ICCS 2009 User Guide for the International Database

Falk Brese
Michael Jung
Plamen Mirazchiyski
Wolfram Schulz
Olaf Zuehlke

ICCS
IEA 09

IEA

ACER

nfer

ROMA TRE
UNIVERSITÀ DEGLI STUDI

Università degli Studi Roma Tre
Laboratorio di Pedagogia sperimentale

ICCS 2009 User Guide for the International Database

Falk Brese
Michael Jung
Plamen Mirazchiyski
Wolfram Schulz
Olaf Zuehlke

Università degli Studi Roma Tre
Laboratorio di Pedagogia sperimentale

The International Association for the Evaluation of Educational Achievement, known as IEA, is an independent, international consortium of national research institutions and governmental research agencies, with headquarters in Amsterdam. Its primary purpose is to conduct large-scale comparative studies of educational achievement with the aim of gaining more in-depth understanding of the effects of policies and practices within and across systems of education.

Contents

List of Tables and Figures

Overview of ICCS 2009

1.1 Overview of the ICCS 2009 International Database and User Guide

The International Civic and Citizenship Education Study (ICCS) 2009 studied the ways in which countries prepare their young people to undertake their roles as citizens. ICCS 2009 was based on the premise that preparing students for citizenship roles involves helping them develop relevant knowledge and understanding and form positive attitudes toward being a citizen and participating in activities related to civic and citizenship education. These notions were elaborated in the ICCS 2009 framework, which was the first publication to emerge from ICCS 2009 (Schulz, Fraillon, Ainley, Losito, & Kerr, 2008).

The reports of results from ICCS 2009 (Schulz, Ainley, Fraillon, Kerr & Losito, 2010a & 2010b; Kerr, Sturman, Schulz & Burge, 2010; Schulz, Ainley, Friedman & Lietz, 2011) document variations among countries in relation to a wide range of different civic-related learning outcomes, actions, and dispositions. They also describe to what extent those outcomes are related to characteristics of countries, and the associations of these outcomes with student characteristics and school contexts. ICCS 2009 considered six research questions concerned with the following:

- Variations in civic knowledge
- Changes in content knowledge since 1999
- Students' interest in engaging in public and political life and their disposition to do so
- Perceptions of threats to civil society
- Features of education systems, schools, and classrooms related to civic and citizenship education
- Aspects of students' backgrounds related to the outcomes of civic and citizenship education.

ICCS 2009 gathered data from more than 140,000 Grade 8 (or equivalent) students in more than 5,300 schools from 38 countries. These student data were augmented by data from more than 62,000 teachers in those schools and by contextual data collected from school principals and the study's national research centers.

ICCS 2009 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 ICCS 2009 data, it is necessary to have an understanding of the characteristics of the study, which are described fully in the *ICCS 2009 Technical Report* (Schulz, Ainley, & Fraillon, forthcoming). It is intended, therefore, that this ICCS 2009 International Database (IDB) User Guide be used in conjunction with the ICCS 2009 Technical Report. Whereas the ICCS 2009 IDB User Guide describes the organization and content of the ICCS 2009 International Database, the *ICCS 2009 Technical Report* provides the rationale for the techniques used and for the variables created.

1.2 Analyzing the ICCS 2009 Data

The ICCS 2009 International Database offers researchers and analysts a rich environment for examining student achievement in civic knowledge in an international context. This includes:

- Extensive data on civic knowledge achievement providing in-depth study of the quality of education in terms of learning outcomes
- Comparable data for 38 countries from around the world providing an international perspective from which to examine educational practices and student outcomes in civic and citizenship education

- Comparable regional data for 24 countries from the European region, 6 countries from the Latin American region and 5 countries from the Asian region that allow investigations on educational practices and student outcomes in civic and citizenship education in a regional context
- Student achievement in civic knowledge linked to questionnaire information from students and school principals, providing policy-relevant contextual information on the antecedents of achievement
- Data from the teacher questionnaire that provide additional contextual information about the organization and culture of sampled schools as well as data on general and civic-specific aspects of teaching
- Achievement scales on a common metric that link ICCS 2009 to the CIVED 1999 sub scale "content knowledge" providing for analysis of changes in civic knowledge from CIVED 1999 to ICCS 2009.

The ICCS 2009 database is quite 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 ICCS 2009 complex multi-stage sample design and its use of imputed scores (also known as plausible values).

The ICCS 2009 student target population was students in the grade that represents eight years of schooling counted from International Standard Classification of Education (ISCED) Level 1, provided that the average age of students in this grade was 13.5 years or above at the time of the assessment (usually Grade 8). If the average age of students in that grade was below 13.5 years, the following grade (Grade 9 in all cases) became the target population.

The target population for the ICCS 2009 teacher survey was defined as all teachers teaching regular school subjects to the students in the target grade at each sampled school. It included only those teachers who were teaching the target grade during the testing period and who had been employed at school since the beginning of the school year.

To obtain accurate and representative samples, ICCS 2009 used a two-stage sampling procedure whereby a random sample of schools is selected at the first stage and one or two intact target grade classes in the case of students or a random sample of teachers from the target grade is sampled at the second stage. This is an effective and efficient sampling approach, but the resulting student sample has a complex structure that must be taken into consideration when analyzing the data. In particular, sampling weights need to be applied and a variance estimation technique such as the jackknife repeated replication needs to be used to estimate sampling variances correctly.[1]

In addition, ICCS 2009 uses Item Response Theory (IRT) scaling to summarize student achievement on the assessment and to provide accurate measures of changes from previous assessments. The ICCS 2009 IRT scaling approach used multiple imputation—or "plausible values"—methodology to obtain proficiency scores in civic knowledge for all students. Because each imputed score is a prediction based on limited information, it almost certainly includes some error. To allow analysts to incorporate this error into analyses of the ICCS 2009 achievement data, the ICCS 2009 International Database provides five separate imputed scores for the civic knowledge 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 information on standard errors incorporating both sampling and imputation error.[2]

---

1 More details on the sampling design and its implementation are provided in Chapter 6 of the *ICCS 2009 Technical Report* (Schulz et al., forthcoming).

2 More details on plausible values can be found in Chapter 11 of the *ICCS 2009 Technical Report* (Schulz et al., forthcoming).

IEA has developed the International Database (IDB) Analyzer software (IEA, 2010) specifically for analyzing ICCS 2009 international data files. Used in conjunction with SPSS,[3] this software helps users analyze the ICCS 2009 achievement data by conducting each analysis separately on each plausible value, averaging the resulting statistics, and applying the jackknife algorithm to provide appropriate standard errors for each statistic. It also simplifies management of the ICCS 2009 International Database by providing a module for selecting subsets of countries and variables, and merging files for analysis.

1.3   Contents of the ICCS 2009 IDB User Guide

This ICCS 2009 IDB User Guide describes the content and format of the data in the ICCS 2009 international database. In addition to this introduction, the ICCS 2009 IDB User Guide includes the following four chapters.

- Chapter 2 describes the structure and content of the ICCS 2009 International Database.
- Chapter 3 introduces the use of weighting and variance estimation variables for analyzing the ICCS 2009 data.
- Chapter 4 introduces the IEA International Database (IDB) Analyzer software (IEA, 2010) and presents examples of analyses of the ICCS 2009 data using this software in conjunction with SPSS.
- Chapter 5 explains how to implement the types of analyses described in Chapter 4 using the SAS (2002) statistical software system and the SAS programs and macros provided with the ICCS 2009 International Database.

The ICCS 2009 IDB User Guide is accompanied by five supplements.

- Supplement 1 includes the international version of all international questionnaires administered in ICCS 2009, along with the questionnaires from the regional module instruments. It will help the user understand what questions were asked and which variable names were used to record the responses in the international database.

- Supplement 2 provides details on all national adaptations that were applied to the national version of all ICCS 2009 international questionnaires, including the questionnaire sections of the regional module instruments. Users should refer to this supplement and check for any special adaptations to background and perception variables that could potentially affect the results of analyses.

- Supplement 3 describes how the derived questionnaire variables used for producing tables in the ICCS 2009 international and regional module reports (Schulz, Ailey, Fraillon, Kerr & Losito, 2010b; Kerr, Sturman, Schulz, Burge, 2010) were computed.

- Supplement 4 provides the information about the explicit and implicit stratification for each country used during the school sampling process.

- Supplement 5 contains all released test items from the ICCS 2009 assessment of civic knowledge along with their respective scoring guides.

---

3 This ICCS 2009 IDB User Guide will refer to SPSS; this includes PASW (Predictive Analytics Software), which replaces the older versions of the software known as SPSS.

1.4 Contents of the ICCS 2009 International Database

The ICCS 2009 International Database and all accompanying documentation is available from the IEA Study Data Repository website at http://rms.iea-dpc.org/. The following is a list of the data and documentation and a description of their contents available for download:

- Data: Student, teacher, and school data files in SAS and SPSS format as well as data from the National Context Survey in SPSS format
- Codebooks: Codebook files describing all variables in the ICCS 2009 International Database
- Documentation: This ICCS 2009 IDB User Guide with its Supplements and the *ICCS 2009 Technical Report* (Schulz et al., forthcoming)
- Programs: SPSS and SAS syntax programs and macros to support analyses.

An executable file for installing the IEA IDB Analyzer software to be used for analyzing the ICCS 2009 international data files can be downloaded from the IEA Studies Datasets and Data Analyzers section of the IEA website at http://www.iea.nl/iea_studies_datasets.html.

The ICCS 2009 International Database Files

2.1 Overview

The International Civic and Citizenship Education Study (ICCS) 2009 International Database contains student achievement data as well as student, teacher, and school questionnaire data collected in the 38 countries that participated in ICCS 2009. The database also includes data from the ICCS 2009 National Context Survey providing information on the national context of civic and citizenship education for all participating countries. Additionally, for countries participating in one of the three regional modules included in ICCS 2009, the database contains regional module data. In the case of the European module (24 countries) and the Latin American module (six countries) there are combined achievement and questionnaire data for the regional module whereas for the Asian module (five countries) there is questionnaire data only. Table 2.1 lists all ICCS 2009 countries, along with identifying codes used in the ICCS 2009 International Database. Table 2.1 also indicates participation in a regional module as well as the ICCS 2009 grade for over-time analysis if the country participated in the Civic Education Study (CIVED) in 1999. The database also contains materials that provide additional information on its structure and content. This chapter describes the content of the database and is divided into three major sections corresponding to the different file types and materials included in the database.

Table 2.1: Countries Participating in ICCS 2009

Countries	ISO Codes		ICCS 2009 Participation in Regional Module			Participation in CIVED 1999	Grade for Over-Time Analysis
	Alpha-3	Numeric	European Module	Latin American Module	Asian Module		
Austria	AUT	40	•				
Belgium (Flemish)	BFL	956	•				
Bulgaria	BGR	100	•			•	8
Chile	CHL	152		•		•	8
Chinese Taipei	TWN	158			•		
Colombia	COL	170		•		•	8
Cyprus	CYP	196	•			•	– [a]
Czech Republic	CZE	203	•			•	8
Denmark	DNK	208	•			•	– [a]
Dominican Republic	DOM	214		•			
England	ENG	926	•			•	9[b]
Estonia	EST	233	•			•	8
Finland	FIN	246	•			•	8
Greece	GRC	300	•			•	9
Guatemala	GTM	320		•			
Hong Kong SAR	HKG	344			•	•	– [a]
Indonesia	IDN	360			•		
Ireland	IRL	372	•				
Italy	ITA	380	•			•	8
Korea, Republic of	KOR	410			•		
Latvia	LVA	428	•			•	8
Liechtenstein	LIE	438	•				
Lithuania	LTU	440	•			•	8
Luxembourg	LUX	442	•				
Malta	MLT	470	•				
Mexico	MEX	484		•			
Netherlands	NLD	528	•				
New Zealand	NZL	554					
Norway	NOR	578				•	9
Paraguay	PRY	600		•			
Poland	POL	616	•			•	8
Russian Federation	RUS	643				•	– [a]
Slovak Republic	SVK	703	•			•	8
Slovenia	SVN	705	•			•	9
Spain	ESP	724	•				
Sweden	SWE	752	•			•	9 [b]
Switzerland	CHE	756	•			•	8
Thailand	THA	764			•		

[a] Cyprus, Denmark, Hong Kong SAR, and the Russian Federation did not collect comparable data to establish a link to CIVED 1999, either because of differences in the target population or changes to the CIVED 1999 test items for use in ICCS 2009.

[b] When interpreting the results for England and Sweden, readers need to take into account that students in CIVED 1999 were assessed at different times of the school year.

2.2 ICCS 2009 Data Files

The ICCS 2009 database includes data from all instruments administered to the students, the teachers teaching in the target grade at their school, and their school principals. This includes the student responses to the international achievement items and the responses to the international student, teacher, and school questionnaires as well as responses to the regional module achievement items and questionnaires. These data files also include the achievement scores estimated for participating students, as well as background variables derived for reporting in the ICCS 2009 international reports. Further, National Research Coordinators' responses to the National Context Questionnaire are also part of the international database.

This section describes the contents and format of the ICCS 2009 data files. They are provided in SPSS format (.SAV) and SAS export format (.EXP), except for the data from the National Context Survey which is only available in SPSS format (.SAV). They can be downloaded from the IEA Study Data Repository at http://rms.iea-dpc.org/. Data files are provided for each country that participated in ICCS 2009 and for which internationally comparable data are available. For the four countries (Greece, Norway, Slovenia, and Sweden) that administered an additional grade in ICCS 2009 to measure changes from the CIVED 1999 survey, there are additional grade data files as well. The file names given to the various data file types are shown in Table 2.2. For example, ISGNORC2.SAV is an SPSS file that contains Norway's ICCS 2009 target grade student questionnaire data. For each file type, a separate data file is provided for each participating country, with the exception of Greece and The Netherlands, which did not meet sampling requirements for the teacher survey and therefore no teacher data were released. All data files and the variables they contain are described in the following sections.

Table 2.2 ICCS 2009 Data File Names

File Names	Descriptions
ISG●●●C2	International Student Questionnaire File (Target Grade)
ISA●●●C2	International Student Achievement File (Target Grade)
ISR●●●C2	International Student Reliability File (Target Grade)
ISE●●●C2	European Module Student File (Target Grade)
ISL●●●C2	Latin American Module Student File (Target Grade)
ISS●●●C2	Asian Module Student File (Target Grade)
ITG●●●C2	Teacher Questionnaire File (Target Grade)
ICG●●●C2	School Questionnaire File (Target Grade)
JSG●●●C2	International Student Questionnaire File (Additional Grade)
JSA●●●C2	International Student Achievement File (Additional Grade)
JSR●●●C2	International Student Reliability File (Additional Grade)
JSE●●●C2	European Module Student File (Additional Grade)
NCQICSC2	National Context Questionnaire File

●●● = 3-character alpha-3 country code based on the ISO 3166 coding scheme (see Table 2.1)

2.2.1 Variable Naming Conventions

Achievement Item and Scoring Reliability Variable Naming Conventions

The achievement item variable names of the international test are based on an alphanumeric code (e.g., CI2COM1) consisting of up to eight characters, which adheres to the following rules:

- The first character indicates the general study context. "C" stands for civic and citizenship education.

- The second character "I" indicates that the variable is originally an achievement variable.

- The third character indicates the assessment cycle when the item was first used in ICCS. The item names in the ICCS 2009 assessment consist of either "1" for items used already in CIVED 1999, or "2" for items newly developed for ICCS 2009.

- The fourth and fifth characters indicate the item content for all items developed newly for ICCS 2009. Items reused from CIVED 1999 have a two digit item identifier as in CIVED 1999.

- The sixth character is used for the item type. "M" represents multiple-choice items, while "O" stands for open-ended items.

- The seventh digit represents the number of an item within a unit comprising the same content.

For example, CI2MLM1 is the first part of a multiple-choice item developed for ICCS 2009 and whose unique (content) identifier is ML.

In the scoring reliability files the variable names for the original score, second score, and score agreement variables are based on the same naming convention as for the international achievement item variables shown above. Only the second character in the variable name is used differently in order to differentiate between the three reliability variables:

- The original score variable has the letter "I" as the second character, in accordance with the achievement item naming convention (e.g., CI2PDO1)

- The second score variable has the letter "R" as the second character (e.g., CR2PDO1) and represents the score assigned by the reliability coder in the Reliability file.

- The score agreement variable has the letter "X" as the second character (e.g., CX2PDO1).

The achievement item variable names of the European and Latin American regional module test are explained below in the section about the naming conventions for questionnaire variables.

Questionnaire Variable Naming Conventions

The questionnaire variable names consist of a seven- or eight-character string. The following rules are applied in naming the variables of the international instruments as well as for questionnaire variables from the regional module instruments:

- The first character indicates the reference level. The letter "I" is used for variables that are administered on an international level. The letter "E" is used for variables from the European Module questionnaire, the letter "L" for variables from the Latin American Module questionnaire, and the letter "A" for variables from the Asian Module questionnaire.

- The second character indicates the type of respondent. The letter "C" is used to identify data from school principals, the letter "T" is used for teacher data, and the letter "S" for student data.

- The third character indicates the study cycle: Number "2" identifies ICCS 2009 as the second cycle of an IEA study exclusively focusing on civic and citizenship education (the first cycle refers to CIVED 1999).

- The fourth character indicates whether the variable relates to a background or perceptions question or whether it is a variable from the test section of the European or Latin American regional module instrument. The letter "G" represents questions related to the student background, letter "P" represents questions about student's attitudes, perceptions and behaviors, and letter "T" is used for test items in the regional module instruments.

- The fifth, sixth, seventh, and eighth characters indicate the question number. Their combination is unique to each variable within a questionnaire.

2.2.2 Questionnaire Variable Location Conventions

To identify the location of a questionnaire variable in its corresponding questionnaire, each question was assigned a unique identification code as shown in Table 2.3. 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 I-CQ-02, it refers to Question 2 in the school questionnaire. The question location is part of the information available in the codebook files.

Table 2.3 Questionnaire Variable Location Convention

Questionnaire	Location Code
Student Questionnaire	I-SQ-G- ● ● ● for background questions
	I-SQ-P- ● ● ● for perceptions questions
Teacher Questionnaire	I-TQ- ● ● ● for background and perceptions questions
School Questionnaire	I-CQ- ● ● ● for background and perceptions questions
European Module Questionnaire	E-SQ-1- ● ● ● for test items
	E-SQ-2- ● ● ● for perceptions questions
Latin American Module Questionnaire	L-SQ-1- ● ● ● for test items
	L-SQ-2 ● ● ● for perceptions questions
Asian Module Questionnaire	A-SQ- ● ● ● for perceptions questions

● ● ● = sequential numbering of the questions within a questionnaire

2.2.3 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. Users must read the following section with particular care since the way in which these missing codes are used may have significant consequences for analyses.

"Omitted" Response Codes (SPSS: 9, 99, 999, ...; SAS: .)
  "Omitted" response codes are used for questions or items that a student, teacher, or school principal should have answered but did not, that is, an "omitted" response code is given when an item is left blank. 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. For example, the "omitted" code for a one-digit variable is "9" whereas the "omitted" code for three-digit variables would be "999".

"Invalid" Response Codes (SPSS 7, 97, 997, …; SAS: .I)
  The response to a question is coded as invalid when the question was administered but an invalid response was given. This code is used for uninterpretable responses or in cases where the respondent has chosen more than one option to a multiple-choice question. The length of the invalid response code in the SPSS data files depends on the number of characters needed to represent the variable. For example, the "invalid" code for a one-digit variable is "7" whereas the "invalid" code for two-digit variables would be "97", and for three-digit variables, "997". Invalid codes are not applicable for open-ended items used in the international test instruments.

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

Special codes were given to items that were not administered to distinguish these cases from data that were missing because the respondent did not answer. In general, the "not administered" code was used when an item was not administered, either by design arising from the rotated test design (i.e., not every student was administered the same questions), or unintentionally when a question or item was misprinted or otherwise unavailable to a respondent. The "not administered" code was used in the following cases:

- Achievement item not assigned to the student: all students participating in ICCS 2009 received only one of the seven test booklets. All variables corresponding to items that were not part of the booklet assigned to a student were coded as "not administered."

- Student absent from session: When a student did not attend a particular testing session, for example because of sickness, all variables relevant to that session were coded as "not administered."

- Question or item left out or misprinted: When a particular question or item (or a whole page) was misprinted or otherwise not available to the respondent, the corresponding variable was coded as "not administered."

- Question or item deleted or mistranslated: A question or 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" if it could not be recoded to match as closely as possible the international version.

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

An item was considered not reached in the achievement data files when the item itself and the item preceding it were not answered, and there were no other items completed in the remainder of the booklet. For most purposes, ICCS 2009 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.[4]

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

"Not Applicable" response codes were used for the questionnaire items for which responses were dependent on a filter question. If the filter question was answered such that the subsequent questions would not apply, any follow-up question was coded "not applicable".

The length of the "not applicable" response code in the SPSS data files depends on the number of characters needed to represent the variable. For example, the "not applicable" code for a one-digit variable is "6" whereas the "not applicable" code for two-digit variables would be "96", and for three-digit variables, "996."

2.2.4 ICCS 2009 Student Achievement Data Files (ISA/JSA)

The ICCS 2009 student achievement data files contain the student responses to the individual achievement items in the ICCS 2009 assessments. The student achievement data files are best suited for performing item-level analyses. Achievement scores (plausible values) for the ICCS 2009 achievement scale are available only in the student questionnaire data files.

Students who participated in ICCS 2009 were administered one of seven assessment booklets, each with a series of items.[5] Most of the items were multiple-choice and some were constructed-response. The student achievement data files contain the actual responses to the multiple-choice questions and the scores assigned to the constructed-response items.

---

4 For more detailed information about the scaling procedure for ICCS test items refer to Chapter 11 of the *ICCS 2009 Technical Report* (Schulz et al., forthcoming).

5 The ICCS 2009 booklet design is described in Chapter 2 of the *ICCS 2009 Technical Report* (Schulz et al., forthcoming).

Item Response Code Values

A series of conventions were adopted to code the data included in the ICCS 2009 data files. This section describes these conventions for the achievement items.

The values assigned to each of the achievement item variables also depend on the item format. For multiple-choice items, numerical values from 1 to 4 are used to correspond to the response options A to D, respectively. For these items, the correct response is included in the achievement codebook file. The correct response is marked with an asterisk (*) following the value label of the correct option.

Each of the six open-ended response items had its own scoring guide[6] that used a one-digit scoring scheme. These items had a valid score range of 0 (incorrect response), 1 (partially correct response), and 2 (correct response). Five of the six items were scored so that responses related to two different described conceptual categories are scored as Code 2 and responses related to a single described conceptual category are scored as Code 1. Item CI2WFO2 followed a different scoring logic from that used for the previous five items. For this item the scoring codes reflect a conceptual hierarchy in which either of two categories of response warrant full credit (Code 2) and a different category of response warrants partial credit (Code 1). The "missing" code (Code 9) was used when a student made no attempt to answer a question. This code was only allocated when the entire stimulus, question stem, and question response area were left blank by the student.

2.2.5 ICCS 2009 Within-Country Scoring Reliability Data Files (ISR/JSR)

The ICCS 2009 within-country scoring reliability data files contain data that can be used to investigate the reliability of the ICCS 2009 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. For each constructed-response item in the achievement test, the following three variables are included in the scoring reliability data files:

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

It should be noted that the second score data were used only to evaluate within-country scoring reliability and were not used when computing the achievement scores included in the database and presented in the international reports.

Reliability Variable Score Values

The values contained in both the original score and second score variables are the one-digit diagnostic codes assigned following the ICCS 2009 scoring guides. The score agreement variable may have one of two values, depending on the degree of agreement between the two scorers: Code 0 was assigned if different scores were assigned, Code 1 was assigned in case of agreement between both scorers, and Code 9 was used if the item was coded as omitted by both scorers.

2.2.6 ICCS 2009 Questionnaire Data Files

There are six types of ICCS 2009 questionnaire data files corresponding to the six types of questionnaires administered in ICCS 2009. The student, teacher, and school data files contain the responses to the questions asked in their respective questionnaires. The regional module data files relate to the regional module instruments: As the European and Latin American

---

6 Scoring guides for the released items are provided in Supplement 5 of this *ICCS 2009 IDB User Guide*.

module instruments consist of a questionnaire and a cognitive test part the data files for these regional modules contain student responses to both the perceptional and behavioral questions as well as to the test items. The Asian module questionnaire contains only questionnaire response data.

ICCS 2009 Student Questionnaire Data Files (ISG/JSG)

Students who participated in ICCS 2009 were administered a perceptions questionnaire with questions related to their home background, their values, beliefs, and attitudes, and behaviors relevant to civic and citizenship. The student questionnaire data files contain students' responses to these questions. They also contain students' civic knowledge achievement scores (plausible values) to facilitate analyses of relationships between student background, perceptional characteristics, and achievement.

In addition, the student background data files feature a number of identification variables, tracking variables, sampling and weighting variables, and derived variables that were used for analyses in the international reports. These variables are described later in this chapter.

ICCS 2009 Regional Module Data Files (ISE/JSE; ISL; ISS)

Students from countries who participated in one of the regional modules were administered a regional module instrument in addition to the student test booklet and questionnaire. For the European and for the Latin American regional module the instrument contained a questionnaire and a cognitive test part. The Asian regional module instrument contained only a questionnaire, no cognitive test. The questions in the questionnaire are related to the students' values, beliefs, and attitudes as well as to behaviors relevant to the region. The questionnaire data files contain students' responses to these questions. In addition, the regional module data files for the European and Latin American Module contain student responses to the cognitive test items. As there are only multiple-choice items in the achievement part of these two regional module instruments, numerical values from 1 to 4 are used to correspond to the response options A to D, respectively. For these items, the correct response is included as part of the variable label in the achievement codebook file. The correct response is marked with an asterisk (*) following the value label of the correct option.

Identification variables and derived variables from the regional module data files that were used for analyses in the international reports are described later in this chapter.

ICCS 2009 Teacher Questionnaire Data Files (ITG)

The teachers sampled for ICCS 2009 were administered one questionnaire with questions pertaining to their background, the school environment, and civic and citizenship education at the school in which they teach. As an international option, some countries asked teachers additional questions about the teaching of civic and citizenship education.

In the teacher questionnaire data files, each teacher has a unique identification number (IDTEACH). The IDTEACH uniquely identifies a teacher within a country.

It is important to note that in contrast to some other IEA surveys the teachers in the teacher questionnaire data files constitute a representative sample of target grade teachers in a country. However, student and teacher data must not be merged directly because these two groups constitute separate target populations. Chapter 4 of this ICCS 2009 IDB User Guide describes student-level analyses with teacher data using the IEA IDB Analyzer software.

ICCS 2009 School Questionnaire Data Files (ICG)

The school questionnaire data files contain responses from school principals to the questions in the ICCS 2009 school questionnaires. Although school level analyses (where schools are the units of analysis) can be performed, it is preferable to analyze school level variables as attributes of students or teachers. To perform student or teacher level analyses with school data, the

school questionnaire data files must be merged with the student or teacher questionnaire data files using the country and school identification variables. Details of the merging procedure using the IEA IDB Analyzer are described in Chapter 4 of the ICCS 2009 IDB User Guide.

2.2.7 Data Coding Conventions

A series of conventions were adopted to code the data included in the data files. This section describes these conventions.

Questionnaire Response Code Values

The values assigned to each of the questionnaire 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, and so on. Open-ended questions, such as "the number of students in a school", are coded with the actual number given as a response.

2.2.8 Additional Variables

In ICCS 2009 an achievement scale was produced for the student's civic knowledge. A detailed description of the ICCS 2009 scaling and how the achievement scale was created is available in Chapter 11 of the *ICCS 2009 Technical Report* (Schulz et al., forthcoming). The ICCS 2009 International Database provides five separate estimates of each student's score on that scale. These are contained in the student questionnaire file. 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 the civic knowledge scale are the best available measures of student achievement on that scale in the ICCS 2009 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 ICCS 2009 IDB User Guide.

The achievement score variable names are based on a six-character alphanumeric code where PV1CIV represents the first plausible value and PV5CIV represents the fifth plausible value. In addition to the plausible values for the achievement scales, the ICCS 2009 database includes two achievement scores that were computed as part of the data processing effort: the National Civic Knowledge Scale and the National Civic Knowledge Rasch Scores.

National Civic Knowledge Scale (KNOWLMLE)

The scaling is based on the 15 international cognitive link items that pertain to the CIVED 1999 sub-scale measuring students' civic content knowledge. The maximum likelihood estimates (MLE), which have a mean of 150 and standard deviation of 10 within each country, were derived using the same item parameters as in CIVED 1999 and then transformed to the same scale metric. Scale scores are available only for three out of seven students who responded the link item cluster and only for those 17 national datasets where the student population is comparable with those surveyed in CIVED in 1999. The data can be analyzed using the same sample weights because booklets within schools were randomly allocated so that the students with CIVED 1999 content knowledge scale scores are a random sub-sample of the selected class.

National Civic Knowledge Rasch Scores (NWLCIV)

The national Rasch scores were computed to facilitate the preliminary item analyses that were conducted prior to the ICCS 2009 IRT scaling. The national Rasch scores were standardized to have a mean score of 150 points and a standard deviation of 10 points within each country.

The scaling is based on the 79 adjudicated international cognitive test items and provides nationally comparable results for students' civic knowledge. The weighted likelihood estimates (WLE) were computed using the same international parameters and scores are available only for students who participated in the test. Because each country has the same mean score and dispersion, these scores are not useful for international comparisons.

Summary Scales and Derived Variables

In the ICCS 2009 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 that provided a more comprehensive picture of the construct of interest than the individual variables could on their own.

In the ICCS 2009 reports, a scale is a special type of derived variable that assigns a score value to students on the basis of their responses to the component variables. In ICCS 2009 scales were typically calculated as IRT WLE scores with a mean of 50 and a standard deviation of 10 for equally weighted countries. Records—whether student, teacher, or school—were included in the scale calculation only if there were data for at least two of their indicator variables.

