NCES 2016 036REV

U.S. DEPARTMENT OF EDUCATION

U.S. Program for the International Assessment of Adult Competencies (PIAAC) 2012/2014: Main Study and National Supplement Technical Report

NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

This page intentionally blank.

U.S. Program for the International Assessment of Adult Competencies (PIAAC) 2012/2014: Main Study and National Supplement Technical Report

SEPTEMBER 2016

Jacquie Hogan
Nina Thornton
Lillian Diaz-Hoffmann
Leyla Mohadjer
Tom Krenzke
Jianzhu Li
Wendy VanDeKerckhove
Westat

Kentaro Yamamoto
Lale Khorramdel
ETS

Holly Xie
Stephen Provasnik
Project Officers
National Center for Education Statistics

NCES 2016-036REV
U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

U.S. Department of Education
Dr. John B. King, Jr.
Secretary of Education

Institute of Education Sciences
Ruth Neild
Deputy Director for Policy and Research
Delegated Duties of the Director

National Center for Education Statistics
Peggy G. Carr
Acting Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high-priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high-quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public. Unless specifically noted, all information contained herein is in the public domain.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

NCES, IES, U.S. Department of Education
Potomac Center Plaza (PCP)
550 12th Street, SW, 4th Floor
Washington, DC 20202

September 2016

The NCES Home Page address is https://nces.ed.gov.
The NCES Publications and Products address is http://nces.ed.gov/pubsearch.

This publication is only available online. To download, view, and print the report as a PDF file, go to the NCES Publications and Products address shown above.

This report was prepared for the National Center for Education Statistics under Contract Numbers ED-04-CO-0059/0030 and ED-IES-12-C-0072 with Westat. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

Suggested Citation
Hogan, J., Thornton, N., Diaz-Hoffmann, L., Mohadjer, L., Krenzke, T., Li, J., VanDeKerckhove, W., Yamamoto, K., and Khorramdel, L. (2016). *U.S. Program for the International Assessment of Adult Competencies (PIAAC) 2012/2014: Main Study and National Supplement Technical Report* (NCES 2016-036REV). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Available from http://nces.ed.gov/pubsearch.

Content Contact
Holly Xie
(202) 245-8481
Holly.Xie@ed.gov

CONTENTS

CONTENTS—Continued

CONTENTS—Continued

CONTENTS—Continued

CONTENTS—Continued

List of Tables

CONTENTS—Continued

List of Figures

This page intentionally blank.

1. INTRODUCTION

1.1 Introduction to PIAAC

The Program for the International Assessment of Adult Competencies (PIAAC) is the most comprehensive international survey of adult skills ever undertaken. The survey examines a range of basic skills in the information age and assesses these adult skills consistently across participating countries. In 2011–12, 24 countries participated in the survey (Round 1), and 9 additional countries participated in 2014–15 (Round 2)[1] of the survey. The assessment focuses on the key cognitive and workplace skills necessary for individuals to participate successfully in the economy and society of the 21st century. This multicycle study is a collaboration between the governments of participating countries, the Organization for Economic Cooperation and Development (OECD), and a consortium of various international organizations, referred to as the PIAAC Consortium. This consortium is led by the Educational Testing Service (ETS) and includes the German Institute for International Educational Research (DIPF), the German Social Sciences Infrastructure Services' Centre for Survey Research and Methodology (GESIS-ZUMA), the University of Maastricht's Research Centre for Education and the Labour Market (ROA), the U.S. research company Westat, the International Association for the Evaluation of Educational Achievement (IEA), and the Belgian firm cApStAn.

The study assesses the following key adult skills for the information age: basic reading skills, reading literacy, numeracy, and problem solving in "technology-rich environments" (the OECD term for "on or with a computer"). PIAAC also measures the ability of individuals to use computer and web applications to find, gather, and use information and to communicate with others. The study uses a "Job Requirements Approach" to ask employed adults about the types and levels of a number of specific skills used in the workplace. These include not only the use of reading and numeracy skills on the job but also physical skills (e.g., stamina, manual dexterity), people skills (e.g., public speaking, negotiating, working in a team), and information technology skills (e.g., using spreadsheets, writing computer code). It asks about the requirements of the person's main job in terms of the intensity and frequency of the use of such skills. PIAAC also breaks new ground by being the first to use laptop computers to administer an international assessment of this kind, although some individuals were given a paper-and-pencil version of the assessment if they were unable or refused to take the assessment on the computer.

An important element of the value of PIAAC is its collaborative and international nature. In the United States, the U.S. Department of Education's National Center for Education Statistics (NCES) is collaborating with the U.S. Department of Labor (DOL) on PIAAC. Staff from NCES and DOL are corepresentatives of the United States in PIAAC's international governing body, and NCES has consulted extensively with DOL, particularly on development of the job skills section of the Background Questionnaire (BQ). Internationally, PIAAC has been developed collaboratively by participating countries' representatives from ministries and departments of education and labor and by OECD staff through an extensive series of international meetings and work groups. These international meetings and work groups, assisted by expert panels, researchers, and the PIAAC Consortium's support staff, have developed frameworks used to develop the assessment and BQ and the common standards and procedures for collecting and reporting data, and also guided the development of a common, international "virtual machine" software that administers the assessment uniformly on laptops. All PIAAC countries must follow the common standards and procedures and use the virtual machine software when conducting the

[1] Round 1 countries: Australia, Austria, Canada, Cyprus, Czech Republic, Denmark, England/N. Ireland (U.K.), Estonia, Finland, Flanders (Belgium), France, Germany, Ireland, Italy, Japan, Korea, Netherlands, Norway, Poland, Russian Federation, Slovak Republic, Spain, Sweden, United States. Round 2 countries: Chile, Greece, Indonesia, Israel, Lithuania, New Zealand, Slovenia, Singapore, Turkey.

survey and assessment. As a result, PIAAC can provide a reliable and comparable measure of adult skills in the adult population (ages 16–65) of participating countries.

1.2 PIAAC in the United States

NCES contracted with Westat to work with NCES and the PIAAC Consortium to conduct the study in the United States. Westat's key tasks included (a) development of a Screener to enumerate and select study participants, (b) adaptation of the international BQ and assessment for the United States, (c) instrument translation (as necessary), (d) sample design and selection, (e) data collection, (f) scoring, and (g) the production of reports detailing the results of the PIAAC Field Test, PIAAC 2012 Main Study, and PIAAC 2014 National Supplement. In the United States, PIAAC was fielded under the user-friendly name International Survey of Adult Skills (ISAS); however, this technical report uses the official international name of the study: U.S. PIAAC.

1.2.1 PIAAC Field Test and Main Study (2012)

The U.S. Field Test data were collected between September and November 2010, with 1,510 adults (ages 16–65) interviewed and assessed in 22 primary sampling units (PSUs) across the country. The U.S. Main Study data collection (adults, ages 16–65 years) took place between August 25, 2011 and April 3, 2012. Five-thousand ten (5,010) cases were completed in 80 PSUs across the United States.

1.2.2 PIAAC National Supplement (2014)

The U.S. PIAAC National Supplement 2014 repeated the administration of PIAAC to an additional sample of U.S. adults in order to enhance the U.S. PIAAC Main Study sample. The National Supplement included a sample of participants from different households in the same 80 Main Study PSUs as well as a sample of inmates selected from 98 prisons across the United States. Although 100 prisons were sampled to participate in the survey, two prisons declined participation.

The National Supplement household sample increased the sample size of two key subgroups of interest, unemployed adults (ages 16–65) and young adults (ages 16–34), and added to the sample one new subgroup of older adults (ages 66–74). All procedures and instruments used during the Main Study were employed during the household data collection for the National Supplement. The National Supplement household data collection took place between August 26, 2013, and May 5, 2014. Three-thousand six-hundred sixty (3,660) additional cases were completed.

The National Supplement prison sample was composed of inmates (ages 16–74) incarcerated in state, federal, or private prisons. To specifically capture the experiences and support the analytic needs of this subgroup, the BQ was modified to include questions regarding the respondents' activities in prisons and their access to academic programs and other learning activities. The same assessment used in the household sample was used for participants in the prison sample. The National Supplement prison data collection took place between February 10, 2014, and June 13, 2014. One-thousand three-hundred nineteen (1,319) cases were completed in the 98 prisons.

This technical report replaces the *Program for the International Assessment of Adult Competencies (PIAAC) 2012: U.S. Main Study Technical Report* (NCES 2014-047) as it combines information from both, the U.S. PIAAC 2012 Main Study and U.S. PIAAC 2014 National Supplement data collection efforts. This report includes detailed information on the Main Study and National Supplement sample design, survey instruments used for data collection, the data collection process and quality of the data, weighting, scaling, and data analysis.

2. MEETING PIAAC CONSORTIUM REQUIREMENTS

The PIAAC Consortium oversees all PIAAC activities on behalf of the OECD and provides technical support to all participating countries regarding all aspects of PIAAC. Each country is responsible for conducting PIAAC in compliance with the *Technical Standards and Guidelines* (PIAAC Consortium 2014) provided by the Consortium to ensure that the survey design and implementation yields high-quality and internationally comparable data. The standards are generally based on agreed-upon policies or best practices to be followed when conducting the study, and all participating countries must follow them in order to have their data included in the OECD reports and data products.

To ensure that standards were met by all participating countries, the Consortium set up a comprehensive quality control process to monitor all aspects of the study. Details on the PIAAC quality control process as it pertained to the U.S. Main Study and National Supplement household samples are covered in this chapter. This chapter also documents all of the major PIAAC tasks that required interaction with, approval from, and/or deliverables to the Consortium.

For the Main Study and National Supplement household data collection, Westat complied with all of the *Technical Standards and Guidelines* or received permission for deviations. Deviations were documented on the National Survey Design and Planning Report, described below in 2.1, and agreed upon with NCES.

Minor modifications to the quality control procedures used in the household data collection were made for the National Supplement prison data collection; these are noted in the remaining sections of this report where relevant.

2.1 National Survey Design and Planning Report

The United States was required to document the proposed methods and procedures for adhering to the PIAAC *Technical Standards and Guidelines* in the National Survey Design and Planning Report. This report consists of a series of questions relating to each of the following sections of the *Technical Standards and Guidelines*:

- ethics (chapter 2);
- survey planning (chapter 3);
- sample design and selection (chapter 4);
- survey instruments (chapter 5);
- translation and adaptations (chapter 6);
- information technology standards (chapter 7);
- field management (chapter 8);
- training (chapter 9);
- data collection (chapter 10);

- data processing (chapter 11);
- data file creation (chapter 12);
- steps to ensure data confidentiality and security (chapter 13);
- weighting/estimation (chapter 14); and
- quality assurance and quality control (chapter 15).

The U.S. Main Study National Survey Design and Planning Report was submitted to NCES for review and successfully submitted to the Consortium on February 1, 2011. Each participating country was required to specify the rationale for any deviations from the technical standards for Consortium review and approval. The only deviations to the PIAAC standards for the United States concerned the translation methodology, and the Consortium approved these deviations.[2] No concerns were raised by the Consortium regarding any other aspect of the design and procedures proposed by Westat.

The proposed methods and procedures described in the Main Study National Survey Design and Planning Report were followed for the National Supplement except where revisions were required for the target sample groups of focus; for example, extending the upper age range for respondents to 74 years.

2.2 Sample Selection and Monitoring

Several quality control sampling checks were required by the Consortium to ensure adherence to the PIAAC *Technical Standards and Guidelines*. Completion of the quality control checks for the Field Test served as practice for the Main Study, where they were important in producing high-quality data that are comparable between countries. The United States completed the following (required) sampling documentation and submitted the documentation to the Consortium:

- sampling plans, consisting of a series of questions on the sample design for the Field Test and Main Study, which served as part of the National Survey Design and Planning Report;
- quality control sample selection forms, completed after each stage of sample selection, describing the sample selection process and the characteristics of the sampled units;
- quality control sample monitoring forms, due periodically throughout data collection and used to monitor sample yields and response rates by subgroup; and
- a final sample monitoring form due after data collection created using final edited and cleaned data.

[2] The recommended procedure for developing the national versions of the Background Questionnaire (BQ) was double translation by two independent translators followed by reconciliation. Instead, Westat followed its established procedures for translation of the BQ into Spanish. A member of the Westat Translation Unit translated the BQ into Spanish. This initial translation was then thoroughly reviewed by another member of the unit. Subsequently, all comments were reconciled by the head of the Translation Unit, and the translation underwent a final review by the project translation coordinator and NCES. More detail is provided in section 2.4.

In addition, the United States attended two sampling workshops[3] held by the Consortium to prepare countries for sampling-related activities for the Field Test and Main Study, including sample design, sample selection, confidentiality, weighting, and nonresponse bias analysis.

The same sample selection and monitoring procedures were followed for the National Supplement except where revisions were required for the target sample groups of focus; for example, oversampling of unemployed adults.

2.3 Background Questionnaire Adaptations

The Consortium developed the PIAAC international master version of the Background Questionnaire (BQ), which was the basis for the U.S. national BQ. The Main Study international master was updated by the PIAAC Consortium based on the Consortium's analysis of the PIAAC Field Test data. As a result of the Consortium review, countries were expected to implement a number of changes based on the Consortium changes to the international master version to eliminate problematic items and to reduce the overall length of the BQ. In some cases the changes to the international master also precipitated changes to the U.S. national BQ.

Working in conjunction with NCES, its support staff at the American Institutes for Research (AIR), and the Consortium, Westat developed a set of recommended changes to the U.S. national BQ. These changes were implemented over a 9-month period, including at the Consortium June 2010 meeting in Frankfurt, Germany and the December 2010 meeting in Princeton, New Jersey. At these meetings, recommended changes to the BQ and assessment items were implemented using the Consortium-provided tools and the Item Management Portal. The impact of the international changes to the U.S. adaptations was also reviewed, and revisions were made as needed to these items as well.

For the National Supplement household sample, the U.S. national BQ used during the Main Study was modified slightly. These modifications extend the upper age range for respondents to 74 years.

For the National Supplement prison sample, the U.S. national BQ was adapted based on recommendations of a prison expert panel and AIR support staff and with the approval of NCES. The content of the questionnaire was also informed by the BQ used for the National Assessment of Adult Literacy (NAAL) 2003 prison sample. Prior to finalizing the content of the prison BQ (both English and Spanish versions), cognitive testing was conducted at a local detention center to ensure that respondents could understand the questions, follow the question logic, and navigate the specified questionnaire skip patterns. The incorporated changes were thoroughly reviewed and tested before the final instrument was approved by NCES.

The English and Spanish versions of the household sample and prison sample BQs can be found at https://nces.ed.gov/surveys/piaac/questionnaire.asp.

[3] The first sampling workshop was in March 2009 in Barcelona, Spain. The second sampling workshop was in December 2010 in Princeton, New Jersey.

2.4 Translation of Instruments

Westat's translation unit translated the PIAAC questionnaires into Spanish employing a four-step process intended to ensure that the Spanish-language instruments would be universally comprehensible to the various Spanish-speaking groups in the United States, equivalent to the original English-language instruments in terms of intent and meaning, and written using correct Spanish grammar and syntax. These four steps were as follows:

1. initial translation by a native Spanish speaker with more than 20 years of experience in English-Spanish translation, and experience translating other educational studies (e.g., NAAL);

2. editorial review of the initial translation by a native Spanish speaker with 10 years of experience in English-Spanish translation, as well as work experience in the U.S. educational community;

3. editorial review by a native English speaker with 20 years of experience in English-Spanish translation, with a focus on equivalency of the English and Spanish question items; and

4. final editorial review by a native Spanish-speaking research analyst with 10 years of experience in designing and testing questionnaires and in developing, administering, and monitoring research studies.

Subsequently, a draft of the Spanish-language questionnaires was submitted to NCES for final review. In addition to the Spanish language translation, Westat "translated" PIAAC's international English-language questionnaires into U.S. English and submitted those edits to NCES for final review.

Once the translated text for the questionnaires was approved, Westat incorporated the approved translated text into the XLIFF[4] files that were provided by the Consortium. New text was added and/or replaced using the supplied Open Language Tool (OLT) software to implement the translation, in both English and Spanish versions. As no "text replace" function existed in OLT, the major task for creating the U.S. English version—finding and replacing intricate phrasings and replacing recurring phrases—had to be performed manually for each instance where a change to PIAAC's international English required "translation" (i.e., adaptation to U.S. English) using this OLT interface. For the Spanish version of the questionnaire, consistent changes to text had to be made; therefore, Westat chose to handle the XLIFF files as text and apply more capable text editors for both quantitative and qualitative improvement and easier implementation. However, going outside the OLT software triggered some subsequent handling problems of this XLIFF (e.g., related to Spanish special characters). Once the initial BQ implementation containing Consortium information technology (IT) adjustments to American English language versions was received, Westat conformed to Consortium requirements by using the supplied OLT software for the remaining light editing.

Another challenge for preparing translated versions of the questionnaires was the volume of Spanish text that needed to be re-entered into XLIFF, a tool not intended for large document handling. This required files to be artificially split to fit size limits of the tool.

[4] XLIFF is an XML-based file (XML Localization Interchange File Format) format that enables translators to concentrate on the text to be translated.

The materials prepared for the Main Study were used for the National Supplement household survey, so no additional translation was needed for it. However, the same translation procedures were followed for the National Supplement prison sample BQ and the prison sample brochure and study information flier.

2.5 Interviewer Training

To ensure that interviewers are trained in a consistent fashion across participating countries, the Consortium provided guidelines and training materials to be used by each participating country but allowed countries to make adaptations as necessary. In addition, the Consortium recommended a minimum of 33 hours of interviewer training for all interviewers plus an additional 4 hours of general interviewing techniques training for trainees new to interviewing. Westat followed all interviewer training recommendations, used all training materials provided by the Consortium, and made the adaptations necessary to meet U.S. needs and practices.[5]

To monitor the quality of interviewer training, the Consortium required each country to fill out Interviewer Training Forms within a month of completing training, which the United States did.

For the National Supplement household sample, interviewer training was also conducted in compliance with Consortium standards and guidelines. However, Interviewer Training Forms were not submitted to the Consortium.

For the National Supplement prison sample, interviewers were used who had already successfully completed the household sample interviewer training program. Since these interviewers were previously trained on PIAAC and the procedures to administer the assessment, the prison sample training program focused on training interviewers to conduct the inmate sample selection procedures within selected prison facilities and other procedural changes specific to data collection within the prisons. All changes to the prison sample BQ were also reviewed.

2.6 Data Collection Quality Control Monitoring Process

The quality control monitoring of data collection in the PIAAC Main Study required each participating country to submit a number of forms and to participate in a number of telephone conference calls with the Consortium. The required forms and their respective due dates are listed below:[6]

- one data collection form per month for each of the 2 months leading up to the start of the Main Study data collection (July and August 2011);
- one data collection form for each of the 7 months of the official Main Study data collection (September 2011 through March 2012);

[5] Westat's approach to interviewer training relies on a progressive exposure of interviewers to the questionnaire and training by example. Important points are highlighted as training progresses. The training scripts provided by the Consortium had to be adapted to conform to this approach.

[6] Dates provided are adapted from standard deadlines set for all other countries. These adaptations were necessary due to the U.S. Census moratorium, which imposed a delay in (U.S. PIAAC Main Study) data collection. Due dates for all other countries occurred much earlier.

- one data collection form after data collection ended (April 2012);
- one data collection form after data submission (July 2012); and
- an interviewer debriefing report, which collected interviewers' feedback regarding the training and preparation for data collection and their field experiences (May 2012).

All Consortium-mandated contact and submission requirements for this activity were met for the Main Study.

The Consortium did not conduct quality control monitoring activities for the National Supplement, although activities similar to those monitored during the Main Study were conducted throughout the data collection period and were reported to NCES in monthly progress reports and project meetings.

2.7 Coding and Data Processing

To ensure that coding and data processing tasks were performed in a uniform way within and across participating countries in the Main Study, the Consortium provided training for both scorers and national data managers as follows:

- scoring training in Bologna, Spain, in January 2010, attended by Pearson, the contractor responsible for scoring and data entry in the United States;
- national data manager training in Frankfurt, Germany, in February 2010, attended by Westat; and
- Main Study training in Dublin, Ireland, in June 2011, attended by Westat.

The Consortium provided the Data Management Expert (DME) software to be used to import the TAO[7] interviews, data entry of the scoring process, editing and quality control, importing of coding results, and exporting of the final data files. Westat entered Consortium-approved national adaptations within the DME and followed the Consortium recommendations for the use of the DME.

Questions relating to occupation, education, language, and country of birth were identified for external coding by the Consortium, and coding schemes were provided. Westat followed Consortium guidelines for the coding process as well as for identifying and coding several national adaptation questions. A complete list of the coded variables and the coding schemes used is available in chapter 9.

The coding and data processing procedures used in the Main Study were followed for the National Supplement.

[7] TAO (Testing Assisté par Ordinateur = computer-based testing, supplied by Centre de Recherche Public Henri Tudor, based in Luxembourg) is a Consortium-supplied software made available to execute the BQ, computer-based assessment, and the automated interviewer guide and instructions to be followed when administering the paper-based assessment.

2.8 Data Delivery

As part of the DME, the Consortium provided "back-end data" processing software to standardize data delivery for all participating countries. Each country was to use this software to (1) aggregate country data into a common format, (2) combine assorted data products into one deliverable dataset, and (3) perform some basic quality control checks. This software fulfilled the major requirements of data cleaning and preparation that each country was responsible for performing. Data cleaning and preparation included loading the individual data files created at the conclusion of each interview into the DME, flagging possible data inconsistencies or errors needing review, generating error reports for key data items needing review, and providing the ability to correct data errors found during the review process.

Significant data delivery issues identified by the Consortium required individual review, comment, and possible corrections. This review included comparing the delivered data with the original data files to ensure that files were correctly loaded into the DME. Helpdesk reports and other edit logs were also reviewed to determine if other edits had been correctly applied to problematic cases. Listed below are data delivery issues that were identified for the United States:

- Background Questionnaire log data (BQ data type) were to be included as part of the international delivery, but due to U.S. confidentiality laws they were withheld from the international delivery. See section 9.8 for further description.

- Quality Control checks in the DME were expanded for the Main Study to improve the quality of the data file delivered and to identify issues early in the data processing cycle so that corrections could be made, if needed, prior to sending the files to the Consortium. The DME software included 45 data checks that countries were to perform. These checks included ensuring unique identifiers across data rows; ensuring no orphaned rows by identifiers; range checking where appropriate; ensuring that data rows existed for each identifier for the modules that were comprehensive; and identifying inconsistences between disposition codes and the various assessments. Of these 45 data checks, 22 applied to the paper assessment and reading components booklets. Eleven of the 22 checks were related to paper booklet disposition codes and could not be fully reconciled. Per Consortium direction to alter data as little as possible, the data were delivered without resolving these checks. The unreconciled checks were due to the limited disposition code frame for the core module (see section 9.7) and a set of false-positive checks that arose from the Consortium-implemented check logic that did not account for U.S. adaptations of the BQ. The United States' reporting of these items was accepted by the Consortium (see section 9.8).

- Countries were asked to compare the aggregate distribution of collected data on specific topics (educational attainment, employment status, industry, and occupation) against existing national measures of the same topics. For the United States, the national measure used to compare the aggregate distributions of PIAAC data was the U.S. Census Bureau's Current Population Survey (CPS). The Consortium *Technical Standards and Guidelines* did not specify specific ranges of acceptability; however, based on review and discussions with NCES, Westat determined that comparisons for education, employment, and industry were acceptable. A direct comparison of occupation could not be done because no cross-walk matching PIAAC codes to CPS codes exists. Therefore, Westat reported a statistical analysis of results rather than aggregated direct matching. The Consortium subsequently accepted this approach, noting that most U.S. variation trended with international variations.

- A small number of computer-based assessments were lost due to technical problems, and a few other cases suffered other technical problems and could not be delivered. In the Main Study, these technical problems were less than 1 percent of completed cases. Losses due to technical problems during the National Supplement were also less than 1 percent of completed cases, both for the household sample as well as the prison sample.

- The Consortium requested review of the small number (less than 0.5 percent) of age or gender discrepancies that occurred between screening and the completion of the BQ (Main Study). This is a common data issue when performing enumeration and confirmation in separate instruments, perhaps at different times or with different respondents.

The same data delivery procedures were followed for the National Supplement. For the National Supplement prison sample, additional data checks were designed by Westat and provided to the Consortium data processing contractor to ensure that the specified skip logic was appropriately applied during the data cleaning and preparation process.

2.9 Weighting and Variance Estimation

During the weighting period for the Main Study, the Consortium required each country to report on its weighting process using quality control monitoring forms. Quality checks were developed to review the weighting process and evaluate the potential for nonresponse-related bias in descriptive variables (such as the region of the country and the percentage of minority population). The quality checks were performed after each step in the weighting process. These included the following:

- reviewing the distribution of weights at each stage to identify any missing or extreme values;

- computing the weighted frequencies of important survey characteristics after each weighting adjustment to show how each adjustment affected the estimates for key survey variables—in addition, weighted frequencies were compared to reliable external totals, such as Current Population Survey (CPS) estimates of the population age 16 to 65 by race/ethnicity;

- reviewing a random listing of records for abnormalities;

- producing the mean, median, minimum, and maximum of weights and checking for each replicate weight after each weight adjustment; and

- after the final weights were produced, producing preliminary standard errors and design effects on survey variables as a check on the replicate weights.

