PN-ABK-262

GENDER

Gender and Education Reporting System

User's Guide

Developed by
Research Triangle Institute
with
Harvard University

B • R • I • D • G • E • S

Basic Research and Implementation in Developing Educational Systems

sponsored by
Office of Education
Bureau of Science and Technology
U.S. Agency for International Development

need path statement

PN-ABK-262
ISN 15264

C:\GENDER>path = C:\host\

get into subdir. type host

host screen - go to conf. verify that default data site is WID

Screen

gender 4x outputs console

Save

THE GENDER AND EDUCATION REPORTING SYSTEM

G E N D E R

USER'S GUIDE AND DESCRIPTION

0 = missing value

Prepared by

Scott Moreland
Gordon Cressman
Luis Cubeddu
Research Triangle Institute

and

Jim Tarvid
Logical Support

March 1989

Basic Research and Implementation in DevelopinG Education Systems (BRIDGES)

Sponsored by

The Office of Education, Bureau of Science and Technology
and
The Office of Women in Development
Bureau for Program, Policy and Coordination

U.S. Agency for International Development
Contract Number 5824-A-5076

Table of Contents

Chapter 1. Introduction to the GENDER System 1

Introduction 1
GENDER Modules 1
Using GENDER on the Microcomputer 2
Hardware Environment 2
Software Environment 2

Chapter 2. The WID International Database 3

Variables in the Database 3
Input Variables 3
A. Control Variables 3
B. Input Variables for Projection Modules 4
C. Output Variables of Projection Modules 4
D. Other Variables and Indicators 5
Imperfections in the Data 5

Chapter 3. Using GENDER's Software 6

Fundamentals of the GENDER System 6
GENDER Modules: POPPROJ 6
Module Description 6
Using POPPROJ: Variable Specification 7
Target Year 7
ZONE. Dummy Variable for Regions: ARAB, ASIA, SSAF 7
Total Fertility Rate 7
Mortality 8
Base Year Population 8
Sex Ratio 8
Target Toggle 8
GENDER Modules: EDUPROJ 8
Module Description 8
Using EDUPROJ: Variable Specification 10
Target Year 10
Target Toggle 10
Entrance Age 10
Population (Base Year and Projected) 10
Historic Enrollment 10
Historic Gross Enrollment Rates 11
Historic Data for Number of Schools and Teachers 11
Base Toggle 11

Chapter 3. Using GENDER's Software (continued)

GENDER Modules: QUERY 11
Module Description 11
To Run QUERY 11
Figure 1 Main Tree Structure of QUERY 12
Making Selections and Moving from Menu to Menu 12
Main Menu 13
1. Select which Database to Search 13
2. Search the Database 13
Overview of the Search Process 13
Clear all Search Results 14
Select the Search Variable 14
Set the Search Range 14
Edit the Search Controls 15
Search the Database 16
View the Search Results 16
Edit Cumulative Search Results 16
3. Use Search Results to Extract Data 17
4. Extract Data From the Database 17
Overview of the Extract Process 17
Select the Variables to Extract 18
Edit the Extract Controls 18
Extract Data from the Database 18
Quit and Return to the Main Menu 19
5. Quit and Return to DOS or Host 19
Using Databases Created with Extract in Host 19
Producing Predeifned Tables and Graphs 20
Table 1 List of Predefined Tables 20
Table 2 List of Predefined Graphs 20
Table 3 Educational Cohort Components 22
Table 4 Demographic Variables 26

Chapter 4. Technical Description of the Projection Methods Used in POPPROJ and EDUPROJ 39

1. Population Projection (POPPROJ) 39
2. Education Projection (EDUPROJ) 39
Enrollment Projection 40
Estimation of Transition Rates: The Standard Method 41
Efficiency Indicators Based on Cohort Analysis 42
Basic Concept 42
Efficiency and Other Educational Indicators 43
Indicators Based on Enrollment Data 44
Projection of Required Teachers and Classrooms 44
Estimation of Disparity Indices 45

Appendices

Appendix 1
Notes on Installation of Host and Gender 46

Appendix 2
WID Database: Variable List and Characteristics 47

Chapter 1. Introduction to the GENDER System

Introduction

GENDER is designed to allow education planners and policymakers to keep track of and to project educational indicators and primary enrollment in the education systems of developing countries. Particular attention is given to gender disparities in access to and retention in education systems. GENDER consists of an international cross section database with an accompanying software system. Based on a time series of enrollment data, and population data for all countries, the user can:

1. Estimate the transition rates within a primary school system.
2. Project those rates into the future using various approaches.
3. Project population every 5 years for 5-year age groups.
4. Calculate various educational gender disparity indices.
5. Search and/or Extract particular variables and data values.

The projection methods used in GENDER have been used for several years in developing country settings and are currently being utilized by the System for Tracking Educational Progress (STEP), developed through the BRIDGES project. Methodologies are described in greater detail under the Technical Description section of this manual (Chapter 4). In this section we describe the implementation of the system on the microcomputer and how to use it.

GENDER Modules

GENDER uses a modular approach. This gives the user the flexibility to run modules independently and to concentrate on those modules that are of particular interest. GENDER has three modules:

1. The module **POPPROJ** allows the user to project population by gender and 5-year age groups, every five years, for all 70 countries.

2. The module **EDUPROJ** estimates the historical education transition rates, projects those rates forward for a specified period of time and determines enrollment in the primary school system.

3. The module **QUERY** allows the user to search and extract particular variables and data values, while creating new data bases.

Using GENDER 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 defacto standard for social science analysis and simulation modeling, GENDER has been implemented for the IBM/PC, IBM/XT, IBM/AT or IBM/PS2 microcomputer and compatibles. GENDER requires:

1. IBM microcomputer or compatible.
2. 256K of Random Access Memory (512K RAM is recommended).
3. 6 MB of Hard Disk space.
4. Numeric Coprocessor.
5. Printer (Optional).
6. Color Monitor (Optional).
7. Graphics Display Adaptor (Optional).

Software Environment

GENDER requires Host a software shell developed for modeling applications. Host was developed by the Research Triangle Institute with funds provided by USAID through the Integrated Population Development Planning (INPLAN) and the BRIDGES projects. An introduction to the facilities of Host, that it is necessary to know in order to use GENDER, is given in a separate manual entitled "The Host User's Guide". Both Host and its User's Guide are distributed with GENDER. In this document, no attempt is made to provide instruction in using Host.

Installation of both Host and GENDER is described in more detail in the Appendix 1.

Chapter 2. The WID International Database

GENDER is supplied with a database consisting presently of a cross-section of some 70 USAID-funded countries for some 60 variables. A list of all of the variables, including their technical definitions as Host variables, is given in the Appendix. All variables, except certain ones designed to control execution of the various modules, are disaggregated by gender and country.

While the GENDER system is distributed with some 60 variables, the trained user can add as many variables as desired provided enough room exists on the user's hard disk. In addition, with the EDIT facility of Host the user can change existing data entries or add new ones (fill in gaps in the database). Conversely, with the "Extract" facility of GENDER's "QUERY" module the user can reduce the database to a subset of the WID database.

Variables in the Database

In the sections below we briefly introduce the main variables in the WID database. The discussion is organized around the types of variables and their functions with respect to the projection modules.

Input Variables

There are two types of input variables in both the POPPROJ and the EDUPROJ module. Some variables were created to provide options for running the application modules. These are called "operation" or "control" variables. The other type of input variables require data values that represent assumptions for running the modules. More detail is provided in the next section on using the modules.

A. Control Variables

1. ZONE: Dummy variable to specify World regions.
2. TARGYEAR: Target Year for Population Projections.
3. TARGTOGL: Determines whether target life expectancy and total fertility rates should be used.
4. BASETOGL: The Base Toggle variable indicates whether the user or the module EDUPROJ must determine the base enrollment years.
5. BASEYEAR: If the Base Toggle variable is "on" (value=1) then the user must specify each country's base enrollment year.

B. Input Variables for Projection Modules (Historic or Actual Data)

1. TOT_FER: Total Fertility Rate.
2. LEXP: Life Expectancy.
3. POP5: Base-year 5-year age group population.
4. SEXRATIO: Ratio of Boys to Girls at Birth.
5. TARGLEXP: Target Life Expectancy.
6. TARGTFR: Target Total Fertility.
7. TARENTR: If the Target Toggle variable is "on" (value=1) then the user must specify the target entrance rate.
8. TARDROP: If the Target Toggle variable is "on" then the user must define the target dropout rate.
9. TARREP: If the Target Toggle variable is "on" then the user must define the target repeat rate.
10. ENTAGE: Age of Entrance to the School system.
11. POP5: Population by five-year age groups for the historic period (and for as many periods in the future the user wishes to project)
12. TEACHERS1: Number of Teachers by gender (Historic Data).
13. SCHOOLS1: Number of Schools (Historic data).
14. ENROL1: Base Year enrollment by grade and gender.

C. Output Variables of Projection Modules

Also referred to as projected variables, output variables are calculated in the projection process of either the POPPROJ module or the EDUPROJ module. To view projection results use Host's EDIT, GRAPH or TABULATE option. Output variables include:

1. ASFR: Age Specific Fertility Rates.
2. POP5: Five-year age group population data, every five years.
3. BASEYEAR: If the Base Toggle variable is "off" (value=0) then the module EDUPROJ defines the base enrollment years for all 70 countries.
4. NUMGRAD: Number of primary school grades.
5. LEXP_DIS: Disparity index for life expectancy.
6. ENTPOP: Entrance population.
7. SCHLPOP: School age population.
8. ENTRT: Entrance rate.
9. REPEATRT: Repeat rate.
10. DROPRT: Dropout rate.
11. ONTIME: Number of on time graduates (not repeated).
12. GRADRATE: Number who graduate out of 1000 entrants.
13. SYGRAD: Student/years per graduates not counting student/years "wasted" by dropouts.
14. SYGRAD2: Student/years per graduate counting student/years "wasted" by dropouts.
15. IORATE: The Input/Output ratio.
16. ENROL: enrollment by grade and gender.
17. GER: Gross enrollment Rate by gender in Primary School.

18. TEACHPRJ: Projected number of teachers.
19. SCHLPROJ: Projected number of schools.
20. DINDPROJ: Projected Disparity Index in Primary School. (Harvard's)
21. DINDPRJ2: Projected Disparity Index in Primary School. (RTI's)
22. DISIND: Disparity Index for historic data.
23. DISIND2: Disparity Index for historic data using alternate method.

D. Other Variables and Indicators

In addition to those variables which are of direct use in the projections of the GENDER system, there are a number of other indicators which are included in the WID data base. These are listed below.

1. CP_RT_UN Completion Rate - United Nations
2. CP_RT_WB Completion Rate - World Bank
3. CUR3AREA Current Secondary Expenditures by Area
4. CURLEV Current Expenditures by Level
5. CURRENT Current Educational Expenditures
6. DEP_RAT Dependency Ratio
7. ED_ATT Educational Attainment
8. ENR2AREA Secondary Enrollment by Area
9. ENR2FLD Secondary Enrollment by Field
10. ENR3FLD Tertiary Enrollment by Field
11. ENROL2 Secondary Enrollment
12. FLFPR Female Labor Force Participation Rate
13. GDP_GR GDP Growth Rate (1980-1986)
14. GNP_PC Gross National Product per Capita
15. GR_RAT Population Growth Rate
16. ILLIT Illiteracy Rate
17. IMRMF Infant Mortality Rate (Male and Female)
18. NETENRRT Net Enrollment Rate
19. REPEAT2 Number of Secondary Repeaters

Imperfections in the Data

It is important to point out that, as with <u>any</u> projections, those made with the GENDER system are only as good as the base year data and assumptions about future mortality, fertility, school entrance, and drop out and repeater rates. The base year values of many of these parameters are either unknown in many countries, or badly estimated in many others. It was beyond the scope of this project to re-estimate the base year population of dozens of countries, or to question their demographic assumptions. In the case of the education system, for example, repetition rates are badly underestimated in many Latin-American countries, and it was impossible to re-estimate them all: each country would have required a research project of its own. Thus, the projections need to be taken as the best possible given the data situation, rather than the best possible in some absolute sense.

Chapter 3. Using GENDER's Software

Fundamentals of the GENDER System

In running GENDER there are three fundamental tasks that the user performs with the help of Host:

a) data entry and editing,
b) running modules, and
c) 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 database and/or to change existing values. The Host RUN facility allows the user to execute a GENDER module (POPPROJ, EDUPROJ or QUERY) 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. Alternatively, the Host GRAPH facility allows one to view the variables in predefined graphs. Each of these facilities is described in more detail in the Host User's Guide.

GENDER Modules: POPPROJ

Module Description

POPPROJ is a "standard" cohort-component population projection module which provides five-year age group population projections on a five year basis. It can be used as a stand alone population package, because it contains the basic elements of standard population projection packages. In the context of the GENDER system, it is designed to provide the module EDUPROJ with the school-age population of all 70 countries in the database. Thus this module is designed primarily as a service module for EDUPROJ.

Because it projects for all countries in the database, its execution time can be lengthy. For example, a twenty year projection will take a little over one minute per country so that a projection for all countries would take over an hour. If one is interested in projections for only a few countries, the appropriate variables can be extracted from the database using the QUERY module and the projections performed on the resulting reduced database. This would save execution time.

Users of course do not have to use the population projections supplied by POPPROJ for the EDUPROJ projections; they can always supply their own and enter them via Host's EDIT facility.
The INPUT data for POPPROJ are:

1. ZONE: Dummy variable to specify World regions.
2. TOT_FER: Country specific Total Fertility Rate.
3. LEXP: Country specific Life Expectancy.

4. POP5: Base-year 5-year age group population.
5. SEXRATIO: Ratio of Boys to Girls at Birth.
6. TARGYEAR: Target Year for Population Projections.
7. TARGTOGL: Determines whether target life expectancy and total fertility rates should be used.
8. TARGLEXP: Target Life Expectancy.
9. TARGTFR: Target Total Fertility.

Note: The base year for the population projection is currently fixed at 1985. Data for years prior to 1985 have been supplied with the WID database.

The OUTPUT data for POPPROJ are:

1. ASFR: Age Specific Fertility Rates for all countries.
2. POP5: Five-year age group population data, every five years.

Using POPPROJ: Variable Specification

Target Year

In POPPROJ the execution period is not specified with the the RUN facility of Host, but rather through the GENDER system variable labeled TARGYEAR (target year). TARGYEAR specifies the target year for population projections in the POPPROJ module as well as for education projections in the EDUPROJ module. When projecting populations (POPPROJ), the value of TARGYEAR must be greater than the population base year (1985).