In addition to the scale indices, the ICCS 2009 International Database contains other indices that were derived by simple recoding or arithmetical transformation of original questionnaire variables.

Supplement 3 to this ICCS 2009 IDB User Guide provides a description of all derived variables (scale scores and indices) included in the international database. For further information about the scaling procedure for questionnaire items, refer to Chapter 12 of the *ICCS 2009 Technical Report* (Schulz et al., forthcoming).

Weighting and Variance Estimation Variables

To calculate population estimates and correct jackknife variance estimates, sampling and weighting variables are provided in the data files. Further details about weighting and variance estimation are provided in Chapter 3 of this ICCS 2009 IDB User Guide.

The following weight variables are included in the ICCS 2009 International Database:

TOTWGTS	Total student weight
SENWGTS	Senate student weight
WGTFAC1	School base weight
WGTADJ1S	School non-participation adjustment for the student survey
WGTFAC2S	Class base weight
WGTADJ2S	Class non-participation adjustment
WGTADJ3S	Student non-participation adjustment
TOTWGTT	Total teacher weight
SENWGTT	Senate teacher weight
WGTADJ1T	School non-participation adjustment for the teacher survey
WGTFAC2T	Teacher base weight
WGTADJ2T	Teacher non-participation adjustment
WGTADJ3T	Teacher multiplicity adjustment
TOTWGTC	Total school weight
WGTADJ1C	School non-participation adjustment for school level data analyses
TCERTAN	Indicator if teacher was sampled with certainty

The availability of these weight variables in the data files is shown below in Table 2.4.

Table 2.4 Location of Weighting Variables in the ICCS 2009 International Database

Weighting Variables	Data File Types						
	ISA JSA	ISG JSG	ITG	ICG	ISE JSE	ISL	ISS
TOTWGTS	•	•			•	•	•
SENWGTS	•	•			•	•	•
WGTFAC1		•	•	•			
WGTADJ1S		•					
WGTFAC2S		•					
WGTADJ2S		•					
WGTADJ3S		•					
TOTWGTT			•				
SENWGTT			•				
WGTADJ1T			•				
WGTFAC2T			•				
WGTADJ2T			•				
WGTADJ3T			•				
TOTWGTC				•			
WGTADJ1C				•			
TCERTAN			•				

The following variance estimation variables (or "jackknife variables") are included in the ICCS 2009 International Database. The actual replicate weights are computed together with the analysis results and are not part of the data files.

JKZONES Jackknife zone to which the students in a school are assigned

JKREPS Jackknife replicate to which the students in a school are assigned

JKZONET Jackknife zone to which the teachers in a school are assigned

JKREPT Jackknife replicate to which the teachers in a school are assigned

JKZONEC Jackknife zone to which a school is assigned for school level data analysis

JKREPC Jackknife replicate to which a school is assigned for school level data analysis

The availability of the variance estimation variables in the data files is shown in Table 2.5 below.

Table 2.5 Location of Variance Estimation Variables in the ICCS 2009 International Database

Variance Estimation Variables	Data File Types						
	ISA JSA	ISG JSG	ITG	ICG	ISE JSE	ISL	ISS
JKZONES	•	•			•	•	•
JKREPS	•	•			•	•	•
JKZONET			•				
JKREPT			•				
JKZONEC				•			
JKREPC				•			

Structure and Design Variables in ICCS 2009 Data Files

Besides the variables used to store responses to the questionnaires and achievement booklets, the ICCS 2009 data files also contain variables meant to store information that identifies and describes the respondents and design information required to properly analyze the data.

Identification Variables

In all ICCS 2009 data files, several identification variables provide information identifying countries, students, teachers, or schools. These variables are also used to link cases between the different data file types.

IDCNTRY

IDCNTRY is an up to three-digit numeric country identification code based on the ISO 3166 classification as shown in Table 2.1. This variable should always be used as the first linking variable whenever files are linked within or across countries.

COUNTRY

COUNTRY is a three-digit alphanumeric country identification code based on the ISO 3166 classification as shown in Table 2.1.

IDPOP

IDPOP identifies the grade and is set to "2" for the ICCS 2009 target grade (representing 8 years of schooling) and "3" for the additional grade (representing 9 years of schooling).

IDGRADE

IDGRADE identifies the tested grade of the participating students. In ICCS 2009, the value is "8" for most countries.

IDSCHOOL

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

CSYSTEM

CSYSTEM is an identification code that uniquely identifies each participating school in a country. This variable was introduced for data processing purposes.

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 across countries using the combination of IDCNTRY and IDCLASS.

IDSTUD

IDSTUD is an eight-digit identification code that uniquely identifies each sampled student within 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 across countries using the combination of IDCNTRY and IDSTUD.

SSYSTEM

SSYSTEM is an identification code that uniquely identifies each participating student in a country. This variable was introduced for data processing purposes.

IDBOOK

IDBOOK identifies the specific assessment booklet that was administered to each student. The booklets are assigned a numerical value from 1 through 7.

IDSCORA

IDSCORA uniquely identifies the scorer who scored the constructed-response items for the main scoring.

IDSCORR

IDSCORR uniquely identifies the scorer who scored the constructed-response items for the reliability scoring.

RELBOOK

RELBOOK is an indicator for the inclusion of the students' booklet into the reliability sample. It is set to "0" if the booklet is not part of the reliability sample and it is set to "1" if the booklet is part of the reliability sample.

IDSTRATE and 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 of this ICCS 2009 IDB User Guide.

IDTEACH

IDTEACH is a six-digit identification code that uniquely identifies the sampled teacher within a country. The variable IDTEACH has a hierarchical structure and is formed by concatenating the IDSCHOOL variable and a two-digit sequential number identifying the sampled teacher within a school. Teachers can be uniquely identified across countries using the combination of IDCNTRY and IDTEACH.

TYSTEM

TYSTEM is an identification code that uniquely identifies each participating teacher in a country. This variable was introduced for data processing purposes.

Table 2.6 shows in which data files the various identification variables are located.

Table 2.6 Location of Identification Variables in the ICCS 2009 International Database

Identification Variables	Data File Types							
	ISA JSA	ISR JSR	ISG JSG	ITG	ICG	ISE JSE	ISL	ISS
IDCNTRY	●	●	●	●	●	●	●	●
COUNTRY	●	●	●	●	●	●	●	●
IDPOP	●	●	●	●	●	●	●	●
IDGRADE			●					
IDSCHOOL	●	●	●	●	●	●	●	●
CSYSTEM					●			
IDCLASS	●	●	●			●	●	●
IDSTUD	●	●	●			●	●	●
SSYSTEM	●	●	●			●	●	●
IDBOOK	●	●	●					
IDSCORA	●							
IDSCORR		●						
IDSTRATE	●		●	●	●	●	●	●
IDSTRATI	●		●	●	●	●	●	●
IDTEACH				●				
TSYSTEM				●				

Tracking Variables

Information about students, teachers, and schools provided by the survey tracking forms[7] or used otherwise in the process of within-school sampling is stored in the tracking variables.

ITADMINI

    ITADMINI is the position of the test administrator of the test session as an attribute for each student. Code "1" is used for national center staff, code "2" for teachers from the school but not from the selected class, and code "3" is used for test administrators belonging to a group other than "1" or "2".

ITDATEM and ITDATEY

    ITDATEM and ITDATEY represent the month and year of testing for each student.

ITEXCLUD

    ITEXCLUD is an indicator for the exclusion of students. Because all students meeting any of the exclusion criteria were dropped from the ICCS 2009 International Database only students who were not excluded (code "9") remain.

ITPART1

    ITPART1 is an indicator for participation in the achievement test session for each student. It is set to "2" for students who were absent in the test session. ITPART1 is set to "3" for students participating in the achievement test session.

ITPART2

    ITPART2 is an indicator for participation in the questionnaire session for each student. It is set to "2" for students who were absent in the questionnaire session. ITPART2 is set to "3" for students participating in the questionnaire session.

ITPARTR

    ITPARTR is an indicator for participation in the regional module session for each student. It is set to "0" for students belonging to a country not participating in any of the regional modules. For students who were absent from the regional module session, code "2" is assigned. ITPARTR is set to "3" for students participating in the regional module session.

CPART

    CPART is the final participation indicator for each school principal. It is set to "3" (questionnaire returned completed) for all records in the database.

SPART

    SPART represents the final participation indicator for each student. It is set to "3" for all records in the database, indicating that a student participated either in the questionnaire or the achievement session or both. Students who returned only their regional module questionnaire are not represented in the database.

TPART

    TPART is the final participation indicator for each teacher. It is set to "3" (questionnaire returned completed) for all records in the database.

ITRM

    This variable indicates whether a student has been assigned to participate in any of the regional modules. Code "0" means that a student's country did not participate in a regional module. The variable is set to "1" for students from countries participating in the European module, to "2" for students from countries participating in the Latin American module, and to "3" for students from Asian module countries.

---

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

ITMODE

ITMODE represents the administration mode of the teachers' or principals' questionnaire in the data source: This variable indicates whether the teacher or principal completed the questionnaire on-line (code "1") or on paper (code "2").

ITMODEW

ITMODEW is an indicator for the questionnaire mode of teachers and principals originally documented in the within-school sampling process. The variable has been made consistent with ITMODE and the same codes are used.

ITMODET

ITMODET is an indicator for the default questionnaire mode for teachers at the school level. The variable is set to "1" if the default mode for teachers was originally set to online and it is set to "2" if the default mode was originally set to paper.

ITMODEO

ITMODEO represents the original default questionnaire mode for teachers. The variable is set to "1" if the default mode was originally set to online and it is set to "2" if the default mode was originally set to paper.

INICS09

INICS09 is an indicator for the inclusion of a school, student, or teacher in the database. It is set to "1" for all records.

Table 2.7 shows in which data files the various tracking variables are located.

Table 2.7 Location of Tracking Variables in the ICCS 2009 International Database

Tracking Variables	Data File Types						
	ISA JSA	ISG JSG	ITG	ICG	ISE JSE	ISL	ISS
ITADMINI		●					
ITDATEM		●					
ITDATEY		●					
ITEXCLUD	●	●			●	●	●
ITPART1		●					
ITPART2		●					
ITPARTR		●					
CPART				●			
SPART	●	●			●	●	●
TPART			●				
ITRM		●					
ITMODE			●	●			
ITMODEW			●	●			
ITMODET				●			
ITMODEO			●				
INICS09	●	●	●	●	●	●	●

Database Creation Variables

Information about the version number of the ICCS 2009 International Database and its creation at the IEA DPC is contained in the database creation variables. They are included in all data files.

VERSION

Throughout the data processing process a system of version numbers for the database was used. The version number of the ICCS 2009 final database is "32".

DPCDATE

This code indicates the date on which the data file was produced at the IEA DPC.

2.2.9 ICCS 2009 National Context Survey Data File

The National Context Survey data file contains the responses of participating countries that National Research Coordinators provided to the ICCS 2009 National Context Questionnaire. The National Context Survey was designed to systematically collect relevant data on the structure of the education system, education policy, civic and citizenship education, teacher qualifications for civic and citizenship education, and the extent of current debates and reforms in this area. The survey also collected data on processes at the national level regarding assessment of and quality assurance in civic and citizenship education and in school curriculum approaches. The National Context Questionnaire was administered online using the IEA Survey System developed at the IEA Data Processing and Research Center (IEA DPC).

The National Context Survey data file (NCQICSC2.SAV) is available in SPSS format and contains data for all 38 countries participating in ICCS 2009.

2.3 ICCS 2009 Codebook Files

All information related to the structure of the ICCS 2009 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 national context survey data file.

The naming convention for codebook files is as follows:

- The first three characters of the filename are identical to those in the file names shown in Figure 2.2.
- The next three characters identify the files as ICCS codebooks and are always "ICS".
- The seventh and eighth characters are always "C2" to indicate the ICCS 2009 study cycle.
- The codebook files are provided in two different formats, indicated by the three-character file extension. The extension .SDB stands for standard dBase format and those files can be used with the WinDEM software provided by the IEA DPC to countries for data capture. The extension .PDF identifies the codebooks files in Adobe PDF format.

Codebook files in standard dBase format may be read using Microsoft Excel, or any standard database or spreadsheet program. The codebook files describe the contents and structure of the ICCS 2009 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.

2.4 ICCS 2009 Program Files

The ICCS 2009 International Database contains SPSS syntax files to perform the variable recodes required for the proper execution of example analyses using the IEA IDB Analyzer. They are described in Chapter 4 of this ICCS 2009 IDB User Guide. There are additional SPSS syntax files available in the database to compute derived variables. They are referred to in Supplement 3 of this ICCS 2009 IDB User Guide.

The ICCS 2009 International Database also includes a number of SAS programs and macros designed to facilitate the manipulation of the ICCS 2009 data files and conduct proper statistical analyses, taking into account the jackknife algorithm and the presence of plausible values. These are described in Chapter 5 of this ICCS 2009 IDB User Guide.

Both SPSS and SAS programs are part of the ICCS 2009 International Database and are available on the IEA study data webpage at http://rms.iea-dpc.org/.

Weights and Variance Estimation

3.1 Overview

This chapter gives a brief introduction to the use of weighting and variance estimation variables in the International Civic and Citizenship Education Study (ICCS) 2009. The names and locations of these variables in the ICCS 2009 International Database are described and their specific roles in student, teacher, and school analyses are explained. Examples demonstrating the importance of using the appropriate weighting and variance estimation techniques are given.

3.2 Sampling Weights

All data in the ICCS 2009 International Database are derived from randomly drawn samples of schools, students, and teachers. Of course, the study results should be valid not only for the sampled units, but for the entire educational system that participated in the ICCS 2009 study. In order to make correct inferences about the educational systems, the complex nature of the sampling design implemented in ICCS 2009 needs to be taken into account. Details about the sampling design are reported in Chapter 6 of the *ICCS 2009 Technical Report* (Schulz et al., forthcoming).

The ICCS 2009 sampling design called for different selection probabilities at the school level and at the within-school sampling level. Sampling weights reflect and compensate the disproportional selection probabilities of the schools, the students, and the teachers. If any unit of response had a small selection probability, a large weight would compensate, and vice versa. Given that some sampled schools, students, and teachers refused to participate in ICCS 2009, it was necessary to adjust the sampling weights for the sample size loss. Thus, the sampling weights were multiplied by non-response adjustments. The final (total) weights are the product of weight factors and adjustment factors that reflect the selection probabilities and the non-response patterns at all levels of analysis. Details about weighting and adjustments are reported in the weighting chapter of the *ICCS 2009 Technical Report* (Schulz et al., forthcoming).

3.2.1 Weight Variables in the ICCS 2009 International Database

Each record in the ICCS 2009 International Database contains data for one or more variables that concern weighting. The last character of the variable name indicates the data type (S = Student, T = Teacher, C = School). The weights and weighting factors differ depending on the type of data. Only the value of the school base weight (variable WGTFAC1) is identical in all three types of datasets, since it does not depend on the data type.

Student Weight Variables

Table 3.1 shows the student weight variables that are part of the ICCS 2009 International Database.

Table 3.1 Student Weight Variables

Variable	Description	Source Files
TOTWGTS	Total student weight	ISA, ISE, ISG, ISL, ISS, JSA, JSE, JSG
SENWGTS	Senate student weight	ISG, ISA, ISE, ISL, ISS, JSA, JSE, JSG
WGTFAC1	School base weight	ISG
WGTADJ1S	School non-participation adjustment for the student survey	ISG
WGTFAC2S	Class base weight	ISG
WGTADJ2S	Class non-participation adjustment	ISG
WGTADJ3S	Student non-participation adjustment	ISG

Teacher Weight Variables

Table 3.2 shows the weight variables in the teacher data files in the ICCS 2009 International Database.

Table 3.2 Weight Variables in Teacher Data Files

Variable	Description	Source Files
TOTWGTT	Total teacher weight	ITG
SENWGTT	Senate teacher weight	ITG
WGTFAC1	School base weight	ITG
WGTADJ1T	School non-participation adjustment for the teacher survey	ITG
WGTFAC2T	Teacher base weight	ITG
WGTADJ2T	Teacher non-participation adjustment	ITG
WGTADJ3T	Teacher multiplicity adjustment	ITG

School Weight Variables

Table 3.3 shows the weight variables in the school data files of the ICCS 2009 International Database.

Table 3.3 Weight Variables in School Data Files

Variable	Description	Source Files
TOTWGTC	Total school weight	ICG
WGTFAC1	School base weight	ICG
WGTFAC1C	School non-participation adjustment for school-level data analyses	ICG

3.2.2 Selecting the Appropriate Weight Variable

For analyzing the ICCS 2009 data, it is important that the appropriate weights are selected. The decision which weight to choose depends on the type of data used for analysis, the level of analysis and the number of countries involved.

Single-Level Analysis

For analyses concerning one data type only, different weights must be applied depending on the type of data:

- For student-level analyses, TOTWGTS should be used
- For teacher-level analyses, TOTWGTT should be used
- For school-level analyses, TOTWGTC should be used.

When the IEA IDB Analyzer is used for data analysis, the software automatically selects these variables.

Please note that ICCS 2009 is conceptually a student and teacher survey, and was not designed as a school survey. Although it is possible to perform school-level analyses, the sampling precision of the estimates is expected to be poor and all statements concerning school-level data alone can be made only with a high degree of uncertainty.

Merging Files from Different Levels

If researchers plan to analyze data from more than one level and plan to merge data of different data types, they must choose the correct weight carefully.

- The variable TOTWGTS should be used for analyzing student data with added school data. This type of analysis of disaggregated data is straightforward with the IEA IDB Analyzer. The software merges school-level data to the student data and selects the correct weight automatically. This way, school information becomes an attribute of the student

and the user can analyze information from both files. A sample research question could be: "What is the percentage of students studying at schools with a female headmaster?"

- Analyzing combined teacher data and school data should be performed in the same way; TOTWGTT is the variable of choice. As for student data, the IEA IDB Analyzer takes care of the correct selection. In this type of analysis, school information becomes an attribute of the teacher. A sample research question could be: "What is the percentage of teachers working at schools with a female headmaster?"

- If student or teacher information is regarded as an attribute of school information, this cannot be handled easily with the IEA IDB Analyzer. The researcher must use other software (e.g., SPSS or SAS) to aggregate the student or teacher data and to merge the resulting information with the school file.

- For aggregating student data within schools, within-school weights (which are the product of class and student-level weight factors WGTFAC2S x WGTADJ2S x WGTADJ3S), should be used. However, for all ICCS 2009 countries except Liechtenstein and Luxembourg, all students in the same school share the same within-school weight. For this reason, it is possible not to use any weights at all for aggregating data within the schools of the remaining countries.

- "Within-school teacher weights" (defined as the product of the teacher-level weight factors WGTFAC2T x WGTADJ2T x WGTADJ3T) should be used for aggregating teacher data within the school. Omitting this weighting step will lead to incorrect results for any ICCS 2009 country.

- After aggregation, the student or teacher file can be merged with the school file (with IDSCHOOL as the key variable). When this step is completed, the data can be processed further with the IEA IDB Analyzer. TOTWGTC should be used for school-level data analysis. A sample question is: "What is the percentage of schools in which more than 50% of the tested students do not speak the language of the test at home?"

It is neither possible nor meaningful to combine files of student and teacher data directly. These two groups constitute separate target populations. A sampled student may never have been taught by a sampled teacher, and a sampled teacher may never have taught a sampled student. However, it is possible to aggregate teacher data at the school level and then treat the result as a contextual attribute of the student data. Similarly, it is possible to aggregate student data at the school level and then treat the result as an attribute of the teacher data.

Multi-level Analysis

Working with aggregated or disaggregated data poses some methodological problems (for details, see Snijders & Bosker, 1999). In order to use the full potential of the data, it is possible to perform multi-level analyses with specialized software packages (e.g., HLM or Mplus). For this type of analysis, users must compute the appropriate weights themselves.

- At Level 1 (student level), the analyst should apply a "within-school student weight" as the product of the class and student level weight factors (WGTFAC2S x WGTADJ2S x WGTADJ3S). If the teachers constitute Level 1, the analyst should apply a "within-school teacher weight" as the product of the teacher level weight factors (WGTFAC2T x WGTADJ2T x WGTADJ3T).

- At Level 2 (school level), the user should calculate a "school weight". For student data analysis, this is the product of the variables WGTFAC1 and WGTADJ1S; for teacher level analysis, this is the product of WGTFAC1 and WGTADJ1T. Users should ensure that the software used for multi-level analysis normalizes the weights, that is, makes the sum of the weights equal to the sample size. Users should not use the variable TOTWGTC from

the school files, as non-response adjustments made to school questionnaire data may make these values slightly different from the correct ones.

Given the small number of schools in Liechtenstein and Luxembourg, multi-level analyses are not recommended for these countries.

Analyses of Groups of Countries

Thus far, the discussion has focused on analysis of data from one country at a time. However, all the above statements also hold true when more than one country is analyzed. Some caution must be exercised when international averages are calculated, however. If an international average is computed directly using TOTWGTS, TOTWGTT, or TOTWGTC, larger countries will contribute more to the average than smaller countries, which may not be the intention of the researcher.

Instead of performing weighted analyses across groups of countries, users must conduct weighted analyses separately for each country and calculate an average of the results afterwards. This is true regardless of whether single-level data, aggregated or disaggregated data, or multi-level data files are used for analysis.

Users of the IEA IDB Analyzer do not need to worry about the issue of international averages, since the software performs the correct calculations automatically. For calculating an international mean, the IEA IDB Analyzer first calculates national means using the TOTWGT variables and then averages the results over the countries that contribute to the international mean.

Some researchers familiar with IEA data prefer using senate weights for calculating international averages. Since the request to include senate-weight variables in the ICCS 2009 International Database was made repeatedly, SENWGT variables are included in the student and the teacher data. Countries that did not meet the international sampling requirements did not receive a SENWGT, and should not contribute to an international average.

Please note that ICCS 2009 does not recommend using SENWGT variables for calculating international averages. If sub-groups of the population are analyzed (e.g., boys and girls), the use of senate weights may yield incorrect results. Also, if data are missing from a variable of analysis, SENWGT will give incorrect results.

3.2.3 Example for Analyzing Weighted Data

Not using weights in data analysis can lead to severely biased results. The following example illustrates the importance of using weights when conducting research with ICCS 2009 data.

A researcher may be interested in the average civic knowledge in Chile (variables PVCIV01-05 in ISG file). Using unweighted data (e.g., in SPSS), the mean of each plausible value is calculated and an average of the five values is calculated. Figure 3.1 shows that this average score is 493.83.

Figure 3.1 Example of Unweighted Analysis in SPSS

Descriptive Statistics

	N	Mean
CIVIC KNOWLEDGE - 1ST PV	5192	493,8852
CIVIC KNOWLEDGE - 2ND PV	5192	494,1029
CIVIC KNOWLEDGE - 3RD PV	5192	493,2203
CIVIC KNOWLEDGE - 4TH PV	5192	493,5182
CIVIC KNOWLEDGE - 5TH PV	5192	494,4289
Valid N (listwise)	5192	

⟶ average: 493.83

But using weighted data with the IEA IDB Analyzer, as in Figure 3.2, shows that in Chile, the correct estimate for civic knowledge is actually only 483.03.

Figure 3.2 Example of Weighted Analysis Using the IEA IDB Analyzer

N of Cases	Sum of TOTWGTT	Percent	Percent (s.e.)	TAGE (Mean)	TAGE (s.e.)
5192	258422	100,00	,00	483,03	3,54

The large difference between the unweighted and the weighted result can be explained by the ICCS 2009 sampling design for Chile. The proportion of students from private schools in the ICCS 2009 school sample is higher than their proportion in the student population. The sample was selected this way intentionally in order to allow the Chilean researchers to make more accurate statements about this group of students. In order to balance out the disproportionate sample allocation, students from private schools were assigned smaller weights than students from the remaining school types. Since on average students from private schools perform better than students from other school types, omitting weights leads to an over-estimate of the students' performance in Chile. The sampling weights compensate for that disproportional school sample allocation.

3.3   Variance Estimation

Since all information in ICCS 2009 is based upon sample data, analysts should report the precision of the population estimates. Due to the complex sampling design used in ICCS 2009, it is not possible to calculate standard errors or to perform significance tests with standard software packages. While these programs implicitly assume that the data is derived from a simple random sample, the ICCS 2009 student and teacher data come from a two-stage stratified cluster sample (each school being regarded as a "cluster" of students or teachers). Any method for estimating sampling variance must take this difference into account.

The ICCS 2009 International Database contains variables that allow for the use of a variance estimation method known as the Jackknife Repeated Replication (JRR). These variables are referred to as "jackknife zones" and as "jackknife replicates". The JRR method was implemented in the IEA IDB Analyzer software (for details about the JRR technique used in ICCS 2009, please refer to Chapter 13 of the *ICCS 2009 Technical Report* (Schulz et al., forthcoming).

3.3.1 Variance Estimation Variables in the ICCS 2009 International Database

Table 3.4 shows student-level variance estimation variables (or "jackknife variables") that are included in the ICCS 2009 International Database.

Table 3.4 Student-level Variance Estimation Variables

Variable	Description	Source Files
JKZONES	Jackknife zone to which students of a school are assigned	ISA, ISE, ISG, ISL, ISS, JSA, JSE, JSG
JKREPS	Jackknife replicate to which students of a school are assigned	ISA, ISE, ISG, ISL, ISS, JSA, JSE, JSG

Table 3.5 shows the jackknife variables included for teachers.

Table 3.5 Teacher-level Jackknife Variables

Variable	Description	Source Files
JKZONET	Jackknife zone to which teachers of a school are assigned	ITG
JKREPT	Jackknife replicate to which teachers of a school are assigned	ITG

Table 3.6 shows the school-level jackknife variables found in the ICCS 2009 International Database.

Table 3.6 School-level Jackknife Variables

Variable	Description	Source Files
JKZONEC	Jackknife zone to which a school is assigned for school- level data analysis	ICG
JKREPC	Jackknife replicate to which a school is assigned for school- level data analysis	ICG

3.3.2 Selecting the Appropriate Variance Estimation Variables

Different variance estimation variables must be applied depending on the type of data:

- For all student-level analyses, JKZONES and JKREPS should be used
- For all teacher-level analyses, JKZONET and JKREPT should be used
- For all school-level analyses, JKZONEC and JKREPC should be used.

Even for the same school, the variables at different levels of analysis can differ from each other and thus are not interchangeable. Just as with weights, researchers should ensure the correct jackknife variables are chosen when working with aggregated datasets. The level of analysis (student, teacher, or school) determines which variable to choose.

When calculations are performed with the IEA IDB Analyzer, the correct variables will be selected automatically. However, researchers may choose to use specialized software for types of data analysis that go beyond the range of the IEA IDB Analyzer's capabilities. In this case, researchers have to specify the jackknife variables according to the requirements of the software. Usually, "-zone" variables must be specified as "stratum" or "strata" variables, while the "rep" variables are commonly referred to as "cluster" variables.

3.3.3 Example for Variance Estimation

Not using the jackknife variables in data analysis will lead to incorrect estimations of sampling precision. The following example illustrates the importance of using the JRR technique for research and analysis with ICCS 2009 data.

A researcher may be interested in the average teacher age (variable TAGE) in Chile. Using SPSS, the researcher finds that the (weighted) average teacher age is about 44 years and the standard error seems to be close to 0.05 years (see Figure 3.3).

Figure 3.3 Example of Incorrect Variance Estimation in SPSS

Descriptive Statistics

	N	Mean	
	Statistic	Statistic	Std. Error
TEACHER'S AGE	43585	43,98	,052
Valid N (listwise)	43585		

But using the JRR technique with the IEA IDB Analyzer, we find that the correct estimate for the standard error is more than seven times as large (see Figure 3.4).

Figure 3.4 Example of Correct Variance Estimation using the IEA IDB Analyzer

N of Cases	Sum of TOTWGTT	Percent	Percent (s.e.)	TAGE (Mean)	TAGE (s.e.)
1744	43585	100,00	,00	43,98	,40

The standard methods of the SPSS base version cannot handle weights correctly for sampling variance estimation, nor can it take the clustered data structure into account. This means that not only standard errors, but also all analyses that contain significance tests will be incorrect unless specialized software is used.

Analyzing the ICCS 2009 Data Using the IEA IDB Analyzer

4.1 Overview

This chapter describes the use of the IEA International Database Analyzer software (IEA, 2010) for analyzing the International Civic and Citizenship Education Study (ICCS) 2009 international data files. Sample analyses will illustrate the capabilities of the IEA IDB Analyzer to compute a variety of statistics, including percentages of students in specified subgroups, average civic knowledge in those subgroups, correlations, regression coefficients, and percentages of students reaching certain proficiency levels. The examples use student, teacher, and school background data to replicate some of the ICCS 2009 results included in the *ICCS 2009 International Report* (Schulz, Ainley, Fraillon, Kerr, & Losito, 2010b), as well as other useful analyses for investigating policy-relevant research questions.

The IEA IDB Analyzer uses the SPSS data files from the ICCS 2009 International Database. Additionally, an SPSS syntax file (Syntax_ISGALLC2.SPS) will be needed to recode certain variables used in the example analyses presented later in this chapter.

Developed by the IEA Data Processing and Research Center (IEA DPC), the IEA International Database Analyzer (IEA IDB Analyzer) is software that uses the Statistical Package for the Social Sciences (SPSS, 2010) as an engine for performing computations using IEA data. The IEA IDB Analyzer creates syntax files reflecting the settings users can define by means of a graphical user interface. The syntax files produced can be used for combining SPSS data files from IEA's large-scale assessments and conduct analyses using SPSS without actually writing programming code. The SPSS syntax generated by the IEA IDB Analyzer takes into account information from the sampling design when computing statistics and the corresponding standard errors. In addition, the SPSS syntax generated makes use of the plausible values for calculating estimates of achievement scores and their corresponding standard errors, combining both sampling and imputation variance.