Westat performed all required checks and submitted all the required forms within the expected time frame for the Main Study.

For the National Supplement, the same weighting and variance estimation procedures were followed; however, no forms were required by or submitted to the Consortium.

3. SAMPLE DESIGN

The PIAAC Consortium specified *Technical Standards and Guidelines* for all aspects of the sample design, including the identification of the target population, the creation of the sampling frame, and the sample size requirements and the sample selection methods. All countries were required to submit sample design plans detailing these aspects to the Consortium for approval several months before data collection. Also, countries were required to complete quality control sample selection forms, which collected sampling information for each stage of selection. These were designed to capture aggregated information necessary for verifying that the sample was representative of the target population and that sampling was conducted in an unbiased and randomized way. The sample design and selection for the U.S. Main Study were performed in accordance with these procedures. The National Supplement also followed the PIAAC *Technical Standards and Guidelines*, where applicable, although no quality control forms were due to the Consortium.

This chapter describes the sample design and selection for the U.S. Main Study and National Supplement. Section 3.1 describes the household sample, and section 3.2 covers the National Supplement prison component. Each section provides an overview of the design: the sample design, sample frames, and sample selection procedures. Each section also includes information on the sample size requirements and coverage issues, including the initial sample size given the assumed response rate and eligibility rate at each stage of data collection. Quality control methods employed during sample selection are provided along with procedures used to monitor the sample during data collection.

3.1 Household Sample

The U.S. PIAAC household sample consisted of 8,670 respondents from two administrations: the Main Study and National Supplement. The Main Study targeted noninstitutionalized adults between the ages of 16 and 65 (inclusive). This was accomplished through a four-stage area sample, consisting of 80 primary sampling units (PSUs), 901 segments, 9,468 dwelling units (DUs), and 6,100 sampled persons, resulting in 5,010 respondents to the survey. The National Supplement provided additional samples of adults for the unemployed (age 16 to 65), young adults (age 1 to 34), and an age 66 to 74 cohort. The sample was selected from the same 80 PSUs using a dual-frame approach for selecting DUs. One frame consisted of existing DU lists created from the same segments selected during the Main Study, and a second frame consisted of purchased postal addresses from 80 census tracts selected with a high concentration of unemployed from within the 80 PSUs. Between the two frames, there were 16,535 DUs selected, resulting in 3,660 respondents.

The sample design for the Main Study is described in section 3.1.1 and for the National Supplement in section 3.1.2. Sample sizes and a summary of quality control procedures are provided in sections 3.1.3 and 3.1.4, respectively. Composite weights were produced so that national estimates can be generated for the combined sample. The process for the combining the samples is found in chapter 8.

The target population for the U.S. PIAAC household sample included only persons living in households or group quarters; it excluded all other persons (such as persons living in shelters, the incarcerated, military personnel who live in barracks or bases, or persons who live in institutionalized group quarters, such as hospitals or nursing homes). The target population included full-time and part-time members of the military who did not reside in military barracks or military bases, adults in other noninstitutional collective DUs, such as workers' quarters or halfway homes, and adults who lived at school in student group quarters, such as a dormitory, fraternity, or sorority (refer to section 3.1.1.4 for more information

about selection procedures for college students in dormitories). Consistent with PIAAC international standards and guidelines, adults who were unable to complete the assessment because of a hearing impairment, blindness/visual impairment, or physical disability were in scope (that is, they were part of the target population); however because the assessment did not offer accommodations for physical disabilities, they were excluded from response rate computations. Adults were included regardless of citizenship, nationality, or language.

Persons temporarily in the country were eligible depending upon how long they had been in the country. The household respondent was asked in the Screener how many people lived in the dwelling and had no usual place of residence elsewhere. Those who thought of the household as their primary place of residence, or spent most of the year in the household even though they may have another residence, were listed as eligible household members. The list included persons who usually stay in the household but were temporarily away on business, vacation, in a hospital, or living at school.

3.1.1 Main Study Sample Design and Selection

To arrive at a minimum of 5,000 completed cases among noninstitutionalized persons age 16–65, a four-stage, stratified area probability sample was selected as follows:

- 80 PSUs consisting of counties or groups of contiguous counties;
- 901 secondary sampling units, or segments, consisting of 2000 Decennial Census blocks or block groups;
- 9,468 DUs; and
- 6,100 individuals within DUs resulting in 5,010 respondents[8] to the survey.

Random sampling methods were used, with known probabilities of selection at each sampling stage.

During the fourth stage of selection, a Screener interview was used to identify the eligible persons within selected DUs. A sampling algorithm was implemented within the computer-assisted personal interviewing (CAPI) system to select one or two sample persons among those identified to be eligible. Once selected, the Background Questionnaire (BQ) interview was completed. Upon completion of the BQ, the respondents were provided either the paper-and-pencil or computer-based assessment, based on whether they reported having any previous computer experience during the BQ interview or whether they refused the computer-based assessment as well as their performance on the computer technology (ICT) core instrument, conducted after the BQ.

Following the completion of the assessment, a monetary incentive of $50 was paid to each respondent. The incentive was also paid to those adults who attempted to complete an assessment but were legitimately not able to complete it because they lacked sufficient fluency in English or Spanish or had a physical or mental disability that precluded responding to the assessment. Respondents who refused to continue with the assessment were not compensated.

[8] There were 5,011 completed cases according to the definition PIAAC Technical Standard 4.3.3, and 5,010 receiving final sampling weights for analysis. Those receiving final sampling weights included BQ completes and those who lacked the literacy skills to complete the BQ. We use the sample size according to the final sampling weights through this technical report.

The PIAAC Main Study target population consisted of noninstitutionalized adults age 16 to 65 who resided in the United States at the time of interview, where age was determined during the Screener questionnaire. Although the PIAAC sample selection methods presented here give all eligible persons a known probability of selection, and although procedures were implemented to include any missed structures and hidden DUs, almost all surveys are subject to some amount of undercoverage. One known source of undercoverage in PIAAC resulted from the selection of one segment within a gated city.[9] Since field staff could not gain entry into the segment to construct a DU sampling frame, and would subsequently be prevented from contacting selected DUs, this segment was eliminated from the sample. The result is undercoverage of approximately 0.08 percent of the target population.[10]

To achieve the targeted number of completed assessments (5,000) for the PIAAC Main Study, assumptions were made regarding the rates of occupancy of the selected DUs, the eligibility of household members, and the level of cooperation of the selected individuals. Tables of actual eligibility rates and response rates, and also sample sizes are provided in section 3.1.3.

3.1.1.1 Primary Sampling Units

In the first stage of sampling, 80 PSUs were selected for the PIAAC study. This stage of selection involved the formation of PSUs leading to the creation of the PSU sampling frame. The selection process included the stratification of the PSUs as well as the selection of one PSU per stratum with probabilities proportionate to a measure of size (MOS).

3.1.1.1.1 Frame

The PSUs consisted of single counties or groups of contiguous counties. The Census Bureau provides a list of all counties in the United States with the most recent (2008, at the time of PSU design) resident population estimates by age group and sex. These estimates were adjusted to remove adults living in institutions or military quarters. Counties not meeting the minimum MOS criterion based on these population estimates for the targeted age group were combined with adjacent counties respecting metropolitan/micropolitan statistical area definitions, state boundaries, and the travel distance for data collectors until the minimum size criterion was met. Counties meeting the minimum size criterion served as PSUs. This resulted in five types of PSUs as follows:

- single counties;
- two or more counties within the same metropolitan Core Based Statistical Area (CBSA);[11]
- two or more counties within the same micropolitan CBSA;

[9] A gated city differs from a gated community in that the gated area encompasses an entire municipality rather than just a small neighborhood or development. Municipal authorities refused to allow the PIAAC listers access after repeated refusal conversion attempts by the field manager and the Westat project management team.

[10] Estimate obtained by taking the ratio of the segment population in households (adjusted to the 2010 Census) to the total population in households according to the 2010 Census.

[11] See *Federal Register*, Vol. 75, No. 123, June 28, 2010, for more information.

- two or more counties not part of a larger statistical area; and
- a combination of counties that are not part of a larger statistical area and those that are part of a micropolitan CBSA.

PSUs were formed based on a county-level input file and some PSU parameters. The main parameters were the minimum MOS, the maximum distance or area within a PSU, and the formation objective (i.e., to minimize distance). The formation algorithm started by designating each county as a resolved or an unresolved county. A resolved county is one that met all the requirements of becoming a PSU by itself, and an unresolved county is one that failed to meet one or more of the requirements. An unresolved county was merged with one or more of the contiguous unresolved counties, if available. Otherwise, an unresolved county was merged with a contiguous resolved county. If there was more than one choice for merging, the resulting PSU was chosen with the lowest end-to-end distance (in miles). If an unresolved county could not be merged with any of the contiguous counties because of other parameter restrictions, then that county was left "unresolved," which was a situation that was handled manually (outside of the automated process). Once the formation process was complete, a PSU-level file was created by aggregating the relevant county-level variables of the counties within a PSU.

As mentioned above, the objective for the PIAAC 2012 PSU formation process was to minimize the distance (i.e., maximum travel distance within a PSU), subject to the following constraints:

- The minimum population size (i.e., the minimum MOS) in a PSU was 15,000.
- The maximum distance between the two farthest corners of a PSU was 100 miles, generally. A distance of more than 100 miles was allowed in some special cases (i.e., some single counties exceed this criterion).
- A county within a metropolitan CBSA was not combined with counties outside that area, except in some special cases.
- A PSU was formed within a state boundary.

The result was a frame consisting of 1,949 PSUs, with characteristics as shown in table 3-1. Most PSUs consisted of one county.

Table 3-1. PIAAC characteristics of the PSU sample frame

Characteristic	Minimum	1st Quartile	Median	3rd Quartile	95th Percentile	Maximum
End-to-end distance (miles)	10	37	48	66	136	1,435
Area (square miles)	25	527	802	1,379	6,106	288,144
Number of counties per PSU	1	1	1	2	4	9
Estimated civilian noninstitutionalized population age 15–64	15,009	22,524	35,385	81,076	420,995	6,669,325

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012.

Each PSU on the frame was assigned an MOS equal to the estimate of the noninstitutionalized population age 15–64 within the PSU. This was derived from the Census Bureau population estimates available for the county noninstitutionalized resident population in that age group. While the MOS included those of

age 15 and excluded those of age 65, this was not expected to impact the sample since the population rates for single age groups should be fairly constant across counties.

3.1.1.1.2 Selection

Four PSUs with the largest MOS were selected with probability equal to one before stratification using a certainty cutoff determined from probability proportionate to size sampling. Such PSUs are referred to as self-representing. The nonself-representing PSUs on the frame were grouped into major strata. The major strata were based on state-level small area estimates (SAE) of the percentage of the population lacking Basic Prose Literacy Skills[12] and whether the PSU was part of a metropolitan area, as shown in table 3-2.

Table 3-2. PIAAC characteristics of the PSU major strata

Major strata	Metropolitan status	State small area estimate (SAE) of percentage of the population lacking Basic Prose Literacy Skills	State ranking on SAE of percentage lacking Basic Prose Literacy Skills	Number of nonself-representing PSUs in frame	Total measure of size in stratum	Number of minor strata
Overall				1,945	185,583,235	76
A	NonMeSA	6–7	1–10	207	5,497,387	2
B	NonMeSA	8–9	11–20	178	5,321,891	2
C	NonMeSA	10–12	21–30	160	4,829,189	2
D	NonMeSA	13–16	31–44	332	10,779,966	4
E	NonMeSA	17–23	45–51	166	5,086,448	2
F	MeSA, but not CSA	6–8	1–13	60	4,229,566	2
G	MeSA, but not CSA	9–12	15–30	96	10,261,359	4
H	MeSA, but not CSA	13	31–35	42	4,966,365	2
I	MeSA, but not CSA	14–16	36–44	51	4,610,072	2
J	MeSA, but not CSA	17–20	45–49	70	14,355,553	6
K	MeSA, but not CSA	22–23	50–51	18	5,239,463	2
L	MeSAs in CSAs	6–7	1–10	67	9,047,361	4
M	MeSAs in CSAs	8	11–14	63	9,416,494	4
N	MeSAs in CSAs	9–15	16–39	253	44,045,481	18
O	MeSAs in CSAs	16	41–44	30	4,166,443	2
P	MeSAs in CSAs	17	45–46	54	10,011,661	4
Q	MeSAs in CSAs	19–20	47–49	44	9,542,202	4
R	MeSAs in CSAs	22–23	50–51	54	24,176,335	10

NOTE: "NonMeSA" means that all counties in the PSU are not part of a Metropolitan Statistical Area; "MeSA, but not Combined Statistical Area (CSA)" means that all counties in the PSU are part of a Metropolitan Statistical Area, but are not part of a CBSA; "MeSAs in CSAs" means that all counties are part of a Metropolitan Statistical Area that is part of a CBSA. See *Federal Register, Vol. 75, No. 123, June 28, 2010,* for more information.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012.

[12] U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, 2003 National Assessment of Adult Literacy (NAAL).

Once major strata were identified, substrata (minor strata) were formed via a nested stratification process, as discussed in Krenzke and Haung (2009), using auxiliary variables related to the expected proficiency scores. An extensive search was conducted for county-level variables for the National Assessment of Adult Literacy (NAAL) SAE task (Mohadjer et al. 2009), and the key predictors of literacy proficiency were related to race/ethnicity, poverty, English-speaking ability, and educational attainment. An evaluation of the 2003 NAAL PSU strata was conducted (Krenzke and Haung 2009), and using the key predictors in the SAE process as stratifiers helped reduce the between-PSU variance. Based on these results, the indirect estimate of the percentage lacking Basic Prose Literacy Skills was included as the evaluation variable, while forming explicit strata using the most recent demographic estimates (the SAE predictors) from the Census Bureau. Table 3-3 shows the variables used to form the minor strata within each major stratum. Strata were close to equal MOS to reduce the variation in interviewer workload.

Once the strata were formed, one nonself-representing PSU was selected per stratum with probability proportionate to its MOS. The resulting 80 self-representing and nonself-representing PSUs were diverse in terms of literacy skills, geographic region of the country, and urbanicity of the PSU, as well as diverse in educational attainment, spoken-English ability, race/ethnicity, and poverty status.

The probability of selecting PSU *i* in stratum *h* is

$$P_{hi} = \frac{m_h \times MOS_{hi}}{\sum_{i \in h} MOS_{hi}},$$

where

m_h = Number of PSUs to be sampled in stratum *h*; and

MOS_{hi} = MOS for PSU *i* in stratum *h*.*

Table 3-3. PIAAC variables used to form the PSU minor strata: 2012

Major strata	Variables used to form minor strata
A	Percentage of the population age 25 and older with some college or more
B	Percentage of the population age 25 and older with some college or more
C	Percentage of the population age 25 and older with some college or more
D	Percentage of the population below 150 percent of poverty; percentage of the population age 15–64 that is non-Hispanic Black
E	Percentage of the population age 15–64 that is Hispanic
F	Percentage of the population age 25 and older with some college or more
G	Percentage of the population age 25 and older with some college or more; percentage of the population age 15–64 that is non-Hispanic Black
H	Percentage of the population age 15–64 that is Hispanic
I	Percentage of the population age 25 and older with some college or more
J	Percentage of the population age 15–64 that is Hispanic; percentage of the population age 5 and up that are English speakers; percentage of the population age 15–64 that is non-Hispanic Black; percentage of the population age 25 and older with some college or more
K	Percentage of the population age 25 and older with some college or more
L	Percentage of the population age 25 and older with some college or more; percentage of the population below 150 percent of poverty

See notes at end of table.

Table 3-3. PIAAC variables used to form the PSU minor strata: 2012—Continued

Major strata	Variables used to form minor strata
N	*PSUs in the Northeast and Midwest regions with the estimated percentage of the population lacking Basic Prose Literacy Skills (BPLS) is less than 9 percent:* Percentage of the population age 25+ with some college or more; percentage of the population age 5 and up that are English speakers; percentage of the population age 15–64 that is non-Hispanic Black
N	*PSUs in the Northeast and Midwest regions with the estimated percentage of the population lacking BPLS is less than 9 percent:* Percentage of the population age 25+ with some college or more; percentage of the population age 5 and up that are English speakers
N	*PSUs in the South and West regions with the estimated percentage of the population lacking BPLS is less than 10.3 percent:* Percentage of the population age 25+ with some college or more; percentage of the population age 5 and up that are English speakers; percentage of the population age 15–64 that is non-Hispanic Black
N	*PSUs in the South and West regions with the estimated percentage of the population lacking BPLS is greater than or equal to 10.3 percent:* Percentage of the population age 5 and up that are English speakers; percentage of the population age 15–64 that is non-Hispanic Black
O	Percentage of the population age 5 and up that are English speakers
P	Percentage of the population age 15–64 that is White or other (non-Hispanic, non-Black); percentage of the population age 25+ with some college or more
Q	Percentage of the population age 25+ with some college or more; percentage of the population 5 and up that are English speakers
R	*PSUs with the estimated percentage of the population lacking BPLS is less than 15 percent:* Percentage of the population age 25+ with some college or more; percentage of the population below 150 percent of poverty
R	*PSUs with the estimated percentage of the population lacking BPLS is greater than or equal to 15 percent:* Percentage of the population age 15–64 that is non-Hispanic Black; percentage of the population age 25+ with some college or more; percentage of the population below 150 percent of poverty

NOTE: Percentage of the population age 25+ with some college or more and the percentage of the population age 5 and up that are English speakers are current year estimates obtained from Claritas, 2009. All other data are from the Census Bureau's 2008 population estimates (as of July 1, 2008). Where noted, population estimates were available for the age range 15–64 for counties, instead of the target population 16–65 age range.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012.

3.1.1.2 Secondary Sampling Units

For the second stage of sampling, a probability proportionate to size sample of 901 segments from within the 80 sampled PSUs was selected.[13] The segments were formed using 2000 Decennial Census block definitions and they were selected using a preassigned sampling rate that lead to a self-weighting sample of DUs as described in section 3.1.1.3.

[13] During the design phase of the survey, it was determined that 900 segments would be sufficient to reach the goal of 5,000 completed cases based on Census Bureau population estimates for a point between the 2000 and 2010 Decennial Censuses. The segment sampling rate was calculated as $900/\sum MOS$, where the MOS was based on data from the 2000 census and the result was the selection of one more segment than targeted as a result of the incongruent estimates between sample design and selection.

3.1.1.2.1 Frame

Due to the timing of the data collection and the listing effort, the frame of segments was created within the selected PSUs using the Census 2000 Summary File 1 (SF 1) block[14] data, which were the most recent decennial census data available at the time.

Housing or demographic data from the last Decennial Census become outdated late in the decade. Data for local areas with considerable growth or demographic shifts since the census-taking are affected the most, and intercensal estimates are not available at the required level for segment formation. Using outdated data such as the MOS could result in considerable differences between the number of ultimate sampling units and expected counts, and increased variation in the number of sampling units across segments.

Varying approaches have been proposed to deal with this issue, including those that employ a different sampling approach and use alternative sources of data for the MOS other than (or in combination with) the usual detailed Decennial Census data. Two methods used to update MOS late in the decade include building permit sampling (Bell et al. 1999) and a two-phase segment sampling approach (Montaquila et al. 1999; Mohadjer, Montaquila, and Sherris 2002; Dohrmann, Harding, and Li 2008).

Given the restrictive timeline for creating the segment sampling frame, a different approach was used for PIAAC. Rather than employing different field methods to improve the segment sampling frame, counts of United States Postal Service (USPS) addresses within geographic areas were used to update the segment MOS. USPS counts were incorporated only in areas in which they appeared to more accurately reflect the number of DUs in 2010 (as compared to the 2000 Decennial Census data).[15] For these segments, the number of DUs in the segment were adjusted upward so that the segment MOS for sampling would be as accurate as possible (Dohrmann, Li, and Mohadjer 2011).

The segments consisted of at least 60 DUs in area blocks or combinations of two or more nearby blocks. Within each PSU, the block data from the SF 1 files were sorted by tract, block group, and block number before creating the segments. Blocks with no DUs and no population were included so that all areas, presumably some of which contained DUs constructed after the 2000 Decennial Census, were involved in the formation process. The result was a segment frame consisting of 218,000 segments.

Once segments were formed, the number of DUs in each segment was compared with counts of residential addresses from the USPS.[16] Additionally, segment-level predicted values from a model (Montaquila, Hsu, and Brick 2011) developed to determine areas for which the USPS lists are expected to have good coverage were also calculated.[17, 18]

---

[14] Blocks are very fine partitions of the United States, formed using visible semipermanent features such as roads, railroad tracks, mountain ridges, bodies of water, and power lines. The only invisible boundaries used are county, state, and national boundaries. Minor civil division boundaries and property lines are ignored. A block group is a small group of contiguous blocks. A tract is a collection of contiguous block groups all within the same county.

[15] USPS address counts tend to undercover areas without residential mail delivery, other rural areas, and areas with large proportions of their population living in noninstitutionalized group quarters.

[16] Addresses were obtained from the USPS's November 2010 Computerized Delivery Sequence File.

[17] For details about the match rate model, refer to Montaquila, Hsu, and. Brick (2011).

[18] The model uses mostly segment-level characteristics (including the ratio of USPS DU counts to census DU counts, urbanicity, mobility, occupancy rate, etc.), which are available from the ACS and decennial census, to predict how well the USPS addresses can cover a segment. For example, a predicted value of 0.85 means that the USPS addresses can cover 85 percent of the actual DUs in a segment. These predicted values, referred to here as "match rate values," help to determine areas for which the USPS DU counts may be the most accurate.

After careful examination of the estimated growth according to those USPS counts and the match rate values from the model, it was decided that adjustments would be made only in the following circumstances:

- in counties for which:
  - The county-level count of USPS residential addresses exceeded the number of DUs according to the 2000 Decennial Census, and
  - The number of DUs according to the 2005–2009 ACS also exceeded the number of DUs according to the 2000 Decennial Census;
- in those segments with segment-level growth of over 20 percent indicated by USPS DU counts compared to the 2000 Decennial Census; and
- in those segments with match rate values larger than 0.85.

Dormitory units are not included in the census count of DUs and are not in the target population of this survey since students living in dorms were sampled through their permanent residences. Since these units may be included in the USPS DU counts, only segments with zero dormitory population (according to the 2000 Decennial Census) were adjusted.

About 10 percent of segments in the frame met all of the above criteria. For these segments, the 2000 Decennial Census segment-level DU counts were adjusted by the following factor:

$$\min\left\{\sqrt{\frac{\text{USPS count}}{\text{2000 census count}}}, 10\right\}.$$

The square root and maximum value of 10 were used in the factor to dampen the effect of USPS counts on MOS (to be conservative in our adjustment). The adjustment factor ranged from 1.096 to 10, with the MOS of about a quarter of the segments inflated by more than 44 percent.

An evaluation of this adjustment showed that using the USPS counts to update the segment MOS improved the MOS accuracy. It further showed that while more segments could have benefited from the adjustment, adjusting all segments in this manner would not have resulted in improved MOS accuracy overall. Finally, using the conservative MOS adjustment shown above resulted in more segments having increased MOS accuracy.

3.1.1.2.2 Selection

The segments were stratified by PSU and selected with probability proportionate to size, with size being based on the segment MOS. The systematic selection used a sorted list based on the geographic sequencing of the segments within the PSU to ensure spatial representation, which also provides a good representation of a variety of demographic subgroups.

The conditional probability of selecting segment *j* from PSU *i* in stratum *h* is

$$CP_{hij} = \frac{q\times(\frac{MOS_{hij}}{P_{hi}})}{\sum_{hij}(MOS_{hij}/P_{hi})} = \frac{MOS_{hij}/P_{hi}}{I_{SSU}},$$

where

q = Total number of segments to be sampled;

MOS_{hij} = MOS for segment j of PSU i in stratum h; and

I_{SSU} = Sampling interval for the selection of segments.

3.1.1.3 Dwelling Units

The third stage of sampling for the PIAAC Main Study involved sampling DUs from listings of addresses in each selected segment to arrive at the targeted number of completed assessments. As mentioned in section 3.1.1, more DUs were selected than needed to guard against unexpected occurrences in the field.

3.1.1.3.1 Dwelling Unit Sampling Frame

All DUs within each selected segment were listed by trained listers. The listing sheets were prepared by the listers and included the information in table 3-4. Table 3-5 provides the quantiles for the number of DUs listed in each selected segment.

Table 3-4. Information on the listing sheets

Information	Description
Listed by	Contains the name of the lister who conducted the listing procedures
Name of city, township, etc.	Contains the name of the city or town
Zip code(s)	Contains the 5-digit zip code and the 4-digit extension when available
PSU #	Contains the PSU ID
Segment #	Contains the segment ID
MS	Missed structure flag for the segment
Line #	Contains a consecutive number for each listed dwelling unit
House #	House number
Street name	Street name
Apartment #	Apartment number
Description of location	Contains comments that may help to describe the location
Group quarters	Contains a flag to indicate whether the dwelling unit is a group quarter
Remarks	Contains information that is special to identify the listed DUs
Hidden DU flag	Contains a flag to identify the hidden DUs

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012.