Since the population projections are performed every five years and TARGYEAR will accept any value, including years which are not in five year increments after the base year, will cause POPPROJ to be executed for the NEXT five year increment beyond TARGYEAR. For example if the user specifies TARGYEAR = 1991, then POPPROJ will project the population through 1995.

ZONE. Dummy Variable for Regions: ARAB, ASIA, SSAF

The user must define the particular world region a country belongs to. For example, in the case of Egypt, the user would enter a 1 under the region labeled ARAB (ASIA and SSAF will contain zeros as default value). ZONE is used to determine the model pattern of fertility for each country.

Total Fertility Rate

Total fertility rates for each country must be specified by editing the variable TOT_FER. TOT_FER data are required input for the base population year (1985). It is assumed by POPPROJ that TOT_FER is expressed in percentage terms. TOT_FER along with the ZONE variable are used to calculate age specific fertility rates.

Mortality

Mortality is specified by entering the future life expectancies in the variable LEXP. LEXP data is also required input for the base population year (1985). GENDER uses the United Nation's (UN) general life tables in order to determine survivorship rates.

Base Year Population

The user must enter the base year population by editing the variable POP5. In the current configuration of the WID database five-year age group population data must be available for 1985, for each country.

Sex Ratio

The sex ratio at birth must be specified in the variable SEXRATIO. POPPROJ assumes this will not change over the projection period, and so this needs to be entered only once. (Example value: 1.05 boys/girls.)

Target Toggle

The user must determine whether to use "target" life expectancy and total fertility rates by entering a 1 (yes) or a zero (no) in the variable TARGTOGL. If targeting is involved then the computer performs a linear interpolation between the base and target values (in the target year) for intermediate values of the variable. If the user's response is positive, then the target rates (TARGLEXP for life expectancy and TARGTFR for fertility) must also be specified. Should the user's response be negative, or no response is provided, then the target life expectancy and total fertility are held equal to the baseyear value and kept constant through out the projections. The variable TARGYEAR specifies the target year.

GENDER Modules: EDUPROJ

Module Description

The EDUPROJ module estimates base year educational transition rates and projects enrollment at the primary level for a specified period. While the POPPROJ module projects every five years, EDUPROJ provides annual projections. In doing so, it also calculates the primary school-age population on an annual basis. EDUPROJ has two components:

a. It estimates from base year enrollment data and population data the historic educational transition rates: entrance, repeater and dropout rates;

b. Using the standard grade-transition method it projects basic educational indicators and gender disparities in education through the period indicated by TARGYEAR.

Teachers 1
×
Enroll

The INPUTS of the EDUPROJ module are:

1. BASETOGL: The base toggle variable indicates whether the user or the module EDUPROJ must determine the base enrollment years. (value = 1 means that user determines the base year; 0 means that EDUPROJ does.)
2. BASEYEAR: If the base toggle variable is "on" (value=1) then the user must specify each country's base. enrollment year in this variable.
3. TARGYEAR: Target year (Must be less than or equal to the target year specified in POPPROJ).
4. TARGTOGL: Target toggle variable determines whether the target transition rates should be used.
5. TARENTR: If the target toggle variable is "on" (value=1) then the user must specify the target entrance rate.
6. TARDROP: If the target toggle variable is "on" then the user must define the target dropout rate.
7. TARREP: If the target toggle variable is "on" then the user must define the target repeat rate.
8. ENTAGE: Age of entrance to the school system.
9. POP5: Population by five-year age groups for the historic period (and for as many periods in the future the user wishes to project)
10. TEACHERS1: Number of teachers by gender (historic Data).
11. SCHOOLS1: Number of schools (historic data).
12. ENROL1: Base year enrollment by grade and gender.
13. LEXP: Life expectancy.
14. GRENRRT: Gross enrollment rate (historic Data).
15. REPEAT1: Primary repeaters.

The OUTPUTS of the EDUPROJ module are:

1. BASEYEAR: If the base toggle variable is "off" (value=0) then the module EDUPROJ defines the base enrollment years for all 70 countries.
2. NUMGRAD: Number of primary school grades.
3. DISIND: Disparity index for historic enrollment data.
4. DISIND2: Disparity index for historic enrollment data. (alternate method)
5. LEXP_DIS: Disparity index for life expectancy.
6. ENTPOP: Entrance population.
7. SCHLPOP: School age population.
8. ENTRT: Entrance rate.
9. REPEATRT: Repeat rate.
10. DROPRT: Dropout rate.
11. ONTIME: Number of on time graduates (not repeated).
12. GRADRATE: Number who graduate out of 1000 entrants.
13. SYGRAD: Student/years per graduates not counting student/years "wasted" by dropouts.

14. SYGRAD2: Student/years per graduate counting student/years "wasted" by dropouts.
15. IORATE: The Input/Output ratio.
16. ENROL: enrollment by grade and gender.
17. GER: Gross enrollment rate by gender in primary School.
18. TEACHPRJ: Projected number of teachers.
19. SCHLPROJ: Projected number of schools.
20. DINDPROJ: Projected disparity index in primary school.
21. DINDPRJ2: Projected disparity index in primary school. (alternate index)

Using EDUPROJ: Variable Specification

Target Year

While executing EDUPROJ from the RUN facility of Host, the user need not specify the number of periods for which to execute the module. True also for POPPROJ, the target year variable (TARGYEAR) will specify the number of years for which to project. TARGYEAR is shone by the two modules POPPROJ and EDUPROJ. In EDUPROJ the target year must be greater than the base population year (1985), but less than or equal to the value specified for used POPPROJ.

Target Toggle

The user must determine whether to use "target" repeat and dropout rates by entering a 1 (yes) or a zero (no) in the variable TARGTOGL. This variable is also shone with POPPROJ. If the user's response is positive, then the target rate values must be specified.

Entrance Age

For each country the user must specify the nominal entrance age of children into the first year of primary school.

Population (Base Year and Projected)

Five-year age group population data (POP5) must be available for the base year as well as for the number of periods for which enrollment is to be projected. For example if we wish to project educational enrollment through 1997, then we must have population data for 1985 and 2000. Normally, this data will have been supplied through running POPPROJ.

Historic Enrollment

In order to project the basic education indicators, historic enrollment data must exist for two years in a row. GENDER's WID database contains historic enrollment data for the period 1980 through 1985, however not all years contain data for all countries.

Historic Gross Enrollment Rates

In order to calculate the historic indices for education disparity, gross enrollment rate data must be available. (GRENRRT)

Historic Data for Number of Schools and Teachers

Assuming that the ratios of teacher to student and school to student remain equal through out the projection, EDUPROJ calculates projected schools and teachers.

Base Toggle

The BASETOGL variable indicates who, whether the user or the EDUPROJ program itself, will determine each country's base enrollment year. If the specified value is 1 then the user defines the base year for all 70 countries. On the other hand, if the value is zero (default) then EDUPROJ will search and determine the "base year" for all countries. The program will find the last two years for which consecutive enrollment data exist and will set the base year as the first of those two years. The purpose of this facility is to allow the user to override the program's choice of the base year.

In the current database the first year for which data are available is 1980. This means the base year must be 1980 or later. Recall that the base year for the population projection is currently fixed at 1985. In cases when the EDUPROJ base year is <u>prior</u> to 1985, EDUPROJ will input necessary population variables by interpolating between 1980 and 1985 population values.

GENDER Modules: QUERY

Module Description

The QUERY module has two basic functions. First, it allows searches of particular variables and data values from a database and identifies which countries fit the criteria for the search. Secondly, QUERY will extract a selected list of variables for selected countries (and levels of other dimensions such as education level, sex, etc.) and store them into a new user-defined database. The new database would be a subset of the original and the criteria for extraction can also correspond to those of a search.

To Run QUERY

As with any Host module to run the QUERY module the user selects the RUN option of Host and then specifies the QUERY module as the GENDER module to execute. QUERY can also be run directly from DOS by typing "QUERY". The other choice to be made is that of a database. Normally this will be the master WID database which forms part of the GENDER system. While, as

explained below, it is possible for the user to create new databases from the original master WID database, unless one of these is the current Host default database, it should not be chosen at the RUN option; within QUERY there is a facility for selecting other databases.

In addition, the QUERY module will run with ANY Host database so that the experienced user can create their own master databases.

As mentioned, QUERY can also be run from DOS simply by typing the word "Query" from the subdirectory in which it is found. In this case, QUERY will choose as its default database the one specified as the default database in the current system environment file of Host (set with Host's CONFIGURE facility).

QUERY differs from the previous two GENDER system modules in that it is an interactive module. The user must make decisions regarding what variables to search and extract, and must be sure to follow QUERY's defined guidelines and procedures listed below. QUERY uses a "nested menu" approach to guide user choice. The "tree structure" of the menu is provided in Figure 1.

Figure 1

Menu Tree Structure of QUERY

```
Main Menu──┬
           ├─Select database
           ├─Search the database──┬
           │                      ├─Clear all search results
           │                      ├─Select the search variable
           │                      ├─Set the search range
           │                      ├─Edit the search controls
           │                      ├─Search the database
           │                      ├─View the database
           │                      ├─Edit the cumulative search results
           │                      └─Quit and return to Main Menu
           ├─Use search results to extract data
           ├─Extract data from the database──┬
           │                                 ├─Select the variables to extract
           │                                 ├─Edit the extract controls
           │                                 ├─Extract data from the database
           │                                 └─Quit and return to the Main Menu
           └─Quit and return to DOS or Host
```

Making Selections and Moving from Menu to Menu

As can be seen from Figure 1, there is a logical order to the menus in terms of the various choices which must be made before an action is taken.

Once a choice is made it is registered by pressing the ENTER key. If a choice is made which is invalid, or is out of the logical order, (e.g., a search is asked for but no variable has been selected to search on) QUERY will normally display an error message or will not accept the choice.

In registering some choices by pressing ENTER, control sometimes automatically passes to the next logical menu selection. In other cases, the user must return to the previous menu level by pressing the escape (Esc) key. As a general rule, the user should press the ENTER key to ensure that his or her choice is registered in the computer.

Main Menu

As can be seen from Figure 1, the Main Menu contains five options:

1. Select Which Database to Search

QUERY, by default, will automatically choose the database defined in the Host "RUN" option when QUERY is run from Host or from the Host "system environment" file when run from DOS. However, the user may select an alternate database to use during a QUERY session. To change the database, select the database from the list presented and press ENTER. Note that the currently active database is indicated in the lower left hand corner of the screen.

2. Search the Database

Overview of the Search Process

The search option allows the user to perform one of the two main functions of QUERY. Search allows the user to search records in the database which correspond to particular criteria which are specified. An example of a search would be to identify all countries whose gross enrollment rate (GER) for girls in primary school in 1980 was between 50 and 75. This can be eventually combined, for example, with a second search for countries whose gross national product per capita in the same year was between $500 and $1000.

In order to understand how the search process works it is necessary to understand how variables in a database are constructed. Almost all variables have at least one dimension and in the case of the WID database they have normally two (period and country). See the Appendix for a list of the variables and their dimensions. In conducting a search, three types of choices are made:

1. the variable on which to search;
2. the search range of values of the variable (minimum and maximum); and
3. the particular levels of each of the variable's dimensions across which the search will be performed (the "search controls").

In the example cited above, the selected variable is GER, the "search range" is a minimum of 50 and a maximum of 75, and the "search controls" would be set to confining the search to education level equal to primary, sex equal to female, and period equal to 1980. The user may select more than one level of a search control (e.g., education levels primary and secondary), but this may not lead to interpretable results since the same country may be selected twice (one for each level of the education system).

While normally the user will be interested in identifying which COUNTRIES correspond to the specified criteria (i.e., the country dimension can be considered the "record") this does not have to be the case; QUERY will act across ANY Host database dimension. For example, one might wish to know in which YEARS the infant mortality rate was below 100 for African countries. In this case a search would be initiated on the infant mortality rate, confined to African countries but across all years in the database.

Below, each of the sub-menus of the Search Menu and their functions are discussed.

Clear all Search Results

This is used to restart the entire search process. This may sometimes be desired because during a search session the levels of some search controls (dimensions) are "turned on" or "off" as a result of search operations. The user may change their mind and wish to start over again. This selection will clear all the previous results and leave all search controls "on".

Select the Search Variable

This entry allows the user to specify which variable will form the basis of the search. Selecting a variable is obligatory for the search process and all of the submenu items below this one will display an error message if no variable had been selected. Only one search variable at a time can be selected.

To choose a variable simply move the cursor to the desired variable and press the ENTER key. Control will then return to the Search Menu and the cursor will be positioned on the next submenu item (Set the search range). To change the search variable, simply re-enter this submenu and choose another variable. Note that the currently active search variable is indicated at the bottom of the screen.

If the escape (Esc) key is pressed, no (new) variable will be selected and control will return to the Search Menu.

Set the Search Range

Here the user specifies the range of values over which the search will be performed. Normally this will involve the setting of maximum and minimum values. For example, as indicated above, a search may wish to be

made for countries with a GER of between 50 (the minimum) and 75 (the maximum).

Since the search range is inclusive of the maximum and the minimum values, users should be careful particularly when 0 is a bound on the range. For example, if one is searching for occurrences of a variable below 0 then the maximum should be a very small negative number (e.g.,-0.000001) rather than 0.

In the case of a "string variable" , such as REGION, the user will enter the "string" or word over which the search would be confined. For example, if the search is to be confined to Africa to user would enter "AFRICA" as both the maximum and the minimum.

Note that if the escape key is pressed BEFORE the ENTER key in this process then control is returned to the search menu and the entry of the maximum or minimum (depending on where the cursor was when the escape key was pressed) may not be registered.

The search range remains valid until it is changed by re-entering this menu or until a new search variable is selected.

Edit the Search Controls

In addition to the selection of a variable, a search also involves selection (or deselection) of the dimensions of the variable over which the search will be confined. This process is referred to as setting the search controls. For example, if the search is over the variable "life expectancy" which has as dimensions country, period and sex, the user may wish to confine the search to values in the year 1980 for females; in this case the search controls for period will be set to 1980 only and for sex to female. Country would not be set since the search would be set over the levels of this dimension.

In entering this submenu the user is presented with a list of those dimensions which pertain to the search variable. The user selects the dimensions or search controls which they wish to set and is subsequently presented with a list of the levels of the dimension of the variable - male, female, for example. To set the control for the dimension simply move the cursor to the dimension levels which are not wanted and use the space bar to "toggle" it on or off. Several function keys are also available to aid in selection as explained on the screen. Note that all levels of all dimensions are "on", prior to a search or to the setting of the search controls so that the selection process is essentially a "deselection" process. These selections will remain in force until otherwise changed.

