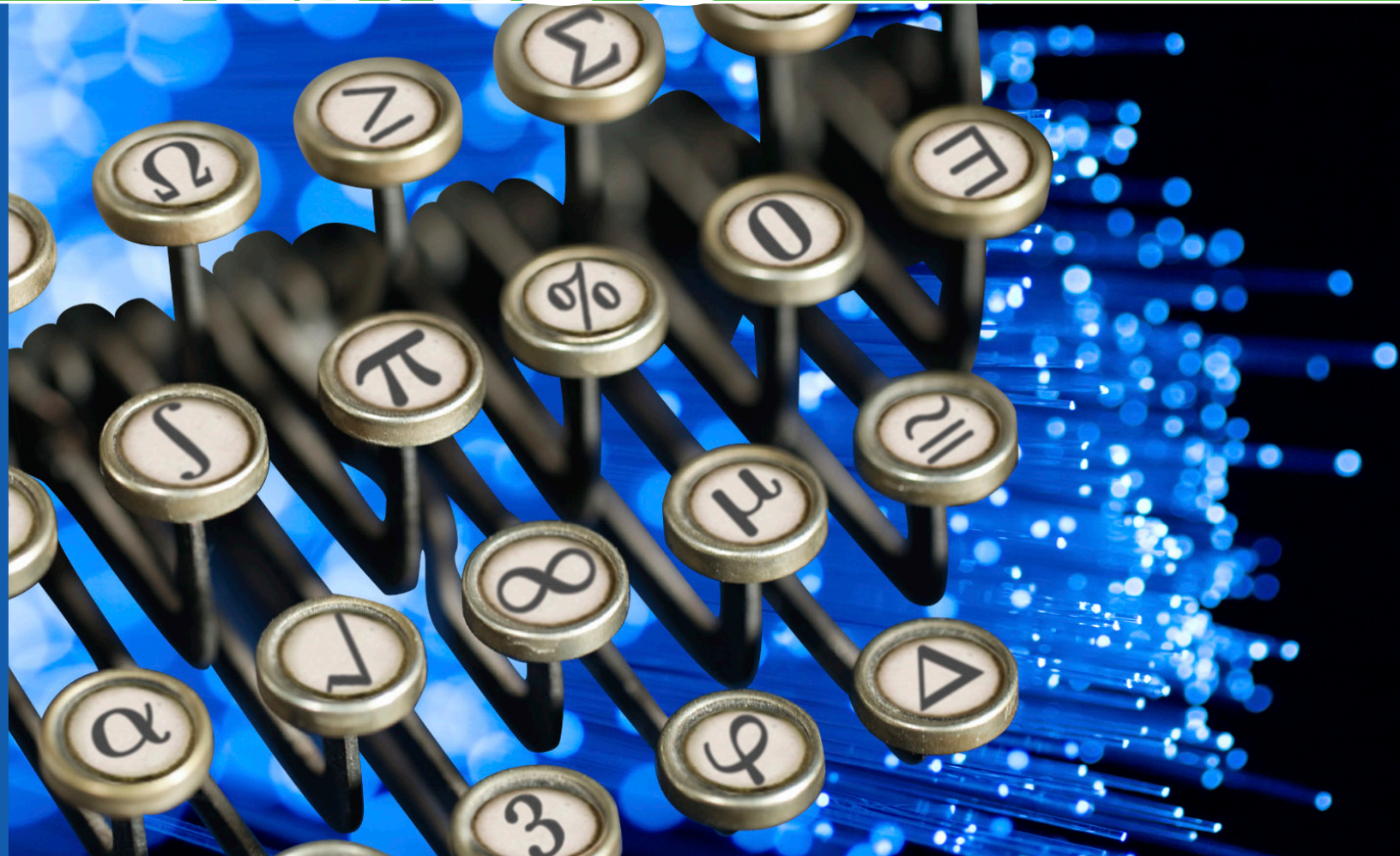


TRENDS IN INTERNATIONAL MATHEMATICS AND SCIENCE STUDY

TIMSS



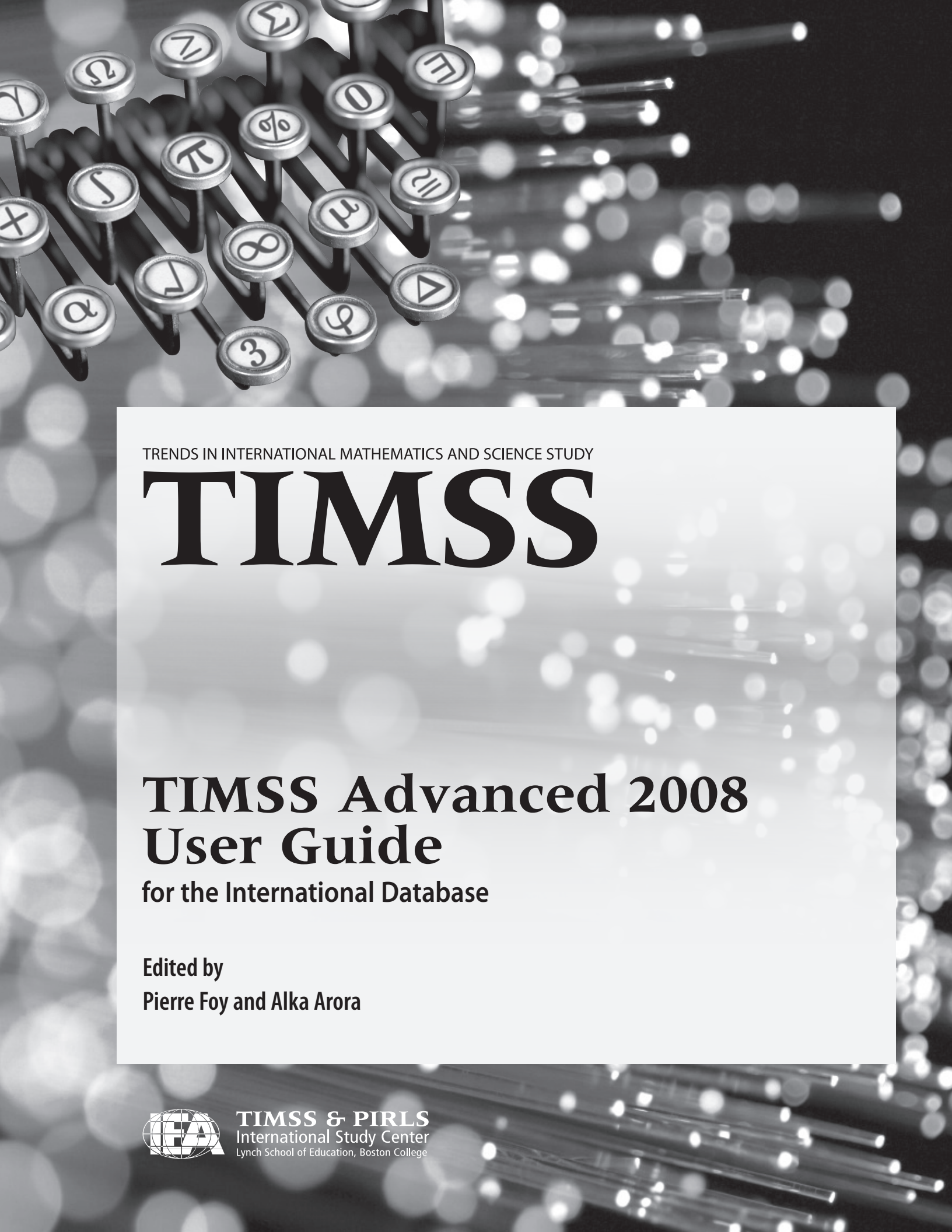
TIMSS Advanced 2008 User Guide

for the International Database

Edited by
Pierre Foy and
Alka Arora



TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College



TRENDS IN INTERNATIONAL MATHEMATICS AND SCIENCE STUDY

TIMSS

TIMSS Advanced 2008 User Guide

for the International Database

Edited by

Pierre Foy and Alka Arora



TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College

Copyright © 2009 International Association for the Evaluation of Educational Achievement (IEA)

TIMSS Advanced 2008 User Guide for the International Database

Edited by Pierre Foy and Alka Arora

Publisher: TIMSS & PIRLS International Study Center,
Lynch School of Education, Boston College

Library of Congress Catalog Card Number: 2009902654

ISBN: 1-889938-57-2

For more information about TIMSS contact:

TIMSS & PIRLS International Study Center

Lynch School of Education

Boston College

Chestnut Hill, MA 02467

United States

tel: +1-617-552-1600

fax: +1-617-552-1203

e-mail: timss@bc.edu

<http://timssandpirls.bc.edu>

Boston College is an equal opportunity, affirmative action employer.

Contents

Chapter 1

Introduction..... 1

Chapter 2

Using the IEA IDB Analyzer to Analyze the
TIMSS Advanced 2008 International Database..... 7

Chapter 3

Analyzing the TIMSS Advanced 2008 International Database
Using SAS.....39

Chapter 4

The TIMSS Advanced 2008 International Database Files.....79

Appendix

Organizations and Individuals Responsible for
TIMSS Advanced 2008.....113



Chapter 1

Introduction

1.1 Overview of the TIMSS Advanced 2008 User Guide and International Database

TIMSS¹ Advanced measures trends in advanced mathematics and physics achievement at the final year of secondary school in participating countries around the world, as well as monitoring curricular implementation and identifying promising instructional practices. TIMSS Advanced has assessed advanced mathematics and physics in 1995 and 2008, with planning underway for 2016. TIMSS Advanced collects a rich array of background information to provide comparative perspectives on trends in achievement in the context of different educational systems, school organizational approaches, and instructional practices.

To support and promote secondary analyses aimed at improving advanced mathematics and physics education at the end of secondary school, the TIMSS Advanced 2008 international database makes available to researchers, analysts, and other users the data collected and processed by the TIMSS Advanced project. This database comprises student achievement data as well as student, teacher, school, and curricular background data for 10 countries. The database includes data from 38,731 students, 2,629 teachers, 2,128 school principals, and the National Research Coordinators of each country involved in the 2008 assessments. The TIMSS Advanced 2008 international database also includes an updated release of the TIMSS Advanced 1995 database since the 1995 assessments were rescaled to conform to the scaling methodology now in use. All participating countries gave the IEA permission to release their national data.

For countries that participated in the 1995 assessments, TIMSS Advanced 2008 provides trend measures in advanced mathematics and physics achievement

¹ The Trends in International Mathematics and Science Study is a project of the International Association for the Evaluation of Educational Achievement (IEA) and together with PIRLS, the Progress in International Reading Literacy Study, constitutes IEA's regular cycle of core studies.

from 1995 to 2008. In countries new to the study, the 2008 results can help policy makers and practitioners assess their comparative standing and gauge the rigor and effectiveness of their advanced mathematics and physics programs. Details of the assessments conducted in 2008 can be found in the *TIMSS Advanced 2008 International Report* (Mullis, Martin, Robitaille, & Foy, 2009).

TIMSS Advanced 2008 was an ambitious and demanding study, involving complex procedures for assessing students' achievement, drawing student samples, and analyzing and reporting the data. In order to work effectively with the TIMSS Advanced data, it is necessary to have an understanding of the characteristics of the study, which are described fully in the *TIMSS Advanced 2008 Technical Report* (Arora, Foy, Martin, & Mullis, 2009). It is intended, therefore, that this User Guide be used in conjunction with the Technical Report. Whereas the User Guide describes the organization and content of the database, the Technical Report provides the rationale for the techniques used and for the variables created.

1.2 Analyzing the TIMSS Advanced 2008 Data

The TIMSS Advanced 2008 International Database offers researchers and analysts a rich environment for examining student achievement in advanced mathematics and physics at the end of secondary school in an international context. This includes:

- Extensive data on advanced mathematics and physics achievement in the final year of secondary school providing in-depth study of the quality of education in terms of learning outcomes
- Comparable data for 10 countries from around the world providing an international perspective from which to examine educational practices and student outcomes in advanced mathematics and physics
- Student achievement in advanced mathematics and physics linked to questionnaire information from students, teachers, school principals, and curriculum experts, providing policy-relevant contextual information on the antecedents of achievement
- Achievement scales on a common metric that link the 1995 and 2008 assessment cycles, providing for analysis of trends in advanced mathematics and physics achievement

The TIMSS Advanced database is complex, which can make analyzing the data challenging for users. In particular, two of the more complicated issues that need to be addressed are the complex multi-stage sample design of TIMSS Advanced and its use of imputed scores (also known as plausible values) to represent student achievement in advanced mathematics and physics.

The TIMSS Advanced 2008 target populations are all advanced mathematics students and all physics students in the final year of secondary school in each participating country. To obtain accurate and representative samples, TIMSS Advanced 2008 used a two-stage sample design whereby a random sample of schools was selected at the first stage and advanced mathematics and physics classes were sampled at the second stage. This is a very effective and efficient sampling approach, but the resulting student samples have a complex structure that must be taken into consideration when analyzing the data. In particular, sampling weights need to be applied and a re-sampling technique such as the jackknife employed to estimate sampling variances correctly.²

In addition, TIMSS Advanced 2008 used Item Response Theory (IRT) scaling to summarize student achievement on the assessment and to provide accurate measures of trends from the previous assessments. The TIMSS Advanced IRT scaling approach used multiple imputation—or “plausible values”—methodology to obtain proficiency scores of students in advanced mathematics and physics. Each student record in the TIMSS Advanced 2008 international database contains imputed scores in either advanced mathematics or physics. Because each imputed score is a prediction based on limited information, it almost certainly includes some small amount of error. To allow analysts to incorporate this error into analyses of the TIMSS Advanced achievement data, the TIMSS Advanced 2008 database provides five separate imputed scores for each scale. Each analysis should be replicated five times, using a different plausible value each time, and the results combined into a single result that includes a standard error incorporating both sampling and imputation error.³

IEA has developed the International Database (IDB) Analyzer software (IEA, 2009) specifically for analyzing data from complex studies like TIMSS Advanced 2008. Used in conjunction with SPSS (SPSS, 2008), this software helps users analyze the TIMSS Advanced achievement data by conducting each

² More details on the sample design and its implementation are provided in Chapter 4 (LaRoche, Zuehlke, & Joncas, 2009) of the *TIMSS Advanced 2008 Technical Report*.

³ More details on plausible values can be found in Chapter 8 (Foy, Galia, & Li, 2009) of the *TIMSS Advanced 2008 Technical Report*.

analysis separately on each plausible value, aggregating the resulting statistics, and applying the jackknife algorithm to provide appropriate standard errors for each statistic. It also simplifies the management of the TIMSS Advanced database by providing a module for selecting subsets of countries and variables, and merging files for analysis.

1.3 Contents of the TIMSS Advanced 2008 User Guide

This User Guide describes the format and contents of the data files and materials in the TIMSS Advanced 2008 international database and presents example analyses. In addition to this introduction, the User Guide includes the following three chapters:

- Chapter 2 introduces the IEA IDB Analyzer and presents examples of analyses of the TIMSS Advanced 2008 data using this software in conjunction with SPSS.
- Chapter 3 explains how to implement the analyses described in Chapter 2 using the SAS (SAS, 2002) statistical software system and the SAS programs and macros included with the database.
- Chapter 4 describes the structure and content of the database.

The User Guide is accompanied by four supplements.

- Supplement 1 comprises the international version of all TIMSS Advanced 2008 background questionnaires.
- Supplement 2 describes all adaptations to the questions in the background questionnaires made by the TIMSS Advanced 2008 participants.
- Supplement 3 describes how indices and other derived variables were constructed for reporting the TIMSS Advanced 2008 data.
- Supplement 4 describes the sampling stratification variables for all participating countries.

1.4 Contents of the TIMSS Advanced 2008 International Database DVD

This User Guide is included as a PDF document on the DVD that contains the TIMSS Advanced 2008 international database, along with advanced

mathematics and physics student achievement data files and student, teacher, and school background questionnaire data files. The DVD also includes support materials and an updated release of the 1995 database since the 1995 assessments were rescaled using the IRT scaling methodology now in use. Exhibit 1.1 displays the DVD's folder structure and provides a description of its contents.

Exhibit 1.1 Contents of the TIMSS Advanced 2008 International Database DVD

1_UserGuide_and_Database	This User Guide with its four supplements		
	TA08_Database	SPSS_Data	TIMSS Advanced 2008 student, teacher, and school data files in SPSS format
		SAS_Data	TIMSS Advanced 2008 student, teacher, and school data files in SAS format
		Codebooks	Codebook files describing all variables in the TIMSS Advanced 2008 international database
		Programs	SAS and SPSS programs and macros
		Almanacs	Data almanacs with summary statistics for all TIMSS Advanced 2008 items and background variables
		Items	TIMSS Advanced 2008 item information files, IRT item parameters, and released items
		TCMA	National item selection data for the Text-Curriculum Matching Analysis (see Appendix C of International Report for more details)
		Curriculum	TIMSS Advanced 2008 curriculum questionnaires data files
	TA95_Database	SPSS_Data	TIMSS Advanced 1995 student and school data files in SPSS format ⁴
		SAS_Data	TIMSS Advanced 1995 student and school data files in SAS format
		Codebooks	Codebook files describing all variables in the TIMSS Advanced 1995 international database
		Programs	SAS and SPSS programs and macros
		Almanacs	Data almanacs with summary statistics for all TIMSS Advanced 1995 items and background variables
		Items	TIMSS Advanced 1995 rescaled IRT item parameters
2_IEA_IDB_Analyzer	Executable file for installing the IEA IDB Analyzer		
3_Publications	PDF versions of all TIMSS Advanced 2008 publications, including the <i>TIMSS Advanced 2008 Assessment Frameworks</i> , <i>TIMSS Advanced 2008 International Report</i> , and <i>TIMSS Advanced 2008 Technical Report</i>		

⁴ The TIMSS Advanced 1995 database includes only student and school data since it did not administer a teacher background questionnaire in 1995.

References

Arora, A., Foy, P., Martin, M.O., & Mullis, I.V.S. (Eds.) (2009). *TIMSS Advanced 2008 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Foy, P., Galia, J., & Li, I. (2009). Scaling the data from the TIMSS advanced 2008 assessments. Chapter 8 in A. Arora, P. Foy M.O. Martin, and I.V.S. Mullis (Eds.), *TIMSS Advanced 2008 technical report* (pp. 143-206). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

IEA (2009). *IEA international database analyzer* (version 2.0) [Computer software and manual]. Hamburg: IEA Data Processing and Research Center.

LaRoche, S., Zuehlke, O., & Joncas, M. (2009). TIMSS Advanced 2008 sampling. Chapter 4 in I.V.S. Mullis, M.O. Martin, P. Foy, & A. Arora (Eds.), *TIMSS Advanced 2008 technical report* (pp. 51-92). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., Robitaille, D.F., & Foy, P. (2009). *TIMSS advanced 2008 international report: Findings from IEA's study of achievement in advanced mathematics and physics in the final year of secondary school*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

SAS Institute (2002). *SAS system for Windows* (version 9.1). Cary, NC: SAS Institute.

SPSS Inc. (2008). *SPSS for Windows* (version 16.0). Chicago, IL: SPSS Inc.



Chapter 2

Using the IEA IDB Analyzer to Analyze the TIMSS Advanced 2008 International Database

2.1 Overview

This chapter describes the use of the IEA International Database (IDB) Analyzer software (IEA, 2009) for analyzing the TIMSS Advanced 2008 data. Used in conjunction with SPSS (SPSS, 2008), the IEA IDB Analyzer provides a user-friendly interface to easily merge the various data file types of the TIMSS Advanced 2008 database and seamlessly takes into account the sampling information and the multiple imputed achievement scores to produce accurate statistical results.

Example analyses will illustrate the capabilities of the IEA IDB Analyzer to compute a variety of statistics, including percentages of students in specified subgroups, mean student achievement in those subgroups, correlations, regression coefficients, and percentages of students reaching benchmark levels. The examples use student, teacher, and school background data to replicate some of the TIMSS Advanced 2008 results included in the *TIMSS Advanced 2008 International Report* (Mullis, Martin, Robitaille, & Foy, 2009), as well as other useful analyses for investigating policy-relevant research questions.

With a basic knowledge of the TIMSS Advanced 2008 international database, users should be able to perform statistical analyses with the IEA IDB Analyzer. The IEA IDB Analyzer also will accept the TIMSS Advanced 1995 database¹ released along with the 2008 database. A more detailed description of the data files contained in the international database—their structure and contents—is given in Chapter 4, along with descriptions of all the supporting documentation provided on the DVD.

¹ This chapter presents analysis examples using data from the 2008 database, which can be adapted and applied to the data from the 1995 database. SPSS programs for the 1995 database also are available on the DVD.

2.2 The IEA IDB Analyzer

Developed by the IEA Data Processing and Research Center (DPC), the IEA IDB Analyzer is a plug-in for SPSS, a well-known statistical analysis system. The IEA IDB Analyzer enables users to combine SPSS data files from IEA's large-scale assessments and conduct analyses using SPSS without actually writing programming code. The IEA IDB Analyzer generates SPSS syntax that takes into account information from the sample design in the computation of statistics and their standard errors. In addition, the generated SPSS syntax makes appropriate use of plausible values for calculating estimates of achievement scores and their standard errors, combining both sampling variance and imputation variance.

The IEA IDB Analyzer consists of two modules—a merge module and an analysis module, which are executed as independent applications. The merge module is used to create analysis datasets by combining data files of different types and from different countries, and selecting subsets of variables for analysis. The analysis module provides procedures for computing various statistics and their standard errors. Both modules can be accessed by the using the START menu in Windows:

Start ⇒ All Programs ⇒ IEA ⇒ IDB Analyzer ⇒ Merge Module
⇒ Analysis Module

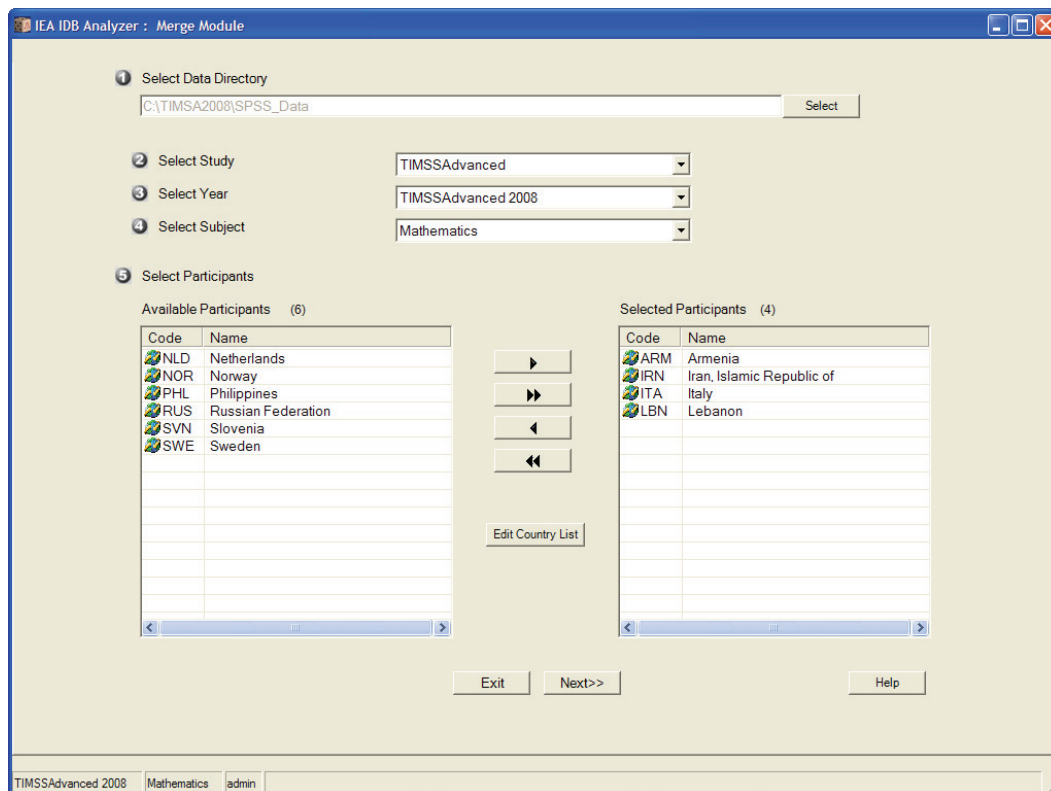
2.3 Merging Files with the IEA IDB Analyzer

The IEA IDB Analyzer uses the SPSS data files located in the “SPSS_Data” folder of the database DVD. The TIMSS Advanced 2008 database consists of a number of file types that are disseminated separately for each country. In addition to allowing users to combine like datasets from more than one country for cross-country analyses, the merge module allows for the combination of data from different sources (e.g., student, teacher, and school) into one SPSS dataset for subsequent analyses. Before doing any statistical analyses with the TIMSS Advanced international database, users should copy the contents of the DVD to an alternate location, either on their computer or on a server. For the purposes of this chapter, we will assume that all SPSS files on the DVD have been copied to the “C:\TIMSA2008\SPSS_Data” folder.

The following steps will create an SPSS data file with data from multiple countries and multiple file types:

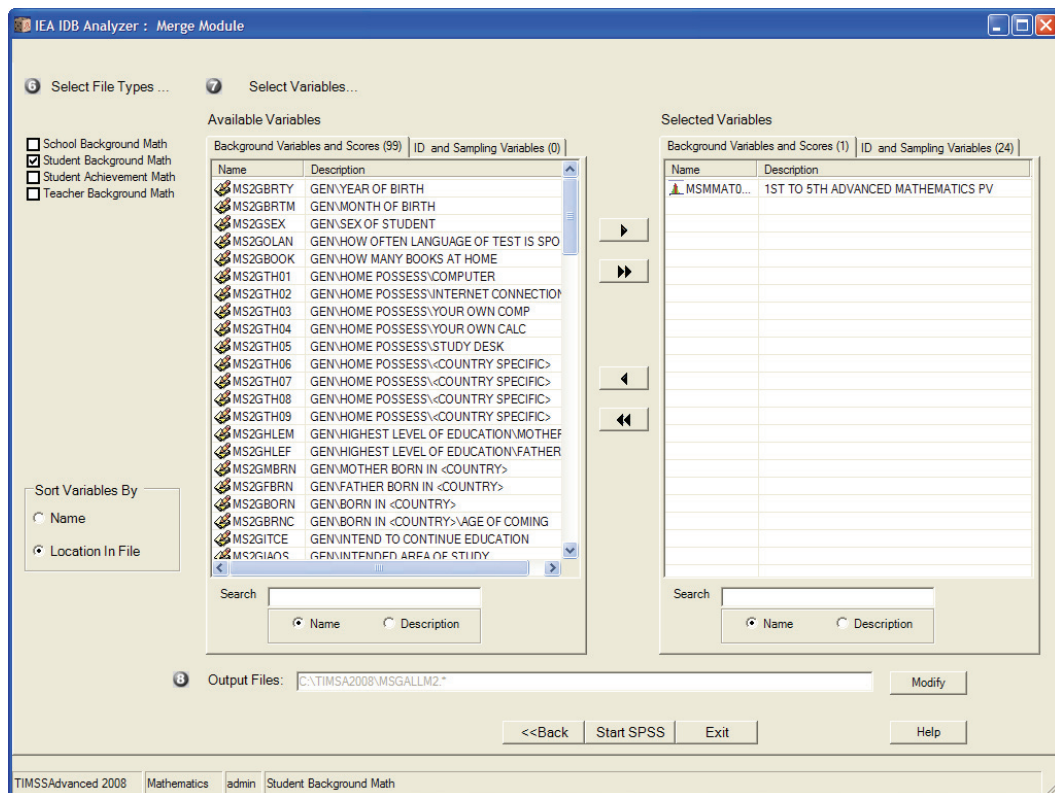
- 1) Open the merge module of the IEA IDB Analyzer.
- 2) In the **Select Data Directory** field, click the **Select** button to browse to the folder where all SPSS data files are located. In Exhibit 2.1, all SPSS data files are located in the “C:\TIMSA2008\SPSS_Data” folder. The program will automatically recognize and complete the **Select Study**, **Select Year**, and **Select Subject** fields and list all countries available in this folder as possible candidates for merging. If the folder contains data from more than one IEA study, or from more than one grade or subject, the IEA IDB Analyzer provides drop-down menus to select the appropriate files for analysis. In Exhibit 2.1, advanced mathematics from TIMSS Advanced 2008 is selected.