The IEA IDB Analyzer consists of two modules: the merge module and the analysis module, which are executed as independent applications. The merge module is used to create analysis datasets by combining data files of different types or from different countries, and selecting subsets of variables for analysis. The analysis module provides procedures for computing various statistics and their standard errors for variables of the user's interest. These procedures can be applied for a country as well as for specific subgroups within a country. Both modules may be accessed using the START menu in Windows:

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

4.2 Scoring the Individual ICCS 2009 Achievement Items Using SPSS

The current section describes how the original answers from students can be scored. The original answers on multiple-choice items are located in the achievement data files (ISA/JSA). The ICCS 2009 data already contains variables for each student's civic knowledge achievement as a set of plausible values. Those are the preferred scores to be used for analysis. The current section describes how to score the multiple-choice items in case item-level analysis is desired. Students' responses to the individual multiple-choice items need to be recoded into score points according to a scheme that specifies the correct option for each one of these items.

Two types of items were administered as part of the ICCS 2009 assessment. There were multiple-choice items, where students were asked to select one out of four options as the correct response. Numbers 1 through 4 represent response Options A through D, respectively, in the achievement data files (ISA/JSA). There also were constructed-response items, in which students

were asked to write a text response to a question, rather than choosing an answer from a list of options. Constructed-response items were worth a total of zero, one, or two score points. Scorers from the national centers were trained to use the scoring guides to score the answers to these questions. The numbers 0 through 2 are used to represent the scored responses to these items and also represent their point values: "0" for an incorrect response, "1" for a partially correct response, and "2" for a correct response. For both types of items, special codes are set aside to represent missing data either as "not administered", "omitted", "not reached" or "invalid". Responses to multiple-choice items must be converted to their appropriate score levels ("1" for correct, and "0" for incorrect and missing responses), as must responses coded to the special missing codes, in order to carry out specific item-level analyses. Database users can get an overview of the correct responses for this item type from the ISA/JSA codebooks and data files, in which the correct response option is marked with an asterisk "*". Constructed-response items were scored in advance by the national centers in each country.

The ICCS 2009 International Database includes an SPSS program (ISASCRC2.SPS) that allows researchers to recode the items from the achievement data files to their score level. The program consists of a macro called "SCOREIT" and a syntax line to call this macro so that all the items in the specified data files are scored. The macro will convert the response option codes for multiple-choice items to dichotomous score levels (0 or 1) based on each item's scoring key. It will also convert the special missing codes as either incorrect (0) or missing. By default, the "not administered" response code is left as missing and the "omitted" and "not reached" response codes as incorrect. These default settings can be modified within the "SCOREIT" macro, depending on the requirements of the researcher's item-level analyses. For example, "not reached" responses were treated as missing for the purpose of calibrating the ICCS 2009 items, whereas they were treated as incorrect when scoring the results of individual countries and deriving achievement scores for students. To use the SCOREIT macro, researchers will need to adapt the program code in the ISASCRC2.SPS program using the following steps:

1) Open the SPSS program file ISASCRC2.SPS

2) Specify the path where the SPSS data files are located in the "LET! LIBDAT!" statement

3) List all the countries of interest in the parameter "COUNTRY". By default, all ICCS 2009 countries are listed

4) Submit the edited code for processing.

The program recodes the items and saves the results in SPSS data files that consist of "ISC" instead of "ISA" as the first three characters of the file name. To treat "not reached "responses as missing rather than incorrect, replace the following statement (which appears twice in the program):

```
( !NR = 0 )
```

with this statement:

```
( !NR = SYSMIS )
```

Figure 4.1 shows a condensed version of the SPSS program that scores the international achievement items.

Figure 4.1 Example of ISASCRC2 SPSS Program for Converting Item Response Codes to Their Score Level

```
DEFINE SCOREIT (TYPE = !CHAREND('/') /
ITEM = !CHAREND('/') /
RIGHT = !CHAREND('/') /
NR = !CHAREND('/') /
NA = !CHAREND('/') /
OM = !CHAREND('/') /
OTHER = !CHAREND('/') ) .

. . .
!ENDDEFINE .
DEFINE DOIT (COUNTRY = !CHAREND('/') ) .
!LET !LIBDAT = !UNQUOTE("C:\ICCS2009\Data\SPSS_Data\") .

. . .
SCOREIT TYPE = MC / ITEM = < List of multiple-choice items where A is correct > .
SCOREIT TYPE = MC / ITEM = < List of multiple-choice items where B is correct > .
SCOREIT TYPE = MC / ITEM = < List of multiple-choice items where C is correct > .
SCOREIT TYPE = MC / ITEM = < List of multiple-choice items where D is correct > .
SCOREIT TYPE = CR / ITEM = < List of constructed-response items > .
!ENDDEFINE .
DOIT COUNTRY = < List of ICCS 2009 countries > .
```

The achievement items available in the European Module data files (ISE/JSE) and the Latin American Module data files (ISL) may be scored using the procedure outlined above in scoring the ICCS 2009 international achievement items.

For the European Module data, the ICCS 2009 International Database includes an SPSS program (ISESCRC2.SPS) that allows researchers to recode the achievement items included in the European Module into their score level. The results will be saved in SPSS data files that consist of "ESC" instead of "ISE" as the first three characters of the file name.

For the Latin American Module data, the ICCS 2009 International Database includes another SPSS program (ISLSCRC2.SPS) that allows researchers to recode the achievement items included in the Latin American Module into their score level. The results will be saved in SPSS data files that consist of "LSC" instead of "ISL" as the first three characters of the file name.

Please note that in both the European Module and the Latin American Module questionnaire there were only multiple-choice items and no constructed-response items used. Both recoding programs (ISESCRC2.SPS and ISLSCRC2.SPS) refer only to multiple-choice item recoding.

4.3   Merging Files with the IEA IDB Analyzer

The ICCS 2009 data files are disseminated separately for each country and by file type. In addition to allowing users to combine data from the same file type from more than one country for cross-country analyses, the merge module allows for combinations of data from different levels, for example, merging student and school data into single SPSS dataset. This will allow analysis of the student data in relation to certain characteristics of the school using the IEA IDB Analyzer Analysis Module later.

Table 4.1 provides an overview of possible combinations of data file types that the ICCS 2009 design allows to be merged at different levels. The grey-shaded cells on the diagonal represent the merges for the same file type. As the table shows, the school background file can be merged with every other file type. Teacher background files can be merged only with themselves (i.e., teacher background files from different countries) and school background files. Merging teacher background files with student files (background, achievement, and regional module data files) is not possible. The reason for this lies in the study's sample design—the ICCS 2009 teacher sample includes all teachers from the students' target grade. The sample includes both teachers who did and who did not teach the sampled students, so that teacher data cannot be directly linked to student data.[8] Also, the user will not be able to merge data from different regional modules (European, Latin American, and Asian) because students only answered the module questionnaire designed for the region in which they live (e.g., European students only completed the European module questionnaire). Finally, data from different grade levels cannot be merged.

Only one set of school and teacher background data files exist per school. The first character of school and teacher data files begin is always "I" because the school principal is the same person and the sampled teachers are the same, regardless of the population sampled. School background data can be merged with both target and additional grade student data. There was no teacher sample for the additional grade.

When merging a regional module file with another file type, IEA IDB Analyzer will always display a warning for files not found for specific countries. In general it will list all countries that did not administer the specific regional module questionnaire—European countries do not have Latin American and Asian Regional Modules, for example. The other warning indicates that a country belongs to a region, but does not use the corresponding regional module. For example, there are European countries that decided not to use the European module.

Table 4.1 Possible Merges Between Different File Types in ICCS 2009

	ISA/JSA	ISG/JSG	ITG	ICG	ISE/JSE	ISL/JSL	ISS/JSS
ISA/JSA	X	X		X	X	X	X
ISG/JSG	X	X		X	X	X	X
ITG			X	X			
ICG	X	X	X	X	X	X	X
ISE/JSE	X	X		X	X		
ISL/JSL	X	X		X		X	
ISS/JSS	X	X		X			X

4.3.1 Merging Data from Different Countries

The following examples on merging ICCS 2009 data files use target grade data (usually Grade 8, file names starting with "I"). Merging the additional grade data files (Grade 9, filenames starting with "J") follows exactly the same procedure.

Merging the files from different countries on a single level is simple. The same steps apply for merging school background, teacher background, or any other file types. The following example will create an SPSS data file with student background data from all countries:

---

8 For more details on the ICCS 2009 sampling strategy and procedures see Chapter 2 of the *ICCS 2009 Technical Report* (Schulz et al., forthcoming).

1) Open the merge module of the IEA IDB Analyzer (Start ⇨ All Programs ⇨ IEA ⇨ IDB Analyzer ⇨ Merge Module).

2) In the **Select Data Source Directory** field, browse to the folder where all SPSS data files are located. For example, in Figure 4.2, all SPSS data files are located in the "C:\ICCS2009\Data" folder. The program will automatically recognize and complete the **Study Type**, **Survey Type**, and **Grade Type** 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, study cycle, or from more than one grade, the IEA IDB Analyzer will prompt users to select files from the desired study and grade for analysis. If the data in the folder is only from one study, cycle, and grade, IEA IDB Analyzer will populate these fields automatically. In Figure 4.2, the ICCS 2009 Grade 8 is selected.

3) Select the countries of interest from the **Available Participants** list. To select multiple countries, hold the CTRL key on the keyboard when selecting the countries, and then press the single-arrow button ▶ to move them in the **Selected Participants** list on the right. In the current example all countries participating in the ICCS 2009 assessment are selected for merging simply by pressing the double-arrow button ▶▶. Figure 4.2 shows the IEA IDB Analyzer screen after selecting all countries for merging.

Figure 4.2 IEA IDB Analyzer Merge Module: Selecting Countries
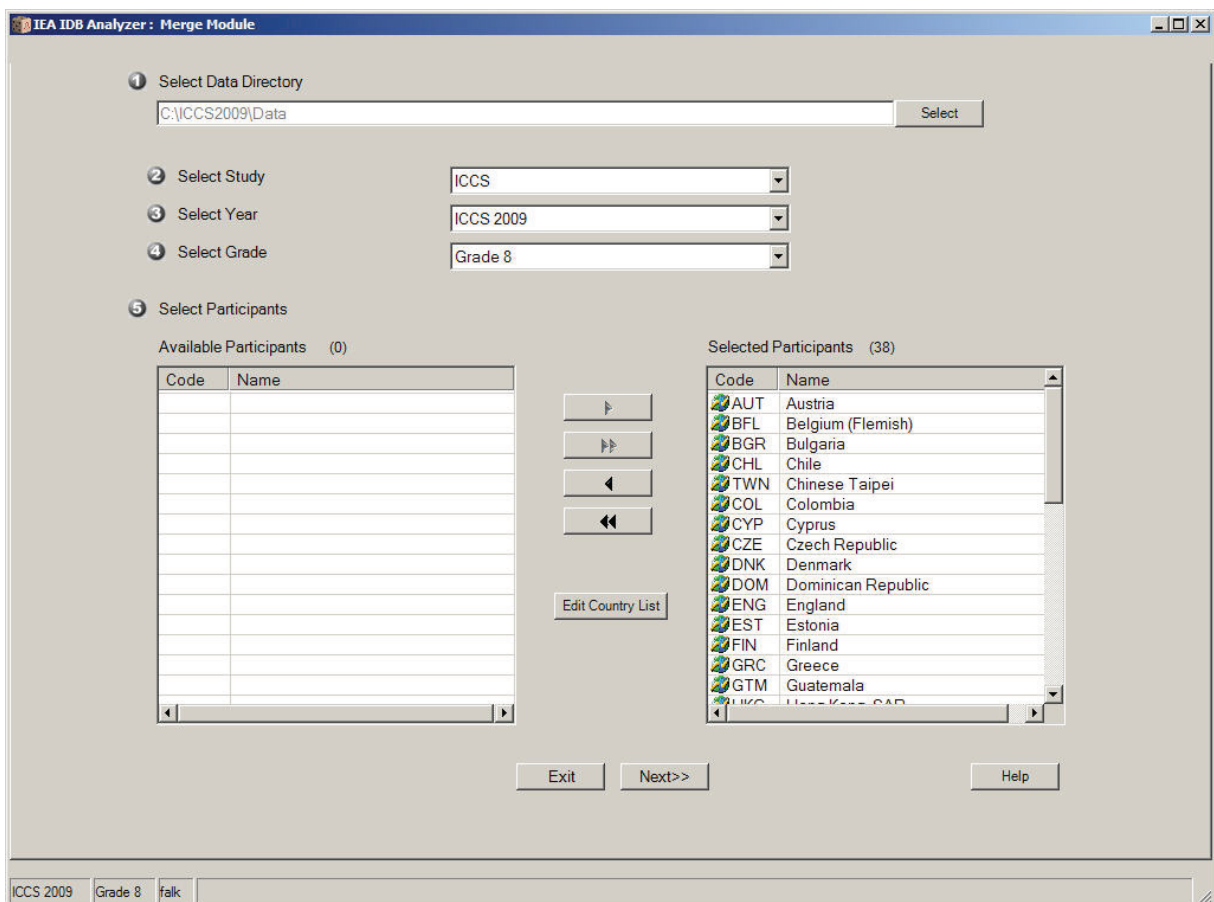
4) Press the **Next>>** button to proceed. The software will open the second window of the merge module, as shown in Figure 4.3, to select the file types and variables to be included in the merged data file.

5) Select the file types for merging by checking the appropriate boxes to the left of the window. In the current example only the International **Student Questionnaire File** is selected (see Figure 4.3).

6) Select the desired variables from the list of background variables available in the left panel. You can select and move separate variables from the **Available Variables** to the **Selected Variables** list by holding the Control key, pressing the left mouse button and then clicking the arrow button ▶. If you want to select all variables and move them in the **Selected Variables** list, use the double-arrow key ▶▶. In our example all student variables will be used for merging. Please note that the IEA IDB Analyzer automatically selects all achievement scores, identification, and sampling variables.

7) Specify the desired name of the merged data file and the folder where it will be stored in the **Output Files** field. The IEA IDB Analyzer 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 Figure 4.3, the data file ISGALLC2.SAV and the syntax file ISGALLC2.SPS are stored in the "C:\ICCS2009\Work" folder. The merged data file will contain student background data with the variables shown in the **Selected Variables** panel to the right.

8) Click on the **Start SPSS** button. The IEA IDB Analyzer will give a warning if it is about to overwrite an existing file with the same name in the specified folder. The IEA IDB Analyzer creates the syntax file with the specified name, stores it in the specified folder and opens it in an SPSS Syntax Editor window (Figure 4.4) ready for execution. The syntax file must be executed by opening the **Run** menu of SPSS and clicking the **All** option.

Figure 4.3 IEA IDB Analyzer Merge Module: Selecting File Types and Variables
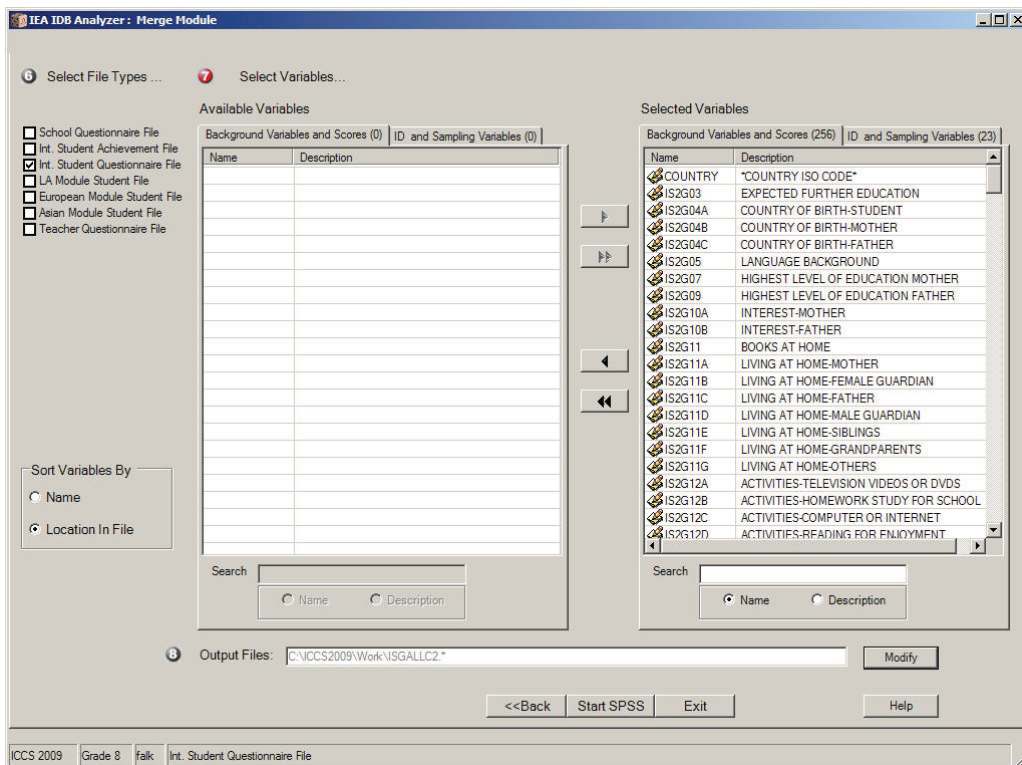
Figure 4.4 SPSS Syntax Editor with Merge Syntax Produced by the IEA IDB Analyzer Merge Module
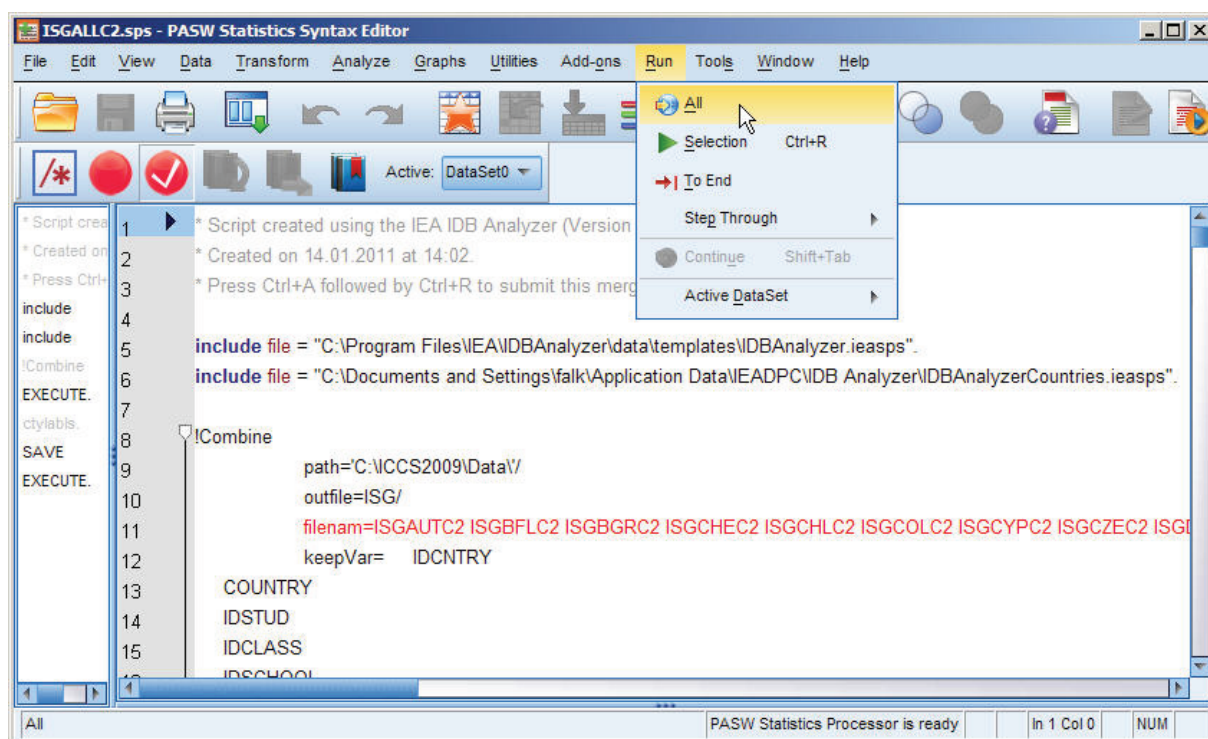
Users should check the resulting SPSS output file for warnings that might indicate that the merge process was not performed correctly.

4.3.2 Merging Student Background and Regional Module Files

Student background files contain contextual variables related to students' background characteristics, perceptions, and behaviors. The regional modules files contain variables addressing specific regional issues and aspects of civic and citizenship education. As the use of the regional modules instrument was optional, not all countries participated. Some European countries, for example, decided not to use the European Module.

Merging the student background data files with the regional module files can give researchers the chance to enrich the student-level analyses with variables that are specific for certain region of the world.

To merge student background data with regional module data, perform Steps 1 to 4 as described in Section 4.3.1. Then, 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 on the **International Student Questionnaire 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. The plausible values, ID, and sampling variables are selected automatically and listed in the right panel.

2) Select the variables of interest and press the right arrow button ▶ to move these variables into the right panel.

3) Select either the **Latin American, European, or Asian Module Student File**. Based on your country selection, the IEA IDB Analyzer might display a warning that certain countries do not have data for the selected Regional Module. Close the warning message and select the variables of interest from the Background Variables and Scores panel in the same manner as described in Steps 1 and 2.

4) Specify the desired name of the merged data file and the folder where it will be stored in the **Output Files** field. The IEA IDB Analyzer will create an SPSS syntax file (*.SPS) of the same name and in the same folder with the code necessary to perform the merge.

5) Click on 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** option.

4.3.3 Merging School and Student Data Files

The ICCS 2009 school samples were designed to optimize the student samples and the student-level estimates. It is preferable to analyze school variables as attributes of students, rather than as elements in their own right. However, the school samples are representative probability samples of schools within each participating country. Therefore it is possible to compute weighted numbers of schools with particular characteristics for providing reasonable estimates of percentages and means across the populations of schools in each country.

To merge the school and student background data files, select both the **School Questionnaire File** and **Student Questionnaire File** types. The variables of interest to be included in the merged data file need to be selected separately by file type using the same set of instructions as described in Section 4.3.2. The ID and sampling variables will be selected automatically. Note that when merging student and school data, only the total student weight (TOTWGTS) variable will be included in the merged file, but not the total school weight (TOTWGTC). An analysis using school variables weighting the data using the total student weight will not allow the researcher to make inferences for the schools themselves. The interpretation of results will be about students who study in schools with certain characteristics. For example, if we use merged student and school data and use the principals' gender as a grouping variable, the total student weight will be selected as weighting variable. The results then will be interpreted as percentages of students who study in schools where the school principal is male or female, for example, "In Austria 67% of students study in schools with male principals, and 33% in schools with female ones."

4.3.4 Merging School and Teacher Data Files

Merging the school and teacher data files follows the same procedure as merging the school and student data files. School data will be disaggregated to the teacher level by adding the respective school level variables to each teacher record. To merge teacher background and school background data files, perform Steps 1 to 4 as described in Section 4.3.1. Then, 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 on the **Teacher Questionnaire File** type so that it appears checked and highlighted. The ID and sampling variables are selected automatically and are listed in the right panel.

2) Select the variables of interest and press the right arrow button ▶ to move these variables into the right panel.

3) Click on the **School Questionnaire File** type. Based on the country selection, the IEA IDB Analyzer might display a warning that certain countries do not have data for the selected Regional Module. Close the warning message and select the variables of interest from the **Background Variables and Scores** panel in the same manner as described in Steps 1 and 2.

4) Specify the desired name of the merged data file and the folder where it will be stored in the **Output Files** field. The IEA IDB Analyzer will create an SPSS syntax file (*.SPS) of the same name and in the same folder with the code necessary to perform the merge.

5) Click on 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** option.

4.3.5 Merging Data Files for the Sample Analyses

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

ISGALLC2.SAV   Merge the student background (ISG) data files for all countries

ITGALLC2.SAV   Merge the teacher background (ITG) data files for all countries

ISG&ICGALLC2.SAV   Merge the school background (ICG) and student background (ISG) data files for all countries

4.4 Performing Analyses with the IEA IDB Analyzer

The analysis module of the IEA IDB Analyzer is used to analyze any files created using the merge module. The analysis module can perform the following statistical procedures:

Percentages and Means
  Computes percentages, means, and standard deviations for selected variables by subgroups defined by grouping variable(s)

Percentages Only
  Computes percentages by subgroups defined by grouping variable(s)

Regression
  Computes regression coefficients for selected variables to predict a dependent variable by subgroups defined by grouping variable(s)

Benchmarks
  Computes percentages of students meeting a set of user-specified achievement proficiency levels by subgroups defined by grouping variable(s)

Correlations
  Computes means, standard deviations, and correlation coefficients for selected variables by subgroups defined by grouping variable(s)

Percentiles
  Computes the score points that separate a given proportion of the distribution of scores, by subgroups defined by the grouping variable(s)

All statistical procedures offered within the analysis module of the IEA IDB Analyzer make use of appropriate sampling weights. Standard errors are computed using the jackknife repeated replicate (JRR) method (see Schulz, Ainley, & Fraillon, forthcoming). Percentages and means, regressions, and correlations may be specified with or without achievement scores. To conduct analyses using achievement scores, select the **With Achievement Scores** option from the **Select Analysis Type** panel. When achievement scores are used, the analyses are performed using all five plausible values and the calculated standard errors include both sampling and imputation error.

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 IDCNTRY as a grouping variable. Additional variables may be selected from the available list. If the **Exclude Missing from Analysis** option is checked, only cases that have non-missing values in the grouping variables will be used in the analysis.

Analysis Variables

This is a list of variables for which means or percentages are to be computed, or the 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.

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.

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, analysts must check the **With Achievement Scores** option in the **Select Analysis Type** panel and select the achievement scores of interest in the Achievement Scores section.

Benchmarks

These are the values that will be used as cut points of the achievement distribution for computing the percentages of students meeting the specified proficiency levels. Although it is best to specify a single proficiency level at a time as a cut point, more can be specified with a space between them.

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 TOTWGTS. When analyzing teacher data TOTWGTT must be used.

Jackknifing Variables

These are the variables that capture the assignment of cases to sampling zones (JKZONES for student and JKZONET for teacher file) and determine whether the case is to be dropped or have its weight doubled (JKREPS for student and JKREPT for teacher files) when computing the sets of replicate weights. The IEA IDB Analyzer automatically uses these variables to compute the 75 sets of replicate weights that are used in all analysis types. This setting cannot be changed.

4.5 Performing Analyses with Student-Level Variables

Many analyses of the ICCS 2009 data may be undertaken using student-level data only. This section presents examples of actual analyses used to produce tables for the *ICCS 2009 International Report* (Schulz et al., 2010b), including examples of percentages only, percentages and means, regression analyses, computing percentages of students reaching proficiency levels, and conducting correlation analysis.

4.5.1 Student-Level Analysis without Achievement Scores

The first example replicates an analysis of students' reported age at the time of testing. The results, presented in Table 3.10 of the *ICCS 2009 International Report* (Schulz et al., 2010b), are reproduced here in Figure 4.5. The example will focus on the results presented in the third data column—the average age at the time of testing. The example reports average ages (with their appropriate standard errors), and therefore computes means without achievement scores.

Figure 4.5 Table of Example Student-Level Analysis without Achievement Scores Taken from the ICCS *2009 International Report (Table 3.10)*

Table 3.10: Country averages for civic knowledge, years of schooling, average age, Human Development Index, and percentile graph

Country	Years of schooling	Average age	Civic Knowledge		HDI
			Average scale score		
Finland	8	14.7	576 (2.4) ▲		0.96
Denmark †	8	14.9	576 (3.6) ▲		0.96
Korea, Republic of[1]	8	14.7	565 (1.9) ▲		0.94
Chinese Taipei	8	14.2	559 (2.4) ▲		0.94
Sweden	8	14.8	537 (3.1) ▲		0.96
Poland	8	14.9	536 (4.7) ▲		0.88
Ireland	8	14.3	534 (4.6) ▲		0.97
Switzerland †	8	14.7	531 (3.8) ▲		0.96
Liechtenstein	8	14.8	531 (3.3) ▲		0.95
Italy	8	13.8	531 (3.3) ▲		0.95
Slovak Republic[2]	8	14.4	529 (4.5) ▲		0.88
Estonia	8	15.0	525 (4.5) ▲		0.88
England ‡	9	14.0	519 (4.4) ▲		0.95
New Zealand †	9	14.0	517 (5.0) ▲		0.95
Slovenia	8	13.7	516 (2.7) ▲		0.93
Norway †	8	13.7	515 (3.4) ▲		0.97
Belgium (Flemish) †	8	13.9	514 (4.7) ▲		0.95
Czech Republic †	8	14.4	510 (2.4) ▲		0.90
Russian Federation	8	14.7	506 (3.8)		0.82
Lithuania	8	14.7	505 (2.8)		0.87
Spain	8	14.1	505 (4.1)		0.96
Austria	8	14.4	503 (4.0)		0.96
Malta	9	13.9	490 (4.5) ▼		0.90
Chile	8	14.2	483 (3.5) ▼		0.88
Latvia	8	14.8	482 (4.0) ▼		0.87
Greece	8	13.7	476 (4.4) ▼		0.94
Luxembourg	8	14.6	473 (2.2) ▼		0.96
Bulgaria	8	14.7	466 (5.0) ▼		0.84
Colombia	8	14.4	462 (2.9) ▼		0.81
Cyprus	8	13.9	453 (2.4) ▼		0.91
Mexico	8	14.1	452 (2.8) ▼		0.85
Thailand †	8	14.4	452 (3.7) ▼		0.78
Guatemala[1]	8	15.5	435 (3.8) ▼		0.70
Indonesia	8	14.3	433 (3.4) ▼		0.73
Paraguay[1]	9	14.9	424 (3.4) ▼		0.76
Dominican Republic	8	14.8	380 (2.4) ▼		0.78

Countries not meeting sampling requirements

Hong Kong SAR	8	14.3	554 (5.7)		0.94
Netherlands	8	14.3	494 (7.6)		0.96

Percentiles of performance: 5th 25th 75th 95th
Mean and confidence interval (±2SE)

▲ Achievement significantly higher than the ICCS average
▼ Achievement significantly lower than the ICCS average

Notes:
( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.
† Met guidelines for sampling participation rates only after replacement schools were included.
‡ Nearly satisfied guidelines for sample participation only after replacement schools were included.
[1] Country surveyed the same cohort of students but at the beginning of the next school year.
[2] National Desired Population does not cover all of International Desired Population.