While most searches will involve the setting of the search controls, this step is not an obligatory one. However, interpretation of the results may be difficult unless the search controls are set and in fact the search may have little meaning. For example, searching on a periodic variable without picking a particular year may result in the same country being

counted several times during the search. Also, the value of a search variable may fall outside of the search range in one year and within it in another. In this case, QUERY would select the country during the search process even though it contains some values outside of the search range.

As with other submenus, one returns to the Search Menu by pressing the escape key.

Search the Database

Once the variable has been selected, the search range specified, and the search controls set, the search can be carried out. Selection of this submenu item will implement the search. As the search is carried out a count is given of the number of occurrences of the selected variable within the search range and according to the search controls. In addition, the mean, standard deviation, and minimum and maximum found are reported.

Once performed, control is returned to the Search Menu and the cursor is placed on the next submenu which allows review of the results of the search.

View the Search Results

Viewing the search results means seeing which levels of the (unrestricted dimensions of the) search variable have values which fall within the search range. Upon entering this menu item, the user is presented with a list of the dimensions pertaining to the search variable. The levels of some of these may have been previously restricted by the search controls operation. As discussed above, most of the time the user will be interested in seeing which countries meet the search criteria for the search variable. To see which countries meet the criteria, simply move the cursor to the country dimension and press enter. Countries which meet the criteria will be highlighted and have a small diamond sign next to them.

For example, if the search is over the variable GRENRTE (gross enrollment rate by level) and the search controls for sex and level had previously been set at female primary, the three dimensions, sex, level and country will appear on the screen. Selecting sex will reveal the "female" level highlighted and selecting level will reveal "primary" level highlighted. The country dimension will have these countries highlighted which meet the search range criteria.

Edit Cumulative Search Results

As indicated above, several searches may be performed during a QUERY session. These searches are carried out serially, i.e., after another. This submenu allows one to see the cumulative results of previous searches and to combine a just completed search with previous ones.

Operations which can be performed in this part of QUERY essentially amount to set operations with sets of dimension levels which correspond to

the search criteria. The set operations are performed using the indicated function keys.

Set operations which are allowed are intersection, union, and difference. Consider two examples. Suppose we wish to find those countries where the GER for females in 1980 is greater than 70 and where the GNP_PC for 1980 is less than $500. This would be performed by combining the results of a search on GER and one on GNP_PC where the appropriate search controls and ranges had been set. The combination would be performed with the intersection operation since this would list countries where both criteria held. If, on the other hand, one were interested in countries where either the GER > 70 or GNP_PC < $500 then the combination would be performed with the union operation.

Once an operation is performed and the user exits the submenu the results become the "cumulative" search results and are available for further operations with the results of subsequent searches. The operations are always pairwise. Hence if the user has conducted three searches and combined in some way the results of the first two, he can combine the results of search three with the cumulative results but cannot combine the results of search one and three.

3. Use Search Results to Extract Data

The two main functions of QUERY, search and extract, are logically different operations. However, the two can be linked via the use of the "Use Search Results to Extract Data" menu item. Entering this menu and confirming with a "Yes" will simply transfer the cumulative search results, in terms of the dimensions levels which are "on", for use in the Extract menu.

This feature is useful if the user wishes to create a file which is a subset of the master database and when the countries (or records) in the new database correspond to certain criteria. For example, the user may wish to create a file with entries for the variables for only African countries. In this case, a search over the variable REGION would be performed for Africa and the search results transferred to the extraction process.

4. Extract Data From the Database

Overview of the Extract Process

The extraction process consists of the creation of a new database which is stored in a new file. The new database will consist of a subset of the database which was chosen at the beginning of the session (the first menu item of the Main Menu or the database specified when QUERY was run). The process consists of selecting variables, setting controls on the dimensions, naming the target database file and carrying out the extraction process.

Select the Variables to Extract

The first decisions that are made pertain to which variables are to be in the new database. Entering this submenu the user simply "toggles" on and off the variables which are wanted in the new database. Function keys are also available to speed up the selection process. Unlike the selection of a search variable, extract allows one to select as many variables as wanted. At least one variable must, however, have been selected.

Although search results may be used to extract data, the variables chosen under extract need not be the same as those chosen as search variables. For example, a search may be performed for all countries with GNP per capita below a certain level in Africa, but the user may wish to establish a new database with variables for life expectancy and fertility.

Edit the Extract Controls

As with the search process, once a variable or variables have been selected, the controls on their dimension levels need to be specified. Here, the process is similar to that in the search process with a few differences. First, some controls may already be set off if the user has transferred some search results; these can always be edited and changed at this stage. Second, the list of dimensions which the user is presented with is the list of <u>all</u> dimensions in the source database rather than only those pertaining to the extract variables. Some of these, therefore, may not be relevant.

Care should be taken to ensure that at least ONE level of each of the pertinent dimensions is on. If this is not the case the extract process will automatically stop when the problem is encountered. Also, since variables often share the same dimensions, the user should always confirm that the extract controls correspond to the levels desired. Of course controls on the same dimensions cannot be different for different variables to be extracted. That is, one cannot extract the GER for males and life expectancy for females.

Extract Data From the Database

Once the extract variables and controls have been selected and set, the extract process can be carried out. The user is prompted to confirm that they want to continue and then is prompted to specify a name and description of the database. A description is useful because it can contain information about, for example, the criteria for extraction of the data or the search criteria used to select the countries.

The extraction process can take several minutes depending on how many variables were selected and how much data is to be extracted. During the process, the screen will inform the user of which variable it is currently extracting. As indicated above, if the setting of the extract controls was insufficient to allow extraction of a variable, a message is displayed on the screen and the process stops.

As with any menu or submenu, use the escape key to return to a higher level in the menu tree.

Quit and Return to the Main Menu

Once the extraction process is finished, the user can return to the Main Menu by selecting this item. At this point, the user can start another search or conduct another extraction.

5. Quit and Return to DOS or Host

To terminate a QUERY session select the bottom item on the Main Menu and press ENTER. Control will be passed to either to DOS or to Host depending upon where the user entered QUERY.

Using Databases Created with Extract in Host

On leaving QUERY, if the user has used extract, the current database will be the newly created database. It will have only the variables selected with data corresponding to the extract controls. For all dimensions except PERIOD, there will only be levels of the dimensions corresponding to those specified in the extract controls. However, the database will continue to allow for the original number of levels of the period dimension although there will only be data for the periods requested.

For example, suppose that the gross enrollment rate for all countries for females in primary has been extracted. In this case, while the dimensions for sex and education level will be retained in the definition of the variable in the extracted file, only the levels for females and primary will be in the database.

Database selection in GENDER works differently than in other Host applications because of the fact that extract creates a new database. As mentioned above, on exiting QUERY the current database will be changed automatically to the one created during the extract process. While the original database will still be viewable on the directory of the hard disk and still exists, only the CURRENT database (as defined in the CONFIGURE menu of Host as the default database) will be usable. **To view or use any other Host databases in the GENDER system, the user must re-enter QUERY, select the desired database using the "Select database" option, and then quit QUERY.** This will re-set the current database to the selected one. Unlike other Host applications, it is not possible to change the current database with the Host configure facility. Of course after this re-setting process, any OTHER databases will also not be immediately usable outside of QUERY.

To see any of the data in a database, the user can use the powerful TABLULATE and GRAPH facilities of Host. These allow the user to define as many tables and graphs as desired and to display them on the console, in a file, or to print them on the printer.

Producing Predefined Tables and Graphs

With the aid of the Host software, users of the GENDER system can produce predefined tables and graphs of the system variables. Of particular interest in this regard are the output variables of the projection modules POPPROJ and EDUPROJ. While GENDER is distributed with a table and graph library, trained Host users can produce their own tables and graphs.

In Tables 1 and 2 a list of the tables and graphs which are distributed with GENDER can be found. Sample of tables and graphs are also presented below.

Table 1

List of Predefined Tables

1.	ASFR:	Age Specific Fertility Rates
2.	COHORT:	Educational Cohort components
3.	COHORT2:	Educational Cohort components
4.	DEMOGTBL:	Demographic Variables
5.	DINDPROJ:	Projected Dispartiy Index (Primary)
6.	DROPRT:	Dropout Rate by Grade
7.	ENROLHIS:	Primary School Total Enrollment (Historic)
8.	ENROLPOP:	School Population vs. Enrollment (Total)
9.	ENROLPRJ:	Primary School Enrollment by Grade
10.	ENROLTBL:	Primary School Total Enrollment
11.	ENTERRT:	Primary School Entrance Rate
12.	ENTPOP:	School Entrance Population
13.	GERPROJ:	Primary School Gross Enrollment Rate
14.	HISTGER:	Gross Enrollment Rate (Historic)
15.	NETFLOW:	Net Flow Rates (Historic)
16.	POP5:	Population (5 year age groups)
17.	POPTOTAL:	Total Population
18.	REPEATRT:	Repetition Rate
19.	SCHLPROJ:	Primary Schools
20.	SOCIOVAR:	Socioeconomic Variables
21.	TEACHPRJ:	Primary School Teachers
22.	TFR:	Total Fertility Rate

Table 2

List of Predefined Graphs

1.	COHORT:	Graduation Rates
2.	DINDEX:	Primary School Disparity Index
3.	ENROL:	Primary School Enrollment
4.	ENROLCMP:	Primary School Enrollment (Comparison)
5.	ENROLPOP:	Enrollment vs. School Enrollment Population
6.	ENTPOP:	Entrance Population
7.	ENTRATE:	Entrance Rates
8.	FLFPR:	Female Labor Force Participation

9. GDP_GR: GDP Growth Rate
10. GER Gross Enrollment Rate
11. GERCOMP: Gross Enrollment Rate (Comparison)
12. IMRMF: Infant Mortality Rate
13. POP: Population
14. RATES: Dropout and Repeater Rates
15. SCHOOLS: Primary Schools
16. SCPOPCMP: School Age Population (Comparison)

Table 3

Educational Cohort Components

GRADRATE : Graduates (COHORT)
IORATE : Input output ratio (COHORT)
ONTIME : On time graduates (COHORT)
SYGRAD2 : Student years per graduate (COHORT)
YRSSCHL : Average years in primary school (COHORT)

PERIOD : 1985

COUNTRY/ BY SEX	GRADRATE			IORATE			ONTIME			SYGRAD2		
	FEMALE	MALE	TOTAL	FEMALE	MALE	TOTAL	FEMALE	MALE	TOTAL	FEMALE	MALE	TOTAL
BANGLADESH	335.68	137.01	194.74	1.557	2.923	2.259	335.7	137.0	194.7	7.79	14.61	11.29
BELIZE	0.00	0.00	0.00	0.000	0.000	0.000	0.0	0.0	0.0	0.00	0.00	0.00
BOLIVIA	384.80	451.35	419.53	1.671	1.534	1.594	384.8	451.4	419.5	13.37	12.27	12.75
BOTSWANA	895.99	815.33	864.32	1.121	1.194	1.150	569.9	504.2	541.7	7.84	8.36	8.05
BURKINA FASO	718.55	733.20	734.24	1.461	1.400	1.412	209.6	236.9	228.6	8.77	8.40	8.47
BURMA	0.00	0.00	0.00	0.000	0.000	0.000	0.0	0.0	0.0	0.00	0.00	0.00
BURUNDI	833.73	837.60	835.83	1.325	1.295	1.307	210.7	240.4	228.6	7.95	7.77	7.84
CAMEROON	132.83	154.70	144.38	7.482	6.764	7.077	15.4	15.7	15.6	52.38	47.35	49.54
CAPE VERDE	287.35	314.20	300.50	3.357	3.151	3.252	35.7	37.4	36.6	20.14	18.91	19.51
CENTRAL AFRICAN REPUBLIC	282.83	637.94	489.98	3.296	2.056	2.371	16.5	37.6	28.6	19.78	12.33	14.23
CHAD	0.00	0.00	0.00	0.000	0.000	0.000	0.0	0.0	0.0	0.00	0.00	0.00
COMOROS	0.00	0.00	0.00	0.000	0.000	0.000	0.0	0.0	0.0	0.00	0.00	0.00
CONGO	791.85	762.56	777.02	1.645	1.777	1.710	71.7	58.5	64.7	9.87	10.66	10.26
COSTA RICA	683.69	624.51	652.68	1.235	1.306	1.271	683.7	624.5	652.7	7.41	7.84	7.62
CYPRUS	11.77	11.74	11.77	73.229	73.403	73.301	11.3	11.4	11.3	*****	*****	*****
DJIBOUTI	752.08	804.86	782.25	1.127	1.099	1.110	752.1	804.9	782.2	6.76	6.59	6.66
DOMINICAN REPUBLIC	0.00	0.00	141.14	0.000	0.000	2.933	0.0	0.0	116.7	0.00	0.00	17.60
ECUADOR	0.00	0.00	574.54	0.000	0.000	1.381	0.0	0.0	351.0	0.00	0.00	8.29
EGYPT	877.91	867.04	871.45	1.069	1.078	1.074	877.9	867.0	871.4	6.41	6.47	6.45
EL SALVADOR	345.63	279.01	310.42	1.783	2.076	1.923	203.8	161.3	181.2	16.04	18.69	17.31
EQUATORIAL GUINEA	0.00	0.00	0.00	0.000	0.000	0.000	0.0	0.0	0.0	0.00	0.00	0.00
FIJI	523.46	489.88	507.72	1.680	1.786	1.730	423.5	357.8	389.9	13.44	14.29	13.84
GAMBIA	716.61	894.83	834.08	1.401	1.276	1.311	255.2	298.1	283.8	8.41	7.66	7.87
GHANA	0.00	0.00	0.00	0.000	0.000	0.000	0.0	0.0	0.0	0.00	0.00	0.00
GRENADA	0.00	0.00	140.48	0.000	0.000	3.710	0.0	0.0	34.0	0.00	0.00	29.68
GUATEMALA	0.00	0.00	287.02	0.000	0.000	1.815	0.0	0.0	287.0	0.00	0.00	10.89
GUINEA	365.89	435.64	411.13	2.337	2.114	2.184	37.8	58.6	51.1	14.02	12.69	13.10
GUINEA BISSAU	238.39	270.13	256.11	4.705	4.571	4.630	4.5	7.3	6.3	28.23	27.43	27.78
HAITI	143.13	159.94	151.93	3.207	2.967	3.076	58.3	70.2	64.4	19.24	17.80	18.45
HONDURAS	463.27	395.15	428.11	1.680	1.864	1.768	218.3	171.1	193.4	10.08	11.19	10.61