Exhibit 2.1 IEA IDB Analyzer Merge Module: Selecting Countries



- 3) Select the countries of interest from the **Available Participants** list. To select multiple countries, hold the CTRL key of the keyboard when selecting the countries. Click the **right arrow** button (▶) to move the selected countries to the **Selected Participants** panel. Click the **double right arrow** button (▶▶) to select all countries. In Exhibit 2.1, Armenia, Iran, Italy, and Lebanon are selected.
- 4) Click the **Next>>** button to proceed to the next step: selecting files and variables. The software will open the second window of the merge module, as shown in Exhibit 2.2, to select the files and the variables to be included in the merged data file.

Exhibit 2.2 IEA IDB Analyzer Merge Module: Selecting File Types and Variables



- 5) Select the files for merging by checking the appropriate boxes to the left of the window. For example, in Exhibit 2.2, the student background data files are selected.

- 6) Select the variables required from the list of background variables available in the left panel. Note that Supplement 1 provides the variable names for all questions in the background questionnaires. Variables are selected by clicking them and then clicking the **right arrow** button. Clicking the **double right arrow** button selects all variables. The search tool at the bottom of the left panel also can be used to locate variables. All achievement scores and all identification and sampling variables are selected automatically by the IEA IDB Analyzer.
- 7) Specify the desired name of the merged data file and the folder where it will be stored in the **Output Files** field by clicking the **Define/Modify** button. The IEA IDB Analyzer also will create an SPSS syntax file (*.SPS) of the same name and in the same folder with the code necessary to perform the merge. In the example shown in Exhibit 2.2, the merged file MSGALLM2.SAV and the syntax file MSGALLM2.SPS will both be created and stored in the “C:\TIMSA2008” folder.²
- 8) Click the **Start SPSS** button to create the SPSS syntax file and open it in an SPSS syntax window ready for execution. The syntax file can be executed by opening the **Run** menu of SPSS and selecting the **All** menu option—or alternatively by pressing CTRL-A and CTRL-R. The IEA IDB Analyzer will display a warning if it is about to overwrite an existing file in the specified folder.

Once SPSS has completed its execution, it is important to check the SPSS output window for possible warnings or error messages. Any warnings or error messages should be examined carefully as they might indicate that the merge process was not performed properly and the resulting merged data file might not be as expected.

Merging Teacher and Student Data Files

The teachers in the TIMSS Advanced 2008 international database do not constitute representative samples of teachers in the participating countries. Rather, they are the teachers of nationally representative samples of students. Therefore, analyses with teacher data should be made with students as the

² As a general rule, merged files created by the merge module should not be saved in the same folder where the source files are located.

units of analysis and reported in terms of students who are taught by teachers with a particular attribute.

Teacher data are analyzed by linking the students to their teachers. The student-teacher linkage data files are used for this purpose and the IEA IDB Analyzer will make use of them automatically. To merge the teacher and student background data files, select both file types in the second window of the IEA IDB Analyzer merge module. The variables of interest need to be selected separately for both file types, as follows:

- 1) Click the **Student Background** file type so that it appears checked and highlighted. The **Background Variables and Scores** listed in the left panel will include all available variables from the student background data files.
- 2) Select the variables of interest from the left panel and click the **right arrow** button to move these variables into the **Selected Variables** panel to the right. Click the **double right arrow** button to select all available variables.
- 3) Next, click the **Teacher Background** file type, selecting the variables of interest from the **Background Variables and Scores** panel to the left in the same manner.
- 4) Specify the folder and merged data file name in the **Output Files** field, as described earlier.
- 5) Click the **Start SPSS** button to create the SPSS syntax file that will produce the required merged data file, which can then be run by opening the **Run** menu of SPSS and selecting the **All** menu option.

Merging School and Student Data Files

Because TIMSS Advanced 2008 has representative samples of schools, it is possible to compute appropriate statistics with schools as units of analysis. However, the school samples were designed to optimize the student samples and the student-level statistics. For this reason, it is preferable to analyze school-level variables as attributes of students, rather than as units of analysis in their own right. Therefore, analyzing school data should be done by linking the students to their schools.

To merge the school and student background data files, select both the **School Background** and **Student Background** file types in the second window of the IEA IDB Analyzer merge module. The variables of interest to be included in

the merged data file are selected separately by file type, as was described earlier and using the same set of instructions.

Merged Data Files for the Examples

To carry out the analysis examples presented in this chapter, the following merged data files should be created with all available background variables and scores selected:

MSGALLM2.SAV Merge the advanced mathematics student background data files for all countries

MTGALLM2.SAV Merge the advanced mathematics teacher and student background data files for all countries

MCGALLM2.SAV Merge the advanced mathematics school and student background data files for all countries

2.4 Performing Analyses with the IEA IDB Analyzer

The analysis module of the IEA IDB Analyzer can perform statistical analyses on any files created using the merge module. The following statistical procedures are available in the analysis module of the IEA IDB Analyzer:

Percentages only

Compute percentages by subgroups

Percentages and Means

Compute percentages, means, and standard deviations for selected variables by subgroups

Regression

Compute regression coefficients for selected independent variables to predict a dependent variable by subgroups

Correlations

Compute means, standard deviations, and correlation coefficients for selected variables by subgroups

Achievement Benchmarks

Compute percentages of students meeting a set of user-specified achievement benchmarks, in particular the TIMSS international achievement benchmarks, by subgroups

All statistical procedures offered within the analysis module of the IEA IDB Analyzer make appropriate use of sampling weights and standard errors are computed using the jackknife repeated replication (JRR) method (see LaRoche, Zuehlke, & Joncas, 2009). Percentages, means, regressions, and correlations may be specified with or without achievement scores. When achievement scores are used, the analyses are performed five times—once for each plausible value—and the results are aggregated to produce accurate estimates of achievement and standard errors that incorporate both sampling and imputation errors. To conduct analyses using achievement scores, check the **With Achievement Scores** option from the **Select Analysis Type** panel of the analysis module.

The IEA IDB Analyzer requires the selection of variables for a number of purposes:

Grouping Variables

This is a list of variables to define subgroups. The list must consist of at least one grouping variable. By default, the IEA IDB Analyzer includes the variable IDCNTY used to distinguish the participating countries. Additional variables may be selected from the available list. If the **Exclude Missing from Analysis** option from the **Select Analysis Type** panel is checked, which is done by default, only cases that have non-missing values in the grouping variables will be used in the analysis. If it is not checked, missing values become reporting categories

Analysis Variables

This is a list of variables for which means are to be computed, or a list of independent variables for a regression analysis. More than one analysis variable can be selected. To compute means for achievement scores, it is necessary to check the **With Achievement Scores** option in the **Select Analysis Type** panel and select the achievement scores of interest in the **Achievement Scores** section.

Achievement Scores

This section is used to identify the set of plausible values to be used when achievement scores are the analysis variable for computing percentages and means, or the dependent variable in a regression analysis. The **With Achievement Scores** option in the **Select Analysis Type** panel should be checked before specifying achievement scores in the **Achievement Scores** section.

Dependent Variable

This is the variable to be used as the dependent variable when a regression analysis is specified. Only one dependent variable can be listed. To use achievement scores as the dependent variable, it is necessary to check the **With Achievement Scores** option in the **Select Analysis Type** panel and select the achievement scores of interest in the **Achievement Scores** section.

Achievement Benchmarks

They are the values that will be used as cut points on an achievement scale for computing the percentages of students meeting the specified benchmarks. Multiple cut points can be specified, each separated by a blank space.

Weight Variable

This is the sampling weight that will be used in the analysis. The IEA IDB Analyzer automatically selects the appropriate weight variable for analysis based on the file types included in the merged data file. Generally, this will be TOTWGT. MATWGT will be used when analyzing advanced mathematics teacher data and PHYWGT when analyzing physics teacher data. Chapter 4 provides more information on the TIMSS Advanced sampling weights.

Jackknifing Variables

They are the variables that capture the assignment of cases to sampling zones (JKZONE) and whether a case is to be dropped or have its weight doubled (JKREP) when computing the sets of replicate weights needed by the jackknife repeated replication method. The IEA IDB Analyzer automatically selects these variables and they cannot be changed.

2.5 TIMSS Advanced Analyses with Student-Level Variables

Many analyses of the TIMSS Advanced 2008 international database can be undertaken using only student-level data. This section presents examples of actual analyses that produced selected exhibits from the *TIMSS Advanced 2008 International Report*. These examples will compute percentages and means of subgroups, percentages of students reaching the TIMSS Advanced international benchmarks and also perform regression analyses.

The first example computes means for a straightforward continuous variable, whereas the second example computes means of achievement scores. In both cases, the IEA IDB Analyzer uses the sampling weights and implements the jackknife repeated replication method to compute appropriate sampling errors. In the second example, where achievement plausible values are used, the IEA IDB Analyzer effectively performs the computations five times—once for each plausible value—and aggregates the results to produce accurate estimates of mean achievement and standard errors that incorporate both sampling and imputation errors.

Student-Level Analysis

In our first example, we will replicate an analysis of advanced mathematics students' reported age at the time of testing. The results, presented in Exhibit 2.1 of the *TIMSS Advanced 2008 International Report*, are reproduced here in Exhibit 2.3. This example will focus on the results presented in the fourth data column—average age at time of testing.

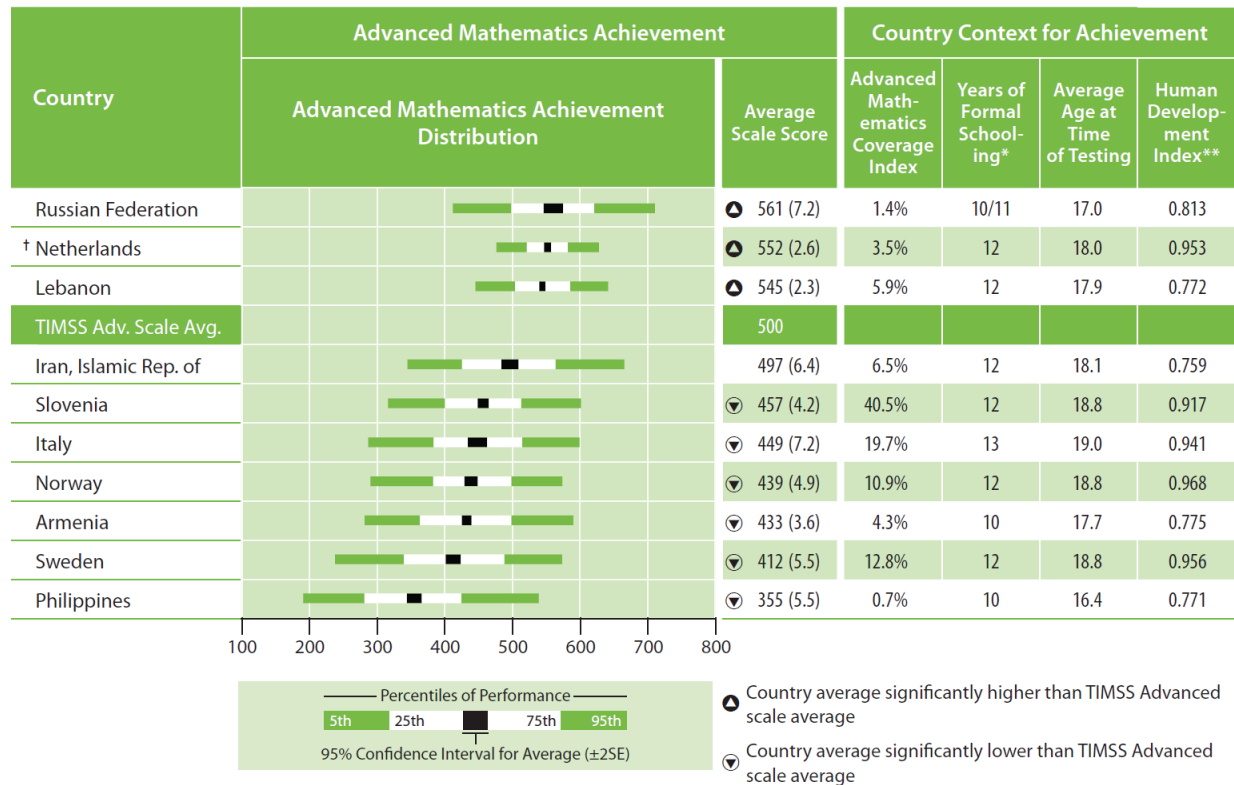
We need to undertake a number of steps to replicate the results in this exhibit. After reviewing the advanced mathematics student background data codebook (the codebooks are described in Section 4.4 of Chapter 4), we identify the student background variable MSDAGE as the variable that reports the age of advanced mathematics students at the time of testing.

After creating the merged data file MSGALLM2, the analysis module of the IEA IDB Analyzer is used to perform the analysis in the following steps:

- 1) Open the analysis module of the IEA IDB Analyzer.
- 2) Select the merged data file MSGALLM2 as the **Analysis File** by clicking the **Select** button.

Exhibit 2.3 Exhibit of Example Student-Level Analysis Taken from the TIMSS Advanced 2008 International Report (Exhibit 2.1)

Exhibit 2.1 TIMSS Advanced 2008 Distribution of Achievement in Advanced Mathematics

TIMSS Advanced 2008
Advanced Mathematics

SOURCE: IEA TIMSS Advanced 2008 ©

* Represents years of schooling counting from the first year of primary or basic education (first year of ISCED Level 1).

** Taken from United Nations Development Programme's *Human Development Report 2007/2008*, p.229-232.

† Met guidelines for sample participation rates only after replacement schools were included (see Appendix A).

() Standard errors appear in parentheses.

- 3) Select **Percentages and Means** as the **Analysis Type**. The **With Achievement Scores** option should remain unchecked since achievement scores are not used in this analysis. However, the **Exclude Missing from Analysis** option should be checked, which is done by default, to exclude cases that have missing values in the grouping variables.
- 4) The variable IDCNTRY is selected automatically as **Grouping Variables**. No additional grouping variables are needed for this analysis.
- 5) Specify the analysis variables. To activate this section, click the **Analysis Variables** radio button. For our example, the variable MSDAGE is selected from the list of available variables and moved to the **Analysis Variables** field by clicking the **right arrow** button in this section.

- 6) The **Weight Variable** is automatically selected by the software. As this example analysis uses student background data, TOTWGT is selected by default. The **Jackknifing Variables** JKZONE and JKREP also are selected by default.
- 7) Specify the name and folder of the output files in the **Output Files** field by clicking the **Define/Modify** button. The IEA IDB Analyzer will use this name and folder to create three output files: an SPSS syntax file that contains the code for performing the analysis, an SPSS data file with the results, and an Excel file with these same results.
- 8) Click the **Start SPSS** button to create the SPSS syntax file and open it in an SPSS syntax window. The syntax file can be executed by opening the **Run** menu of SPSS and selecting the **All** menu option. If necessary, the IEA IDB Analyzer prompts users to confirm the overwriting of existing files.

Exhibit 2.4 shows the completed analysis module for this example. The results are displayed in Exhibit 2.5, though only the first four countries are displayed to conserve space (this will be done for all analysis examples in this chapter).

Exhibit 2.4 IEA IDB Analyzer Setup for Example Student-Level Analysis

IEA IDB Analyzer: Analysis Module

1 Analysis File: C:\TIMSA2008\MSGALLM2.sav [Select]

2 Select Analysis Type

Percentages and Means Percentages only Regression Correlations Benchmarks Clear all Selections

With Achievement Scores Exclude Missing From Analysis Cumulative With Analysis Variable

3 Select Variables

Name	Description
MS2GBR1Y	GEN\YEAR OF BIRTH
MS2GBR1M	GEN\MONTH OF BIRTH
MS2GSEX	GEN\SEX OF STUDENT
MS2GOLAN	GEN\HOW OFTEN LANGUAGE OF TEST IS SPOKEN
MS2GBOOK	GEN\HOW MANY BOOKS AT HOME
MS2GTH01	GEN\HOME POSSESS\COMPUTER
MS2GTH02	GEN\HOME POSSESS\INTERNET CONNECTION
MS2GTH03	GEN\HOME POSSESS\YOUR OWN COMP
MS2GTH04	GEN\HOME POSSESS\YOUR OWN CALC
MS2GTH05	GEN\HOME POSSESS\STUDY DESK
MS2GTH06	GEN\HOME POSSESS\<COUNTRY SPECIFIC>
MS2GTH07	GEN\HOME POSSESS\<COUNTRY SPECIFIC>
MS2GTH08	GEN\HOME POSSESS\<COUNTRY SPECIFIC>
MS2GTH09	GEN\HOME POSSESS\<COUNTRY SPECIFIC>
MS2GHLEM	GEN\HIGHEST LEVEL OF EDUCATION\MOTHER
MS2GHLEF	GEN\HIGHEST LEVEL OF EDUCATION\FATHER
MS2GMBRN	GEN\MOTHER BORN IN <COUNTRY>
MS2GFBRN	GEN\FATHER BORN IN <COUNTRY>
MS2GBORN	GEN\BORN IN <COUNTRY>
MS2GBRNC	GEN\BORN IN <COUNTRY>\AGE OF COMING
MS2GITCE	GEN\INTEND TO CONTINUE EDUCATION
MS2GIAOS	GEN\INTENDED AREA OF STUDY
MS2GTUSC	GEN\USE COMPUTER\TIME

Grouping Variables

Name	Description
IDCOUNTRY	COUNTRY ID

Analysis Variables

Name	Description
MSDAGE	AGE OF STUDENT

Achievement Scores

Name	Description
------	-------------

Dependent Variable

Name	Description
------	-------------

Achievement Benchmarks

Name	Description
------	-------------

Weight Variable

Name	Description
TOTWGT	TOTAL STUDENT WEIGHT

Jackknifing Variables

Name	Description
JKZONE	JACKKNIFE ZONE

Number Of Decimals: 2

4 Output Files: C:\TIMSA2008\MSDAGE.* [Modify]

[Start SPSS] [Exit] [Help]

Exhibit 2.5 reports each country's average for the MSDAGE variable. The countries are identified in the first column. The second column reports the number of students with valid data. The third column reports the sum of weights of the sampled students, followed by the percentage, mean, and standard deviation, each accompanied by its jackknife standard error. The last column reports the percentage of missing values. From the first line of results, Armenia has valid data for 858 students and these sampled students represent a population of 2,684 students. Advanced mathematics students in Armenia were, on average, 17.68 years old at the time the TIMSS Advanced 2008 assessment took place, with a standard error of 0.02.

Exhibit 2.5 Output for Example Student-Level Analysis


Average for MSDAGE by (IDCNTRY)									PAGE	1
COUNTRY ID	N of Cases	Sum of TOTWGT	Percent	Percent (s.e.)	MSDAGE (Mean)	MSDAGE (s.e.)	Std.Dev.	Std.Dev. (s.e.)	Percent Missing	
Armenia	858	2684	1.13	.04	17.68	.02	.48	.01	.00	
Iran	2425	111298	46.79	1.70	18.08	.02	.43	.04	.00	
Italy	2143	119162	50.10	1.77	18.99	.02	.59	.03	.00	
Lebanon	1612	4702	1.98	.08	17.95	.03	.79	.03	.00	

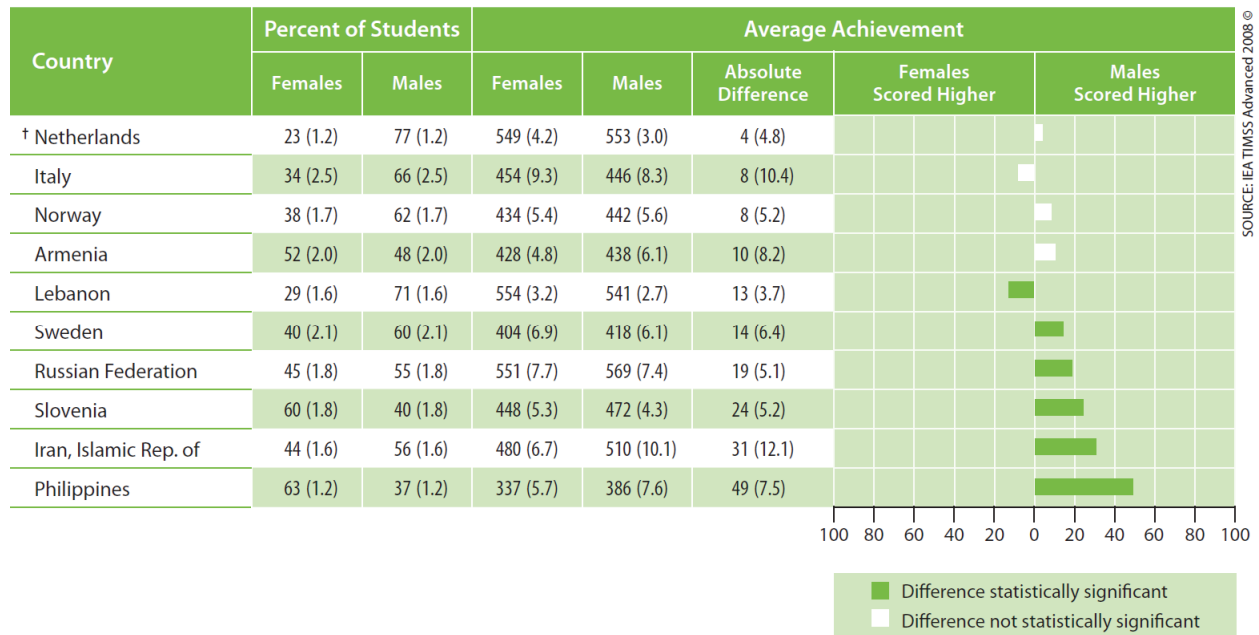
Student-Level Analysis with Achievement Scores

In our second example, we want to replicate another set of results presented in the *TIMSS Advanced 2008 International Report*. We are interested in investigating the relationship between students' gender and advanced mathematics achievement. These results, presented in Exhibit 2.4 of the *TIMSS Advanced 2008 International Report*, are repeated here in Exhibit 2.6. Since the results in this exhibit are based on plausible values, we need to make sure they are included when creating the input file, and also to indicate that this analysis will make use of achievement scores.

After reviewing the appropriate codebook, we observe that the variable ITSEX contains categorical information on the gender of students, and this variable is found in the student background data files. The **Percentages and Means** analysis type with activation of the **With Achievement Scores** checkbox will compute the percentages and mean achievement scores based on plausible values and their respective standard errors.

Exhibit 2.6 Exhibit of Example Student-Level Analysis with Achievement Scores Taken from the TIMSS Advanced 2008 International Report (Exhibit 2.4)

Exhibit 2.4 TIMSS Advanced 2008 Average Achievement in Advanced Mathematics by Gender 



† Met guidelines for sample participation rates only after replacement schools were included (see Appendix A).

() Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

The analysis module of the IEA IDB Analyzer is used to perform the analysis in the following steps:

- 1) Open the analysis module of the IEA IDB Analyzer.
- 2) Select the merged data file MSGALLM2 as the **Analysis File** by clicking the **Select** button.
- 3) Select **Percentages and Means** as the **Analysis Type**.
- 4) Check the **With Achievement Scores** box.
- 5) Add the variable ITSEX as a second **Grouping Variable**.
- 6) Specify the achievement scores to be used for the analysis. To activate this section, click the **Achievement Scores** radio button. Select the variable MSMMAT01-05 from the list of available variables and move it to the **Achievement Scores** field by clicking the **right arrow** button in this section.

- 7) The **Weight Variable** is automatically selected by the software. As this example analysis uses student background data, TOTWGT is selected by default. The **Jackknifing Variables** JKZONE and JKREP also are selected by default.
- 8) Specify the name and folder of the output files in the **Output Files** field by clicking the **Define/Modify** button.
- 9) Click the **Start SPSS** button to create the SPSS syntax file and open it in an SPSS syntax window. The syntax file can be executed by opening the **Run** menu of SPSS and selecting the **All** menu option. If necessary, the IEA IDB Analyzer prompts users to confirm the overwriting of existing files.

Exhibit 2.7 displays the analysis module with the proper settings for this example analysis. The output for this example is shown in Exhibit 2.8.

Exhibit 2.7 IEA IDB Analyzer Setup for Example Student-Level Analysis with Achievement Scores

The screenshot shows the 'IEA IDB Analyzer: Analysis Module' window. The 'Analysis File' is set to 'C:\TIMSA2008\MSGALLM2.sav'. Under 'Select Analysis Type', 'Percentages and Means' is selected, along with 'With Achievement Scores' and 'Exclude Missing From Analysis'. In the 'Select Variables' section, a list of variables is shown on the left, and several are moved to the right-hand panels: 'IDCOUNTRY' and 'ITSEX' are in the 'Grouping Variables' panel; 'MSMMAT01' is in the 'Achievement Scores' panel; 'TOTWGT' is in the 'Weight Variable' panel; and 'JKZONE' is in the 'Jackknifing Variables' panel. The 'Number Of Decimals' is set to 2. The 'Output Files' field is 'C:\TIMSA2008\MATGENDER.*'. At the bottom, there are buttons for 'Start SPSS', 'Exit', and 'Help'.

In Exhibit 2.8, each country's results are displayed on two lines, one for each value of the ITSEX variable. The countries are identified in the first column

and the second column describes the category of ITSEX being reported. The third column reports the number of students with valid data and the fourth the sum of weights of the sampled students. The next two columns report the percentage of students in each category and its standard error, followed by the estimated mean advanced mathematics achievement and its standard error. The standard deviation of advanced mathematics achievement and its standard error are reported in the last two columns. From the first two lines of results, 52.19% of advanced mathematics students in Armenia are girls, and 47.81% are boys. The mean achievement for advanced mathematics girls is 427.91 (standard error of 4.80) and 438.12 for boys (standard error of 6.07).

Exhibit 2.8 Output for Example Student-Level Analysis with Achievement Scores

Average for MSMMAT0 by IDCNTY ITSEX									
COUNTRY ID	SEX OF STUDENT	N of Cases	Sum of TOTWGT	Percent	Percent (s.e.)	MSMMAT0 (Mean)	MSMMAT0 (s.e.)	Std.Dev	Std.Dev (s.e.)
Armenia	GIRL	467	1401	52.19	1.96	427.91	4.80	89.17	3.26
	BOY	391	1283	47.81	1.96	438.12	6.07	101.38	4.75
Iran	GIRL	956	49158	44.17	1.59	479.73	6.65	82.07	4.45
	BOY	1469	62140	55.83	1.59	510.26	10.06	108.28	4.02
Italy	GIRL	727	40719	34.17	2.49	454.01	9.33	87.55	5.97
	BOY	1416	78444	65.83	2.49	446.08	8.28	99.20	4.42
Lebanon	GIRL	460	1363	29.00	1.56	553.74	3.19	56.16	2.37
	BOY	1152	3339	71.00	1.56	541.06	2.66	61.90	2.04

Student-Level Regression Analysis

This section presents a regression analysis using variables from the merged data file MSGALLM2. In this example, we will examine gender as a predictor of the advanced mathematics students' age at the time of testing (MSDAGE). This will allow us to determine if the age difference between girls and boys is statistically significant.

For this example, the values of the variable ITSEX are recoded into variable REGSEX by running the special SPSS syntax file SYNTAX_MSGALLM2.SPS shown in Exhibit 2.9 and provided on the DVD. REGSEX has a value of zero for girls and 1 for boys. By using REGSEX, the regression intercept—or constant—will be the estimated mean age of girls, whereas the regression slope—the REGSEX estimate column in the SPSS output—will be the estimated difference in mean age for boys.³

³ This form of variable recoding—known as “dummy coding”—makes the interpretation of regression coefficients easier. It transforms a regression analysis into an analysis of variance to test for differences among groups.

Exhibit 2.9 Example SPSS Program to Recode Variables for Student-Level Regression Analysis

```

* Compute new variable REGSEX from ITSEX .
get file = "<datpath>MSGALLM2.SAV" .

compute REGSEX = ITSEX - 1 .

value labels
  REGSEX 0 'GIRL'
         1 'BOY ' .

variable labels
  REGSEX "Recoded ITSEX (Girls = 0; Boys = 1)" .

save outfile = "<datpath>MSGALLM2.SAV" .