Table 3-5. Selected percentiles of the number of dwelling units listed per segment

Percentile	Number of dwelling units
100 (Max)	328
99	283
95	249
90	223
75 Q3	166
50 (Median)	104
25 Q1	74
10	63
5	59
1	38
0 (Min)	0

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012.

3.1.1.3.2 Selection of Dwelling Units

Given the actual number of listed DUs and derived sampling rates for each segment, line numbers were selected from the listing sheets to identify the DUs. The DUs were stratified explicitly by segment and selected systematically using predefined sampling rates assigned for each segment in order to arrive at a self-weighting sample of DUs.

The conditional probability of selecting housing unit *k* from segment j of PSU i in stratum h is

$$CP_{hijk} = \frac{r}{P_{hi} \times CP_{hij}} = \frac{r \times I_{SSU}}{MOS_{hij}},$$

where $r = d/D$, and where

d = Total number of housing units to be sampled; and

D = Total number of housing units in the target population.

The overall probability of selecting housing unit *k* in segment *j* of PSU *i* of stratum *h* is

$$P_{hijk} = P_{hi} \times CP_{hij} \times CP_{hijk} = r.$$

The DU sample size in a segment is

$$d_{hij} = CP_{hijk} \times D_{hij} = r \times I_{SSU} \times \frac{D_{hij}}{MOS_{hij}},$$

where

d_{hij} = Number of housing units to be sampled in segment *j* of PSU *i* of stratum *h;* and

D_{hij} = Number of housing units in segment *j* of PSU *i* of stratum *h*.

3.1.1.3.3 Procedures for Selecting Missed Structures and Hidden Dwelling Units

The missed structure and hidden DU procedures were developed to correct for any undercoverage that occurred during the listing operation. Procedures were implemented during data collection to handle any DUs identified through the missed structure and hidden DU procedures.

For the missed structure procedure, interviewers looked for entire structures missed during the listing operation within segments where the first line number (first dwelling unit listed) within each segment was selected. The segments designated for the quality check were selected at a rate such that the inclusion of all units found retained the self-weighting feature of the sample. Twenty-seven DUs were added through the missed structure procedure and received the probability of selection associated with the dwelling unit with the first line number in the segment.

For each selected structure (e.g., DU), the hidden DU procedure was applied at the time of screening. The hidden DU procedure involved looking for DUs within a structure not included in the listed DUs during the listing operation. Forty DUs were added in the PIAAC Main Study through the hidden DU procedure, and each received the probability of selection associated with the selected structure for which the hidden DU was found.

3.1.1.4 Persons

The fourth stage of selection involved enumerating the age-eligible household members (aged 16 to 65) for each selected household (occupied dwelling unit).

3.1.1.4.1 Person Selection Frame

The enumeration and selection of persons was performed using a CAPI system. Using the Screener instrument, information that included age and gender of persons in the household was collected and a systematic sample of eligible persons selected. Household members away at college and staying in college dormitories were considered to be part of their family's household.

3.1.1.4.2 Person Selection

One person was selected at random within households with three or fewer eligible persons, and two persons were selected if the household had four or more eligible persons. The design involves the selection of two persons in households with a large number of eligible persons to reduce the variation in the resulting sampling weights. Taking only one eligible person per household from households with a considerable number of eligible persons causes substantially different weights and consequently unduly increases the variance of the sample estimates.

Students residing in dormitories were sampled through their permanent residence. Under this design the dormitories were excluded from the listing procedure to ensure that such students had only one chance of selection. If a college student was selected as a respondent from the sampled household, there were two options: (1) depending on the location and the availability of nearby staff, an interview was attempted at

the student's dormitory, or (2) an interview was scheduled with the student at the sampled residence, at a time when he or she was home from college.

The conditional probability of selecting person l from housing unit k of segment *j* in PSU *i* within stratum *h* is

$$CP_{hijkl} = \frac{n_{hijk}}{N_{hijk}}.$$

The overall probability of selecting person *l* from housing unit *k* of segment *j* in PSU *i* within stratum *h* is

$$P_{hijkl} = P_{hi} \times CP_{hij} \times CP_{hijk} \times CP_{hijkl} = r \times \frac{n_{hijk}}{N_{hijk}},$$

where

n_{hijk} = Number of persons to be sampled from housing unit *k* of segment *j* in PSU *i* within stratum *h*; and

N_{hijk} = Total number of eligible persons in housing unit *k* of segment *j* in PSU *i* within stratum *h*.

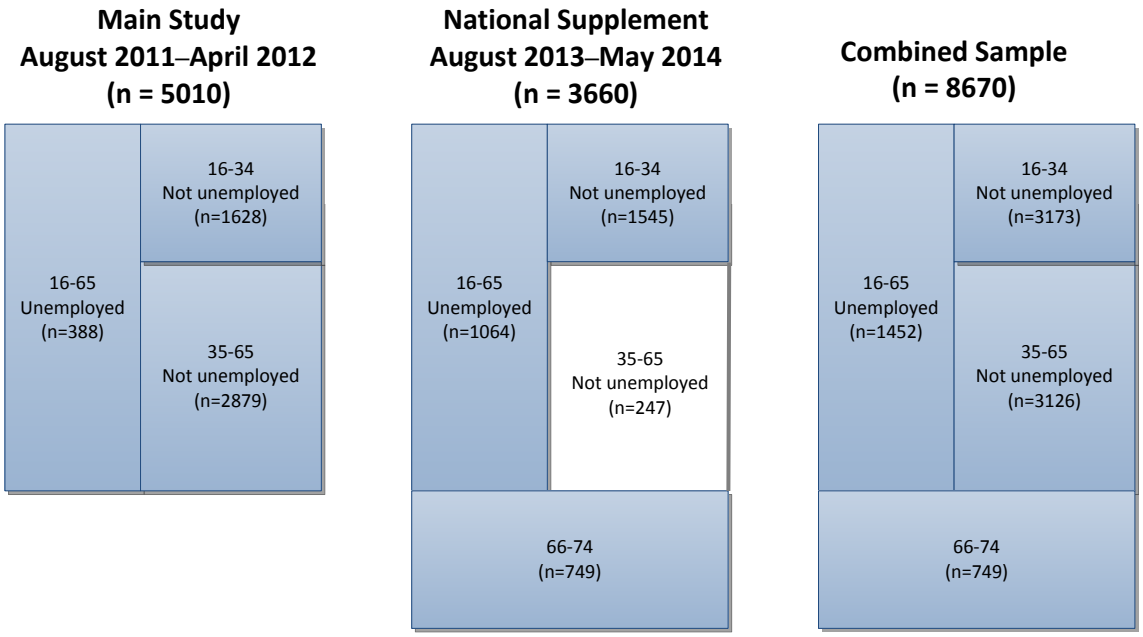
Once the selection process was complete, the Screener data were read into the virtual machine platform system and then the BQ interview conducted. The Information and Computer Technology (ICT) Core assessment occurred after the interview, and the virtual machine system proceeded with the Core assessment and subsequently the main assessment.

3.1.2 National Supplement Sample Design and Selection

The purpose of the PIAAC National Supplement household sample was to provide additional samples of adults for the unemployed (age 16 to 65), young adults (age 16 to 34), and an age 66 to 74 cohort (regardless of employment status). The target population for the National Supplement's household-based sample consisted of noninstitutionalized adults, 16 to 74 years old,[19] who resided in the United States at the time of interview, *excluding* adults 35–65 who were either employed or not in the labor force as determined by the Screener interview. Figure 3-1 illustrates the samples from the Main Study and National Supplement.

[19] Age is determined during the Screener questionnaire.

Figure 3-1. Illustration of the components of the combined household sample



NOTE: Shading indicates inclusion in the target population. As a result of employment status misclassification (discussed in section 3.1.3), the final National Supplement household sample included 247 adults ages 35 to 65 who were not unemployed, as shown in the unshaded portion. The numbers in parentheses indicate the number of respondents (cases assigned final weights). The components do not sum to the overall total because of 115 respondents in the Main Study sample and 55 respondents age 16 to 65 in the National Supplement sample for whom final employment status was unknown.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

The sample selection method for the National Supplement household sample differed from the Main Study sample design. Given the sample size goal for unemployed and the low prevalence of unemployed adults in the population, a dual-frame approach was implemented, which is a more efficient method of sampling rare populations. The dual-frame approach consisted of an area sample and a list sample.

Under this approach, an area sample of DUs was selected from the same PSUs and segments selected for the Main Study. The DU frame consisted of the PIAAC Main Study listings after removing the DUs previously released. One or more persons from the National Supplement household sample target population was sampled within a household, as described further in section 3.1.2.2.3.

To obtain the oversample of unemployed adults, the frame was supplemented with a list of DUs from high unemployment census tracts. Within each of the PSUs, five high unemployment tracts were identified and one was randomly selected for the National Supplement list sample. The USPS address list was purchased for each of the sampled tracts, and a sample of DUs was taken from these lists. Within the sampled DUs, only those who were unemployed were eligible for selection.

Specifically, to arrive at a minimum of 3,600 completed cases for the National Supplement, the four-stage, stratified area frame probability sample involved the following steps:

- 80 PSUs previously selected for the Main Study consisting of counties or groups of contiguous counties;
- 896 secondary sampling units (SSUs or segments) previously selected for the Main Study consisting of census blocks or block groups;
- 9,579 DUs; and
- 3,617 individuals within DUs resulting in 2,790 respondents to the survey.

The list sample involved the following steps:

- 80 PSUs previously selected for the Main Study consisting of counties or groups of contiguous counties;
- 80 SSUs consisting of census tracts;
- 6,956 DUs; and
- 951 individuals within DUs resulting in 870 respondents to the survey.

Random sampling methods were used, with known probabilities of selection at each sampling stage. The National Supplement household sample design resulted in a sample that is not stand-alone, but is nationally representative when combined with the Main Study. Combining the National Supplement household sample with the Main Study sample provided larger subgroup sample sizes that produced estimates of higher precision for the subgroups of interest. A description of the method used to combine the samples can be found in chapter 8.

The Main Study and the National Supplement household sample each consisted of four data collection stages. During the fourth stage of selection, a Screener interview was used to identify the eligible persons within selected DUs. A sampling algorithm was implemented within the CAPI system to select one or more sample persons among those identified to be eligible. Once selected, the BQ interview was completed. Upon completion of the BQ, respondents were provided either the paper-and-pencil or computer-based assessment based whether they reported having any previous computer experience during the BQ interview or whether they refused the computer-based assessment as well as their performance on the computer technology (ICT) Core instrument, conducted after the BQ.

A monetary incentive of $5 was paid to household representatives who completed the Screener for the National Supplement. (No monetary incentive was paid in the Main Study for completing the Screener.) The Screener incentive for the National Supplement was intended to help reduce nonresponse to a slightly longer Screener than that used in the Main Study. Specifically, the National Supplement's Screener included various questions about unemployment status that were not in the Main Study's Screener. As in the Main Study, following the completion of the assessment, an additional monetary incentive of $50 was paid to each respondent. The incentive was also paid to those adults who attempted to complete an assessment but were legitimately not able to complete it for reasons of language barriers or physical or mental disabilities. Respondents who refused to continue with the assessment were not compensated.

3.1.2.1 Area Sampling Frame and Selection

3.1.2.1.1 Primary Sampling Units and Secondary Sampling Units

The same PSUs and segments selected for the PIAAC Main Study were used for the area sample component of the National Supplement. For details about the selection of PSUs and segments, refer to sections 3.1.1.1 and 3.1.1.2.

3.1.2.1.2 Dwelling Units

The third stage of sampling involved sampling DUs from listings of addresses in each selected segment to arrive at the targeted number of completed cases.

Dwelling Unit Sampling Frame

Since all DUs within each selected segment had been listed by trained listers in the Main Study, line numbers were selected from the same listing sheets to identify the DUs for the National Supplement. The DUs that had been released in the Main Study were not eligible to be selected for the National Supplement sample and were removed from the sampling frame. Table 3-6 provides the quantiles for the number of DUs listed in each selected segment excluding the DUs released in the Main Study.

Table 3-6. Selected percentiles of the number of dwelling units listed per segment, excluding the DUs released in the PIAAC Main Study

Percentile	Number of dwelling units
100 (Max)	298
99	272
95	235
90	212
75 Q3	155
50 (Median)	93
25 Q1	64
10	54
5	50
1	41
0 (Min)	5

NOTE: Percentiles are based on the count of DUs within the 896 segments with eligible DUs.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Selection of Dwelling Units

The DUs were stratified by segment, and selected systematically using predefined sampling rates assigned for each segment in order to arrive at a self-weighting sample of DUs. More DUs were selected than needed to guard against unexpected occurrences in the field.

The probability of a DU not being selected into the Main Study is[20]

$$P'_{hijk} = P_{hi} \times CP_{hij} \times \left(1 - CP_{hijk}\right) = \frac{MOS_{hij}}{I_{SSU}} \times \left(1 - \frac{r \times I_{SSU}}{MOS_{hij}}\right) = \frac{MOS_{hij}}{I_{SSU}} - r.$$

Let $r' = d'/D$, where, d' = total number of DUs to be sampled for the National Supplement area sample, D = total number of DUs in the target population. Within each segment, selection from DUs not released in the Main Study was done with the conditional probability of selecting DU *k* from SSU *j* of PSU *i* in stratum *h* as a function of the selection rate r':

$$CP_{hijk}^{NS} = \frac{r'}{\frac{MOS_{hij}}{I_{SSU}} - r}$$

For the National Supplement area sample, the overall probability of selecting DU *k* in segment *j* of PSU *i* of stratum *h* is

$$P_{hijk} = P_{hi} \times CP_{hij} \times (1 - CP_{hijk}) \times CP_{hijk}^{NS} = r'.$$

For the Main Study, the missed structure and hidden DU procedures were developed to correct for any undercoverage that occurred during the listing operation (see section 3.1.1.3.3). The hidden DU procedure was continued for the National Supplement household sample. The missed structure procedure was not necessary for the National Supplement since the segments for the operation had already been selected and the procedure already implemented.

3.1.2.1.3 Persons

The fourth stage of selection involved enumerating the age-eligible household members (age 16 to 74) and asking about employment status for those who were 16–65 years old.

Person Selection Frame

As in the Main Study, the enumeration and selection of persons was performed using a CAPI system. Using the Screener instrument, the household respondent was asked to enumerate people who lived in the dwelling and had no usual place of residence elsewhere. Those who thought of the household as their primary place of residence or who spent most of the year in the household even though they may have had another residence were listed as eligible household members. The list included persons who usually stayed in the household but were temporarily away on business, vacation, in a hospital or living at school. For each enumerated person, age and gender were collected. Then only for those who were eligible based on age, a short series of questions were asked to determine their employment status.

[20] This is a simplified formula. In practice, the actual Main Study probabilities were used, which reflect changes to the subsampling procedures as described in section 8.1.2.1.1.

Person Selection

Prior to selection, the enumerated individuals were stratified into the following three sampling domains:

- Domain 1—Unemployed, 16–65 years old;
- Domain 2—Not unemployed 16–34 years old and unknown employment status 16–65 years old; and
- Domain 3—66–74 years old regardless of employment status.

Households without an individual in one of the above domains were classified as ineligible. If employment status could not be determined in the Screener, and the age was between 16 and 65, the eligibility status of the individual was unknown. Such persons were included in Domain 2, as shown above, and given a chance of selection. Selection rates were assigned for each domain such that the overall target sample sizes could be achieved.

For the area sample, the sampling rule was as follows for sampling domains:

- For unemployed persons (age 16 to 65), all persons up to 4 were selected.
- For adults (age 16 to 65) with unknown employment status or young adults (16 to 34 years old) who were not unemployed, a predetermined rate was applied to determine if any selection occurred; if so, one person was selected from this group. The predetermined rate was given to each household in order to reduce the number of completes to the targeted amount for domain 2 which is more prevalent than the others. Based on experience from the first two months of data collection, it was decided to decrease the predetermined rate for this group in the third sample release. In the fourth release,[21] the rate was reset to equal the predetermined rate.
- For adults age 66 to 74, all persons up to 2 were selected.

The sampling rules within households were based on sample sizes, instead of sampling rates, to control the maximum number selected within a household. The sampling rules allowed for up to seven persons selected within a household, although in practice the number selected never exceeded five. About 0.7 percent of the households that responded to the Screener had more than two persons selected.

The conditional probability of selecting person l in sampling domain g from housing unit k of release group f and segment j in PSU i within stratum h is

$$CP_{hijfkgl} = r_{fg} \times \frac{n_{hijkg}}{N_{hijkg}},$$

where

r_{fg} = Rate at which selection occurs within a household for sampling domain g and release group f,

[21] Refer to section 3.1.3 for more information on sample release groups.

n_{hijkg} = Number of persons to be sampled from housing unit *k* of SSU *j* in PSU *i* within stratum *h* and sampling domain *g*, and

N_{hijkg} = Total number of eligible persons in housing unit *k* of SSU *j* in PSU *i* within stratum *h* and group *g*.

The values of r_{fg} are provided in table 3-7.

Table 3-7. Rate at which within-household selection occurs for sampling domain g and release group f ($\boldsymbol{r_{fg}}$) for the area sample

Release group (*f*)	Sampling domain (*g*)	$\boldsymbol{r_{fg}}$
Any	Unemployed (age 16–65)	1
1, 2 or 4	Adults (age 16–65) with unknown employment status or young adults (age 16–34) who are not unemployed	0.68343
3	Adults (age 16-65) with unknown employment status or young adults (age 16–34) who are not unemployed	0.52945
Any	Adults (age 66–74)	1
Any	Other	0

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

The overall probability of selecting person *l* in domain *g* from housing unit *k* of release group *f* and segment *j* in PSU *i* within stratum *h* is

$$P_{hijfkgl} = P_{hi} \times CP_{hij} \times (1 - CP_{hijk}) \times CP_{hijk}^{NS} \times CP_{hijfkgl} = r' \times r_{fg} \times \frac{n_{hijkg}}{N_{hijkg}}.$$

Once the selection process was complete, the Screener data were read into the virtual machine (VM) platform system and then the BQ interview conducted. The computer-based assessment (CBA) Core Stage 1 occurred after the interview, and the VM system proceeded with the Core assessment and subsequently the main assessment, as described in section 5.3.3.

3.1.2.2 List Sampling Frame and Selection

3.1.2.2.1 Primary Sampling Units and Secondary Sampling Units

The 80 PSUs for the list sample were the same as those for the Main Study and area sample. A description of the PSU selection can be found in section 3.1.1.1. The second stage of sampling involved the identification of five high unemployment census tracts within each of the 80 sampled PSUs. One of the five tracts in each PSU was selected with probability proportionate to the unemployed population.

Secondary Sampling Unit Sampling Frame

For the second stage of selection, the initial frame consisted of all Census 2010 tracts within the 80 sampled PSUs, with tract-level population data from the Census 2010 Summary File 1 (SF 1) and the American Community Survey (ACS) 2007–2011. In addition, counts of USPS addresses were obtained for each of the tracts. As stated above, the list sample was not a nationally representative sample but instead was designed to supplement the PIAAC sample with additional unemployed adults. The final frame of tracts was constructed to meet this goal.

First, small census tracts were removed from the frame. A tract was considered small if the ACS 2007–2011 population age 16 to 64 was less than 500 or if the number of addresses on the USPS list was under 250. Typically, secondary sampling units that do not meet a minimum MOS criterion are combined with other units, as was done for the Main Study. However, combining tracts would have resulted in geographic areas that were too large to be operationally feasible for data collection.

Second, a match rate model (Montaquila, Hsu and Brick 2011), like that described in section 3.1.1.2, was fitted for each of the tracts. The match rate provides an indicator of the coverage of the USPS address list. The USPS address lists have been shown to have over a 90 percent coverage rate overall; however, this can be considerably lower for some rural areas (Kalton, Kali, and Sigman 2014). Given that the USPS list was to serve as the frame of DUs for the list sample, the coverage rate was taken into consideration when constructing the frame by excluding tracts where it was expected to be low. Therefore, tracts with a match rate under 0.5 were removed from the frame, except in two PSUs where there were less than five tracts with a match rate over this value.

Third, the geographic overlap of the tracts and the Main Study segments was determined using mapping files. The Main Study segments were based on Census 2000 geography, and so a segment was not always fully contained within a Census 2010 tract. Tracts that had over 50 percent of the area overlapping with one or more Main Study segments were removed. This was done to reduce the possibility of a Main Study or a National Supplement area sample respondent being reselected for the survey.

Within each PSU, the resulting tracts were sorted by the ACS 2007–2011 proportion unemployed among adults age 16 to 64. The five tracts in each PSU with the highest proportion of unemployed adults were retained. This served as the frame for the selection of second-stage units for the list sample.

Selection of Secondary Sampling Units

From the frame of five tracts in each PSU, one tract was selected with probability proportionate to size, with size being equal to the unemployed population age 16 to 64. The actual unemployed population during the data collection period was unknown and difficult to estimate given the instability in the unemployment rates. Estimates from ACS 2007–2011 were used, which covered the period before and after the start of the recession and were the most recent tract-level data available.

The conditional probability of selecting tract *j* from the five tracts in PSU *i* in stratum *h* is

$$CP_{hij}^{LS} = \frac{MOS_{hij}}{\sum_j MOS_{hij}}.$$

3.1.2.2.2 Dwelling Units

The third stage of sampling involved selecting DUs. The USPS address list was purchased for each of the sampled tracts, and a random sample of DUs was taken from these lists using an Address-Based Sampling (ABS) approach, with the goal of attaining additional completed assessments for unemployed adults. As in the Main Study and area sample, more DUs were selected than needed to guard against unexpected occurrences in the field.

Dwelling Unit Sampling Frame

For the selection of DUs, residential addresses were obtained from the USPS Computerized Delivery Sequence (CDS) file through a third-party vendor. The lists included street addresses (city-style, rural route, and highway contract), along with the carrier route information. Qualified vendors received updated lists from the USPS on a bimonthly or weekly basis and attached additional information from outside commercial databases. The vendor assigned Census 2010 geography to each address based on street-level geocoding, where possible; otherwise, the centroid of the zip+4 or zip code was used. The final frame for the list sample DUs consisted of any addresses that geocoded into the sampled tracts. Table 3-8 provides the quantiles for the number of DUs geocoded into each selected tract.

Table 3-8. Selected percentiles of the number of dwelling units on the list sample frame per tract

Percentile	Number of dwelling units
100 (Max)	3220
99	3220
95	2720
90	2547
75 Q3	1948
50 (Median)	1514
25 Q1	1118
10	810
5	671
1	423
0 (Min)	423

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Selection of Dwelling Units

The next step was to take a systematic sample of DUs in each tract, where the DUs were sorted geographically by zip code and carrier route information to best avoid sampling neighboring households. The sampling rate within a tract was preassigned based on the goals of achieving a balanced workload across tracts while avoiding excessive design effects due to unequal weights.

The conditional probability of selecting housing unit *k* from tract *j* of PSU i in stratum *h* is

$$CP_{hijk}^{LS} = r_{hij},$$

where r_{hij} = preassigned sampling rate in tract *j* of PSU *i* in stratum *h*.

The overall probability of selecting housing unit *k* in tract *j* of PSU *i* of stratum *h* is

$$P_{hijk}^{LS} = P_{hi} \times CP_{hij}^{LS} \times CP_{hijk}^{LS}.$$

The initial DU sample size in a tract is

$$d_{hij} = CP_{hijk}^{LS} \times D_{hij},$$

where D_{hij} = number of housing units in tract *j* of PSU *i* of stratum *h*.

The selection was done independently of the National Supplement area sample and Main Study sample, and sampled DUs were compared to those in the other two samples to ensure that the DU was not visited twice. Dwelling units identified as duplicates were removed from the list sample. Given that the list sample serves only as a supplemental sample, no missed structure or hidden DU procedures were implemented.

Procedures for Selecting Drop Units

Drop units occur when multiple housing units receive their mail at a single drop point. The USPS lists contain the drop point address and the count of housing units associated with the drop point. Typically there are two or three units associated with a drop point, as in a duplex, but in some instances there can be a large number of units, such as apartments in a high-rise. Each drop unit was treated as a separate record during sampling. About 1 percent of the selected addresses for the list sample were drop units. The base address of the drop point was known, but not the specific address of the drop unit. Therefore, a special procedure was needed to determine which unit(s) to interview.

For any cases flagged as a drop unit, interviewers were instructed to list all units at the base address. The home office then randomly selected the unit(s) to be interviewed. If the actual number of drop units at a drop point differed from the expected number, the selection probability was adjusted by the following factor:

$$f_{hijk} = m_{hijk} \times \frac{M'_{hijk}}{M_{hijk}},$$

where

m_{hijk} = the number of drop units selected at the drop point associated with dwelling unit *k* in tract *j* of PSU *i* of stratum *h*;

M'_{hijk} = the expected number of drop units at the drop point associated with dwelling unit *k* in tract *j* of PSU *i* of stratum *h*; and

M_{hijk} = the actual number of drop units at the drop point associated with dwelling unit *k* in tract *j* of PSU *i* of stratum *h*.

