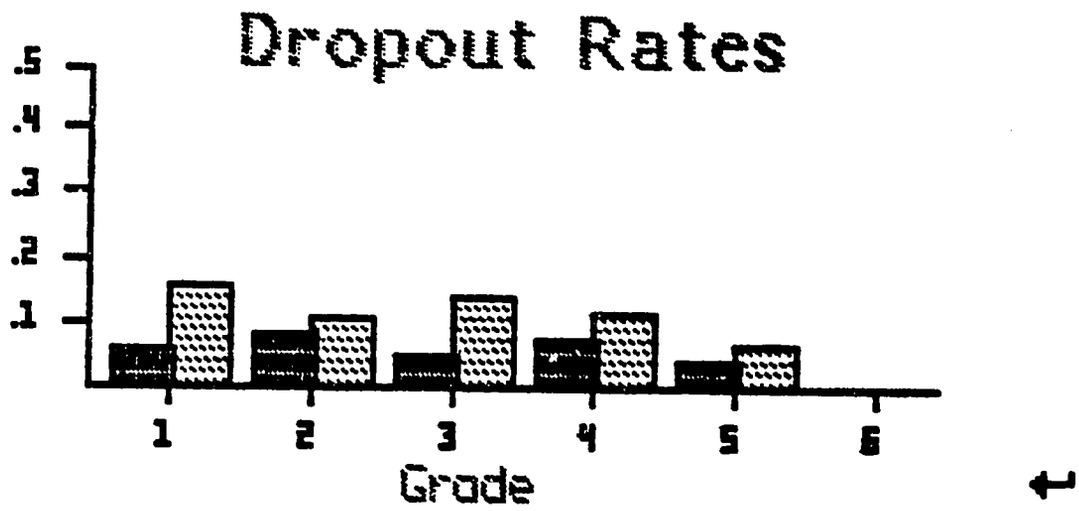

Enrolment and Population
ESTIMATE
Rates
PROJRATE
Projected Rates
EDPROJ
COHORT
Projected Enrolment
Efficiency Measures

ESTIMATE

Methods of Estimation

- Schiefelbein method
  - internal consistency with observed enrolment
  - higher repeater and lower dropout rates
- Standard method


Repeater Rates
.5
.4
.3
.2
.1
1
2
3
4
5
6
Grade
Schiefelbein
Standard


Dropout Rates
.5
.4
.3
.2
.1
1
2
3
4
5
6
Grade

ESTIMATE

INPUT

- Enrolment by grade and age
- Population by age
- Age at entrance
- Enrolment for last age group

Schiefelbein Method:

- Repetiton rate of last grade
- % of nonpromoted who are dropouts
- Disaggregation of enrolment

OUTPUT

- Repeater rates by grade
- Dropout rates by grade
- Entrance rate of new students in first grade
- Age-specific enrolment rates

ESTIMATE

INPUT

- Enrolment by grade and age
- Population by age
- Age at entrance
- Enrolment for last age group

Standard Method:

- Number of repeaters by grade

OUTPUT

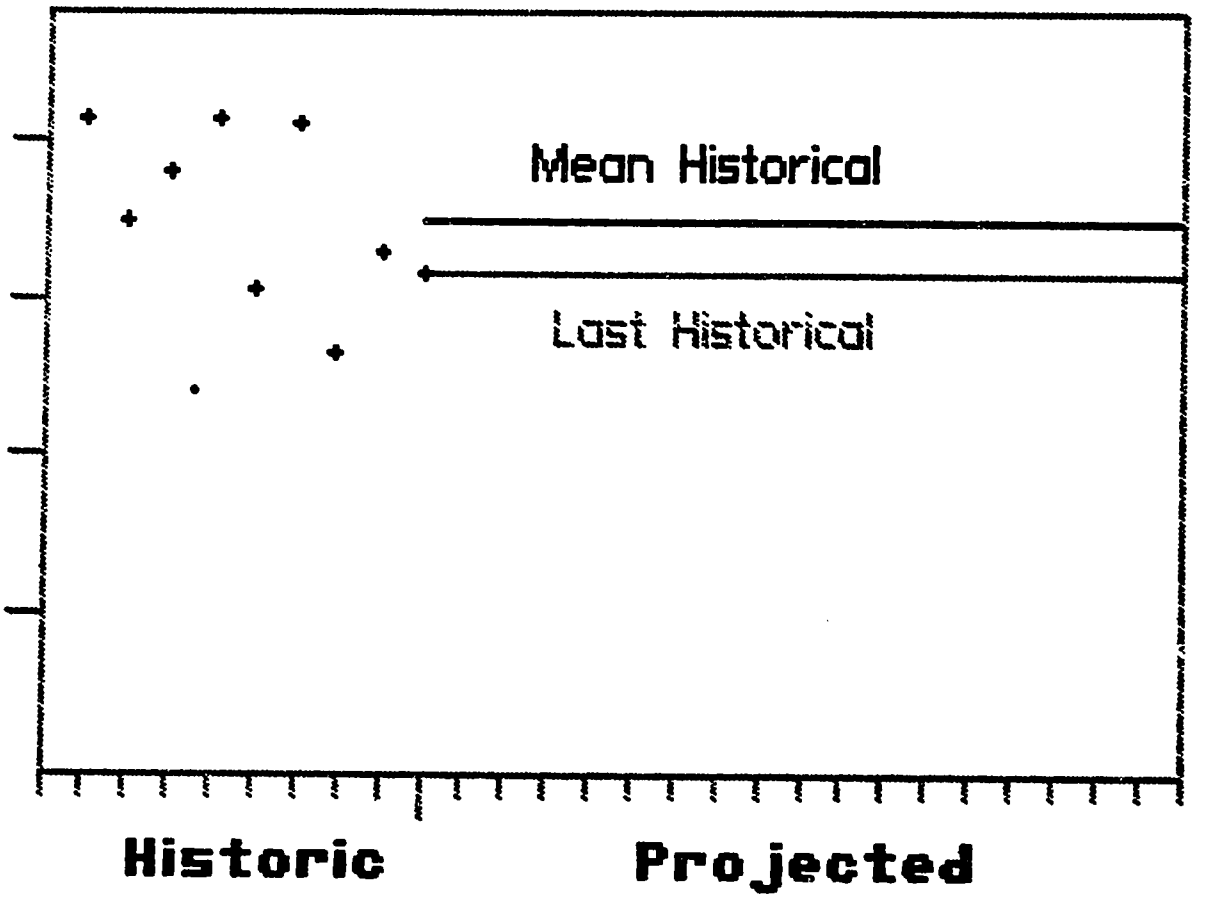
- Repeater rates by grade
- Dropout rates by grade
- Entrance rate of new students in first grade
- Age-specific enrolment rates