```

The parameter <datpath> in the SPSS syntax shown in Exhibit 2.9 needs to be edited to specify the location of the MSGALLM2.SAV data file.

The example regression analysis is performed by the analysis module of the IEA IDB Analyzer using the following steps:

- 1) Open the analysis module of the IEA IDB Analyzer.
- 2) Specify the data file MSGALLM2 as the **Analysis File** by clicking the **Select** button, after having run the SPSS syntax file SYNTAX_MSGALLM2.SPS to create the variable REGSEX.
- 3) Select Regression as the **Analysis Type**. Make sure that the **With Achievement Scores** option is not checked.
- 4) The variable IDCNTRY is selected automatically as **Grouping Variables**. No additional grouping variables are needed for this analysis.
- 5) Click the **Analysis Variables** radio button to activate the section and select REGSEX as the analysis variable. This is done by selecting REGSEX from the list of available variables and moving it to the **Analysis Variables** field by clicking the **right arrow** button in this section.
- 6) Click the **Dependent Variable** radio button. Select the variable MSDAGE from the list of available variables and move it to the **Dependent Variable** field by clicking the **right arrow** button in this section.
- 7) The **Weight Variable** is automatically selected by the software. As this example analysis uses student background data, TOTWGT is selected by default. The **Jackknifing Variables** JKZONE and JKREP also are selected by default.

- 8) Specify the name and folder of the output files in the **Output Files** field by clicking the **Define/Modify** button.
- 9) Click the **Start SPSS** button to create the SPSS syntax file and open it in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** menu option. If necessary, the IEA IDB Analyzer prompts users to confirm the overwriting of existing files.

Exhibit 2.10 illustrates the completed analysis module for this example regression analysis. The output for this example is shown in Exhibit 2.11.

Exhibit 2.10 IDB-Analyzer Set-Up for Example Student-Level Regression Analysis

The screenshot shows the 'IEA IDB Analyzer: Analysis Module' window. The 'Analysis File' is 'C:\TIMSA2008\MSGALLM2.sav'. Under 'Select Analysis Type', 'Regression' is selected. In 'Select Variables', 'REGSEX' is chosen as the dependent variable and 'MSDAGE' as the analysis variable. The 'Output Files' field is 'C:\TIMSA2008\AGEGENDER.*'. The 'Start SPSS' button is highlighted.

From the first line of results in Exhibit 2.11, the estimated mean age of advanced mathematics girls in Armenia, labeled “Constant (estimate)”, is 17.76 years, with a standard error of 0.02. The advanced mathematics boys are an estimated 0.17 years younger than the girls, as shown in the column labeled “REGSEX (estimate)”. With an estimated standard error of 0.03, this difference is statistically significant at the 95% confidence level.

Exhibit 2.11 Output for Example Student-Level Regression Analysis

Predictors: REGSEX / Predicted: MSDAGE							PAGE	1
COUNTRY ID	N of Cases	Mult_RSQ	Constant (estimate)	Constant (s.e.)	REGSEX (estimate)	REGSEX (s.e.)	REGSEX (t-test)	
Armenia	858	.03	17.76	.02	-.17	.03	-6.01	
Iran	2425	.00	18.11	.02	-.04	.03	-1.40	
Italy	2143	.01	18.90	.02	.15	.04	3.80	
Lebanon	1612	.01	17.85	.04	.14	.04	3.57	

Student-Level Regression Analysis with Achievement Scores

The next example of a student-level regression analysis will examine gender as a predictor of advanced mathematics achievement based on its five plausible values (MSMMAT01 through MSMMAT05). This example also will use the recoded variable REGSEX, this time to determine if the difference in mean achievement between girls and boys is statistically significant. The results of this analysis replicate the significance test applied to the gender differences in Exhibit 2.4 of the International Report (see Exhibit 2.6 of this chapter).

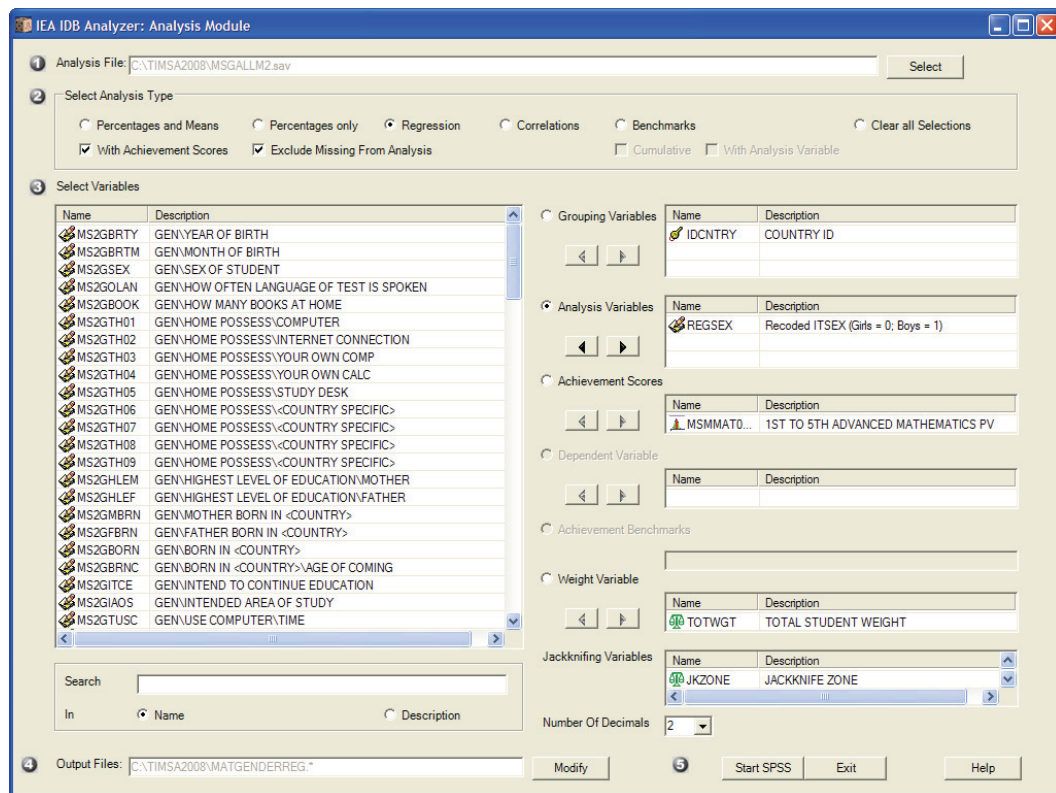
The regression analysis is performed by the analysis module of the IEA IDB Analyzer using the following steps:

- 1) Open the analysis module of the IEA IDB Analyzer.
- 2) Specify the data file MSGALLM2 as the **Analysis File** by clicking the **Select** button, after having run the SPSS syntax file SYNTAX_MSGALLM2.SPS to create the variable REGSEX.
- 3) Select **Regression** as the **Analysis Type**.
- 4) Check the **With Achievement Scores** box.
- 5) The variable IDCNTRY is selected automatically as **Grouping Variables**. No additional grouping variables are needed for this analysis.
- 6) Click the **Analysis Variables** radio button to activate the section. Select the variable REGSEX from the list of available variables and move it to the **Analysis Variables** field by clicking the **right arrow** button in this section.
- 7) Click the **Achievement Scores** radio button. Select the variable MSMMAT01-05 from the list of available variables and move it to the **Achievement Scores** field by clicking the **right arrow** button in this section.

- 8) The **Weight Variable** is automatically selected by the software. As this example analysis uses student background data, TOTWGT is selected by default. The **Jackknifing Variables** JKZONE and JKREP also are selected by default.
- 9) Specify the name and folder of the output files in the **Output Files** field by clicking the **Define/Modify** button.
- 10) Click the **Start SPSS** button to create the SPSS syntax file and open it in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** menu option. If necessary, the IEA IDB Analyzer prompts users to confirm the overwriting of existing files.

Exhibit 2.12 shows the analysis module with the proper settings for this example regression analysis with achievement scores. The output is displayed in Exhibit 2.13.

Exhibit 2.12 IDB-Analyzer Set-Up for Example Student-Level Regression Analysis with Achievement Scores



From the first line of results in Exhibit 2.13, the estimated mean achievement of advanced mathematics girls in Armenia is 427.91, with a standard error of 4.80; the same results shown in Exhibit 2.8. The estimated mean achievement of advanced mathematics boys in Armenia is 10.22 points higher than for girls. With an estimated standard error of 8.22, this difference is not statistically significant at the 95% confidence level.

Exhibit 2.13 Output for Example Student-Level Regression Analysis with Achievement Scores

Predictors: REGSEX / Predicted: MSMMAT0							PAGE	1
COUNTRY ID	N of Cases	Mult_RSQ	Constant (estimate)	Constant (s.e.)	REGSEX (estimate)	REGSEX (s.e.)	REGSEX (t-test)	
Armenia	858	.00	427.91	4.80	10.22	8.22	1.24	
Iran	2425	.02	479.73	6.65	30.53	12.07	2.53	
Italy	2143	.00	454.01	9.33	-7.93	10.36	-.77	
Lebanon	1612	.01	553.74	3.19	-12.68	3.75	-3.39	

Calculating Percentages of Students Reaching Benchmarks

This section describes how to use the IEA IDB Analyzer to perform analyses of student achievement in relation to the TIMSS achievement benchmarks. As an example, we will compute the percentages of students reaching the three TIMSS Advanced 2008 international benchmarks of advanced mathematics achievement (advanced, high, and intermediate) using the merged MSGALLM2 data file. These results, presented in Exhibit 3.2 of the *TIMSS Advanced 2008 International Report*, are repeated here in Exhibit 2.14.

This example is performed by the analysis module of the IEA IDB Analyzer using the following steps:

- 1) Open the analysis module of the IEA IDB Analyzer.
- 2) Specify the data file MSGALLM2 as the **Analysis File** by clicking the **Select** button.
- 3) Select **Benchmarks** as the **Analysis Type**.
- 4) Check the **Cumulative** box in the **Select Analysis Type** panel to get cumulated percentages of students reaching the international benchmarks.
- 5) The variable IDCNTRY is selected automatically as **Grouping Variables**. No additional grouping variables are needed for this analysis.

Exhibit 2.14 Example Exhibit of Benchmark Analysis Taken from the TIMSS Advanced 2008 International Report (Exhibit 3.2)

Exhibit 3.2 **Percent of Students Reaching the TIMSS Advanced 2008 International Benchmarks of Mathematics Achievement**

Country	Percent of Students Reaching the International Benchmarks			TIMSS Advanced Mathematics Coverage Index
	Advanced Benchmark (625)	High Benchmark (550)	Intermediate Benchmark (475)	
Russian Federation	24 (2.9)	55 (3.2)	83 (2.2)	1.4%
Iran, Islamic Rep. of	11 (1.8)	29 (3.0)	56 (2.8)	6.5%
Lebanon	9 (1.2)	47 (1.9)	88 (1.3)	5.9%
† Netherlands	6 (0.8)	52 (2.8)	95 (1.1)	3.5%
Italy	3 (1.0)	14 (2.0)	41 (3.0)	19.7%
Slovenia	3 (0.5)	14 (1.4)	41 (2.4)	40.5%
Armenia	2 (0.8)	13 (1.6)	33 (2.0)	4.3%
Norway	1 (0.4)	9 (1.0)	35 (2.2)	10.9%
Sweden	1 (0.4)	9 (1.2)	29 (1.9)	12.8%
Philippines	1 (0.3)	4 (0.7)	13 (1.5)	0.7%

TIMSS Advanced 2008
Advanced Mathematics

SOURCE: IEA TIMSS Advanced 2008 ©

† Met guidelines for sample participation rates only after replacement schools were included (see Appendix A).

() Standard errors appear in parentheses.

- 6) Click the **Achievement Scores** radio button. Select the variable MSMMAT01-05 from the list of available variables and move it to the **Achievement Scores** field by clicking the **right arrow** button in this section.
- 7) Click the **Achievement Benchmarks** radio button to activate this section and specify the TIMSS Advanced 2008 international benchmarks, which are 475, 550, and 625, respectively as intermediate, high, and advanced. Enter these three values in the input field, each separated by a blank space.
- 8) The **Weight Variable** is automatically selected by the software. As this example analysis uses student background data, TOTWGT is selected by default. The **Jackknifing Variables** JKZONE and JKREP also are selected by default.
- 9) Specify the name and folder of the output files in the **Output Files** field by clicking the **Define/Modify** button.

- 10) Click the **Start SPSS** button to create the SPSS syntax file and open it in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** menu option. If necessary, the IEA IDB Analyzer prompts users to confirm the overwriting of existing files.

Exhibit 2.15 displays the completed analysis module for this example. The output is displayed in Exhibit 2.16.

Exhibit 2.15 IDB-Analyzer Set-Up for Example Benchmark Analysis

IEA IDB Analyzer: Analysis Module

1 Analysis File: C:\TIMSA2008\MSGALLM2.sav [Select]

2 Select Analysis Type

Percentages and Means Percentages only Regression Correlations Benchmarks Clear all Selections

With Achievement Scores Exclude Missing From Analysis Cumulative With Analysis Variable

3 Select Variables

Name	Description
IDPUNCH	PUNCHER ID
IDSCHOOL	SCHOOL ID
IDCLASS	CLASS ID
IDGRADE	GRADE ID
IDSTUD	STUDENT ID
ITBIRTHM	DATE OF BIRTH/MONTH
ITBIRTHY	DATE OF BIRTH/YEAR
ITSEX	SEX OF STUDENT
IDBOOK	ACHIEVEMENT TEST BOOKLET
ITADMINI	TEST ADMINISTRATORS POSITION
ITDATEM	DATE OF TESTING/MONTH
ITDATEY	DATE OF TESTING/YEAR
ITLANG	LANGUAGE OF TESTING
ITPART	FINAL PARTICIPATION INDICATOR
IDPOP	POPULATION ID
IDGRADER	STANDARDIZED GRADE ID
IDSTRATE	EXPLICIT STRATUM CODE
IDSTRATI	IMPLICIT STRATUM CODE
MS2GBRTY	GEN\YEAR OF BIRTH
MS2GBRTM	GEN\MONTH OF BIRTH
MS2GSEX	GEN\SEX OF STUDENT
MS2GOLAN	GEN\HOW OFTEN LANGUAGE OF TEST IS SPOKEN
MS2GBOOK	GEN\HOW MANY BOOKS AT HOME

Grouping Variables

Name	Description
IDCOUNTRY	COUNTRY ID

Analysis Variables

Name	Description
MSMMAT0...	1ST TO 5TH ADVANCED MATHEMATICS PV

Achievement Scores

Name	Description
475 550 625	

Dependent Variable

Name	Description

Achievement Benchmarks

475 550 625

Weight Variable

Name	Description
TOTWGT	TOTAL STUDENT WEIGHT

Jackknifing Variables

Name	Description
JKZONE	JACKKNIFE ZONE

Number Of Decimals: 2

4 Output Files: C:\TIMSA2008\MATBENCHMARK.* [Modify] [Start SPSS] [Exit] [Help]

From the first few lines of results in Exhibit 2.16, 32.86% of the advanced mathematics students in Armenia are at or above the intermediate international benchmark of 475, with a standard error of 2.03; 12.68% of advanced mathematics students reached the high international benchmark, with a standard error of 1.56; 2.16% of the advanced mathematics students reached the advanced international benchmark, with a standard error of 0.77.

Exhibit 2.16 Output for Example Benchmark Analysis

Percent reaching benchmarks (475 550 625) of MSMMAT0					PAGE	1
COUNTRY ID	Performance Group	N of Cases	Sum of TOTWGT	Percent	Percent (s.e.)	
Armenia	Above 475	263	882	32.86	2.03	
	Above 550	90	340	12.68	1.56	
	Above 625	13	58	2.16	.77	
Iran	Above 475	1407	62593	56.24	2.76	
	Above 550	726	32160	28.90	2.95	
	Above 625	272	11949	10.74	1.76	
Italy	Above 475	815	49028	41.14	2.99	
	Above 550	266	16453	13.81	1.98	
	Above 625	50	3435	2.88	.97	
Lebanon	Above 475	1421	4122	87.68	1.34	
	Above 550	772	2209	46.98	1.94	
	Above 625	154	431	9.17	1.15	

Computing Correlations with Background Variables and Achievement Scores

In addition to the analyses described above, the IEA IDB Analyzer can compute correlations among background variables, and between background variables and achievement scores. While no examples are presented here, the steps for conducting these analyses are the same as those described previously: select the grouping variables, the analysis variables, the achievement scores (if necessary), and confirm the weight and sampling variables. The output will display, for each group defined by the grouping variables, the correlation coefficients for each possible pair of variables. When using only background variables, the diagonal and the elements above the diagonal of the correlation matrix are displayed with their respective standard errors. When using achievement scores, a single column is displayed containing the correlations between each of the background variables specified in the model and the achievement scores selected.

2.6 TIMSS Analyses with Teacher-Level Variables

Analyses with teacher background data seek to make statements about students whose teachers have a given characteristic, rather than about teachers with a given characteristic. As our example of an analysis using teacher background data, we will investigate the percentage of advanced mathematics students according to the age of their advanced mathematics teachers. The results of such an analysis are presented in Exhibit 5.1 of the *TIMSS Advanced 2008 International Report*, which is reproduced here in Exhibit 2.17.

Exhibit 2.17 Exhibit of Example Teacher-Level Analysis Taken from the TIMSS Advanced 2008 International Report (Exhibit 5.1)

Exhibit 5.1 **Advanced Mathematics Teachers' Gender, Age, and Number of Years Teaching**

Country	Percent of Students by Teacher Characteristics						Average Number of Years Teaching	
	Gender		Age				Teaching Altogether	Teaching Mathematics at the Advanced Level
	Female	Male	29 Years or Under	30–39 Years	40–49 Years	50 Years or Older		
Armenia	76 (4.6)	24 (4.6)	0 (0.0)	10 (2.4)	44 (5.4)	46 (5.4)	25 (0.9)	13 (1.2)
Iran, Islamic Rep. of	33 (2.5)	67 (2.5)	6 (2.3)	49 (3.8)	31 (3.4)	13 (2.5)	17 (0.6)	9 (0.3)
Italy	54 (5.4)	46 (5.4)	2 (1.4)	10 (3.0)	43 (4.7)	45 (4.6)	22 (0.9)	12 (0.8)
Lebanon	10 (1.5)	90 (1.5)	3 (0.9)	13 (1.9)	20 (2.0)	65 (2.4)	27 (0.5)	25 (0.5)
Netherlands	14 (3.5)	86 (3.5)	3 (2.0)	10 (2.1)	20 (4.8)	67 (5.2)	27 (1.1)	17 (1.1)
Norway	18 (3.9)	82 (3.9)	1 (0.7)	8 (2.3)	19 (4.2)	73 (4.3)	27 (0.9)	26 (0.9)
Philippines	63 (4.4)	37 (4.4)	25 (4.2)	36 (4.4)	25 (4.5)	14 (3.8)	14 (1.0)	5 (0.5)
Russian Federation	90 (2.7)	10 (2.7)	1 (0.6)	13 (3.0)	36 (5.2)	51 (5.2)	26 (0.8)	12 (0.8)
Slovenia	76 (5.1)	24 (5.1)	4 (1.9)	34 (5.8)	32 (5.9)	30 (5.5)	18 (1.1)	14 (0.7)
Sweden	19 (3.8)	81 (3.8)	2 (1.1)	18 (4.0)	22 (3.3)	58 (4.0)	22 (1.0)	9 (0.7)

TIMSS Advanced 2008
Advanced Mathematics

SOURCE: IEA TIMSS Advanced 2008 ©

Data provided by teachers.

An "s" indicates data are available for at least 50% but less than 70% of the students.

() Standard errors appear in parentheses.

We will use the **Percentages only** analysis type to estimate the percentages of students in the reporting categories of teachers' age since we are not concerned with student achievement in this analysis.

As in previous examples, we first identify the variables relevant to the analysis in the appropriate files and review the documentation for any specific national adaptations to the questions of interest (Supplements 1 and 2). In the advanced mathematics teacher background data files, we find that variable MT2GAGE contains the information on the age of advanced mathematics teachers.

The merged data file MTGALLM2 will be used for this example. Note that one of the steps in reproducing the results shown in Exhibit 2.17 is to combine response categories 1 and 2 and response categories 5 and 6 of the variable MT2GAGE in order to match the results presented in the International Report, where teachers are categorized into four groups: 29 years or under, 30 to 39 years, 40 to 49 years, and 50 years or older. The SPSS syntax shown in Exhibit 2.18 is used to recode MT2GAGE into a new variable NEWAGE and is available in the SPSS syntax file SYNTAX_MTGALLM2.SPS on the DVD. The

parameter <datpath> needs to be edited to specify the location of the MTGALLM2.SAV data file.

Exhibit 2.18 Example SPSS Program to Recode Variables for Teacher-Level Analysis

```
* Compute new variable NEWAGE from MT2GAGE .
get file = "<datpath>MTGALLM2.SAV" .
recode MT2GAGE (1,2=1) (3=2) (4=3) (5,6=4) (else=sysmis) into NEWAGE .

value labels
  NEWAGE 1 '29 YEARS OR UNDER'
          2 '30 TO 39 YEARS'
          3 '40 TO 49 YEARS'
          4 '50 YEARS OR OLDER' .

variable labels
  NEWAGE "Recoded MT2GAGE Teacher Age" .

save outfile = "<datpath> MTGALLM2.SAV " .
```

The example teacher-level analysis is performed by the analysis module of the IEA IDB Analyzer using the following steps:

- 1) Open the analysis module of the IEA IDB Analyzer.
- 2) Specify the data file MTGALLM2 as the **Analysis File** by clicking the **Select** button.
- 3) Select **Percentages only** as the **Analysis Type**.
- 4) Add the variable NEWAGE as a second **Grouping Variable**.
- 5) The **Weight Variable** is automatically selected by the software. As this example analysis uses advanced mathematics teacher background data, MATWGT is selected by default. The **Jackknifing Variables** JKZONE and JKREP also are selected by default.
- 6) Specify the name and folder of the output files in the **Output Files** field by clicking the **Define/Modify** button.
- 7) Click the **Start SPSS** button to create the SPSS syntax file and open it in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** menu option. If necessary, the IEA IDB Analyzer prompts users to confirm the overwriting of existing files.

Exhibit 2.19 displays the analysis module with the proper settings for this example analysis. The output is shown in Exhibit 2.20.

Exhibit 2.19 IDB-Analyzer Set-Up for Example Teacher-Level Analysis

The screenshot shows the 'IEA IDB Analyzer: Analysis Module' window. The 'Analysis File' is 'C:\TIMSA2008\MTGALLM2.sav'. Under 'Select Analysis Type', 'Percentages only' is selected. Under 'Select Variables', a list of variables is shown with 'NEWAGE' selected. On the right, 'Grouping Variables' includes 'IDCOUNTRY' and 'NEWAGE'. 'Weight Variable' is set to 'MATWGT'. 'Jackknifing Variables' includes 'JKZONE'. The 'Number Of Decimals' is set to 2. At the bottom, there are buttons for 'Start SPSS', 'Exit', and 'Help'.

Each country's results in Exhibit 2.20 are presented on four lines, one for each value of the NEWAGE variable. The results are presented in the same manner as in previous examples, with countries identified in the first column and the second column describing the categories of NEWAGE. From the first three lines of results, 10.30% of students in Armenia were taught by teachers 30 to 39 years, 43.74% by teachers 40 to 49 years, and 45.96% by teachers 50 years or older. The appropriate standard errors also are presented in Exhibit 2.20. Note that Armenia does not display a row for teachers 29 years or under because there were none in their national sample.

Exhibit 2.20 Output for Example Teacher-Level Analysis

Percentages by (IDCOUNTRY NEWAGE)					PAGE 1
COUNTRY ID	Recorded MT2GAGE Teacher Age	N of Cases	Sum of MATWGT	Percent	Percent (s.e.)
Armenia	30 TO 39 YEARS	92	273	10.30	2.43
	40 TO 49 YEARS	355	1160	43.74	5.39
	50 YEARS OR OLDER	401	1218	45.96	5.41
Iran	29 YRS OR UNDER	87	7047	6.33	2.35
	30 TO 39 YEARS	1163	54802	49.24	3.79
	40 TO 49 YEARS	880	34893	31.35	3.44
	50 YEARS OR OLDER	295	14556	13.08	2.54
Italy	29 YRS OR UNDER	35	2101	1.77	1.38
	30 TO 39 YEARS	189	12108	10.19	3.04
	40 TO 49 YEARS	920	50623	42.62	4.67
	50 YEARS OR OLDER	991	53950	45.42	4.65
Lebanon	29 YRS OR UNDER	37	118	2.55	.89
	30 TO 39 YEARS	247	614	13.27	1.89
	40 TO 49 YEARS	362	908	19.62	1.95
	50 YEARS OR OLDER	1173	2990	64.57	2.44

2.7 TIMSS Analyses with School-Level Variables

When performing analyses with school background data, the data are analyzed to make statements about students attending schools with a given characteristic, rather than about schools with a given characteristic. Our example of an analysis using school background data will compute the percentages of advanced mathematics students who attend schools with a high, medium, and low index of principals' perception of school climate. We can use the variable MCDGPPSC for this purpose. We also will calculate the mean advanced mathematics achievement at each level of principals' perception of school climate. The results of this analysis are presented in Exhibit 6.5 of the *TIMSS Advanced 2008 International Report*, which is reproduced here in Exhibit 2.21.

For this analysis, we will use the **Percentages and Means** analysis type of the IEA IDB Analyzer with the **With Achievement Scores** option checked. The variable MCDGPPSC in the school background data files contains information on the index of principals' perception of school climate. As MCDGPPSC is a derived variable, we can refer to Supplement 3 to see how it was created. We also should review the documentation on national adaptations to its component variables in Supplement 2.

Exhibit 2.21 Exhibit of Example School-Level Analysis Taken from the TIMSS Advanced 2008 International Report (Exhibit 6.5)

Exhibit 6.5 Index of Principals' Perceptions of School Climate (PPSC)

TIMSS Advanced 2008
Advanced Mathematics

Country	High PPSC		Medium PPSC		Low PPSC	
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement
Philippines	53 (5.3)	354 (7.8)	43 (4.5)	365 (9.8)	4 (2.2)	303 (28.8)
Slovenia	25 (3.6)	506 (6.1)	68 (4.4)	447 (6.0)	6 (2.6)	405 (8.3)
Iran, Islamic Rep. of	25 (4.0)	528 (14.0)	59 (5.2)	496 (9.1)	16 (3.7)	449 (6.4)
Lebanon	25 (2.0)	558 (4.6)	59 (2.1)	543 (3.1)	16 (1.3)	525 (3.7)
Sweden	18 (4.9)	438 (11.0)	73 (4.8)	411 (6.2)	10 (3.3)	381 (9.0)
Russian Federation	13 (3.4)	605 (18.6)	81 (3.8)	559 (6.7)	6 (2.0)	494 (24.7)
Norway	7 (2.7)	441 (11.2)	90 (3.3)	440 (5.1)	4 (1.9)	403 (58.4)
Italy	3 (1.8)	481 (45.6)	60 (5.1)	458 (8.7)	37 (5.0)	431 (12.4)
Armenia	2 (0.1)	~ ~	83 (0.4)	436 (4.1)	15 (0.4)	420 (5.3)
Netherlands	1 (0.7)	~ ~	68 (5.5)	555 (3.7)	31 (5.5)	547 (3.7)

SOURCE: IEA TIMSS Advanced 2008 ©

Based on principals' responses to the following aspects of school climate in their schools: teachers' job satisfaction, teachers' opportunities for professional development, teachers' understanding of the school's curricular goals, teachers' degree of success in implementing the school's curriculum, teachers' expectations for student achievement, parental support for student achievement, parental involvement in school activities, students' regard for school property, and students' desire to do well in school. Average is computed across the nine statements based on a 5-point scale: 1 = Very High, 2 = High, 3 = Medium, 4 = Low, 5 = Very Low. High level indicates students whose principals' perception of their school climate was very positive (average is less than or equal to 2). Medium level indicates

students whose principals' perception of their school climate was moderately positive (average is greater than 2 and less than 3). Low level indicates students whose principals' perception of their school climate was not so positive (average is greater than or equal to 3).

() Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

A tilde (~) indicates insufficient data to report achievement.

An "r" indicates data are available for at least 70% but less than 85% of the students.

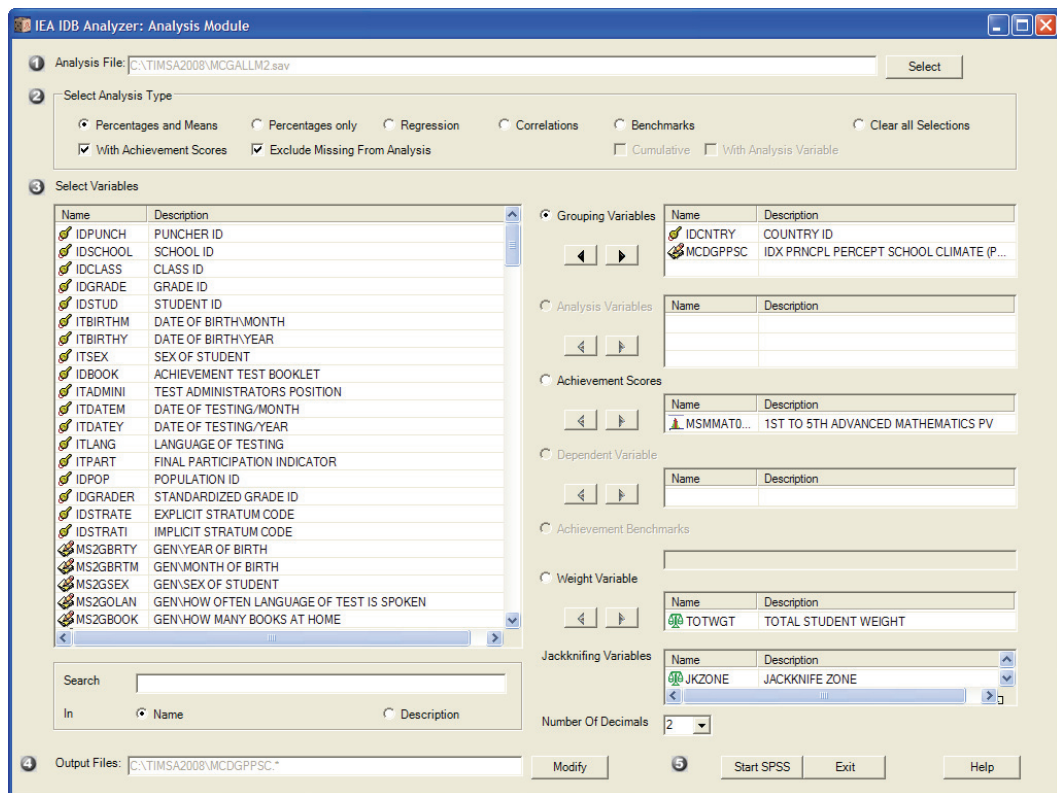
The merged data file MCGALLM2 will be used for this example. This school-level analysis is performed by the analysis module of the IEA IDB Analyzer using the following steps:

- 1) Open the analysis module of the IEA IDB Analyzer.
- 2) Specify the data file MCGALLM2 as the **Analysis File** by clicking the **Select** button.
- 3) Select **Percentages and Means** as the **Analysis Type**.
- 4) Check the **With Achievement Scores** box.
- 5) Add the variable MCDGPPSC as a second **Grouping Variable**.
- 6) Click the **Achievement Scores** radio button. Select the variable MSMMAT01-05 from the list of available variables and move it to the **Achievement Scores** field by clicking the **right arrow** button in this section.

- 7) The **Weight Variable** is automatically selected by the software. As this example analysis uses school background data linked to student background data, TOTWGT is selected by default. The **Jackknifing Variables** JKZONE and JKREP also are selected by default.
- 8) Specify the name and folder of the output files in the **Output Files** field by clicking the **Define/Modify** button.
- 9) Click the **Start SPSS** button to create the SPSS syntax file and open it in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** menu option. If necessary, the IEA IDB Analyzer prompts users to confirm the overwriting of existing files.

Exhibit 2.22 shows the analysis module with the proper settings for this example analysis. The output is shown in Exhibit 2.23.

Exhibit 2.22 IDB-Analyzer Set-Up for Example School-Level Analysis



Each country's results in Exhibit 2.23 are displayed on three lines, one for each value of the MCDGPPSC variable. The results are presented in the same manner as in previous examples, with countries identified in the first column and the second column describing the categories of MCDGPPSC. From the first three lines of results, 1.83% of advanced mathematics students in Armenia attend schools with a high level of principals' perception of school climate, 82.68% attend schools with a medium level, and 15.50% attend schools with a low level of principals' perception of school climate. Also, the estimated mean achievement of advanced mathematics students in schools with a high level of principals' perception of school climate is 409.25 (standard error of 16.21), whereas the estimated mean achievement of advanced mathematics students in schools with medium and low levels of principals' perception of school climate are 435.74 (standard error of 4.14) and 419.84 (standard error of 5.35), respectively.

Exhibit 2.23 Output for Example School-Level Analysis

Average for MSMMAT0 by IDCNTY MCDGPPSC									PAGE	1
COUNTRY ID	IDX PRNCPL PERCEPT SCHOOL CLIMATE (PPSC)	N of Cases	Sum of TOTWGT	Percent	Percent (s.e.)	MSMMAT0 (Mean)	MSMMAT0 (s.e.)	Std.Dev	Std.Dev (s.e.)	
Armenia	HIGH	23	49	1.83	.08	409.25	16.21	69.41	16.64	
	MEDIUM	661	2219	82.68	.41	435.74	4.14	97.70	3.15	
	LOW	174	416	15.50	.40	419.84	5.35	82.93	5.22	
Iran	HIGH	532	26902	24.89	3.97	527.72	14.04	98.61	5.75	
	MEDIUM	1390	63435	58.69	5.21	496.02	9.13	98.46	4.54	
	LOW	440	17745	16.42	3.67	448.74	6.38	72.61	3.35	
Italy	HIGH	71	3656	3.07	1.78	481.37	45.58	75.47	28.27	
	MEDIUM	1260	71625	60.11	5.15	458.18	8.74	95.28	5.80	
	LOW	812	43881	36.82	5.01	430.75	12.36	94.30	5.23	
Lebanon	HIGH	412	1131	24.59	1.96	558.22	4.64	61.24	3.17	
	MEDIUM	902	2715	59.04	2.07	542.89	3.09	60.36	2.05	
	LOW	261	753	16.37	1.29	524.50	3.73	52.82	3.73	

References

IEA (2009). *IEA international database analyzer* (version 2.0) [Computer software and manual]. Hamburg: IEA Data Processing and Research Center.

LaRoche, S., Zuehlke, O., & Joncas, M. (2009). TIMSS Advanced 2008 sampling. Chapter 4 in I.V.S. Mullis, M.O. Martin, P. Foy, & A. Arora (Eds.), *TIMSS Advanced 2008 technical report* (pp. 51-92). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., Robitaille, D.F., & Foy, P. (2009). *TIMSS advanced 2008 international report: Findings from IEA's study of achievement in advanced mathematics and physics in the final year of secondary school*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

SPSS Inc. (2008). *SPSS for Windows* (version 16.0). Chicago, IL: SPSS Inc.

A close-up, black and white photograph of typewriter keys. The keys are arranged in a grid, with some keys in the foreground being sharp and others in the background being blurred. The keys have various symbols and letters on them, including Greek letters like Omega (Ω), Pi (π), and Alpha (α), as well as numbers like 7 and 0. The lighting creates highlights on the metallic surfaces of the keys.

Chapter 3

Analyzing the TIMSS Advanced 2008 International Database Using SAS

3.1 Overview

Although users of the TIMSS Advanced 2008 international database are encouraged to use the IEA IDB Analyzer in conjunction with SPSS because it is easy to use and deals effectively with the complexity of the TIMSS Advanced 2008 data, this chapter also presents some basic examples of analyses that can be performed with the TIMSS Advanced 2008 international database using the SAS statistical analysis system (SAS, 2002) and the SAS programs and macros provided on the DVD. The SAS macros use sampling weights and a jackknifing algorithm to deal with the TIMSS complex sample design and take into account plausible values when analyzing student achievement.

Although some familiarity with the structure of the TIMSS Advanced 2008 database will be helpful, the analyses presented in this chapter are simple in nature, and are designed primarily to familiarize users with the various data files and their structure, as well as the variables to be used in most analyses. Chapter 4 provides a more detailed description of the data files contained in the international database—their structure and contents, along with detailed information on all the supporting materials provided on the DVD.

The examples in this chapter compute percentages of students in specified subgroups, average advanced mathematics achievement in those subgroups, and appropriate standard errors for these statistics. Additional examples compute regression coefficients and their standard errors. The example analyses, using student, teacher and school data, replicate some of the analyses that are included in the *TIMSS Advanced 2008 International Report* (Mullis, Martin, Robitaille, & Foy, 2009). Users are encouraged to practice analyzing the TIMSS Advanced 2008 data by replicating some of the exhibits presented in the International Report.

Before doing any statistical analyses with the TIMSS Advanced 2008 international database, users should copy the contents of the TIMSS Advanced 2008 database DVD to an alternate location, either on their computer or on a server. For the purposes of this chapter, we will assume that all SAS files on the DVD have been copied to the “C:\TIMSA2008\SAS_Data” folder.

All SAS programs presented in this chapter are available on the DVD. They can be adapted to perform a variety of analyses with some basic knowledge of the SAS language. With a little experience and some practice with these programs, users should be able to make the necessary modifications to obtain the desired results. The example SAS programs invoke SAS macros that will be described in this chapter. Although users will be expected to modify the example programs, there is no need to make any changes within the SAS macros. The DVD also includes an updated release of the TIMSS Advanced 1995 international database and SAS programs to analyze its data.¹

3.2 SAS Programs and Macros

The “Programs” folder on the TIMSS Advanced 2008 international database DVD includes a number of SAS programs needed to process the SAS data files, compute survey results, and carry out example analyses. This chapter gives detailed instructions on how to adapt and make use of them.

CONVERT.SAS

This SAS program converts the SAS Export files found on the DVD into SAS data files. All programs and macros described in this chapter require that the SAS Export files be converted into SAS data files.

MSASCRM2.SAS, PSASCRM2.SAS²

These two SAS programs can be used to convert the response codes to the achievement items to their corresponding score levels.

JOIN.SAS

This SAS program combines files of the same type from more than one country.

¹ This chapter presents analysis examples using data from the 2008 database, which can be adapted and applied to the data from the 1995 database.

² SPSS versions of these two programs also are available on the DVD.

JACKGEN.SAS (and SAMPLEJACKGEN.SAS)

This SAS macro program is used to compute weighted percentages of students within defined subgroups, along with their means on a specified continuous variable. This macro generates replicate weights and computes standard errors using the jackknife repeated replication (JRR) methodology. The analysis variable can be any continuous variable. When computing mean achievement scores with plausible values, the macro JACKPV.SAS should be used.

JACKPV.SAS (and SAMPLEJACKGPV.SAS)

This SAS macro program is used to compute weighted percentages of students within defined subgroups, along with their mean achievement on a scale using the available plausible values. This macro generates replicate weights and computes standard errors using the jackknife repeated replication (JRR) and multiple imputation methodologies. This macro should be used when achievement plausible values are needed in an analysis.

JACKREG.SAS (and SAMPLEJACKREG.SAS)

This SAS macro program is used to compute weighted regression coefficients and their standard errors within defined subgroups. This macro can be used with any analysis variable, but is not appropriate for analyzing achievement with plausible values.

JACKREGP.SAS (and SAMPLEJACKREGP.SAS)

This SAS macro program is used to compute weighted regression coefficients and their standard errors within defined subgroups when using achievement plausible values as the dependent variable.

Each of the four SAS macro programs above has a corresponding sample program that calls its respective macro and prints out the results. These sample programs are discussed later in this chapter.

EXAMPLE1.SAS, EXAMPLE2.SAS, EXAMPLE3.SAS, EXAMPLE4.SAS

They are the SAS programs used in the example analyses presented in this chapter.

3.3 Converting the SAS Export Files

The DVD provides a program called CONVERT.SAS that converts the SAS Export files provided on the DVD into SAS data files. This conversion is necessary since all the SAS macros and SAS programs presented in this chapter require the use of SAS data files. Two versions of this program are available for the 2008 and 1995 databases.

To convert SAS Export files into SAS data files, users should apply the following steps:

- 1) Open the SAS program file CONVERT.SAS.
- 2) At the beginning of the program, specify the data file type in the parameter “TYPE”.
- 3) Specify the path where the SAS Export files are located in the parameter “EXPPATH”.
- 4) Specify the folder where the converted SAS data files will be located in the parameter “DATPATH”.
- 5) List all the countries of interest in the parameter “COUNTRY”
By default, all TIMSS Advanced 2008 countries are listed and the program will automatically select the appropriate list by population based on the file type specified.
- 6) Submit the edited code for processing.

An example of the CONVERT program is presented in Exhibit 3.1. It converts the SAS Export files of type MSG for all countries. For this example, all SAS Export files are located in the “C:\TIMSA2008\SAS_Data” folder and the converted SAS data files also will be located in this folder.

Users are advised to run the CONVERT program for all countries and all file types. The file types for advanced mathematics are MCG, MSA, MSG, MSR, MST, and MTG. For physics, the file types are PCG, PSA, PSG, PSR, PST, and PTG. The file types are described in Chapter 4. In principle, this program needs to be run only once for each file type and should be one of the first things users do with the TIMSS Advanced 2008 international database before moving on to any data analyses, in particular the data analysis examples in this chapter.

Exhibit 3.1 Example of CONVERT Program Used to Convert SAS Export Files into SAS Data Files

```

%LET TYPE = MSG ;

%LET EXPPATH = C:\TIMSA2008\SAS_Data\ ;

%LET DATPATH = C:\TIMSA2008\SAS_Data\ ;

%MACRO DOIT ;

    %LET COUNTRY = < List of TIMSS Advanced 2008 countries > ;

    %LET I = 1 ;
    %DO %WHILE(%LENGTH(%SCAN(&COUNTRY,&I))) ;
        %LET CTRY = %SCAN(&COUNTRY,&i) ;

        PROC CIMPORT FILE = "&EXPPATH&TYPE&CTRY.M2.EXP"
                    DATA = "&DATPATH&TYPE&CTRY.M2" ;

        %LET I = %EVAL(&I + 1) ;
    %END ;

%MEND DOIT ;

%DOIT ;

```

3.4 Scoring Individual TIMSS Advanced 2008 Items

Student achievement in TIMSS Advanced is represented by a set of five plausible values for each population and these are the preferred scores for any analysis of student achievement. However, analyzing performance on individual items may be of interest to some users. Carrying out such analyses requires that the individual items in the TIMSS Advanced 2008 database be assigned their correctness score levels, rather than the actual response options selected by students for multiple-choice items, or the two-digit codes given to students' responses to constructed-response items. The DVD provides SAS programs to perform this task and two versions are available for the 2008 and 1995 databases. SPSS versions are also available.

For multiple-choice items, numbers 1 through 5 are used to represent response options A through E, respectively, in the TIMSS Advanced 2008 achievement data files. These responses need to be converted to their appropriate score level ("1" for correct and "0" for incorrect) based on each multiple-choice item's correct response key. For constructed-response items, worth a total of 1 or 2 points, two-digit codes are used to represent the students' written responses in the TIMSS Advanced database (see Section 4.3.1 of Chapter 4). These codes also need to be recoded to represent the correct point values of the responses—either 0, 1, or 2 points.

For both types of items, special codes are set aside to represent missing data as either “Not Administered”, “Omitted”, or “Not Reached”. These special missing codes also must be recoded in order to carry out specific item-level analyses. By default, the not-administered response code is left as missing and the omitted and not-reached response codes are coded as incorrect. These default settings can be modified within the programs, depending on the requirements of the item-level analyses. For example, not-reached responses were treated as missing for the purpose of calibrating the TIMSS Advanced 2008 items, whereas they were treated as incorrect when deriving achievement scores for students.

The DVD includes two SAS programs—MSASCRM2.SAS for the advanced mathematics items and PSASCRM2.SAS for the physics items—which will recode the responses to individual items from the achievement data files to their appropriate score levels. To score each individual TIMSS Advanced 2008 item, the program code in the MSASCRM2 and PSASCRM2 programs needs to be adapted. Users should do the following steps:

- 1) Open the SAS program file MSASCRM2.SAS or PSASCRM2.SAS.
- 2) Specify the folder where the SAS data files are located in the “LIBNAME” statement.
- 3) List all the countries of interest in the parameter “COUNTRY”.
By default, all TIMSS Advanced 2008 countries are listed.
- 4) Submit the edited code for processing.

Each program uses the student achievement data files as input (MSA/PSA), recodes the individual items and saves the results in SAS data files that have “MSC/PSC” as the first three characters in their file names. Exhibit 3.2 shows a condensed version of the MSASCRM2 and PSASCRM2 programs to score the individual TIMSS Advanced 2008 items.

Exhibit 3.2 Example of MSASCRM2/PSASCRM2 Programs for Converting Individual Item Response Codes to their Score Level

```
LIBNAME LIBDAT "C:\TIMSA2008\SAS_Data\" ;

%LET COUNTRY = < List of TIMSS Advanced 2008 countries > ;

%LET ARIGHT = < List of multiple-choice items where A is correct > ;
%LET BRIGHT = < List of multiple-choice items where B is correct > ;
%LET CRIGHT = < List of multiple-choice items where C is correct > ;
%LET DRIGHT = < List of multiple-choice items where D is correct > ;
%LET ERIGHT = < List of multiple-choice items where E is correct > ;
%LET CONSTR = < List of constructed-response items > ;

%MACRO SCOREIT (ITEM, TYPE, RIGHT, NR, NA, OM, OTHER) ;

    . . .

%MEND SCOREIT ;

%MACRO DOIT ;

    . . .

    DO OVER ARIGHT ; %SCOREIT (ARIGHT, "MC", 1, .R, .A, ., .I) ; END ;
    DO OVER BRIGHT ; %SCOREIT (BRIGHT, "MC", 2, .R, .A, ., .I) ; END ;
    DO OVER CRIGHT ; %SCOREIT (CRIGHT, "MC", 3, .R, .A, ., .I) ; END ;
    DO OVER DRIGHT ; %SCOREIT (DRIGHT, "MC", 4, .R, .A, ., .I) ; END ;
    DO OVER ERIGHT ; %SCOREIT (ERIGHT, "MC", 5, .R, .A, ., .I) ; END ;
    DO OVER CONSTR ; %SCOREIT (CONSTR, "CR", , .R, .A, ., .I) ; END ;

    . . .

%MEND DOIT ;

%DOIT ;
```

If not reached responses are to be treated as missing rather than incorrect, users should replace the following statement (which appears twice in the programs):

```
IF &ITEM = &NR THEN SCORE = 0 ;
```

with this statement:

```
IF &ITEM = &NR THEN SCORE = . ;
```

3.5 Joining the TIMSS Advanced 2008 Data Files

The TIMSS Advanced 2008 international database contains separate data files for each country. The DVD provides a SAS program called JOIN.SAS that joins individual country data files of a particular type into a single aggregated data file, facilitating joint analyses involving more than one country. Two versions of this program are available for the 2008 and 1995 databases. This program, however, can only join SAS data files of the same type. The JOIN

program can be used for the following data file types: MCG/PCG, MSA/PSA, MSC/PSC, MSG/PSG, MSR/PSR, MST/PST, and MTG/PTG. To create a SAS data file with more than one country's data, users should do the following:

- 1) Open the SAS program file JOIN.SAS.
- 2) At the beginning of the program, specify the data file type in the parameter "TYPE".
- 3) Specify the folder where the SAS data files are located in the LIBDAT statement.
- 4) List all the countries of interest in the parameter "COUNTRY".
By default, all TIMSS Advanced 2008 countries are listed.
- 5) Submit the edited code for processing.

An example of the JOIN program is displayed in Exhibit 3.3. It joins the advanced mathematics student background data files (MSG) of all countries. All country data files are located in the "C:\TIMSA2008\SAS_Data" folder for the sake of this example. The resulting data file, MSGALLM2, will be saved in this folder as well.

Exhibit 3.3 Example of JOIN Program Used to Join SAS Data Files for More Than One Country

```
%LET TYPE = MSG ;

LIBNAME LIBDAT "C:\TIMSA2008\SAS_Data\" ;

%MACRO DOIT ;

  %LET COUNTRY = < List of TIMSS Advanced 2008 countries > ;

  DATA &TYPE.ALLM2 ;
    SET %LET I = 1 ;
        %DO %WHILE (%LENGTH(%SCAN(&COUNTRY,&I))) ;
          %LET CTRY = %SCAN(&COUNTRY,&I) ;
          LIBDAT.&TYPE&CTRY.M2
          %LET I = %EVAL(&I + 1) ;
        %END ; ;

  PROC SORT DATA = &TYPE.ALLM2 OUT = LIBDAT.&TYPE.ALLM2 ;
    BY &SORTVARS ;

%MEND DOIT ;

%DOIT ;
```

3.6 SAS Macros to Compute Statistics and their Standard Errors

This section describes the four SAS macros needed to compute specific statistics with their correct standard errors, along with sample SAS programs to demonstrate their use. Users are encouraged to modify the sample SAS programs and familiarize themselves with their functioning. However, the four SAS macros do not require any modifications.

Each SAS macro serves a specific analytical purpose. These macros ensure that analyses of the TIMSS Advanced 2008 data are done properly. Sampling weights are used and standard errors are computed using the jackknife repeated replication (JRR) method. Furthermore, achievement scores are based on sets of five plausible values that take into account the measurement error arising from the test design and the IRT scaling methodology. The macros that make use of plausible values effectively perform five analyses—one for each plausible value—and aggregate the results to produce accurate estimates of achievement and standard errors that incorporate both sampling and imputation errors.

The sample SAS programs presented in this section all use as input the SAS data file MSGALLM2, which contains the advanced mathematics student background data files of all participating countries. In all sample programs, the parameter <datpath> must be edited to specify the folder where the MSGALLM2 file is located.

Computing Means and Their Standard Errors (JACKGEN)

The JACKGEN macro is used to compute percentages and means of continuous variables with their JRR standard errors. We will demonstrate its use with a sample SAS program that calls the macro JACKGEN to compute the percentages of students within specified subgroups and their mean on a variable of choice. The macro also computes the appropriate standard errors for the percentages and means. However, this macro is not appropriate for analyzing achievement means based on plausible values; the JACKPV macro should be used for this purpose.

The JACKGEN macro is a self-contained program, located in the program file JACKGEN.SAS, and should not be modified. It essentially computes sets of replicate weights using the sampling and weighting variables, aggregates the

data by subgroups using the replicate weights, and then computes and stores the desired statistics in a SAS working file called FINAL.

The macro JACKGEN is included in a SAS program by issuing the following command:

```
%INCLUDE "<macpath>JACKGEN.SAS" ;
```

where <macpath> points to the folder where the SAS macro program JACKGEN.SAS is located. The macro requires that several parameters be specified as input when it is invoked:

- | | |
|------|--|
| WGT | The sampling weight to be used in the analysis. Generally, TOTWGT should be used. MATWGT should be used when analyzing advanced mathematics teacher data and PHYWGT when analyzing physics teacher data. |
| JKZ | The variable that captures the assignment of cases to sampling zones. The name of this variable in all TIMSS Advanced data files is JKZONE. |
| JKR | The variable that captures whether the case is to be dropped or have its weight doubled for each set of replicate weights. The name of this variable in all TIMSS Advanced data files is JKREP. |
| NJKZ | The number of replicate weights to be generated when computing the JRR standard errors. The value of NJKZ should be set to 75, the maximum possible value across all participating countries. |
| CVAR | The list of variables that are to be used to define the subgroups. The list can consist of one or more variables. We recommend that users always include IDCNTRY as the first classification variable. |
| DVAR | The variable for which means are to be computed. Only one variable can be listed and it should be a continuous variable. Plausible values of achievement scores should not be specified here. |

INFILE The name of the data file that contains the data being analyzed. If the folder is included as part of the file name, the name of the file must be enclosed in quotes. It is important to emphasize that this data file must include only those cases that are of interest in the analysis. If users want to have specific cases excluded from the analysis, for example students with missing data, this should be done prior to invoking the macro.

The JACKGEN macro is invoked by a SAS program using the conventional SAS notation for invoking macros. This involves listing the macro name followed by the list of parameters in parenthesis, each separated by a comma. For example, the JACKGEN macro invoked with the following statement:

```
%JACKGEN (TOTWGT, JKZONE, JKREP, 75, IDCNTY ITSEX, MSDAGE, MSGALLM2) ;
```

will compute the mean age (MSDAGE) of advanced mathematics students by gender (ITSEX) and their standard errors within each country (IDCNTY), using the weighting variable TOTWGT. It also will compute the percentages of boys and girls and their standard errors within each country. The data will be read from the data file MSGALLM2 and the standard errors will be computed based on 75 sets of replicate weights.

The results of the JACKGEN macro are stored in a SAS working file called FINAL, which is stored in the default folder used by SAS. The following variables are contained in this results file:

Classification Variables

All classification variables are kept in the results file. In the example invocation above, there are two classification variables: IDCNTY and ITSEX. There is one record in the results file for each subgroup defined by the categories of the classification variables.

N

This variable contains the number of valid cases for each subgroup defined by the classification variables. In the example, it is the number of boys and girls with valid data in each country's sample.

Weight Variable

The weight variable contains the sum of weights within each subgroup defined by the classification variables. In the example, this variable is called TOTWGT since TOTWGT was specified as the weighting variable. This variable will be an estimate of the total population within each subgroup.

MNX

This variable contains the estimated means of the specified analysis variable by subgroup.

MNX_SE

This variable contains the JRR standard errors of the estimated means by subgroup.

PCT

This variable contains the estimated percentages of students in each subgroup for the last classification variable listed. In the example it is the percentage of boys and girls within each country.

PCT_SE

This variable contains the JRR standard errors of the estimated percentages.

The contents of the FINAL file can be printed using the SAS PRINT procedure. The sample SAS program that invokes the JACKGEN macro and a printout of the results are presented in Exhibit 3.4. This program is available on the database DVD in the program file called SAMPLEJACKGEN.SAS. It produces the mean ages for advanced mathematics boys and girls in all countries, although the exhibit shows the results only for the first four countries.

From the first two lines of the results shown in Exhibit 3.4, there are 467 girls in the Armenia advanced mathematics sample representing 1,401 girls in the whole population. The mean age for advanced mathematics girls in Armenia is estimated to be 17.76 with a standard error of 0.02. Girls made up 52.19% of Armenia's advanced mathematics student population. Conversely, Armenia sampled 391 advanced mathematics boys representing 1,283 boys in the whole

population. The estimated mean age for advanced mathematics boys in Armenia is 17.59 with a standard error of 0.02. Boys made up 47.81 percent of Armenia's advanced mathematics student population.

Exhibit 3.4 Sample SAS Program Invoking the SAS Macro JACKGEN and Results

```
LIBNAME TA08 "<datpath>" ;
%INCLUDE "<macpath>JACKGEN.SAS" ;
DATA MSGALLM2 ;
  SET TA08.MSGALLM2 ;
  WHERE NMISS (ITSEX, MSDAGE) = 0 ;
PROC FORMAT LIBRARY = WORK ;
  VALUE COUNTRY
    < list TIMSS Advanced 2008 country formats > ;
  VALUE SEX
    1 = 'GIRL'
    2 = 'BOY' ;
%JACKGEN (TOTWGT, JKZONE, JKREP, 75, IDCNTY ITSEX, MSDAGE, MSGALLM2) ;
PROC PRINT DATA = FINAL NOOBS ;
  VAR IDCNTY ITSEX N TOTWGT MNX MNX_SE PCT PCT_SE ;
  FORMAT IDCNTY COUNTRY. ITSEX SEX. N 6.0 TOTWGT 10.0
    MNX MNX_SE PCT PCT_SE 6.2 ;
```

IDCNTY	ITSEX	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
ARMENIA	GIRL	467	1401	17.76	0.02	52.19	1.96
ARMENIA	BOY	391	1283	17.59	0.02	47.81	1.96
IRAN	GIRL	956	49158	18.11	0.02	44.17	1.59
IRAN	BOY	1469	62140	18.06	0.02	55.83	1.59
ITALY	GIRL	727	40719	18.90	0.02	34.17	2.49
ITALY	BOY	1416	78444	19.04	0.03	65.83	2.49
LEBANON	GIRL	460	1363	17.85	0.04	29.00	1.56
LEBANON	BOY	1152	3339	17.99	0.03	71.00	1.56

Computing Achievement Means and Their Standard Errors (JACKPV)

The JACKPV macro computes percentages and mean achievement scores using plausible values. It makes use of the sampling weights, the jackknifing algorithm to compute sampling variances, and the five plausible values to compute imputation variances. It effectively performs five analyses—one for each plausible value—and aggregates the results to produce accurate estimates of mean achievement and standard errors that incorporate both sampling and imputation errors.

