The following briefly describes the PISA 2003 data sets and how to access the data using SAS or SPSS. It is highly recommended that the user refer first to the *PISA 2003 Data Analysis Manual* for detailed information on these analysis issues. That report is available for downloading at (http://www.pisa.oecd.org/document/18/0,2340,en_32252351_32236173_35016146_1_1_1_1,00.html). The international data may be downloaded from this same site.

PISA 2003 Data Sets

The PISA database contains three data sets: the school questionnaire file (usa_schl.dat), the student questionnaire file (usa_stud.dat) and the assessment items file (usa_assesm.dat). The data are in ASCII format. Associated control programs are included on this website to assist the user in reading the data to produce SAS data sets and SPSS system files. Since the data are hierarchical (students are clustered within schools) each student record contains identification variables that enable the user to merge the school data with the student data. The school data may be merged to the student data using the variable SCHOOLID.

The contents of the PISA 2003 files are described below:

usa_stud.dat. This file contains: student and school identification variables; student responses to the questionnaire; derived index scores; mathematics, science and reading performance scores; student sampling weights; and, (Fay) replicate weights. There are 5,456 cases in the student file.

usa_schl.dat. This file contains: the school identification variable; school responses to the school questionnaire; derived school index scores; and, the school sampling weight. There are 274 cases in the school file.

usa_assesm.dat. This file contains data on student responses to each item of the assessment. There 5,456 cases in the student assessment file.

Accessing the U.S. Data through the SAS/SPSS Control Files

SAS and SPSS control files are available on this site. These files contain code to create SAS/SPSS datasets ready for analysis. Users will need to make some minor edits to the code prior to running it. The most basic edits required are changes to the paths pointing the raw data files and the

path where the output dataset will be saved. The naming convention for the programs is to take the root name of the dataset and replace the "dat" with "sas" or "sps" respectively. For example, in order to create a SAS dataset from the PISA 2003 student background file, usa_stud.dat, the user must modify and run the program PISA2003 student.sas.