3.1.2.2.3 Persons

The fourth stage of selection involved enumerating the eligible household members for each selected household. The construction of the frame was the same as in the area sample, except that only unemployed adults aged 16 to 65 were eligible. The study specifications called for the selection of up to four unemployed persons (16 to 65 years old) within a household.

The conditional probability of selecting person *l* from housing unit *k* of tract *j* in PSU *i* within stratum *h* is

$$CP_{hijkl} = \frac{n_{hijk}}{N_{hijk}}.$$

The overall probability of selecting person *l* from housing unit *k* of tract *j* in PSU *i* within stratum *h* is

$$P_{hijkl} = P_{hi} \times CP_{hij} \times CP_{hijk} \times CP_{hijkl},$$

where

n_{hijk} = Number of persons to be sampled from housing unit *k* of segment *j* in PSU *i* within stratum *h*; and

N_{hijk} = Total number of eligible persons in housing unit *k* of segment *j* in PSU *i* within stratum *h*.

Once the selection process was complete, data collection proceeded as in the Main Study and National Supplement area sample.

3.1.3 Sample Sizes and Response Rates

To achieve the targeted number of completed cases for the PIAAC Main Study (5,000) and National Supplement (3,600) household sample, assumptions were made regarding the rates of occupancy of the selected DUs, the eligibility of household members, the level of cooperation of the selected individuals, and the misclassification rate. Tables 3-9 and 3-10 provide a summary of the actual rates experienced. The samples experienced a slightly higher occupancy rate and rate of two-person households than assumed and a lower than assumed eligibility rate. The difference in eligibility rates between the three samples reflects the differences in the target populations. The list sample eligibility rate is lowest, since only unemployed adults (age 16 to 65) were eligible.

Employment status misclassification occurred in the National Supplement when a person was identified as unemployed in the Screener by the household representative but changed classification based on the sampled person's responses to the BQ items. The employment status misclassification rate represents the net loss in anticipated unemployed respondents due to this misclassification. If the household representative and the sampled person were not the same person, there tended to be more discrepancies in their responses to the employment items. In addition, there could have been a time lag between the Screener and BQ, resulting in different reference periods, and the sampled person's employment status tended to be more susceptible to change over time. The employment status could also have differed between the Screener and BQ because of misreporting or based on the responses to the four additional

employment items in the BQ. While some misclassification of age was experienced as well, this had a negligible impact on sample sizes.

Table 3-9. PIAAC household sample occupancy and eligibility rates

Component	Main Study (percent)	National Supplement	
		Area sample (percent)	List sample (percent)
Screener—occupancy rate	86.4	86.7	86.5
Screener—eligibility rate	81.8	44.0	15.2
Rate of Screeners completed with two or more sample persons	7.3	18.8	17.5
Rate of employment status misclassification	†	31.9	34.0

† Not applicable.
NOTE: Rates are unweighted.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Table 3-10. PIAAC household sample weighted response rates

Component	Main Study (percent)	National Supplement		Main Study and National Supplement combined (percent)
		Area sample (percent)	List sample (percent)	
Screener	86.5	81.4	84.8	84.7
BQ	82.2	78.1	92.9	80.9
Assessment	99.0	98.5	98.8	98.8
Overall	70.3	62.6	77.9	67.8

NOTE: Technical problems with the computer-assisted BQ and Assessment, which were provided by the Consortium for use in the Main Study and National Supplement, are excluded from the numerator of response rate computations to be consistent with NCES standards, although deviating from PIAAC *Technical Standards and Guidelines*.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Table 3-11 provides a summary of the actual sample sizes, and table 3-12 provides sample sizes by census region and metropolitan status. The sample sizes by component were shown in figure 3-1.

Table 3-11. Sample sizes for the PIAAC household sample

Sample	Main Study	National Supplement			Main Study and National Supplement combined
		Area sample	List sample	Area and list combined	
PSUs	80	80	80	80	80
Selected SSUs (segments or tracts)	901	901	80	981	981
Selected SSUs with eligible dwelling units	896	896	80	976	976
Selected dwelling units (after removing duplicates)[1]	15,580	14,067	11,532	25,599	41,179
Dwelling units released	9,401	9,555	6,947	16,502	25,903

See notes at end of table.

Table 3-11. Sample sizes for the PIAAC household sample—Continued

Sample	Main Study	National Supplement			Main Study and National Supplement combined
		Area sample	List sample	Area and list combined	
Dwelling units added through quality control of listings	67	24	9	33	100
Dwelling units screened[2]	9,468	9,579	6,956	16,535	26,003
Eligible households screened	5,686	2,986	788	3,774	9,460
Sample persons	6,100	3,617	951	4,568	10,668
Background Questionnaires[3]	5,010	2,790	870	3,660	8,670
Completed cases[4]	5,011	2,781	869	3,650	8,661
Assessments (with reading components)	4,835	2,673	847	3,520	8,355
Assessments (without reading components)	4,842	2,679	846	3,525	8,367
Background Questionnaires not completed as a result of technical problems	20	9	2	11	31
Assessments not completed as a result of technical problems	11	15	2	17	28

[1] Initially 11,605 DUs were selected for the list sample. However, 73 of these DUs were found to be duplicates of those sampled in the area sample or PIAAC Main Study, and these cases were dropped prior to the release of the sample. See section 3.1.2.2.2.
[2] Includes released DUs plus DUs added through the quality control of listings.
[3] Cases receiving a final weight for analysis.
[4] Defined by PIAAC *Technical Standards and Guidelines* 4.3.3.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Table 3-12. Distribution of PSUs, segments, DUs, and persons and numbers of completed Background Questionnaires for the PIAAC household sample, by census region and metropolitan status

Domain	PSUs	Segments	DUs	Persons	Background Questionnaires
Total	80	981	26,003	10,668	8,670
Census region[22]					
Northeast	18	226	5,684	2,289	1,753
Midwest	17	217	5,822	2,265	1,882
South	31	378	10,250	4,319	3,555
West	14	160	4,247	1,795	1,480
MeSA status					
NonMeSA	11	173	4,445	1,689	1,398
MeSA	69	808	21,558	8,979	7,272

NOTE: NonMeSA: All counties in the PSU are not part of a Metropolitan Statistical Area; MeSA: All or some of the counties in the PSU are part of a Metropolitan Statistical Area. See *Federal Register, Vol. 75, No. 123, June 28, 2010,* for more information.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

[22] Details about the definition of census region can be found at http://www2.census.gov/geo/docs/maps-data/maps/reg_div.txt.

A sample of 15,580 DUs was selected initially for the Main Study and then divided into subsamples. The entire sample was systematically assigned (with a random start) to several subsamples. Initially, only the first subsample of 7,195 DUs was released. The yield from this release was monitored and used to project estimates of the total yield from this group. Based on these figures, the second group of 1,439 DUs, and then a third group of 767 DUs, was released (see table 3-13).

Table 3-13. PIAAC household sample release group distribution

		National Supplement	
Release group	Main Study	Area sample	List sample
1 (Initial release)	7,195	6,949	4,610
2	1,439	0	1,535
3	767	1,912	802
4	†	694	0

† Not applicable.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

For the National Supplement, a sample of 14,067 DUs was initially selected for the area sample and 11,532 for the list sample. Each sample was then systematically assigned (with a random start) to several subsamples. At the start of data collection, only the first subsample of 11,559 DUs (6,949 from the area sample and 4,610 from the list sample) was released. The yield from this release was monitored and used to project estimates of the total yield from this group, as well as the yield by domain. To improve the yield for unemployed adults, the second release consisted of all 1,535 remaining list sample cases in tracts with an unemployment rate over 15 percent. In the third release, a random sample of 1,912 DUs was released from the area sample, along with a random sample of 802 DUs from list sample tracts with between 10 percent and 15 percent unemployment. The final release consisted of another random group of 694 DUs from the area sample (see table 3-13).

3.1.4 Quality Control of Sample Selection and Sample Monitoring

In surveys with multistage designs such as PIAAC, it is essential to monitor the sampling procedures closely to uncover and correct any errors in the preparation of sampling specifications, computer programming, or clerical operations. The Consortium stipulated several quality control procedures in the PIAAC's *Technical Standards and Guidelines*. These procedures were followed strictly throughout the U.S. PIAAC Main Study sample selection process.

For example, after the PSU, segment, and dwelling unit samples were selected, the expected sample yield was computed to check that it was satisfactory and in line with the sample design expectations. In addition, the quality control sample selection forms required by the Consortium were completed on a flow basis after the sample at each stage was selected. Prior to data collection, the within-household selection algorithm was tested in the CAPI system. A test file was created and processed through the system and checked to ensure that the sample was selected as expected.

While data were collected by the field staff, the following two major tasks were implemented to ensure that the sample requirements were met:

- monitoring and calculating response rates; and
- monitoring sample yield (number and demographic distribution of cases complete and number of cases worked).

Sample monitoring began the first week of the data collection. During the data collection period, a weekly progress report for sample monitoring was produced that contained the number and demographic distribution of cases sampled, cases worked, and cases completed, as well as response rates, for each data collection stage.

Sample yield checks were made by comparing actual sample yield distributions to expected distributions. The information was reported by age and gender (for all stages but the Screener) and some key demographic and geographic subgroup characteristics, such as educational attainment, income, urban/ rural, race/ethnicity, and census region.

The sample design and selection for the National Supplement also adhered to the PIAAC *Technical Standards and Guidelines*, where applicable. The above quality control procedures were implemented in the National Supplement, but no quality control forms were due to the Consortium.

3.2 Prison Sample

This section describes the sample design for the sample of incarcerated adults in the PIAAC supplement Prison Study. An overview of the design is provided in section 3.2.1. Section 3.2.2 focuses on the sample size requirements and coverage issues, including the initial sample size given response rate and eligibility rate assumptions at each stage of data collection. Section 3.2.3 discusses the sample design, sample frames, and sample selection procedures.

3.2.1 Overview

The PIAAC Supplemental Prison Study had a target of a minimum of 1,200 completed cases, including at least 240 females and at least 960 males. In order to achieve this goal, a two-stage, stratified sample was selected with 100 sampled prisons selected in the first stage, among which 80 were all-male or coed prisons and 20 were all-female prisons. Due to higher than expected eligibility and response rates, 1,546 eligible inmates were selected within participating prisons, resulting in 1,319 respondents (cases that received a final weight for analysis) to the survey.

Random sampling methods were used, with known probabilities of selection at each sampling stage. During the second stage of selection, a sampling algorithm was implemented within the CAPI system to select the inmates among those identified to be eligible. Once selected, the BQ interview was completed. Upon completion of the BQ, respondents were provided either the paper-and-pencil or computer-based assessment, based on their performance on the information and computer technology (ICT) Core instrument, conducted after the BQ.

3.2.2 Sample Sizes

The target population of the PIAAC Supplemental Prison Study was inmates age 16 to 74 from federal and state prisons in the United States. At the urging of the PIAAC Prison Expert Group, the following types of facilities and institutions were excluded:

- private facilities not primarily for state or federal inmates;
- military facilities;
- Immigration and Customs Enforcement (ICE) facilities;
- Bureau of Indian Affairs facilities;
- facilities operated by or for local government, including those housing state prisoners;
- facilities operated by the United States Marshals Service;
- hospital wings and wards reserved for state prisoners;
- facilities that hold only juveniles; and
- community corrections facilities (such as halfway-houses, boot camps, weekend programs, and other entities in which individuals are locked up overnight).

To achieve the targeted number of completed assessments (1,200) for the PIAAC Prison Study, assumptions were made regarding ineligibility (prisons that had changed their status), prison nonresponse, and inmate nonresponse. Table 3-14 provides a summary of the actual eligibility and response rates. Table 3-15 provides a summary of the actual sample sizes. Of the 100 sampled prisons, one all-male prison and one all-female prison were closed before the interviews started and thus became ineligible. Two all-male prisons refused to participate in the survey and were each assigned a substitute prison with a similar geographic location, security level, type, and size. One sampled prison was converted from an all-female institution to an all-male institution in 2012. As a result, there were 98 participating prisons, of which 80 were male/coed and 18 were female-only. The prison response rate was 98 percent without substitute prisons and 100 percent with substitute prisons.

Table 3-14. PIAAC Supplemental Prison Study eligibility and response rates (weighted)

Component	Rate
Prison eligibility rate	98.0
Prison response rate (without substitute prisons)	98.0
Prison response rate (with substitute prisons)	100.0
Background Questionnaire—response rate (weighted)	85.8
Assessment—response rate (weighted)	97.7
Overall—response rate (weighted)	82.2

NOTE: Data affected by technical problems with the computer-assisted BQ and Assessment, which were provided by the Consortium for use in the Main Study and National Supplement, were excluded from the numerator of response rate computations to be consistent with NCES standards, although deviating from PIAAC *Technical Standards and Guidelines*.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Table 3-15. PIAAC Supplemental Prison Study sample sizes

Sample	Sample size
Selected prisons	100
Eligible prisons	98
Participating prisons	98
Sample inmates	1,546
Background Questionnaires	1,315
Completed cases[1]	1,303
Cases receiving a final weight for analysis	1,319
Assessments	1,274
Background Questionnaires not completed as a result of technical problems	1
Assessments not completed as a result of technical problems	5

[1] Defined by PIAAC *Technical Standards and Guidelines* 4.3.3.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Of the 1,546 sampled inmates, 1,315 completed the BQ. Of the 231 inmates who did not complete the BQ, 4 were unable to do so because of a literacy-related barrier: either the inability to communicate in English or Spanish (the two languages in which the BQ was administered) or a learning or mental disability. The final response rate for the BQ, which included respondents who completed it and respondents who were unable to complete it because of a literacy-related barrier, was 85.8 percent weighted.

Of the 1,315 inmates who completed the BQ, 1,274 completed the assessment. An additional 10 were unable to complete the assessment for literacy-related reasons, and 1 had a missing paper booklet. The final response rate for the overall assessment was 97.7 percent weighted.

The overall weighted response rate for the prison sample was 82.2 percent (treating substitute prisons as nonresponse). The final prison reporting sample consisted of 1,319 respondents, including 1,315 respondents who completed the BQ plus the 4 respondents who were unable to complete the BQ for literacy-related reasons.

3.2.3 Sample Design, Sampling Frames, and Sample Selection

3.2.3.1 Selection of Prisons

In the first stage of sampling, 100 prisons were selected for the PIAAC prison study. The selection process included the stratification of prisons by gender composition as well as the selection of prisons with probabilities proportionate to a MOS as described in section 3.2.3.1.2.

3.2.3.1.1 Frame

The Prison Study sampling frame was created from two data sources: the most recent Bureau of Justice Statistics Census of State and Federal Adult Correctional Facilities (referred to in the following text as the Prison Census) and the most recent Directory of Adult and Juvenile Correctional Departments,

Institutions, Agencies, and Probation and Parole Authorities available from the American Correctional Association (ACA) (referred to in the following text as the ACA Directory).

At the time of building the sampling frame, the most recent Prison Census had been conducted in 2005. The facility universe for that census was developed from the Census of State and Federal Adult Correctional Facilities conducted in 2000. As defined for the Prison Census, the target population includes the following types of state and federal adult correctional facilities: prisons; prison farms; reception, diagnostic, and classification centers; facilities primarily for parole violators and other persons returned to custody; road camps; forestry and conservation camps; youthful offender facilities (except in California); vocational training facilities; drug and alcohol treatment facilities; and state-operated local detention facilities in Alaska, Connecticut, Delaware, Hawaii, Rhode Island, and Vermont. Facilities were included in the enumeration if they had these characteristics:

- staffed with federal, state, local, or private employees;
- holding inmates primarily for state or federal authorities;
- physically, functionally, and administratively separate from other facilities; and
- operational on December 30, 2005.

The 2005 Prison Census excluded the following types of institutions:

- private facilities not primarily for state or federal inmates;
- military facilities;
- Immigration and Customs Enforcement (ICE) facilities;
- Bureau of Indian Affairs facilities;
- facilities operated by or for local government, including those housing state prisoners;
- facilities operated by the United States Marshals Service;
- hospital wings and wards reserved for state prisoners; or
- facilities that hold only juveniles.

Even though they contain inmates up to age 21, juvenile facilities were excluded from the PIAAC prison sample for two reasons: (1) to remain consistent with the facilities listed in the 2005 Prison Census and (2) to promote cost efficiency because it would not be cost effective to visit these facilities to sample the small number of inmates 16 years of age and older compared with those in the state or federal correctional facilities. (The 2012 American Correctional Association directory showed that in 2011 there were 23,713 juveniles under supervision while there were over 1.5 million adult inmates.)

The ACA Directory, published in 2012, was used to update the sampling frame constructed based on the 2005 Prison Census. The result was a frame consisting of 1,038 prisons, with characteristics as shown in table 3-16.

Table 3-16. PIAAC characteristics of the PSU sample frame

Facility characteristic	Number of prisons
Total	1,038
Gender composition	
All-male	906
All-female	98
Coed	34
Security	
Supermax	40
Max/close/high	351
Medium	396
Minimum/low	251
Type	
Federal	106
State	761
Other[1]	71
Region	
Northeast	160
Midwest	191
South	518
West	169

[1] Other prisons, such as prisons managed under contracts with private contractors that contain federal and state inmates.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Each prison on the frame was assigned an MOS equal to the size of average daily population (ADP). This information was obtained from the Prison Census or the ACA Directory.

3.2.3.1.2 Selection

The prison sampling frame was first stratified by whether or not they housed females only. Within each stratum, the frame was sorted to implicitly stratify the facilities, first by census region (Northeast, Midwest, South and West), then by facility type (Federal and Others), security level (Supermax/Max/ Close/High, Medium, and Minimum/Low/Other), and size (ADP). This process resulted in a list of prisons in which units with similar characteristics were adjacent to each other in the sorted list.

The prisons were systematically selected from the sampling frame with probabilities proportional to the MOS (prison ADP). The probability of facility *i* being selected from stratum *h* is given by

$$P_{hi} = \frac{a_h \hat{x}_{hi}}{\sum_{i=1}^{A_h} \hat{x}_{hi}},$$

where

A_h = number of prisons on the sampling frame in stratum h,

a_h = number of prisons selected for the sample in stratum h, and

$\hat{x}_{hi}$ = estimated number of inmates in facility i in stratum h as it appears on the sampling frame.

The sample contained 80 all-male/coed prisons and 20 all-female prisons, with all-female prisons being selected with higher probabilities than male/coed prisons. The characteristics of the sampled prisons are shown in table 3-17.

Table 3-17. PIAAC characteristics of the sampled prisons

Facility characteristic	Number of prisons
Total	100[1]
Gender composition	
All-male/Coed	80
All-female	20
Security	
Supermax/max/close/high	55
Medium	30
Minimum/low	15
Type	
Federal	9
Others	91
Region	
Northeast	13
Midwest	19
South	45
West	23

[1] Two prisons were closed before the study began.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

3.2.3.2 Selection of Inmates Within Facilities

For the second stage of sampling, a systematic sample of 1,546 inmates from within the 98 participating prisons was selected.

3.2.3.2.1 Frame

The interviewers were provided a sampling frame by the prison personnel at the time of the visit (a list of all inmates occupying a bed the night before inmate sampling was conducted), except for the Federal Bureau of Prisons (BOP)-operated prisons. The sampling frames for the BOP prisons were based on rosters of inmates maintained by the BOP a week before the visit.

3.2.3.2.2 Selection

Inmates were selected with a probability inversely proportional to the prison's population size so that the product of the first- and second-stage selection probabilities was constant. While this sample design was intended to provide a constant overall probability of selection across all inmates, in practice, the overall selection probabilities for sampled inmates varied because of differences between the anticipated and actual sizes of the inmate populations and because of constraints on the sample size per prison. They also varied due to the desired oversampling of female-only prisons.

The conditional probability of inmate *j* being selected in prison *i* from stratum *h*, where *h* differentiates female-only and other prisons, is given by

$$CP_{hij} = \frac{b_{hi}}{\hat{x}'_{hi}},$$

where

b_{hi} = the expected number of inmates to be selected in prison *i* from stratum *h*; and,

$\hat{x}'_{hi}$ = the updated inmate population of prison *i* from stratum *h*, obtained through a telephone call to the facility after its selection into the sample.

The expected number of inmates to be selected in prison *i* from stratum *h* is calculated as

$$b_{hi} = \frac{R_{hi}}{k_h} b_h,$$

where

b_h = the average inmate sample size inflated for anticipated nonresponse;

R_{hi} = $\frac{\hat{x}'_{hi}}{\hat{x}_{hi}}$; and,

k_h = $\frac{\sum_{i=1}^{a'_h} R_{hi}}{a'_h}$

and where

a'_h = the number of participating prisons in stratum *h*.

The expression for k_h is the mean of the R_{hi} values. Note that k_h is equal to 1 if $\hat{x}'_{hi} = \hat{x}_{hi}$ for all sampled prisons (i.e., the number of inmates on the frame is equal to the number of inmates in the prison as determined through telephone contact). If all $\hat{x}'_{hi} > \hat{x}_{hi}$, then k_h will be greater than 1, and b_{hi} will tend to be an average of size b_h.

In order to retain equal workloads in prisons and equal probabilities (among those in female-only prisons, and among those in other prisons), the expression of b_{hi} was substituted in the formula for the conditional probability, CP_{hij}, which gives,

$$CP_{hij} = \frac{R_{hi} b_h}{k_h \hat{x}'_{hi}} = \frac{\hat{x}'_{hi}}{\hat{x}_{hi}} \frac{a'_h}{\sum_{i=1}^{a'_h} R_{hi}} \frac{b_h}{\hat{x}'_{hi}} = \frac{b_h a'_h}{\hat{x}_{hi} \sum_{i=1}^{a'_h} R_{hi}}$$

The overall selection probability of an inmate is thus

$$P_{hij} = P_{hi} CP_{hij} = \frac{a_h \hat{x}_{hi}}{\sum_{i=1}^{A_h} \hat{x}_{hi}} \frac{b_h a'_h}{\hat{x}_{hi} \sum_{i=1}^{a'_h} R_{hi}} = \frac{a_h b_h a'_h}{\sum_{i=1}^{a'_h} R_{hi} \sum_{i=1}^{A_h} \hat{x}_{hi} \, a'_h}$$

Note that P_{hij} is constant across all inmates within stratum h.

An upper bound on the number of selected inmates was set to constrain the inmate sample size per prison. This upper bound was dictated by the practical limitations of interviewing a large number of inmates in each prison. If a prison's expected sample size exceeded the upper bound, it was truncated to the upper bound, and the sample sizes for the other prisons were inflated to yield the total expected inmate sample size. This iterative process continued until there were no prisons with an expected inmate sample size greater than the upper bound. In addition, a lower bound on the number of selected inmates was set to justify the cost of traveling to prisons with a small number of interviews.

The average inmate sample size, b_h, was set to 14.9 for males and 16.8 for females to achieve 1,200 completed interviews based on the anticipated nonresponse rates. Later, it was increased on two subsequent occasions, as 16.1 for males and 18.0 for females, and then to 17.3 for males and 19.2 for females to boost the yield.

Inmates in each non-BOP-operated prison were selected from a list of inmates occupying a bed the previous night. The interviewers received forms and instructions to follow when sampling inmates from the lists. The interviewers had a laptop computer preprogrammed with a sampling algorithm. The statisticians assigned both the random number and the sampling interval to the prison before the fieldwork began; these values were preloaded into the sampling algorithm. The facility name, location, security level, type (federal, state, private), and gender composition (male only, female only, mixed) were also loaded on the laptop. The interviewers were required to verify all sampling information since it was used in the sample weighting process. Since the BOP-operated prisons provided the rosters of inmates a week before the visit, the home office selected the sample using the same computer algorithm as described above.

The algorithm required that the interviewer enter the total number of inmates on the list. After the number of inmates was entered, the sampling algorithm compared the value with a preloaded acceptable range. In general, the acceptable range was computed to be within 10 percent of the expected inmate population. If the number of inmates fell outside the acceptable range, a message appeared on the laptop instructing the interviewer to contact the home office to receive a new sampling rate. After the interviewer entered the appropriate sampling rate, the laptop displayed the number of inmates to be sampled and the selected line

numbers. The interviewer circled these line numbers on the list, and those inmates were selected. If a selected inmate was released the day of sampling, or before interviewing began, he or she was considered a nonrespondent to the survey.[23]

3.2.4 Quality Control of Sample Selection and Sample Monitoring

The quality control procedures in the PIAAC *Technical Standards and Guidelines* were strictly followed throughout the prison sample selection process. For example, after the frame was created, the frame counts by different prison characteristics were compared to the publications by the Bureau of Justice Statistics. After the prison and inmate samples were selected, the expected sample yields were computed to check that they were satisfactory and in line with the sample design expectations. In addition, although the sample was not part of the international comparisons, the quality control sample selection forms required by the PIAAC Consortium were completed on a flow basis after the sample at each stage was selected. Prior to data collection, the within-prison selection algorithm was tested in the CAPI system. A test file was created and processed through the system and checked to ensure that the sample was selected as expected.

While data were being collected by the field staff, two major tasks were implemented to ensure that the sample yield requirements were achieved:

- monitoring and calculating response rates; and
- monitoring sample yield (number and demographic distribution of cases completed and number of cases worked).