A second sample program demonstrates the use of the JACKPV macro, which computes the percentages of students within specified subgroups and their

mean achievement scores. The SAS macro also computes the appropriate standard errors for those percentages and achievement means.

The JACKPV macro is a self-contained program, located in the program file JACKPV.SAS, and should not be modified. It essentially computes sets of replicate weights using the sampling and weighting variables, aggregates the data by subgroups using the replicate weights, and then computes and stores the desired statistics in a SAS working file called FINAL. The macro aggregates data across all plausible values to obtain the correct results.

The SAS macro JACKPV is included in a SAS program by issuing the following command:

```
%INCLUDE "<macpath>JACKPV.SAS" ;
```

where <macpath> points to the folder where the SAS macro program JACKPV.SAS is located. The macro requires that several parameters be specified as input when it is invoked:

- | | |
|------|--|
| WGT | The sampling weight to be used in the analysis. Generally, TOTWGT should be used. MATWGT should be used when analyzing advanced mathematics teacher data and PHYWGT when analyzing physics teacher data. |
| JKZ | The variable that captures the assignment of cases to sampling zones. The name of this variable in all TIMSS Advanced data files is JKZONE. |
| JKR | The variable that captures whether the case is to be dropped or have its weight doubled for each set of replicate weights. The name of this variable in all TIMSS Advanced data files is JKREP. |
| NJKZ | The number of replicate weights to be generated when computing the JRR standard errors. The value of NJKZ should be set to 75, the maximum possible value across all participating countries. |
| CVAR | The list of variables that are to be used to define the subgroups. The list can consist of one or more variables. We recommend that users always include IDCNTRY as the first classification variable. |

- ROOTPV** The variable root used to identify the set of plausible values for the achievement score of interest. It corresponds to the first seven characters of the plausible values variable name. For example, the root of the advanced mathematics plausible values is MSMMAT0.
- NPV** The number of plausible values that will be used for the analysis. Generally, it is set to five to use all five plausible values for analysis.
- INFILE** The name of the data file that contains the data being analyzed. If the folder is included as part of the file name, the name of the file must be enclosed in quotes. It is important to emphasize that this data file must include only those cases that are of interest in the analysis. If users want to have specific cases excluded from the analysis, for example students with missing data, this should be done prior to invoking the macro.

The JACKPV macro is invoked by a SAS program using the conventional SAS notation for invoking macros. This involves listing the macro name followed by the list of parameters in parenthesis, each separated by a comma. For example, the JACKPV macro invoked with the following statement:

```
%JACKPV (TOTWGT, JKZONE, JKREP, 75, IDCNTY ITSEX, MSMMAT0, 5, MSGALLM2) ;
```

will compute the mean advanced mathematics achievement based on its five plausible values (MSMMAT01 through MSMMAT05) by gender (ITSEX) within each country (IDCNTY) and their standard errors, using the weighting variable TOTWGT. The macro uses all five plausible values to compute these statistics. It also will compute the percentages of boys and girls within each country, and their standard errors. The data will be read from the data file MSGALLM2 and the standard errors will be computed based on 75 sets of replicate weights.

The results of the JACKPV macro are stored in a SAS working file called FINAL, which is stored in the default folder used by SAS. The following variables are contained in this results file:

Classification Variables

All classification variables are kept in the results file. In this example, there are two classification variables: IDCNTY and ITSEX. There is one record in the

results file for each subgroup defined by the categories of the classification variables.

N

This variable contains the number of valid cases for each subgroup defined by the classification variables. In the example, it is the number of boys and girls with valid data in each country's sample.

Weight Variable

The weight variable contains the sum of weights within each subgroup defined by the classification variables. In the example, this variable is called TOTWGT since TOTWGT was specified as the weighting variable. This variable will be an estimate of the total population within each subgroup.

MNPV

This variable contains the estimated mean achievement by subgroup, based on the plausible values.

MNPV_SE

This variable contains the JRR standard errors of the estimated mean achievement by subgroup, based on the plausible values.

PCT

This variable contains the estimated percentages of students in each subgroup for the last classification variable listed. In the example it is the percentage of boys and girls within each country.

PCT_SE

This variable contains the JRR standard errors of the estimated percentages.

The contents of the FINAL file can be printed using the SAS PRINT procedure. The sample SAS program that invokes the JACKPV macro and a printout of the results are shown in Exhibit 3.5. This program is available on the DVD in the program file called SAMPLEJACKPV.SAS. It produces the

mean advanced mathematics achievement for boys and girls in all countries, although Exhibit 3.5 gives the results only for the first four countries.

Exhibit 3.5 Sample SAS Program Invoking the SAS Macro JACKPV and Results

```
LIBNAME TA08 "<datpath>" ;
%INCLUDE "<macpath>JACKPV.SAS" ;
DATA MSGALLM2;
  SET TA08.MSGALLM2 ;
  WHERE NMISS (ITSEX) = 0 ;
PROC FORMAT LIBRARY = WORK ;
  VALUE COUNTRY
    < list TIMSS Advanced 2008 country formats > ;
  VALUE SEX
    1 = 'GIRL'
    2 = 'BOY' ;
%JACKPV (TOTWGT, JKZONE, JKREP, 75, IDCNTY ITSEX, MSMMAT0, 5, MSGALLM2) ;
PROC PRINT DATA = FINAL NOOBS ;
  VAR IDCNTY ITSEX N TOTWGT MNPV MNPV_SE PCT PCT_SE ;
  FORMAT IDCNTY COUNTRY. ITSEX SEX. N 6.0 TOTWGT 10.0
    MNPV MNPV_SE PCT PCT_SE 6.2 ;
```

IDCNTY	ITSEX	N	TOTWGT	MNPV	MNPV_SE	PCT	PCT_SE
ARMENIA	GIRL	467	1401	427.91	4.80	52.19	1.96
ARMENIA	BOY	391	1283	438.12	6.07	47.81	1.96
IRAN	GIRL	956	49158	479.73	6.65	44.17	1.59
IRAN	BOY	1469	62140	510.26	10.06	55.83	1.59
ITALY	GIRL	727	40719	454.01	9.33	34.17	2.49
ITALY	BOY	1416	78444	446.08	8.28	65.83	2.49
LEBANON	GIRL	460	1363	553.74	3.19	29.00	1.56
LEBANON	BOY	1152	3339	541.06	2.66	71.00	1.56

From the first two lines of the results presented in Exhibit 3.5, the mean achievement of advanced mathematics girls in Armenia is estimated to be 427.91 with a standard error of 4.80. The mean achievement of advanced mathematics boys in Armenia is estimated to be 438.12 with a standard error of 6.07.

Computing Regression Coefficients and Their Standard Errors (JACKREG)

The JACKREG macro performs a multiple linear regression between a dependent variable and a set of independent variables. A third sample program demonstrates the use of the JACKREG macro, which computes the regression coefficients and their JRR standard errors. This macro is not appropriate for

regression analyses using achievement scores as the dependent variable. The JACKREGP macro should be used for this purpose.

The JACKREG macro is a self-contained program, located in the program file JACKREG.SAS, and should not be modified. It computes sets of replicate weights using the sampling and weighting variables, performs a linear regression by subgroup and each set of replicate weights, and then computes and stores the desired statistics in a SAS working file called REG.

The SAS macro JACKREG is included in a SAS program by issuing the following command:

```
%INCLUDE "<macpath>JACKREG.SAS" ;
```

where <macpath> points to the specific folder where the SAS macro program JACKREG.SAS is located. The macro requires that several parameters be specified as input when it is invoked:

- | | |
|------|--|
| WGT | The sampling weight to be used in the analysis. Generally, TOTWGT should be used. MATWGT should be used when analyzing advanced mathematics teacher data and PHYWGT when analyzing physics teacher data. |
| JKZ | The variable that captures the assignment of cases to sampling zones. The name of this variable in all TIMSS Advanced data files is JKZONE. |
| JKR | The variable that captures whether the case is to be dropped or have its weight doubled for each set of replicate weights. The name of this variable in all TIMSS Advanced data files is JKREP. |
| NJKZ | The number of replicate weights to be generated when computing the JRR standard errors. The value of NJKZ should be set to 75, the maximum possible value across all participating countries. |
| CVAR | The list of variables that are to be used to define the subgroups. The list can consist of one or more variables. We recommend that users always include IDCNTRY as the first classification variable. |
| XVAR | The list of independent variables used as predictors in the regression model. The independent variables can be either continuous or categorical, such as ITSEX for example. |

- DVAR** The dependent variable to be predicted by the list of independent variables specified in XVAR. Only one variable can be listed and plausible values of achievement scores should not be specified here.
- INFILE** The name of the data file that contains the data being analyzed. If the folder is included as part of the file name, the name of the file must be enclosed in quotes. It is important to emphasize that this data file must include only those cases that are of interest in the analysis. If users want to have specific cases excluded from the analysis, for example students with missing data, this should be done prior to invoking the macro.

The JACKREG macro is invoked by a SAS program using the conventional SAS notation for invoking macros. This involves listing the macro name followed by the list of parameters in parenthesis, each separated by a comma. For example, the JACKREG macro invoked with the following statement:

```
%JACKREG (TOTWGT, JKZONE, JKREP, 75, IDCNTY, REGSEX, MSDAGE, MSGALLM2) ;
```

will perform a linear regression with gender (REGSEX) as a predictor of the advanced mathematics students' age at the time of testing (MSDAGE), using the weighting variable TOTWGT. It will compute the regression coefficients and their standard errors. The data will be read from the data file MSGALLM2 and the standard errors will be computed based on 75 replicate weights.

The results of the JACKREG macro are stored in a SAS working file called REG, which is stored in the default folder used by SAS. The following variables are contained in this results file:

Classification Variables

All classification variables are kept in the results file. In this example, there is a single classification variable IDCNTY. There is one record in the results file for each subgroup defined by the categories of the classification variables.

N

This variable contains the number of valid cases for each subgroup defined by the classification variables. In the example, it is the number of students with valid data in each country's sample.

MULT_RSQ

The squared multiple correlation coefficient (R^2) for the regression model applied in each subgroup.

SS_RES, SS_REG, SS_TOTAL

The residual, regression, and total weighted sums of squares for the regression model applied in each subgroup.

Regression Coefficients and Standard Errors (B## and B##.SE)

The regression coefficients for the intercept and the predictor variables with their respective standard errors. The regression coefficients are numbered sequentially, starting with zero (B00) for the intercept, based on the order the predictor variables are specified in the parameter XVAR.

The contents of the REG file can be printed using the SAS PRINT procedure. The sample SAS program that invokes the JACKREG macro and a printout of the results are displayed in Exhibit 3.6. This program is available on the DVD in the program file called SAMPLEJACKREG.SAS. It performs a linear regression in each country, with the variable REGSEX as a predictor of the advanced mathematics students' age at the time of testing (MSDAGE). The exhibit displays the results for the first four countries.

The regression performed by the sample program uses the independent variable REGSEX, which is a “dummy-coded” version of ITSEX, such that the value zero represents the girls and the value 1 represents the boys. By performing this recoding, the intercept B00 will be the estimated mean age of advanced mathematics girls, whereas the regression coefficient B01 will be the estimated difference in the mean age for boys. This will allow us to determine if the difference in mean ages between girls and boys is statistically significant.

From the first line of the results displayed in Exhibit 3.6, the estimated mean age of advanced mathematics girls in Armenia (B00) is 17.76 years, with a standard error of 0.02. The advanced mathematics boys in Armenia are an estimated 0.17 years younger (B01) than the girls, with a standard error of 0.03.

Exhibit 3.6 Sample SAS Program Invoking the SAS Macro JACKREG and Results

```
LIBNAME TA08 "<datpath>" ;
%INCLUDE "<macpath>JACKREG.SAS" ;

DATA MSGALLM2 ;
  SET TA08.MSGALLM2;

  WHERE NMISS (ITSEX, MSDAGE) = 0 ;

  REGSEX = ITSEX - 1 ;

PROC FORMAT LIBRARY = WORK ;

VALUE COUNTRY
  < list TIMSS Advanced 2008 country formats > ;

%JACKREG (TOTWGT, JKZONE, JKREP, 75, IDCNTY, REGSEX, MSDAGE, MSGALLM2) ;

PROC PRINT DATA = REG NOOBS ;
  VAR IDCNTY N MULT_RSQ SS_TOTAL SS_REG B00 B00_SE B01 B01_SE ;
  FORMAT IDCNTY COUNTRY. N 6.0 MULT_RSQ 5.3 SS_TOTAL SS_REG 10.0
    B00 B00_SE B01 B01_SE 6.2 ;
```

IDCNTY	N	MULT_RSQ	SS_TOTAL	SS_REG	B00	B00_SE	B01	B01_SE
ARMENIA	858	0.032	606	19	17.76	0.02	-0.17	0.03
IRAN	2425	0.002	20946	50	18.11	0.02	-0.04	0.03
ITALY	2143	0.014	41329	583	18.90	0.02	0.15	0.04
LEBANON	1612	0.006	2924	19	17.85	0.04	0.14	0.04

Computing Regression Coefficients and Their Standard Errors with Achievement Scores (JACKREGP)

The JACKREGP macro is used to perform a multiple linear regression between a set of plausible values as the dependent variable and a set of independent variables. It computes the regression coefficients and their JRR standard errors, making use of the sampling weights, the jackknifing algorithm to compute sampling variances, and the five plausible values to compute imputation variances. It effectively performs five regression analyses—one for each plausible value—and aggregates the results to produce accurate estimates of the regression coefficients and standard errors that incorporate both sampling and imputation errors. We present a fourth sample program to demonstrate the use of the JACKREGP macro.

The JACKREGP macro is a self-contained program, located in the program file JACKREGP.SAS, and should not be modified. It computes sets of replicate weights using the sampling and weighting variables, performs a multiple linear regression by subgroups and each set of replicate weights, and then computes and stores the desired statistics in a SAS working file called REG.

The SAS macro JACKREGP is included in a SAS program by issuing the following command:

```
%INCLUDE "<macpath>JACKREGP.SAS" ;
```

where <macpath> points to the specific folder where the SAS macro program JACKREGP.SAS is located. The macro requires that several parameters be specified as input when it is invoked:

- | | |
|--------|--|
| WGT | The sampling weight to be used in the analysis. Generally, TOTWGT should be used. MATWGT should be used when analyzing advanced mathematics teacher data and PHYWGT when analyzing physics teacher data. |
| JKZ | The variable that captures the assignment of cases to sampling zones. The name of this variable in all TIMSS Advanced data files is JKZONE. |
| JKR | The variable that captures whether the case is to be dropped or have its weight doubled for each set of replicate weights. The name of this variable in all TIMSS Advanced data files is JKREP. |
| NJKZ | The number of replicate weights to be generated when computing the JRR standard errors. The value of NJKZ should be set to 75, the maximum possible value across all participating countries. |
| CVAR | The list of variables that are to be used to define the subgroups. The list can consist of one or more variables. We recommend that users always include IDCNTRY as the first classification variable. |
| XVAR | The list of independent variables used as predictors in the regression model. The independent variables can be either continuous or categorical, such as ITSEX for example. |
| ROOTPV | The variable root used to identify the set of plausible values for the achievement score of interest. It corresponds to the first seven characters of the plausible values variable name. For example, the root of the advanced mathematics plausible values is MSMMAT0. |

INFILE The name of the data file that contains the data being analyzed. If the folder is included as part of the file name, the name of the file must be enclosed in quotes. It is important to emphasize that this data file must include only those cases that are of interest in the analysis. If users want to have specific cases excluded from the analysis, for example students with missing data, this should be done prior to invoking the macro.

The JACKREGP macro is invoked by a SAS program using the conventional SAS notation for invoking macros. This involves listing the macro name followed by the list of parameters in parenthesis, each separated by a comma. For example, the JACKREGP macro invoked with the following statement:

```
%JACKREGP (TOTWGT, JKZONE, JKREP, 75, IDCNTY, REGSEX, MSMMAT0, 5, MSGALLM2) ;
```

will perform a linear regression with gender (REGSEX) as a predictor of advanced mathematics achievement based on its five plausible values (MSMMAT01 through MSMMAT05), using the weighting variable TOTWGT. It will compute the regression coefficients and their standard errors. The data will be read from the data file MSGALLM2 and the standard errors will be computed based on 75 replicate weights.

The results of the JACKREGP macro are stored in a SAS working file called REG, which is stored in the default folder used by SAS. The following variables are contained in this results file:

Classification Variables

All classification variables are kept in the results file. In this example, there is a single classification variable IDCNTY. There is one record in the results file for each subgroup defined by the categories of the classification variables.

N

This variable contains the number of valid cases for each subgroup defined by the classification variables. In the example, it is the number of students with valid data in each country's sample.

MULT_RSQ

The squared multiple correlation coefficient (R^2) for the regression model applied in each subgroup.

SS_RES, SS_REG, SS_TOTAL

The residual, regression, and total weighted sums of squares for the regression model applied in each subgroup.

Regression Coefficients and Standard Errors (B## and B##.SE)

The regression coefficients for the predictor variables and the intercept with their respective standard errors. The regression coefficients are numbered sequentially, starting with zero (B00) for the intercept, based on the order the predictor variables are specified in the parameter XVAR.

The contents of the REG file can be printed using the SAS PRINT procedure. The sample SAS program invoking the JACKREGP macro and a printout of the results are presented in Exhibit 3.7. This program is available on the DVD in the program file called SAMPLEJACKREGP.SAS. It performs a linear regression in each country, with the variable REGSEX as a predictor of advanced mathematics achievement. The exhibit displays the results for the first four countries.

The regression performed by our sample program uses the variable REGSEX that was defined in our previous example. By using REGSEX, the intercept B00 will be the estimated mean achievement of advanced mathematics girls, whereas the regression coefficient B01 will be the estimated difference in the mean achievement of advanced mathematics boys. This will allow us to determine if advanced mathematics achievement is significantly different between girls and boys.

From the first line of the results shown in Exhibit 3.7, the estimated mean advanced mathematics achievement of girls in Armenia (B00) is 427.91 with a standard error of 4.80. Note that these are the same results obtained from the JACKPV sample program (Exhibit 3.5). The boys have an estimated mean advanced mathematics achievement 10.22 points (B01) higher than girls, with a standard error of 8.22.

Exhibit 3.7 Sample SAS Program Invoking the SAS Macro JACKREGP and Results

```
LIBNAME TA08 "<datpath>" ;
%INCLUDE "<macpath>JACKREGP.SAS" ;

DATA MSGALLM2 ;
  SET TA08.MSGALLM2;

  WHERE NMISS (ITSEX) = 0 ;

  REGSEX = ITSEX - 1 ;

PROC FORMAT LIBRARY = WORK ;

  VALUE COUNTRY
    < list TIMSS Advanced 2008 country formats > ;

%JACKREGP (TOTWGT, JKZONE, JKREP, 75, IDCNTY, REGSEX, MSMMAT0, 5, MSGALLM2) ;

PROC PRINT DATA = REG NOOBS ;
  VAR IDCNTY N MULT_RSQ SS_TOTAL SS_REG B00 B00_SE B01 B01_SE ;
  FORMAT IDCNTY COUNTRY. N 6.0 MULT_RSQ 5.3 SS_TOTAL SS_REG 10.0
    B00 B00_SE B01 B01_SE 6.2 ;
```

IDCNTY	N	MULT_RSQ	SS_TOTAL	SS_REG	B00	B00_SE	B01	B01_SE
ARMENIA	858	0.003	24406781	76142	427.91	4.80	10.22	8.22
IRAN	2425	0.024	1085520382	25648735	479.73	6.65	30.53	12.07
ITALY	2143	0.002	1085936938	1770334	454.01	9.33	-7.93	10.36
LEBANON	1612	0.009	17254056	157640	553.74	3.19	-12.68	3.75

3.7 TIMSS Advanced Analyses with Student-Level Variables

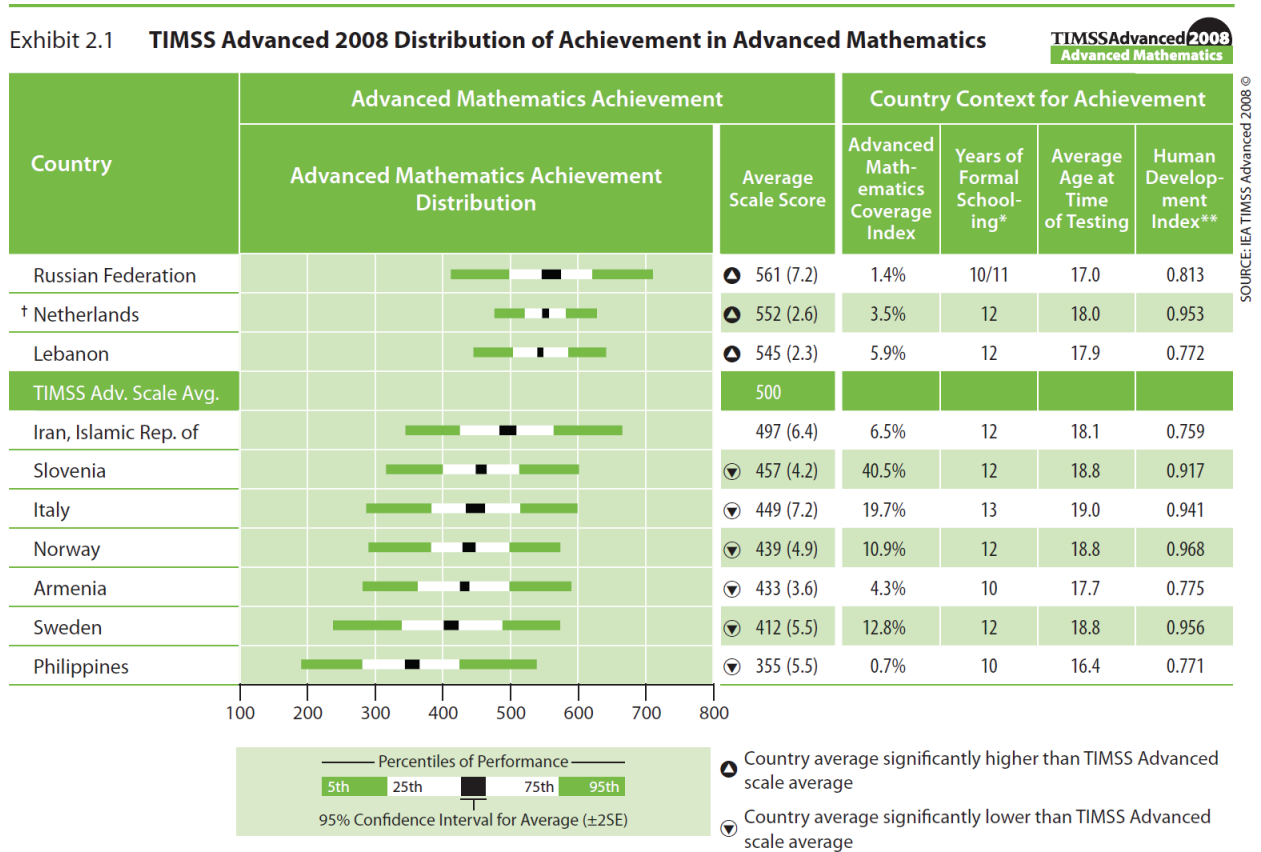
Many analyses of the TIMSS Advanced 2008 data can be undertaken using only student-level data. Examples in the previous sections illustrated the use and functionality of the SAS macros. This section presents examples of actual analyses that produced selected exhibits in the *TIMSS Advanced 2008 International Report* (Mullis, Martin, Robitaille, & Foy, 2009), using SAS programs provided on the DVD.

The first example computes means for a straightforward continuous variable, whereas the second example computes means of achievement scores. Both examples use the sampling weights and implement the jackknife repeated replication method to compute appropriate sampling errors. The second example, which uses achievement plausible values, effectively performs the computations five times—once for each plausible value—and aggregates the results to produce accurate estimates of mean achievement and standard errors that incorporate both sampling and imputation errors.

Student-Level Analysis

In this first example, we wish to replicate the analysis of advanced mathematics students’ reported age at the time of testing. The results, presented in Exhibit 2.1 of the *TIMSS Advanced 2008 International Report*, are reproduced here in Exhibit 3.8. This example will focus on the results presented in the fourth data column—average age at time of testing.

Exhibit 3.8 Exhibit of Example Student-Level Analysis Taken from the TIMSS Advanced 2008 International Report (Exhibit 2.1)



* Represents years of schooling counting from the first year of primary or basic education (first year of ISCED Level 1).
 ** Taken from United Nations Development Programme's *Human Development Report 2007/2008*, p.229-232.

† Met guidelines for sample participation rates only after replacement schools were included (see Appendix A).
 () Standard errors appear in parentheses.

We need to undertake a number of steps to replicate the results in this exhibit. After reviewing the advanced mathematics student background data codebook (the codebooks are described in Section 4.4 of Chapter 4), we identify the

student background variable MSDAGE as the variable that reports the age of students at the time of testing.

We then proceed to read from the student background data files our variable of interest (MSDAGE), the student sampling weight (TOTWGT), the variables that contain the jackknife replication information (JKZONE and JKREP), and the variable containing the country identification code (IDCNTRY). In this analysis, we will use the data for all available countries. We used the JOIN program, described earlier in this chapter, to join the student background data files for all countries into a single file called MSGALLM2.

The SAS program used to perform this first example is presented in Exhibit 3.9 and is included on the DVD under the name EXAMPLE1.SAS. The results obtained from this program are displayed in Exhibit 3.10, although only the results of the first four countries, sorted alphabetically, are shown for the sake of conciseness. Note that one of the steps in this program is to select only those students who have non-missing data in our variables of interest MSDAGE.

Exhibit 3.9 Example SAS Program to Perform Student-Level Analysis (EXAMPLE1.SAS)

```
LIBNAME TA08 "<datpath>" ;
%INCLUDE "<macpath>JACKGEN.SAS" ;
DATA MSGALLM2 ;
  SET TA08.MSGALLM2;
  WHERE NMISS (MSDAGE) = 0 ;
PROC FORMAT LIBRARY = WORK ;
  VALUE COUNTRY
    < list TIMSS Advanced 2008 country formats > ;
%JACKGEN (TOTWGT, JKZONE, JKREP, 75, IDCNTRY, MSDAGE, MSGALLM2) ;
PROC PRINT DATA = FINAL NOOBS ;
  VAR IDCNTRY N TOTWGT MNX MNX_SE PCT PCT_SE ;
  FORMAT IDCNTRY COUNTRY. N 6.0 TOTWGT 10.0
    MNX MNX_SE PCT PCT_SE 6.2 ;
```

Users should perform the following steps to carry out basic student-level analyses with the student background data files:

- 1) Identify the variables of interest in the student background data files and note any specific national adaptations to the variables.
- 2) Retrieve the relevant variables from the student background data files, including classification variables, analysis variables, identification

variables, sampling and weighting variables, and any other variables used in the selection of cases.

- 3) Perform any necessary variable transformations or recodes.
- 4) Use the macros JACKGEN and JACKREG with the appropriate parameters.
- 5) Specify the location of the data files in the parameter <datpath> and of the macros in the parameter <macpath>.
- 6) Print the results file.