COUNTRY/ BY SEX	GRADRATE FEMALE	GRADRATE MALE	GRADRATE TOTAL	IORATE FEMALE	IORATE MALE	IORATE TOTAL	ONTIME FEMALE	ONTIME MALE	ONTIME TOTAL	SYGRAD2 FEMALE	SYGRAD2 MALE	SYGRAD2 TOTAL
INDIA	0.00	0.00	0.00	0.000	0.000	0.000	0.0	0.0	0.0	0.00	0.00	0.00
INDONESIA	631.26	652.07	754.35	1.282	1.254	1.284	631.3	652.1	433.7	7.69	7.52	7.70
ISRAEL	950.72	910.51	931.13	1.027	1.048	1.037	950.7	910.5	931.1	8.21	8.39	8.30
JAMAICA	838.01	801.42	821.66	1.180	1.245	1.211	648.8	594.5	621.6	7.08	7.47	7.26
JORDAN	963.93	926.58	955.99	1.077	1.094	1.084	720.2	682.3	708.8	6.46	6.57	6.50
KENYA	499.31	542.18	522.49	1.429	1.305	1.360	499.3	542.2	522.5	10.01	9.14	9.52
LEBANON	0.00	0.00	0.00	0.000	0.000	0.000	0.0	0.0	0.0	0.00	0.00	0.00
LESOTHO	516.46	367.55	446.62	1.807	2.220	1.965	98.0	59.2	78.4	12.65	15.54	13.76
LIBERIA	0.00	0.00	0.00	0.000	0.000	0.000	0.0	0.0	0.0	0.00	0.00	0.00
MADAGASCAR	0.00	0.00	318.35	0.000	0.000	1.731	0.0	0.0	318.3	0.00	0.00	8.65
MALAWI	278.85	399.16	349.92	2.307	1.832	1.976	62.4	97.9	82.8	18.45	14.65	15.81
MALI	317.70	378.09	355.27	2.887	2.504	2.633	37.1	48.4	44.0	17.32	15.02	15.80
MAURITANIA	0.00	0.00	0.00	0.000	0.000	0.000	0.0	0.0	0.0	0.00	0.00	0.00
MAURITIUS	990.32	984.28	987.77	1.006	1.006	1.006	990.3	984.3	987.8	6.04	6.03	6.03
MOROCCO	662.19	705.29	687.87	1.505	1.514	1.511	236.2	218.7	225.1	7.52	7.57	7.56
MOZAMBIQUE	152.38	263.73	212.47	4.601	2.961	3.503	40.0	82.8	62.3	18.40	11.85	14.01
NEPAL	332.86	330.77	330.85	1.542	1.586	1.575	332.9	330.8	330.8	7.71	7.93	7.88
NIGER	809.10	793.68	799.18	1.108	1.118	1.114	809.1	793.7	799.2	6.65	6.71	6.69
OMAN	921.13	931.01	928.17	1.190	1.242	1.222	460.4	354.2	390.3	7.14	7.45	7.33
PAKISTAN	380.74	626.97	551.99	1.616	1.150	1.245	380.7	627.0	552.0	8.08	5.75	6.22
PANAMA	828.06	798.29	812.92	1.212	1.288	1.250	442.7	332.2	381.7	7.27	7.73	7.50
PERU	521.79	646.42	584.60	1.320	1.163	1.233	521.8	646.4	584.6	7.92	6.98	7.40
PHILIPPINES	696.70	605.01	648.29	1.201	1.295	1.247	641.9	558.0	597.6	7.20	7.77	7.48
RWANDA	387.76	453.48	422.15	2.012	1.725	1.851	161.1	183.5	173.1	16.09	13.80	14.81
SENEGAL	754.67	871.09	829.42	1.448	1.297	1.345	218.6	272.4	251.7	8.69	7.78	8.07
SEYCHELLES	919.59	943.14	941.65	1.047	1.025	1.027	919.6	943.1	941.7	9.42	9.23	9.24
SIERRA LEONE	0.00	0.00	0.00	0.000	0.000	0.000	0.0	0.0	0.0	0.00	0.00	0.00
SOMALIA	500.65	474.98	482.87	1.386	1.380	1.383	500.6	475.0	482.9	11.09	11.04	11.06
SRI LANKA	842.81	813.47	850.20	1.125	1.162	1.235	842.8	813.5	468.4	6.75	6.97	7.41
SUDAN	686.26	681.43	683.81	1.181	1.204	1.194	686.3	681.4	683.8	7.09	7.23	7.17
SWAZILAND	673.72	636.32	654.50	1.401	1.489	1.444	288.3	205.3	243.3	9.81	10.42	10.11
THAILAND	0.00	0.00	825.61	0.000	0.000	1.097	0.0	0.0	825.6	0.00	0.00	6.58
TOGO	224.78	431.46	342.41	4.194	2.502	2.984	13.6	31.1	23.0	25.17	15.01	17.91
TUNISIA	10.77	16.32	13.85	81.055	57.178	65.386	2.6	3.2	2.9	*****	*****	*****
TURKEY	928.72	926.09	927.32	1.135	1.140	1.138	603.5	589.1	595.8	5.68	5.70	5.69
UGANDA	643.53	814.81	756.34	1.419	1.220	1.273	275.2	353.0	326.0	9.93	8.54	8.91
YEMEN	493.46	445.93	450.82	1.379	1.507	1.492	493.5	445.9	450.8	8.28	9.04	8.95
ZAIRE	625.42	522.49	605.87	1.433	1.498	1.427	179.6	147.0	172.5	8.60	8.99	8.56
ZAMBIA	727.00	883.82	833.43	1.250	1.093	1.136	651.7	782.6	742.3	8.75	7.65	7.95
ZIMBABWE	692.44	778.54	739.98	1.205	1.107	1.148	655.7	734.3	699.3	8.44	7.75	8.03

COUNTRY/ BY SEX	YRSSCHL FEMALE	MALE	TOTAL
BANGLADESH	2.6	2.0	2.2
BELIZE	0.0	0.0	0.0
BOLIVIA	5.1	5.5	5.4
BOTSWANA	7.0	6.8	7.0
BURKINA FASO	6.3	6.2	6.2
BURMA	0.0	0.0	0.0
BURUNDI	6.6	6.5	6.6
CAMEROON	7.0	7.3	7.2
CAPE VERDE	5.8	5.9	5.9
CENTRAL AFRICAN REPUBLIC	5.6	7.9	7.0
CHAD	0.0	0.0	0.0
COMOROS	0.0	0.0	0.0
CONGO	7.8	8.1	8.0
COSTA RICA	5.1	4.9	5.0
CYPRUS	6.0	6.0	6.0
DJIBOUTI	5.1	5.3	5.2
DOMINICAN REPUBLIC	1.0	1.0	2.5
ECUADOR	1.0	1.0	4.8
EGYPT	5.6	5.6	5.6
EL SALVADOR	5.5	5.2	5.4
EQUATORIAL GUINEA	0.0	0.0	0.0
FIJI	7.0	7.0	7.0
GAMBIA	6.0	6.9	6.6
GHANA	0.0	0.0	0.0
GRENADA	1.0	1.0	4.2
GUATEMALA	1.0	1.0	3.1
GUINEA	5.1	5.5	5.4
GUINEA BISSAU	6.7	7.4	7.1
HAITI	2.8	2.8	2.8
HONDURAS	4.7	4.4	4.5
INDIA	0.0	0.0	0.0
INDONESIA	4.9	4.9	5.8
ISRAEL	7.8	7.6	7.7
JAMAICA	5.9	6.0	6.0
JORDAN	6.2	6.1	6.2
KENYA	5.0	5.0	5.0
LEBANON	0.0	0.0	0.0
LESOTHO	6.5	5.7	6.1
LIBERIA	0.0	0.0	0.0

COUNTRY/ BY SEX	YRSSCHL FEMALE	MALE	TOTAL
MADAGASCAR	1.0	1.0	2.8
MALAWI	5.1	5.8	5.5
MALI	5.5	5.7	5.6
MAURITANIA	0.0	0.0	0.0
MAURITIUS	6.0	5.9	6.0
MOROCCO	5.0	5.3	5.2
MOZAMBIQUE	2.8	3.1	3.0
NEPAL	2.6	2.6	2.6
NIGER	5.4	5.3	5.3
OMAN	6.6	6.9	6.8
PAKISTAN	3.1	3.6	3.4
PANAMA	6.0	6.2	6.1
PERU	4.1	4.5	4.3
PHILIPPINES	5.0	4.7	4.9
RWANDA	6.2	6.3	6.3
SENEGAL	6.6	6.8	6.7
SEYCHELLES	8.7	8.7	8.7
SIERRA LEONE	0.0	0.0	0.0
SOMALIA	5.5	5.2	5.3
SRI LANKA	5.7	5.7	5.3
SUDAN	4.9	4.9	4.9
SWAZILAND	6.6	6.6	6.6
THAILAND	1.0	1.0	5.4
TOGO	5.7	6.5	6.1
TUNISIA	7.0	7.5	7.2
TURKEY	5.3	5.3	5.3
UGANDA	6.4	7.0	6.7
YEMEN	4.1	4.0	4.0
ZAIRE	5.4	4.7	5.2
ZAMBIA	6.4	6.8	6.6
ZIMBABWE	5.8	6.0	5.9