Sample monitoring began the first week of data collection. During the data collection period, a weekly progress report for sample monitoring was produced that contained the number and demographic distribution of cases sampled, cases worked, and cases completed, as well as response rates, for each data collection stage.

Sample yield checks were made by comparing actual sample yield distributions to expected distributions. The information was reported by age, gender, race/ethnicity, and some key prison characteristics such as census region, security, facility type, and ADP.

The conditional probability of selecting segment *j* from PSU *i* in stratum *h* is

$$CP_{hij} = \frac{q \times (\frac{MOS_{hij}}{P_{hi}})}{\sum_{hij}(MOS_{hij}/P_{hi})} = \frac{MOS_{hij}/P_{hi}}{I_{SSU}},$$

where

q = Total number of segments to be sampled;

MOS_{hij} = MOS for segment *j* of PSU *i* in stratum *h*; and

I_{SSU} = Sampling interval for the selection of segments.

[23] Seventeen sampled inmates were released or transferred to another facility before the interview and assessment could be completed.

This page intentionally blank.

4. LISTING

4.1 Overview

After primary sampling units (PSUs) and segments were selected for the Main Study, a field listing operation was implemented to create a sampling frame of dwelling units (DUs) for the third stage of sampling. Listers were hired and trained as part of a 2-day training to construct a list of all DUs within the segment boundaries, using tract and segment maps created by the Westat mapping department. Nine hundred one segments (901) were selected and listed for PIAAC.

The National Supplement household sample was selected in part from the listing data collected for the Main Study; no additional listing was conducted. Instead, a dual-frame approach was implemented to more efficiently sample the rare populations required (oversampling of unemployed subgroup) and consisted of an area sample and a list sample.

The area sample was selected from the listing data that had been collected for the Main Study after removing the DUs previously released for the Main Study. This frame was then supplemented with a list sample of addresses purchased from the U.S. Postal Service (USPS) from high unemployment census tracts.

The remainder of this section describes the listing activities that were conducted prior to the start of the Main Study data collection.

4.2 Listing Materials

PIAAC listing materials were developed based on Westat's corporate listing training package. Minor modifications and adaptions were made to these materials to incorporate training points, clarifications, and examples to highlight the listing protocol to be followed to support the PIAAC sample design.

4.3 Staff Training

Lister training included a total of 15 hours, including a home study packet, classroom training, and a field listing exercise and review. Listers received a home study packet to complete 1 week before training. The packet included the Westat Listing Manual and the Westat Listing Manual Exercise Booklet. This manual and the related exercise booklet were designed to help listers become familiar with listing terms and activities before they arrived at training. It included an introduction to listing, descriptions and examples of the materials that would be used, a step-by-step overview, and information about listing special areas and structures. Written exercises were included to ensure that the listers completed the required components of the home study packet. The home study took approximately 2 hours to complete.

Five 2-day in-person lister training sessions were held at Westat's offices in Rockville, Maryland, February 28–March 5, 2011. Multiple training sessions, comprising 18–20 trainees per group, were held to accommodate the field listing exercise in a nearby Rockville neighborhood, without overwhelming the residents. There were no more than two training groups in progress at any one time—one in which listers were completing the first day of classroom training and one in which they were preparing for and

debriefing from the field listing practice exercise completed on the second day of training. Each training session was led by a lead trainer and supported by an assistant trainer, the field manager, and the region's supervisor. The lead trainers were very familiar with Westat's standard listing procedures and the PIAAC specific guidelines.

Of a total of 96 listers hired, 92 attended the in-person training sessions and 89 successfully completed training.

4.4 Listing Operations

Field listing work was undertaken during March and April 2011. Listing activities were completed on a flow basis, and finalized segments were sent to the Westat home office weekly for review and quality control checks described in section 4.5.

Throughout the field period, Westat statisticians and survey operations staff worked together to support the listing process by monitoring the ongoing listing fieldwork and responding to issues as needed. In addition, they worked with listers to complete segment chunking. "Chunking" is a procedure to reduce the burden of listing large sampled segments (generally more than 300 DUs) by dividing the segment into chunks. A chunk was selected with probability proportionate to the estimated number of DUs within the chunk, and listing was conducted within the selected chunk.

Hard-copy materials were stored in and shipped to and from the listers via the Westat field room. In addition to handling the materials, ongoing field room procedures included receipt and quality control of listed segments; recording the selected DUs and related data items on the hard-copy listing sheets; and quality control review of the keyed data file.

Field listing activities were completed for all 901 sampled segments, except for one gated city. A gated city differs from a gated community in that the gated area encompasses an entire municipality rather than just a small neighborhood or development. Municipal authorities refused to allow the PIAAC listers access after repeated refusal conversion attempts by the field manager and the Westat project management team.

4.5 Quality Control

Throughout the listing field period a number of quality control activities were undertaken to ensure that listing procedures were being correctly followed. Quality control activities for listing included the following:

- The first two segments completed by each lister were reviewed by the field manager to ensure that all procedures were followed correctly. As needed, the field manager, supervisor, and lister had a conference call to review problems found during the quality control check. Segment folders with issues were returned to the lister for "repair work" before additional assignments were made.

- Home office/field room staff conducted a detailed review of every completed segment to ensure that all activities were undertaken. Issues were documented on a Listing

Segment Problem Sheet, which was then used to provide feedback to the lister or to make corrections to the listing sheets and maps if needed.

- The following quality control checks and verifications were completed for each segment:
  - all materials returned in the segment folder;
  - listing line numbers recorded accurately;
  - DUs listed in the correct order following the appropriate directions on the map;
  - each DU listed on a separate line;
  - recorded descriptions of rural routes if needed;
  - no duplicate listings;
  - route clearly marked on the segment map; and
  - "No DUs" or listing line numbers recorded appropriately on the segment map.
- The field room also tracked each assigned segment and documented the receipt of each segment folder upon completion.
- Completed segments were reviewed to determine if the number of DUs listed was not within the expected range. In cases where the range expectations were not met, the field room notified the statisticians, who reviewed the segment and advised further quality control measures as appropriate to ensure that the lister canvassed the proper area.
- After the final successful review of the listed segment, the statisticians selected the sampled DUs and the field room documented this information on the hard-copy listing sheets.
- Finally, quality control checks were completed after sampled DUs were keyed to ensure that selected DUs were correctly keyed into the survey control file that would be used during data collection.

This page intentionally blank.

5. DATA COLLECTION

5.1 Introduction

U.S. PIAAC Main Study data collection occurred between August 25, 2011, and April 3, 2012. The goal was to interview and assess 5,000 adults in 80 primary sampling units (PSUs) across the country. The National Supplement was an administration of PIAAC to an additional sample of U.S. adults in order to supplement the U.S. PIAAC Main Study sample. The goal was to interview and assess 3,600 adults in the 80 Main Study PSUs and 1,200 adults inmates incarcerated in 100 state, federal, and private prisons. The National Supplement household data collection occurred between August 26, 2013, and May 5, 2014, and the National Supplement prison data collection occurred between February 10, 2014, and June 13, 2014.

In the Main Study and National Supplement household samples, each sampled household was administered a Screener to determine the eligibility of household members. Within households, each sample person selected was administered (1) an in-person Background Questionnaire (BQ), and (2) either a computer-based assessment or a paper-and-pencil assessment (for those who could not or would not use a computer).

In the National Supplement prison sample, inmates were sampled from a list of inmates provided by the prison and then each sample person selected was administered (1) an in-person BQ, and (2) either a computer-based assessment or a paper-and-pencil assessment (for those who could not or would not use a computer).

A description of the development process as well as additional details about these instruments is provided below.

5.2 Development of the Background Questionnaire and Direct Assessment

In the United States, the Department of Education's National Center for Education Statistics (NCES) collaborated with the U.S. Department of Labor (DOL) on PIAAC. Staff from NCES and DOL are corepresentatives of the United States on PIAAC's international governing body, and NCES has consulted extensively with DOL, particularly on the development of the job skills section of the BQ. Internationally, PIAAC has been developed collaboratively by participating countries' representatives from both ministries or departments of education and labor and by OECD staff through an extensive series of international meetings and work groups. These international meetings and work groups, assisted by expert panels, researchers, and the PIAAC Consortium's support staff, developed the framework used to design the BQ and Direct Assessment and guided the development of a common, international virtual machine software to administer the Direct Assessment uniformly on laptops. All PIAAC countries followed the common standards and procedures and used the virtual machine software when conducting the survey and assessment.

The same procedures and instruments used during the Main Study were employed during the household data collection for the National Supplement.

The BQ for the National Supplement prison sample was specifically tailored to collect information related to the needs and experiences of incarcerated adults. Adaptations to the household questionnaire were

implemented in order to adequately capture information from the prison population. These included deleting questions from the household questionnaire that would be irrelevant to respondents in prison as well as the addition of questions that addressed respondents' specific activities in prison, such as participation in academic programs and English as a Second Language (ESL) classes; experiences with prison work assignments; involvement in nonacademic programs, such as life skills and employment readiness classes; and educational attainment and employment prior to incarceration.

The same Direct Assessment used in the household sample was used for participants in the prison sample.

5.3 Data Collection Instruments

The U.S. PIAAC Main Study and the National Supplement household sample interviews were conducted by means of three distinct instruments: the Screener, the BQ, and the Direct Assessment.

The National Supplement prison sample interview was conducted by means of three distinct instruments as well: the Screener, the BQ, and the Direct Assessment. However, the Screener and BQ instruments used for the prison sample differed from the versions used for the household sample as described in the sections below.

5.3.1 Screener

The Main Study and the National Supplement household sample Screener was a Blaise-based[24] computer-assisted personal interviewing (CAPI) instrument. The household Screener was used to determine the eligibility of household members to participate in the study. Interviewers were required to collect the first name, age, and gender of each household member. After this information was collected, the CAPI system conducted a within-household sampling procedure to select the sample persons (SPs) to participate in the study.

In the Main Study Screener household members who were age 16–65 were eligible to be selected, with up to two sample persons being selected in households with four or more qualifying adults.

The household Screener used for the National Supplement differed slightly from the Main Study version. The National Supplement Screener included questions regarding each person's employment status and the within-household sampling procedure selected eligible household members based on employment status as well as age (16–74). Although this Screener allowed for the selection of up to seven sample persons per household, less than 1 percent of the households had more than two persons selected. See section 3.1.2.1.3 for more information regarding the sample person selection rules.

The household Screeners (Main Study and National Supplement) also collected race and ethnicity data for the sample persons as well as a phone number for subsequent validation of the interviewer's work. Additionally, the Screener included a question designed to determine if the sampled address contained any hidden dwelling units (DUs). Found DUs were reviewed by the home office and added to the sample.

[24] Blaise is a powerful and flexible software package used to develop instruments for CAPI, interactive editing, high-speed data entry, and data manipulation capabilities. Blaise is used widely in Europe and North America by government statistical agencies, private research firms, and universities to support the development of interviewing instruments and related capabilities.

The National Supplement prison sample Screener was also a Blaise-based CAPI instrument. The prison Screener was a much-abbreviated version of the household Screener used to collect the sampled person's facility ID number (inmate ID), gender, age, and race/ethnicity.

5.3.2 Background Questionnaire

The household sample BQ had 10 sections (A-J), which included questions about the sample person's education, work experience, language use, technology use and skills, and literacy practices. A brief summary of each section in the household BQ follows:

A. **General Information.** Questions that determined the month and year of sample person's birth and the sample person's gender;

B. **Education and Training.** Questions relating to the sample person's formal and nonformal education and training;

C. **Current Status and Work History.** Questions about the sample person's employment status at the time of the interview and his or her work history;

D. **Current Work.** Questions about the sample person's current occupation and earnings;

E. **Last Job.** Questions for sample persons who are not currently working but who had recent work experience in the last 12 months or had left paid work longer than 12 months ago;

F. **Skills Used at Work.** Collected information about the frequency of use of a number of generic skills used in the workplace, including communication, presentation, and team-working skills;

G. **Literacy, Numeracy, and ICT Skills Used at Work.** Questions about skill practices at work—specifically reading, writing, mathematics, and information and communication technologies (ICT) activities;

H. **Literacy, Numeracy, and ICT Skills Used in Everyday Life.** Questions about skill practices in everyday life—specifically reading, writing, mathematics, and ICT activities;

I. **About Yourself.** Questions about sample person's attitudes, health, and civic activities; and

J. **Background Information.** Questions about general household information, such as the number and age of children, partner's job status, as well as questions about the sample person's background, including country of birth, linguistic familiarity, and parental background.

The prison sample BQ was a variation of the household sample questionnaire tailored to address the needs and experiences of incarcerated adults.[25] Specifically, the prison sample BQ focused on collecting information about various educational and training activities in prison, such as participation in academic programs and ESL classes, experiences with prison jobs, and involvement in vocational training and nonacademic programs such as employment readiness classes. Several questions were adapted from the Main Study household BQ but referred to experiences in prison. A majority of the prison-specific questions were adopted from the National Assessment of Adult Literacy (NAAL) 2003 prison BQ.

The prison sample BQ had 10 sections (A-J), which included questions about the sample person's education, work experience, language use, technology use and skills, and literacy practices. Prison-specific questions are included with sections B, D, and J.

A brief summary of each section in the prison BQ follows:

A. **General Information.** Questions that determined the month and year of sample person's birth and the sample person's gender;

B. **Education and Training.** Questions relating to the sample person's formal and nonformal education and training and includes questions regarding both past education as well as education while in prison including nonacademic programs, training, and other activities;

C. **Present Work Experience.** Questions about the sample person's work experience while in prison, including skills used at work in prison;

D. **Work History Prior to Incarceration.** Questions about the sample person's occupation and earnings prior to incarceration;

E. **Work Experience Prior to Incarceration.** Questions for sample persons who are not currently working but who had recent work experience in the last 12 months or had left paid work longer than 12 months ago;

F. **Work Responsibilities Prior to Incarceration.** Collected information about the frequency of use of a number of generic skills used in the workplace, including communication, presentation, and team-working skills;

G. **Literacy, Numeracy, and ICT Skills Used at Work Prior to Incarceration.** Questions about skill practices at work – specifically, reading, writing, mathematics, and information and communication technologies (ICT) activities;

H. **Literacy, Numeracy, and ICT Skills Used in Everyday Life In Prison.** Questions about skill practices in everyday life – specifically reading, writing, mathematics, and ICT activities;

[25] The prison-specific questions were developed based on the recommendations of an expert panel as well as a review of the questions in the National Center for Education Statistics' (NCES) previous prison study, which was conducted as a component of the 2003 NAAL.

I. **About Yourself.** Questions about sample person's attitudes, health, and civic activities; and

J. **Background Information.** Questions about general household information, such as the number and age of children, partner's job status, as well as questions about the sample person's background, including country of birth, linguistic familiarity, and parental background.

Both BQs are available in English and Spanish on the NCES website.

5.3.2.1 Background Questionnaire Adaptation and Translation

The Main Study followed the Consortium-prepared guidelines for adaptation and translation. The final adapted national English version was reviewed by the Consortium to verify adherence to the specified design and ensure consistency with international practice. A parallel BQ Adaptation Spreadsheet (BQAS) documentation of changes was produced and submitted to cApStAn, the Consortium member that managed all linguistic aspects of this multinational survey.[26]

Westat relied on its corporate translation unit to implement the Spanish translation of the BQ for the Main Study. The translation unit made the changes required by the Consortium and approved by NCES, following the translation process described in section 2.4. The translated Main Study BQ was submitted to the Consortium.

The Consortium did not review the prison BQ changes for the National Supplement; however, the process followed was the same as that used for the Main Study and Westat again relied on its corporate translation unit to implement the Spanish translation of the prison BQ as required.

5.3.3 Direct Assessment

The Direct Assessment included several sections. The first section was the Core, which was a self-administered series of tasks to determine whether the sample person could use the computer to complete the assessment and had sufficient literacy and numeracy skills to undertake either the computer or paper-based assessment. The Core consisted of three components: the computer-based assessment (CBA) Core Stage 1, the CBA Core Stage 2, and the paper-based assessment (PBA) Core.

The CBA Core Stage 1 was designed to determine whether a sample person had the basic set of skills needed to complete the computer-based assessment. It was administered to all sample persons who were willing to take the assessment on the computer. Those who refused to take the self-administered CBA Core Stage 1 were routed to the PBA Core. The CBA Core Stage 1 measured the following six skills:

1. clicking;

2. typing;

[26] A significant number of wording changes from International English to American English were processed prior to the field test as translations following a request by ROA (Research Centre for Education and the Labor Market). Originally the plan was to treat these as adaptations.

3. selecting from a pull-down menu;

4. scrolling and clicking;

5. dragging and dropping text; and

6. highlighting text.

To pass the CBA Core Stage 1, a sample person needed to perform at least three of the first five Core tasks correctly, plus the sixth task (highlighting text). Because many of the assessment items required highlighting of text, the sample person had to correctly complete the highlighting task in the CBA Core Stage 1 to be directed to the CBA Core Stage 2.

CBA Core Stage 1 questions were automatically scored by the computer, and sample persons who passed the CBA Core Stage 1 continued onto the CBA Core Stage 2. Sample persons who did not pass the CBA Core Stage 1 were routed to the PBA Core.

The CBA Core Stage 2 and the PBA Core both measured basic literacy and numeracy necessary to undertake the assessment. CBA Core Stage 2 consisted of six tasks and the PBA Core consisted of eight tasks, and in both cases these tasks measured basic literacy and numeracy skills.

CBA Core Stage 2 questions were automatically scored by the computer, and sample persons who passed the CBA Core Stage 2 continued on to the computer-based assessment, which was self-administered and measured literacy, numeracy, and problem solving in technology-rich environments. Sample persons who did not pass the CBA Core Stage 2 were routed directly to the paper-based reading components portion of the assessment.

PBA Core questions were scored by the interviewer and scores were entered into the computer to determine if the sample person passed the PBA Core. Sample persons who passed the PBA Core continued on to the paper-based assessment, which was also self-administered and measured literacy, numeracy, and reading components. Sample persons who did not pass the PBA Core were routed directly to the paper-based reading components portion of the assessment.

The same Direct Assessment was used in the National Supplement, both for the household and prison samples, without any changes.

5.3.4 Instrument Timings

The average administration time for the PIAAC Main Study instruments is displayed in table 5-1.

Table 5-1. Average administration time, by instrument—Main Study household sample

Instrument	Average administration time (minutes)
Screener	4.24
Background Questionnaire	36.80
Core – Computer-based	6.80
Core – Paper-based	12.14
Direct Assessment – Computer-based	47.12
Direct Assessment – Paper Booklet 1	24.90
Direct Assessment – Paper Booklet 2	25.11
Direct Assessment – Reading Components	18.69

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012.

The average administration time for the PIAAC National Supplement household sample instruments is displayed in table 5-2.

Table 5-2. Average administration time, by instrument—National Supplement household sample

Instrument	Average administration time (minutes)
Screener	5.67
Background Questionnaire	38.79
Core – Computer-based	7.32
Core – Paper-based	10.59
Direct Assessment – Computer-based	45.01
Direct Assessment – Paper Booklet 1	28.10
Direct Assessment – Paper Booklet 2	26.17
Direct Assessment – Reading Components	18.94

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC) National Supplement, 2014.

The average administration time for the PIAAC National Supplement prison sample instruments is displayed in table 5-3.

Table 5-3. Average administration time, by instrument—National Supplement Prison Sample

Instrument	Average administration time (minutes)
Screener	2.55
Background Questionnaire	39.84
Core – Computer-based	7.53
Core – Paper-based	8.54
Direct Assessment – Computer-based	48.92
Direct Assessment – Paper Booklet 1	31.09
Direct Assessment – Paper Booklet 2	25.92
Direct Assessment – Reading Components	16.82

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC) National Supplement, 2014.

5.4 Field Staff Training

Westat adapted the training materials provided by the Consortium and also prepared U.S.-specific training materials for locating households and respondents, the household Screener instrument, and administrative procedures. All training was conducted using a Trainer's Guide, which included all lecture scripts, role-playing exercises, and written exercises. The Trainer's Guide was used to ensure standardization of the materials presented to interviewers.

Five training sessions were held for the PIAAC Main Study: a train-the-trainers/supervisor training, an initial interviewer training, and three attrition trainings, as described below.

Three training sessions were held for the National Supplement household sample: a train-the-trainers/ supervisor training, an initial interviewer training, and an attrition training, as described below.

One training session was held for the National Supplement prison sample as described below.

5.4.1 Main Study Train-the-Trainers/Supervisor Training

The Main Study train-the-trainers/supervisor training session was held at Westat on July 25–30, 2011. The purpose of this session was threefold: to test the training scripts and procedures, to train the trainers on how to conduct the interviewer training, and to train supervisors on study procedures.

Eleven lead trainers, six data display staff,[27] eleven supervisors, and two field managers were trained during this training session. Revisions were made to the Trainer's Guide and training procedures after this training session.

Supervisors received a supervisor manual during training. This document, used exclusively by the field supervisors, the field managers, and the study manager, covered all of the study procedures and the use of reports for monitoring work in the study regions.

5.4.2 Main Study Interviewer Training

Interviewers received a home study packet to complete a week before training. The home study packet used was based on the version designed by the Consortium but adapted to U.S. study-specific content. It included an introduction covering the study history, design, and purpose of PIAAC, as well as guidance on issues related to obtaining respondent cooperation. A written exercise was included to ensure that the interviewers completed the required components of the home study packet. The home study took approximately 2 hours to complete.

The Main Study interviewer training session was held August 18–23, 2011, at the Renaissance Hollywood Hotel in Hollywood, California. There were 11 training rooms, one for each region in the study. Each room had a lead trainer, the region's supervisor serving as assistant trainer, a data display person from Westat or an experienced interviewer to fill that role, and an average of 17 trainees. Lead

[27] In addition, five experienced interviewers acted as data display operators during the main interviewer training. Consequently only six data display operators were trained during the train-the-trainers/supervisor training session.

trainers were experienced trainers able to motivate trainees and keep the training on schedule. They were very familiar with general interviewing techniques, specific materials, and CAPI applications. Westat technical staff (approximately one for every two rooms), familiar with the interviewer systems and study questionnaires, were available at all times during the training to troubleshoot and answer questions.

Training scripts were adapted from a package provided by the Consortium. Training featured 39 hours of in-person training, including the following topics:

- general interviewing techniques;
- gaining respondent cooperation;
- locating the sampled DUs;
- administering the Screener;
- BQ and Direct Assessment; and
- administrative and quality control procedures.

The basic concepts of the instruments were taught through interactive lectures that increased in complexity as the training progressed. The trainees were led through the CAPI instruments and were called on to play the role of the interviewer while the trainer acted as the respondent. The trainer stopped frequently to explain a question more fully or to make a particular point about a question or its administration.

Towards the end of training, trainees practiced administering the interview via the following:

- **Role-Playing Exercises.** The trainers arranged trainees in pairs, taking into consideration their strengths and weaknesses. Within each pair, one trainee took the role of the interviewer while the other played the respondent, using a prepared script. They then reversed roles. Training staff observed the pairs, correcting the interviewers if needed.
- **Paid Respondent Practice.** The trainees conducted practice interviews with paid respondents. This allowed the trainees to practice in an unscripted situation and provided a more realistic experience of what they would encounter with "real" respondents.

Interviewers were provided with an interviewer manual during training. This document contained an overview of the study and explained data collection procedures and protocols to interviewers.

A total of 186 trainees registered for interviewer training. In addition, trainees had the opportunity to attend CAPI lab evening sessions to improve their CAPI skills or to catch up on missed sessions due to travel delays. Two trainees and a supervisor were dismissed for poor performance before the end of training.

The bilingual interviewers hired were paired up to complete the role-play exercises for the BQ in Spanish.

5.4.3 Main Study Attrition Trainings

As a result of interviewer attrition, two additional Main Study trainings were held in December 2011 and January 2012.[28] A total of nine people were trained in these sessions, one of whom did not complete training due to poor performance. These trainings were held at Westat. One experienced lead trainer and one assistant trainer conducted these trainings based on the same training package used in August 2011. However, the training progressed significantly faster due to the smaller number of people being trained.

5.4.4 National Supplement Train-the-Trainers/Supervisor Training

The train-the-trainers/supervisor training session was held at Westat on July 29–August 3, 2013. Nine lead trainers, nine supervisors, and three field managers successfully completed this training. The purpose of this session was to train the trainers on how to conduct the interviewer training and to train supervisors on study procedures. The supervisor manual was provided and reviewed during training which covered all of the study procedures and the use of reports to monitor the field work.

5.4.5 National Supplement Household Sample Interviewer Training

The National Supplement interviewer training session and materials followed a similar format and approach as that implemented for the Main Study (see section 5.4.2). The interviewer training was held August 17–23, 2013, at the Hyatt Regency Bethesda in Bethesda, Maryland. There were nine training rooms, one for each region in the study. Each room had a lead trainer, with the region's supervisor serving as assistant trainer, and these trainers were supported by a team of technical and operations staff, including the project managers. One hundred thirty-seven (137) field interviewers successfully completed this training. Lead trainers were experienced trainers able to motivate trainees and keep the training on schedule.