In Exhibit 3.10, each country's mean value for the MSDAGE variable is reported for all sampled students. The countries are identified in the first column and the second column reports the number of students with valid data. The third column reports the sum of weights of the sampled students, followed by the mean for MSDAGE and its standard error. The last two columns report the weighted percentage of students in the population and its standard error. For this example, the weighted percentages are of little use as they are the proportion each country represents among all participating countries. From the first line, Armenia has valid data for 858 students and these sampled students represent a population of 2,684 students. Advanced mathematics students in Armenia were, on average, 17.68 years old at the time the TIMSS Advanced 2008 assessment took place, with a standard error of 0.02.

Exhibit 3.10 Output for Example Student-Level Analysis (EXAMPLE 1)

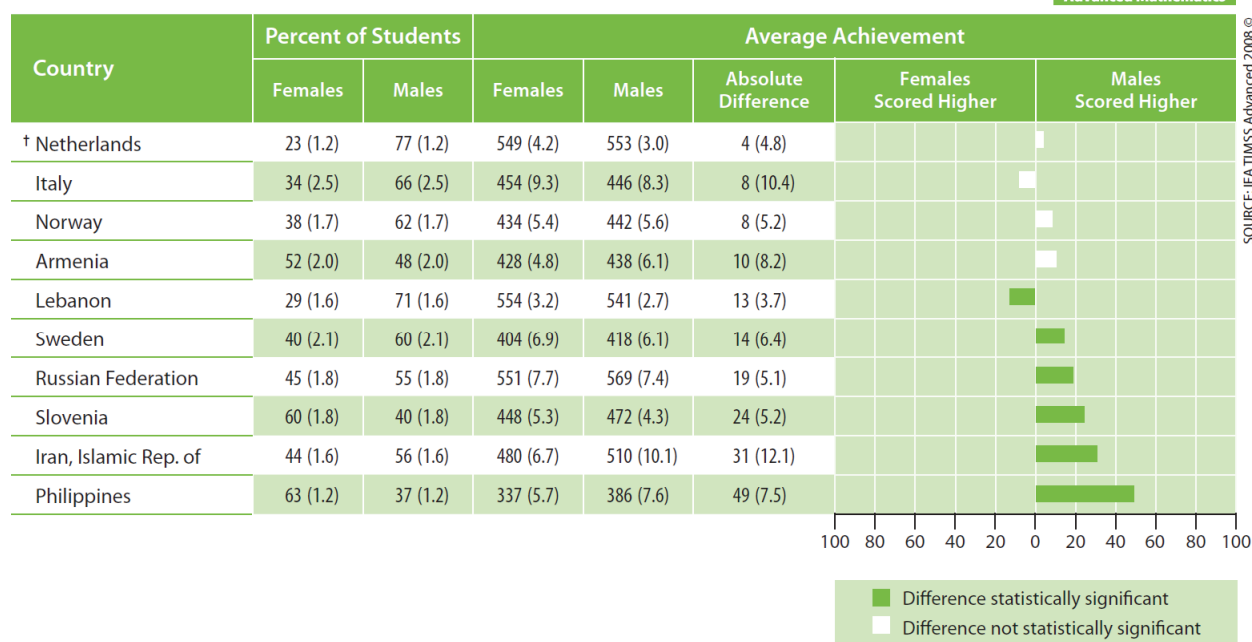
IDCNTRY	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
ARMENIA	858	2684	17.68	0.02	0.84	0.03
IRAN	2425	111298	18.08	0.02	34.76	1.09
ITALY	2143	119162	18.99	0.02	37.21	1.63
LEBANON	1612	4702	17.95	0.03	1.47	0.05

Student-Level Analysis with Achievement Scores

Our second example replicates another set of results presented in the *TIMSS Advanced 2008 International Report*. We will investigate the relationship between advanced mathematics students' gender and advanced mathematics achievement. These results, presented in Exhibit 2.4 of the *TIMSS Advanced 2008 International Report*, are repeated here in Exhibit 3.11. Since the results in this exhibit are based on plausible values, we will use the macro JACKPV.

Exhibit 3.11 Exhibit of Example Student-Level Analysis with Achievement Scores Taken from the TIMSS Advanced 2008 International Report (Exhibit 2.4)

Exhibit 2.4 TIMSS Advanced 2008 Average Achievement in Advanced Mathematics by Gender



† Met guidelines for sample participation rates only after replacement schools were included (see Appendix A).

() Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

After reviewing the appropriate codebook, we observe that the variable ITSEX in the student background data files contains information on the gender of students. We then proceed to read from the student background data files our variable of interest (ITSEX), the five plausible values of advanced mathematics achievement (MSMMAT01 through MSMMAT05), the student sampling weight (TOTWGT), the variables that contain the jackknifing information (JKZONE and JKREP), and the country identification variable (IDCNTRY). Again, we will use the student background data of all available countries contained in the file MSGALLM2.

The SAS program that implements this second example is presented in Exhibit 3.12 and is included on the DVD under the name EXAMPLE2.SAS. Note that one of the steps in this program is to select only those students who have non-missing data in our variable of interest ITSEX. The results obtained from this program are shown in Exhibit 3.13. For the sake of conciseness, only the results of the first four countries, sorted alphabetically, are shown.

Exhibit 3.12 Example SAS Program to Perform Student-Level Analysis with Achievement Scores (EXAMPLE2.SAS)

```
LIBNAME TA08 "<datpath>" ;
%INCLUDE "<macpath>JACKPV.SAS" ;

DATA MSGALLM2 ;
  SET TA08.MSGALLM2;

  WHERE NMISS (ITSEX) = 0 ;

PROC FORMAT LIBRARY = WORK ;

VALUE COUNTRY
  < list TIMSS Advanced 2008 country formats > ;

VALUE SEX
  1 = 'GIRL'
  2 = 'BOY' ;

%JACKPV (TOTWGT, JKZONE, JKREP, 75, IDCNTY ITSEX, MSMMAT0, 5, MSGALLM2) ;

PROC PRINT DATA = FINAL NOOBS ;
  VAR IDCNTY ITSEX N TOTWGT MNPV MNPV_SE PCT PCT_SE ;
  FORMAT IDCNTY COUNTRY. ITSEX SEX. N 6.0 TOTWGT 10.0
  MNPV MNPV_SE PCT PCT_SE 6.2 ;
```

To carry out student-level analyses with the student background data files and achievement scores, users should perform the following steps:

- 1) Identify the variables of interest in the student background data files and note any specific national adaptations to the variables.
- 2) Retrieve the relevant variables from the student background data files, including the plausible values of achievement, classification variables, identification variables, sampling and weighting variables, and any other variables used in the selection of cases.
- 3) Perform any necessary variable transformations or recodes.
- 4) Use the macros JACKPV and JACKREGP with the appropriate parameters.
- 5) Specify the location of the data files in the parameter <datpath> and of the macros in the parameter <macpath>.
- 6) Print the results file.

In Exhibit 3.13, each country's results are displayed on two lines, one for each value of the variable ITSEX. The countries are identified in the first column and the second column describes the category of ITSEX being reported. The third column reports the number of students with valid data and the fourth the

sum of weights of the sampled students. The next two columns report the estimated mean advanced mathematics achievement and its standard error, followed by the percentage of students in each category and its standard error. From the first two lines, the mean achievement for advanced mathematics girls in Armenia is 427.91 (standard error of 4.80) and is 438.12 (standard error of 6.07) for boys. An estimated 52.19% of advanced mathematics students in Armenia are girls, and 47.81% are boys.

Exhibit 3.13 Output for Example Student-Level Analysis with Achievement Scores (EXAMPLE 2)

IDCNTRY	ITSEX	N	TOTWGT	MNPV	MNPV_SE	PCT	PCT_SE
ARMENIA	GIRL	467	1401	427.91	4.80	52.19	1.96
ARMENIA	BOY	391	1283	438.12	6.07	47.81	1.96
IRAN	GIRL	956	49158	479.73	6.65	44.17	1.59
IRAN	BOY	1469	62140	510.26	10.06	55.83	1.59
ITALY	GIRL	727	40719	454.01	9.33	34.17	2.49
ITALY	BOY	1416	78444	446.08	8.28	65.83	2.49
LEBANON	GIRL	460	1363	553.74	3.19	29.00	1.56
LEBANON	BOY	1152	3339	541.06	2.66	71.00	1.56

3.8 TIMSS Analyses with Teacher-Level Variables

The teachers in the TIMSS Advanced 2008 international database do not constitute representative samples of teachers in the participating countries. Rather, they are the teachers of nationally representative samples of students. Therefore, analyses with teacher data should be made with students as the units of analysis and reported in terms of students who are taught by teachers with a particular attribute.

When analyzing teacher data, it is first necessary to link the students to their respective teachers. The student-teacher linkage data files (MST/PST) were created for this purpose. Student achievement scores (plausible values), jackknife replication information, and teacher weighting variables—MATWGT for advanced mathematics teachers and PHYWGT for physics teachers—appropriate for conducting analyses with teacher variables, are found in the student-teacher linkage data files in order to simplify the merging process for analyses that link teacher background variables to student achievement. For such analyses, it is only necessary to merge the teacher background data files (MTG/PTG) with the student-teacher linkage data files. For analyses linking teacher variables to student background variables, it is

also necessary to merge the student background data files (MSG/PSG) with the teacher background data files after having been combined with the student-teacher linkage data files.

As our example of an analysis using teacher background data, we will investigate the age of the TIMSS Advanced 2008 advanced mathematics teachers. The results of such an analysis are presented in Exhibit 5.1 of the *TIMSS Advanced 2008 International Report*, which is reproduced here in Exhibit 3.14. Although the results in this exhibit do not include any achievement, we will use the macro JACKPV to estimate the percentages we want.

Exhibit 3.14 Exhibit of Example Teacher-Level Analysis Taken from the TIMSS Advanced 2008 International Report (Exhibit 5.1)

Exhibit 5.1 **Advanced Mathematics Teachers' Gender, Age, and Number of Years Teaching**

Country	Percent of Students by Teacher Characteristics						Average Number of Years Teaching	
	Gender		Age				Teaching Altogether	Teaching Mathematics at the Advanced Level
	Female	Male	29 Years or Under	30–39 Years	40–49 Years	50 Years or Older		
Armenia	76 (4.6)	24 (4.6)	0 (0.0)	10 (2.4)	44 (5.4)	46 (5.4)	25 (0.9)	13 (1.2)
Iran, Islamic Rep. of	33 (2.5)	67 (2.5)	6 (2.3)	49 (3.8)	31 (3.4)	13 (2.5)	17 (0.6)	9 (0.3)
Italy	54 (5.4)	46 (5.4)	2 (1.4)	10 (3.0)	43 (4.7)	45 (4.6)	22 (0.9)	12 (0.8)
Lebanon	10 (1.5)	90 (1.5)	3 (0.9)	13 (1.9)	20 (2.0)	65 (2.4)	27 (0.5)	25 (0.5)
Netherlands	14 (3.5)	86 (3.5)	3 (2.0)	10 (2.1)	20 (4.8)	67 (5.2)	27 (1.1)	17 (1.1)
Norway	18 (3.9)	82 (3.9)	1 (0.7)	8 (2.3)	19 (4.2)	73 (4.3)	27 (0.9)	26 (0.9)
Philippines	63 (4.4)	37 (4.4)	25 (4.2)	36 (4.4)	25 (4.5)	14 (3.8)	14 (1.0)	5 (0.5)
Russian Federation	90 (2.7)	10 (2.7)	1 (0.6)	13 (3.0)	36 (5.2)	51 (5.2)	26 (0.8)	12 (0.8)
Slovenia	76 (5.1)	24 (5.1)	4 (1.9)	34 (5.8)	32 (5.9)	30 (5.5)	18 (1.1)	14 (0.7)
Sweden	19 (3.8)	81 (3.8)	2 (1.1)	18 (4.0)	22 (3.3)	58 (4.0)	22 (1.0)	9 (0.7)

Data provided by teachers.

(1) Standard errors appear in parentheses.

An "s" indicates data are available for at least 50% but less than 70% of the students.

SOURCE: IEA TIMSS Advanced 2008 ©

Conducting analyses with teacher data requires a few extra steps. As before, we first proceed to identify the variables relevant to the analysis in the appropriate files, and review the documentation for any specific national adaptations to the questions of interest (Supplements 1 and 2). Since we are using a teacher-level variable, we need to use the teacher background data files and the student-teacher linkage data files to find the variables. From the teacher background

data files, we need the variable that contains the information on the advanced mathematics teachers' age (MT2GAGE), the variable that identifies the country (IDCNTRY), and the two teacher identification variables (IDTEACH and IDLINK) that will allow us to link the teacher data to the student data.

We then proceed to retrieve the necessary information from the advanced mathematics student-teacher linkage data files. From these files, we need the country identification (IDCNTRY) and the two teacher identification variables (IDTEACH and IDLINK) needed to link the teacher data to the student data. We also need the jackknife replication variables (JKZONE and JKREP), the advanced mathematics teacher weighting variable (MATWGT), and the advanced mathematics achievement plausible values (MSMMAT01 through MSMMAT05). Although we are only interested in estimating percentages, the advanced mathematics achievement plausible values are required input for the JACKPV macro. This could be of analytical interest, providing some insight into the relationship between advanced mathematics students' achievement and the age of their teachers.

The two file types are merged and the resulting merged file is then input to the JACKPV macro. The merging is done using the combination of identification variables IDCNTRY, IDTEACH, and IDLINK. The combination of values for these three variables is unique within the teacher background data files, but is repeated in the student-teacher linkage data files as many times as needed to link a teacher to all students in a classroom. After the files are merged, the JACKPV macro is invoked and the results can be printed.

For this analysis, we will again use the data for all available countries, making use of an aggregated teacher background data file, MTGALLM2, and an aggregated student-teacher linkage data file, MSTALLM2. These aggregated files can be created with the JOIN macro.

The SAS program that executes this third example is presented in Exhibit 3.15 and is included on the DVD under the name EXAMPLE3.SAS. The results obtained from this program are displayed in Exhibit 3.16, edited to show only the first four countries, alphabetically, for the sake of conciseness. Note that one of the steps in this program is to select only those students who have non-missing data in our variable of interest MT2GAGE. A second step consists of combining response categories 1 and 2 and response categories 5 and 6 of the variable MT2GAGE in order to match the results presented in Exhibit 3.14,

where teachers are categorized into four groups: 29 years or under, 30 to 39 years, 40 to 49 years, and 50 years or older.

Exhibit 3.15 Example SAS Program to Analyze Teacher Variables (EXAMPLE3.SAS)

```
LIBNAME TA08 "<datpath>" ;

%INCLUDE "<macpath>JACKPV.SAS" ;

PROC SORT DATA = TA08.MTMALLM2 OUT = MTMALLM2;
  BY IDCNTRY IDTEACH IDLINK ;

PROC SORT DATA = TA08.MSTALLM2 OUT = MSTALLM2;
  BY IDCNTRY IDTEACH IDLINK ;

DATA MERGED ;
  MERGE MTMALLM2 (IN = INMTM)
        MSTALLM2 (IN = INMST) ;
  BY IDCNTRY IDTEACH IDLINK ;
  IF INMTM AND INMST ;

DATA MERGED ;
  SET MERGED ;

  IF NMISS (MT2GAGE) = 0 ;

  SELECT (MT2GAGE) ;
  WHEN (1,2) NEWAGE = 1 ; * 29 YEARS OR UNDER ;
  WHEN (3)   NEWAGE = 2 ; * 30 TO 39 YEARS   ;
  WHEN (4)   NEWAGE = 3 ; * 40 TO 49 YEARS   ;
  WHEN (5,6) NEWAGE = 4 ; * 50 YEARS OR OLDER ;
  OTHERWISE  NEWAGE = . ;
END ;

PROC FORMAT LIBRARY = WORK ;

  VALUE COUNTRY
  < list TIMSS Advanced 2008 country formats > ;

  VALUE AGE
  1 = '29 YEARS OR UNDER'
  2 = '30 TO 39 YEARS   '
  3 = '40 TO 49 YEARS   '
  4 = '50 YEARS OR OLDER' ;

%JACKPV (MATWGT, JKZONE, JKREP, 75, IDCNTRY NEWAGE, MSMMAT0, 5, MERGED) ;

PROC PRINT DATA = FINAL NOOBS ;
  VAR IDCNTRY NEWAGE N MATWGT MNPV MNPV_SE PCT PCT_SE ;
  FORMAT IDCNTRY COUNTRY. NEWAGE AGE. N 6.0 MATWGT 10.0
  MNPV MNPV_SE PCT PCT_SE 6.2 ;
```

To perform analyses with the teacher background data files, users should execute the following steps:

- 1) Identify the variables of interest in the teacher background data files and note any specific national adaptations to the variables.
- 2) Retrieve the relevant variables from the advanced mathematics teacher background data files, including analysis variables, classification variables, identification variables (IDCNTRY, IDTEACH, and IDLINK), and any other variables used in the selection of cases.

- 3) Retrieve the relevant variables from the student-teacher linkage data files, including plausible values of achievement, classification variables, identification variables (IDCNTRY, IDSTUD, IDTEACH, and IDLINK), sampling (JKZONE and JKREP) and weighting (MATWGT or PHYWGT) variables, and any other variables used in the selection of cases.
- 4) Merge the advanced mathematics teacher background data files with the student-teacher linkage data files using the variables IDCNTRY, IDTEACH and IDLINK.
- 5) If student background variables also are needed, merge the student background data files with the merged student-teacher data files from the previous step using the variables IDCNTRY and IDSTUD.
- 6) Perform any necessary variable transformations or recodes.
- 7) Use the macros JACKGEN and JACKREG, or JACKPV and JACKREGP if plausible values are involved, with the appropriate arguments and parameters.
- 8) Specify the location of the data files in the parameter <datpath> and of the macros in the parameter <macpath>.
- 9) Print the results file.

Exhibit 3.16 Output for Example Teacher Variable Analysis (EXAMPLE 3)

IDCNTRY	NEWAGE	N	MATWGT	MNPV	MNPV_SE	PCT	PCT_SE
ARMENIA	30 TO 39 YEARS	92	273	443.09	15.42	10.30	2.43
ARMENIA	40 TO 49 YEARS	355	1160	424.78	13.95	43.74	5.39
ARMENIA	50 YEARS OR OLDER	401	1218	436.11	13.25	45.96	5.41
IRAN	29 YEARS OR UNDER	87	7047	475.35	29.33	6.33	2.35
IRAN	30 TO 39 YEARS	1163	54802	487.06	8.91	49.24	3.79
IRAN	40 TO 49 YEARS	880	34893	513.88	10.27	31.35	3.44
IRAN	50 YEARS OR OLDER	295	14556	502.76	20.16	13.08	2.54
ITALY	29 YEARS OR UNDER	35	2101	445.08	150.54	1.77	1.38
ITALY	30 TO 39 YEARS	189	12108	423.97	19.60	10.19	3.04
ITALY	40 TO 49 YEARS	920	50623	444.07	10.74	42.62	4.67
ITALY	50 YEARS OR OLDER	991	53950	458.05	12.70	45.42	4.65
LEBANON	29 YEARS OR UNDER	37	118	575.76	29.75	2.55	0.89
LEBANON	30 TO 39 YEARS	247	614	529.85	4.78	13.27	1.89
LEBANON	40 TO 49 YEARS	362	908	549.29	5.50	19.62	1.95
LEBANON	50 YEARS OR OLDER	1173	2990	545.32	2.93	64.57	2.44

In Exhibit 3.16, each country's results are shown on four lines, one for each value of the recoded NEWAGE variable. The results are presented much in the

same manner as in previous examples, where the countries are identified in the first column and the second column describes the category of NEWAGE being reported. From the first three lines of results, 10.30% of students in Armenia were taught by teachers 30 to 39 years, 43.74% by teachers 40 to 49 years, and 45.96% by teachers 50 years or older. The appropriate standard errors also are presented in Exhibit 2.20. Note that Armenia does not display a row for teachers 29 years or under because there were none in their national sample.

3.9 TIMSS Analyses with School-Level Variables

Because TIMSS Advanced 2008 has representative samples of schools, it is possible to compute reasonable statistics with schools as units of analysis. However, the school samples were designed to optimize the student samples and the student-level estimates. For this reason, it is preferable to analyze school-level variables as attributes of students, rather than as units of analysis in their own right. Therefore, analyzing school data should be done by linking the students to their schools.

Our example of an analysis using school background data will compute the percentages of advanced mathematics students who attend schools with a high, medium, and low index of principals' perception of school climate. We can use the variable MCDGPPSC for this purpose. We also will calculate the mean advanced mathematics achievement at each level of principals' perception of school climate. The results of this analysis are presented in Exhibit 6.5 of the *TIMSS Advanced 2008 International Report*, which is reproduced here in Exhibit 3.17.

The variable MCDGPPSC in the school background data files contains information on the index of principals' perception of school climate. As MCDGPPSC is a derived variable, we can refer to Supplement 3 to see how it was created. We also should review the documentation on national adaptations to its component variables in Supplement 2.

Since we are using a school-level variable, we need to use the school background data files and the student background data files to find the variables of interest. From the school background data files, we need the variable that contains the information on the index of principals' perception of school climate (MCDGPPSC) and the identification variables IDCNTRY and IDSCHOOL that will allow us to link the school data to the student data.

Next, we retrieve the variables of interest from the student background data files. We need the country and school identification variables (IDCNTRY and IDSCHOOL) necessary to merge the school data to the student data. We also need the jackknife replication variables (JKZONE and JKREP), the student weighting variable (TOTWGT), and the advanced mathematics achievement plausible values (MSMMAT01 through MSMMAT05).

Exhibit 3.17 Exhibit of Example School-Level Analysis Taken from the TIMSS Advanced 2008 International Report (Exhibit 6.5)

Exhibit 6.5 Index of Principals' Perceptions of School Climate (PPSC)

TIMSS Advanced 2008
Advanced Mathematics

Country	High PPSC		Medium PPSC		Low PPSC	
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement
Philippines	53 (5.3)	354 (7.8)	43 (4.5)	365 (9.8)	4 (2.2)	303 (28.8)
Slovenia	25 (3.6)	506 (6.1)	68 (4.4)	447 (6.0)	6 (2.6)	405 (8.3)
Iran, Islamic Rep. of	25 (4.0)	528 (14.0)	59 (5.2)	496 (9.1)	16 (3.7)	449 (6.4)
Lebanon	25 (2.0)	558 (4.6)	59 (2.1)	543 (3.1)	16 (1.3)	525 (3.7)
Sweden	18 (4.9)	438 (11.0)	73 (4.8)	411 (6.2)	10 (3.3)	381 (9.0)
Russian Federation	13 (3.4)	605 (18.6)	81 (3.8)	559 (6.7)	6 (2.0)	494 (24.7)
Norway	7 (2.7)	441 (11.2)	90 (3.3)	440 (5.1)	4 (1.9)	403 (58.4)
Italy	3 (1.8)	481 (45.6)	60 (5.1)	458 (8.7)	37 (5.0)	431 (12.4)
Armenia	2 (0.1)	~ ~	83 (0.4)	436 (4.1)	15 (0.4)	420 (5.3)
Netherlands	1 (0.7)	~ ~	68 (5.5)	555 (3.7)	31 (5.5)	547 (3.7)

SOURCE: IEA TIMSS Advanced 2008 ©

Based on principals' responses to the following aspects of school climate in their schools: teachers' job satisfaction, teachers' opportunities for professional development, teachers' understanding of the school's curricular goals, teachers' degree of success in implementing the school's curriculum, teachers' expectations for student achievement, parental support for student achievement, parental involvement in school activities, students' regard for school property, and students' desire to do well in school. Average is computed across the nine statements based on a 5-point scale: 1 = Very High, 2 = High, 3 = Medium, 4 = Low, 5 = Very Low. High level indicates students whose principals' perception of their school climate was very positive (average is less than or equal to 2). Medium level indicates

students whose principals' perception of their school climate was moderately positive (average is greater than 2 and less than 3). Low level indicates students whose principals' perception of their school climate was not so positive (average is greater than or equal to 3).

() Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

A tilde (~) indicates insufficient data to report achievement.

An "r" indicates data are available for at least 70% but less than 85% of the students.

We then proceed to merge the school data with the student data using the variables IDCNTRY and IDSCHOOL and use the macro JACKPV to obtain the percentages of advanced mathematics students and their mean achievement scores within each category of the variable MCDGPPSC for each country. For this analysis, we will use the data for all available countries, making use of an aggregated school file MCGALLM2 and an aggregated student file MSGALLM2. These aggregated files can be created with the JOIN macro.

The SAS program that implements this fourth example is presented in Exhibit 3.18 and is included on the DVD under the name EXAMPLE4.SAS. The results of this program are displayed in Exhibit 3.19, edited to show only the first four countries, alphabetically, for the sake of brevity. Note that one of the steps in this program is to select only those students who have non-missing data in our variable of interest MCDGPPSC.

Exhibit 3.18 Example SAS Program for School Variable Analysis (EXAMPLE4.SAS)

```
LIBNAME TA08 "<datpath>" ;
%INCLUDE "<macpath>JACKPV.SAS" ;
PROC SORT DATA = TA08.MCGALLM2 OUT = MCGALLM2 ;
  BY IDCNTRY IDSCHOOL ;
PROC SORT DATA = TA08.MSGALLM2 OUT = MSGALLM2 ;
  BY IDCNTRY IDSCHOOL ;
DATA MERGED ;
  MERGE MCGALLM2 (IN = INMCG)
        MSGALLM2 (IN = INMSG) ;
  BY IDCNTRY IDSCHOOL ;
  IF INMCG AND INMSG ;
DATA MERGED ;
  SET MERGED ;
  IF NMISS (MCDGPPSC) = 0 ;
PROC FORMAT LIBRARY = WORK ;
  VALUE COUNTRY
    < list TIMSS Advanced 2008 country formats >
  VALUE PPSC
    1 = 'HIGH'
    2 = 'MEDIUM'
    3 = 'LOW' ;
%JACKPV (TOTWGT, JKZONE, JKREP, 75, IDCNTRY MCDGPPSC, MSMMAT0, 5, MERGED) ;
PROC PRINT DATA = FINAL NOOBS ;
  VAR IDCNTRY MCDGPPSC N TOTWGT MNPV MNPV_SE PCT PCT_SE ;
  FORMAT IDCNTRY COUNTRY. MCDGPPSC PPSC. N 6.0 TOTWGT 10.0
         MNPV MNPV_SE PCT PCT_SE 6.2 ;
```

In general, to perform analyses using the school background data files, you should do the following:

- 1) Identify the variables of interest in the school and student background data files and note any specific national adaptations to the variables.
- 2) Retrieve the relevant variables from the school background data files, including analysis variables, classification variables, identification variables (IDCNTRY and IDSCHOOL), and any other variables used in the selection of cases.

- 3) Retrieve the relevant variables from the student background data files, including plausible values of achievement, classification variables, identification variables (IDCNTRY and IDSCHOOL), sampling (JKZONE and JKREP) and weighting (TOTWGT) variables, and any other variables used in the selection of cases.
- 4) Merge the school background data files with the student background data files using the variables IDCNTRY and IDSCHOOL.
- 5) Perform any necessary variable transformations or recodes.
- 6) Use the macros JACKGEN and JACKREG, or JACKPV and JACKREGP if plausible values are involved, with the appropriate arguments and parameters.
- 7) Specify the location of the data files in the parameter <datpath> and of the macros in the parameter <macpath>.
- 8) Print the results file.