To replicate the results in this table, analysts must review the student background data codebook and identify the student background variable SAGE as the numeric variable reporting the age of students at the time of testing.

After creating the merged data file for the analysis, the analysis module of the IEA IDB Analyzer will perform the analysis in the following steps:

1) Open the analysis module of the IEA IDB Analyzer.

2) Select the merged data file ISGALLC2.SAV as the **Analysis File**.

3) Select **Percentages and Means** as the **Analysis Type**. Note that there are two options available to check: **With Achievement Scores** and **Exclude Missing from Analysis**. Since no achievement scores are used in this analysis, only Exclude Missing from Analysis should be checked. This option is checked 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, SAGE 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 software automatically defines the **Weight Variable**. As this example analysis uses student background data, TOTWGTS is selected by default. The **Jackknifing Variables** JKZONES and JKREPS also are selected by default.

7) Specify the name and folder of the output files in the **Output Files** field. The IEA IDB Analyzer will use this name and folder to create three an SPSS syntax file that contains the code for performing the analysis. After running the syntax file it will create an SPSS data file and an Excel file with the results.

8) Press the **Start SPSS** button to create the SPSS syntax file. The file will open in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** option. If necessary, the IEA IDB Analyzer will prompt you to confirm overwriting already existing files.

Figure 4.6 shows the how the IEA IDB Analyzer analysis module window looks when all information is completed. The results are displayed in Figure 4.7, although only the first eight countries are displayed to save space (this will be done for all analysis examples). Note that IEA IDB Analyzer also displays the international average statistics for all countries included in the analysis.

In this example, each country's average for the SAGE variable is reported for all sampled students. The countries are identified in the first column. The second column reports the number of valid cases. The third column reports the sum of weights of the sampled students, followed by the percent, mean, and standard deviation, each accompanied by its jackknife standard error. The last column reports the percent of missing values. The first line in Figure 4.7 shows that in Austria valid data were available for 3,135 students and these sampled students represent a population of 81,859 students. Austrian students were, on average, 14.36 years old at the time they took the ICCS 2009 test, with a standard error of 0.02. In total 7.53% of Austrian students did not report their age at the time of testing.

Figure 4.6 IEA IDB Analyzer Setup for Example Student-Level Analysis without Plausible Values
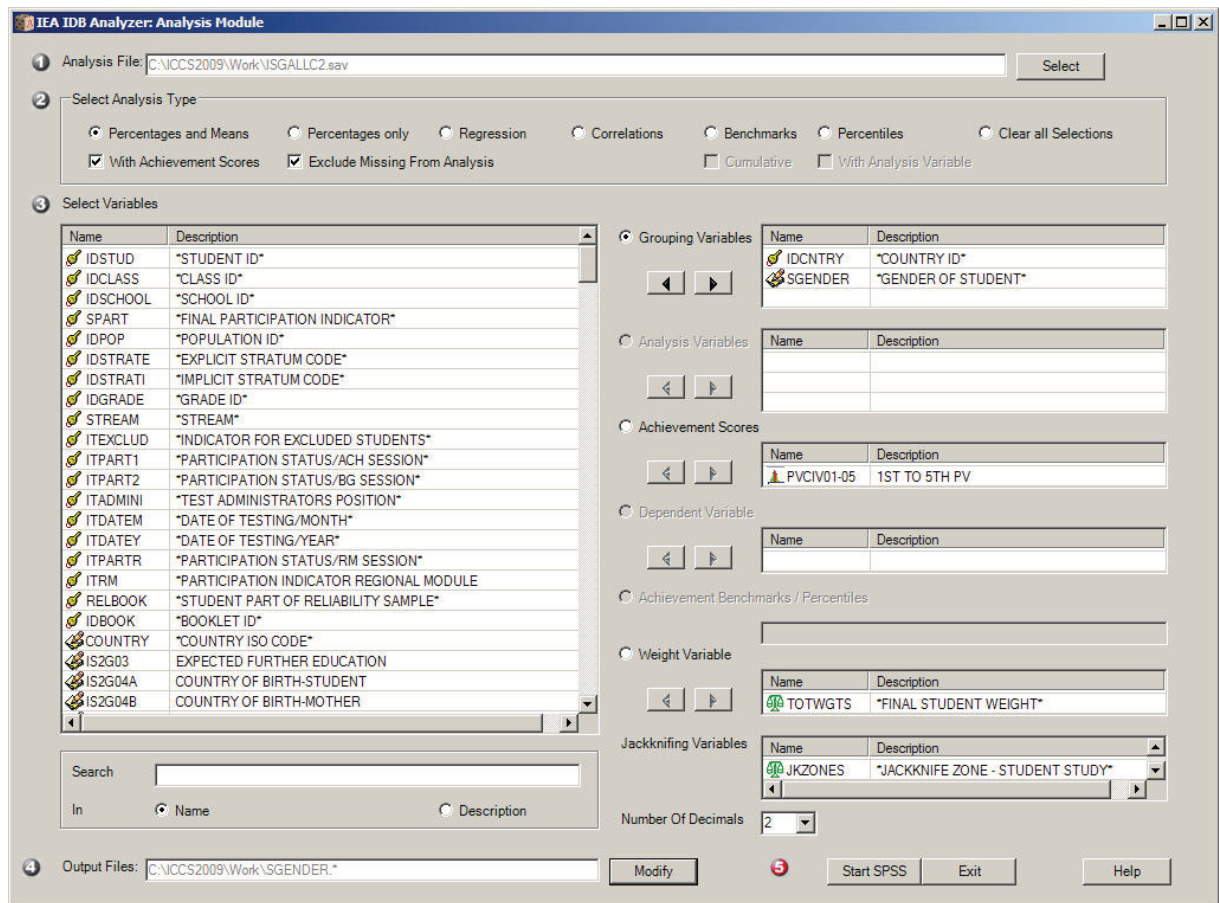
Figure 4.7 Output for Example Student-Level Analysis without Achievement Scores

Average for SAGE by (IDCNTRY)									PAGE 1
COUNTRY ID	N of Cases	Sum of TOTWGTS	Percent	Percent (s.e.)	SAGE (Mean)	SAGE (s.e.)	Std.Dev	Std.Dev. (s.e.)	Percent Missing
Austria	3135	81859	.64	.02	14.36	.02	.54	.02	7.53
Bulgaria	3197	62405	.49	.02	14.69	.01	.49	.02	1.81
Chile	5131	255497	2.00	.06	14.19	.02	.64	.02	1.13
Chinese Taipei	5155	302974	2.37	.04	14.20	.00	.31	.00	.22
Colombia	6064	644327	5.03	.15	14.40	.03	1.02	.02	2.64
Cyprus	3025	8400	.07	.00	13.86	.01	.42	.01	5.32
Czech Republic	4590	94960	.74	.02	14.40	.01	.48	.01	.86
Denmark	4326	59830	.47	.01	14.90	.01	.39	.01	3.86
. . .									
x.International Average	.	.	2.63	.02	14.41	.00	.54	.00	.

4.5.2 Student-Level Analysis with Achievement Scores

The second example replicates another set of results presented in the *ICCS 2009 International Report* (Schulz et al., 2010b), the relationship between students' gender and civic knowledge. The latter is represented by a set of five plausible values. These results, presented in Table 3.13 of the *ICCS 2009 International Report* (Schulz et al., 2010b), are repeated here in Figure 4.8. Since the results in this table are based on plausible values, analysts must include the values when creating the file using the merge module and indicate that the analysis will make use of achievement scores.

Figure 4.8 Table of Example Student-Level Analysis with Achievement Scores Taken from the ICCS 2009 International Report *(Table 3.13)*

Table 3.13: Gender differences in civic knowledge

Country	Mean Scale Score Females	Mean Scale Score Males	Difference (males–females)	Gender Difference
Guatemala[1]	435 (4.2)	434 (4.3)	-2 (3.7)	
Colombia	463 (3.1)	461 (4.0)	-3 (4.1)	
Belgium (Flemish) †	517 (5.3)	511 (5.6)	-6 (5.8)	
Switzerland †	535 (3.0)	528 (5.5)	-7 (4.6)	
Denmark †	581 (3.4)	573 (4.5)	-8 (3.5)	
Luxembourg	479 (2.8)	469 (3.4)	-10 (4.5)	
Liechtenstein	539 (6.4)	526 (6.2)	-12 (10.4)	
Chile	490 (4.3)	476 (4.2)	-14 (4.8)	
Austria	513 (4.6)	496 (4.5)	-16 (4.7)	
Slovak Republic[2]	537 (5.4)	520 (4.4)	-18 (4.2)	
Czech Republic †	520 (3.0)	502 (2.4)	-18 (2.8)	
Italy	540 (3.4)	522 (3.9)	-18 (3.3)	
Indonesia	442 (3.9)	423 (3.5)	-19 (3.0)	
Spain	514 (4.2)	496 (4.8)	-19 (3.6)	
England ‡	529 (6.1)	509 (6.1)	-20 (8.5)	Females Score Higher / Males Score Higher
Russian Federation	517 (4.3)	496 (3.8)	-21 (3.4)	
Sweden	549 (3.4)	527 (4.2)	-21 (4.5)	
Ireland	545 (4.8)	523 (6.0)	-22 (6.2)	
Korea, Republic of[1]	577 (2.4)	555 (2.3)	-22 (3.0)	
Norway †	527 (3.7)	504 (4.5)	-23 (4.4)	
Mexico	463 (3.2)	439 (3.1)	-24 (2.9)	
Dominican Republic	392 (2.8)	367 (2.7)	-25 (2.7)	
Bulgaria	479 (5.2)	454 (6.1)	-26 (5.3)	
Chinese Taipei	573 (2.7)	546 (2.7)	-26 (2.5)	
Finland	590 (2.9)	562 (3.5)	-28 (4.3)	
Paraguay[1]	438 (4.1)	408 (3.9)	-29 (4.6)	
Slovenia	531 (2.6)	501 (3.9)	-30 (4.0)	
Latvia	497 (3.7)	466 (5.0)	-30 (3.7)	
New Zealand †	532 (5.9)	501 (6.4)	-31 (7.5)	
Greece	492 (4.8)	460 (5.1)	-32 (4.5)	
Poland	553 (4.5)	520 (5.5)	-33 (4.3)	
Estonia	542 (4.8)	509 (4.9)	-33 (3.9)	
Malta	507 (7.7)	473 (3.6)	-34 (8.2)	
Lithuania	523 (2.9)	488 (3.4)	-35 (3.0)	
Cyprus	475 (2.7)	435 (3.2)	-40 (3.7)	
Thailand †	474 (3.9)	426 (4.5)	-48 (4.5)	
ICCS average	**511 (0.7)**	**489 (0.7)**	**-22 (0.8)**	

Countries not meeting sample requirements

Country	Mean Scale Score Females	Mean Scale Score Males	Difference (males–females)	Gender Difference
Hong Kong SAR	564 (6.5)	543 (8.3)	-21 (9.8)	
Netherlands	497 (6.6)	490 (10.4)	-7 (7.9)	

Gender Difference scale: (-100) (-50) 0 50 100

■ Gender difference statistically significant at 0.05 level
□ Gender difference not statistically significant

Notes:
( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.
†   Met guidelines for sampling participation rates only after replacement schools were included.
‡   Nearly satisfied guidelines for sample participation only after replacement schools were included.
[1] Country surveyed the same cohort of students but at the beginning of the next school year.
[2] National Desired Population does not cover all of International Desired Population.

The codebooks show that the variable SGENDER contains categorical information on the gender of the student, and that this variable is found in the student background data files.

The **Percentages and Means** analysis type with the **With Achievement Scores** option activated computes percentages and mean achievement scores based on plausible values and their respective standard errors. Analysts must select the BSGALLM4.SAV data file using these steps:

1) Open the analysis module of the IEA IDB Analyzer.

2) Select the merged data file ISGALLC2.SAV as the **Analysis File**.

3) Select **Percentages and Means** as the **Analysis Type**. By default, the program will exclude records with missing grouping variables from the analysis.

4) Check the **With Achievement Scores** box.

5) Add the variable SGENDER 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 variable PVCIV01-05 from the list of available variables (this set of plausible values should be the only set available) and move it to the analysis variables field by clicking the right arrow button ▶ in this section.

7) The software automatically defines the **Weight Variable**. This sample analysis uses student background data, so 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.

9) Click the **Start SPSS** button to create the SPSS syntax file. The file will open 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 will prompt to confirm overwriting already existing files.

Figure 4.9 displays the analysis module with the proper settings for this sample analysis. The output for the set-up is shown in Figure 4.10.

Figure 4.9 IEA IDB Analyzer Setup for Example Student-Level Analysis with Achievement Scores
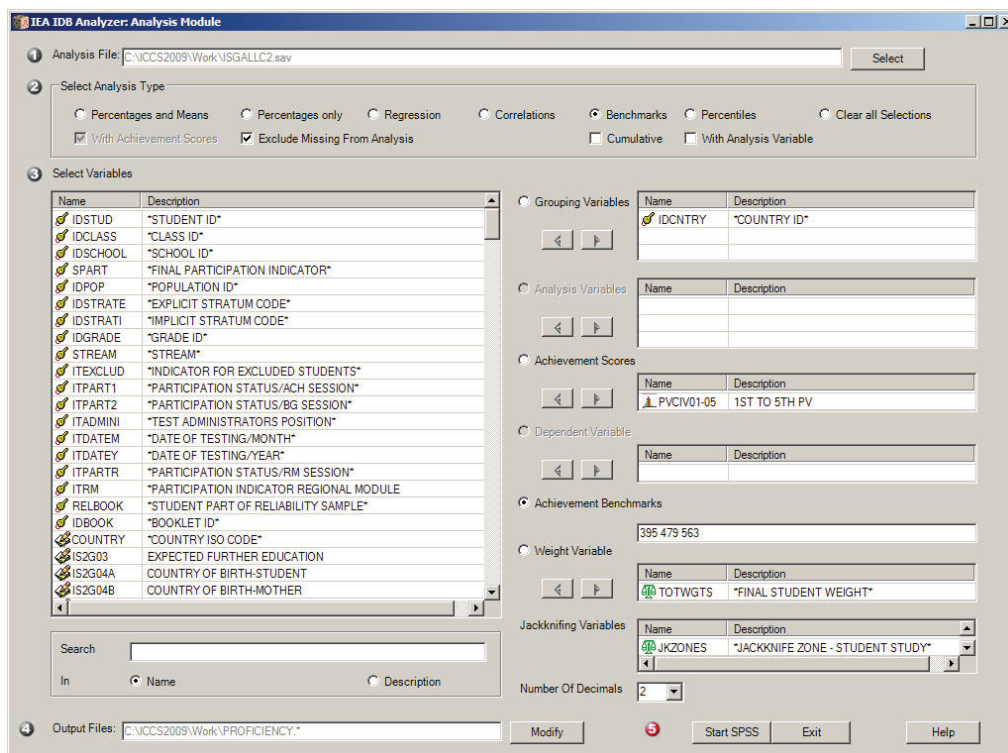
In this example, each country's results are presented on two lines, one for each gender, that is, the values of the SGENDER variable. The countries are identified in the first column and the second column describes the category of SGENDER being reported. The third column reports the number of valid cases and the fourth the sum of weights of the sampled students. The next two columns report the percentage of students in each category and the standard error, followed by the estimated mean civic knowledge achievement and the standard error. The standard deviation of the achievement scores and the standard error are reported in the last two columns.

The first two lines of Figure 4.10 show that in Austria 49.93% of the target population students are girls and 50.07% are boys. The mean civic knowledge is 512.60 (standard error of 4.59) for girls and 496.47 (standard error of 4.51) for boys.

Figure 4.10 Output for Example Student-Level Analysis with Achievement Scores

Std.Dev. *COUNTRY ID*	*GENDER OF STUDENT*	N of Cases	Sum of TOTWGTS	Percent	Percent (s.e.)	PVCIV (Mean)	PVCIV (s.e.)	Std.Dev	(s.e.)
Austria	BOY	1553	41734	50.07	1.41	496.47	4.51	100.05	2.43
	GIRL	1637	41624	49.93	1.41	512.60	4.59	92.07	2.31
Bulgaria	BOY	1590	30431	48.21	1.65	453.51	6.13	105.75	3.13
	GIRL	1642	32687	51.79	1.65	479.30	5.21	103.02	3.87
Chile	BOY	2510	126397	49.17	1.45	476.23	4.20	88.55	1.92
	GIRL	2651	130659	50.83	1.45	489.83	4.26	85.91	2.39
Chinese Taipei	BOY	2670	155929	51.58	.57	546.12	2.75	96.32	1.49
	GIRL	2474	146348	48.42	.57	572.55	2.73	89.21	1.81
Colombia	BOY	2877	315242	47.77	1.37	460.63	4.05	81.87	2.25
	GIRL	3315	344694	52.23	1.37	463.41	3.06	79.85	1.89
Cyprus	BOY	1548	4322	50.35	.62	434.81	3.20	93.28	1.89
	GIRL	1540	4261	49.65	.62	475.08	2.74	87.59	1.86
Czech Republic	BOY	2492	51610	54.00	.97	502.14	2.45	86.56	1.58
	GIRL	2128	43966	46.00	.97	519.96	3.03	87.30	1.57
Denmark	BOY	2092	28592	47.37	.82	573.35	4.45	101.89	2.38
	GIRL	2271	31771	52.63	.82	581.44	3.44	96.47	1.63
. . .									
x.International Average	BOY	.	.	49.75	.24	490.59	.79	91.21	.43
	GIRL	.	.	50.25	.24	512.45	.71	85.88	.43

Average for PVCIV by IDCNTRY SGENDER — PAGE 1

4.5.3 Student-Level Regression Analysis

The IEA IDB Analyzer is able to calculate multiple linear regressions between dependent variables and a set of independent variables. This section demonstrates an example for a regression analysis with achievement scores using student-level variables selected in the merged data file ISGALLC2.SAV.

The IEA IDB Analyzer can also be used to compute regression analyses without achievement scores, but no example is given here, as the steps are similar to those described for a regression analysis with achievement scores. The difference is that instead of selecting **Achievement Scores**, a **Dependent Variable** should be selected as an outcome.

4.5.4 Student-Level Regression Analysis with Achievement Scores

This example will look at gender as a predictor of civic knowledge achievement. The linear regression analysis will use the variable SGENDER as the predictor of the five plausible values for civic knowledge (PVCIV01 through PVCIV05), using the weighting variable TOTWGTS. The data will come from the merged data file ISGALLC2.SAV and the standard errors will be computed based on 75 sets of replicate weights. The previous example computed the mean achievement between girls and boys. This example will test whether the differences between them are statistically significant. The current example replicates Table 3.13 from the *ICCS 2009 International Report* (Schulz et al., 2010b), shown in Figure 4.8 (see chapter 4.5.2). The mean achievement for girls and boys is represented in the second and third column and the mean score differences and the indication of whether these differences are statistically significant is in the last columns of the table, represented by bars.

For this example, the values of the variable SGENDER are recoded into variable REGGENDER. This recoded variable is created by running the special SPSS syntax file Syntax_ ISGALLC2.SPS and is provided in Figure 4.11. By using this recoded variable, the intercept or constant will be the estimated average civic knowledge achievement for girls, whereas the regression coefficient (REGGENDER [estimate]) shows the estimated difference in civic knowledge achievement score points of boys compared to girls. A t-test will determine if the average civic knowledge achievement is significantly different between girls and boys.

Figure 4.11 Example SPSS Program to Recode Variable SGENDER for Student-Level Regression Analysis

```
GET FILE = "<datapath> ISGALLC2.SAV".

* Create new variable REGGENDER from SGENDER.
RECODE SGENDER (MISSING=SYSMISS) (0=1) (1=0) INTO REGGENDER.
VALUE LABELS REGGENDER
        '0' 'Girl'
        '1' 'Boy' .

VARIABLE LABELS
    REGGENDER "Recoded SGENDER (Girls = 0; Boys = 1)".

EXECUTE.

SAVE OUTFILE = "<datapath> ISGALLC2.sav".
```

The analysis module of the IEA IDB Analyzer will perform the sample regression analysis using the following steps (the completed analysis window shown in Figure 4.12):

1)   Open the analysis module of the IEA IDB Analyzer.

2)   Specify the data file ISGALLC2.SAV as the **Analysis File,** after having run the SPSS syntax file Syntax_ ISGALLC2.SPS to create the variable REGGENDER.

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 and select REGGENDER as the analysis variable. To do this, select REGGENDER from the list of available variables and moving it into the **Analysis Variables** field by clicking the right arrow button in this section.

7) Click the **Achievement Scores** radio button. Select the variable PVCIV01-05 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 software automatically defines the **Weight Variable**. As this sample analysis uses student background data, TOTWGTS is selected by default. The **Jackknifing Variables** JKZONES and JKREPS also are selected by default.

9) Specify the name and folder of the output files in the **Output Files** field.

10) Click the **Start SPSS** button to create the SPSS syntax file. The file will open in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** option. If necessary, the IEA IDB Analyzer will prompt for confirmation before overwriting already existing files.

Figure 4.12 IDB Analyzer Setup for Example Student-Level Regression Analysis with Achievement Scores

The results of this analysis are presented in Figure 4.13. The first line of results shows that in Austria the estimated mean civic knowledge achievement of target grade girls, labeled "Constant (estimate)", is 512.60, with a standard error of 4.59. Austrian target grade boys have an estimated mean civic knowledge achievement 16.14 points (REGGENDER (estimate)) lower than Austrian girls with standard error of 4.66. The estimated t-test value is -3.47 (REGGENDER (t-test)) which in absolute value is greater than 1.96, indicating that this difference is statistically significant at a 95% confidence level. The statistical significance and insignificance in Table 3.13 in the *ICCS 2009 International Report* (Schulz et al., 2010b) are marked a different color.

Figure 4.13 Output for Example Student-Level Regression Analysis with Achievement Scores

| Predictors: REGGENDER / Predicted: PVCIV | | | | | | PAGE   1 | |
COUNTRY ID	N of Cases	Mult_RSQ	Constant (estimate)	Constant (s.e.)	REGGENDER (estimate)	REGGENDER (s.e.)	REGGENDER (t-test)
Austria	3190	.01	512.60	4.59	-16.14	4.66	-3.4
Bulgaria	3232	.02	479.30	5.21	-25.79	5.32	-4.8
Chile	5161	.01	489.83	4.26	-13.61	4.78	-2.8
Chinese Taipei	5144	.02	572.55	2.73	-26.42	2.53	-10.43
Colombia	6192	.00	463.41	3.06	-2.78	4.06	-.6
Cyprus	3088	.05	475.08	2.74	-40.27	3.67	-10.9
Czech Republic	4620	.01	519.96	3.03	-17.82	2.79	-6.3
Denmark	4363	.00	581.44	3.44	-8.10	3.51	-2.30
. . .							
x.International Average	.	.	512.45	.71	-21.86	.83	-26.31

4.5.5 Calculating Percentages of Students Reaching Proficiency Levels

This section describes the IEA IDB Analyzer's ability to perform benchmark analyses, which will compute the percentages of students reaching specified proficiency levels on an achievement scale and within specified subgroups, along with appropriate standard errors.

As an example, we will compute the percentages of students who did not reach the three ICCS 2009 international proficiency levels of civic knowledge achievement (Level 1 is 395 to 478 score points; Level 2 is 479 to 562 score points; Level 3 is 563 score points and above) using the merged ISGALLC2.SAV data file. These results, presented in Table 3.12 of the ICCS 2009 International Report (Schulz et al., 2010b), are repeated in Figure 4.14.

Figure 4.14 Example Table of Proficiency Levels Analysis Taken from the ICCS 2009 International
Report *(Table 3.12)*

Table 3.12: Percentages of students at each proficiency level across countries

Country	Below Level 1 (less then 395 score points)		Level 1 (from 395 to 479 score points)		Level 2 (from 479 to 563 score points)		Level 3 (563 score points and more)		
Finland	2	(0.3)	10	(0.7)	30	(1.2)	58	(1.3)	
Denmark †	4	(0.5)	13	(0.8)	27	(1.1)	56	(1.6)	
Korea, Republic of[1]	3	(0.3)	12	(0.6)	32	(0.9)	54	(1.1)	
Chinese Taipei	5	(0.4)	15	(0.8)	29	(1.0)	50	(1.3)	
Liechtenstein	8	(1.4)	18	(1.9)	30	(2.4)	45	(2.0)	
Ireland	10	(1.1)	20	(1.4)	29	(1.2)	41	(1.8)	
Poland	9	(1.0)	19	(1.1)	31	(1.0)	41	(2.0)	
Sweden	8	(0.8)	21	(0.9)	32	(1.1)	40	(1.4)	
Italy	7	(0.7)	20	(1.0)	35	(1.0)	38	(1.5)	
Slovak Republic[2]	7	(0.9)	22	(1.4)	34	(1.4)	37	(2.2)	
Switzerland †	6	(0.8)	21	(1.5)	37	(1.3)	37	(1.8)	
Estonia	8	(1.1)	22	(1.3)	34	(1.4)	36	(2.1)	
New Zealand †	14	(1.2)	22	(1.5)	28	(1.4)	35	(2.1)	
England ‡	13	(1.2)	22	(0.9)	31	(1.2)	34	(1.6)	
Norway †	11	(0.9)	24	(1.1)	33	(1.1)	32	(1.3)	
Slovenia	9	(0.9)	25	(1.1)	36	(1.2)	30	(1.2)	
Belgium (Flemish) †	8	(1.2)	24	(1.7)	39	(1.6)	29	(2.1)	
Austria	15	(1.4)	25	(1.2)	32	(1.2)	29	(1.4)	
Czech Republic †	10	(0.7)	27	(1.0)	36	(1.1)	28	(1.1)	
Spain	11	(1.3)	26	(1.3)	37	(1.5)	26	(1.8)	
Russian Federation	10	(0.9)	29	(1.5)	36	(1.2)	26	(1.8)	
Lithuania	9	(0.8)	28	(1.2)	39	(1.2)	24	(1.3)	
Malta	17	(1.6)	26	(1.8)	33	(1.9)	24	(2.3)	
Greece	22	(1.7)	28	(1.3)	29	(1.1)	21	(1.4)	
Bulgaria	27	(1.8)	26	(1.5)	27	(1.6)	20	(1.9)	
Chile	16	(1.3)	33	(1.2)	32	(1.3)	19	(1.1)	
Luxembourg	22	(1.2)	30	(1.0)	29	(0.8)	19	(0.6)	
Latvia	15	(1.6)	33	(1.3)	35	(1.7)	16	(1.4)	
Cyprus	28	(1.0)	32	(1.0)	27	(1.0)	13	(0.9)	
Colombia	21	(1.3)	36	(1.0)	32	(1.1)	11	(0.8)	
Mexico	26	(1.3)	36	(1.1)	27	(1.0)	10	(0.8)	
Thailand †	25	(1.6)	38	(1.4)	29	(1.6)	8	(1.1)	
Paraguay[1]	38	(1.9)	35	(1.6)	20	(1.2)	7	(0.7)	
Guatemala[1]	30	(1.7)	42	(1.6)	22	(1.4)	5	(1.2)	
Indonesia	30	(1.9)	44	(1.5)	22	(1.3)	3	(0.7)	
Dominican Republic	61	(1.6)	31	(1.3)	7	(0.6)	1	(0.2)	
ICCS average	**16**	**(0.2)**	**26**	**(0.2)**	**31**	**(0.2)**	**28**	**(0.2)**	

■ Below Level 1   ■ Level 1
■ Level 2   ☐ Level 3

Countries not meeting sampling requirements

Hong Kong SAR	7	(1.2)	14	(1.4)	30	(1.5)	50	(2.6)	
Netherlands	15	(2.7)	28	(2.4)	33	(2.3)	24	(3.0)	

Notes:

Countries ranked in descending order by percentages in Level 3.

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

†   Met guidelines for sampling participation rates only after replacement schools were included.

‡   Nearly satisfied guidelines for sample participation only after replacement schools were included.

[1]   Country surveyed the same cohort of students but at the beginning of the next school year.

[2]   National Desired Population does not cover all of International Desired Population.

Researchers may use the analysis module of IEA IDB Analyzer to replicate this example using the steps described below. Figure 4.15 shows the completed analysis window.

1) Open the analysis module of the IEA IDB Analyzer.

2) Specify the data file ISGALLC2.SAV as the **Analysis File**.

3) Select Benchmarks as the **Analysis Type**.

4) The variable IDCNTRY is selected automatically as **Grouping Variables**. No additional grouping variables are needed for this analysis.

5) Click the **Achievement Scores** radio button. Select the variable PVCIV01-05 from the list of available variables and move it to the **Achievement Scores** field by clicking the right arrow button ▶ in this section.

6) Click the **Achievement Benchmarks** radio button to activate this section and specify the ICCS 2009 international benchmarks, which are 395, 479, and 563 as Level 1, Level 2, and Level 3, respectively. Enter these four values in the input field, each separated by a blank space.

7) The software automatically defines the **Weight Variable**. As this example analysis uses student background data, TOTWGTS is selected by default. The **Jackknifing Variables** JKZONES and JKREPS also are selected by default.

8) Specify the name and folder of the output files in the **Output Files** field.