PROJRATE

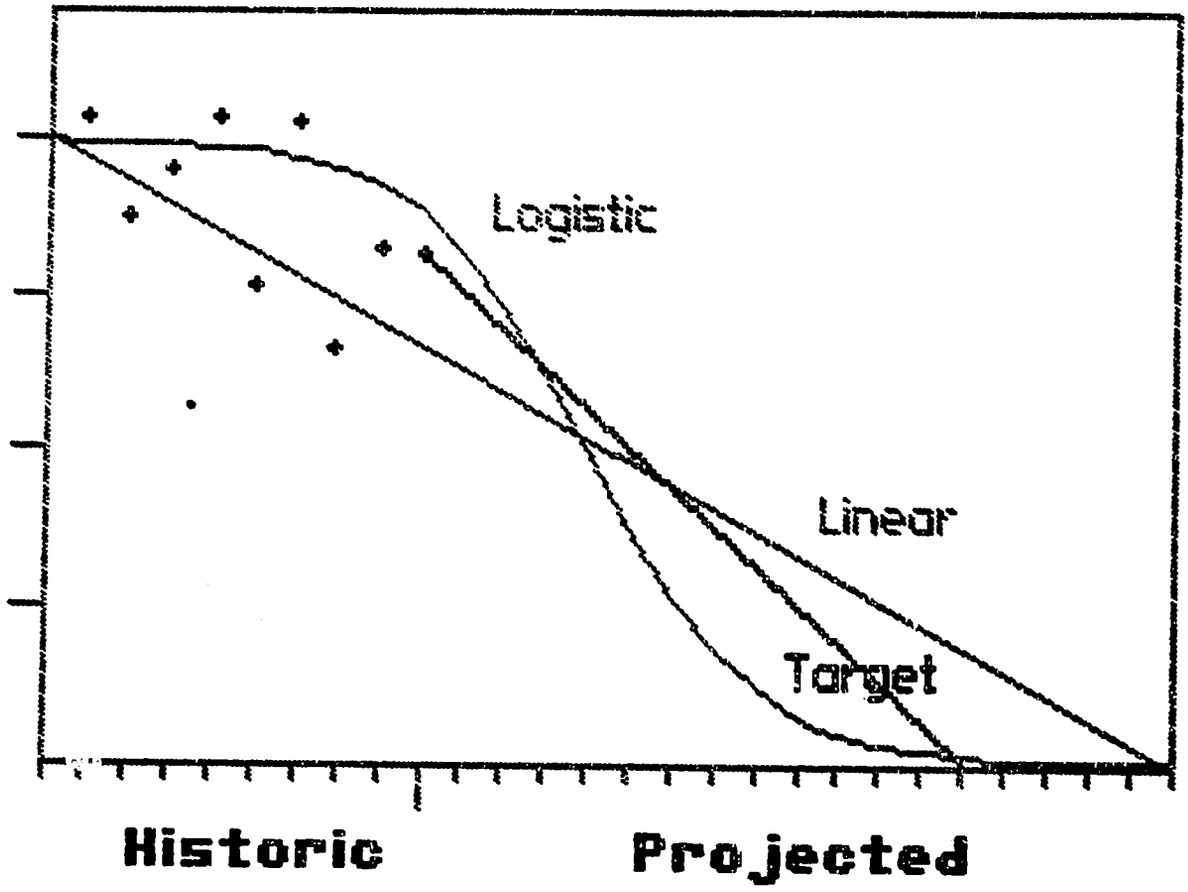
Methods for projections

- Fixed future rates
  - last historic rate
  - average historic rate
- Variable future rates
  - logistic curve
  - linear trend
  - target levels and date

PROJRATE: Fixed Future Rates



PROJRATE: Variable Future Rates



1

PROJRATE

INPUT

- Repeater rates
- Dropout rates
- Entrance rate
- Indicators for:
  - which periods are historic
  - which historic periods to use in projections
- Target rates (optional)