Exhibit 3.19 Output for Example School Variable Analysis (EXAMPLE 4)

IDCNTRY	BCDGPPSC	N	TOTWGT	MNPV	MNPV_SE	PCT	PCT_SE
ARMENIA	HIGH	23	49	409.25	16.21	1.83	0.08
ARMENIA	MEDIUM	661	2219	435.74	4.14	82.68	0.41
ARMENIA	LOW	174	416	419.84	5.35	15.50	0.40
IRAN	HIGH	532	26902	527.72	14.04	24.89	3.97
IRAN	MEDIUM	1390	63435	496.02	9.13	58.69	5.21
IRAN	LOW	440	17745	448.74	6.38	16.42	3.67
ITALY	HIGH	71	3656	481.37	45.58	3.07	1.78
ITALY	MEDIUM	1260	71625	458.18	8.74	60.11	5.15
ITALY	LOW	812	43881	430.75	12.36	36.82	5.01
LEBANON	HIGH	412	1131	558.22	4.64	24.59	1.96
LEBANON	MEDIUM	902	2715	542.89	3.09	59.04	2.07
LEBANON	LOW	261	753	524.50	3.73	16.37	1.29

In Exhibit 3.19, each country's results are presented on three lines, one for each value of the MCDGPPSC variable. The results are presented much in the same manner as in previous examples, where the countries are identified in the first column and the second column describes the category of MCDGPPSC being reported. From the first three lines, 1.83% of advanced mathematics students in Armenia attend schools with a high level of principals' perception of school climate, 82.68% attend schools with a medium level, and 15.50% attend schools with a low level of principals' perception of school climate. Also, the estimated mean achievement of advanced mathematics students in Armenia attending

schools with a high level of principals' perception of school climate is 409.25 (standard error of 16.21), whereas the estimated mean achievement of advanced mathematics students in schools with a medium and low level of principals' perception of school climate is 435.74 (standard error of 4.14) and 419.84 (standard error of 5.35), respectively.

References

Mullis, I.V.S., Martin, M.O., Robitaille, D.F., & Foy, P. (2009). *TIMSS advanced 2008 international report: Findings from IEA's study of achievement in advanced mathematics and physics in the final year of secondary school*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

SAS Institute (2002). *SAS system for Windows* (version 9.1). Cary, NC: SAS Institute.

A close-up, black and white photograph of typewriter keys, showing various characters like Greek letters and numbers. The keys are arranged in a grid, with some in sharp focus and others blurred in the background.

Chapter 4

The TIMSS Advanced 2008 International Database Files

4.1 Overview

The TIMSS Advanced 2008 international database contains achievement data and student, teacher and school background data collected in the 10 countries that took part in TIMSS Advanced 2008. The database also contains materials that provide additional information on its structure and contents. This chapter describes the contents of the database and is divided into nine major sections corresponding to the different file types and materials included in the database, as described in Exhibit 1.1 of Chapter 1. The database also includes data and relevant materials from TIMSS Advanced 1995¹ since the 1995 achievement scores were rescaled. Exhibit 4.1 lists all the countries that participated in the TIMSS Advanced assessments of 1995 and 2008, along with the identifying codes used in the international database.

4.2 TIMSS Advanced 2008 User Guide

The TIMSS Advanced 2008 database includes a copy of this *TIMSS Advanced 2008 User Guide* in printable PDF format. It also includes a series of four supplements to the User Guide, also in PDF format. The User Guide and its supplements are located in the “1_UserGuide_and_Database” folder of the DVD.

¹ The TIMSS Advanced 1995 database includes only student and school data since it did not administer a teacher background questionnaire in 1995.

Exhibit 4.1 Countries Participating in TIMSS Advanced 1995 and 2008

Countries	ISO Codes		Advanced Mathematics		Physics	
	Alpha	Numeric	2008	1995	2008	1995
Armenia	ARM	051	●		●	
Australia	AUS	036		●		●
Austria	AUT	040		●		●
Canada	CAN	124		●		●
Cyprus	CYP	196		●		●
Czech Republic	CZE	203		●		●
Denmark	DNK	208		●		●
France	FRA	250		●		●
Germany	DEU	276		●		●
Greece	GRC	300		●		●
Iran, Islamic Republic of	IRN	364	●		●	
Israel	ISR	376		●		●
Italy	ITA	380	●	●	●	
Latvia	LVA	428				●
Lebanon	LBN	422	●		●	
Lithuania	LTU	440		●		
Netherlands	NLD	528	●		●	
Norway	NOR	578	●		●	●
Philippines	PHL	608	●			
Russian Federation	RUS	643	●	●	●	●
Slovenia	SVN	705	●	●	●	●
Sweden	SWE	752	●	●	●	●
Switzerland	CHE	756		●		●
United States	USA	840		●		●

Supplement 1 – International Version of the TIMSS Advanced 2008 Background and Curriculum Questionnaires

Supplement 1 includes the international version of all background and curriculum questionnaires administered in TIMSS Advanced 2008. It is a good reference guide to understand what questions were asked and the variable names under which the responses are recorded in the international database.

Supplement 2 – National Adaptations of International Background Questionnaires

Supplement 2 provides details on all national adaptations that were applied to the national version of all TIMSS Advanced 2008 background questionnaires. Users should refer to this supplement for any special adaptations to background variables that could potentially affect the results of analyses.

Supplement 3 – Variables Derived from the Student, Teacher, and School Questionnaire Data

Supplement 3 describes how the derived background variables used for producing exhibits in the *TIMSS Advanced 2008 International Report* (Mullis, Martin, Robitaille, & Foy, 2009) were computed.

Supplement 4 – Sampling Stratification Information

Supplement 4 provides the labels assigned to the national explicit and implicit strata defined during the sampling process.

4.3 TIMSS Advanced 2008 Data Files

The TIMSS Advanced 2008 international database includes the actual data from all instruments administered to the students, their teachers, and their school principals. This includes the student responses to the achievement items and the responses to the student, teacher, and school background questionnaires. These data files also include the achievement scores estimated for participating students, as well as background variables derived for reporting in the *TIMSS Advanced 2008 International Report*. National Research Coordinators' responses to the curriculum questionnaires are also part of the international database and are described later in this chapter.

The TIMSS Advanced 2008 international database DVD also includes an updated release of the TIMSS Advanced 1995 database. All 1995 data files have the same attributes as their 2008 counterparts. However, the 1995 database includes only student and school data since no teacher questionnaire was administered in 1995.

This section describes the format and contents of the TIMSS Advanced data files. With the exception of the curriculum data files, they are provided in SAS export format (.EXP) and SPSS format (.SAV) in the "SAS_Data" and "SPSS_Data" folders of the DVD, respectively. Data files are provided for each

country that participated in TIMSS Advanced 2008 and for which internationally comparable data are available. Each participating country—with the exception of the Philippines that participated only in the advanced mathematics assessment—has two set of files: one set related to the students that participated in the advanced mathematics assessment and a second set related to the students participating in the physics assessment. The file names given to the various data file types are shown in Exhibit 4.2. For example, MSGNORM2.SAV is an SPSS file that contains Norway’s TIMSS Advanced 2008 advanced mathematics student background data. For each file type, a separate data file is provided for each participating country. All data files and the variables they contain are described in the following sections.

Exhibit 4.2 TIMSS Advanced 2008 Data File Names

File Names	Descriptions
MCG●●●M2	Advanced mathematics school background data files
MSA●●●M2	Advanced mathematics student achievement data files
MSG●●●M2	Advanced mathematics student background data files
MSR●●●M2	Advanced mathematics within-country reliability scoring data files
MST●●●M2	Advanced mathematics student-teacher linkage files
MTG●●●M2	Advanced mathematics teacher background data files
PCG●●●M2	Physics school background data files
PSA●●●M2	Physics student achievement data files
PSG●●●M2	Physics student background data files
PSR●●●M2	Physics within-country reliability scoring data files
PST●●●M2	Physics student-teacher linkage files
PTG●●●M2	Physics teacher background data files

●●● = Three-character country abbreviation based on the ISO 3166 coding scheme (see Exhibit 4.1).

4.3.1 TIMSS Advanced 2008 Student Achievement Data Files (MSA/PSA)

The TIMSS Advanced 2008 student achievement data files contain the student responses to the individual achievement items in the TIMSS Advanced 2008 assessments. The student achievement data files are best suited for performing

item-level analyses. Achievement scores (plausible values) for the TIMSS Advanced 2008 achievement scales are available in the student achievement data files, as well as in the student background data files and student-teacher linkage data files.

Students who participated in TIMSS Advanced 2008 were administered one of eight assessment booklets.² Some of these items were multiple-choice items and some were constructed-response items. The student achievement data files contain the actual responses to the multiple-choice questions and the scores assigned to the constructed-response items.

Variable Naming Convention for Achievement Items

The achievement item variable names are based on an 8-character alphanumeric code (e.g., MA23131A), which adheres to the following rules:

- The first and second characters are either “MA” for advanced mathematics items, or “PA” for physics items.
- The third character indicates the assessment cycle when the item was first used in TIMSS Advanced. The items in the 2008 assessments have code “1” for items produced in 1995 and “2” for new items in 2008.
- The fourth character is always “3” to identify TIMSS Advanced items from items of the regular TIMSS assessments.
- The fifth through seventh characters are a three-digit number used to uniquely identify the items.
- The eighth character indicates the item part, and only appears when required. It is generally a letter from “A” to “E”, depending how many parts there are to a particular item.

For example, MA23131A is the first part of an advanced mathematics item produced in 2008 and whose unique sequence number is 131.

Codes for Responses to the Achievement Items

A series of conventions were adopted to code the data included in the TIMSS Advanced data files. The following sections describe these conventions for the achievement items.

² The TIMSS Advanced 2008 booklet design is described in Chapter 3 of the *TIMSS Advanced 2008 Assessment Frameworks* (Garden, et al., 2006).

The values assigned to each of the achievement item variables depend on the item format. For multiple-choice items, numerical values from 1 through 5 are used to correspond to the response options A through E, respectively. For these items, the correct response is included as part of the variable label in the achievement codebook file and SAS and SPSS programs are included as part of the international database to score these items.

Each constructed-response item had its own scoring guide³ that relied on a two-digit scoring scheme to provide diagnostic information. The first digit designated the correctness level of the response: 2 for a 2-point response, 1 for a 1-point response, and 7 for an incorrect response. The second digit, combined with the first, represented a diagnostic code used to identify specific types of approaches, strategies, or common errors and misconceptions in responding to the item. A second digit of 0 through 5 was used for pre-defined international codes at each correctness level, while a second digit of 9 corresponded to “other” types of responses that fell within the appropriate correctness level, but did not fit any of the pre-defined international codes. A special code, 99, was used for completely blank responses.

Codes for Missing Values

A subset of the values for each variable type was reserved for specific codes related to different categories of missing data. We recommend that users read the following section with particular care since the way in which these missing codes are used may have major consequences for analyses.

Omitted Response Codes (SAS: . ; SPSS: 9, 99)

“Omitted” response codes were used for items that a student should have answered but did not. An omitted response code also was given when an item was left blank or when two or more response options were checked for a multiple-choice item.

3 Scoring guides for the released items are provided in the “Items” folder of the DVD.

Not Administered Response Codes (SAS: .A ; SPSS: sysmis)

Special codes were given to items that were “Not Administered” to distinguish these cases from data that were missing due to non-response. In general, the not administered code was used when an item was not administered, either by design arising from the rotation of items across the assessment booklets, or unintentionally when an item was misprinted or otherwise unavailable for a student to respond. The not-administered code was used in the following cases:

- Achievement item not assigned to the student: all students participating in TIMSS advanced received only one of the eight test booklets. All variables corresponding to items that were not present in a student’s assigned booklet were coded as “Not Administered.”
- Student absent from session: When a student was not present for a particular testing session, all variables relevant to that session were coded as “Not Administered.”
- Item left out or misprinted: When a particular item (or a whole page) was misprinted or otherwise not available to the respondent, the corresponding variable was coded as “Not Administered.”
- Item deleted or mistranslated: An item identified during translation verification or item review as having a translation error, such that the nature of the question was altered, or as having poor psychometric properties, was coded as “Not Administered.”

Not Reached Response Codes (SAS: .R ; SPSS: 6, 96)

An item was considered “Not Reached” when the item itself and the item immediately preceding it were not answered, and there were not other items completed in the remainder of the assessment booklet. For most purposes, TIMSS Advanced 2008 treated the not-reached items as incorrect responses, except during the item calibration step of the IRT scaling, when not-reached items were considered to have not been administered (see Foy, Galia, & Li, 2009).

TIMSS Advanced Achievement Scores

Achievement scales were produced for advanced mathematics and physics. A detailed description of the TIMSS Advanced 2008 scaling and how these achievement scales were created is given in Chapter 8 (Foy, Galia, & Li, 2009)

of the *TIMSS Advanced 2008 Technical Report*. For each achievement scale, the TIMSS Advanced 2008 database provides five separate estimates of each student's score on that scale. The five estimated scores are known as "plausible values" and the variability between them encapsulates the uncertainty inherent in the scale estimation process.

The plausible values for any given scale are the best available measures of student achievement on that scale in the TIMSS Advanced 2008 international database, and should be used as the outcome measure in any study of student achievement. Plausible values can be readily analyzed using the IEA IDB Analyzer and the SAS programs described in this User Guide.

The variables that contain the plausible values of advanced mathematics achievement are labeled MSMMAT01 through MSMMAT05 in the TIMSS Advanced 2008 international database. The plausible values of physics achievement are labeled PSPPHY01 through PSPPHY01.

In addition to the plausible values for the achievement scales, the TIMSS Advanced 2008 database includes three interim achievement scores that were computed as part of the data processing effort.

Raw Scores

MSMSCT, PSPSCT

Number of score points obtained by a student on the advanced mathematics items, or physics items, in his or her assigned booklet.

After the achievement items were scored (1 for correct or 0 for incorrect for multiple choice items; 0, 1, or 2 points for constructed-response items), raw scores were computed by adding the number of points obtained by each student over all the items in the student's assessment booklet. Because a raw score is dependent on the number of items and score points in a student's assessment booklet, and since this number varies from booklet to booklet, the raw scores are not comparable across booklets, and so may be of limited utility. Their main value in the database is as a validity check for analysts who wish to apply a different scoring approach to the TIMSS Advanced items.

Standardized Raw Scores

MSMSTDR, PSPSTDR

Standardized advanced mathematics and physics raw scores.

Because of the difficulty in making any comparisons across the test booklets using only the number of raw score points obtained on a set of items, raw scores were standardized by booklet to provide a simple score that could be used in comparisons across the TIMSS Advanced 2008 booklets for preliminary analyses. Each standardized score was computed so that the weighted mean score within each booklet in a country was equal to 50, and the weighted standard deviation was equal to 10. Despite this standardization, comparisons across booklets have limited validity because of differences in difficulty across the assessment booklets, although every effort was made to ensure all booklets were similar in difficulty.

National Rasch Scores

MSMNRSC, PSPNRSC

National Rasch advanced mathematics and physics scores.

The national Rasch scores were computed to facilitate preliminary item analyses that were conducted prior to the TIMSS Advanced 2008 IRT scaling. Their main purpose was to provide preliminary measures of overall advanced mathematics and physics achievement that could be used as criterion variables in evaluations of item discrimination. The national Rasch scores were standardized to have a mean score of 150 points and a standard deviation of 10 points within each country. Because each country has the same mean score and dispersion, these scores are not useful for international comparisons.

TIMSS Advanced International Benchmarks of Achievement

To help users of the TIMSS Advanced achievement results understand what performance on the advanced mathematics and physics achievement scales signifies in terms of the advanced mathematics and physics students in the final year of secondary school know and can do, TIMSS Advanced identified three points on the advanced mathematics and physics scales to serve as international benchmarks. As shown in Exhibit 4.3, the TIMSS Advanced

international benchmark scores are 625, 550, and 475, corresponding to the Advanced International Benchmark, the High International Benchmark, and the Intermediate International Benchmark, respectively. TIMSS Advanced used a technique known as scale anchoring⁴ to summarize and describe student achievement at these three points on the scale. The *TIMSS Advanced 2008 International Report* presents the results of this scale anchoring, and reports the percentage of students in each country reaching each of the international benchmarks.

Exhibit 4.3 TIMSS Advanced 2008 International Benchmarks for Advanced Mathematics and Physics Achievement

Scale Scores	International Benchmarks
625	Advanced International Benchmark
550	High International Benchmark
475	Intermediate International Benchmark

To assist analysts in using the international benchmarks in secondary analyses, the TIMSS Advanced 2008 international database contains a set of variables indicating which international benchmark the students reached. There is a benchmark variable for each plausible value of the advanced mathematics and physics scales. The five benchmark variables for advanced mathematics are labeled MSMIBM01 through MSMIBM05; PSPIBM01 through PSPIBM05 are the five benchmark variables for physics. The codes used for all the benchmark variables are described in Exhibit 4.4.

Exhibit 4.4 TIMSS Advanced 2008 International Benchmark Variable Codes

Codes	Descriptions
1	Student performed below the Intermediate International Benchmark
2	Student performed at or above the Intermediate International Benchmark but below the High International Benchmark
3	Student performed at or above the High International Benchmark but below the Advanced International Benchmark
4	Student performed at or above the Advanced International Benchmark

⁴ The scale anchoring procedure is described in Chapter 8 (Foy, Galia, & Li, 2009) of the *TIMSS Advanced 2008 Technical Report*.

4.3.2 TIMSS Advanced 2008 Within-Country Scoring Reliability Data Files (MSR/PSR)

The TIMSS Advanced 2008 within-country scoring reliability data files contain data that can be used to investigate the reliability of the TIMSS Advanced constructed-response item scoring. The scoring reliability data files contain one record for each booklet that was double scored during the within-country scoring reliability exercise (see Johansone, 2009). For each constructed-response item in the achievement test, the following three variables are included in the scoring reliability data files:

- Original Score (two-digit score assigned by the first scorer);
- Second Score (two-digit score assigned by the second scorer);
- Score Agreement (degree of agreement between the two scorers).

It should be noted that the Second Score data are located only in the MSR/PSR files and were used only to evaluate within-country scoring reliability. They were not used in computing the achievement scores included in the database and presented in the international report.

Scoring Reliability Variable Naming Convention

The variable names for the Original Score, Second Score, and Score Agreement variables are based on the same naming convention as that for the achievement item variables shown earlier. The second character in the variable name differentiates between the three reliability variables:

- The Original Score variable has the letter “A” as the second character, in accordance with the achievement item naming convention (e.g., MA23131A)
- The Second Score variable has the letter “R” as the second character (e.g., MR23131A)
- The Score Agreement variable has the letter “I” as the second character (e.g., MI23131A).

Reliability Variable Score Values

The values contained in both the Original Score and Second Score variables are the two-digit diagnostic codes assigned using the TIMSS Advanced 2008 scoring guides. The Score Agreement variable may have one of three values,

depending on the degree of agreement between the two scorers, as described in Exhibit 4.5.

Exhibit 4.5 TIMSS Advanced 2008 Score Agreement Variable Codes

Codes	Descriptions
0	Identical codes (both digits in the original and second scores)
1	Identical score levels, but different diagnostic codes (first digit of both scores are the same; second digits are different)
2	Different score levels (first digit of both scores are different)

In general, the data in the Original Score variables are identical to those contained in the student achievement data files. In some cases, however, the response scores for specific items were recoded after a review of the international item statistics revealed inconsistencies in the original scoring guides or showed that the original scores were not functioning as desired. The recoded score values were used in computing the achievement scores reflected in the international report and are present in the student achievement data files. In contrast, the Original Score variables in the scoring reliability data files contain the original unrecoded response scores. This was done so that the scoring reliability measures indicated in the Score Agreement variables were based on the original scoring guides used during the constructed-response scoring sessions conducted in each country.

4.3.3 TIMSS Advanced 2008 Background Questionnaire Data Files

There are five types of TIMSS Advanced 2008 background questionnaire data files: four of them—student, teacher, school, and curriculum—corresponding to the four types of background questionnaires administered in TIMSS Advanced 2008 and the fifth one used to link the student and teacher background data. The first four types of data files contain the responses to the questions asked in their respective background questionnaires.

TIMSS Advanced 2008 Student Background Data Files (MSG/PSG)

Students who participated in TIMSS Advanced 2008 were administered a background questionnaire with questions related to their home background, school experiences, and attitudes to advanced mathematics and physics. The

student background data files contain students' responses to these questions. They also contain students' advanced mathematics and physics achievement scores (plausible values) to facilitate analyses of relationships between student background characteristics and achievement.

There were two versions of the student questionnaire: one for advanced mathematics students and one for physics students. Students completed the student questionnaire that was appropriate for the assessment booklet they received—either advanced mathematics or physics. The responses of students to the advanced mathematics student questionnaire are found in the MSG files and the responses of students to the physics student questionnaire are in the PSG files.

The student background data files also contain a number of identification variables, tracking variables, sampling and weighting variables, and derived variables that were used for producing exhibits in the International Report. These variables are described later in this chapter.

TIMSS Advanced 2008 Teacher Background Data Files (MTG/PTG)

The advanced mathematics and physics teachers of the students that were sampled for TIMSS Advanced 2008 were administered a questionnaire with questions pertaining to their background and their teaching practices in the classes of the sampled students. Each teacher was asked to respond to a questionnaire for each class taught that contained sampled students. The teacher background data files contain one record for each of the classes taught either by an advanced mathematics or a physics teacher. If a teacher taught more than one class, they were expected to complete only one part A (general background questions) and a separate part B (class-specific questions) for each class they taught. In some cases, although the teacher was to respond to more than one questionnaire, responses to only one were obtained. In these cases, there were as many records entered in the teacher background data file as classes were taught by the teacher, and the background information in part A from the completed questionnaire was entered into these teacher records.

There were two types of teacher questionnaires—one for the advanced mathematics teachers and one for the physics teachers. The responses of teachers to the advanced mathematics teacher questionnaire are found in the MTG files and the responses of teachers to the physics teacher questionnaire are in the PTG files.

In the teacher background data files, each teacher has a unique identification number (IDTEACH) and a link number (IDLINK) that is specific to the class taught by the teacher and to which the information in the data record corresponds. The IDTEACH and IDLINK combination uniquely identifies, within a country, a teacher teaching a specific class. So, for example, students linked to teachers identified by the same IDTEACH but different IDLINK are taught by the same teacher but in different classes. Consequently, the teacher background data files cannot be merged directly with the student data files and they do not contain sampling and weighting information nor achievement scores.

It is important to note that the teachers in the teacher background data files do not constitute a representative sample of teachers in a country, but rather are the teachers who taught a representative sample of students. The teacher data, therefore, should be thought of as attributes of the students to which they are linked, and should be analyzed only in conjunction with the student-teacher linkage data files. This User Guide describes student-level analyses of teacher data with the student-teacher linkage data files using the IEA IDB Analyzer software in Chapter 2, as well as using SAS programs and macros in Chapter 3.

TIMSS Advanced 2008 School Background Data Files (MCG/PCG)

The school background data files contain school principals' responses to the questions in the TIMSS Advanced school background questionnaire. Although school-level analyses where the schools are the units of analysis can be performed, it is preferable to analyze school-level variables as attributes of students. To perform student-level analyses with school data, the school background data files must be merged with the student background data files using the country and school identification variables. Details of the merging procedure using the IEA IDB Analyzer or SAS programs are described in Chapters 2 and 3 of this User Guide, respectively.

There was a single TIMSS Advanced school background questionnaire for all schools, whether they were sampled for advanced mathematics, physics, or both. Consequently, the TIMSS Advanced school files for advanced mathematics (MCG) and physics (PCG) have identical structures. The variable names differ, however; the variables with the school principals' responses start with the letter "M" in the MCG files, whereas they begin with the letter "P" in the PCG files.

TIMSS Advanced 2008 Student-Teacher Linkage Data Files (MST/PST)

The TIMSS Advanced 2008 student-teacher linkage data files contain information required to link the student and teacher data files. The student-teacher linkage data files contain one entry per student-teacher linkage combination in the data. For instance, if two teachers are linked to a student, there are two entries in the file corresponding to that student. The sole purpose of the student-teacher linkage data files is to link teacher-level data with student-level data to perform appropriate student-level analyses where teacher characteristics become attributes of the students.

TIMSS Advanced 2008 Curriculum Data Files

The TIMSS Advanced 2008 curriculum questionnaire data files contain the responses provided by the National Research Coordinators of the participating countries to the TIMSS Advanced 2008 curriculum questionnaires. There are two separate curriculum questionnaire data files for the two populations—advanced mathematics and physics. These files are available as Excel files in the “Curriculum” folder of the DVD.

Variable Naming Convention for Background Questions

The variable naming convention for questions in the background questionnaires is based on an eight-character string. The following rules are applied in naming the background variables:

- The first character is either “M” for advanced mathematics data, or “P” for physics data.
- The second character indicates the type of respondent. The letter “C” is used to identify data from the school principals, the letter “T” is used for teacher data, the letter “S” for student data, and the letter “U” for curriculum data.
- The third character is used to indicate the survey cycle when a background question was first introduced. All background variables use the number “2” to represent TIMSS Advanced 2008 as the second cycle of studies. Alternatively, the letter “D” is used to identify derived variables.

- The fourth character is used to indicate the subject or topic to which a background question refers. The following letters are used:
 - G General question (not subject specific)
 - M Question related to advanced mathematics
 - P Questions related to physics
- The fifth through eighth characters of all background questionnaire variables are used to assign a unique and concise label to each question.

Variable Location Convention for Background Questions

To identify the location of a background variable in its corresponding background questionnaire, each question was assigned a unique identification code as shown in Exhibit 4.6. This unique code is followed by the sequence number of the question within the questionnaire. For example, if the location of a variable is given as SQ3-06, it refers to question 6 in the student background questionnaire. This convention is followed in the data almanacs and in the description of the variables included in Supplement 2 and Supplement 3 to this User Guide.

Exhibit 4.6 Background Questionnaire Variable Location Convention

Questionnaire	Location Code
Student Questionnaire	SQ3-●●●● for questions common to both questionnaires SQM3-●●●● for questions unique to advanced mathematics questionnaire SQP3-●●●● for questions unique to physics questionnaire
Teacher Questionnaire	TQM3-●●●● for advanced mathematics questionnaire TQP3-●●●● for physics questionnaire
School Questionnaire	SCQ3-●●●● for all school questions
Curriculum Questionnaire	CQM3-●●●● for advanced mathematics questionnaire CQP3-●●●● for physics questionnaire

●●●● = sequential numbering of the question location in the questionnaire

Data Coding Conventions

A series of conventions also were adopted to code the data included in the data files. The following sections describe these conventions.

Codes for Responses to Background Questions

The values assigned to each of the background variables depend on the item format and the number of options available. For categorical questions, sequential numerical values are used to correspond to the response options available. The numbers correspond to the sequence of appearance of the response options. For example, the first response option is represented with a 1, the response option with a 2, etc. Open-ended questions, such as “the number of students in a class”, are coded with the actual number given as a response.

Codes for Missing Values

A subset of values for each variable type is reserved for specific codes related to various categories of missing data in the background variables.

Omitted Response Codes (SAS: . ; SPSS: 9, 99, 999, ...)

“Omitted” response codes were used for questions that a student, teacher, or school principal should have answered but did not. The length of the omitted response code given to a variable in the SPSS data files depends on the number of characters needed to represent the variable. In all cases, the space necessary to represent the variable is filled with 9’s. For questionnaire data, no distinction was made between items left blank and items with invalid answers, such as checking two or more response options in a categorical question, or unreadable or uninterpretable responses to open-ended questions.

Not Administered Response Codes (SAS: .A ; SPSS: sysmis)

Special codes were given to items that were “Not Administered” to distinguish these cases from data that were missing due to non-response. In general, the not administered code was used when a questionnaire was not completed or a question was not administered, such as when a question was left out of the instrument or misprinted. The not-administered code was used in the following cases:

- Question left out or misprinted: When a particular question (or a whole page) was misprinted, or otherwise not available to the respondent, the corresponding variables were coded as “Not Administered.”
- Background questions removed: Variables corresponding to questions in the student, teacher, or school background questionnaires that were

considered not applicable in some countries were not included in the national versions of the questionnaires. These questions were coded as “Not Administered.”

- Background questions mistranslated or not internationally comparable: In some cases, questions in the international version of the questionnaires were mistranslated or modified to fit the national context. Whenever possible, modified questions were recoded to match as closely as possible the international version. When this was not possible, modified questions were recoded as “Not Administered.”

Not Applicable Response Codes (SAS: .B ; SPSS: 6, 96, 996, ...)

“Not Applicable” response codes were used for the background questionnaire items for which responses were dependent on a filter question. Generally, a “No” response to a filter question lead to any follow-up questions being coded as not applicable since there were no appropriate responses to these follow-up questions.

Summary Indices and Derived Variables

In the TIMSS Advanced questionnaires, there were often several questions asked about various aspects of a single construct. In these cases, responses to the individual items were combined to create a derived variable which provided a more comprehensive picture of the construct of interest than the individual variables could on their own.