5.4.6 National Supplement Household Sample Attrition Training

An additional interviewer training was held November 4–8, 2013, at the Doubletree by Hilton in Silver Spring, Maryland. Thirty-two interviewers were trained during this session. One experienced lead trainer and one assistant trainer conducted these trainings based on the same training package used in August 2013.

5.4.7 National Supplement Prison Sample Training

The National Supplement prison sample data collection training was held at Westat on January 22–24, 2014. The staff trained for the prison sample data collection also worked on the household sample data collection and therefore were very experienced with the majority of study procedures/instruments. Thus these staff required a much shorter training than would normally be required for staff with no prior study

[28] The attrition trainings sessions were held December 9–12, 2011; December 12–15, 2011; and January 10–13, 2012.

knowledge. The training approach was similar to that implemented for the household sample in terms of training style and staffing, logistical arrangements, and trainee evaluation.

The content of the prison sample training focused on specific procedures for data collection within a prison setting. Content focused on an overview of the negotiations process that was completed prior to training; procedures working in the prisons, including security, precautions, and facility staff; inmate sampling procedures and exercises; introduction to and hands-on practice with the prison Screener and BQ; and special procedures that may be required for administering the computer-based assessment within the prison setting.

Thirty-nine field interviewers attended and 38 successfully completed the prison sample training. Additionally, two field managers and two of the field supervisors were trained.

5.5 Approach to Data Collection

5.5.1 Field Management Organization

For the Main Study, two field managers supervised ten regional supervisors.[29] Field managers were responsible for an area of the country and managed the work of supervisors in their area. Supervisors were each responsible for a region (consisting of approximately eight PSUs) and oversaw the work of seventeen interviewers, on average, in their region.

For the National Supplement household sample, two field managers supervised nine regional supervisors. As with the Main Study, field managers were responsible for an area of the country and managed the work of the supervisors, regions, and interviewers in their area.

Training and data collection for the National Supplement prison sample started approximately 5 months after the start of household data collection. Just prior to prison sample training, the field management team was restructured. Two of the nine household supervisors were designated to manage the prison data collection only, and seven supervisors continued to manage the household data collection effort requiring the reassignment of some regions and interviewers to new supervisors. Each of the two prison supervisors was assigned approximately 50 prisons in which to schedule and coordinate the data collection effort using the pool of 38 interviewers trained for this effort.

5.5.2 Interviewer Management and Supervision

Supervisors used a web-based interactive Study Management System (SMS) to manage case information, record interim and final case disposition codes, review interviewer productivity, and monitor overall production in their region. Data collection was monitored through reports reflecting case status, weekly reports by field managers and supervisors, and receipt control reports reflecting completed cases received by the office.

[29] Eleven supervisors were hired but one was dismissed at training.

More specifically, supervisors performed the following tasks regarding managing and supervising field interviewers:

- **Set Goals.** Establish clear production and cost goals with interviewers.
- **Assign Cases.** Assign and reassign cases in accordance with cost-efficiency and response rate requirements.
- **Hold Weekly Interviewer Conference Calls.** Hold weekly conferences by telephone with interviewers to review status of each of their cases, find out how much they have worked, review any problem situations, and motivate them to finish on time. Additionally, be available to receive calls from interviewers who have problems throughout the week.
- **Monitor Progress of Data Collection.** Review production reports regularly and implement strategies for meeting production goals and response rates.
- **Review Nonresponse Cases.** Review interviewer production and implement reassignment and conversion procedures.
- **Monitor Interview Cost.** Review, correct, and sign each interviewer's time and expense report for accuracy and authenticity. Ensure that interviewers are working effectively keeping costs to a minimum.
- **Conduct Quality Control.** Review audio recording from each interviewer's third and tenth completed case and provide feedback.

5.6 Obtaining Household Respondent Cooperation

5.6.1 General Factors that Influenced Household Respondent Cooperation

A variety of approaches and outreach materials were used in both the Main Study and the National Supplement household sample to obtain respondent cooperation. A study-specific logo was designed and used on the materials to lend legitimacy and recognizability to the study.

- **Study Identification Badge.** All interviewers were issued a prominent photo identification badge. Interviewers were required to wear it at all times while conducting fieldwork. The badge authenticated their official role in connection with the study.
- **Advance Contact with Selected Households.** All sampled DUs were mailed an introductory letter on NCES letterhead explaining the study, confidentiality, and the importance of participation. A refrigerator magnet featuring the study logo, the website address, and the toll-free respondent hotline number was included with the letter to make the letter more memorable to households.
- **Brochure.** A study brochure aimed at informing sample persons was developed and printed. Upon selection, sample persons were given the study brochure. As the

brochure provided a concise and well laid-out summary of the study, some interviewers chose to use it to gain cooperation at the screening level.

- **Study Website.** A website featuring clearly articulated, detailed information about the study was set up at NCES. This website could be accessed by interested sample persons to ascertain the purpose and legitimacy of the study during the period of study implementation. The website address was included in the brochure and magnet, and all other outreach materials provided to sample persons.

- **Study Toll-Free Telephone Number.** This number provided contact with a Westat home office project staff person who addressed respondent concerns or relayed messages left by respondents or households to field supervisors. It was cited on all survey documents provided to respondents.

- **Pen with Study Logo.** A pen featuring the U.S. International Survey of Adult Skills (ISAS) logo, website address, and toll-free study number was given to respondents as a token of appreciation.

- **Outreach Flyers.** This series of four one-page flyers was aimed at sample persons who were reluctant to participate. Each flyer was tailored to specific types of reluctant respondents and was mailed to SPs who had not participated after several contact attempts by interviewers.

- **Screener Flyer.** A Screener flyer that featured the $5 incentive for completing the Screener was developed and sent to National Supplement households. The $5 Screener incentive was implemented for the National Supplement data collection since a high level of screening was required to identify sample persons in the targeted groups of interest. During the Main Study, no incentives were given to respondents for completing the Screener interview.

- **Postcard.** This card was mailed 1 month prior to the end of data collection reminding all nonresponse households that the study was about to end and of the importance of participation and the $50 incentive payment.

- **Endorsement Letter.** An endorsement letter from former news anchor Tom Brokaw was available for distribution to respondents at the interviewer's discretion.

- **Nonresponse Letters.** Several types of nonresponse letters were used by supervisors to address specific concerns expressed by households or respondents. These were mailed using priority mail to convey the importance of the study.

In addition to the aforementioned outreach materials, Westat employed the following strategies to gain respondent cooperation:

- **Review of Nonfinalized Cases.** Several times during the field period, a thorough review of the contact attempts for each nonfinalized case was conducted by field supervisors, field managers, and the study manager. Cases that showed potential for completion were flagged and tallied, and additional strategies for contacting these households were discussed with interviewers.

- **Switching Cases between Interviewers.** Supervisors reassigned nonresponse cases (including refusals and no contacts) between interviewers within a PSU to see if a different interviewer was able to gain response.

- **Travelers.** Top-producing interviewers were sent to PSUs in which there were not enough interviewers to work the cases or in which special circumstances required a different interviewer.

- **Mail-in Screener.** A simplified version of the Screener was designed and mailed to households who refused the in-person screening procedure.

5.7 Obtaining Prison and Inmate Cooperation

For the National Supplement prison sample data collection, PIAAC-specific procedures for study consent, obtaining cooperation, interviewing, study management, receipt control, data transmission, and shipping of materials followed the same procedures that were used for the Main Study and the National Supplement household sample. Operational procedures specific to the incarcerated sample included obtaining authorization from the appropriate federal and state agencies; negotiations with the wardens at sampled facilities; inmate sampling; and inmate recruitment.

5.7.1 Summary of the Approach

Prisons were sampled approximately 9 months prior to the start of data collection. Permission and cooperation of Federal, state, and correctional facility officials was required before the National Supplement data collection in the prisons could begin. Extensive negotiations were conducted with the selected correctional facilities to secure cooperation for the study, including detailed tracking of facility-specific requirements identified during the negotiations.

Successful prison sample data collection hinged upon an effective prison recruitment and negotiation strategy. State- and prison-level contacts were contacted approximately 6 months prior to the start of data collection. NCES obtained letters of endorsement from its collaborators and members of their Prison Expert Group to be used during this process. Negotiations included both written and telephone contacts to explain the study, asking for permission to contact selected facilities within the agency's jurisdiction, and determining prison contact procedures. Some institutions required Institutional Review Board approval, thus starting many months prior to the start of data collection was vital to ensure prisons were on board at the start of the data collection period.

Once approval was received from the agency with jurisdiction, each prison was contacted to invite them to participate in the study, to obtain the required permissions and clearances for staff to interview the inmates, to identify a prison coordinator who would be the point of contact when interviewers were working in the facility, and to set a date for data collection to occur at the facility. NCES and its partners also provided support, as necessary, to maximize cooperation from individual facilities.

When data collection began at a particular prison, interviewers traveled to the facility to complete the selection of inmates and the interviews with the sampled prisoners. By the time an interviewer entered any correctional institution, the project negotiator had already obtained that facility's approval for participation, established a contact within the facility, and finalized interviewing arrangements. The

facility contact(s) and the project negotiator had also agreed to scheduling, sampling, and security arrangements in advance.

In instances where the sampled inmates could not participate in the interview and assessment due to circumstances such as solitary confinement or administrative segregation, hospitalization, having a court date or a work assignment, the interviewers were instructed to use prison-specific disposition codes to indicate the reason that the interview could not be completed.

5.7.2 Negotiating with Authorities and Enrolling Facilities

The process of contacting and enrolling the selected facilities began in July 2013. The project negotiator conducted several activities to obtain the cooperation of Federal and state authorities. These steps included the following:

- An introductory package including a letter, study brochure, and endorsement letters was sent to a previously identified contact at the State Department of Corrections or Federal agency for which there is a selected facility.

- Follow-up calls were conducted with Federal and state officials to obtain study endorsement and the name of an appropriate facility contact for each sampled prison under their purview.

- During the state-level negotiations, the state/Federal contact was asked if he or she preferred to contact the sampled prisons directly or if the project negotiator should do so. If the state/Federal authority agreed to contact the facility, the project negotiator followed up by phone with the facility to confirm participation. If the project negotiator was responsible for contacting the facility, he or she sent the facility an introductory package.

- After the state contact reviewed the introductory package and agreed to the state's participation, negotiations continued with the warden and key facility staff members to obtain facility participation.

- Once prison participation was granted by the warden, the project negotiator contacted the facility to contact to:

  - develop facility-specific access procedures;

  - specify the facility sampling procedures;

  - establish contacts at the facility, including the identification of a prison coordinator;

  - review security and confidentiality procedures;

  - provide materials for Institutional Review Board, if required;

- establish specific interviewing arrangements and schedules; and
- arrange for any other special considerations required by the facility.

5.8 Quality Control

5.8.1 Quality Control Measures

Quality control was an integral component to the overall success of the study. The fieldwork quality control measures listed below were implemented to ensure that the data collected met high-quality standards.

- **Validation.** Validation ensures that interviews are actually conducted with the sampled person in the sampled households and that cases are assigned the correct final disposition code. All interviewers working on the household data collection had at least 10 percent of their finalized work validated, which included completes, ineligibles, and nonresponse cases. All cases were selected randomly as they were being finalized.[30] The highest producing interviewers were validated at a rate of 15 percent. When the validation process uncovered falsification, all cases worked by these interviewers were validated to check for additional falsification. Nine Main Study interviewers and four National Supplement interviewers were immediately dismissed.[31] Falsified cases were refielded or flagged as appropriate.

- **Recording of Interviews.** High-quality survey research also requires that the interview be conducted in a professional and ethical manner and that interviewers be given a chance to improve their interviewing techniques. Interviewers were required to record their third and tenth interviews. These recordings were reviewed by supervisors who then provided feedback to interviewers about any aspect of their interviewing technique that needed improvement. Interviews conducted in prisons were not permitted to be recorded; however, since prison interviewers were experienced household interviewers, they had previously conducted recordings of their household interviews and received feedback as necessary.

- **Falsification Detection Reports.** Several automated reports were developed to monitor various aspects of the household interviewer behavior that may be indicative of falsification. Reports to monitor the following aspects were developed:
  - interview start time (too early/too late in the day);
  - elapsed time between interviews;
  - BQ administration length;

[30] Nonresponse also affects validation efforts. Substituted cases were sometimes used when it was necessary to reach the 10 percent threshold for an interviewer.

[31] One falsifying interviewer was discovered only after she was released from the project due to a lack of production and therefore could not be fired. This interviewer was, however, removed from the list of potential candidates to work on any other Westat survey.

- multiple interviews completed per day;
- BQ and assessment completed on different dates; and
- proportion of item nonresponse in BQ and in assessment results submitted by interviewer.

- **Review of Completed Case Materials.** Interviewers were required to return completed case materials (case folders, paper exercise booklets, etc.) to the home office on a regular basis. Upon receipt, these materials were reviewed for completeness and accuracy, and feedback was given to field interviewers if necessary.

Quality control was also an integral component of the prison data collection. However, individual validation was not necessary because interviewers worked in pairs and also worked closely with the staff appointed by the warden at each facility. Instead, the quality control of the prison sample focused on the review of the completed case materials described above. Likewise, the falsification detection reports described above were not prepared because they are only relevant to household data collection.

5.8.2 CAPI Helpdesk

The PIAAC CAPI helpdesk was maintained and operated by staff specially trained on the PIAAC instrument and who were well versed with the workflow, instrumentation, and the systems aspect of the Main Study and National Supplement. Field interviewers and supervisors who experienced technical difficulties while administering the interviews or completing systems-related tasks called the helpdesk via a toll-free number for assistance; problems that could not be resolved by helpdesk staff were passed on to PIAAC IT staff for further direction on resolution. The helpdesk staff reported receiving 1,704 calls throughout the period of the Main Study data collection. During the period of data collection for the National Supplement there were 830 calls to the helpdesk.

This page intentionally blank.

6. REDUCING THE RISK OF DATA DISCLOSURE

Over the past decade, concerns about the disclosure of information related to individual survey respondents have increased dramatically. Laws have been passed since the Privacy Act of 1974 to further ensure the protection of confidential data. The most recent of these is the Education Sciences Reform Act of 2002, which explicitly requires that the National Center for Education Statistics (NCES) protect the confidentiality of all those responding to NCES-sponsored surveys. More specifically, NCES Standard 4-2, "Maintaining Confidentiality," provides guidelines for limiting the risk of data disclosure for data released by NCES.[32] This chapter describes the procedures for controlling statistical disclosure for PIAAC in accordance with the guidelines specified in NCES Standard 4-2.

6.1 Overview of PIAAC Data Collection and Dissemination

Several types of data were collected and derived during the PIAAC sampling, data collection, and weighting processes. These variables were reviewed to determine their disclosure risk levels. The confidentiality analysis used a three-step process to reduce disclosure risk: (1) determining the disclosure risk arising from existing external data, (2) coarsening the data, and (3) swapping the data. Westat conducted the risk analysis, coarsening, and data swapping procedures to produce the following files, included in the PIAAC data dissemination products:

- International Public-Use File (PUF)—to include international variables only. This file contains the Main Study sample only and was provided for international comparisons in Round 1 of PIAAC;
- International PIAAC Data Explorer (developed by ETS) and International Database Analyzer (developed by IEA) data tools;
- U.S. PUF for the Main Study and National Supplement household sample—includes both international and U.S. variables;
- U.S. PUF for the prison sample—includes both international and U.S. variables;
- U.S. Restricted-Use File (RUF) for the Main Study and National Supplement household sample—includes both international and U.S. variables; and
- U.S. RUF for the prison sample—includes both international and U.S. variables.

Note that the PUF and RUF for the prison sample contain prison-specific U.S. variables. Some of the general U.S. variables were not included in the prison sample files since they were not applicable to the prison population. The U.S. PUF and RUF for the Main Study and National Supplement household sample can be used for international comparisons while the U.S. PUF and RUF for the prison sample cannot, since the international PIAAC sample excluded the prison population.

Following the NCES guidelines, the RUF contains noncoarsened, swapped data and the PUF contains coarsened, swapped data. In addition, a data tool created for PIAAC followed the confidentiality

[32] Available at http://nces.ed.gov/statprog/2012/pdf/Chapter4.pdf.

procedures established for disseminating data via data tools. The data swapping was done in one step for the RUF and the PUF, thereby ensuring consistency between detailed and coarsened variables.

The approach for the PUFs included categorizing several critical variables with some risk of data disclosure and suppressing several variables with a very high risk of disclosure. These actions were based on extensive initial disclosure risk analyses, which included the following steps:

- identifying personal identifiers, geographic information, and contextual variables (variables that can indirectly identify a geographic area);
- evaluating the existence of other publicly available files;
- evaluating the disclosure risk associated with release of the sampling and variance estimation variables; and
- evaluating the disclosure risk associated with release of key variables (i.e., visible variables) through extensive frequency tables.

One aspect of the disclosure risk analysis for the PUF was a review of each Background Questionnaire (BQ) variable and groups of BQ variables to determine whether any of the data presented a nonnegligible risk of individual disclosure. Several types of variables were available, including variables collected through the survey and assessments as well as variables created during weighting. These variables are summarized below.

The following are the main sources of the variables considered for dissemination for public use:

- Sample Design International File;
- case initialization data from the Screener;
- BQ data;
- cognitive data from the assessment; and
- Weighting International File.

Careful attention was given to the BQ items and combinations of items. Even a very limited amount of demographic detail—such as income, occupation, age, year of immigration to the United States, foreign language spoken, and country of birth—can increase the chance that an individual can be identified. As discussed in section 6.2, personal and geographic identifiers were removed. Section 6.3 presents outcomes from the risk analysis in the form of variable suppression and recodes. Section 6.4 summarizes the final disclosure reduction step, data perturbation in the form of swapping.

6.2 Personal Identifiers and Geographic Identifiers

Any information that might be used to directly identify sample persons and/or sample locations was suppressed from the PUF and RUF. This information included direct personal identifiers such as names, addresses, and telephone numbers. The inclusion of any geographic detail has a large impact on the level of disclosure risk. Review of the available variables indicated that there were several geographic

variables, such as the primary sampling unit (PSU) and segment identifiers embedded in the in-house person ID, county, and zip code. Thus, the person ID included on the PUF and RUF was randomly generated with no linkages to clustering or other geographic or sampling information embedded in the original person ID. Only census region (four levels) and NCES Urban-Centric Locale Codes (four levels)[33] were retained on the PUF; all other geographic indicators were suppressed. The PSU identifier, segment identifier, and a 12-level NCES Urban-Centric Locale Codes were included in the RUF.

6.3 Data Coarsening

In general, data coarsening includes several types of procedures that decrease disclosure risk by reducing the amount of information released. Coarsening approaches include removing direct identifiers, limiting geographic detail, categorizing continuous variables, performing top- and bottom-coding,[34] and recoding values into broader categories. Targeted or local suppression is also performed by removing the sensitive item's value from the record or suppressing the variable from the file. After a thorough disclosure risk analysis had been conducted, the results of frequencies and multiway cross-tabulations were used to guide the coarsening process. First, one-way tabulations were reviewed to determine the categorical variables that would need to be recoded because of the small number of responses in one or more categories. Next, multiway tabulations were reviewed to identify variables that had problematic categories when used in combination with other variables. The following paragraphs further describe the various statistical disclosure control treatments.

Suppressed Variables. A number of variables were suppressed due to low frequency counts (i.e., frequencies of less than 15) because they revealed too much detail about the sample design units and geographical location or because they were used to derive coding variables.

The PSU ID variable and other variables related to the sample design units were suppressed in response to disclosure concerns about being able to indirectly identify the location of the respondents. The variance strata and variance unit variables were retained on the PUF to facilitate variance estimation.

All open-ended variables (e.g., "other, specify" responses) were also suppressed from the PUFs to prevent the possibility of revealing geography or an individual identity from the responses.

Recoded Variables. The process of recoding categorical variables helps to minimize the risk of data disclosure. Candidates for suppression included variables for which adequate protection was not possible without losing the meaning and usefulness of the data.

The process for recoding categorical variables involved grouping different levels of the same variable to create categories with larger frequencies. As a guideline, categories that comprised less than 0.5 percent of the sample were grouped with others. This reduced the risk of identifying a respondent when an attack consists of combining survey variables.

Continuous Variables. Top-coding was performed for some continuous variables. One approach considered was to categorize all continuous variables, since, with top-coding, there is still potential for bias on computations of the average for subgroups or for regression analysis. Categorization protects

[33] The four levels of census region include Northeast, Midwest, South, and West. The four levels of NCES Urban-Centric Locale Codes include city, suburban, town, and rural.

[34] With top-coding, the largest values of a variable are replaced with an upper limit, reducing the appearance of outlier data. Similarly, bottom-coding replaces the smallest values with a lower limit.

against bias and reduces disclosure risk. However, it was decided to use top-code cutoffs for these continuous variables because the analytical value of continuous versions of some variables outweighs the potential disclosure risk. The criterion was that the proportion of cases with values greater than the cutoff was at least 0.5 percent of the weighted number of persons; most variables had proportions between 1 percent and 5 percent for extra protection. The top-coding cutoff was used as the replacement value for the cases with values greater than the cutoff.

While top-coding applies to some continuous variables, the following continuous variables were recoded to be categorical:

- B_D01D (Education – Time elapsed since finished highest qualification);
- B_D03D (Education – Derived Months elapsed since leaving education without completing program);
- B_Q01C1/B_Q01C1US (Education – Highest qualification – Age of finish);
- B_Q03C1 (Education – Uncompleted qualification – Age of dropout);
- B_Q12F (Activities – Last year – On the job training – How many);
- B_Q20A (Activities – Last year – Time spent on activities – Hours);
- C_Q08c1 (Current status/work history – Age when stopped working in last job);
- D_Q05A1 (Current work – Start of work for employer – Age);
- D_Q05B1 (Current work – Start of work for business – Age);
- D_Q07b (Current work – Employees working for you – Count);
- E_Q05A1/E_Q05A1USP (Last job – Start of work for employer – Age);
- E_Q05B1/E_Q05B1USP (Last job – Start of work for business – Age);
- IMYRS (Years in country);
- J_Q03C (Background – Age of the child);
- J_Q03d1 (Background – Age of the youngest child);
- J_Q03d2 (Background – Age of the oldest child);
- J_Q04c1 (Background – Age at time of immigration); and
- LEAVEDU (Respondent's age when leaving formal education).

Derived Race/Ethnicity. The logic shown in table 6-1 was used to create a derived race/ethnicity variable. The RUF included a race/ethnicity variable with more detailed categories which was deemed too risky for the PUF. Hence, the race/ethnicity variable shown in table 6-1 and another one with less detailed categories were derived and included in the PUF.

Table 6-1. Derived variable for race/ethnicity

Variable name	Value	Value label	Derivation
RACETHN_5CAT	1	Hispanic	If J_Q04dUSX1a = 1 then RACETHN_5CAT = 1;
	2	White	Else if J_Q04dUSX1a = 2 and J_Q04dUSX2_01 = 1 and J_Q04dUSX2_02 ne 1 and J_Q04dUSX2_03 ne 1 and J_Q04dUSX2_04 ne 1 and J_Q04dUSX2_05 ne 1 then RACETHN_5CAT = 2;
	3	Black	Else if J_Q04dUSX1a = 2 and J_Q04dUSX2_02 = 1 and J_Q04dUSX2_01 ne 1 and J_Q04dUSX2_03 ne 1 and J_Q04dUSX2_04 ne 1 and J_Q04dUSX2_05 ne 1 then RACETHN_5CAT = 3;
	4	Asian/Pacific Islander	Else if J_Q04dUSX1a = 2 and (J_Q04dUSX2_03 = 1 or J_Q04dUSX2_05 = 1) and J_Q04dUSX2_01 ne 1 and J_Q04dUSX2_02 ne 1 and J_Q04dUSX2_04 ne 1 then RACETHN_5CAT = 4;
	6	Other	Else if J_Q04dUSX1a = 2 and (J_Q04dUSX2_01 in (1,2) or J_Q04dUSX2_02 in (1,2) or J_Q04dUSX2_03 in (1,2) or J_Q04dUSX2_04 in (1,2) or J_Q04dUSX2_05 in (1,2)) then RACETHN_5CAT = 6;

NOTE: Variable names match those provided in the BQ available at http://nces.ed.gov/surveys/piaac/questionnaire.asp.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012.

6.4 Swapping

To ensure that a data intruder could never be sure that the individual is identified, the risk of data disclosure was further reduced on both the PUF and RUF by using the data swapping technique requested by NCES. Data swapping is an NCES requirement that reduces risk by modifying microdata. In data swapping, a probability sampling of records are paired with other records on the file using selected characteristics, and then some identifying variables are swapped between the two records. The sampling rate for PIAAC swapping was designed to protect the confidentiality of the data without affecting the usability of the dataset. This method is an effective way of keeping as much valuable data as possible while not identifying any research participants.

The standard swapping software that has been approved for use on NCES studies was applied to the household sample and prison sample to (1) select target records to be swapped, (2) select swapping partners, and (3) swap the data. Swapping preserves the unweighted frequencies, means, and variances; however, it may affect the weighted distributions and multivariate relationships. The NCES software selected swapping partners with the smallest absolute bias pertaining to a variable of interest. As a check, analysts reviewed pre- and postswapping percentage distributions (unweighted and weighted), as well as correlation analyses and regression models, to examine the relationships between the swapped variables and the key variables, such as educational attainment and computer experience.