Table 4

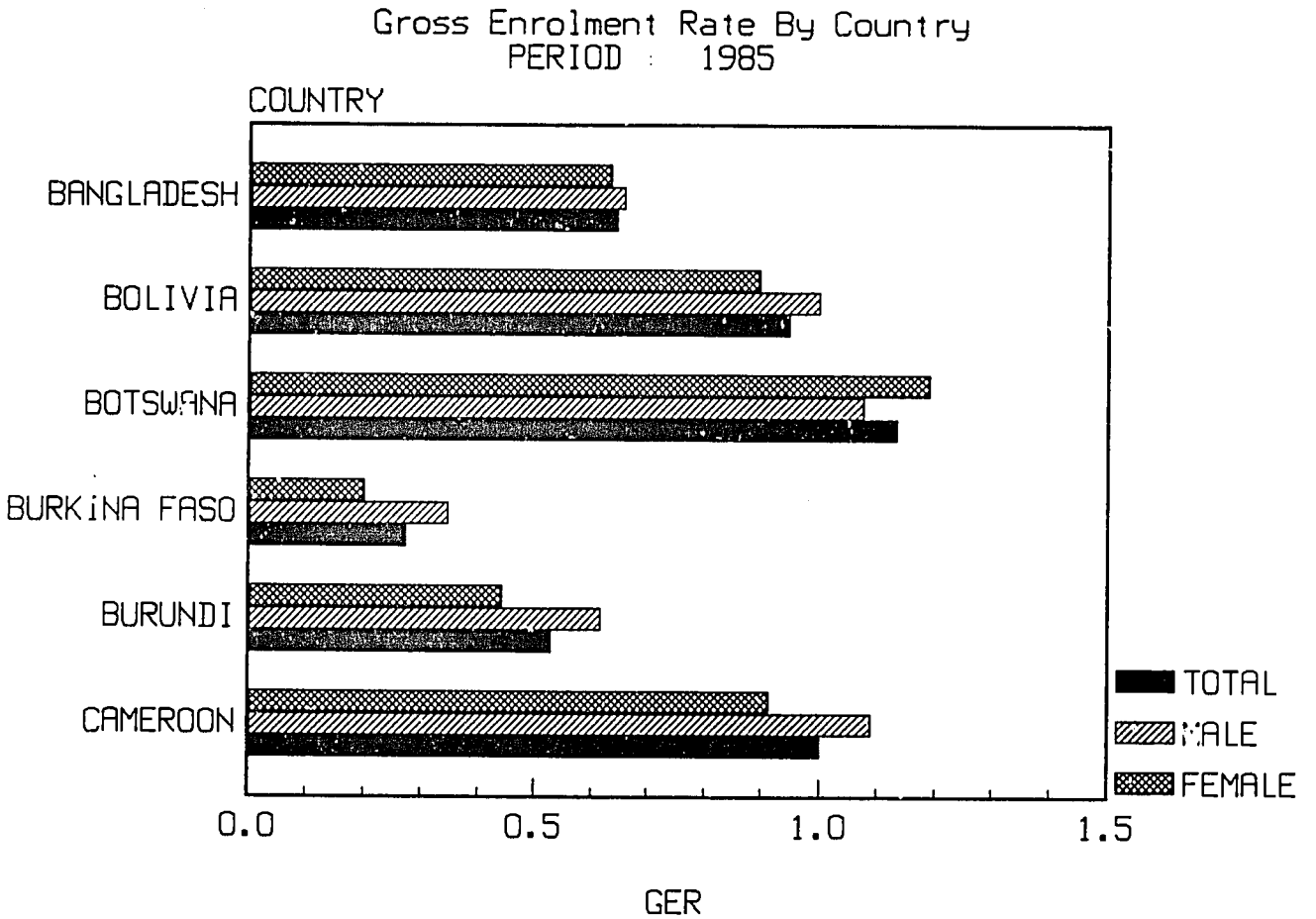
Demographic Variables

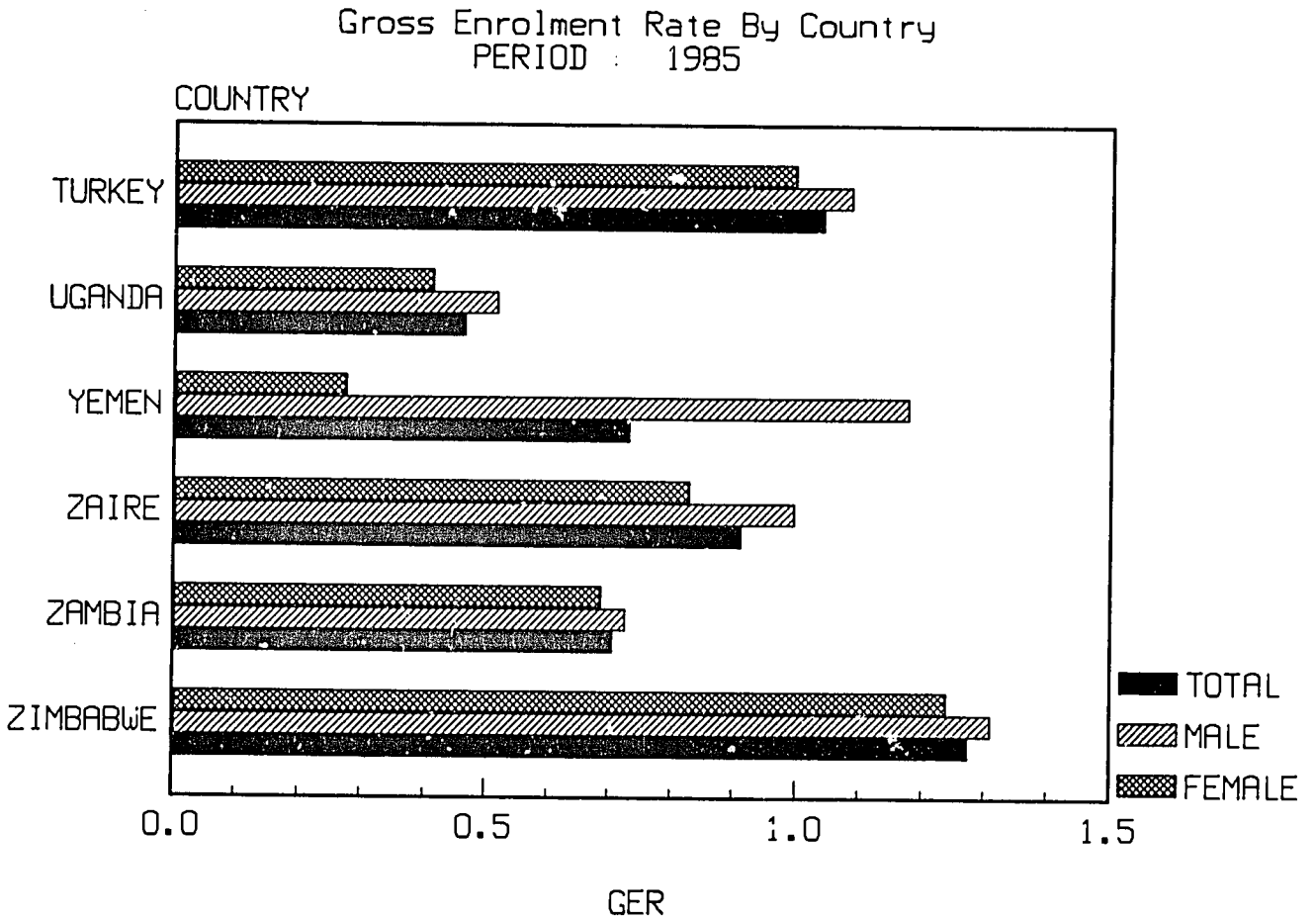
GR_RAT : Population growth rate
TOT_FER : Total fertility rate
IMRMF : Infant mortality rate (Male & Female)
LEXP : Life expectancy
RUR_POP : Proportion rural

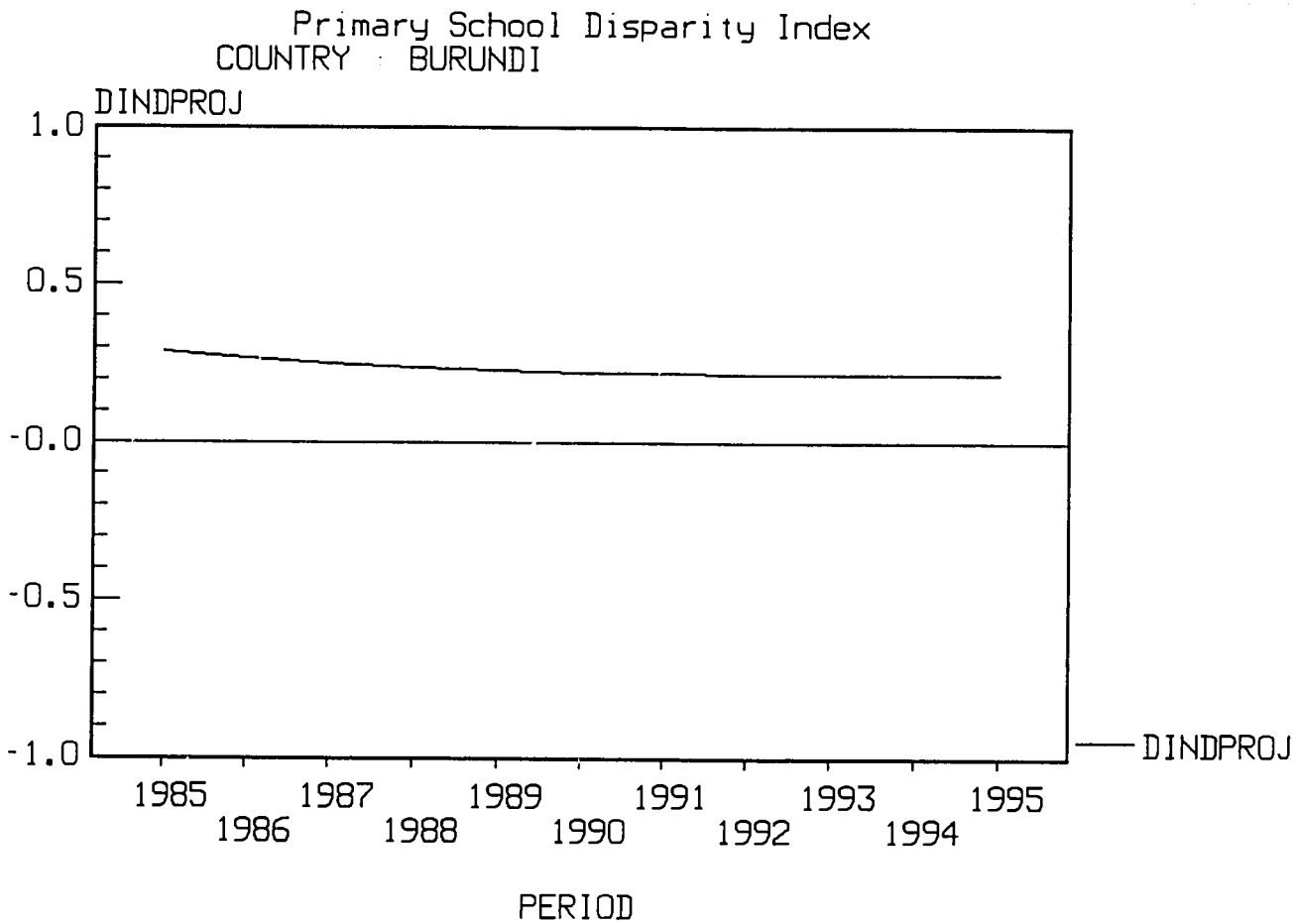
PERIOD : 1980

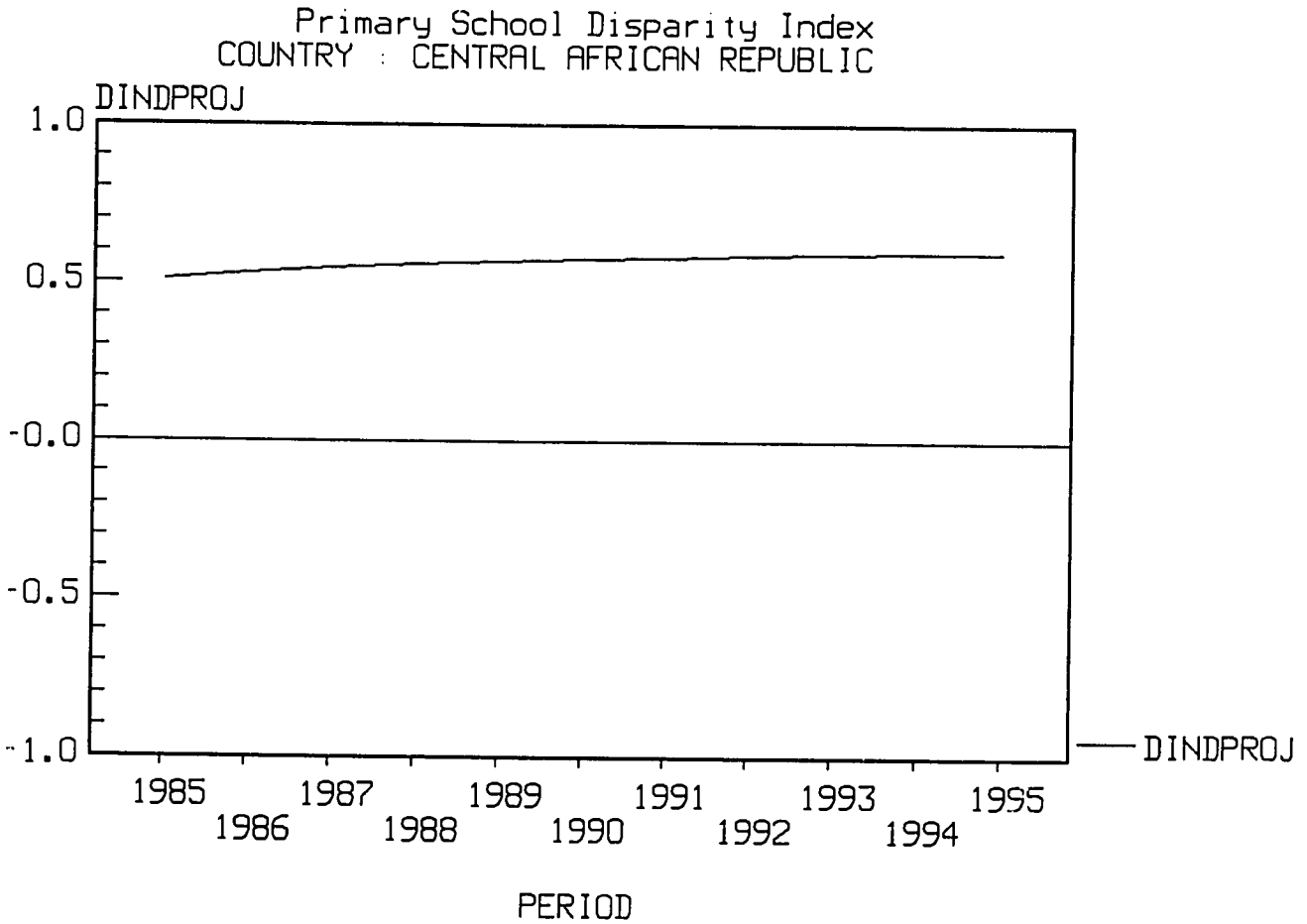
COUNTRY/	GR_RAT	TOT_FER	IMRMF	LEXP	RUR_POP
BANGLADESH	2.56	6.000	371.40	148.410	89.00
BELIZE	1.80	4.756	144.60	195.500	0.00
BOLIVIA	2.62	6.250	388.30	152.590	67.00
BOTSWANA	3.27	6.500	156.10	192.780	0.00
BURKINA FASO	1.56	6.500	469.10	133.380	0.00
BURMA	2.51	5.330	287.50	164.450	73.00
BURUNDI	2.64	6.500	371.50	139.910	98.00
CAMEROON	3.16	6.500	275.40	160.790	65.00
CAPE VERDE	2.02	5.750	167.50	190.060	0.00
CENTRAL AFRICAN REPUBLIC	2.45	5.500	355.20	143.180	59.00
CHAD	2.14	5.500	481.80	130.530	82.00
COMOROS	3.38	7.000	272.80	161.380	0.00
CONGO	3.45	6.000	166.00	190.430	63.00
COSTA RICA	2.61	3.500	49.60	221.400	57.00
CYPRUS	0.66	2.310	46.00	222.710	0.00
DJIBOUTI	0.00	6.800	366.00	141.030	0.00
DOMINICAN REPUBLIC	2.38	4.180	175.90	188.130	49.00
ECUADOR	2.81	5.416	174.90	188.340	55.00
EGYPT	2.20	4.629	310.20	172.210	45.00
EL SALVADOR	2.90	5.556	164.40	190.780	59.00
EQUATORIAL GUINEA	1.84	5.000	481.80	130.530	0.00
FIJI	2.03	3.800	0.00	191.960	0.00
GAMBIA	2.63	6.500	509.40	124.440	0.00
GHANA	3.86	7.000	226.30	182.420	64.00
GRENADA	1.80	3.638	116.80	202.110	0.00
GUATEMALA	2.84	5.166	213.00	179.840	61.00
GUINEA	1.95	6.000	570.20	113.070	81.00
GUINEA BISSAU	1.82	6.000	570.20	113.070	0.00
HAITI	1.70	4.649	346.70	162.710	65.00
HONDURAS	3.35	6.500	283.80	179.320	64.00
INDIA	2.15	4.800	283.00	164.860	78.00
INDONESIA	2.06	4.348	306.10	160.760	80.00
ISRAEL	2.05	3.090	46.00	222.720	11.00
JAMAICA	1.34	3.460	87.80	209.710	50.00
JORDAN	3.67	7.380	165.70	190.480	44.00
KENYA	4.25	8.000	232.00	171.170	86.00