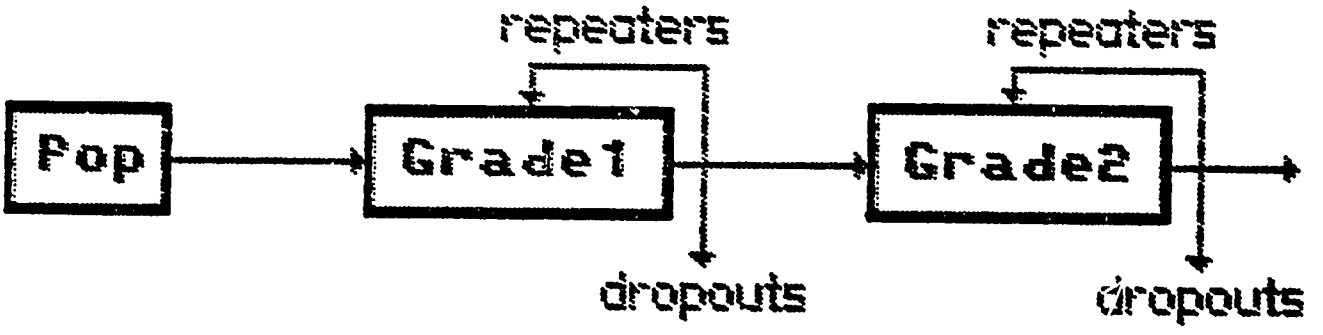
OUTPUT

Projected values of:

- Repeater rates by grade
- Dropout rates by grade
- Entrance rate
- Regression statistics

EDPROJ

Projects enrolment using the grade-transition method



EDPROJ

INPUT

- Repeater rates by grade
- Dropout rates by grade
- Entrance rate
- Population by age
- Oldest school age

OUTPUT

- Enrolment by grade
- Gross enrolment rate

COHORT

A hypothetical cohort of 1000 entrants into the primary school system

---

* Fixed repeater and dropout rates
* Variable repeater and dropout rates

COHORT

INPUT

- Repeater rates by grade
- Dropout rates by grade

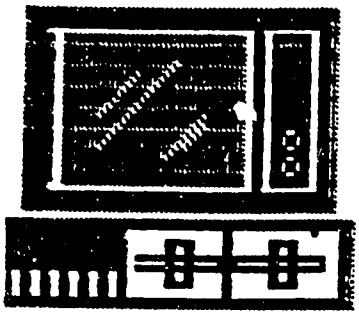
OUTPUT

- Input/output ratio
- Number of graduates
- Average years in school
- Average years to graduate
- Number who graduate on time

POSSIBLE PROJECTION COMBINATIONS

ESTIMATE	PROJRATE					COHORT
	Fixed Future		Variable Future			
	Last	Avg	Logistic	Linear	Target	
Schiefelbein*						Fixed
						Variable
Standard						Fixed
						Variable

PETS runs on a microcomputer



- Process large data files
- Execute complex calculations
- Present results in tables

. . . But you don't have to be a computer programmer to use PETS

PETS operates under

HOST

A Model Development and Support System

Developed by RTI with funding provided by USAID through the Integrated Population and Development Planning project (INPLAN)

PETS

STEPS TO RUN THE SYSTEM

EDIT	Enter and edit the data
RUN	Execute the modules
TABULATE	View the results

↵

MAIN MENU

Configure Define Edit Run Tabulate Graph Instruct Manage Access *DOS Quit

Specify file names, directories, printer and monitor settings.

- Use the arrow keys to highlight your menu choice then press [Enter], or press the key corresponding to the first letter of your menu choice. Pressing the [Esc] key is the same as selecting the choice "Quit".

Variable Name : ENROLAGE
Description : Enrolment by grade and age

EDIT
Data File : GUAT

ENAGE/GRADE PERIOD : 1974 SEX : Both sexes

	Primero	Segundo	Tercero	Cuarto	Fifth
5 yr. olds	0.0000000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000
6 yr. olds	8445.0000000	0.0000000000	0.0000000000	0.0000000000	0.0000000000
7 yr. olds	69969.000000	6197.0000000	0.0000000000	0.0000000000	0.0000000000
8 yr. olds	52453.000000	30018.000000	6417.0000000	0.0000000000	0.0000000000
9 yr. olds	34930.000000	27203.000000	21634.000000	5686.0000000	0.0000000000
10 yr. olds	25754.000000	23300.000000	21622.000000	17356.000000	4955.0000000
11 yr. olds	13885.000000	15656.000000	17177.000000	15070.000000	13874.000000
12 yr. olds	8826.0000000	11627.000000	14366.000000	13895.000000	13393.000000
13 yr. olds	3619.0000000	5597.0000000	8343.0000000	8920.0000000	10716.000000
14 yr. olds	1254.0000000	1885.3333333	3072.0000000	4194.0000000	6614.0000000
15 yr. olds	418.00000000	628.44444444	1024.0000000	1398.0000000	2204.6666667
16 yr. olds	139.33333333	209.48148148	341.33333333	466.00000000	734.88888889
17 yr. olds	46.444444444	69.827160494	113.77777778	155.33333333	244.96296296
18 yr. olds	15.481481481	23.275720165	37.925925926	51.777777778	81.654320988
19 yr. olds	5.1604938272	7.7585733882	12.641975309	17.259259259	27.218106996
20 yr. olds	1.7201646091	2.5861911294	4.2139917695	5.7530864198	9.0727023320