In the *TIMSS Advanced 2008 International Report*, an index is a special type of derived variable that assigns students to one of three levels—high, medium, or low—on the basis of their responses to the component variables. The high category of an index represents the responses that are expected to characterize aspects of a positive learning environment, and the low category those responses that are least supportive of learning.

Records—whether students, teachers or schools—were included in the derived variable calculation only if there were valid data for at least two thirds of the variables involved. For example, if a derived variable was based on six component variables, records that were missing responses to more than two of these were counted as missing on the derived variable. Supplement 3 to the User Guide provides a description of all derived variables included in the international database.

Using Sampling Weights when Analyzing the TIMSS Advanced 2008 Data

An important characteristic of the TIMSS Advanced studies, and one that has crucial implications for data analysis, is that they use data from carefully-drawn random samples of schools, classes, and students to make inferences about the advanced mathematics and physics achievement in the final year of secondary school. For analyses based on these sample data to accurately reflect population attributes, it is necessary that they take the design of the sample into account. This is accomplished in part by assigning a sampling weight to each respondent in the sample, and weighting the respondents by their sampling weight in all analyses. The sampling weights properly account for the sample design, take into account any stratification or disproportional sampling of subgroups, and include adjustments for non-response (see LaRoche, Zuehlke, & Joncas, 2009).

Because the students within each country were selected using probability sampling procedures, the probability of each student being selected as part of the sample is known. The sampling weight is the inverse of this selection probability. In a properly selected and weighted sample, the sum of the weights for the sample approximates the size of the population. In TIMSS Advanced, the sum of the sampling weights of all students in a country is an estimate of the size of the advanced mathematics, or physics, student population in that country. The student sampling weight, known as TOTWGT in the international database, must be used whenever student population estimates are required. The use of TOTWGT ensures that the various subgroups that constitute the sample are properly and proportionally represented in the computation of population estimates, and that the sample size will be inflated to approximate the size of the population.

Because statistics generated from the international database are estimates of national performance based on samples of students, rather than the value that could be calculated if every student in every country had answered every question, it is important to have a way of quantifying the uncertainty associated with these statistics. In TIMSS Advanced, the jackknife procedure is used to provide a robust estimate of the sampling error of each statistic presented in the international report. When used with achievement scores, or plausible values, the jackknife standard errors include both the error component due to sampling variation and the error component due to variation among the five plausible values generated for each student. These

standard errors may be used to create confidence intervals for statistics computed from the TIMSS Advanced data.

The TIMSS Advanced 2008 international database includes the IEA IDB Analyzer software (see Chapter 2) that enables analysts to apply the sampling weights, the jackknife algorithm, and plausible values to a range of analyses of school, teacher, and student variables. It also provides a set of SAS programs and macros (see Chapter 3), which will perform analyses using the sampling weights, the jackknife algorithm, and plausible values.

Sampling Weights in the TIMSS Advanced 2008 Data Files

Several sampling and weighting variables are included in the TIMSS Advanced data files. They are listed and described in Exhibit 4.7, whereas Exhibit 4.8 illustrates the location of the various sampling and weighting variables among the different types of data files. It is important to note that the teacher background data files do not have any sampling and weighting variables.

Although TOTWGT has desirable properties, it also has drawbacks for some analyses. Because TOTWGT sums to the student population in each country, analyses using TOTWGT that combine countries will have proportionately more students from larger countries and fewer from smaller countries, which may not be desirable for some purposes. For cross-country analyses in which each country should be treated equally, TIMSS Advanced provides SENWGT, a transformation of TOTWGT, that results in a weighted sample size of 500 in each country. Additionally, since TOTWGT inflates sample sizes to approximate the population size, software systems that use the actual sample size to compute significance tests will give misleading results for analyses weighted by TOTWGT. HOUWGT, another transformation of TOTWGT, ensures that the weighted sample corresponds to the actual sample size in each country.

The weight variables TOTWGT, SENWGT and HOUWGT are designed for use in student-level analyses from all student-level files. The weight variable SCHWGT is designed for use in school-level analyses where schools are the units of analysis.

Exhibit 4.7 TIMSS Advanced 2008 Sampling and Weighting Variables

Variable Names	Descriptions
TOTWGT	Total student weight – sums to the national population
SENWGT	Student senate weight – sums to 500 in each country
HOUWGT	Student house weight – sums to the student sample size in each country
MATWGT	Advanced mathematics teacher weight
PHYWGT	Physics teacher weight
JKZONE	The sampling zone, or stratum, to which the student’s school is assigned
JKREP	The sampling replicate, or primary sampling unit, to which the student’s school is assigned
JKCZONE	The sampling zone, or stratum, to which the school is assigned
JKCREP	The sampling replicate, or primary sampling unit, to which the school is assigned
WGTFAC1	School weighting factor
WGTADJ1	School weighting adjustment
WGTFAC2	Class weighting factor
WGTADJ2	Class weighting adjustment
WGTFAC3	Student weighting factor
WGTADJ3	Student weighting adjustment

The weight variables MATWGT and PHYWGT are specifically designed for using teacher background data in student-level analyses and are based on TOTWGT. Whereas MATWGT is used for analyses of advanced mathematics teachers, PHYWGT is used for analyses of physics teachers. These teacher weights are located in the student-teacher linkage files (MST/PST), not in the actual teacher background data files (MTG/PTG). Analyses with teacher data will be properly weighted by merging the teacher files with the student-teacher linkage files.

Exhibit 4.8 Location of Sampling and Weighting Variables in the TIMSS Advanced 2008 Database

Sampling and Weighting Variables	Data File Types			
	MSG PSG	MSA PSA	MST PST	MCG PCG
TOTWGT	●	●		
SENWGT	●	●		
HOUWGT	●	●		
MATWGT			●	
PHYWGT			●	
SCHWGT				●
JKREP	●	●	●	
JKZONE	●	●	●	
JKCREP				●
JKCZONE				●
WGTFAC1	●			●
WGTADJ1	●			●
WGTFAC2	●			
WGTADJ2	●			
WGTFAC3	●			
WGTADJ3	●			

The sampling variables beginning with the letters “JK” are used to compute standard errors based on the jackknife repeated replication methodology. All weighting variables beginning with the letters “WGT” are included to provide insight into the multi-stage sampling and weighting methodology applied to the TIMSS data. All weighting variables are described in Chapter 4 (LaRoche, Zuelhke, & Joncas, 2009) of the *TIMSS Advanced 2008 Technical Report*.

Structure and Design Variables in the TIMSS Advanced 2008 Data Files

Besides the variables used to store responses to the background questionnaires and achievement booklets, the TIMSS Advanced 2008 data files also contain

variables meant to store information that identify and describe the respondents and design information required to properly analyze the data.

Identification Variables

In all TIMSS Advanced data files, several identification variables are included that provide information to identify countries, students, teachers, or schools. These variables also are used to link cases between the different data file types. The identification variables have the prefix “ID” and are described below.

IDCNTRY

IDCNTRY is a three-digit country identification code based on the ISO 3166 classification as shown in Exhibit 4.1. This variable should always be used as the first linking variable whenever files are linked within and across countries.

IDPOP

IDPOP identifies the target grade and is always set to “3” in TIMSS Advanced to represent the final year of secondary school.

IDGRADE

IDGRADE identifies the target grade of the participating students. In TIMSS Advanced, the usual values are in the range of 10 through 13.

IDSCHOOL

IDSCHOOL is a four-digit identification code that uniquely identifies the participating schools within each country. The school codes are not unique across countries. Schools across countries can be uniquely identified only with the IDCNTRY and IDSCHOOL combination of linking variables.

IDCLASS

IDCLASS is a six-digit identification code that uniquely identifies the sampled classrooms within a country. The variable IDCLASS has a hierarchical structure and is formed by concatenating the IDSCHOOL variable and a two-digit sequential number identifying the sampled classrooms within a school. Classrooms can be uniquely identified in the database by the combination of IDCNTRY and IDCLASS as linking variables.

IDSTUD

IDSTUD is an eight-digit identification code that uniquely identifies each sampled student in a country. The variable IDSTUD also has a hierarchical structure and is formed by concatenating the IDCLASS variable and a two-digit sequential number identifying all students within each classroom. Students can be uniquely identified in the database by the combination of IDCNTRY and IDSTUD as linking variables.

IDBOOK

IDBOOK identifies the specific assessment booklet that was administered to each student. The booklets are given a numerical value from 1 through 4 for the advanced mathematics booklets and from 5 through 8 for the physics booklets.

IDSTRATE & IDSTRATI

IDSTRATE and IDSTRATI are identification variables generated by the school sampling process. IDSTRATE identifies the explicit strata and IDSTRATI the implicit strata from which the participating schools were sampled. The codes assigned to these two variables vary from country to country and are documented in Supplement 4 to the User Guide.

IDTEACH

IDTEACH is a six-digit identification code that uniquely identifies a teacher within a school. It has a hierarchical structure and is formed by the concatenation of IDSCHOOL and a two-digit sequential number within each school.

IDLINK

IDLINK is a two-digit identification code that uniquely identifies the class for which a teacher answered a questionnaire. The combination of linking variables IDCNTRY, IDTEACH, and IDLINK uniquely identifies all teacher-class combinations in the database.

Exhibit 4.9 shows in which data files the various identification variables are located. It also highlights the combinations of variables used to uniquely identify the records contained in the different data file types. In the student background

and achievement data files, the variables IDCNTRY and IDSTUD provide a unique identification number to identify all students in the database. Since teachers may teach more than one class, the combination of the IDCNTRY, IDTEACH and IDLINK variables in the teacher background data files is needed to uniquely identify all teachers and the classes they teach. Teacher background variables are linked to the appropriate students using the student-teacher linkage data files. The variable IDSCHOOL, contained in all files, is a unique identification number for each school within a country. Combined with IDCNTRY, it can be used to link school background data to corresponding students or teachers.

Exhibit 4.9 Location of Identification Variables in the TIMSS Advanced 2008 Database

Identification Variables	Data File Types				
	MSA PSA	MSG PSG	MST PST	MTG PTG	MCG PCG
IDCNTRY	●	●	●	●	●
IDGRADE	●	●	●	●	●
IDPOP	●	●	●	●	●
IDSCHOOL	●	●	●	●	●
IDCLASS	●	●	●	●	
IDSTUD	●	●	●		
IDBOOK	●	●			
IDSTRATE	●	●			
IDSTRATI	●	●			
IDTEACH			●	●	
IDLINK			●	●	

Tracking Variables

Information about students, teachers, and schools provided by the survey tracking forms⁵ is stored in the tracking variables. These variables have the prefix “IT.” All tracking variables are included in the student background data

⁵ Survey tracking forms are lists of students, teachers, or schools used for sampling and administrative purposes.

files. ITLANG is included in the student achievement and student background data files.

ITSEX

Gender of each student as stated in the Student Tracking Forms.

ITBIRTHM and ITBIRTHY

Month and year of birth of each student as stated in the Student Tracking Forms.

ITDATEM and ITDATEY

Month and year of testing for each student.

ITLANG

Language of testing for each student. It is set to “1” for all countries that tested in a single language. For countries that administered the test in more than one language, additional numerical codes are used that correspond to the order of the testing languages as shown in Supplement 2 to the User Guide.

4.4 TIMSS Advanced 2008 Codebook Files

All information related to the structure of the TIMSS Advanced 2008 data files, as well as the source, format, descriptive labels, and response option codes for all variables, is contained in codebook files. Each data file type in the database is accompanied by a codebook file, with the exception of the curriculum data files. Codebooks also are available for the 1995 database.

The naming convention for codebook files is as follows:

- The first three characters of the filename are in every respect identical to those in the file names shown in Exhibit 4.2.
- The next three characters identify the files as TIMSS Advanced codebooks and are always “TMA”.
- The seventh and eighth characters are always “M2” to indicate TIMSS Advanced 2008 as the second study cycle.

- The three-character file extension is always .SDB, which stands for standard dBase format.

Codebook files are located in the “Codebooks” folder of the DVD and can be read using Excel, or any standard database or spreadsheet program. Codebook files also are provided in printable PDF format. They describe the contents and structure of the TIMSS Advanced data files. Important codebook fields include FIELD_LABL, which contains extended textual information for all variables, QUEST_LOC, which provides the location of questions and achievement items within their respective survey instruments, and FIELD_CODE, which lists all acceptable responses allowed in the database.

4.5 TIMSS Advanced 2008 Achievement Item Information Files

Achievement item information files are provided to enable users of the TIMSS Advanced 2008 database to readily produce summaries of item characteristics. There are separate achievement item information files for advanced mathematics and physics in the database. They are in Excel format in the “Items” folder of the DVD and they include the following information for each item in the TIMSS Advanced 2008 assessments:

- The item’s permanent and unique identifier
- The item’s block and its sequential location within the block
- The item’s label
- The item’s content domain and cognitive domain
- The item’s type, either multiple-choice or constructed-response
- The number of options for a multiple-choice item
- The correct response key for a multiple-choice item
- The item’s point value
- An indicator showing if the item was included in the IRT scaling
- An indicator showing if the item was released after the 2008 assessment

Item-Related Documents

The “Items” Folder also includes various documents related to the TIMSS Advanced 2008 achievement items. It contains PDF versions of the released TIMSS Advanced 2008 advanced mathematics and physics achievement items. The documents include the items themselves with descriptive information and the scoring guides for the constructed-response items.

The folder also includes Excel files with the IRT item parameters estimated from the 1995-2008 concurrent calibration for all TIMSS Advanced items across both study cycles and for the advanced mathematics and physics achievement scales. These same item parameters are presented in Appendix D of the *TIMSS Advanced 2008 Technical Report* (Arora, Foy, Martin, & Mullis, 2009). The international database also includes the IRT item parameters from the rescaling of the 1995 assessments.

4.6 TIMSS Advanced Data Almanac Files

Data almanacs provide weighted summary statistics for all variables in the TIMSS Advanced 2008 data files. There are two basic types of data almanacs: achievement data almanacs for the achievement items and background data almanacs for the background variables. All data almanac files are located in the “Almanacs” folder of the DVD in printable PDF format. Almanacs also are available for the 1995 database.

Achievement Data Almanacs

The achievement data almanacs provide weighted summary statistics for each participating country on each individual achievement item included in the TIMSS Advanced 2008 assessments. There are separate achievement almanacs for advanced mathematics and physics. The achievement data almanacs display for each item its classification in the content and cognitive domains, the item block it belongs to, a brief description of the item, its variable name, whether it is a multiple-choice or constructed-response item, and the correct response key if it is a multiple-choice item. The trend item almanacs provide summary statistics for achievement items used in both the 1995 and 2008 assessments. The achievement data almanac files available in the database are listed in Exhibit 4.10. The data almanacs also display the international averages for each item, with each country weighted equally.

Exhibit 4.10 TIMSS Advanced 2008 Achievement Data Almanacs

Achievement Data Almanacs	Contents
TA08_MAT_ItemAlmanac	Almanac for advanced mathematics items
TA08_MAT_TrendItemAlmanac	Trend almanac for advanced mathematics items
TA08_PHY_ItemAlmanac	Almanac for physics items
TA08_PHY_TrendItemAlmanac	Trend almanac for physics items

There are two types of displays in the achievement data almanacs, depending on whether an item is a multiple-choice item or a constructed-response item. The following statistics are displayed in these almanacs:

- N: The number of students to whom the item was administered.
- DIFF: Percentage of students that responded correctly to a multiple-choice item.
- A, B, C, D and E: The percentage of students choosing each one of the response options for a multiple-choice item.
- Scoring Guide Codes (20, 21, 10, 11, 70, 71, etc.): The percentage of student responses assigned each of the codes in the scoring guide for a constructed-response item.
- OMITTED: The percentage of students that omitted to respond to the item.
- NOT REACHED: The percentage of students that did not reach the item.
- V1, V2: The percentage of students that scored 1 point or better on the item (V1) or 2 points or better (V2).
- 1.GIRL %RIGHT, 2.BOY %RIGHT: The percentage of girls and boys that either got a multiple-choice item right, or obtained the maximum score on a constructed-response item.

Background Data Almanacs

Background data almanac files contain weighted summary statistics for each participating country on each variable in the student, teacher, and school background questionnaires, including derived variables based on these

background variables. Among the statistics reported are mean advanced mathematics and physics achievement by response category. The background data almanacs also display for each variable the question as it was asked, its location in the corresponding questionnaire, and its variable name in the data files. The background data almanac files available in the database are listed in Exhibit 4.11. The background data almanacs also display the international averages for each variable, with each country weighted equally.

Exhibit 4.11 TIMSS Advanced Background Data Almanacs

Background Data Almanacs	Contents
TA08_MAT_StudentAlmanac	Advanced mathematics student background almanac with achievement
TA08_MAT_TeacherAlmanac	Advanced mathematics teacher background almanac with achievement
TA08_MAT_SchoolAlmanac	Advanced mathematics school background almanac with achievement
TA08_PHY_StudentAlmanac	Physics student background almanac with achievement
TA08_PHY_TeacherAlmanac	Physics teacher background almanac with achievement
TA08_PHY_SchoolAlmanac	Physics school background almanac with achievement

There are two types of displays in the background data almanacs, depending on whether the data are categorical (i.e., have a small number of discrete values) or continuous. The almanac tables for categorical variables include:

- The sample size (number of students, teachers or schools included in the sample)
- The number of valid cases (number of students, parents, teachers or schools for whom valid data were obtained)
- The weighted percentages of students corresponding to each valid response option (percentages based only on the students with valid data)

- The weighted percentages of students for whom none of the valid response options were selected, coded as “Not Administered” or “Omitted” (percentages based on the sample size)
- The weighted mean achievement values of students corresponding to each valid response option, as well as the “Not Administered” and “Omitted” codes
- In cases where a variable can be coded as “Not Applicable” because of an earlier filter question, the weighted percentage of students for whom the variable is coded as “Not Applicable” is also displayed, based only on the students with valid data, along with the corresponding weighted mean achievement

The almanac tables for continuous variables include:

- The sample size (number of students, teachers or schools included in the sample)
- The number of valid cases (number of students, parents, teachers or schools for whom valid data were obtained)
- The weighted percentages of students for whom the variable is coded as “Not Administered” or “Omitted” (percentages based on the sample size)
- The weighted mean, mode, minimum, maximum, and the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles across students (based only on the students with valid data)
- In cases where a variable can be coded as “Not Applicable” because of an earlier filter question, the weighted percentage of students for whom the variable is coded as “Not Applicable” is also displayed, based only on the students with valid data

4.7 TIMSS Advanced 2008 Test-Curriculum Matching Analysis Data Files

The Test-Curriculum Matching Analysis (TCMA) was conducted to investigate the appropriateness of the TIMSS Advanced 2008 advanced mathematics and physics tests for students in the participating countries. To that end, participating countries were asked to indicate which items on the TIMSS Advanced 2008 tests were included in their national curricula. Thus,

based on the computed average percent correct, each country was able to see the performance of all countries on the items appropriate for its curriculum, and also the performance of its students on the items judged appropriate for the curriculum in other countries. The analytical method used and the results of the TCMA are presented in Appendix C of the *TIMSS Advanced 2008 International Report*.

The “TCMA” folder of the DVD contains two TCMA data files—one for advanced mathematics and one for physics—in Excel format and printable PDF format, showing which items were selected by each participating country.

4.8 TIMSS Advanced 2008 Program Files

The TIMSS Advanced 2008 international database includes a number of SAS programs and macros designed to facilitate the manipulation of the TIMSS Advanced 2008 data files and conduct proper statistical analyses taking into account the jackknife algorithm and the presence of plausible values. These programs are located in the “SAS_Programs” subfolder of the “Programs” folder on the DVD and are described in Chapter 3.

The “IDB_Recode_Programs” subfolder contains SPSS syntax files to perform variable recodes required for the proper execution of example analyses using the IEA IDB Analyzer. They are described in Chapter 2.

The “SPSS_Programs” subfolder contains two SPSS programs that score the TIMSS Advanced 2008 items, much like their SAS counterparts described in Section 3.4 of Chapter 3.

4.9 TIMSS Advanced 1995 International Database

Because the assessment data from the TIMSS Advanced 1995 international database were rescaled, it is being released with its rescaled scores alongside the TIMSS Advanced 2008 international database and is included on the DVD. Along with the actual TIMSS Advanced 1995 data files, the DVD includes 1995 codebooks, almanacs and SAS and SPSS programs. The DVD also includes the IRT item parameters estimated from the rescaling. All files for the TIMSS Advanced 1995 international database are found in the “TA95_Database” folder of the DVD.

References

Arora, A., Foy, P., Martin, M.O., & Mullis, I.V.S. (Eds.). (2009). *TIMSS Advanced 2008 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Foy, P., Galia, J., & Li, I. (2009). Scaling the data from the TIMSS advanced 2008 assessments. Chapter 8 in I.V.S. Mullis, M.O. Martin, P. Foy, & A. Arora (Eds.), *TIMSS Advanced 2008 technical report* (pp. 143-206). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Garden, R.A., Lie, S., Robitaille, D.F., Angell, C., Martin, M.O., Mullis, I.V.S., Foy, P., & Arora, A. (2006), *TIMSS advanced 2008 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Johansone, I. (2009). TIMSS advanced 2008 Survey Operations Procedures. Chapter 5 in I.V.S. Mullis, M.O. Martin, P. Foy, & A. Arora (Eds.), *TIMSS Advanced 2008 technical report* (pp. 93-108). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

LaRoche, S., Zuehlke, O., & Joncas, M. (2009). TIMSS Advanced 2008 sampling. Chapter 4 in I.V.S. Mullis, M.O. Martin, P. Foy, & A. Arora (Eds.), *TIMSS Advanced 2008 technical report* (pp. 51-92). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., Robitaille, D.F., & Foy, P. (2009), *TIMSS advanced 2008 international report: Findings from IEA's Study of achievement in advanced mathematics and physics in the final year of secondary school*, Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.



Appendix A

Organizations and Individuals Responsible for TIMSS Advanced 2008

Introduction

TIMSS Advanced 2008 was a collaborative effort involving many individuals around the world. This appendix recognizes the individuals and organizations for their contributions. Given the work on TIMSS Advanced 2008 has spanned approximately four years and has involved so many people and organizations, this list may not include all who contributed. Any omission is inadvertent.

Of the first importance, TIMSS Advanced 2008 is deeply indebted to the students, teachers, and school principals who contributed their time and effort to the study.

Management and Coordination

TIMSS Advanced 2008 was conducted by the TIMSS & PIRLS International Study Center at Boston College, which has responsibility for the overall direction and management of IEA's TIMSS and PIRLS projects. Headed by Ina V.S. Mullis and Michael O. Martin, the study center is located in the Lynch School of Education. In carrying out the project, the TIMSS & PIRLS International Study Center worked closely with the IEA Secretariat in Amsterdam, which provided guidance overall and was responsible for verification of all translations produced by the participating countries. The IEA Data Processing and Research Center in Hamburg was responsible for processing and verifying the internal consistency and accuracy of the data submitted by the participants. Statistics Canada in Ottawa was responsible for school and student sampling activities. Educational Testing Service (ETS) in Princeton, New Jersey provided psychometric methodology recommendations

addressing calibration and scaling, and also made available software for scaling the achievement data.

The Project Management Team, comprised of the Directors and Senior Management from the TIMSS & PIRLS International Study Center, the IEA Secretariat, the IEA Data Processing and Research Center, Statistics Canada, and ETS, met twice a year throughout the study to discuss progress, procedures, and schedule. In addition, the Directors of the TIMSS & PIRLS International Study Center met with members of IEA's Technical Executive Group twice yearly to review technical issues.

Dr. Robert Garden from New Zealand was the TIMSS Advanced 2008 Mathematics Coordinator and Dr. Svein Lie, from the University of Oslo, was the TIMSS Advanced 2008 Physics Coordinator. Together with the Physics and Mathematics task force, a panel of internationally recognized experts in mathematics and physics research, curriculum, instruction, and assessment, they provided excellent guidance throughout TIMSS Advanced 2008.

To work with the international team and coordinate within-country activities, each participating country designated one or two individuals to be the TIMSS National Research Coordinator or Co-Coordinators, known as the NRCs. The NRCs had the complicated and challenging task of implementing the TIMSS Advanced 2008 study in their countries in accordance with TIMSS guidelines and procedures. The quality of the TIMSS Advanced 2008 assessment and data depends on the work of the NRCs and their colleagues in carrying out the very complex sampling, data collection, and scoring tasks involved.

Continuing the tradition of truly exemplary work established in other TIMSS assessments, the TIMSS Advanced 2008 NRCs performed their many tasks with dedication, competence, energy, and goodwill, and have been commended by the IEA Secretariat, the TIMSS & PIRLS International Study Center, the IEA Data Processing and Research Center, and Statistics Canada for their commitment to the project and the high quality of their work.

Funding

Funding for TIMSS Advanced 2008 was provided through a generous grant from the Norwegian Ministry of Education, fees from participating countries and through IEA's own resources. The financial support provided by Boston College is gratefully acknowledged.

IEA Secretariat

Hans Wagemaker, Executive Director
Barbara Malak, Manager Membership Relations
Juriaan Hartenberg, Financial Manager

TIMSS & PIRLS International Study Center at Boston College

Ina V.S. Mullis, Co-Director
Michael O. Martin, Co-Director
Pierre Foy, Director of Sampling and Data Analysis
Alka Arora, TIMSS Advanced 2008 Project Coordinator
Ryan Auster, TIMSS Graduate Assistant
Marcie Bligh, Manager of Office Administration
Susan Farrell, Co-Manager of Publications
Joseph Galia, Senior Statistician/Programmer
Christine Hoage, Manager of Finance
Jiefang Hu, TIMSS Graduate Assistant
Ieva Johansone, Survey Operations Coordinator
Isaac Li, Statistician/Programmer
Jennifer Moher, Data Graphics Specialist
Mario Pita, Co-Manager of Publications
Corinna Preuschoff, TIMSS Research Associate
Ruthanne Ryan, Data Graphics Specialist
Steven Simpson, Data Graphics Specialist

IEA Data Processing and Research Center

Dirk Hastedt, Co-Director

Milena Taneva, Manager, TIMSS Advanced Data Processing

Dirk Oehler, Researcher

Tim Daniel, Researcher

Alexander Konn, Programmer

Olaf Zuehlke, Researcher

Statistics Canada

Sylvie LaRoche, Senior Methodologist

Marc Joncas, Senior Methodologist

Educational Testing Service

Matthias Von Davier, Principal Research Scientist

Sampling Referee

Keith Rust, Vice President and Associate Director of the Statistical Group,
Westat, Inc.

TIMSS Advanced 2008 Task Force

Robert Garden, TIMSS Advanced 2008 Mathematics Coordinator

Svein Lie, TIMSS Advanced 2008 Physics Coordinator

Carl Angell, Norway

Wolfgang Dietrich, Sweden

Liv Sissel Gronmo, Norway

Helen Lye, Australia

Torgeir Onstad, Norway

David Robitaille, Canada

TIMSS Advanced 2008 National Research Coordinators (NRCs)

Armenia

Arsen Baghdasaryan
Yerevan State University

Iran, Islamic Republic of

Abdol'azim Karimi
Ministry of Education
Research Institute for Education (RIE)

Italy

Anna Maria Caputo
Instituto Nazionale per la Valuatazione del
Sistema Educativo di Istruzione e di Formazione

Lebanon

Leila Maliha Fayad
Center for Educational Research and
Development
Ministry of Education and Higher Education

Netherlands

Marjolein Drent
Martina Meelissen
Centre for Applied Research in Education
University of Twente

Norway

Liv Sissel Grønmo
University of Oslo, ILS

Philippines

Ester B. Ogena
Science Education Institute
Department of Science and Technology

Merle C. Tan
National Institute for Science and Mathematics
Education Development
University of Philippines

Russian Federation

Galina Kovaleva
Center for Evaluating the Quality of Secondary
General Education
Institute of Content and Methods of Learning
Russian Academy of Education

Slovenia

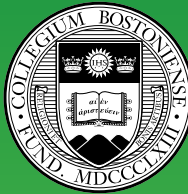
Barbara Japelj Pavesic
Educational Research Institute

Sweden

Marie Eklund
National Agency for Education



TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College



**BOSTON
COLLEGE**



timssandpirls.bc.edu
Copyright © 2009 International Association for the
Evaluation of Educational Achievement (IEA)

ISBN 1-889938-57-2