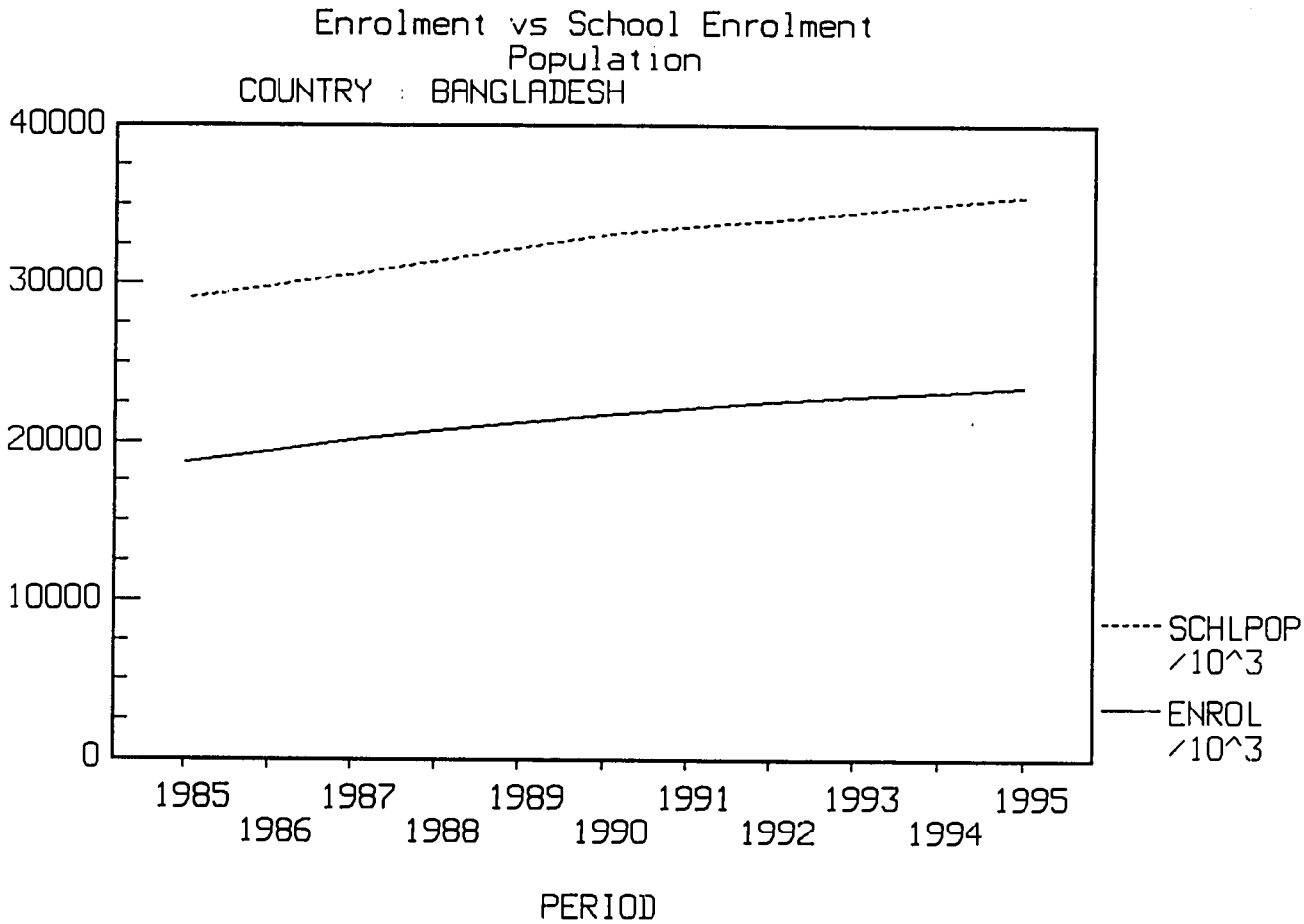
COUNTRY/	GR_RAT	TOT_FER	IMRMF	LEXP	RUR_POP
LEBANON	0.52	3.790	145.80	195.210	23.00
LESOTHO	2.44	5.800	282.40	159.170	95.00
LIBERIA	3.27	6.902	346.20	144.970	67.00
MADAGASCAR	2.92	6.455	340.90	146.140	82.00
MALAWI	3.06	7.600	516.60	132.370	91.00
MALI	2.29	6.500	396.50	135.070	80.00
MAURITANIA	2.04	6.000	391.40	136.040	77.00
MAURITIUS	1.64	2.760	126.10	199.910	0.00
MOROCCO	2.47	5.822	373.50	156.020	59.00
MOZAMBIQUE	2.72	6.490	388.20	136.490	91.00
NEPAL	2.42	6.250	433.90	137.550	95.00
NIGER	3.14	7.000	396.50	135.070	77.00
OMAN	3.78	7.070	370.50	156.730	0.00
PAKISTAN	2.59	5.840	361.20	150.020	72.00
PANAMA	2.12	3.460	76.70	212.830	46.00
PERU	2.30	4.495	145.80	175.550	33.00
PHILIPPINES	2.30	4.200	154.50	193.140	64.00
RWANDA	3.18	8.000	376.40	138.950	96.00
SENEGAL	2.82	6.650	391.40	135.990	75.00
SEYCHELLES	1.03	3.500	92.10	208.570	0.00
SIERRA LEONE	2.18	6.500	570.20	113.070	75.00
SOMALIA	2.83	6.772	399.30	134.550	70.00
SRI LANKA	1.98	3.374	95.50	207.660	73.00
SUDAN	2.71	6.580	363.00	141.620	75.00
SWAZILAND	3.60	6.982	283.00	159.020	0.00
THAILAND	1.89	3.400	173.40	188.710	86.00
TOGO	3.17	6.500	339.60	146.420	73.00
TUNISIA	2.38	4.935	193.80	184.080	48.00
TURKEY	2.12	4.100	169.30	189.110	53.00
UGANDA	3.25	7.000	329.20	147.370	88.00
YEMEN	2.53	6.760	476.00	131.790	90.00
ZAIRE	3.05	6.250	316.80	151.390	66.00
ZAMBIA	3.43	6.750	314.80	151.530	62.00
ZIMBABWE	3.96	7.000	246.30	166.800	77.00

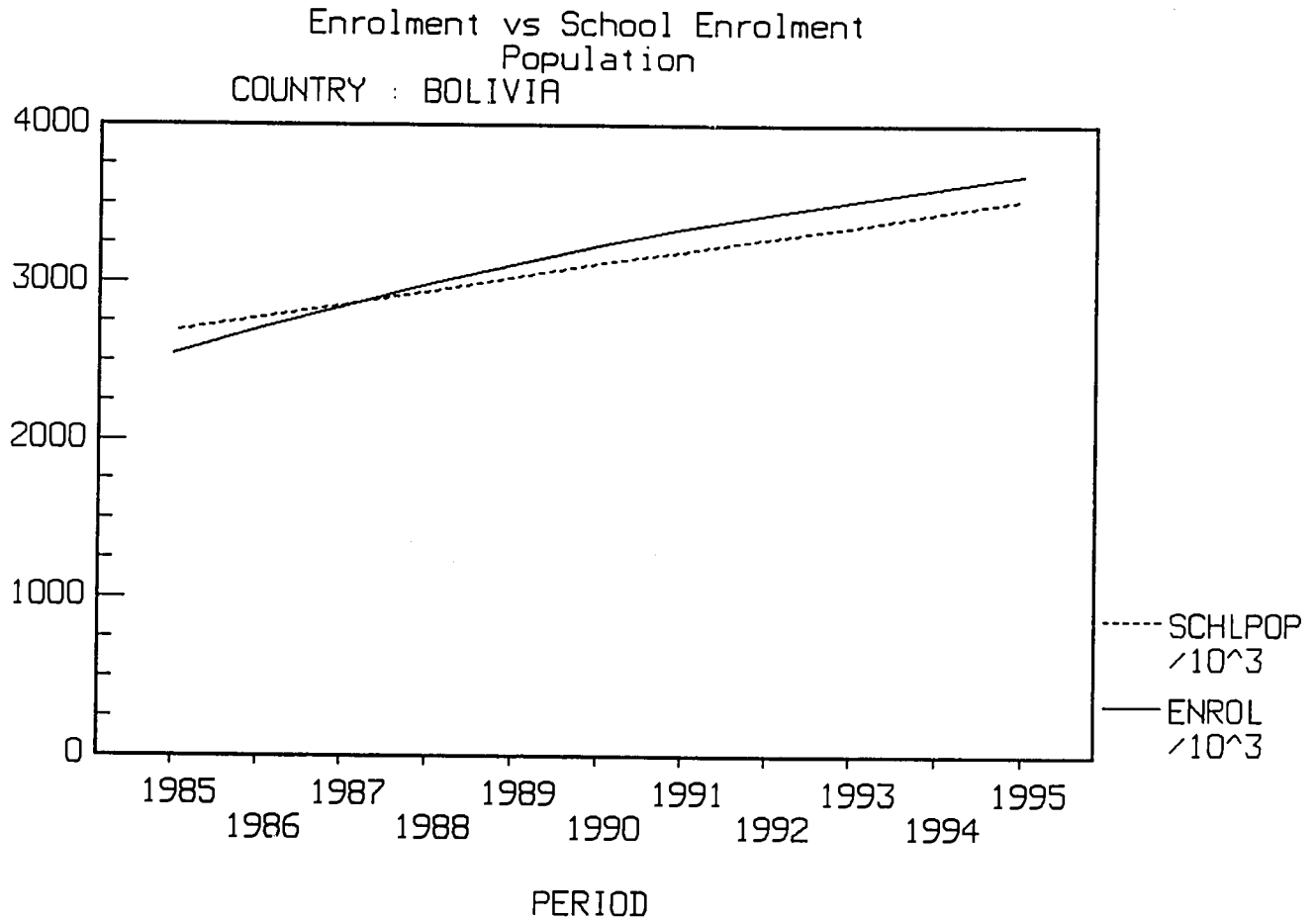

Gross Enrolment Rate By Country
PERIOD : 1985
COUNTRY
BANGLADESH
BOLIVIA
BOTSWANA
BURKINA FASO
BURUNDI
CAMEROON
TOTAL
MALE
FEMALE
0.0
0.5
1.0
1.5
GER

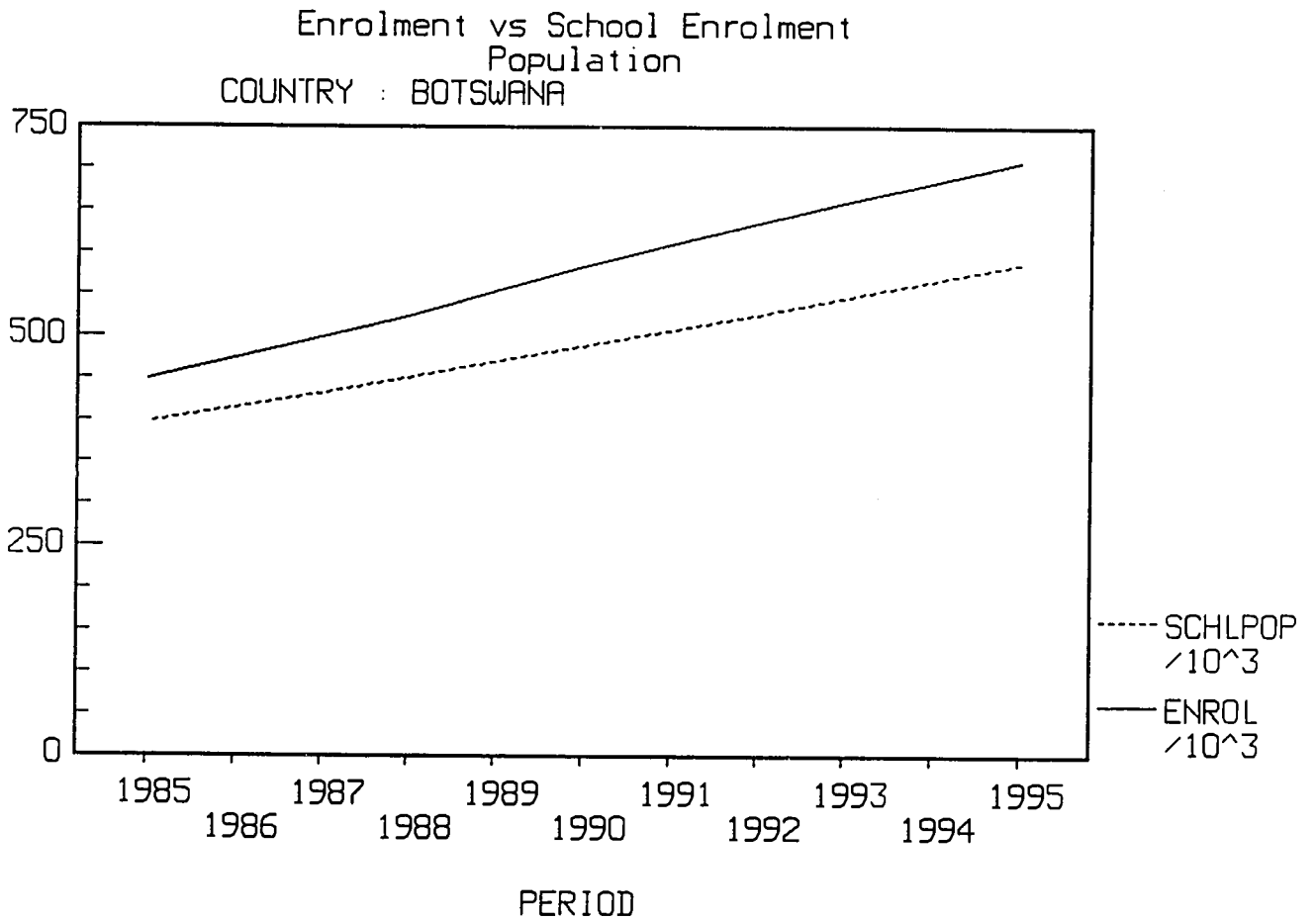

Gross Enrolment Rate By Country
PERIOD : 1985
COUNTRY
TURKEY
UGANDA
YEMEN
ZAIRE
ZAMBIA
ZIMBABWE
0.0
0.5
1.0
1.5
GER
TOTAL
MALE
FEMALE