This page intentionally blank.

7. NONRESPONSE BIAS ANALYSIS (NRBA)

Under ideal situations, every eligible adult in the target population would have a nonzero chance of selection in a national sample, would be located, and would agree to participate in the study. In practice, these circumstances are not realized in any survey population. Response rate is a valuable data quality measure and the most widely used indicator of survey quality. A high response rate increases the likelihood that the survey accurately represents the target population, and a low response rate reflects the possibility of bias in the outcome statistics.

It is well understood that when response rates are low, there is a greater chance for nonresponse bias. The extent of nonresponse bias depends on many survey conditions, including the differential impact that the likelihood of response has on the bias of each of the survey outcomes. It is, therefore, critical to evaluate the potential for nonresponse bias as a quality check on the estimates at the conclusion of the data collection.

There are several ways to reduce the potential for nonresponse bias, including planning and implementing field procedures that obtain a high level of cooperation, and monitoring the distribution of the sample during data collection to ensure steps are taken to reduce the potential for bias as much as possible. Further, if nonresponse rates increase, one needs to actively seek auxiliary data (e.g., age or gender of all sampled persons) to reduce the impact of response propensities on the survey estimates. These auxiliary variables can then be used in weighting adjustments for the purpose of reducing nonresponse bias.

Although sample weight adjustments based on auxiliary data are effective in reducing nonresponse bias, they are not considered as replacements for a vigorous effort to achieve the highest response rate possible.

7.1 Standards and Procedures Relating to NRBA

The PIAAC estimates of literacy-related skills in the United States are subject to potential bias due to nonresponse at various levels of data collection. As mandated by the PIAAC Consortium, the U.S. National Center for Education Statistics (NCES), and the U.S. Office of Management of Budget (OMB), a Nonresponse Bias Analysis (NRBA) is required if response rates are below the guidelines shown in table 7-1.

Table 7-1. Cut-point on response rates requiring a nonresponse bias analysis, by component and organization

Component	(Percent) PIAAC Consortium	NCES	OMB
Data collection phase unit	80	85	†
Overall unit	70	†	80
Item	85	85	70

† Not applicable.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

This section provides the results of a systematic analysis of bias resulting from some selected households or individuals choosing not to participate or from respondents failing to complete all questionnaire items. The analysis is in accordance with both NCES Standard 4.4 and PIAAC TS&G 4.7.

PIAAC had three stages of data collection where unit nonresponse occurred: the Screener, the Background Questionnaire (BQ), and the assessment components. Participation in the BQ was dependent upon the completion of the Screener. Likewise, participation in the assessment was dependent upon completion of the BQ. The final weighted response rates for each data collection stage are shown in table 7-2, including the overall response rate for the survey.

Table 7-2. PIAAC Main Study and National Supplement area sample and List Sample response rates

		National Supplement		Main Study and
	Main Study (percent)	Area sample (percent)	List sample (percent)	National Supplement combined (percent)
Screener	86.5	81.4	84.8	84.7
BQ	82.2	78.1	92.9	80.9
Assessment	99.0	98.5	98.8	98.8
Overall	70.3	62.6	77.9	67.8

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

The assessment response rates were all above the 85 percent response rate requirement. Therefore, the unit-level analysis focused on the Screener and BQ stages. The key subgroups for the analysis are defined in section 7.2.

The basic NRBA was not required for the Prison Study because the weighted response rates for all data collection stages were above the 85 percent response rate requirement (see table 7-3).

Table 7-3. PIAAC National Supplement Prison Study response rates

	Without substitution (percent)	With substitution (percent)
Prison	97.8	100.0
BQ	85.7	85.7
Assessment	97.4	97.4
Overall	81.6	83.5

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

7.1.1 An Extended Analysis

The basic descriptive analysis required by Standard 4.7.5 provides an initial assessment of nonresponse bias (NRB) and is essential in identifying effective weighting variables. However, it has its limitations. The analysis relies on auxiliary variables and does not evaluate the effectiveness of weighting adjustments on reducing NRB. The descriptive analysis assesses the potential for bias in some statistics that may not necessarily be highly related to the final proficiency estimates. The goal of the extended analysis is to assess the effect of weighting adjustments by assessing remaining bias in the key statistics or closely

related variables. For PIAAC, the key statistics are scores measuring proficiency in the components of the assessments.

Multiple analyses to assess NRB are necessary because each analysis has its own limitations. Together, they provide an insight into the patterns and potential for bias. The components of the extended analyses are as follows:

- comparison of estimates before and after weighting adjustments;
- comparison of weighted estimates to external totals;
- correlations of auxiliary variables and proficiency estimates;
- comparison of estimates from alternative weighting adjustments; and
- level-of-effort analysis.

More detail on and the results of each of the extended analyses is found in section 7.3.

7.1.2 Classification of Nonresponse

There are two classes of nonresponse to the BQ: (1) literacy-related reasons and (2) nonliteracy-related reasons. Literacy-related nonrespondents (LRNR) includes persons who were nonrespondents because of insufficient reading or writing skills, insufficient fluency in English or Spanish, or a learning or mental disability that prevented them from responding to the questionnaire. For the Main Study, the BQ response rate is 82 percent if treating the BQ LRNR cases as respondents and 80 percent if treating them as nonrespondents. For the National Supplement, the preliminary BQ response rates are 78 percent (area sample) and 93 percent (list sample) if treating the BQ LRNR cases as respondents and 76 percent (area sample) and 93 percent (list sample) if treating them as nonrespondents. In this nonresponse bias analysis (NRBA), the BQ LRNR cases are sometimes included and sometimes excluded. If included, they are treated as respondents. These BQ LRNR cases were included in the weight calibration process and received nonzero final weights because of their nonignorable reasons for nonresponse; however, plausible values for literacy scores were not generated due to the lack of available information. A national estimate of BQ LRNR cases in the population can be generated, if desired. The weights for the BQ LRNR account for the LRNR at the Screener stage. Those who completed the BQ but did not complete the assessment due to literacy-related reasons are referred to as assessment LRNR. Plausible values were generated for assessment LRNR cases by assuming they would have incorrect responses to the test items. The BQ LRNR cases were excluded from the basic NRBA in section 7.2 because the analysis focused on respondent and nonrespondent cases assumed to be similar (e.g., would have scored at the same level). The BQ LRNR cases are included in selected analysis, where appropriate.

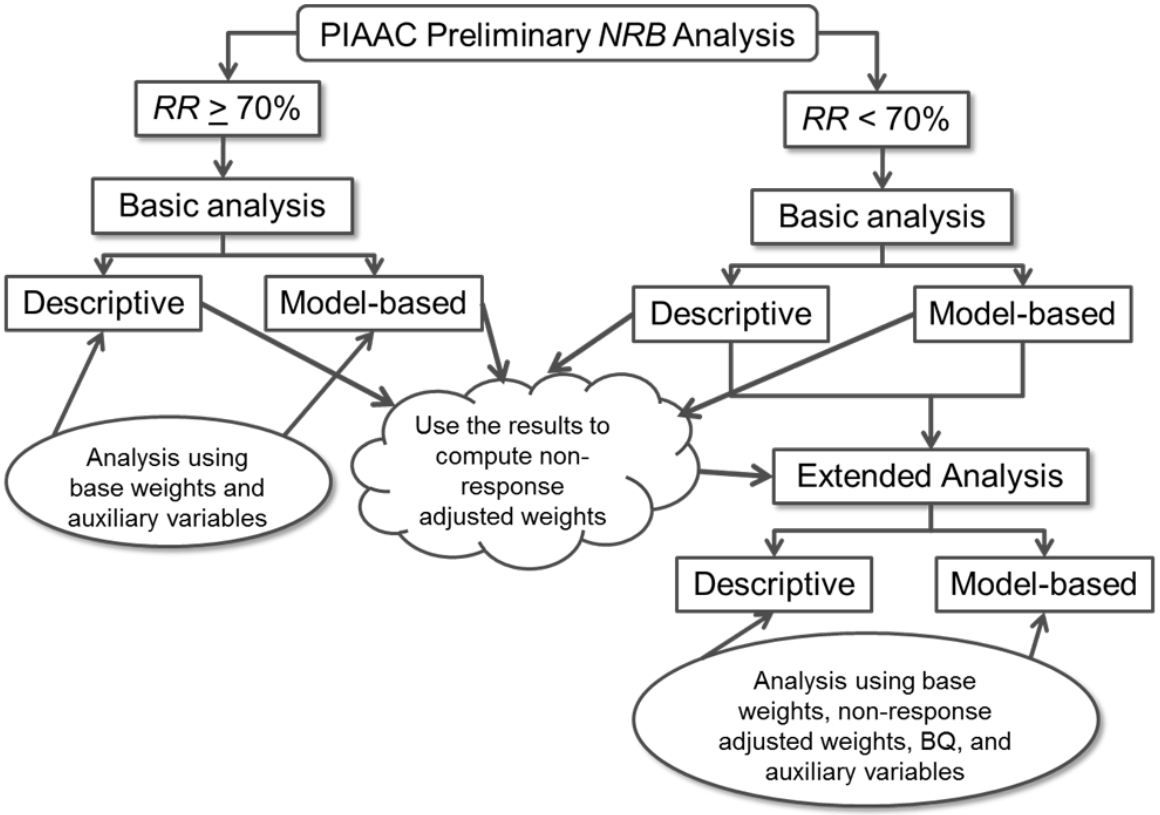
At the BQ stage, nonliteracy related nonrespondents for the NRBA include the following: partial complete/breakoff, refusal, speech impairment, maximum number of calls, temporarily absent or unavailable during the field period, technical problem, hearing impairment, blindness or visual impairment, physical disability, and other disability.

7.1.3 Procedures Used to Reduce Bias Due to Nonresponse

There were several ways to reduce the potential for NRB in PIAAC. First and foremost was implementing field procedures that supported the response rate goal. Second was the careful monitoring of data collection in the field to monitor response rates and take steps to reduce bias as much as possible. Finally, an extensive search for auxiliary data was conducted so that these variables could be used to evaluate the potential for bias in order to reduce the impact of nonresponse on survey estimates during the weighting process.

Figure 7-1 (extracted from the PIAAC Consortium document "PIAAC Reducing Nonresponse Bias and Preliminary Nonresponse Bias Analysis" dated March 10, 2010) provides a schematic of the processing after data collection. The preliminary NRBA occurs during the data collection period and sample monitoring process. The descriptive and model-based (multivariate) basic analyses allow the variables to be selected for weighting, and the weighting process helps to adjust for differential nonresponse across subgroups. An extensive analysis is then used to further investigate the potential for nonresponse after weighting procedures are conducted.

Figure 7-1. Postdata collection activities relating to PIAAC nonresponse bias analysis



SOURCE: OECD PIAAC document "PIAAC Reducing Nonresponse Bias and Preliminary Nonresponse Bias Analysis" dated March 10, 2010.

7.2 Results of the Basic NRBA

This section presents weighted response rates and a set of basic analyses of the potential for nonresponse bias conducted to arrive at the set of variables used in weighting adjustments. The analyses focus on identifying variables that are highly correlated with nonresponse and with the outcome statistics and closely follow the guidelines provided for the basic NRBA by the PIAAC Consortium. Section 7.3 provides a more extensive NRBA focused on evaluating the potential for bias remaining after weighting adjustments are conducted, using the outcome statistic (proficiency scores) following the NCES standards.

Total survey error has two components: variable error (measured through the calculation of variances) and bias. The variance is the first term in the following equation for total survey error in a survey estimate:

$$Total\ survey\ error = variance + bias^2.$$

Bias, the second term in the equation, contains all sources of error other than variable error. A major contributor to bias is nonresponse, that is, the bias owing to the failure of some selected persons in the sample to respond to the survey. Nonresponse bias can be substantial when two conditions hold: (1) the response rate is relatively low, and (2) the difference between the characteristics of respondents and nonrespondents is relatively large.

An estimate for nonresponse bias, assuming that nonresponse is the only source of bias, is expressed in Cochran (1977) as

$$Bias(\bar{y}_R) = (1 - W_R)(\bar{Y}_R - \bar{Y}_N),$$

where W_R is the response rate and $\bar{Y}_R$ and $\bar{Y}_N$ are the mean values of the survey items estimated among the respondents and nonrespondents, respectively. Thus, the estimates from any survey are subject to bias when some selected persons fail to participate in the survey. Because we do not have survey values for nonrespondents, nonresponse bias is not known and can only be estimated.

An alternative model of nonresponse assumes each sampled person has a certain propensity to respond, and NRB in a characteristic is a function of the covariance between the response propensity and the characteristic:

$$Bias(\bar{y}_R) = \frac{\sigma_{yp}}{\bar{p}},$$

where σ_{yp} is the covariance between the outcome variable and response propensity, and $\bar{p}$ is the mean response propensity. Based on this model, NRB is present if a missing response is related to competency, as measured in PIAAC.

The following sections provide insights into the effects of nonresponse on U.S. PIAAC. Unweighted response rates are indicators of the success of the data collection effort. Since weighted response rates are more appropriate in examining the potential effect of nonresponse on population parameters, these are provided in addition to bivariate and multivariate analyses of the potential for nonresponse bias.

Data from respondents were collected through a Screener, a BQ, and an assessment. In nonresponse follow-up, efforts were made to reduce the potential for nonresponse bias by targeting interviewer

resources in areas with low response rates. To identify target areas, a multivariate analysis was conducted using the classification software package Search[35] for the initial nonresponse bias analyses at both the Screener and BQ levels. In the Main Study, the resulting classification tree revealed the domains, as defined by combinations of variables, with the most differential response rates, thereby leading to domains with a high potential for nonresponse bias. Overall, the results of the analysis showed acceptable response rates for most of the cells identified by the program. The analysis, which was conducted for both the Screener and the combined BQ/assessment response rates, identified the PSUs that included the domains with less than a 70 percent response rate. To identify target areas in the National Supplement, three groups were monitored throughout the data collection period: unemployed 16- to 65-year-olds, other 16- to 34-year-olds, and 66- to 74-year-olds. Field activities and resources were focused on these PSUs and areas in the remaining weeks of the data collection.

After data had been collected and weights produced, an analysis was conducted to examine the impact of bias owing to the remaining nonresponding dwelling units (DUs) and persons in the household sample. This analysis used weights created at varying stages of the weighting process. Screener and person-level base weights were used for the calculation of response rates; unknown-eligibility-adjusted weights were used for the other analyses. The analysis is divided into three pieces: first, a discussion of the weighted response rates for the Screener and the BQ; second, the results of the basic nonresponse bias analyses; and third, a summary of the potential for nonresponse bias prior to the weighting adjustments. As mentioned in section 7.2.1, the BQ LRNR cases are excluded from the basic analyses mentioned in section 7.2.2 since the focus is on analyzing nonrespondents and respondents whom we assume would score at the same level.

For the basic NRBA, the three samples (Main Study; National Supplement: area sample; and National Supplement: List Sample) were analyzed separately to inform the separate nonresponse weighting adjustments for each sample. The basic NRBA for each sample was conducted prior to weighting to identify potential nonresponse bias and to select variables for weighting adjustments that would be most effective in reducing the bias. Each sample was weighted for nonresponse and then combined into one sample with final weights calibrated to control totals from the American Community Survey (ACS). The extended NRBA was performed using the single combined sample and the final sample weights.

7.2.1 Response Rates

U.S. PIAAC had two stages of data collection in which appreciable unit nonresponse occurred: the Screener and the BQ. Weighted response rates were computed for each stage. Screener base weights were used in the Screener response rate calculations, and BQ base weights were used for the BQ calculations.

Response rates were calculated following PIAAC Consortium guidelines found in section 4 of the *Technical Standards and Guidelines*.

Table A-1 in appendix A contains response rates for the Screener and the BQ, conditional and unconditional on Screener response, for both the Main Study and National Supplement (area sample and list sample). The results clearly show differential response rates by subgroups. The regional and metropolitan response rates for PIAAC follow the usual patterns found in U.S. surveys. Lower response rates were experienced in the very urban areas that are more prevalent in the Northeastern and Western

[35] Details on the history and development of the software and also some references to early evaluations may be found at http://www.isr.umich.edu/src/smp/search. Details of some computations in Search are provided in Sonquist, Baker, and Morgan (1974).

regions of the country. (One exception was the BQ response rate for the list sample in the National Supplement, which was higher in PSUs that were part of a Metropolitan Statistical Area.) Areas with fewer people born outside the country and/or areas with fewer linguistically isolated people are likely rural areas that typically have higher response rates.

Additionally, areas with higher educational attainment had lower response rates.

For the Main Study, youth and women were more likely to respond. Results for the area sample and list sample were not as straightforward. Among the area sample, those age 35–45 had the highest response rate on BQ, while the opposite was observed in the list sample, for which response rates were the lowest in the 35–45 age group.

However, list sample response rates were higher than area sample response rates in all other age categories. Overall, there was a higher response rate in the list sample; this could be due to the nature of the list sample, which was comprised of unemployed adults. These unemployed persons could have been more available to participate in the survey and could have been more attracted to the monetary incentive. This argument is supported by the fact that in the area sample unemployed adults were more likely than others to respond to the BQ.

A $5 incentive was offered to each responding household in the area and list samples in the National Supplement in order to screen for the subgroups of interest. For all three samples, following the completion of the assessment, a monetary incentive of $50 was paid to each respondent. The incentive was also paid to those adults who attempted to complete an assessment but were legitimately unable to do so because of language barriers, insufficient literacy skills, or disability. Respondents who refused to continue with the assessment were not compensated. Given this incentive, groups with lower income and lower educational attainment according to the most recent census were more likely to respond. This includes the Hispanic population, which are more likely to live in poverty, according to the U.S. Census.[36]

7.2.2 Nonresponse Bias Analysis

To determine the set of variables for the nonresponse bias analysis for the Main Study, a regression tree and regression model were created using data from the U.S. PIAAC field test. The same analysis variables were then used for the National Supplement. Standardized logit scores (similar to proficiency scores) were used as the dependent variable. The predictors were extracted from the following data sources: 2000 Census, ACS 2005–2009 tract-level[37] data, and the Screener (for BQ-level analysis only). The continuous variables from the ACS were recoded into categories of approximately equal sample size. Data from the 2003 National Assessment of Adult Literacy (NAAL) and 2003 Adult Literacy and Lifeskills (ALL) survey were also reviewed to identify characteristics related to literacy skills.

[36] U.S. Census Bureau, Income, Poverty, and Health Insurance Coverage in the United States: 2009, Current Population Reports, P60-238, and Historical tables—table 3, September 2010. See also http://www.census.gov/hhes/www/poverty/poverty.html.

[37] Blocks are very fine partitions of the United States, formed using visible semipermanent features such as roads, railroad tracks, mountain ridges, bodies of water, and power lines. The only invisible boundaries used are county, state, and national boundaries. Minor civil division boundaries and property lines are ignored. A block group is a small group of contiguous blocks. A tract is a collection of contiguous block groups all within the same county.

The variables that were closely related to literacy skills are listed below. These variables from the 2000 Census were available on the PSU sampling frame. The list notes which variables were used only in weighting the Screener.

- census region; and
- indicator of whether the PSU is part of a Metropolitan Statistical Area Region (only used in Screener weighting).

Estimated quartiles of tract-level data from the ACS 2005–2009 for the Main Study and 2008–2012 for the National Supplement were computed for the following:

- percentage of housing units occupied by owner;
- percentage of the population age 25 and older with at least a high school education;
- percentage of the population that is Hispanic or Non-Hispanic Black;
- percentage of the population that is Hispanic (only used in Screener weighting);
- percentage of the population age 18–64 that is unemployed;
- percentage of the population below 150 percent of poverty (only used in Screener weighting);
- percentage of the population that is foreign born;
- percentage of households that are linguistically isolated;
- percentage of the population age 18–64 that is employed;
- percentage of the population age 25 and older with a high school education;
- percentage of the population age 25 and older with some college education; and
- categorized average household size.

The following characteristics were available from the Screener:

- age category (after imputation);[38]
- gender;
- race/ethnicity (after imputation);[39]

[38] For the Main Study, age was imputed for some 0.2 percent of sampled persons; race/ethnicity for 2.1 percent. For the National Supplement, no imputation of age was needed; race/ethnicity was imputed for 1.9 percent of sampled persons.
[39] Ibid.

- language used at screening (not used during the BQ weighting stage); and
- indicator for children under age 16 in household.

For the area sample in the National Supplement, the sampling domain indicator (unemployed age 16 to 65, not unemployed age 16 to 34, and age 66 to74) was also included as an analysis variable.

The basic nonresponse bias analysis for both the Screener and the BQ is presented in the sections that follow.

7.2.2.1 Evaluating Bias Due to Screener Nonresponse

A comparison of Screener respondents and nonrespondents using variables known for both groups provides some indication of the potential for nonresponse bias in survey estimates in the absence of weighting adjustments. A discussion of the chi-square tests that may detect a significant relationship between the response indicator and the analysis variable of interest is found in section 7.2.2.2. Section 7.2.2.3 provides a multivariate analysis of the relationship between the response indicator and analysis variables that may reveal the areas with the greatest potential for bias before weighting adjustments occurred.

7.2.2.2 Screener Bivariate Analysis

The distribution of Screener respondents was compared with the distribution of all eligible sampled DUs for the variables from the PSU sampling frame and the ACS. For the Main Study, weighted percentages and standard errors (SEs) were calculated using the WesVar software (Westat 2007) and full-sample and replicated Screener unknown-eligibility-adjusted base weights. For the National Supplement, weighted percentages and SEs were computed using the SURVEYFREQ procedure in SAS and variance strata and variance units in order to account for the complex sample design. The full-sample weights were adjusted for unknown-eligibility status. To test the significance of the relationship between response status and each of the variables, a Rao-Scott chi-square (RS3) test of independence (Rao and Scott 1984) was performed.

The results of the Rao-Scott chi-square analyses for the three samples are shown in tables A-2a through A-2c in appendix A. Table 7-4 summarizes the results of the significance tests for each sample. Variables in the Screener analysis that were significant at the $\alpha = 0.05$ level are indicated by an X. The difference between "percentage of the population age 25 and older with at least a high school education" and "percentage of the population age 25 and older with a high school education" is that the latter (high school and no more) is the percentage of a subset of the former.

These analyses confirmed what was seen in the response rate analysis above. The percentage of the population age 25 and older with at least a high school education was significant for all three samples. Unlike results from the Main Study and the area sample, in the list sample, the nonresponse rate was much higher in areas with a lower percentage of Hispanic or Non-Hispanic Black and with a higher percentage of employed adults. Respondents and nonrespondents were not significantly different in terms of regional and income-related characteristics in the list sample of unemployed adults (age 16 to 65).

Table 7-4. Summary of results from Screener bivariate analysis

Analysis variables	Main Study	National Supplement	
		Area sample	List sample
Indicator of whether the PSU is part of a Metropolitan Statistical Area		X	
Census region	X	X	
Percentage of housing units occupied by owner			
Percentage of the population age 25 and older with at least a high school education	X	X	X
Percentage of the population that is Hispanic or Non-Hispanic Black		X	X
Percentage of the population that is Hispanic	X	X	
Percentage of the population age 18–64 that is unemployed			
Percentage of the population below 150 percent of poverty	X	X	
Percentage of the population that is foreign born	X	X	
Percentage of households that are linguistically isolated	X	X	
Categorized average household size			
Percentage of the population age 18–64 that is employed			X
Percentage of the population age 25 and older with a high school education	X	X	
Percentage of the population age 25 and older with some college education			

NOTE: Variables that were indicated by an X were significant (at the $\alpha = 0.05$ level) in the bivariate nonresponse bias analysis for Screener.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

7.2.2.3 Screener Multivariate Analysis

The bivariate analysis described above is useful in explaining each variable individually. A multivariate analysis is useful in showing relationships among a number of variables. One approach is to provide a classification tree, which divides a sample into subgroups that best explain differential response rates.

The classification software package, Search, which employs a hierarchical tree algorithm, was used for the initial nonresponse bias analyses at both the Screener and BQ levels. (Search is a freeware program developed and maintained by the University of Michigan.) Cell sizes were limited to 30 or more cases in each analysis. The resulting classification tree reveals the domains, as defined by combinations of variables, with the most differential response rates, thereby leading to domains with the highest potential for nonresponse bias.

Search was run with Screener response status as the dependent variable and the variables from the PSU sampling frame and the ACS as the independent variables.

Main Study

The classification tree for the Screener is summarized in table A-3a in appendix A. Twenty-six cells were formed with weighted response rates ranging from 75.7 percent to 97.0 percent. The lowest response rate was for the group within more rural areas that had each of the following characteristics:

- more than 2.6 percent of the population being foreign born; and
- less than 19.4 percent of the population below 150 percent of the poverty level.

The percentage of foreign born in the segment was the dominant variable in distinguishing response rate groups.

National Supplement: Area Sample

The classification tree for the area sample Screener is presented in table A-3b in appendix A. Twenty-six cells were formed with weighted response rates ranging from 61.0 percent to 93.3 percent. The lowest response rate was for the group within areas that had each of the following characteristics:

- more than 87.88 percent of the population having at least a high school education;
- more than 7.46 percent of the population being foreign born;
- less than 32.61 percent of the population having some college education;
- employment exceeding 71.66 percent;
- West and Northeast segments; and
- 48.67 percent or less of the housing units being occupied by owner.