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

Figure 4.15 IDB Analyzer Set-Up for Example Benchmark Analysis
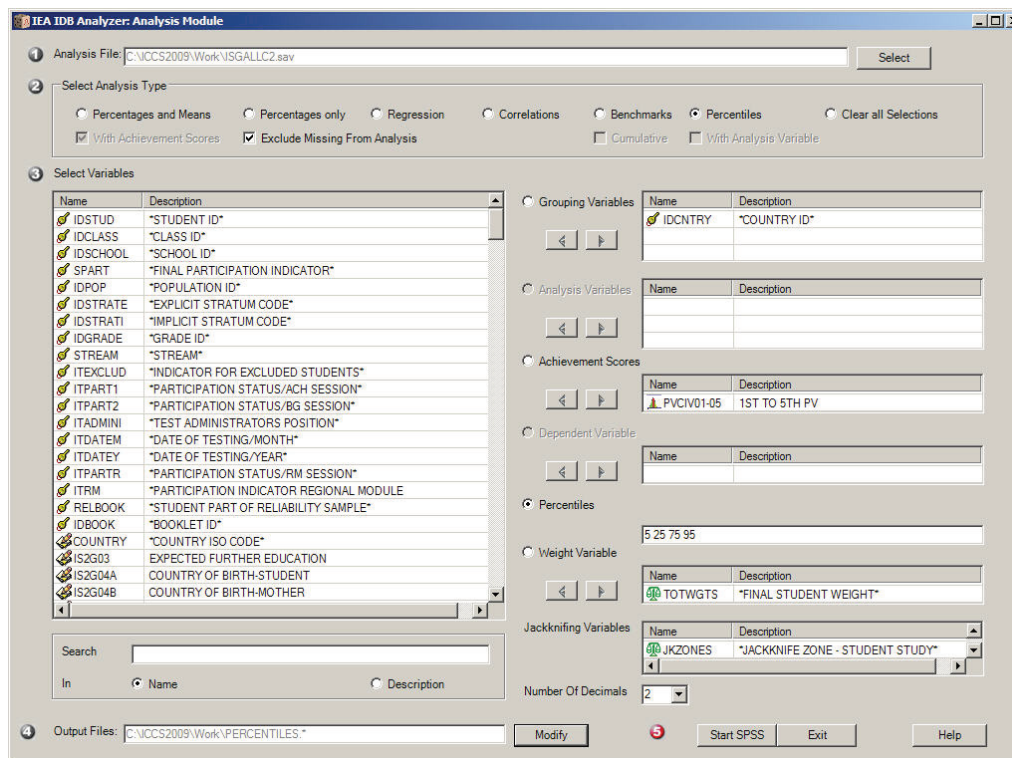
The results of this analysis are presented in Figure 4.16. In Austria 14.56 percent of target grade students are below the Proficiency Level 1 of 395 score points, with a standard error of 1.41 percent. In the next group, Proficiency Level 2, 25.00 percent of students reached scored between 395 and 479 score points, with a standard error of 1.22 percent. In the final group, Proficiency Level 3, 31.52 percent of students scored between 479 and 563 points, with a standard error of 1.25 percent. Finally, 28.92 percent of students scored higher than Proficiency Level 3 (above 563 score points) with standard error of 1.44 percent.

Figure 4.16 Output for Example Benchmark Analysis

Percent within benchmarks (395 479 563) of PV				PAGE	1
COUNTRY ID	Performance Group	N of Cases	Sum of TOTWGTS	Percent	Percent (s.e.)
Austria	1.Below 395	487	12887	14.56	1.41
	2.From 395 to 479	812	22135	25.00	1.22
	3.From 479 to 563	1067	27905	31.52	1.25
	4.Above 563	1018	25600	28.92	1.44
Bulgaria	1.Below 395	820	17026	26.79	1.84
	2.From 395 to 479	875	16708	26.29	1.50
	3.From 479 to 563	880	17106	26.91	1.61
	4.Above 563	681	12716	20.01	1.86
Chile	1.Below 395	737	42230	16.34	1.26
	2.From 395 to 479	1548	84534	32.71	1.19
	3.From 479 to 563	1695	81739	31.63	1.29
	4.Above 563	1213	49918	19.32	1.07
Chinese Taipei	1.Below 395	234	14743	4.86	.44
	2.From 395 to 479	757	46476	15.31	.81
	3.From 479 to 563	1507	89463	29.46	1.01
	4.Above 563	2670	152949	50.37	1.26
Colombia	1.Below 395	1234	140040	21.16	1.33
	2.From 395 to 479	2204	238842	36.09	1.03
	3.From 479 to 563	2020	210983	31.88	1.06
	4.Above 563	746	71922	10.87	.83
Cyprus	1.Below 395	908	2493	28.11	1.00
	2.From 395 to 479	1020	2845	32.07	.96
	3.From 479 to 563	870	2420	27.28	.98
	4.Above 563	396	1113	12.54	.90
Czech Republic	1.Below 395	442	9190	9.59	.69
	2.From 395 to 479	1251	25803	26.94	.97
	3.From 479 to 563	1651	34075	35.58	1.13
	4.Above 563	1287	26712	27.89	1.12
Denmark	1.Below 395	167	2293	3.68	.47
	2.From 395 to 479	579	7920	12.73	.78
	3.From 479 to 563	1246	17108	27.49	1.11
	4.Above 563	2516	34913	56.10	1.60
. . .					
x.International Average	1.Below 395	.	.	15.42	.21
	2.From 395 to 479	.	.	25.71	.21
	3.From 479 to 563	.	.	30.59	.22
	4.Above 563	.	.	28.28	.26

4.5.6 Computing Correlations with Background Variables and Achievement Scores

In addition to the analyses described above, the IEA IDB Analyzer also is able to compute correlations between background variables, and between background variables and achievement scores. The example shown here is a correlation analysis with achievement scores. A correlation analysis between two background variables would follow the same steps. The only difference is that the correlation between two background variables requires adding two variables in the **Analysis Variables** field instead of one.

The *ICCS 2009 International Report* (Schulz et al., 2010b) does not contain the examples of correlation between background variable and achievement scores presented here. The steps of conducting correlation analysis with IEA IDB Analyzer are described below. Figure 4.17 shows the completed analysis window. This example calculates the correlation between students' discussion of political and social issues outside of school (POLDISC) and the civic knowledge achievement score (represented by the five plausible values PVCIV01-05).

1) Open the analysis module of the IEA IDB Analyzer.

2) Specify the data file ISGALLC2.SAV as the **Analysis File**.

3) Select Correlations as the **Analysis Type**. The IDCNTRY (country ID) is selected by default. No other variable needs to be selected for this analysis.

Figure 4.17 IDB Analyzer Setup for Example Correlation Analysis
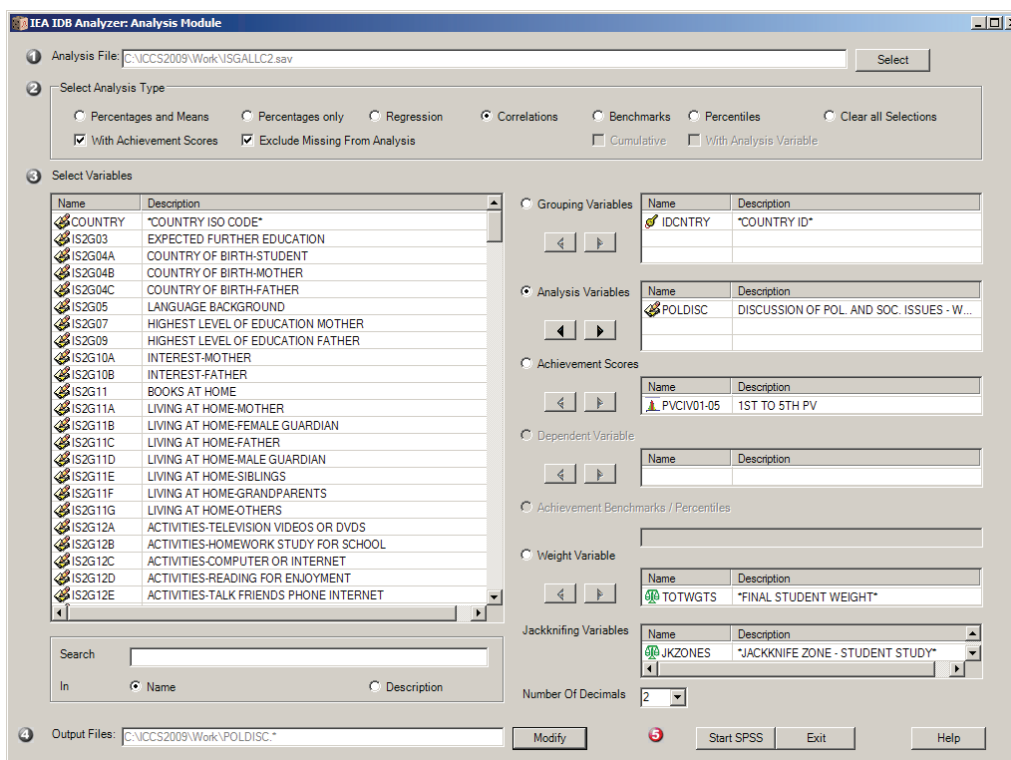
4) Click the **With Achievement Scores** check-box to activate this option.

5) Click on **Analysis Variables** radio button and move the variable POLDISC into this field using the right arrow button ▶.

6) Click on **Achievement Scores** radio button and select PVCIV01-05 as achievement scores. Use the right arrow button ▶ to move it to the corresponding field.

7) The software automatically defines the **Weight Variable**. As this sample analysis uses student background data, TOTWGTS is selected by default. The **Jackknifing Variables** JKZONES and JKREPS also are selected by default.

8) Specify the name and folder of the output files in the **Output Files** field.

9) Click the **Start SPSS** button to create the SPSS syntax file. The file will open in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** option. If necessary, the IEA IDB Analyzer will prompt for confirmation to overwrite existing files.

Figure 4.18 shows the output from the correlation analysis conducted by IEA IDB Analyzer. The output contains a separate matrix for each country. The output shows numeric country codes instead of the actual country names. These numeric codes can be matched against the country names by opening the produced SPSS file, switching to **Variable View**, and clicking on the **Values** column of the first variable (IDCNTRY). The country names are also displayed in the Excel output file. As the SPSS output shows, the correlation between students' discussion of political and social issues outside of school and the civic knowledge achievement score in Austria (IDCNTRY = 40) is 0.1977 with a standard error of 0.0229.

Figure 4.18 Output for Example Correlation Analysis

Correlation matrix for IDCNTRY= 40 *

Variable	Sum of Wgts	Mean	s.e	StdDev	s.e	Correlations and s.e.	
PVCIV	87404.27	504.07	3.930	96.60	1.957	1.0000	.0000
POLDISC	87404.27	50.89	.241	9.85	.175	.1977	.0229

Correlation matrix for IDCNTRY= 100 *

Variable	Sum of Wgts	Mean	s.e	StdDev	s.e	Correlations and s.e.	
PVCIV	61969.88	469.85	4.911	103.76	3.218	1.0000	.0000
POLDISC	61969.88	50.44	.269	10.06	.179	.0561	.0268

Correlation matrix for IDCNTRY= 152 *

Variable	Sum of Wgts	Mean	s.e	StdDev	s.e	Correlations and s.e.	
PVCIV	256373.32	483.73	3.518	87.26	1.538	1.0000	.0000
POLDISC	256373.32	49.33	.227	9.91	.099	.1442	.0191

Correlation matrix for IDCNTRY= 158 *

Variable	Sum of Wgts	Mean	s.e	StdDev	s.e	Correlations and s.e.	
PVCIV	302016.34	559.20	2.448	93.59	1.245	1.0000	.0000
POLDISC	302016.34	49.38	.193	10.39	.091	.1799	.0146

Correlation matrix for IDCNTRY= 170 *

Variable	Sum of Wgts	Mean	s.e	StdDev	s.e	Correlations and s.e.	
PVCIV	638459.08	464.87	2.851	79.87	1.586	1.0000	.0000
POLDISC	638459.08	51.00	.247	10.11	.147	-.0311	.0166

Correlation matrix for IDCNTRY= 196 *

Variable	Sum of Wgts	Mean	s.e	StdDev	s.e	Correlations and s.e.	
PVCIV	8622.53	456.03	2.439	92.08	1.368	1.0000	.0000
POLDISC	8622.53	49.95	.234	10.19	.136	.1389	.0217

Correlation matrix for IDCNTRY= 203 *

Variable	Sum of Wgts	Mean	s.e	StdDev	s.e	Correlations and s.e.	
PVCIV	95112.78	510.95	2.348	86.97	1.323	1.0000	.0000
POLDISC	95112.78	47.64	.163	9.19	.107	.1242	.0190

Correlation matrix for IDCNTRY= 208 *

Variable	Sum of Wgts	Mean	s.e	StdDev	s.e	Correlations and s.e.	
PVCIV	60307.65	577.67	3.530	99.00	1.540	1.0000	.0000
POLDISC	60307.65	50.26	.253	9.93	.123	.3217	.0198

4.5.7 Calculating Percentiles of Student Achievement

To calculate percentiles of achievement scores, select the **Percentiles** analysis type. This computes the percentiles within the distribution of student achievement scores within specified subgroups of students. This analysis type also computes the appropriate standard errors for those percentiles.

This example will compute the percentiles of student achievement scores and their standard errors within each country, using the weighting variable TOTWGTS, as in Table B.1 of Appendix B of the ICCS *2009 International Report* (see Schulz et al., 2010b: 265). The data will be read from the data file ISGALLC2.sav and the standard errors will be computed based on replicate weights.

Figure 4.19 Example Table of Percentiles Analysis Taken from the ICCS 2009 International Report *(Table B.1)*

Table B.1: Percentiles of civic knowledge

Country	5th percentile		25th percentile		75th percentile		95th percentile	
Austria	336	(8.8)	435	(6.9)	574	(4.6)	657	(5.4)
Belgium (Flemish) †	374	(7.0)	459	(8.1)	572	(6.1)	640	(5.5)
Bulgaria	296	(7.5)	389	(8.6)	544	(8.2)	632	(7.4)
Chile	344	(7.2)	420	(5.0)	544	(4.6)	629	(6.3)
Chinese Taipei	397	(5.4)	495	(4.6)	626	(5.3)	705	(5.1)
Colombia	329	(6.1)	405	(4.2)	518	(4.2)	594	(5.0)
Cyprus	304	(5.7)	386	(3.9)	518	(3.8)	607	(6.5)
Czech Republic †	370	(4.9)	447	(3.7)	571	(4.9)	656	(5.2)
Denmark †	410	(7.1)	509	(6.0)	645	(5.6)	736	(5.9)
Dominican Republic	280	(4.0)	333	(5.3)	423	(4.9)	498	(5.0)
England ‡	344	(8.3)	447	(6.6)	592	(6.3)	690	(10.6)
Estonia	371	(9.2)	463	(6.2)	590	(6.4)	671	(8.1)
Finland	433	(7.4)	520	(4.5)	635	(4.7)	710	(4.2)
Greece	317	(6.7)	404	(8.4)	548	(6.5)	635	(7.7)
Guatemala[1]	312	(5.7)	384	(4.8)	485	(6.5)	564	(9.2)
Indonesia	321	(6.4)	385	(4.6)	479	(5.7)	551	(6.0)
Ireland	361	(8.2)	461	(8.4)	607	(6.6)	695	(6.6)
Italy	380	(8.5)	472	(6.0)	593	(4.3)	669	(6.1)
Korea, Republic of[1]	424	(4.3)	512	(4.8)	621	(3.9)	688	(3.9)
Latvia	349	(6.2)	425	(6.3)	538	(5.2)	617	(7.8)
Liechtenstein	380	(20.9)	477	(15.3)	595	(5.6)	682	(9.2)
Lithuania	373	(5.8)	450	(4.8)	561	(4.0)	635	(5.9)
Luxembourg	315	(5.2)	405	(4.2)	542	(3.2)	630	(4.6)
Malta	326	(9.4)	423	(8.5)	560	(5.6)	635	(8.0)
Mexico	321	(5.2)	392	(5.0)	510	(4.8)	591	(5.0)
New Zealand †	333	(8.6)	440	(7.0)	596	(7.3)	693	(7.2)
Norway †	352	(7.0)	450	(6.0)	581	(5.0)	669	(6.7)
Paraguay[1]	280	(6.3)	362	(5.4)	483	(6.1)	575	(4.4)
Poland	371	(6.9)	469	(7.8)	606	(7.1)	695	(6.4)
Russian Federation	370	(4.7)	446	(5.2)	565	(6.2)	647	(8.1)
Slovak Republic[2]	382	(6.4)	466	(5.3)	593	(6.6)	673	(8.0)
Slovenia	372	(5.4)	455	(5.0)	577	(5.0)	660	(6.0)
Spain	358	(8.5)	447	(6.9)	566	(6.4)	639	(5.6)
Sweden	374	(5.5)	468	(4.6)	605	(6.0)	701	(6.5)
Switzerland †	391	(7.5)	476	(5.3)	589	(5.2)	665	(6.4)
Thailand †	327	(6.1)	396	(6.1)	507	(6.5)	579	(7.1)

Countries not meeting sampling requirements

Hong Kong SAR	379	(12.0)	494	(8.4)	621	(5.8)	702	(5.5)
Netherlands	342	(13.8)	431	(10.4)	559	(8.5)	635	(8.7)

Additional grade samples

Greece	351	(8.2)	450	(6.8)	584	(5.7)	666	(4.2)
Norway †	359	(6.9)	469	(6.1)	613	(5.2)	699	(6.7)
Slovenia	390	(4.6)	479	(5.0)	604	(4.6)	686	(5.6)
Sweden	391	(6.2)	502	(5.4)	650	(6.0)	745	(6.5)

Notes:

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

† Met guidelines for sampling participation rates only after replacement schools were included.

‡ Nearly satisfied guidelines for sample participation only after replacement schools were included.

[1] Country surveyed the same cohort of students but at the beginning of the next school year.

[2] National Desired Population does not cover all of International Desired Population.

The steps in the IEA IDB Analyzer required to follow the example are:

1) Open the analysis module of the IEA IDB Analyzer.

2) Specify the data file ISGALLC2.SAV as the **Analysis File**.

3) Select **Percentiles** as the **Analysis Type**. The IDCNTRY (country ID) is selected by default. No other variable needs to be selected for this analysis.

4) Click on **Achievement Scores** radio button and select PVCIV01-05 as achievement scores. Use the right arrow button ▶ to move it to the corresponding field.

5) The software automatically defines the **Weight Variable**. As this sample analysis uses student background data, TOTWGTS is selected by default. The **Jackknifing Variables** JKZONES and JKREPS also are selected by default.

6) Click on the **Percentiles** radio button and specify the percentile points in the distribution. The example uses 5th, 25th, 75th, and 95th percentiles. These need to be typed in increasing order separated by spaces.

7) Specify the name and folder of the output files in the **Output Files** field.

8) Click the **Start SPSS** button to create the SPSS syntax file. The file will open in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** option. If necessary, the IEA IDB Analyzer will prompt for confirmation to overwrite existing files.

Figure 4.20 shows the IDB Analyzer Setup Screen for this analysis. Figure 4.21 shows the SPSS output obtained from SPSS after running the analysis.

Figure 4.20 Analysis Module Setup Screen for Computing Percentiles

Figure 4.21 SPSS Output for Percentiles

APercentiles for PVCIV by IDCNTRY										PAGE 1
COUNTRY ID	N of Cases	Sum of TOTWGTS	p5	p5_se	p25	p25_se	p75	p75_se	p95	p95_se
Austria	3385	88527	336,02	8,81	435,49	6,86	573,66	4,63	656,94	5,47
Bulgaria	3257	63557	295,64	7,43	388,80	8,56	544,22	8,17	632,33	7,36
Chile	5192	258422	343,59	7,23	419,72	5,04	544,43	4,64	629,18	6,30
Chinese Taipei	5167	303632	397,01	5,40	495,01	4,64	625,60	5,32	704,89	5,13
Colombia	6204	661787	329,29	6,15	405,21	4,24	518,14	4,24	594,48	4,98
Cyprus	3194	8872	303,63	5,95	386,49	3,88	518,35	3,93	607,31	6,72
Czech Republic	4630	95781	370,09	4,88	447,06	3,67	570,93	4,88	656,31	5,23
Denmark	4508	62233	410,26	7,19	509,31	5,99	644,74	5,56	736,22	5,93
. . .	.	.								
x.International Average	.	.	352,57	1,28	439,49	1,09	564,21	,94	646,03	1,09

The first few lines of the results displayed in Figure 4.21 show that in Austria the score of the 5th percentile of the score distribution is 336 points; for the 25th percentile, 435 points; for the 75th percentile, 574 points; and for the 95th percentile, 657 points. The corresponding standard errors for these percentiles are 8.8, 6.9, 4.6 and 5.4.

4.6   Performing Analyses with Teacher-Level Data

As already noted, student and teacher data cannot be merged and analyzed together because of the sampling design of ICCS 2009. The sample analysis using teacher background data presented here will investigate the percentage of teachers who report taking part in cultural activities (e.g., theatre, music, or cinema) with any of the target classes they teach. Table 6.2 (column five) of the *ICCS 2009 International Report* (Schulz et al., 2010b) presents the results of such an analysis. Figure 4.22 reproduces this analysis using the **Percent Only** analysis type to estimate the percentages of teachers reporting taking part in cultural activities with target classes.

Figure 4.22 Table of Sample Teacher-Level Analysis Taken from the ICCS 2009 International Report (Table 6.2)

Table 6.2: Principals' reports on participation of target-grade classes in community activities (in national percentages of students)

Percentages of Students Reported To Have Been Involved in …

Country	activities related to the environment and geared to the local area	human rights projects	activities related to underprivileged people or groups	cultural activities (e.g., theater, music, cinema)	multicultural and intercultural initiatives within the <local community>	campaigns to raise people's awareness, such as <World AIDS Day, World No Tobacco Day>	activities related to improving facilities for the <local community>	sports events
Austria	32 (4.2) ►	27 (4.3)	33 (4.6)	87 (3.2) ◄	18 (3.6) ►	65 (4.3)	11 (3.0) ►	84 (3.5)
Belgium (Flemish)†	63 (4.1) ◄	45 (4.8) ◄	68 (4.7) ◄	95 (1.5) ◄	33 (4.8)	73 (3.5) ◄	12 (2.5) ►	88 (2.6) ◁
Bulgaria	46 (4.6)	8 (2.6)	24 (3.5) ►	75 (3.7)	36 (4.8)	76 (3.4) ◄	37 (4.2) ◁	85 (3.1)
Chile	40 (3.8) ►	15 (2.8) ►	35 (3.7)	57 (3.7) ►	31 (3.5)	40 (4.1) ►	9 (1.9) ►	74 (3.5) ▽
Chinese Taipei	34 (4.1) ►	24 (3.9) ►	31 (4.1)	53 (4.1) ►	30 (4.1)	53 (4.8)	35 (4.3) ►	75 (3.6) ▽
Colombia	57 (4.0)	40 (3.3) ►	16 (2.7) ►	55 (3.4) ►	36 (3.4)	41 (3.3) ►	22 (3.2) ►	76 (3.3) ▽
Cyprus	21 (0.2) ►	19 (0.2) ►	11 (0.1) ►	41 (0.3) ◄	26 (0.2) ▽	19 (0.2) ►	13 (0.2) ►	46 (0.3) ►
Czech Republic †	74 (4.1) ►	42 (5.0) ►	34 (4.7)	98 (1.0) ◄	51 (4.8) ◄	77 (4.1) ◄	29 (4.3)	87 (2.9)
Denmark†	22 (3.7) ►	24 (3.8) ►	25 (3.8) ►	80 (3.1)	18 (3.6) ►	18 (3.5) ►	26 (3.8)	74 (3.9) ▽
Dominican Republic ‡	66 (6.7) ◄	38 (5.3) ◄	41 (4.7)	53 (6.2) ►	52 (6.3) ◄	74 (4.3) ◄	30 (4.1)	77 (3.9)
England ‡	49 (5.3)	47 (5.1) ◄	70 (3.9) ◄	89 (3.3) ◄	40 (5.5)	66 (4.7)	24 (4.6)	96 (2.2) ◄
Estonia	76 (3.8) ◄	23 (3.7) ►	15 (2.9) ►	99 (1.1) ◄	40 (3.9)	78 (3.5) ◄	56 (4.7) ◄	99 (0.9) ◄
Finland	39 (3.3) ►	15 (3.2) ►	48 (4.2) ◄	82 (2.9) ◁	28 (3.7)	88 (2.6) ◄	32 (3.9)	86 (2.5)
Greece	25 (3.5) ►	10 (2.8) ►	13 (3.4)	41 (4.1) ◄	11 (2.8) ►	22 (3.4)	6 (2.1) ►	50 (4.9) ►
Guatemala[1]	59 (4.6)	40 (4.8)	30 (4.1)	69 (4.3)	46 (4.8) ◄	44 (4.7) ►	37 (4.7) ◁	90 (2.1) ◁
Indonesia	67 (4.2) ◄	18 (3.1) ►	47 (4.5) ◄	34 (4.1) ►	17 (3.4) ◄	19 (3.6) ◄	34 (4.0)	79 (3.9)
Ireland	40 (3.7) ►	39 (4.6)	33 (4.3)	52 (4.4) ►	18 (3.4) ◄	21 (3.5) ►	10 (2.7) ►	79 (3.9)
Italy	60 (4.3) ◁	66 (3.6) ◄	44 (3.8) ◁	82 (3.1) ◁	47 (3.7) ◄	56 (3.8)	24 (3.6) ►	81 (2.8)
Korea, Republic of[1]	32 (3.6) ►	22 (3.4) ►	32 (3.9)	28 (3.8) ►	16 (3.0) ►	42 (3.8) ►	24 (3.4) ►	38 (4.3) ►
Latvia	43 (4.2) ►	30 (4.1) ►	31 (4.9)	96 (1.8) ◄	47 (4.4) ◄	53 (4.8) ◄	65 (4.2) ◄	98 (1.2) ◄
Liechtenstein	32 (0.4) ►	59 (0.4) ◄	59 (0.4) ◄	87 (0.3) ◄	0 (0.0)	75 (0.4) ◄	13 (0.3) ►	87 (0.4) ◁
Lithuania	55 (4.3) ◄	28 (4.2) ►	20 (3.3) ►	76 (3.4) ►	51 (3.5) ◄	67 (4.1) ◁	63 (3.9) ◄	97 (1.5) ◄
Luxembourg	23 (1.4) ►	32 (2.2) ►	39 (2.3) ►	63 (2.2) ►	35 (2.2)	74 (1.9) ◄	0 (0.0) ►	75 (2.3) ▽
Malta	42 (0.9) ▽	38 (0.9) ◁	48 (0.9) ◄	65 (1.0) ▽	19 (0.6) ►	39 (0.9) ►	13 (0.4) ►	94 (0.1) ◄
Mexico	66 (3.4) ◄	47 (3.7) ◄	32 (3.0) ►	54 (3.4) ►	40 (3.6)	60 (3.2)	32 (3.0)	67 (3.5) ►
New Zealand†	46 (5.1) ◄	40 (5.2) ◄	54 (5.7) ◄	81 (4.2) ◄	51 (4.5) ◄	62 (4.5) ◄	17 (3.9) ▽	97 (0.6) ◄
Norway †	38 (4.8) ►	31 (4.1)	37 (4.5)	90 (2.8) ◄	21 (3.6) ►	57 (5.2)	21 (4.1)	80 (3.3) ◄

Figure 4.22 Table of Sample Teacher-Level Analysis Taken from the ICCS 2009 International Report (Table 6.2) (continued)

Table 6.2: Principals' reports on participation of target-grade classes in community activities (in national percentages of students) (contd.)

Country	activities related to the environment, geared to the local area	human rights projects	activities related to underprivileged people or groups	cultural activities (e.g., theater, music, cinema)	multicultural and intercultural initiatives within the <local community>	campaigns to raise people's awareness, such as <World AIDS Day, World No Tobacco Day>	activities related to improving facilities for the <local community>	sports events
		Percentages of Students Reported To Have Been Involved in …						
Paraguay[1]	82 (3.0) ▲	49 (5.0) ▲	50 (4.2) ▲	84 (3.0) ▲	59 (4.3) ▲	61 (4.2)	53 (4.4) ▲	94 (2.0) ▲
Poland	63 (4.1) ▲	51 (4.3) ▲	50 (4.1) ▲	88 (2.7) ▲	33 (4.3)	92 (2.1) ▲	22 (3.6)	92 (2.2) ▲
Russian Federation	80 (3.1) ▲	36 (3.0)	49 (2.8) ▲	91 (1.9) ▲	42 (3.2) △	81 (2.8) ▲	32 (3.6)	95 (1.2) ▲
Slovak Republic[2]	74 (3.6) ▲	50 (4.5) ▲	34 (4.1)	93 (2.2) ▲	53 (4.5) ▲	63 (4.2)	36 (4.3) △	94 (1.9) ▲
Slovenia	68 (3.4) ▲	49 (4.6) ▲	39 (4.4)	90 (2.2) ▲	46 (3.7) ▲	85 (2.8) ▲	31 (3.4)	89 (2.7) △
Spain	63 (4.3) ▲	52 (4.2) ▲	44 (3.9)	86 (2.3) ▲	34 (4.1)	72 (4.0) ▲	14 (2.9) ▼	76 (3.9)
Sweden	35 (4.1) ▼	47 (4.1) ▲	34 (4.1)	92 (2.2) ▲	27 (3.3) ▽	30 (4.2) ▼	20 (3.5) ▽	81 (3.3)
Switzerland †	38 (6.1) ▼	15 (3.2) ▼	12 (3.2) ▼	85 (3.0) ▲	13 (2.5) ▼	52 (4.8)	13 (2.8) ▼	94 (2.1) ▲
Thailand†	66 (4.3) ▲	45 (4.1) ▲	46 (4.7)	71 (3.5)	59 (4.1) ▲	82 (3.4) ▲	69 (4.4) ▲	92 (2.2) ▲
ICCS average	**50 (0.7)**	**35 (0.6)**	**37 (0.6)**	**74 (0.5)**	**34 (0.6)**	**58 (0.6)**	**27 (0.6)**	**82 (0.5)**

Countries not meeting sampling requirements

Country	activities related to the environment, geared to the local area	human rights projects	activities related to underprivileged people or groups	cultural activities (e.g., theater, music, cinema)	multicultural and intercultural initiatives within the <local community>	campaigns to raise people's awareness, such as <World AIDS Day, World No Tobacco Day>	activities related to improving facilities for the <local community>	sports events
Hong Kong SAR	38 (6.5)	14 (5.1)	34 (6.5)	67 (6.4)	34 (5.5)	45 (7.4)	29 (6.2)	87 (4.9)
Netherlands	25 (9.4)	24 (7.2)	42 (8.8)	82 (7.7)	23 (9.3)	29 (10.3)	16 (5.2)	82 (5.1)

National percentage

▲ More than 10 percentage points above ICCS average
▽ Significantly below ICCS average
△ Significantly above ICCS average
▼ More than 10 percentage points below ICCS average

Notes:

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.
† Met guidelines for sampling participation rates only after replacement schools were included.
‡ Nearly satisfied guidelines for sampling participation only after replacement schools were included.
[1] Country surveyed the same cohort of students but at the beginning of the next school year.
[2] National Desired Population does not cover all of International Desired Population.