F3-Previous Period | F4-Next Period | F5-First Period | F6-Last Period

Currently executing :
RUN M=COHORT D=GUAT S=1 N=5

P R I M A R Y E D U C A T I O N T R A C K I N G S Y S T E M
(P E T S)

Developed by
Research Triangle Institute
with
Harvard University
under the
Basic Research and Implementation
in Developing Education Systems Project
(B R I D G E S)

Now Executing Module COHORT
Press ENTER to continue:

Repetition Rates

REPEATRT: Repetition rate

SEX : Both sexes

PERIOD/GRADE	Primero	Segundo	Tercero	Cuarto	Fifth	Sixth
1983	0.521	0.306	0.262	0.162	0.099	0.050
1984	0.524	0.286	0.240	0.139	0.083	0.050
1985	0.528	0.268	0.220	0.120	0.069	0.050
1986	0.531	0.250	0.201	0.103	0.057	0.050
1987	0.534	0.233	0.183	0.088	0.047	0.050
1988	0.537	0.216	0.166	0.075	0.039	0.050
1989	0.541	0.201	0.150	0.064	0.032	0.050
1990	0.544	0.186	0.136	0.054	0.027	0.050
1991	0.547	0.173	0.123	0.046	0.022	0.050
1992	0.550	0.160	0.111	0.039	0.018	0.050
1993	0.554	0.148	0.100	0.033	0.015	0.050

Enrolment by Grade

ENROL: Enrolment by grade

GRADE/PERIOD	1983	1984	1985	1986	1987
Primero	331230	362501	385184	402220	416504
Segundo	190138	202783	219040	231512	239470
Tercero	147860	160116	174243	191096	205691
Cuarto	107687	120751	133759	148110	164794
Fifth	84259	93365	106918	120919	136135
Sixth	68956	76183	86279	100598	115704
Total	930130	1015698	1105424	1194464	1278298

GRADE/PERIOD	1988	1989	1990	1991	1992
Primero	429489	441954	454303	465748	476827
Segundo	244335	247423	249594	251346	252511
Tercero	216177	223217	228020	231553	234430
Cuarto	180022	191756	200276	206598	211616
Fifth	153526	169744	182655	192346	199740
Sixth	132095	150690	168379	182903	194141
Total	1355644	1424784	1483226	1530494	1569265

GRADE/PERIOD	1993
Primero	487829
Segundo	253366
Tercero	236635
Cuarto	215919
Fifth	205705
Sixth	202919
Total	1602373

Cohort Analysis of 1000 Entrants

PERIOD/	GRADRATE	ONTIME	SYGRAD	SYGRAD2	YRSSCHL
1983	716	126	8.23	9.87	7.07
1984	774	150	8.11	9.35	7.23
1985	820	172	8.02	8.96	7.35
1986	857	194	7.93	8.67	7.43
1987	886	214	7.86	8.44	7.48
1988	909	232	7.79	8.26	7.51
1989	927	249	7.75	8.12	7.53
1990	942	264	7.71	8.01	7.54
1991	953	278	7.67	7.92	7.54
1992	962	290	7.64	7.85	7.54
1993	969	300	7.62	7.79	7.54