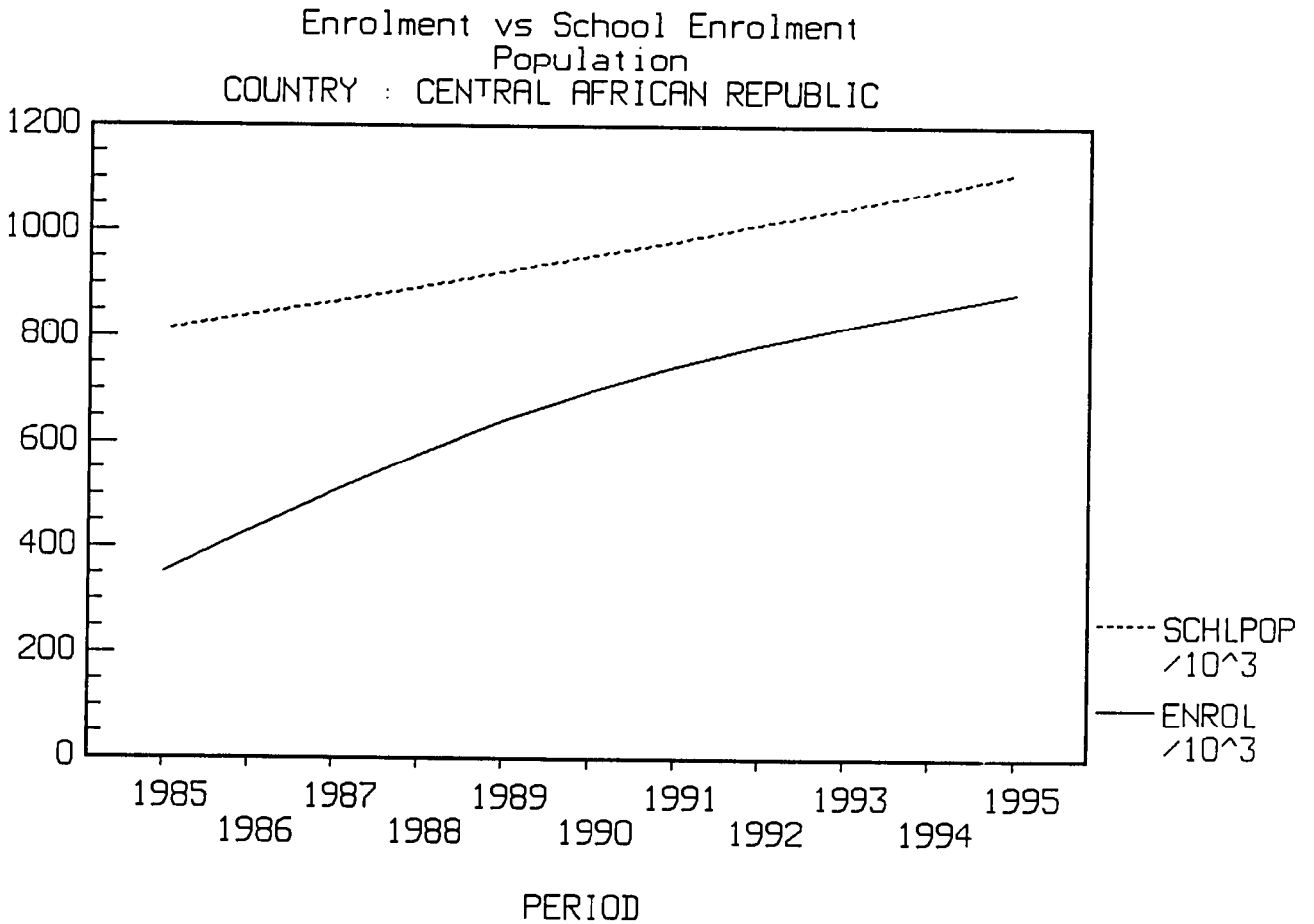

Primary School Disparity Index
COUNTRY BURUNDI
DINDPROJ
1.0
0.5
-0.0
-0.5
-1.0
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
PERIOD
DINDPROJ

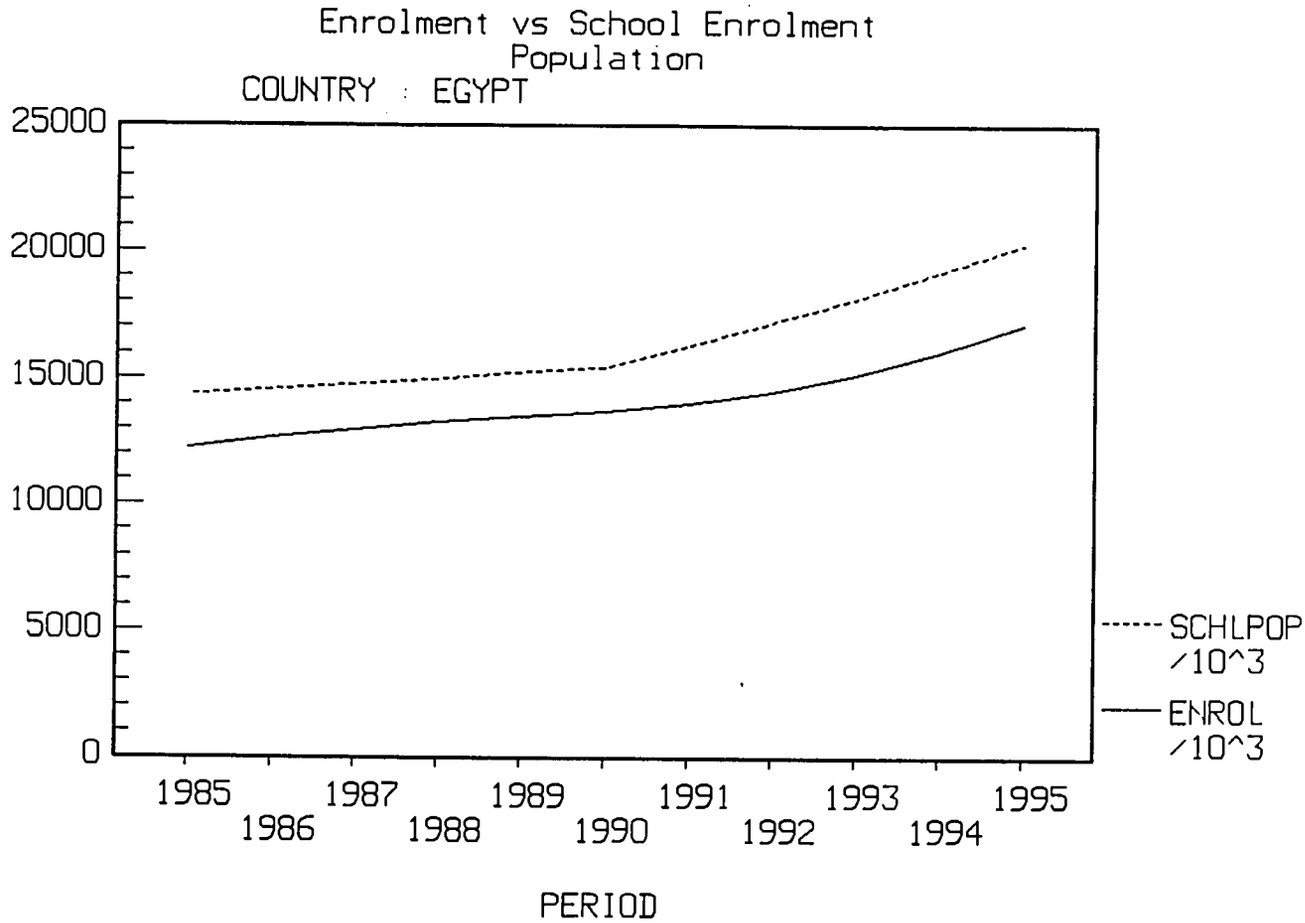

Primary School Disparity Index
COUNTRY : CENTRAL AFRICAN REPUBLIC
DINDPROJ
1.0
0.5
-0.0
-0.5
-1.0
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
PERIOD
DINDPROJ

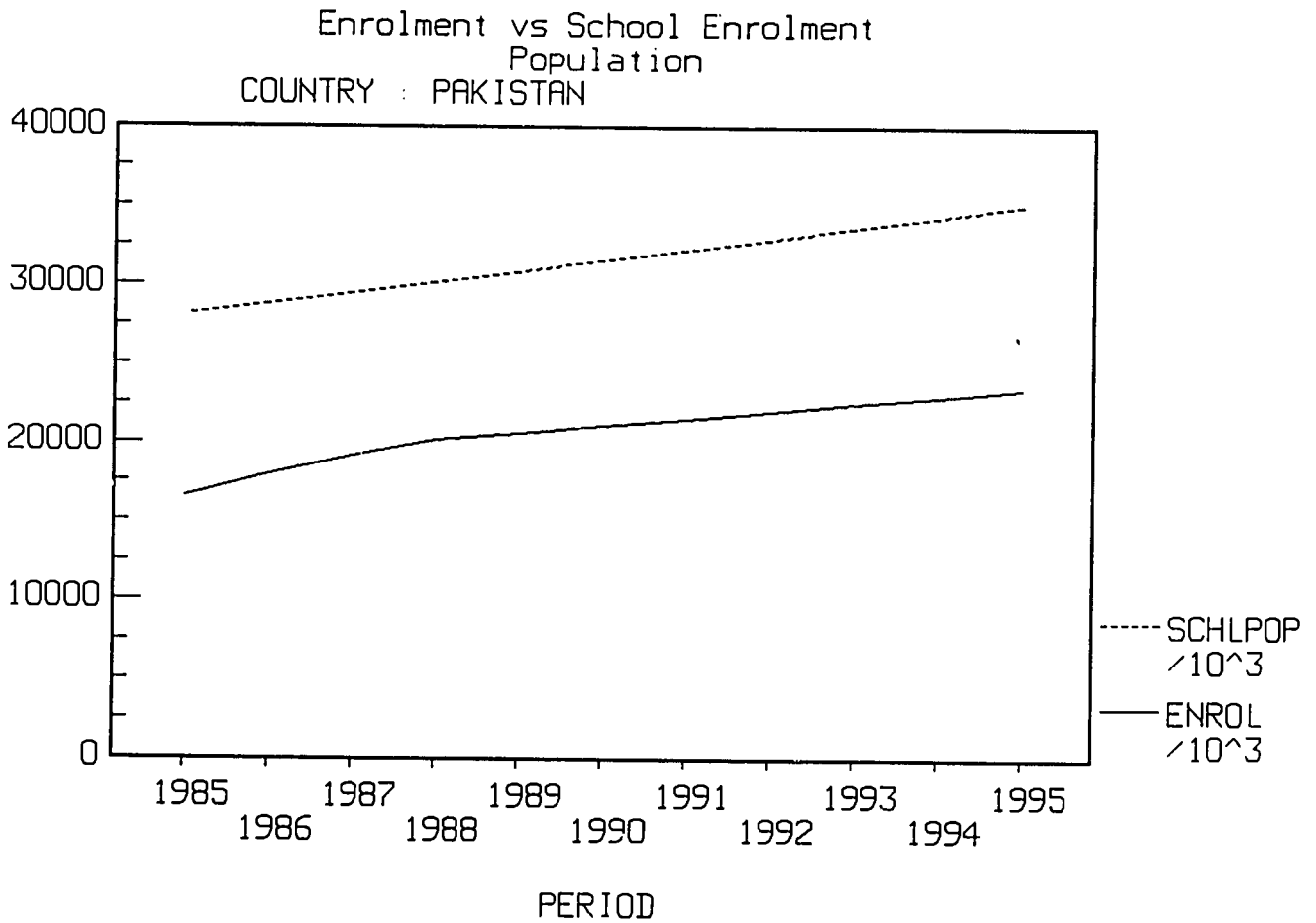

Enrolment vs School Enrolment
Population
COUNTRY : BANGLADESH
40000
30000
20000
10000
0
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
PERIOD
SCHLPOP
/10^3
ENROL
/10^3

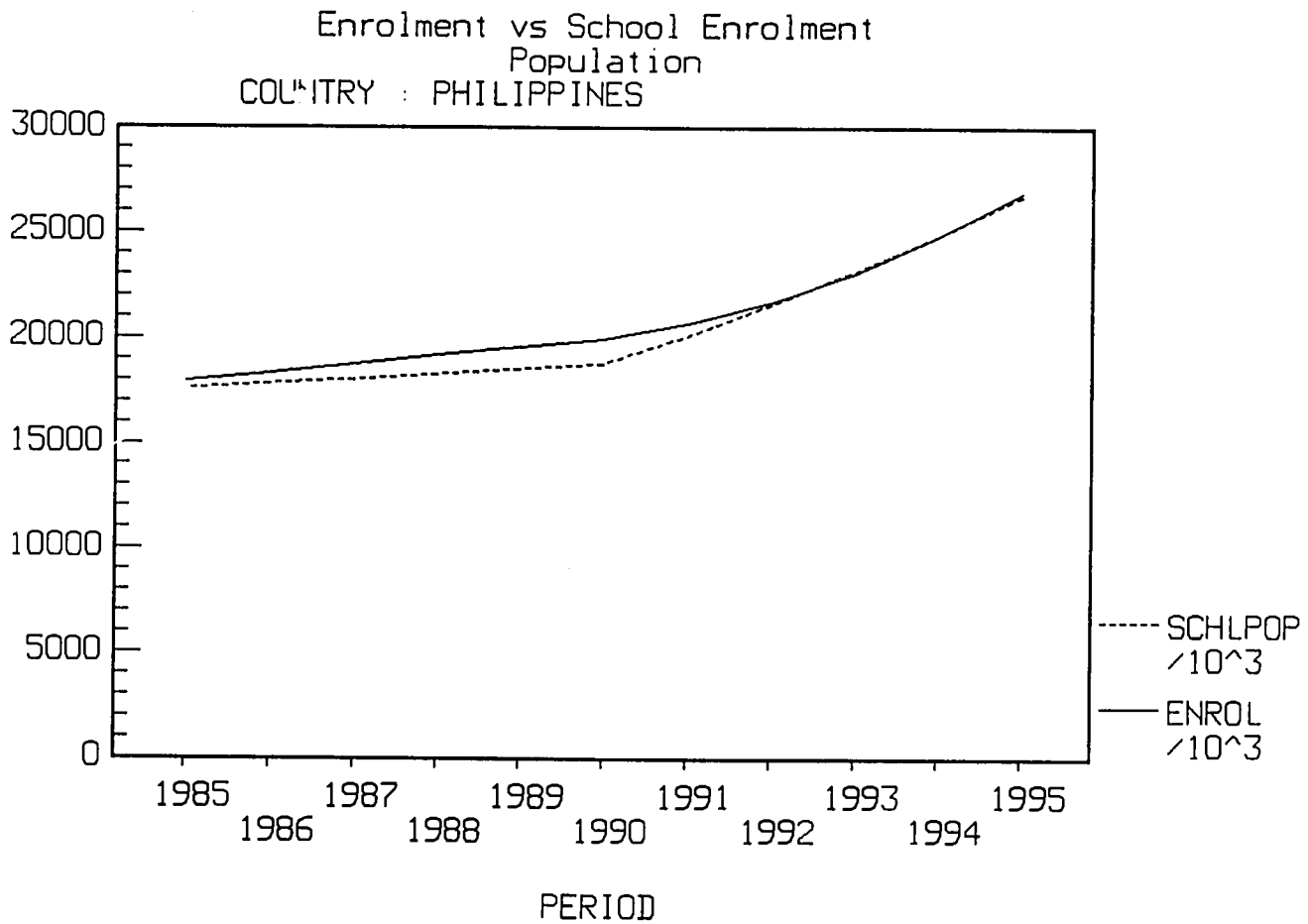

Enrolment vs School Enrolment
Population
COUNTRY : BOLIVIA
4000
3000
2000
1000
0
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
PERIOD
SCHLPOP
/10^3
ENROL
/10^3


Enrolment vs School Enrolment
Population
COUNTRY : BOTSWANA
750
500
250
0
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
PERIOD
SCHLPOP
/10^3
ENROL
/10^3


Enrolment vs School Enrolment
Population
COUNTRY : CENTRAL AFRICAN REPUBLIC
1200
1000
800
600
400
200
0
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
PERIOD
SCHLPOP
/10^3
ENROL
/10^3


Enrolment vs School Enrolment
Population
COUNTRY : EGYPT
25000
20000
15000
10000
5000
0
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
PERIOD
SCHLPOP
/10^3
ENROL
/10^3


Enrolment vs School Enrolment
Population
COUNTRY : PAKISTAN
40000
30000
20000
10000
0
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
PERIOD
SCHLPOP /10^3
ENROL /10^3


Enrolment vs School Enrolment
Population
COUNTRY : PHILIPPINES
30000
25000
20000
15000
10000
5000
0
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
PERIOD
SCHLPOP
/10^3
ENROL
/10^3

Chapter 4. Technical Description of the Projection Methods Used in POPPROJ AND EDUPROJ

1. Population Projection (POPPROJ)

POPPROJ uses the "standard" cohort-component method to project the base year population. The method makes five-year age group population projections for five-year time periods, by country and gender. The major computational steps involve:

1. Derive survival ratios from mortality assumptions. The mortality assumptions are specified in terms of projected life expectancy at birth, which is specified by the user either by entering values for each period or by using the targeting facility. On the basis of the life expectancy at every period, a model life table column is chosen. For the sake of simplicity, the United Nations' standard life table is used. From the life table column of person-years lived by a cohort (Lx), i.e., the number of survivors at different ages in a hypothetical cohort, survival ratios are calculated.