As with previous examples, the first steps are 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 (see Supplement 2 of this ICCS 2009 IDB User Guide). Since this analysis concerns teacher-level data, examine the teacher background data files for the variable that contains the information on teachers' reports on taking part in cultural activities (IT2G15D).

This example uses the merged data file ITGALLC2.SAV. The IEA IDB Analyzer analysis module automatically selects the variable that identifies the country (IDCNTRY) and the variables that contain the sampling information. These will be used to generate the replicate weights for the analysis.

The analysis module of the IEA IDB Analyzer performs the teacher-level analysis. Figure 4.23 shows the completed analysis window.

1) Open the analysis module of the IEA IDB Analyzer.

2) Specify the data file ITGALLC2.SAV as the **Analysis File**.

3) Select Percentages only as the **Analysis Type**.

4) Add the variable IT2G15D as a second **Grouping Variable**.

5) The software automatically defines the **Weight Variable**. As this sample analysis uses only teacher background data, TOTWGTT is selected by default. The **Jackknifing Variables** JKZONET and JKREPT also are selected by default.

6) Specify the name and folder of the output files in the **Output Files** field.

7) Click the **Start SPSS** button to create the SPSS syntax file. The file will open in an SPSS syntax window. Opening the **Run** menu of SPSS and selecting the **All o**ption will execute the syntax file. If necessary, the IEA IDB Analyzer will prompt for confirmation to overwrite existing files.

Figure 4.23 IDB Analyzer Setup for Example Teacher-Level Analysis
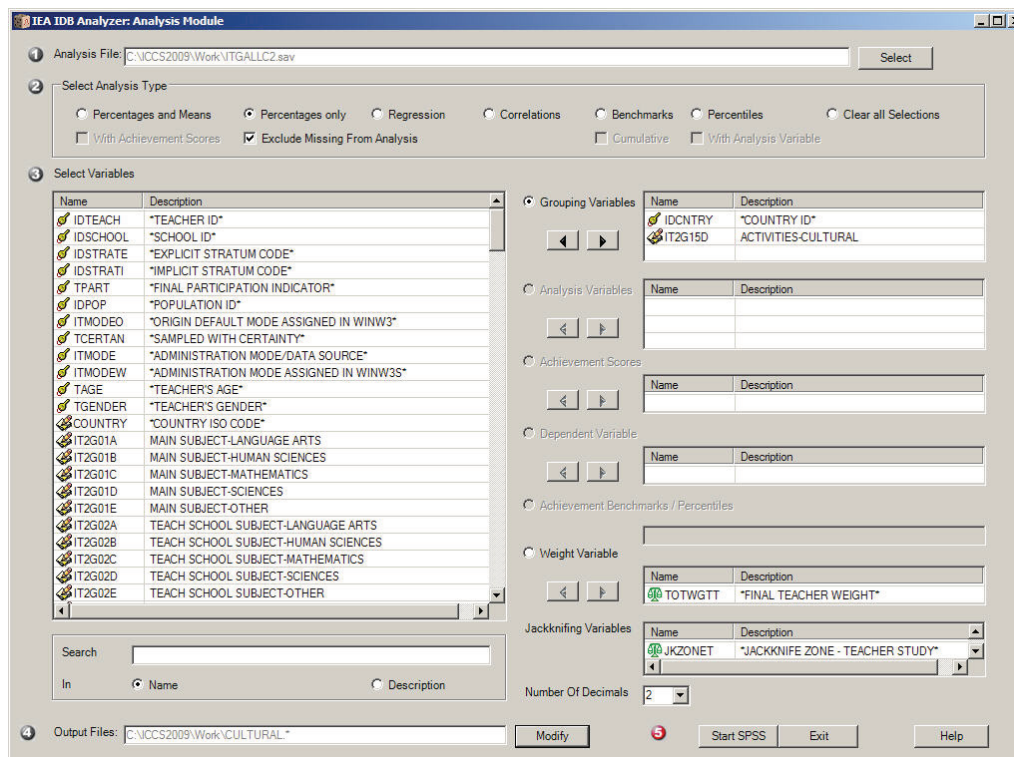
Figure 4.24 presents the results of this analysis. Each country's results are presented on two lines, one for each value of the IT2G15D variable. In this case the selected variable has two categories, yes and no. For categorical variables with more than two categories the output will show one line per category for each single country. The results are presented in the same manner as in the previous examples, with countries identified in the first column and the second column describing the categories of IT2G15D. From Figure 4.24, 63.99 percent of teachers in Austria reported they had taken part in cultural activities with the target classes they teach, while 36.01 percent of the teachers reported they had not. The standard error in both cases is equal to 1.97 percent.

Figure 4.24 Output for Example Teacher-Level Analysis

Percentages by (IDCNTRY IT2G15D)			PAGE 1		
COUNTRY ID	ACTIVITIES-CULTURAL	N of Cases	Sum of TOTWGTT	Percent	Percent (s.e.)
Austria	YES	586	21289	63.99	1.97
	NO	326	11980	36.01	1.97
Bulgaria	YES	1345	12433	73.01	2.19
	NO	491	4596	26.99	2.19
Chile	YES	856	21753	50.45	1.84
	NO	870	21366	49.55	1.84
Chinese Taipei	YES	1223	16212	52.05	1.39
	NO	1133	14934	47.95	1.39
Colombia	YES	1514	77060	75.57	1.87
	NO	484	24916	24.43	1.87
Cyprus	YES	439	1126	49.65	1.81
	NO	461	1142	50.35	1.81
Czech Republic	YES	1119	20910	70.54	1.43
	NO	469	8731	29.46	1.43
Denmark	YES	491	6625	54.67	2.35
	NO	411	5492	45.33	2.35
. . .					
x.International Average	YES	.	.	65.98	.33
	NO	.	.	34.02	.33

4.7 Performing Analyses with School-Level Data

When performing analyses with the merged school-level data, the data are analyzed to make statements about the number or percentages of students attending schools with a given characteristic, rather than about the number or percentages of schools with a given characteristic. An example of a school-level analysis will investigate the percentage of students who attend schools in which civic and citizenship education is taught as a separate subject by teachers in civic education. The example will calculate the average civic knowledge achievement within each of the two categories of the variable. The example presented here is not taken from the *ICCS 2009 International Report* (Schulz et al., 2010b).

The current example will use the ISG&ICGALLC2.SAV data file that contains school- and student-level data merged as described earlier. Note that in merging school- and student-level data only the **Total Student Weight** TOTWGTS and student **Jackknifing Variables** JKZONES and JKREPS are included in the merged file, not the school ones. This analysis will use the **Means and Percentages** analysis type in the IEA IDB Analyzer with the **With Achievement Scores** option checked. The first step is to identify the variables of interest from the appropriate files and review the documentation on specific national adaptations to the questions of interest (see Supplement 2). Variable IC2G16A in the school background data file contains information on the approach to teaching civic and citizenship in school as a separate subject taught by teachers in this subject.

The IEA IDB Analyzer automatically selects the variable identifying the country (IDCNTRY), as well as the variables that contain the sampling information used to generate the replicate weights for the analysis (JKZONES and JKREPS).

The analysis module of the IEA IDB Analyzer will perform the sample school-level analysis. Figure 4.25 shows the completed analysis window.

1) Open the analysis module of the IEA IDB Analyzer.

2) Specify the data file ISG&ICGALLC2.SAV as the **Analysis File**.

3) Select Percentages and Means as the **Analysis Type**.

4) Check the **With Achievement Scores** box.

5) Add the variable IC2G16A as a second **Grouping Variable**.

6) Click the **Achievement Scores** radio button. Select the variable PVCIVT01-05 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 software automatically defines the **Weight Variable**. As this sample analysis uses student background data as well as school background data disaggregated to the student level, TOTWGTS is selected by default. The **Jackknifing Variables** JKZONES and JKREPS also are selected by default.

8) Specify the name and folder of the output files in the **Output Files** field.

9) Click the **Start SPSS** button to create the SPSS syntax file. The file will open in an SPSS syntax window. The syntax file will be executed by opening the **Run** menu of SPSS and selecting the **All** option. If necessary, the IEA IDB Analyzer will prompt for permission to overwrite existing files.

Figure 4.25 IDB Analyzer Set-Up for Example Analysis with School-Level Data
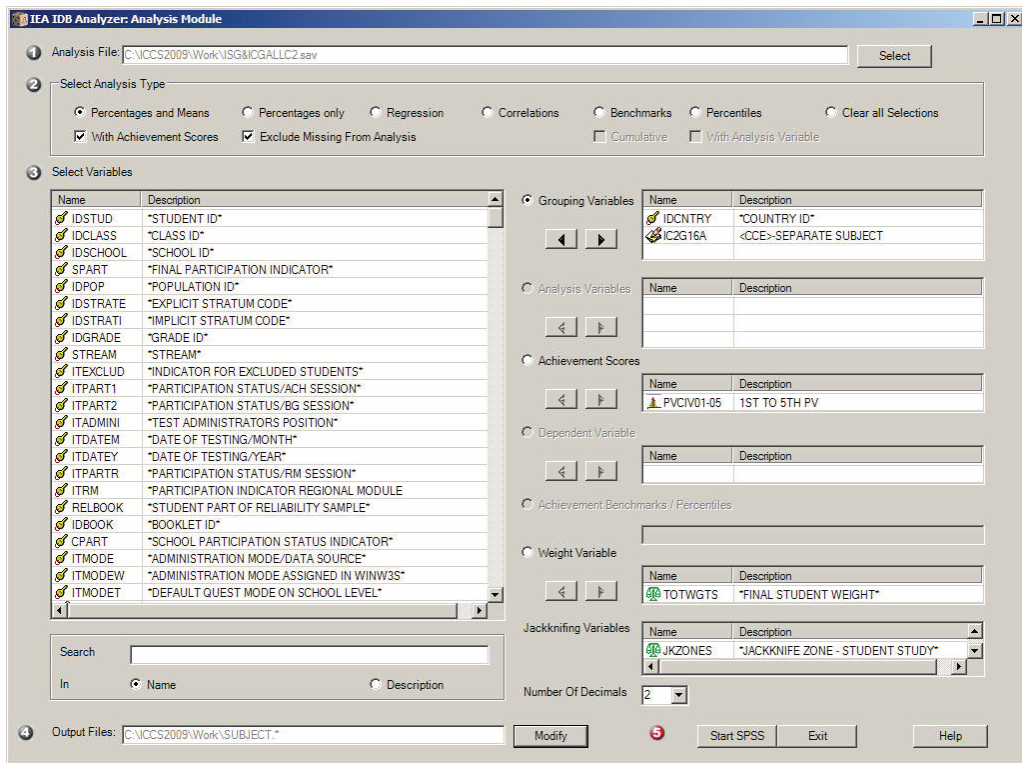
Figure 4.26 presents the results of this analysis. In this example, each country's results are listed on two lines, one for each value of the IC2G16A variable. The results are presented in the same manner as in the previous examples, with countries identified in the first column and the second column describing the categories of IC2G16A (Yes or No). The third and the fourth columns show the number of sampled students in each category and the actual number in the populations they represent. The fifth column represents the percentage of students for each of the two categories of the IC2G16A selected by the principals. The seventh column represents the mean civic knowledge achievement of the students for which the principals selected "Yes" or "No".

Figure 4.26 shows that 22.72 percent of target grade students in Austria attend schools with civic and citizenship education as a separate subject, and 77.28 percent attend schools where it is not. The standard errors of these percentages are 4.29 percent in both cases. The estimated mean civic knowledge achievement of students in schools with civic and citizenship education as a separate subject is 479.00 score points (with a standard error of 10.48 score points), whereas the estimated mean civic knowledge achievement of students in schools where citizenship education is not taught as a separate subject is 510.01 score point (standard error of 4.97 score points).

Figure 4.26 Output for Example Analysis with School-Level Data

Average for PVCIV by IDCNTRY IC2G16A									PAGE 1
COUNTRY ID	<CCE>-SEPARATE SUBJECT	N of Cases	Sum of TOTWGTS	Percent	Percent (s.e.)	PVCIV (Mean)	PVCIV (s.e.)	Std.Dev	Std.Dev. (s.e.)
Austria	YES	617	16237	22.72	4.29	479.00	10.48	99.92	4.20
	NO	2123	55226	77.28	4.29	510.01	4.97	93.74	2.58
Chile	YES	680	30127	11.77	1.98	484.87	11.86	91.44	4.11
	NO	4440	225880	88.23	1.98	481.78	3.65	86.62	1.70
Chinese Taipei	YES	4508	263949	86.93	2.65	562.03	2.82	93.56	1.19
	NO	659	39683	13.07	2.65	536.82	8.27	93.46	2.71
Colombia	YES	1747	187171	28.38	3.62	474.72	7.28	84.20	4.03
	NO	4437	472441	71.62	3.62	456.88	3.31	79.03	1.52
Czech Republic	YES	4159	86031	95.57	1.17	507.91	2.60	86.77	1.37
	NO	197	3990	4.43	1.17	521.53	19.83	87.12	7.16
Denmark	YES	3259	44942	83.53	2.95	578.19	4.04	97.31	1.77
	NO	701	8864	16.47	2.95	579.58	11.22	101.28	4.48
. . .									
x.International Average	YES	.	.	53.14	.60	496.32	1.34	87.88	.73
	NO	.	.	48.47	.62	494.69	1.37	88.07	.66

Analyzing the ICCS 2009 International Database Using SAS

5.1 Overview

Users of the International Civic and Citizenship Education Study (ICCS) 2009 International Database are encouraged to use the IEA IDB Analyzer in conjunction with Statistical Package for the Social Sciences (SPSS) because it is easy to use and deals effectively with the complexity of the ICCS 2009 data. Nevertheless, this chapter presents some basic examples of analyses that can be performed with the ICCS 2009 International Database using the SAS statistical analysis system (SAS, 2002) and the SAS programs and macros provided along with the ICCS 2009 International Database. The SAS macros use sampling weights and a jackknifing algorithm to deal with the complex sample design of ICCS 2009 and take into account plausible values when analyzing student achievement.

Although some familiarity with the structure of the ICCS 2009 database would be helpful, it is not essential. The analyses presented in this chapter are simple in nature and are designed primarily to familiarize users with the various data files, their structure, and the variables used in most analyses. Chapter 2 of this *ICCS 2009 IDB User Guide* 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 documentation provided together with the ICCS 2009 International Database.

The examples in this chapter compute percentages of students in specified subgroups, average civic knowledge achievement in those subgroups, and appropriate standard errors for these statistics. Additional examples compute regression coefficients and their standard errors. The sample analyses, using student, teacher, and school data, replicate some of the analyses that are included in the *ICCS 2009 International Report* (Schulz et al., 2010b). Users are encouraged to practice analyzing the ICCS 2009 data by replicating some of the figures presented in the international reports.

For the purposes of this chapter, analysts must copy all files of the ICCS 2009 International Database to the "C:\ICCS2009\" folder. All SAS programs presented in this chapter are available for download at the IEA Study Data Repository: http://rms.iea-dpc.org/. They can be easily adapted to perform a variety of analyses with even a basic knowledge of the SAS language to obtain the desired results. The example SAS programs invoke SAS macros that this chapter will describe. Although users will likely modify the sample programs, there is no need to make any changes to the SAS macros.

5.2 SAS Programs and Macros

The programs available for download at the IEA Study Data Repository (see above) include a number of SAS programs needed to process the SAS data files, compute survey results, and carry out sample analyses. This chapter gives detailed instructions on how to adapt and make use of them. The following programs are available:

CONVERT.SAS

This SAS program converts the SAS export files published as part of the ICCS 2009 data into SAS data files. All programs and macros described in this chapter require that the SAS export files be converted into SAS data files.

ISASCRC2.SAS, ISESCRC2.SAS, ISLSCRC2.SAS

These SAS programs may be used to convert the achievement item response codes to their corresponding score levels. Achievement items were administered in the Student Achievement Booklets, the European Module Questionnaire, and the Latin American Module Questionnaire.

JOIN.SAS

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

JACKGEN.SAS (and SAMPLEJACKGEN.SAS)

The SAS macro JACKGEN.SAS 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. The sample program SAMPLEJACKGEN.SAS provides an example of how to work with the JACKGEN.SAS macro. When computing mean achievement scores with plausible values, the macro JACKPV.SAS should be used.

JACKPV.SAS (and SAMPLEJACKPV.SAS)

The SAS macro JACKPV.SAS 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 JRR and multiple imputation methodologies. This macro should be used when achievement plausible values are used in an analysis. The sample program SAMPLEJACKPV.SAS provides an example of how to work with the JACKPV.SAS macro.

JACKREG.SAS (and SAMPLEJACKREG.SAS)

The SAS macro JACKREG.SAS 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. The sample program SAMPLEJACKREG.SAS provides an example of how to work with the JACKREG.SAS macro.

JACKREGP.SAS (and SAMPLEJACKREGP.SAS)

The SAS macro JACKREGP.SAS is used to compute weighted regression coefficients and their standard errors within defined subgroups when using achievement plausible values as the dependent variable. The sample program SAMPLEJACKREGP.SAS provides an example of how to work with the JACKREGP.SAS macro.

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

These are the programs used in the sample analyses presented in this chapter.

5.3   Converting the SAS Export Files

The program called CONVERT.SAS converts the SAS Export files 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.

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 ICCS 2009 countries are listed and the program will automatically select the appropriate list based on the file type specified.

6)   Submit the edited code for processing.

Figure 5.1 presents an example of the CONVERT program. This example converts the SAS export files with the Student Questionnaire data (type ISG) to SAS data files for all countries. For this example, all SAS export files are located in the "C:\ICCS2009\Data\SAS_Data" folder and the converted SAS data files will also be located in this folder.

Figure 5.1 Example of CONVERT Program Used to Convert SAS Export Files into SAS Data Files

```
%LET TYPE = ISG ;
%LET EXPPATH = C:\ICCS2009\Data\SAS_Data\ ;
%LET DATPATH = C:\ICCS2009\Data\SAS_Data\ ;
%MACRO DOIT ;
    %LET COUNTRY = < List of ICCS 2009 countries > ;
    %LET I = 1 ;
    %DO %WHILE(%LENGTH(%SCAN(&COUNTRY,&I))) ;
        %LET CTRY = %SCAN(&COUNTRY,&i) ;
        PROC CIMPORT FILE = "&EXPPATH&TYPE&CTRY.C2.EXP"
                     DATA = "&DATPATH&TYPE&CTRY.C2" ;
        %LET I = %EVAL(&I + 1) ;
    %END ;
%MEND DOIT ;
%DOIT ;
```

Users are advised to run the CONVERT program for all countries and all file types. The file types of the target grade data are ICG, ITG, ISA, ISG, ISR, ISS, ISE, and ISL. For analyses using the additional grade data, the file types are JSA, JSG, JSR, and JSE. File types are described in Chapter 2 of this ICCS 2009 IDB User Guide. In principle, this program needs to be run only once for each file type and should be one of the first thing users do with the ICCS 2009 International Database before moving on to any data analyses, more specifically the data analysis examples in this ICCS 2009 IDB User Guide.

5.4   Scoring Individual ICCS 2009 Items

Student achievement in ICCS 2009 is represented by a set of five plausible values for the civic knowledge scale 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 ICCS 2009 database be assigned their correctness score levels, rather than the actual response options selected by students for multiple choice items, or the one-digit codes given to students' responses to constructed response items. A SAS program is available to perform this task.

For multiple choice items, numbers 1 through 4 are used to represent response options A through D, respectively, in the ICCS 2009 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 one or two points, one-digit codes are used to represent the students' written responses in the ICCS 2009 database. These codes do not need to be recoded. They already represent the correct point values of the responses—either zero, one, or two 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 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 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 ICCS 2009 items, whereas they were treated as "incorrect" when deriving achievement scores for students.

The SAS programs ISASCRC2.SAS, ISESCRC2.SAS, and ISLSCRC2.SAS will recode the responses to single items from the achievement, European, and Latin American Module data files to their appropriate score levels. To score each single ICCS 2009 item from the Student Achievement Booklets, the program ISASCRC2.sas needs to be used. To score the achievement items from the first part of the European Module Questionnaire, the program ISESCRC2.SAS needs to be used. Finally, to score the achievement items from the Latin American Module Questionnaire, the program ISLSCRC2.SAS needs to be used. The program code in each of these programs needs to be adapted. Users should perform the following steps:

1) Open the SAS program file ISASCRC2.SAS / ISESCRC2.SAS / ISLSCRC2.SAS.

2) Specify the folder where the SAS data files are located in the "LIBNAME" statement.

3) Specify the desired grade in the parameter "GRADE". By default, the target grade (I) is selected. To change to the additional grade data (J), replace the "I" with a "J". Note that there is only additional grade data for the international student data and the European Module data.

4) List all the countries of interest in the parameter "COUNTRY". By default, all ICCS 2009 countries are listed.

5) Submit the edited code for processing.

Each program uses the student data files that contain achievement data as input (ISA/JSA, ISE/JSE, ISL), recodes the individual items, and saves the results in SAS data files that have "ICA/JCA" instead of "ISA/JSA" as the first three characters in their file names. For the European Module data, the resulting scored SAS data file names will begin with "ICE/JCE", and for the Latin American Module data, "ICL". Figure 5.2 shows a condensed version of the ISASCRC2.SAS program to score the ICCS 2009 international items.

Figure 5.2 Example of ISASCRC2 Program for Converting Individual Item Response Codes to their Score Level

```
LIBNAME LIBDAT "C:\ICCS2009\Data\SAS_Data\" ;
%LET GRADE = I ;
%LET COUNTRY = < List of ICCS 2009 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 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 CONSTR ; %SCOREIT (CONSTR, "CR", , .R, .A, ., .I) ; END ;
 . . .
%MEND DOIT ;
%DOIT ;
```

If "not reached" responses are to be treated as missing rather than as 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 = . ;

5.5   Joining the ICCS 2009 Data Files

The ICCS 2009 International Database contains separate data files for each country. A SAS program called JOIN.SAS is available that joins individual country data files of a particular file type into a single aggregated data file, facilitating joint analyses involving more than one country. 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: ICG, ITG, ISA/JSA, ISC/JSC, ISG/JSG, ISR/JSR, ISS, ISE/JSE, ICE/JCE, ISL, and ICL. 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".

5)   Submit the edited code for processing.

An example of the JOIN program is displayed in Figure 5.3. It joins the target grade student background data files (ISG) of all countries. All country data files are located in the "C:\ICCS2009\Data\SAS_Data" folder for the sake of this example. The resulting data file, ISGALLC2, will also be saved in this folder.

Figure 5.3 Example of JOIN Program Used to Join SAS Data Files for More than One Country

```
%LET TYPE = ISG ;
LIBNAME LIBDAT "C:\ICCS2009\Data\SAS_Data\" ;
%MACRO DOIT ;
    %LET COUNTRY = < List of ICCS 2009 countries > ;
    DATA &TYPE.ALLC2 ;
        SET %LET I = 1 ;
        %DO %WHILE(%LENGTH(%SCAN(&COUNTRY,&I))) ;
            %LET CTRY = %SCAN(&COUNTRY,&I) ;
            LIBDAT.&TYPE&CTRY.C2
            %LET I = %EVAL(&I + 1) ;
        %END ; ;
    PROC SORT DATA = &TYPE.ALLC2 OUT = LIBDAT.&TYPE.ALLC2 ;
        BY &SORTVARS ;
%MEND DOIT ;
%DOIT ;
```

5.6 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 ICCS 2009 data are done properly. Sampling weights are used and standard errors are computed using the 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 ISGALLC2, which contains the target grade student questionnaire data files of all participating countries. In all sample programs, <datpath> must be edited to specify the folder where the ISGALLC2 file is located.

5.6.1 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. This sample SAS program uses 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> indicates the folder where the SAS macro JACKGEN.SAS is located. The macro requires that several parameters be specified as input when it is invoked. These parameters are:

WGT The sampling weight to be used in the analysis. Generally, TOTWGTS should be used for analysis at the student-level. For analysis at the school-level, TOTWGTC should be used and TOTWGTT for teacher level analysis.

JKZ The variable that captures the assignment of cases to sampling zones. The name of this variable is JKZONES in student-level data files, JKZONET in teacher-level data files, and JKZONEC in school-level data files.

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 is JKREPS in student-level data files, JKREPT in teacher-level data files, and JKREPC in school-level data files.

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 quotation marks. It is important to emphasize that this data file must include only those cases that are of interest in the analysis. If users wish to exclude specific cases 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 can be invoked using the following statement:

```
%JACKGEN (TOTWGTS, JKZONE, JKREP, 75, IDCNTRY SGENDER, SAGE,
ISGALLC2) ;
```

The macro will compute the mean age (SAGE) of target grade students by gender (SGENDER) and their standard errors within each country (IDCNTRY), using the weighting variable for student-level analysis TOTWGTS. It will also compute the percentages of boys and girls and their standard errors within each country. The data will be read from the data file ISGALLC2 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 above, there are two classification variables: IDCNTRY 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 TOTWGTS since TOTWGTS 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 Figure 5.4. This program is available in the file called SAMPLEJACKGEN.SAS. It produces the mean ages for target grade boys and girls in all countries. The figure shows the results for only the first four countries.

Figure 5.4 Sample SAS Program Invoking the SAS Macro JACKGEN and Results

```
LIBNAME ICCS2009 "<datpath>" ;
%INCLUDE "<macpath>JACKGEN.SAS" ;
DATA ISGALLC2 ;
    SET ICCS2009.ISGALLC2 ;
    WHERE NMISS (SGENDER, SAGE) = 0 ;
PROC FORMAT LIBRARY = WORK ;
VALUE COUNTRY
        < list ICCS 2009 country formats > ;
    VALUE SEX
        0 = 'BOY'
        1 = 'GIRL' ;
%JACKGEN (TOTWGTS, JKZONES, JKREPS, 75, IDCNTRY SGENDER, SAGE,
ISGALLC2) ;
PROC PRINT DATA = FINAL NOOBS ;
VAR IDCNTRY SGENDER N TOTWGTS MNX MNX_SE PCT PCT_SE ;
FORMAT IDCNTRY COUNTRY. SGENDER SEX. N 6.0 TOTWGTS 10.0
MNX MNX_SE PCT PCT_SE 6.2 ;
```

IDCNTRY	SGENDER	N	TOTWGTS	MNX	MNX_SE	PCT	PCT_SE
AUT	BOY	1520	40758	14.41	0.02	49.85	1.39
AUT	GIRL	1612	41006	14.30	0.02	50.15	1.39
BGR	BOY	1567	29974	14.73	0.02	48.09	1.65
BGR	GIRL	1627	32360	14.65	0.01	51.91	1.65
CHL	BOY	2490	125302	14.25	0.02	49.05	1.44
CHL	GIRL	2640	130154	14.13	0.02	50.95	1.44
TWN	BOY	2670	155929	14.20	0.01	51.58	0.57
TWN	GIRL	2474	146348	14.20	0.01	48.42	0.57

From the first two lines of the results shown in Figure 5.4, there are 1,520 boys in the Austrian target grade sample representing 40,758 boys in the whole population. The mean age for target grade boys in Austria is estimated to be 14.41 with a standard error of 0.02. Boys made up 49.85% of the target grade student population in Austria. Conversely, Austria sampled 1,612 girls representing 41,006 girls in the whole target grade population. The estimated mean age for target grade girls in Austria is 14.30 with a standard error of 0.02. Girls made up 50.15 percent of the target grade student population in Austria.

5.6.2 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 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 may be included in a SAS program by issuing the following command:

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

The term <macpath> indicates 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. These parameters are:

WGT   The sampling weight to be used in the analysis. Generally, TOTWGTS should be used for analysis at student-level. For analysis at school-level, TOTWGTC should be used and TOTWGTT for teacher-level analysis respectively.

JKZ   The variable that captures the assignment of cases to sampling zones. The name of this variable is JKZONES in student-level data files, JKZONET in teacher-level data files, and JKZONEC in school-level data files.

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 is JKREPS in student-level data files, JKREPT in teacher-level data files, and JKREPC in school-level data files.

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.

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 quotation marks. 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 is invoked using the following statement:

```
%JACKPV (TOTWGTS, JKZONE, JKREP, 75, IDCNTRY SGENDER, 5, ISGALLC2) ;
```

The macro will compute the mean achievement of target grade students by gender (SGENDER) within each country (IDCNTRY) and their standard errors, using the weighting variable TOTWGTS. The macro uses all five plausible values to compute these statistics. It will also compute the percentages of boys and girls within each country, and their standard errors. The data will be read from the data file ISGALLC2 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: IDCNTRY 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 TOTWGTS since TOTWGTS 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 Figure 5.5. This program is available in the file called SAMPLEJACKPV.SAS. It produces the mean civic knowledge achievement score for target grade boys and girls in all countries. Figure 5.5 provides the results of this analysis for the first four countries.

Figure 5.5 Sample SAS Program Invoking the SAS Macro JACKPV and Results

```
LIBNAME ICCS2009 "<datpath>" ;
%INCLUDE "<macpath>JACKPV.SAS" ;
DATA ISGALLC2 ;
    SET ICCS2009.ISGALLC2 ;
    WHERE NMISS (SGENDER) = 0 ;
PROC FORMAT LIBRARY = WORK ;
    VALUE COUNTRY
        < list ICCS 2009 country formats > ;
    VALUE SEX
        0 = 'BOY'
        1 = 'GIRL' ;
%JACKPV (TOTWGTS, JKZONES, JKREPS, 75, IDCNTRY SGENDER, 5, ISGALLC2) ;
PROC PRINT DATA = FINAL NOOBS ;
    VAR IDCNTRY SGENDER N TOTWGTS MNPV MNPV_SE PCT PCT_SE ;
    FORMAT IDCNTRY COUNTRY. SGENDER SEX. N 6.0 TOTWGTS 10.0
        MNPV MNPV_SE PCT PCT_SE 6.2 ;
```

IDCNTRY	SGENDER	N	TOTWGTS	MNPV	MNPV_SE	PCT	PCT_SE
AUT	BOY	1553	41734	496.47	4.45	50.07	1.41
AUT	GIRL	1637	41624	512.60	4.74	49.93	1.41
BGR	BOY	1590	30431	453.51	5.94	48.21	1.65
BGR	GIRL	1642	32687	479.30	5.17	51.79	1.65
CHL	BOY	2510	126397	476.23	4.04	49.17	1.45
CHL	GIRL	2651	130659	489.83	4.21	50.83	1.45
TWN	BOY	2670	155929	546.12	2.84	51.58	0.57
TWN	GIRL	2474	146348	572.55	2.66	48.42	0.57

From the first two lines of the results presented in Figure 5.5, the mean civic knowledge score of Grade 8 boys in Austria is estimated to be 496.47 with a standard error of 4.45. The mean civic knowledge score of Grade 8 girls in Austria is estimated to be 512.60 with a standard error of 4.74.

5.6.3 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. For the latter kind of analysis the JACKREGP macro should be used.

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 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" ;
```

In this command, <macpath> indicates 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. These parameters are:

WGT   The sampling weight to be used in the analysis. Generally, TOTWGTS should be used for analysis at student-level. For analysis at the school level, TOTWGTC should be used and for the teacher level, TOTWGTT.

JKZ   The variable that captures the assignment of cases to sampling zones. The name of this variable is JKZONES in student-level data files, JKZONET in teacher-level data files, and JKZONEC in school-level data files.

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 is JKREPS in student-level data files, JKREPT in teacher-level data files, and JKREPC in school-level data files.

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 SGENDER, 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 quotation marks. 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 using the following statement:

```
%JACKREG (TOTWGTS, JKZONES, JKREPS, 75, IDCNTRY, REGSEX, SAGE,
ISGALLC2) ;
```

will perform a linear regression with gender (REGSEX) as a predictor of the target grade students' age at the time of testing (SAGE), using the weighting variable TOTWGTS. It will compute the regression coefficients and their standard errors. The data will be read from the data file ISGALLC2 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 IDCNTRY. 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, and based on the order of 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 Figure 5.6. This program is available in the file called SAMPLEJACKREG.SAS. It performs a linear regression in each country, with the variable REGSEX as a predictor of the target grade students' age at the time of testing (SAGE). Figure 5.6 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 SGENDER, such that the value zero represents the boys and the value one represents the girls, and all missing data are coded as omitted responses. By performing this recoding, the intercept B00 will be the estimated mean age of target grade boys, whereas the regression coefficient B01 will be the estimated increase in mean age for girls. This will determine whether the difference in mean ages between girls and boys is statistically significant.

From the first line of the results displayed in Figure 5.6, the estimated mean age of target grade boys in Austria (B00) is 14.41 years, with a standard error of 0.02. The target grade girls in Austria are an estimated 0.10 years younger (B01) than the boys. With an estimated standard error of 0.02, this difference is statistically significant at a 95% confidence level.

Figure 5.6 Sample SAS Program Invoking the SAS Macro JACKREG and Results

```
LIBNAME ICCS2009 "<datpath>" ;
%INCLUDE "<macpath>JACKREG.SAS" ;
DATA ISGALLC2 ;
    SET ICCS2009.ISGALLC2;
    WHERE NMISS (SGENDER, SAGE) = 0 ;
    SELECT (SGENDER) ;
        WHEN (0) REGSEX = 0 ; * BOYS ;
        OTHERWISE REGSEX = . ; * ALL MISSING DATA SET TO OMITTED ;
    END ;
PROC FORMAT LIBRARY = WORK ;
    VALUE COUNTRY
        < list ICCS 2009 country formats >
%JACKREG (TOTWGTS, JKZONES, JKREPS, 75, IDCNTRY, SGENDER, SAGE,
ISGALLC2) ;
PROC PRINT DATA = REG NOOBS ;
    VAR IDCNTRY N MULT_RSQ SS_TOTAL SS_REG B00 B00_SE B01 B01_SE ;
    FORMAT IDCNTRY COUNTRY. N 6.0 MULT_RSQ 5.3 SS_TOTAL SS_REG 10.0
        B00 B00_SE B01 B01_SE 6.2 ;


IDCNTRY    N    MULT_RSQ   SS_TOTAL   SS_REG   B00  B00_SE    B01   B01_SE
  AUT     3132    0.009      24183      225    14.41  0.02   -0.10   0.02
  BGR     3194    0.006      15039       86    14.73  0.02   -0.07   0.02
  CHL     5130    0.009     104246      954    14.25  0.02   -0.12   0.02
  TWN     5144    0.000      29253        1    14.20  0.01    0.00   0.01
```

5.6.4 *Computing Regression Coefficients and Their Standard Errors with Achievement Scores (JACK-REGP)*

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 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" ;
```

In this command, <macpath> indicates 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. These parameters are:

WGT   The sampling weight to be used in the analysis. Generally, TOTWGTS should be used for analysis at the student level. For analysis at the school level, TOTWGTC should be used and at the teacher level, TOTWGTT.

JKZ   The variable that captures the assignment of cases to sampling zones. The name of this variable is JKZONES in student-level data files, JKZONET in teacher-level data files, and JKZONEC in school-level data files.

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 is JKREPS in student-level data files, JKREPT in teacher-level data files, and JKREPC in school-level data files.

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 SGENDER for example.

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 quotation marks. It is important to emphasize that this data file must include only those cases that are of interest in the analysis. If users wish to exclude specific cases 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 typing the macro name followed by the list of parameters in parenthesis, with each parameter separated by a comma. For example, the JACKREGP macro invoked using the following statement:

```
%JACKREGP (TOTWGTS, JKZONES, JKREPS, 75, IDCNTRY, REGSEX, 5,
ISGALLC2) ;
```

will perform a linear regression with gender (REGSEX) as a predictor of civic knowledge achievement score of target grade students based on its five plausible values (PV1CIV through PV5CIV), using the weighting variable TOTWGTS. It will compute the regression coefficients and their standard errors. The data will be read from the data file ISGALLC2 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 IDCNTRY. 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, and based on the order of 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 Figure 5.7. This program called SAMPLEJACKREGP.SAS is available as part of the ICCS 2009 International Database. It performs a linear regression in each country, with the variable REGSEX as a predictor of the civic knowledge score. Figure 5.7 displays the results for the first four countries.

The regression performed by the sample program uses the variable REGSEX that was defined in the previous example. By using REGSEX, the intercept B00 will be the estimated mean civic knowledge score of target grade boys, whereas the regression coefficient B01 will be the estimated difference in the mean civic knowledge score of girls. This will allow analysts to determine if the target grade civic knowledge score is significantly different between girls and boys.

From the first line of the results shown in Figure 5.7, the estimated mean civic knowledge of target grade boys in Austria (B00) is 496.47, with a standard error of 4.45. Note that these are the same results obtained from the JACKPV sample program (Figure 5.5). The target grade girls in Austria have an estimated mean civic knowledge score of 16.14 points (B01) higher than boys. With an estimated standard error of 4.75, this difference is statistically significant at a 95% confidence level.

Figure 5.7 Sample SAS Program Invoking the SAS Macro JACKREGP and Results

```
LIBNAME ICCS2009 "<datpath>" ;
%INCLUDE "<macpath>JACKREGP.SAS" ;
DATA ISGALLC2 ;
    SET ICCS2009.ISGALLC2;
    WHERE NMISS (SGENDER) = 0 ;
    SELECT (SGENDER) ;
        WHEN (0) REGSEX = 0 ; * GIRLS ;
        WHEN (1) REGSEX = 1 ; * BOYS ;
        OTHERWISE REGSEX =. ; * ALL MISSING DATA SET TO OMITTED ;
    END ;
    PVCIV01 = PV1CIV ;
    PVCIV02 = PV2CIV ;
    PVCIV03 = PV3CIV ;
    PVCIV04 = PV4CIV ;
    PVCIV05 = PV5CIV ;
PROC FORMAT LIBRARY = WORK ;
    VALUE COUNTRY
        < list ICCS 2009 country formats > ;
%JACKREGP (TOTWGTS, JKZONES, JKREPS, 75, IDCNTRY, REGSEX, PVCIV0, 5,
ISGALLC2) ;
PROC PRINT DATA = REG NOOBS ;
    VAR IDCNTRY N MULT_RSQ SS_TOTAL SS_REG B00 B00_SE B01 B01_SE ;
    FORMAT IDCNTRY COUNTRY. N 6.0 MULT_RSQ 5.3 SS_TOTAL SS_REG 10.0
        B00 B00_SE B01 B01_SE 6.2 ;
```

IDCNTRY	N	MULT_RSQ	SS_TOTAL	SS_REG	B00	B00_SE	B01	B01_SE
AUT	3190	0.007	776107080	5444987	496.47	4.45	16.14	4.75
BGR	3232	0.015	697796563	10491171	453.51	5.94	25.79	5.23
CHL	5161	0.006	1967465824	11921907	476.23	4.04	13.61	4.70
TWN	5144	0.020	2664352847	52715928	546.12	2.84	26.42	2.46
COL	6192	0.000	4312376501	1495749	460.63	4.20	2.78	4.11
CYP	3088	0.047	73780108	3480026	434.81	3.21	40.27	3.57

5.7   ICCS 2009 Analyses with Student-Level Variables

Many analyses of the ICCS 2009 data can be undertaken using only student-level data.
Examples in the previous sections illustrate the functioning of the SAS macros. This section
presents examples of actual analyses used to produce the Figures in the *ICCS 2009 International
Report* (Schulz et al., 2010b), using SAS programs also provided as part of the ICCS 2009
International Database.

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.

5.7.1 Student-Level Analysis

The first example replicates the analysis of Grade 8 students' reported age at the time of testing. The results, presented in Figure 3.10 of the *ICCS 2009 International Report* (Schulz et al., 2010b), are reproduced here in Figure 5.8. This example will focus on the results presented in the third column—the average age at the time of testing.

Users need to undertake a number of steps to replicate the results displayed in this figure. By reviewing the student questionnaire data codebook (the codebooks are described in Chapter 2), users can identify the student background variable SAGE as the variable that reports the age of students at the time of testing.

The variable of interest (SAGE), the student sampling weight (TOTWGTS), the variables that contain the jackknife replication information (JKZONES and JKREPS), and the variable containing the country identification code (IDCNTRY) are all included in the student questionnaire data files. This analysis will use the data for all countries available. To prepare the data, the JOIN program described earlier in this chapter has been used to join the student questionnaire data files for all countries into in a single file called ISGALLC2.

Figure 5.9 presents the SAS program used to perform the first example and which is available as part of the database as EXAMPLE1.SAS. Figure 5.10 displays the results obtained from the program for the first four countries. Note that one of the steps in this program is to select only those students who have non-missing data in the variables of interest, SAGE. In general, to perform student-level analyses using the student questionnaire data files, users should do the following:

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 (<datpath>) and the macros (<macpath>).

6) Print the results file.

Figure 5.8 Sample Student-Level Analysis Taken from the ICCS 2009 International Report *(Figure 3.10)*

Table 3.10: Country averages for civic knowledge, years of schooling, average age, Human Development Index, and percentile graph

Country	Years of schooling	Average age	Civic Knowledge (Percentile graph 200–800)	Average scale score		HDI
Finland	8	14.7		576	(2.4) ▲	0.96
Denmark †	8	14.9		576	(3.6) ▲	0.96
Korea, Republic of[1]	8	14.7		565	(1.9) ▲	0.94
Chinese Taipei	8	14.2		559	(2.4) ▲	0.94
Sweden	8	14.8		537	(3.1) ▲	0.96
Poland	8	14.9		536	(4.7) ▲	0.88
Ireland	8	14.3		534	(4.6) ▲	0.97
Switzerland †	8	14.7		531	(3.8) ▲	0.96
Liechtenstein	8	14.8		531	(3.3) ▲	0.95
Italy	8	13.8		531	(3.3) ▲	0.95
Slovak Republic[2]	8	14.4		529	(4.5) ▲	0.88
Estonia	8	15.0		525	(4.5) ▲	0.88
England ‡	9	14.0		519	(4.4) ▲	0.95
New Zealand †	9	14.0		517	(5.0) ▲	0.95
Slovenia	8	13.7		516	(2.7) ▲	0.93
Norway †	8	13.7		515	(3.4) ▲	0.97
Belgium (Flemish) †	8	13.9		514	(4.7) ▲	0.95
Czech Republic †	8	14.4		510	(2.4) ▲	0.90
Russian Federation	8	14.7		506	(3.8)	0.82
Lithuania	8	14.7		505	(2.8)	0.87
Spain	8	14.1		505	(4.1)	0.96
Austria	8	14.4		503	(4.0)	0.96
Malta	9	13.9		490	(4.5) ▼	0.90
Chile	8	14.2		483	(3.5) ▼	0.88
Latvia	8	14.8		482	(4.0) ▼	0.87
Greece	8	13.7		476	(4.4) ▼	0.94
Luxembourg	8	14.6		473	(2.2) ▼	0.96
Bulgaria	8	14.7		466	(5.0) ▼	0.84
Colombia	8	14.4		462	(2.9) ▼	0.81
Cyprus	8	13.9		453	(2.4) ▼	0.91
Mexico	8	14.1		452	(2.8) ▼	0.85
Thailand †	8	14.4		452	(3.7) ▼	0.78
Guatemala[1]	8	15.5		435	(3.8) ▼	0.70
Indonesia	8	14.3		433	(3.4) ▼	0.73
Paraguay[1]	9	14.9		424	(3.4) ▼	0.76
Dominican Republic	8	14.8		380	(2.4) ▼	0.78

Countries not meeting sampling requirements

Hong Kong SAR	8	14.3		554	(5.7)	0.94
Netherlands	8	14.3		494	(7.6)	0.96

Percentiles of performance
5th 25th 75th 95th
Mean and confidence interval (±2SE)

▲ Achievement significantly higher than the ICCS average
▼ Achievement significantly lower than the ICCS average

Notes:

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

† Met guidelines for sampling participation rates only after replacement schools were included.

‡ Nearly satisfied guidelines for sample participation only after replacement schools were included.

[1] Country surveyed the same cohort of students but at the beginning of the next school year.

[2] National Desired Population does not cover all of International Desired Population.

Figure 5.9 Sample SAS Program to Perform Student-Level Analysis (EXAMPLE1.SAS)

```
LIBNAME ICCS2009 "<datpath>" ;
%INCLUDE "<macpath>JACKGEN.SAS" ;
DATA ISGALLC2 ;
    SET ICCS2009.ISGALLC2;
    WHERE NMISS (SAGE) = 0 ;
PROC FORMAT LIBRARY = WORK ;
    VALUE COUNTRY
        < list ICCS 2009 country formats > ;
%JACKGEN (TOTWGTS, JKZONES, JKREPS, 75, IDCNTRY, SAGE, ISGALLC2) ;
PROC PRINT DATA = FINAL NOOBS ;
    VAR IDCNTRY N TOTWGTS MNX MNX_SE PCT PCT_SE ;
    FORMAT IDCNTRY COUNTRY. N 6.0 TOTWGTS 10.0
        MNX MNX_SE PCT PCT_SE 6.2 ;
```

Figure 5.10 reports each country's mean value for the SAGE variable for all sampled students. The first column identifies the countries and the second column reports the number of valid cases. The third column reports the sum of weights of the sampled students, followed by the mean for SAGE 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, Austria has valid data for 3,135 students and these sampled students represent a population of 81,859 students. Students in Austria were, on average, 14.36 years old at the time the ICCS 2009 assessment took place, with a standard error of 0.02.

Figure 5.10 Output for Example Student-Level Analysis (Example 1)

IDCNTRY	N	TOTWGTS	MNX	MNX_SE	PCT	PCT_SE
AUT	3135	81859	14.36	0.02	0.64	0.02
BGR	3197	62405	14.69	0.01	0.49	0.02
CHL	5131	255497	14.19	0.02	2.00	0.06
TWN	5155	302974	14.20	0.00	2.37	0.04

5.7.2 Student-Level Analysis with Achievement Scores

The second example replicates another set of results presented in the *ICCS 2009 International Report* (Schulz et al., 2010b), the relationship between target grade students' gender and civic knowledge. These results, presented in Figure 3.13 of the *ICCS 2009 International Report*, are repeated here in Figure 5.11. Since the results in this figure are based on plausible values, the example will use the macro JACKPV. The codebook indicates that the variable SGENDER in the student questionnaire data files contains information on students' gender.

The student questionnaire data files contain the variable of interest (SGENDER), the five plausible values of civic knowledge (PV1CIV through PV5CIV), the student sampling weight (TOTWGTS), the variables that contain the jackknifing information (JKZONES and JKREPS), and the country identification variable (IDCNTRY). The example uses data from all available countries contained in the file ISGALLC2.

Figure 5.12 presents the SAS program used to implement the second example. It is available as part of the database as EXAMPLE2.SAS. Note that one of the steps in this program is to select only those students who have non-missing data in the variable of interest, SGENDER. The results obtained from this program are shown in Figure 5.13. For the sake of conciseness, only the results of the first four countries, sorted alphabetically, are shown.

To perform student-level analyses using the student questionnaire data files and civic knowledge scores, users should do the following:

1) Identify the variables of interest in the student questionnaire data files and note any specific national adaptations to the variables.

2) Retrieve the relevant variables from the student questionnaire data files, including the plausible values of civic knowledge, 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 (<datpath>) and the macros (<macpath>).

6) Print the results file.

Figure 5.11 Sample Student-Level Analysis with Civic Knowledge Scores Taken from the ICCS 2009 International Report (Figure 3.13)

Table 3.13: Gender differences in civic knowledge

Country	Mean Scale Score Females	Mean Scale Score Males	Difference (males–females)	Gender Difference
Guatemala[1]	435 (4.2)	434 (4.3)	-2 (3.7)	
Colombia	463 (3.1)	461 (4.0)	-3 (4.1)	
Belgium (Flemish) †	517 (5.3)	511 (5.6)	-6 (5.8)	
Switzerland †	535 (3.0)	528 (5.5)	-7 (4.6)	
Denmark †	581 (3.4)	573 (4.5)	-8 (3.5)	
Luxembourg	479 (2.8)	469 (3.4)	-10 (4.5)	
Liechtenstein	539 (6.4)	526 (6.2)	-12 (10.4)	
Chile	490 (4.3)	476 (4.2)	-14 (4.8)	
Austria	513 (4.6)	496 (4.5)	-16 (4.7)	
Slovak Republic[2]	537 (5.4)	520 (4.4)	-18 (4.2)	
Czech Republic †	520 (3.0)	502 (2.4)	-18 (2.8)	
Italy	540 (3.4)	522 (3.9)	-18 (3.3)	
Indonesia	442 (3.9)	423 (3.5)	-19 (3.0)	
Spain	514 (4.2)	496 (4.8)	-19 (3.6)	
England ‡	529 (6.1)	509 (6.1)	-20 (8.5)	
Russian Federation	517 (4.3)	496 (3.8)	-21 (3.4)	
Sweden	549 (3.4)	527 (4.2)	-21 (4.5)	
Ireland	545 (4.8)	523 (6.0)	-22 (6.2)	
Korea, Republic of[1]	577 (2.4)	555 (2.3)	-22 (3.0)	
Norway †	527 (3.7)	504 (4.5)	-23 (4.4)	
Mexico	463 (3.2)	439 (3.1)	-24 (2.9)	
Dominican Republic	392 (2.8)	367 (2.7)	-25 (2.7)	
Bulgaria	479 (5.2)	454 (6.1)	-26 (5.3)	
Chinese Taipei	573 (2.7)	546 (2.7)	-26 (2.5)	
Finland	590 (2.9)	562 (3.5)	-28 (4.3)	
Paraguay[1]	438 (4.1)	408 (3.9)	-29 (4.6)	
Slovenia	531 (2.6)	501 (3.9)	-30 (4.0)	
Latvia	497 (3.7)	466 (5.0)	-30 (3.7)	
New Zealand †	532 (5.9)	501 (6.4)	-31 (7.5)	
Greece	492 (4.8)	460 (5.1)	-32 (4.5)	
Poland	553 (4.5)	520 (5.5)	-33 (4.3)	
Estonia	542 (4.8)	509 (4.9)	-33 (3.9)	
Malta	507 (7.7)	473 (3.6)	-34 (8.2)	
Lithuania	523 (2.9)	488 (3.4)	-35 (3.0)	
Cyprus	475 (2.7)	435 (3.2)	-40 (3.7)	
Thailand †	474 (3.9)	426 (4.5)	-48 (4.5)	
ICCS average	**511 (0.7)**	**489 (0.7)**	**-22 (0.8)**	

Countries not meeting sample requirements

Hong Kong SAR	564 (6.5)	543 (8.3)	-21 (9.8)	
Netherlands	497 (6.6)	490 (10.4)	-7 (7.9)	

Gender Difference scale markers: (-100) (-50) 0 50 100. Labels: Females Score Higher / Males Score Higher.

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

† Met guidelines for sampling participation rates only after replacement schools were included.

‡ Nearly satisfied guidelines for sample participation only after replacement schools were included.

[1] Country surveyed the same cohort of students but at the beginning of the next school year.

[2] National Desired Population does not cover all of International Desired Population.

■ Gender difference statistically significant at 0.05 level
□ Gender difference not statistically significant

Figure 5.12 Example SAS Program to Perform Student-Level Analysis with Achievement Scores (EXAMPLE2.SAS)

```
LIBNAME ICCS2009 "<datpath>" ;
%INCLUDE "<macpath>JACKPV.SAS" ;
DATA ISGALLC2 ;
    SET ICCS2009.ISGALLC2;
    WHERE NMISS (SGENDER) = 0 ;
PROC FORMAT LIBRARY = WORK ;
    VALUE COUNTRY
        < list ICCS 2009 country formats > ;
    VALUE SEX
        0 = 'BOY'
        1 = 'GIRL' ;
%JACKPV (TOTWGTS, JKZONES, JKREPS, 75, IDCNTRY SGENDER, 5, ISGALLC2) ;
PROC PRINT DATA = FINAL NOOBS ;
    VAR IDCNTRY SGENDER N TOTWGTS MNPV MNPV_SE PCT PCT_SE ;
    FORMAT IDCNTRY COUNTRY. SGENDER SEX. N 6.0 TOTWGTS 10.0
        MNPV MNPV_SE PCT PCT_SE 6.2 ;
```

Figure 5.13 displays each country's results on two lines, one for each value of the variable SGENDER. The countries are identified in the first column and the second column describes the category being reported, SGENDER. The third column reports the number of valid cases and the fourth the sum of weights of the sampled students. The next two columns report the estimated mean civic knowledge and its standard error, followed by the percentage of students in each category and its standard error. From the first two lines, the mean civic knowledge score in Austria for boys is 496.47 (standard error of 4.45) and 512.60 (standard error of 4.74) for girls. An estimated 50.07% of students in Austria are boys, and 49.93% are girls.

Figure 5.13 Output for Example Student-Level Analysis with Civic Knowledge Scores (Example 2)

IDCNTRY	SGENDER	N	TOTWGTS	MNPV	MNPV_SE	PCT	PCT_SE
AUT	BOY	1553	41734	496.47	4.45	50.07	1.41
AUT	GIRL	1637	41624	512.60	4.74	49.93	1.41
BGR	BOY	1590	30431	453.51	5.94	48.21	1.65
BGR	GIRL	1642	32687	479.30	5.17	51.79	1.65
CHL	BOY	2510	126397	476.23	4.04	49.17	1.45
CHL	GIRL	2651	130659	489.83	4.21	50.83	1.45
TWN	BOY	2670	155929	546.12	2.84	51.58	0.57
TWN	GIRL	2474	146348	572.55	2.66	48.42	0.57

5.8   ICCS 2009 Analyses with Teacher-Level Variables

The teachers in the ICCS 2009 International Database constitute representative samples of target grade teachers in participating countries. The next sample analysis will use teacher questionnaire data, focusing on teachers' confidence in teaching civic and citizenship education, more specifically the percentages of teachers who are confident or very confident in teaching human rights. The first column of Table 6.18 of the *ICCS 2009 International Report* (Schulz et al., 2010b) presents the results of such an analysis. Figure 5.14 shows the same information. The macro JACKGEN will estimate the percentages needed. Note that only teachers who reported teaching civic related subjects answered this question.

As in the previous analyses, users must 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). Since the focus is on a teacher-level variable, users will need to use the teacher questionnaire data files, which will yield the variable that contains information on the target grade teachers' confidence in teaching human rights (IT2G28A), the variable that identifies the country (IDCNTRY), and the teacher identification variable (IDTEACH). Users will also need the jackknife replication variables (JKZONET and JKREPT) and the teacher weighting variable (TOTWGTT).

Figure 5.14 Sample Teacher-Level Analysis Taken from the ICCS 2009 International Report (Table 6.18)

Table 6.18: Teachers' confidence in teaching civic and citizenship education

| Country | Percentages of Teachers Who Are Confident or Very Confident in Teaching: | | | | | | |
	Human rights	Different cultural and ethnic groups	Voting and elections	The economy and business	Rights and responsibilities at work	The global community and international organizations	The environment
Bulgaria	89 (2.6)	90 (2.7) △	81 (3.3)	47 (4.5) ▼	90 (2.8)	80 (4.7)	89 (2.1)
Chile	94 (2.3)	92 (2.2) △	89 (3.1)	67 (4.1)	93 (2.4) △	86 (4.3) ▲	89 (3.1)
Chinese Taipei	92 (1.7)	90 (1.7) △	97 (1.3) ▲	78 (3.4) ▲	94 (1.8) △	81 (2.8) △	89 (2.2)
Colombia	98 (1.5) △	86 (3.3)	91 (2.8)	54 (3.8)	96 (0.9) △	69 (4.0)	95 (1.6) △
Cyprus	95 (2.7)	86 (4.2)	78 (5.2)	38 (5.9) ▼	84 (4.7)	73 (5.2)	92 (3.3)
Czech Republic †	96 (1.4) △	80 (3.0)	90 (1.9) △	62 (3.6)	87 (2.5)	80 (3.1)	90 (1.7)
Dominican Republic	93 (2.8)	88 (3.5)	85 (4.1)	62 (5.8)	90 (3.3)	64 (5.5) ▼	92 (3.1)
Finland	83 (1.8) ▽	73 (2.3) ▼	65 (1.9) ▼	50 (2.4) ▼	74 (2.1) ▼	53 (2.4) ▼	87 (1.4) ▽
Indonesia	96 (2.0) △	87 (2.6)	89 (2.6)	78 (3.4) ▲	91 (2.9)	80 (3.6)	95 (2.1)
Ireland ‡	94 (1.8)	78 (3.0)	86 (2.4)	69 (3.2) △	92 (1.4) △	88 (2.0) ▲	96 (1.2) △
Italy	98 (0.5) △	94 (0.8) ▲	87 (1.3)	39 (2.2) ▼	82 (1.9) ▽	86 (1.6) ▲	92 (1.2)
Korea, Republic of	67 (3.8) ▼	58 (3.4) ▼	75 (2.5) ▼	54 (4.0)	80 (2.3) ▽	52 (3.5) ▼	83 (2.1) ▽
Latvia	94 (1.9)	74 (3.2) ▼	83 (3.5)	65 (4.3)	86 (3.4)	64 (4.2) ▼	89 (3.2)
Liechtenstein	85 (7.5)	82 (7.4)	84 (7.5)	66 (9.6)	47 (9.6) ▼	77 (8.8)	82 (7.7)
Lithuania	89 (2.4)	88 (3.0)	82 (3.5)	57 (5.1)	81 (3.2)	91 (2.1) ▲	93 (1.9)
Malta	87 (3.2)	85 (2.9)	73 (3.9) ▼	40 (4.3) ▼	89 (3.0)	63 (4.6) ▼	95 (2.4)
Mexico	95 (1.9)	79 (3.9)	86 (3.5)	59 (4.4)	98 (1.1) ▲	62 (4.9) ▼	97 (1.7) △
Paraguay	97 (1.6) △	91 (3.1) △	96 (1.6) △	67 (5.1)	98 (1.5) ▲	81 (3.4)	100 (0.4) △
Poland	100 (0.0) △	89 (3.4)	97 (1.2) ▲	84 (3.7) ▲	87 (2.9)	90 (3.2) ▲	91 (2.4)
Russian Federation	98 (0.8) △	78 (2.5) ▽	95 (1.2) △	72 (3.1) ▲	93 (1.9) △	79 (2.3)	95 (1.5) △
Slovak Republic[1]	97 (1.1) △	76 (2.9) ▽	85 (2.7)	68 (4.0)	82 (3.0)	68 (3.8)	94 (2.0)
Slovenia	91 (1.8)	83 (1.5)	77 (1.7) ▽	32 (2.5) ▼	81 (1.9) ▽	63 (2.2) ▼	91 (1.2)
Spain	98 (1.3) △	94 (1.9) ▲	90 (2.7)	55 (4.3)	90 (2.3)	88 (2.7) ▲	91 (2.2)
Sweden †	99 (0.7) △	90 (1.8) △	97 (1.0) ▲	80 (2.9) ▲	85 (2.5)	93 (1.7) ▲	86 (2.2) ▽
Thailand †	88 (3.8)	84 (3.3)	95 (2.6) △	68 (4.3)	95 (2.6) △	67 (4.0) ▽	98 (1.1) △
ICCS average	93 (0.5)	84 (0.6)	86 (0.6)	60 (0.9)	87 (0.6)	75 (0.8)	92 (0.5)

Countries not meeting sampling requirements

Austria	94 (1.7)	78 (4.3)	96 (2.0)	55 (4.3)	65 (4.7)	79 (4.2)	73 (4.2)
Belgium (Flemish)	61 (2.7)	72 (2.8)	55 (2.2)	33 (2.3)	38 (2.3)	54 (2.7)	77 (2.4)
Denmark	93 (1.6)	86 (2.0)	83 (2.6)	54 (3.7)	72 (2.5)	74 (3.3)	76 (3.0)
England	83 (2.2)	80 (2.2)	73 (2.7)	51 (2.8)	72 (2.4)	70 (2.5)	87 (2.0)
Hong Kong SAR	63 (2.8)	66 (3.5)	67 (2.9)	46 (3.3)	78 (2.6)	56 (3.2)	79 (2.2)
New Zealand	96 (1.3)	97 (1.3)	91 (2.2)	57 (3.6)	89 (2.4)	87 (2.6)	94 (1.8)
Norway	96 (1.0)	85 (3.8)	94 (1.8)	71 (7.7)	83 (7.1)	84 (6.9)	95 (1.3)
Switzerland	85 (3.0)	73 (4.5)	91 (2.8)	59 (3.8)	50 (3.6)	72 (5.0)	85 (3.4)

National percentage

▲ More than 10 percentage points above ICCS average   △ Significantly above ICCS average
▽ Significantly below ICCS averag   ▼ More than 10 percentage points below ICCS average

Notes:
( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.
† Met guidelines for sampling participation rates only after replacement schools were included.
‡ Nearly satisfied guidelines for sample participation only after replacement schools were included.
[1] National Desired Population does not cover all of International Desired Population.