2. Next, the survivors aged 5+ at the end of the interval are determined from the population at the beginning of the period and the survival ratios calculated above. For the first time period, the population by five-year age groups are entered by the user.

3. Next, the age specific fertility rates are determined from fertility assumptions. The fertility assumptions are in the form of projected total fertility rates, and are either entered by the user for each period or by using the interpolation facility. We next compute the age specific fertility rates from the assumed total fertility rates, using a polynomial function fitted to the data tabulated in World Population Prospects As Assessed in 1973, Table 2. (UN Publication, Sales No. E.76.XIII.4.)

4. Then the births by gender are calculated based on the basis of the fertility rates, the number of women by age at the beginning and end of the interval, (from step 2) and the sex ratio at birth.

5. Finally, the survivors aged 0-4 from projected births and first-year age group survival ratios are calculated.

6. To get the single-year school-age population (needed to calculate the entrants into the school system) Beers coefficients are used to disaggregate across ages. Annual data are derived from the five yearly data.

2. Education Projection (EDUPROJ)

The EDUPROJ module combines elements from the System for Tracking Educational Progress (STEP), described in the STEP User's Manual. In constructing EDUPROJ, elements from the STEP modules PROJRATE, ESTIMATE, EDPROJ and COHORT were used.

Enrollment Projection

The enrollment projections are made using the standard grade-transition method. The method at the primary school level, is straightforward, and consists of two simple steps.

First, the entry grade enrollment is projected as the sum of the product of last year's first grade enrollment in first grade times the repetition rate (to get the number of repeaters), plus the product of an entrance rate and the population of school-entrance age (to get the number of new students in first grade). The population of school-entrance age comes from the population projection made within the GENDER system.

Second, for all other grades, enrollment is projected as the sum of last year's enrollment in the same grade times the repetition rate from that grade, plus the product of last year's enrollment in the previous grade times the promotion rate from that grade.

All of these variables are disaggregated by gender and by country in the GENDER model.

This methodology can be used with either fixed or varying repetition, entrance, and dropout rates. In either case, it is necessary to project the repetition and dropout rates themselves before projecting enrollment.

Key indicators can then be estimated. In the GENDER system, the most important are the gross enrollment ratio and disparity indices. The gross enrollment ratio is simply the ratio of population in school to population of school age, disaggregated by gender, of course. The disparity indices are calculated in two alternative ways. First, an index called DINDPROJ is calculated as the difference between male and female gross enrollment ratios, divided by the male gross enrollment ratio. Another, DINDPRJ2, is the difference between male and female gross enrollment ratios, divided by the overall enrollment ratio.

Of the many available methods for projecting enrollment, 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 primary school level, is straightforward, and consists of two basic equations:

The first equation applies to grade one only:

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

a. E_{1+t} is enrollment in grade 1 at time t.
b. $r_{1,t}$ is the repetition rate in grade 1 at time t.
c. n_{t+1} is the proportion of the population of age entage ("entrance age") that enrolls in grade 1 for the first time at time t+1.
d. $POP_{entage,t+1}$ is the population of age entage at time t+1.

Note: All these variables can be disaggregated by gender and country within the GENDER system, but both were left out of technical description to simplify the notation.

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 years enrollment population 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.

The second equation applies to the second through the last grade:

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

Note: The symbols are analogous to those above, and dg,t is the dropout rate from grade g at time t. enrollment in year t+1 in any grade g comes from two sources:

1. A proportion $r_{g,t}$ of last year's enrollment $E_{g,t}$, which repeats the same course.

2. A proportion $(1-r_{g-1,t}-d_{g-1,t})$ of last year's enrollment in the previous grade, $E_{g-t,t}$, that is prompted forward. (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 enrollment.

Estimation of Transition Rates: The Standard Method

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

The repetition rates are estimated as follows:

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

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 enrollment 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 one is:

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

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 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 die. This makes it unnecessary to refer to a death rate in making the enrollment projections.

The entrance (or first grade enrollment) rate, n_t, is calculated using the estimated repeater rate and the population and enrollment data:

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

Efficiency Indicators Based on Cohort Analysis

Basic Concepts

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 enrollment 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 enrollment 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 year 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 dropout 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 dropout 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 dropouts is divided by the number of graduates. In a six year system with repetition and dropout 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. The 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 student/years per dropout, weighed respectively by the number of graduates and the number of dropouts. Note that the number of dropouts and the number of graduates must sum to 1000.

Indicators Based on Enrollment Data

Gross Enrollment Ratio. This is the ratio of enrollment 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 enrollment to actual or projected population, it has to be derived either from historical data or from an enrollment projection. It is not based on the analysis of an idealized cohort.

Net Enrollment 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 Enrollment 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 Enrollment Ratio and the Gross Enrollment 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.

Projection of Required Teachers and Classrooms

GENDER provides two elementary measures of resource requirements for projected enrollment at the primary level. The projected number of primary schools, SCHLPROJ, is projected as a function of the base period number of schools per student times the current number of students enrolled in primary:

$$SCHLPROJ_t = \left(\frac{SCHOOLS_o}{\sum_s \sum_g E_{g,o}} \right) \sum_s \sum_g E_{s,g,t}$$

Similarly, the number of required teachers is projected as a function of the base period teacher-pupil ratio times the current number of students in primary. A somewhat different aspect from cohort is normally projected in models of this type is that <u>required</u> teachers are projected by <u>gender</u>.

$$TEACHPROJ_{s,t} = \left(\frac{TEACHER1_{s,o}}{\sum_s \sum_g E_{s,g,o}} \right) \sum_s \sum_g E_{s,g,t}$$

Estimation of Disparity Indices

The GENDER system, while it keeps track of basic education indicators, focuses on determining the evolution of gender differentials in the education systems of developing countries.

Enrollment disparity is calculated two ways:

a) DIS_IND = 1 - (GER [Female] / GER [Male]).
b) DIS_IND2 = (GER [Male] - GER [Female]) / (GER [TOTAL]).

The variables DINDPROJ and DINDPROJ2 are the same as variables DIS_IND and DIS_IND2 respectively, but projected into the future.

Life Expectancy Disparity:

LExp_Dis = (LEXP [Male] - LEXP [Female]) / (LEXP [TOTAL]).

Appendix 1

Notes on Installation of Host and GENDER

Installing Host

The GENDER system runs under Host, a microcomputer modeling system developed by RTI for various applications. To install the Host system, you must use the Host INSTALL routine which is found on the first diskette of the Host system. Host also requires a commercial graphing program called Graph-in-the-Box, if you want to have graphs. We suggest you place Host in its own subdirectory on your hard disk and establish a path to it. Remember to edit your config.sys file as indicated in the Host INSTALL program. Host takes about 2 MB of space.

Installing GENDER

To copy GENDER to your hard disk, use the DOS RESTORE command:

C:\> md gender.

c\>: RESTORE A: C: /S or C\>: RESTORE A: C:\GENDER\ (if the GENDER subdirectory already exists).

This will place GENDER in its own subdirectory. GENDER takes 3-4 MB of space. To go into GENDER, invoke Host from the GENDER subdirectory.

GENDER Files

In addition to the Host system files, GENDER is distributed with the following files:

QUERY.COM	Modules
EDUPROJ.COM	Modules
POPPROJ.COM	Modules
WID.DAT	Master Database
WID.IDX	Master Database
WID.DDX	Master Database
WID.VAR	Master Database
WID.DIM	Master Database
TABLE.LIB	Pre-defined Tables (Table Library)
TABLE.IDX	Pre-defined Tables (Table Library)
GRAPH.LIB	Pre-defined Graphs (Graph Library)
GRAPH.IDX	Pre-defined Graphs (Graph Library)
LABEL.LIB	Label Library
LABEL.IDX	Label Library

Appendix 2

WID Database: Variable List and Characteristics

Name	Size (bytes)	Type	Periodic	Number of Dimensions	Dimension Names	Maximum Levels
ASFR	5600	real	Yes	2	COUNTRY	100
					FERTAGE	7
BASETOGL	2	integer	No	0	NONE	NONE
BASEYEAR	200	integer	No	1	COUNTRY	100
CP_RT_UN	2400	real	Yes	2	COUNTRY	100
					SEX	3
CP_RT_WB	2400	real	Yes	2	COUNTRY	100
					SEX	3
CUR2AREA	2400	real	Yes	2	COUNTRY	100
					AREA2	3
CURLEV	2400	real	Yes	2	COUNTRY	100
					LEVEL	3
CURRENT	800	real	Yes	1	COUNTRY	100
DEP_RAT	800	real	Yes	1	COUNTRY	100
DINDPRJ2	800	real	Yes	1	COUNTRY	100
DINDPROJ	800	real	Yes	1	COUNTRY	100
DIS_IND	2400	real	Yes	2	COUNTRY	100
					LEVEL	3
DIS_IND2	2400	real	Yes	2	COUNTRY	100
					LEVEL	3
DROPRT	24000	real	Yes	3	COUNTRY	100
					SEX	3
					GRADE1	10
ED_ATT	14400	real	Yes	3	COUNTRY	100
					SEX	3
					ATT_LEV	6
ENR2AREA	7200	real	Yes	3	COUNTRY	100
					SEX	3
					AREA2	3
ENR2FLD	28800	real	Yes	3	COUNTRY	100
					SEX	3
					FIELD2	12
ENR3FLD	52800	real	Yes	3	COUNTRY	100
					SEX	3
					FIELD3	22
ENROL	24000	real	Yes	3	COUNTRY	100
					SEX	3
					GRADE1	10
ENROL1	24000	real	Yes	3	COUNTRY	100
					SEX	3
					GRADE1	10
ENROL2	24000	real	Yes	3	COUNTRY	100
					SEX	3
					GRADE2	10
ENTAGE	200	integer	No	1	COUNTRY	100

Name	Size (bytes)	Type	Periodic	Number of Dimensions	Dimension Names	Maximum Levels
ENTPOP	2400	real	Yes	2	COUNTRY	100
					SEX	3
ENTRT	2400	real	Yes	2	COUNTRY	100
					SEX	3
FLFPR	800	real	Yes	1	COUNTRY	100
GDP_GR	800	real	No	1	COUNTRY	100
GER	2400	real	Yes	2	COUNTRY	100
					SEX	3
GNP_PC	800	real	Yes	1	COUNTRY	100
GRADRATE	2400	real	Yes	2	COUNTRY	100
					SEX	3
GRENRRT	7200	real	Yes	3	COUNTRY	100
					SEX	3
					LEVEL	3
GR_RAT	800	real	Yes	1	COUNTRY	100
ILLIT	2400	real	Yes	2	COUNTRY	100
					SEX	3
IMRMF	2400	real	Yes	2	COUNTRY	100
					SEX	3
IORATE	2400	real	Yes	2	COUNTRY	100
					SEX	3
LEXP	2400	real	Yes	2	COUNTRY	100
					SEX	3
LEXP_DIS	800	real	Yes	1	COUNTRY	100
NETENRRT	7200	real	Yes	3	COUNTRY	100
					SEX	3
					LEVEL	3
NUMGRAD	200	integer	Yes	1	COUNTRY	100
ONTIME	2400	real	Yes	2	COUNTRY	100
					SEX	3
POP5	43200	real	Yes	3	COUNTRY	100
					SEX	3
					AGE5	18
POPSCH	800	real	Yes	1	COUNTRY	100
REGION	25600	string	No	1	COUNTRY	100
REPEAT1	24000	real	Yes	3	COUNTRY	100
					SEX	3
					GRADE1	10
REPEAT2	24000	real	Yes	3	COUNTRY	100
					SEX	3
					GRADE2	10
REPEATRT	24000	real	Yes	3	COUNTRY	100
					SEX	3
					GRADE1	10
RUR_POP	800	real	Yes	1	COUNTRY	100
SCHLPOP	2400	real	Yes	2	COUNTRY	100
					SEX	3

Name	Size (bytes)	Type	Periodic	Number of Dimensions	Dimension Names	Maximum Levels
SCHLPROJ	800	real	Yes	1	COUNTRY	100
SCHOOLS1	800	real	Yes	1	COUNTRY	100
SEXRATIO	8	real	No	0	NONE	NONE
SYGRAD	2400	real	Yes	2	COUNTRY	100
					SEX	3
SYGRAD2	2400	real	Yes	2	COUNTRY	100
					SEX	3
TARDROP	24000	real	No	3	COUNTRY	100
					SEX	3
					GRADE1	10
TARENTRT	2400	real	No	2	COUNTRY	100
					SEX	3
TARGLEXP	2400	real	No	2	COUNTRY	100
					SEX	3
TARGTFR	800	real	No	1	COUNTRY	100
TARGTOGL	1	byte	No	0	NONE	NONE
TARGYEAR	2	integer	No	0	NONE	NONE
TARREP	24000	real	No	3	COUNTRY	100
					SEX	3
					GRADE1	10
TEA2AREA	7200	real	Yes	3	COUNTRY	100
					SEX	3
					AREA2	3
TEACHER1	2400	real	Yes	2	COUNTRY	100
					SEX	3
TEACHER2	2400	real	Yes	2	COUNTRY	100
					SEX	3
TEACHPRJ	2400	real	Yes	2	COUNTRY	100
					SEX	3
TOT_FER	800	real	Yes	1	COUNTRY	100
YRSSCHL	2400	real	Yes	2	COUNTRY	100
					SEX	3
ZONE	600	integer	No	2	COUNTRY	100
					ZONE	3