This analysis will use data for all available countries, making use of the teacher questionnaire data file, ITGALLC2. This file can be created with the JOIN macro.

The SAS program that executes this third example is presented in Figure 5.15 and is part of the database as EXAMPLE3.SAS. Figure 5.16 displays the results obtained from this program, 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 teachers who have non-missing data in the variable of interest, IT2G28A. A second step consists of combining response categories 1 and 2 and response categories 3 and 4 of the variable IT2G28A in order to match the results presented in Figure 5.14, where teachers are categorized into two groups: teachers who report being very confident or quite confident in teaching human rights and teachers who report being not very confident or not confident at all.

In general, to perform analyses using the teacher questionnaire data files, users should do the following:

1) Identify the variables of interest in the teacher questionnaire data files and note any specific national adaptations to the variables.

2) Retrieve the relevant variables from the teacher questionnaire data files, including analysis variables, classification variables, identification variables (IDCNTRY, IDTEACH), sampling (JKZONET and JKREPT) and weighting (TOTWGTT) 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 arguments and parameters.

5) Specify the location of the data files (<datpath>) and the macros (<macpath>).

6) Print the results file.

Figure 5.15 Sample SAS Program to Analyze Teacher Variables (EXAMPLE3.SAS)

```
LIBNAME ICCS2009 "C:\ICCS2009\Data\SAS_Data\" ;
%INCLUDE "C:\ICCS2009\JACKGEN.SAS" ;
DATA ITGALLC2 ;
    SET ICCS2009.ITGALLC2 ;
    IF NMISS(IT2G28A) = 0;
    SELECT (IT2G28A) ;
        WHEN (1,2) NEW28A = 1 ;  * Confident in teaching human rights ;
        WHEN (3,4) NEW28A = 2 ;  * Not Confident in teaching human
rights ;
        OTHERWISE  NEW28A = . ;  * All other responses ;
      END ;
PROC FORMAT LIBRARY = WORK ;
    VALUE COUNTRY
        < list ICCS 2009 country formats > ;
    VALUE NEW28A
        1 = 'Confident'
        2 = 'Not conf.' ;
%JACKGEN (TOTWGTT, JKZONET, JKREPT, 75, IDCNTRY NEW28A, IDCNTRY,
ITGALLC2) ;
PROC PRINT DATA = FINAL NOOBS ;
    VAR IDCNTRY NEW28A N TOTWGTT PCT PCT_SE ;
    FORMAT IDCNTRY COUNTRY. NEW28A NEW28A. N 6.0 TOTWGTT 10.0
        PCT PCT_SE 6.2 ;
```

In Figure 5.16, each country's results are shown on four lines, one for each value of the recoded IT2G28A variable. The results are presented in much the same manner as in previous examples, where the countries are identified in the first column and the second column describes the category of IT2G28A being reported. Looking at the third and fourth line of the output, 88.60% of the target grade teachers in Bulgaria that teach civic related subjects are confident in teaching human rights, 11.40% of the target grade teachers in Bulgaria that teach civic related subjects are not confident in teaching human rights. The standard error for both estimates is 2.60%.

Figure 5.16 Output for Example Teacher Variable Analysis (Example 3)

```
IDCNTRY     NEW28A          N       TOTWGTT     PCT      PCT_SE

   AUT      Confident      115       4087      94.07      1.71
   AUT      Not Conf.        9        258       5.93      1.71
   BGR      Confident      174       2270      88.60      2.60
   BGR      Not Conf.       30        292      11.40      2.60
   CHL      Confident      212       5962      94.04      2.33
   CHL      Not Conf.       10        378       5.96      2.33
   TWN      Confident      379       1980      92.48      1.66
   TWN      Not Conf.       31        161       7.52      1.66
```

5.9   ICCS 2009 Analyses with School-Level Variables

Because ICCS 2009 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 elements in their own right. Therefore, analyzing school data should be done by linking the students to their schools. An example of an analysis using school questionnaire data will compute the percentages of students in schools where principals report on participation of target grade classes in human rights projects. Variable IC2G06B will serve for this purpose. Figure 6.2 of the ICCS 2009 International Report (Schulz et al., 2010b) displays the results of this analysis, as does Figure 5.17 here.

Figure 5.17 Sample School-Level Analysis Taken from the ICCS 2009 International Report (Table 6.2)

Table 6.2: Principals' reports on participation of target-grade classes in community activities (in national percentages of students)

Percentages of Students Reported To Have Been Involved in …

Country	activities related to the environment and geared to the local area	human rights projects	activities related to underprivileged people or groups	cultural activities (e.g., theater, music, cinema)	multicultural and intercultural initiatives within the <local community>	campaigns to raise people's awareness, such as <World AIDS Day, World No Tobacco Day>	activities related to improving facilities for the <local community>	sports events
Austria	32 (4.2) ▶	27 (4.3)	33 (4.6)	87 (3.2) ◀	18 (3.6) ▶	65 (4.3)	11 (3.0) ▶	84 (3.5)
Belgium (Flemish)†	63 (4.1) ◀	45 (4.8) ◀	68 (4.7) ◀	95 (1.5) ◀	33 (4.8)	73 (3.5) ◀	12 (2.5) ▶	88 (2.6) △
Bulgaria	46 (4.6)	8 (2.6) ◀	24 (3.5) ▶	75 (3.7) ▶	36 (4.8)	76 (3.4) ◀	37 (4.2) △	85 (3.1)
Chile	40 (3.8) ▶	15 (2.8) ▶	35 (3.7)	57 (3.7) ▶	31 (3.5)	40 (4.1) ▶	9 (1.9) ▶	74 (3.5) ▽
Chinese Taipei	34 (4.1) ▶	24 (3.9) ▶	31 (4.1)	53 (4.1) ▶	30 (4.1)	53 (4.8)	35 (4.3)	75 (3.6) ▽
Colombia	57 (4.0) ▶	40 (3.3) ▶	16 (2.7) ▶	55 (3.4) ▶	36 (3.4)	41 (3.3) ▶	22 (3.2)	76 (3.3) ▽
Cyprus	21 (0.2) ▶	19 (0.2)	11 (0.1) ▶	41 (0.3) ◀	26 (0.2) ▽	19 (0.2) ▶	13 (0.2) ▶	46 (0.3) ▶
Czech Republic †	74 (4.1) ◀	42 (5.0)	34 (4.7)	98 (1.0) ◀	51 (4.8) ◀	77 (4.1) ◀	29 (4.3)	87 (2.9)
Denmark†	22 (3.7) ▶	24 (3.8) ▶	25 (3.8) ▶	80 (3.1) ▶	18 (3.6) ◀	18 (3.5) ▶	26 (3.8)	74 (3.9) ▽
Dominican Republic	66 (6.7) ◀	38 (5.3)	41 (4.7)	53 (6.2) ▶	52 (6.3) ◀	74 (4.3) ◀	30 (4.1)	77 (3.9)
England ‡	49 (5.3)	47 (5.1)	70 (3.9) ◀	89 (3.3) ◀	40 (5.5)	66 (4.7)	24 (4.6)	96 (2.2) ◀
Estonia	76 (3.8) ◀	23 (3.7)	15 (2.9) ▶	99 (1.1) ◀	40 (3.9)	78 (3.5) ◀	56 (4.7) ◀	99 (0.9) ◀
Finland	39 (3.3) ▶	15 (3.2)	48 (4.2) ◀	82 (2.9) ▶	28 (3.7)	88 (2.6) ◀	32 (3.9)	86 (2.5)
Greece	25 (3.5) ▶	10 (2.8) ▶	13 (3.4) ▶	41 (4.1) ◀	11 (2.8) ▶	22 (3.4) ▶	6 (2.1) ▶	50 (4.9)
Guatemala[1]	59 (4.6)	40 (4.8)	30 (4.1)	69 (4.3) ◀	46 (4.8) ◀	44 (4.7) ▶	37 (4.7) △	90 (2.1) △
Indonesia	67 (4.2) ◀	18 (3.1) ▶	47 (4.5) ◀	34 (4.1) ▶	17 (3.4) ▶	19 (3.6) ▶	34 (4.0)	79 (3.9)
Ireland	40 (3.7) ▶	39 (4.6)	33 (4.3)	52 (4.4) ▶	18 (3.4) ▶	21 (3.5) ▶	10 (2.7) ▶	79 (3.9)
Italy	60 (4.3) △	66 (3.6) ◀	44 (3.8) △	82 (3.1) △	47 (3.7) ◀	56 (3.8)	24 (3.6)	81 (2.8)
Korea, Republic of[1]	32 (3.6) ▶	22 (3.4) ▶	32 (3.9)	28 (3.8) ▶	16 (3.0) ▶	42 (3.8) ▶	24 (3.4) ▶	38 (4.3) ▶
Latvia	43 (4.2)	30 (4.1)	31 (4.9)	96 (1.8) ◀	47 (4.4) ◀	53 (4.8)	65 (4.2) ◀	98 (1.2) ◀
Liechtenstein	32 (0.4) ▶	59 (0.4) ◀	59 (0.4) ◀	87 (0.3) ◀	0 (0.0)	75 (0.4) ◀	13 (0.3) ▶	87 (0.4) △
Lithuania	55 (4.3)	28 (4.2)	20 (3.3) ▶	76 (3.4) ▶	51 (3.5) ◀	67 (4.1) △	63 (3.9) ◀	97 (1.5) ◀
Luxembourg	23 (1.4) ▶	32 (2.2)	39 (2.3)	63 (2.2) ▶	35 (2.2)	74 (1.9) ◀	0 (0.0) ▶	75 (2.3) ▽
Malta	42 (0.9) ▽	38 (0.9) △	48 (0.9) ◀	65 (1.0) ▽	19 (0.6) ▶	39 (0.9) ▶	13 (0.4) ▶	94 (0.1) ◀
Mexico	66 (3.4) ◀	47 (3.7) ◀	32 (3.0)	54 (3.4) ◀	40 (3.6)	60 (3.2)	32 (3.0) ◀	67 (3.5) ▶
New Zealand†	46 (5.1)	40 (5.2)	54 (5.7) ◀	81 (4.2) ◀	51 (4.5) ◀	62 (4.5)	17 (3.9) ▽	97 (0.6) ◀
Norway †	38 (4.8) ▶	31 (4.1)	37 (4.5)	90 (2.8) ◀	21 (3.6) ▶	57 (5.2)	21 (4.1)	80 (3.3) ◀

Table 6.2. Principals' reports on participation of target-grade classes in community activities (in national percentages of students) (contd.)

Country	Percentages of Students Reported To Have Been Involved in …							
	activities related to the environment, geared to the local area	human rights projects	activities related to underprivileged people or groups	cultural activities (e.g., theater, music, cinema)	multicultural and intercultural initiatives within the \<local community>	campaigns to raise people's awareness, such as \<World AIDS Day, World No Tobacco Day>	activities related to improving facilities for the \<local community>	sports events
Paraguay[1]	82 (3.0) ▲	49 (5.0) ▲	50 (4.2) ▲	84 (3.0) ▲	59 (4.3) ▲	61 (4.2)	53 (4.4) ▲	94 (2.0) ▲
Poland	63 (4.1) ▲	51 (4.3) ▲	50 (4.1) ▲	88 (2.7) ▲	33 (4.3)	92 (2.1) ▲	22 (3.6)	92 (2.2) ▲
Russian Federation	80 (3.1) ▲	36 (3.0)	49 (2.8) ▲	91 (1.9) ▲	42 (3.2) △	81 (2.8) ▲	32 (3.6)	95 (1.2) ▲
Slovak Republic[2]	74 (3.6) ▲	50 (4.5) ▲	34 (4.1)	93 (2.2) ▲	53 (4.5) ▲	63 (4.2) ▲	36 (4.3) △	94 (1.9) ▲
Slovenia	68 (3.4) ▲	49 (4.6) ▲	39 (4.4)	90 (2.2) ▲	46 (3.7) ▲	85 (2.8) ▲	31 (3.4)	89 (2.7) △
Spain	63 (4.3) ▲	52 (4.2) ▲	44 (3.9)	86 (2.3) ▲	34 (4.1)	72 (4.0) ▲	14 (2.9) ▼	76 (3.9)
Sweden	35 (4.1) ▼	47 (4.1) ▲	34 (4.1)	92 (2.2) ▲	27 (3.3) ▽	30 (4.2) ▼	20 (3.5) ▽	81 (3.3)
Switzerland †	38 (6.1) ▼	15 (3.2) ▼	12 (3.2) ▼	85 (3.0) ▲	13 (2.5) ▼	52 (4.8)	13 (2.8) ▼	94 (2.1) ▲
Thailand†	66 (4.3) ▲	45 (4.1) ▲	46 (4.7)	71 (3.5)	59 (4.1) ▲	82 (3.4) ▲	69 (4.4) ▲	92 (2.2) ▲
ICCS average	**50 (0.7)**	**35 (0.6)**	**37 (0.6)**	**74 (0.5)**	**34 (0.6)**	**58 (0.6)**	**27 (0.6)**	**82 (0.5)**

Countries not meeting sampling requirements

Country								
Hong Kong SAR	38 (6.5)	14 (5.1)	34 (6.5)	67 (6.4)	34 (5.5)	45 (7.4)	29 (6.2)	87 (4.9)
Netherlands	25 (9.4)	24 (7.2)	42 (8.8)	82 (7.7)	23 (9.3)	29 (10.3)	16 (5.2)	82 (5.1)

National percentage

▲ More than 10 percentage points above ICCS average   △ Significantly above ICCS average

▽ Significantly below ICCS average   ▼ More than 10 percentage points below ICCS average

Notes:

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

† Met guidelines for sampling participation rates only after replacement schools were included.

‡ Nearly satisfied guidelines for sample participation only after replacement schools were included.

[1] Country surveyed the same cohort of students but at the beginning of the next school year.

[2] National Desired Population does not cover all of International Desired Population.

Since the analysis uses a school-level variable, the school questionnaire data files and the student questionnaire data files will identify the variables. Within the school questionnaire data files are the variable that contains the principals' reports on the participation of target grade classes in human rights projects (IC2G06B) and the identification variables IDCNTRY and IDSCHOOL that will allow linking of the school data to the student data.

Next, users must retrieve the variables of interest from the student questionnaire data files. The country and school identification variables (IDCNTRY and IDSCHOOL) are necessary to merge the school data with the student data. The analysis also uses the jackknife replication variables (JKZONES and JKREPS), the student weighting variable (TOTWGTS), and the civic knowledge plausible values (PV1CIV through PV5CIV).

The analysis then merges the school data with the student data using the variables IDCNTRY and IDSCHOOL, and using the macro JACKPV to obtain the percentages of students and their mean civic knowledge scores within each category of the variable BCDGAS for each country. This analysis will use the data for all available countries, making use of an aggregated school file ICGALLC2 and an aggregated student file ISGALLC2. These aggregated files can be created with the JOIN macro.

The SAS program that implements this fourth example is presented in Figure 5.18 and is part of the database as EXAMPLE4.SAS. The results of this program are displayed in Figure 5.19, edited to show only the first four countries in alphabetical order. Note that one of the steps in this program is to select only those students who have non-missing data in the variable of interest IC2G06B.

In general, to perform analyses using the school questionnaire data files, analysts should do the following:

1) Identify the variables of interest in the school and student questionnaire data files and note any specific national adaptations to the variables.

2) Retrieve the relevant variables from the school questionnaire 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 questionnaire data files, including plausible values for civic knowledge, classification variables, identification variables (IDCNTRY and IDSCHOOL), sampling (JKZONES and JKREPS) and weighting (TOTWGTS) variables, and any other variables used in the selection of cases.

4) Merge the school questionnaire data files with the student questionnaire 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 (<datpath>) and the macros (<macpath>).

8) Print the results file.

Figure 5.18 Example SAS Program for School Variable Analysis (EXAMPLE4.SAS)

```
LIBNAME ICCS2009 "<datpath>" ;
%INCLUDE "<macpath>JACKPV.SAS" ;
PROC SORT DATA = ICCS2009.ICGALLC2 OUT = ICGALLC2;
    BY IDCNTRY IDSCHOOL ;
PROC SORT DATA = ICCS2009.ISGALLC2 OUT = ISGALLC2;
    BY IDCNTRY IDSTUD ;
DATA MERGED ;
    MERGE ICGALLC2 (IN = INICG)
    ISGALLC2 (IN = INISG) ;
    BY IDCNTRY IDSCHOOL ;
    IF INICG AND INISG ;
DATA MERGED ;
    SET MERGED ;
    IF NMISS (IC2G06B) = 0 ;
    SELECT (IC2G06B) ;
        WHEN (1,2) NEW06B = 1 ;  * STUDENTS HAVE BEEN INVOLVED IN ;
        WHEN (3,4,5) NEW06B = 2 ;  * STUDENTS HAVE NOT BEEN INVOLVED IN ;
        OTHERWISE  NEW06B = . ;
    END ;
PROC FORMAT LIBRARY = WORK ;
    VALUE COUNTRY
        < list country formats >
    VALUE NEW06B
        1 = 'Students have been involved in'
        2 = 'Students have not been involved in' ;
%JACKPV (TOTWGTS, JKZONES, JKREPS, 75, IDCNTRY NEW06B, 5, MERGED) ;
PROC PRINT DATA = FINAL NOOBS ;
    VAR IDCNTRY NEW06B N TOTWGTS MNPV MNPV_SE PCT PCT_SE ;
    FORMAT IDCNTRY COUNTRY. NEW06B 1.0 N 6.0 TOTWGTS 10.0
        MNPV MNPV_SE PCT PCT_SE 6.2 ;
```

In Figure 5.19, each country's results are presented on three lines, one for each value of the IC2G06B variable. The results are presented in much the same manner as in previous examples, where the countries are identified in the first column and the second column describes the category of IC2G06B being reported. From the first two lines, 27.08% (standard error of 4.30) of target grade students in Austria attend schools where principals report participation of target-grade classes in human rights projects, and 72.92% (standard error of 4.30) of target grade students attend schools where principals' report no participation of target grade classes in human rights projects. The estimated mean civic knowledge of target grade students in Austria in schools where principals' report participation of target grade classes in human rights projects is 510.74 (standard error of 8.71), whereas the estimated mean civic knowledge of target grade students in schools where principals' report no participation of target grade classes in human rights projects is 498.84 (standard error of 5.91).

Figure 5.19 Output for Example School Variable Analysis (Example 4)

IDCNTRY	NEW06B	N	TOTWGTS	MNPV	MNPV_SE	PCT	PCT_SE
AUT	1	752	19104	510.74	8.71	27.08	4.30
AUT	2	1962	51442	498.84	5.91	72.92	4.30
BGR	1	209	4825	481.74	26.03	7.81	2.56
BGR	2	2957	56932	463.93	5.11	92.19	2.56
CHL	1	750	37351	488.52	11.85	14.87	2.75
CHL	2	4294	213845	480.39	4.15	85.13	2.75
TWN	1	1204	70830	555.61	5.42	23.82	3.92
TWN	2	3854	226531	559.43	3.27	76.18	3.92

Organizations and Individuals Responsible for ICCS 2009

Introduction

The International Civic and Citizenship Education Study (ICCS) 2009 was a collaborative effort involving hundreds of individuals around the world. This appendix recognizes the individuals and organizations for their contributions. Given that the work on ICCS 2009 has spanned approximately five 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, ICCS 2009 is deeply indebted to the students, teachers, and school principals who contributed their time and effort to the study.

The international study center and its partner institutions

The international study center for ICCS 2009 is located at the Australian Council for Educational Research (ACER). Center staff at ACER were responsible for the design and implementation of the study in close cooperation with the center's partner institutions NFER (National Foundation for Educational Research, Slough, United Kingdom), LPS (Laboratorio di Pedagogia Sperimentale at the Roma Tre University, Rome, Italy), the IEA Data Processing and Research Center (DPC), and the IEA Secretariat.

Staff at ACER

John Ainley, *project coordinator*
Wolfram Schulz, *research director*
Julian Fraillon, *coordinator of test development*
Tim Friedman, *project researcher*
Naoko Tabata, *project researcher*
Maurice Walker, *project researcher*
Eva Van De Gaer, *project researcher*
Anna-Kristin Albers, *project researcher*
Corrie Kirchhoff, *project researcher*
Renee Chow, *data analyst*
Louise Wenn, *data analyst*

Staff at NFER

David Kerr, *associate research director*
Joana Lopes, *project researcher*
Linda Sturman, *project researcher*
Jo Morrison, *data analyst*

Staff at LPS

Bruno Losito, *associate research director*
Gabriella Agrusti, *project researcher*
Elisa Caponera, *project researcher*
Paola Mirti, *project researcher*

International Association for the Evaluation of Educational Achievement (IEA)

IEA provides overall support with respect to coordinating ICCS 2009. The IEA Secretariat in Amsterdam, The Netherlands, is responsible for membership, translation verification, and quality control monitoring. The IEA Data Processing and Research Center (DPC) in Hamburg, Germany, is mainly responsible for sampling procedures and the processing of ICCS 2009 data.

Staff at the IEA Secretariat

Hans Wagemaker, *executive director*

Barbara Malak, *manager membership relations*

Dr Paulina Koršňáková, *senior administrative officer*

Jur Hartenberg, *financial manager*

Staff at the IEA Data Processing and Research Center (DPC)

Heiko Sibberns, *co-director*

Dirk Hastedt, *co-director*

Falk Brese, *ICCS coordinator*

Michael Jung, *researcher*

Olaf Zuehlke, *researcher (sampling)*

Sabine Meinck, *researcher (sampling)*

Eugenio Gonzalez, *consultant to the Latin American regional module*

ICCS project advisory committee (PAC)

PAC has, from the beginning of the project, advised the international study center and its partner institutions during regular meetings.

PAC members

John Ainley (*chair*), ACER, Australia

Barbara Malak, IEA Secretariat

Heiko Sibberns, IEA Technical Expert Group

John Annette, University of London, United Kingdom

Leonor Cariola, Ministry of Education, Chile

Henk Dekker, University of Leiden, The Netherlands

Bryony Hoskins, Center for Research on Lifelong Learning, European Commission

Rosario Jaramillo F., Ministry of Education, Colombia (2006–2008)

Margarita Peña B., Colombian Institute for the Evaluation of Education (2008–2010)

Judith Torney-Purta, University of Maryland, United States

Lee Wing-On, Hong Kong Institute of Education, Hong Kong SAR

Christian Monseur, University of Liège, Belgium

Other project consultants

Aletta Grisay, University of Liège, Belgium

Isabel Menezes, Porto University, Portugal

Barbara Fratczak-Rudnicka, Warszaw University, Poland

ICCS sampling referee

Jean Dumais from Statistics Canada in Ottawa was the sampling referee for ICCS. 2009 He provided invaluable advice on all sampling-related aspects of the study.

National research coordinators (NRCs)

The national research coordinators (NRCs) played a crucial role in the development of the project. They provided policy- and content-oriented advice on the development of the instruments and were responsible for the implementation of ICCS 2009 in participating countries.

Austria

Günther Ogris

SORA Institute for Social Research and Analysis, Ogris & Hofinger GmbH

Belgium (Flemish)

Saskia de Groof

Center of Sociology, Research Group TOR, Free University of Brussels (Vrije Universiteit Brussel)

Bulgaria

Svetla Petrova

Center for Control and Assessment of Quality in Education, Ministry of Education and Science, Bulgaria

Chile

Marcela Ortiz

Unidad de Curriculum y Evaluación, Ministerio de Educación

Chinese Taipei

Meihui Liu

Department of Education, Taiwan Normal University

Colombia

Margarita Peña

Instituto Colombiano para la Evaluación de la Educación (ICFES)

Cyprus

Mary Koutselini

Department of Education, University of Cyprus

Czech Republic

Petr Soukup

Institute for Information on Education

Denmark

Jens Bruun

Department of Educational Anthropology, The Danish University of Education

Dominican Republic

Ancell Scheker

Director of Evaluation in the Ministry of Education

England

Julie Nelson

National Foundation for Educational Research

Estonia

Anu Toots

Tallinn University

Finland

Pekka Kupari

Finnish Institute for Educational Research, University of Jyväskylä

Greece

Georgia Polydorides

Department of Early Childhood Education

Guatemala

Luisa Muller Durán

Dirección General de Evaluación e Investigación Educativa (DIGEDUCA)

Hong Kong SAR

Wing-On Lee

Hong Kong Institute of Education

Indonesia
Diah Haryanti
Balitbang Diknas, Depdiknas

Ireland
Jude Cosgrove
Educational Research Centre, St Patrick's College

Italy
Genny Terrinoni
INVALSI

Republic of Korea
Tae-Jun Kim
Korean Educational Development Institute (KEDI)

Latvia
Andris Kangro
Faculty of Education and Psychology, University of Latvia

Liechtenstein
Horst Biedermann
Universität Freiburg, Pädagogisches Institut

Lithuania
Zivile Urbiene
National Examination Center

Luxembourg
Joseph Britz
Ministère de l'Éducation Nationale

Romain Martin
University of Luxembourg

Malta
Raymond Camilleri
Department of Planning and Development, Education Division

Mexico
María Concepción Medina
Mexican Ministry of Education

Netherlands
M. P. C. van der Werf
GION, University of Groningen

New Zealand
Kate Lang
Sharon Cox
Comparative Education Research Unit, Ministry of Education

Norway
Rolf Mikkelsen
University of Oslo

Paraguay
Mirna Vera
Dirección General de Planificación

Poland
Krzysztof Kosela
Institute of Sociology, University of Warsaw

Russia

Peter Pologevets
Institution for Education Reforms of the State University Higher School of Economics

Slovak Republic

Ervin Stava
Department for International Measurements, National Institute for Certified Educational Measurements NUCEM

Slovenia

Marjan Simenc
University of Ljubljana

Spain

Rosario Sánchez
Instituto de Evaluación, Ministerio de Educación y Ciencia

Sweden

Marika Sanne
Fredrik Lind
The Swedish National Agency for Education (Skolverket)

Switzerland

Fritz Oser
Universität Freiburg, Pädagogisches Institut

Thailand

Siriporn Boonyananta
The Office of the Education Council, Ministry of Education

Somwung Pitiyanuwa
The Office for National Education Standards and Quality Assessment

References

Kerr, D., Sturman, L., Schulz, W., & Bethan, B. (2010). *ICCS 2009 European Report. Civic knowledge, attitudes and engagement among lower secondary school students in twenty-four European countries*. Amsterdam, the Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

IEA (2011). *International Database Analyzer* (version 2.0). Hamburg, Germany: International Association for the Evaluation of Educational Achievement Data Processing and Research Center.

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

Schulz, W., Ainley, J. & Fraillon, J. (Eds.) (forthcoming). *ICCS 2009 Technical Report*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Schulz, W., Fraillon, J., Ainley, J., Losito, B., & Kerr, D. (2008). *International Civic and Citizenship Education Study: Assessment framework*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Schulz, W., Fraillon, J., Ainley, J., Kerr, D. & Losito, B. (2010a). *Initial Findings from the IEA International Civic and Citizenship Study*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Schulz, W., Ainley, J., Fraillon, J., Kerr, D. & Losito, B. (2010b). *ICCS 2009 International Report. Civic knowledge, attitudes and engagement among lower secondary school students in thirty-eight countries*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Schulz, W., Ainley, J., Friedman, T. & Lietz, P. (2011). *ICCS 2009 Latin American Report: Civic knowledge and attitudes among lower secondary students in six Latin American countries*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement (IEA).

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel Analysis*. London: Sage Publications.

SPSS Inc. (2010). *SPSS for Windows* (version 18.0). Chicago, IL: SPSS Inc.