The percentage of the population age 25 and older with at least a high school education in the segment was the dominant variable in distinguishing response rate groups.

National Supplement: List Sample

The classification tree for the list sample is summarized in table A-3c in appendix A. Twenty cells were formed with weighted response rates ranging from 61.4 percent to 99.4 percent. The lowest response rate was for the group within segments that had each of the following characteristics:

- less than 16.39 percent of the population being Hispanic or Non-Hispanic Black;
- the percentage of the population that is Hispanic ranging from 2.00 percent to 15.72 percent;
- average household size of more than 2.56; and
- less than 87.89 percent of the population having at least a high school education.

The percentage of Hispanic or Non-Hispanic Black in the segment was the dominant variable in distinguishing response rate groups.

Although the classification tree is useful for dissecting the sample into fine groups of DUs with response patterns as different as possible, it should be used with caution. Since the software only incorporates the full-sample weights, and not replicated weights, in the analysis Search does not take the complex design of the sample into account. Consequently, the significance level of the test may be lower than the $\alpha = 0.05$ level indicated. If the appropriate significance level could be used, then the tree might have fewer significant response cells. Thus, the trees summarized in tables A-3a through A-3c in appendix A give a

conservative picture because any indication of nonresponse bias shown by the Search results may be overstated.

Logistic regression models are also useful in identifying significant effects on response propensity. Screener response status was used as the binary dependent variable, and the PSU sampling frame and ACS variables were used as the predictors, weighted using the Screener base weights adjusted for unknown-eligibility according to PIAAC guidelines.

Main Study

The main effects model was processed using WesVar. An F-test was performed on each variable to determine whether it was significantly related to response propensity. The results of the logistic regression analysis are presented in table A-4a in appendix A. Only the percentage of the population below 150 percent of the poverty level was significant. This confirms the earlier results indicating that there are fewer nonrespondents in the higher poverty levels.

National Supplement: Area Sample and List Sample

For each of the area sample and the list sample, the main effects model was first processed using the LOGISTIC procedure with the stepwise selection method in SAS. The LOGISTIC results indicated the independent variables that should be included in the final model. A subsequent analysis using SURVEYLOGISTIC identified independent variables that were statistically significant when the complex sample design is taken into account.

The results of the logistic regression analyses are shown in tables A-4b and A-4c in appendix A for the area sample and list sample, respectively. For the area sample, two variables were significant at predicting the Screener response propensity: percentage of the population age 25 and older with at least a high school education and percentage of the population that was foreign born. Both were negatively related to response propensity, given the other variables in the model. For the list sample, the percentage of the population that was Hispanic or Non-Hispanic Black had a significantly positive relationship with response propensity; the percentage of the population that was Hispanic and the percentage of the population age 25 and older with at least a high school education had a significantly negative relationship with response propensity, given the other variables in the model. The results confirm the earlier results in terms of the significant predictors of Screener nonresponse.

7.2.2.4 Evaluating Bias Owing to Nonresponse to the BQ

A comparison of BQ respondents and nonrespondents, using variables known for both groups, provided some indication of the potential for nonresponse bias prior to the weighting adjustments. Section 7.2.2.5 describes chi-square tests of the relationship between response indicator and the analysis variable of interest. Section 7.2.2.6 provides a multivariate analysis of the relationship between response indicator and analysis variables that may reveal the areas with the greatest potential for bias before weighting adjustments.

7.2.2.5 BQ Bivariate Nonresponse Bias Analysis

The distribution of BQ respondents was compared with the distribution of all eligible sampled persons for the variables from the PSU sampling frame, the ACS, and the Screener. For the Main Study, weighted percentages and standard errors (SEs) were calculated using the WesVar software and using replicated BQ base weights. For the National Supplement, weighted percentages and SEs were computed using the SURVEYFREQ procedure in SAS and variance strata and variance units in order to account for the complex sample design. To test the significance of the relationship between the response status and each of the variables, a Rao-Scott chi-square (RS3) test of independence was performed. For the Main Study, the analysis used the person-level base weights that had undergone adjustment for unknown eligibility at the Screener level. For the National Supplement, the analyses used the person-level base weights without any unknown eligibility adjustment.

The results of the Rao-Scott chi-square analysis are shown in table A-5a through A-5c in appendix A. Table 7-5 summarizes the results of the significance tests for each sample. Variables in the BQ analysis that were significant at the $\alpha = 0.05$ level were indicated by an X.

Table 7-5. Summary of results from Background Questionnaire bivariate analysis

		National Supplement	
Analysis variables	Main Study	Area sample	List sample
Indicator of whether the PSU is part of a Metropolitan Statistical Area			X
Census region	X	X	
Percentage of housing units occupied by owner			
Percentage of the population age 25 and older with at least a high school education	X	X	
Percentage of the population that is Hispanic or Non-Hispanic Black		X	
Percentage of the population that is Hispanic			
Percentage of the population age 18–64 that is unemployed			
Percentage of the population below 150 percent of poverty	X	X	X
Percentage of the population that is foreign born			
Percentage of households that are linguistically isolated			
Categorized average household size			
Percentage of the population age 18–64 that is employed	X	X	
Percentage of the population age 25 and older with a high school education			
Percentage of the population age 25 and older with some college education			
Race/ethnicity (after imputation)		X	
Age category (after imputation)	X	X	
Language used at screening			
Indicator for children under age 16 in household	X	X	
Gender	X		
Sampling domain (area sample only)		X	

NOTE: Variables that were indicated by an X were significant (at the $\alpha = 0.05$ level) in the bivariate nonresponse bias analysis for BQ.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

The analyses in table 7.5 confirmed what was seen in the response rate analysis in section 7.2.1. The percentage of the population below 150 percent of poverty was the only variable significant for all three samples.

7.2.2.6 BQ Multivariate Nonresponse Bias Analysis

As was done for the Screener analysis, a multivariate analysis was performed to investigate the relationship among a number of variables using the person-level base weights adjusted for unknown eligibility at the Screener level. The classification software package Search (described earlier) was run with BQ response status as the dependent variable, and the variables from the PSU sampling frame, the ACS, and the Screener as the independent variables.

Main Study

The classification tree is summarized in table A-6a in appendix A. Twenty-six cells were formed, with weighted response rates ranging from 58.5 percent to 93.1 percent. The lowest response rate was for a combination of the following characteristics:

- Hispanics age 26 and older;
- with no children under age 16 in the household;
- not living in the Northeastern United States;
- living in segments with unemployment exceeding 4.8 percent; and
- less than 5.1 percent of the population being linguistically isolated.

The presence of children under age 16 in the household was the dominant variable in distinguishing response rate groups.

National Supplement: Area Sample

The classification tree for the area sample is presented in table A-6b in appendix A. Twenty-six cells were formed with weighted response rates ranging from 37.7 percent to 91.8 percent. The lowest response rate was for a combination of the following characteristics:

- with no children under age 16 in the household;
- not unemployed (ages 16 to 34) or older (ages 66 to 74);
- living in census tracts in which the employment rate exceeds 64.53 percent;
- living in the Northeastern United States;

- living in census tracts in which more than 2.42 percent of the population is foreign born;
- persons age 25 to 34 or older than 55; and
- living in census tracts in which the unemployment rate is 4.48 percent or less.

The presence of children under age 16 in the household was the dominant variable in distinguishing response rate groups.

National Supplement: List Sample

The classification tree for the list sample is presented in table A-6c in appendix A. Fifteen cells were formed with weighted response rates ranging from 79.0 percent to 100.0 percent. The lowest response rate was for the combination of the following characteristics:

- living in a Metropolitan Statistical Area;
- female;
- living in the Western and Northeastern United States;
- living in census tracts in which less than 28.57 percent of the population has a high school education; and
- with no children under age 16 in the household.

The indicator of whether a sampled person resided in a Metropolitan Statistical Area was the dominant variable in distinguishing response rate groups. The list sample had one empty variance unit and therefore the corresponding stratum's contribution to the variance estimate was not included when using the Taylor series (linearization) method, which may be a slight underestimate of the true variance.

As discussed earlier (section 7.2.2.3), the trees summarized in tables A-6a through A-6c in appendix A give a conservative picture. Any indication of nonresponse bias shown by the Search results may be overstated since the software does not take into account the complex design of the sample. To further investigate the multivariate relationships between the variables and the response propensity, a logistic regression model was also fit using BQ response status as the binary dependent variable, and the PSU sampling frame, ACS, and Screener variables as the predictors.

Main Study

The results of the logistic regression analysis are presented in table A-7a in appendix A. Only the age and gender of the respondent and whether there is a person age 16 or younger in the household were found significant at the $\alpha = 0.05$ level. This again confirms what was found in the Main Study bivariate analyses: younger persons are more available to participate in an in-person household survey, as are those with children ages 16 and younger, and women.

National Supplement: Area Sample and List Sample

The results of the logistic regression analyses for the BQ response indicator are shown in tables A-7b and A-7c in appendix A for the area sample and list sample, respectively. For the area sample, five variables were significant at predicting the BQ response propensity: sampling domain; percentage of the population age 25 and older with at least a high school education; indicator for children under the age of 16; census region; and percentage of households that are linguistically isolated.[40] For the list sample, only the indicator of whether the PSU is part of a Metropolitan Statistical Area and census region were significant. The list sample had one empty variance unit and therefore the corresponding stratum's contribution to the variance estimate was not included when using the Taylor series (linearization) method. The results confirm the earlier results in terms of the significant predictors of BQ nonresponse.

7.3 Results of the Extended NRBA

Section 7.2 presented an initial assessment of nonresponse bias (NRB) that is essential in identifying effective variables for the weighting process. However, this basic descriptive analysis has its limitations since it does not reflect the effect of weighting adjustments on NRB and the extent of bias remaining after nonresponse adjustments are conducted. Brief descriptions of these types of extended analyses are provided below. To gain further insights into the potential for NRB, the first plausible value for literacy scores is used in some of the extended analyses. Note that multiple analyses to assess NRB are necessary because each analysis has its own limitations. Together, they provide an insight into the patterns and potential for bias. Results are summarized at the end of the section. Key subgroups used in the analysis were discussed at the beginning of section 7.2. In general, unless noted otherwise, the procedures followed are described more fully in appendix D.

With the exception of the comparison of estimates by weighting stage (described in section 7.3.1), the extended analysis is performed using the combined household sample and the final weights. No extended NRBA was performed for the Prison Study given the high response rate.

7.3.1 Comparison of Estimates Before and After Weight Adjustments

The basic analysis described in section 7.2 compared the base-weighted estimates for respondents to the base-weighted full sample estimates. Because the basic analysis was done prior to the weighting process, the inverse of the overall probability of selection was used as the person base weight. That analysis was extended to include a third stage—the respondent estimates, using weights adjusted for nonresponse. In the extended analysis, the person base weights incorporated the Screener nonresponse adjustment. To describe the analysis in other words, estimates of auxiliary variables based on the full sample can be compared to estimates based on the respondents before and after nonresponse weighting adjustments. If the full sample estimates are closer to the adjusted estimates than the base weighted estimates, this indicates that bias in the auxiliary variables was reduced through the weighting process. To the extent that the auxiliary variables are related to proficiency, this could indicate a reduction in bias in the proficiency estimates.

[40] Definition from U.S. Census Bureau: A linguistically isolated household is one in which no member 14 years old and over (1) speaks only English or (2) speaks a non-English language and speaks English "very well."

As mentioned in section 7.1.2, some analyses included the literacy-related nonrespondents (LRNR) to the BQ, for whom age and gender were collected from a household representative, and others will not. In general, such cases are excluded when the analysis involves literacy scores, since literacy scores were not generated for BQ LRNR cases. Table B-1 provides estimated percentages and standard errors for subgroups using base weight and the weights after nonresponse adjustment in weighting. In this analysis, the BQ LRNR cases are included.

Two comparisons were involved: one between the base-weighted full (BWF) sample and base-weighted respondent (BWR) sample and the other between the base-weighted full (BWF) sample and the nonresponse-adjusted (NRA) sample. Because of the substantial overlap between the two groups for each comparison and the resulting impact on covariances and statistical tests, it was decided to take another approach. In general, the *t* test procedure showed significant differences when the absolute differences were very small. The general procedure used to determine if there was an important indication of bias was to first determine if the subgroup percentage for the BWR or NRA sample moved by more than two standard errors from the BWF sample, where the standard error was computed from the BWF sample. The "number of standard errors moved" is represented by the relative difference column in table B-1. For the Main Study and National Supplement area sample, this approach revealed two age subgroups with estimated BWR percentages that moved two standard errors away from the estimates for the BWF sample. However, the NRA estimates for them were less than two standard errors away from the BWF sample, indicating that nonresponse bias in these estimates has been reduced through nonresponse adjustment in weighting. This, in turn, indicates the potential for bias is at a low level.

7.3.2 Comparison of Weighted Estimates to External Totals

Another extended analysis compares estimates from PIAAC to estimates from an external source different from that used in the weighting process. The PIAAC estimates were produced using the final weights adjusted for nonresponse and calibrated to control totals from the 2012 ACS 1-year Public Use Micro Sample (PUMS) data. Care was taken to choose external source estimates that measured the same characteristic for a similar time period. As an alternate external source of estimates, totals were generated from the Current Population Survey's 2012 Annual Social and Economic Supplement. The external source estimates were subject to error, and the variance of these estimates was taken into account when making this comparison. Standard errors for the CPS estimates were derived from the generalized variance function formula using the parameters in table 4 of *Source and Accuracy of Estimates for Income, Poverty, and Health Insurance Coverage in the United States: 2012*.

The results are shown in table B-2. A 95 percent confidence interval for the difference between the PIAAC estimate and the CPS estimate is used to determine whether that difference is significantly different from zero. In the case of the age variable, the confidence interval for the difference includes 0.0 for all age groups except for the 19 to 25, the 36 to 40, and the 41 to 45 year age groups. The actual differences for these three age groups were less than one percentage point in each case. For educational attainment the estimates for the percent with less than a high school education are not significantly different; the estimated percentages for the other three education levels are statistically significant but within approximately 3 percentage points of each other. For gender, race/ethnicity, and census region, the estimates are not significantly different. Other variables were unavailable due to the BQ LRNR cases not having BQ data. In general, the comparison with alternate external data yielded no important indications of potential bias due to nonresponse.

7.3.3 Correlations between Weighting Adjustment Variables and Competency Measures

The analyses described thus far rely on auxiliary variables and do not directly measure bias in the competency estimates. As mentioned earlier, potential bias found in the auxiliary variables is indicative of bias in the competency estimates to the extent that the auxiliary variables and competency estimates are correlated. Thus, correlations between the auxiliary variables and competency data can be computed to evaluate this relationship. For variables used in the weighting adjustments, a low correlation with literacy implies that using the variable in the weighting adjustments did little to reduce NRB. On the other hand, a high correlation with literacy implies a potentially high reduction in NRB. For variables not used in the weighting adjustments, a high correlation with literacy may indicate potential bias in the literacy estimates, unless they are highly correlated with other variables used in weighting, or have no correlation with response status.

The disadvantage of using correlations to evaluate NRB is that the correlations are based on respondents only, and the relationship between competency and the auxiliary variables might be different for nonrespondents. However, this is less of a concern if the relationship can be confirmed using outside sources.

Table B-3 provides the correlations between literacy score and key variables. The BQ LRNR cases are excluded from this analysis since literacy score is not available for them. The correlations (r) were computed as the square root of the R^2 values from a weighted analysis of variance. Employment status ($r = 0.19$) was not used in raking since there was a lack of consistent external data that matched the PIAAC employment variable. Employment status also could not be used in the nonresponse adjustment for the Main Study because it was unknown for nonrespondents. However, it was collected in the Screener for the National Supplement and used in the BQ nonresponse adjustment for that sample. In addition, the categorized percent of the population age 18–64 that was unemployed ($r = 0.15$) was used in both Screener and BQ nonresponse adjustment for both the Main Study and the National Supplement area sample. Among the variables used in weighting, those with the highest correlation with literacy proficiency were education attainment ($r = 0.54$), race/ethnicity ($r = 0.35$), percentage of the population with at least a high school education ($r = 0.36$), and percentage of the population below 150 percent of poverty ($r = 0.31$). Collectively, the BQ nonresponse adjustment cells had a correlation of $r = 0.30$. The correlations across the seven raking dimensions ranged from $r = 0.29$ to 0.58. When the BQ nonresponse adjustment cells were considered with the seven raking dimensions, the correlation was $r = 0.65$.

7.3.4 Comparison of Estimates Using Alternative Weighting Adjustments

For this evaluation, an auxiliary variable is recalibrated to known totals, and estimates of the key statistics are compared before and after the reweighting. Reweighting can be useful as an evaluation tool when:

- the variable was not used in weighting (because it was of a low quality);
- the variable had broad categories and perhaps more detailed variables would have been beneficial;
- the variable is correlated with the outcome measure; and
- the variable is correlated with response propensity.

Any differences between estimates using the official survey weights and the reweighted weights reflect NRB, but if there is not a large change in the estimates, it is further confirmation that NRB may not be a concern, especially if the alternative external totals do not share the same level of quality on the timing of those used in the original weighting process.

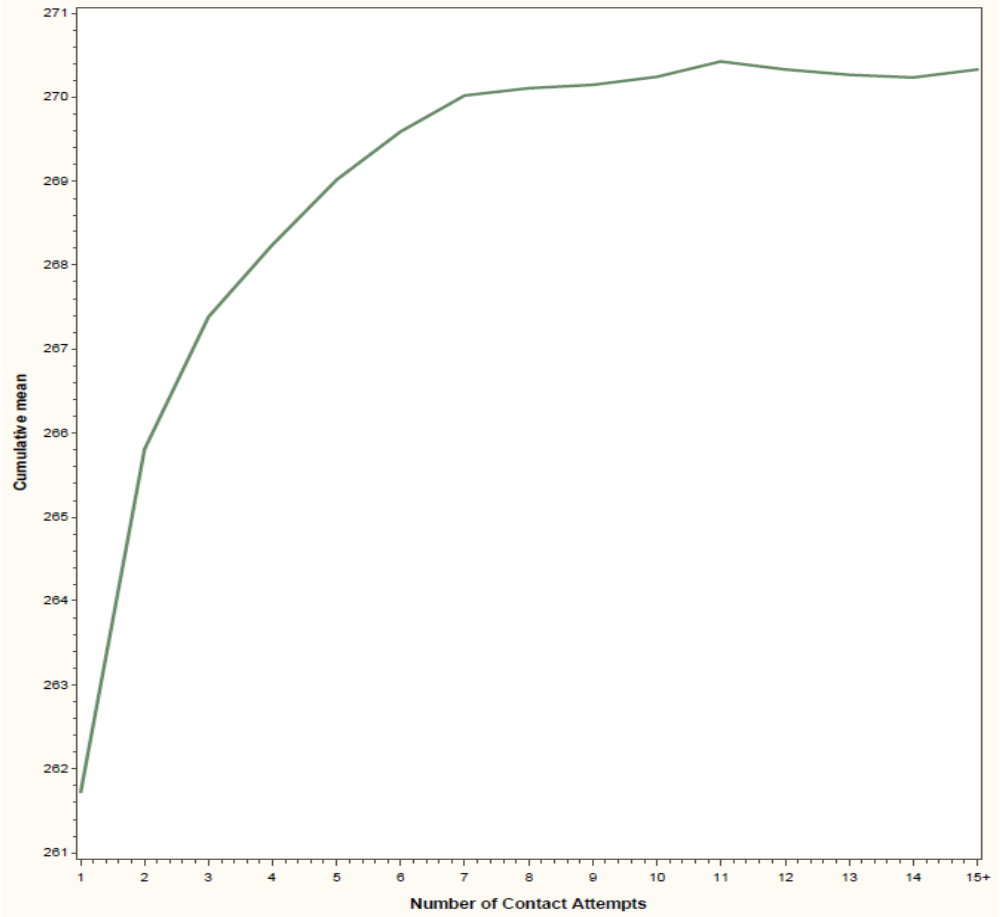
For this analysis, the final weights were recalibrated to 2012 CPS totals with more detailed age categories. The age categories used in the analysis were 16–18, 19–25, 26–30, 31–35, 36–40, 41–45, 46–50, 51–55, 56–60, 61–65 and 66–74. The categories used in the PIAAC weighting process were 16–25, 26–35, 36–45, 46–55, 56–65 and 66–74. Table B-4 provides the estimates and standard errors from the final weights and from the recalibrated weights for key subgroups. Comparison of the estimates, by the weighting approach, does not show any statistically significant differences.

7.3.5 Comparison of Estimates by Level of Effort

This analysis is helpful in evaluating the potential for NRB for differences that cannot be captured through adjusting for known demographics. Significant differences between the competency levels of respondents obtained after a small number of contact attempts ("low effort") and those obtained after a large number of contact attempts ("high effort") imply that the bias could have been potentially reduced through the additional attempts. However, to the extent that differences by level of effort reflect differences between respondents and nonrespondents, the findings indicate that some level of NRB might still be present (depending on the magnitude of the nonresponse rate).

Figure 7-2 shows the average literacy score for cases completed after one contact attempt, and then, cumulatively, after two contact attempts, three contact attempts, etc. The graph shows that average literacy scores were much lower for cases that only required one attempt, and then rose considerably with additional attempts. One possible conclusion from the plot is that perhaps the easiest respondents to find at home scored lower, and that the more difficult to locate, the more different the cases became. This would indicate that the additional attempts helped to reduce the bias in the estimates. Similar results are found when performing this analysis by subgroups, such as age groups.

Figure 7-2. Average literacy score by cumulative contact attempts



SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

7.3.6 Conclusion

The extended NRBA investigated the impact of weighting adjustments, and used the resulting literacy score to gain insights into potential bias through a variety of analyses. The conclusions from the extended NRBA for the household sample are the following:

- An examination of the impact of weighting adjustments indicates the potential for bias is at a low level.
- The comparison with alternate external data yields no important indications of potential bias in outcome statistics due to nonresponse.
- Collectively, the BQ nonresponse adjustment cells had a correlation of $r = 0.30$ with the first plausible value for literacy. The correlations across the seven raking dimensions ranged from $r = 0.29$ to 0.58. When also considering the BQ nonresponse adjustment cells with the seven raking dimensions, the correlation was $r = 0.65$. This

indicates that the weighting variables were moderately correlated with the survey's outcome and likely were effective in reducing bias due to nonresponse.

- When more finely detailed age groupings were used to recalibrate the weights (using the same alternate external data source for the control totals as during the comparison with alternate external data, above), no statistically significant differences were found between the final-weighted estimates and the reweighted estimates.

- Further analysis shows that low-effort respondents (respondents obtained after a small number of contact attempts) scored lower, and that the high-effort respondents (more difficult to locate and/or to respond) scored higher. One plausible assumption is that nonrespondents would score at about the same level as high-effort respondents. The extra level of effort provided different proportions of certain characteristics of respondents.

7.4 Item Nonresponse

7.4.1 Standards on Item Nonresponse Bias

PIAAC Standard 4.7.7 states that "Countries will be required to compute item response rates and conduct an item nonresponse bias analysis for any BQ items with response rates below 85 percent."

Similarly, NCES Standard 2.2.4 states that a nonresponse bias analysis is required at any stage of a data collection with a unit response rate less than 85 percent. If the item response rate is below 85 percent for any items used in a report, a nonresponse bias analysis is also required for each of those items (excluding individual test items). The extent of the analysis must reflect the magnitude of the nonresponse.

7.4.2 Weighted Item Response Rate Computation

The unweighted item response rates are computed by dividing the number of respondents to an item by the total number of unit respondents, excluding those with a valid skip for that item. The weighted item response rates account for the selection probabilities of sampled persons and weighting adjustments (i.e., unknown eligibility adjustment, nonresponse adjustment, and benchmarking). The final weights *(SPFWT0)* were used to compute weighted BQ item response rates. The weighted item response rates were calculated by dividing the number of sampled persons who responded to an item by the number of sampled persons who were eligible to answer the item, adjusting by the final weights. The weighted item response rate can be expressed as:

$$WR_{(item)} = \frac{\sum_{i=1}^{NI} FW_i}{\sum_{i=1}^{N} FW_i}$$

where

$WR(item)$ = weighted item response rate,

FW_i = final weight for respondent i,

NI = total number of persons with a valid response to the item, and

N = total number of persons for which a response to the item is required.

The numerator is the sum of the final weights for all sampled respondents for which a valid response to an item is obtained. Items with responses of "Refused (RF)" were considered nonresponses. The appropriate treatment of "Don't Know" (DK) responses depends on the particular BQ item. In some cases DK is a response that is more informative than just a missing value. If DK responses are of analytic interest, they were treated as responses.

Two sets of weighted item response rates were computed, one treating DK as nonresponse and the other treating DK as valid response. Item NRB analysis is warranted if both weighted response rates fall below 85 percent for a BQ item.

The denominator is the sum of the final weights for all sampled persons for which a response to an item is required. The denominator included completed BQ cases (i.e., DISP_CIBQ (01)), excluding those with a valid skip for the item. Literacy-related nonrespondents were excluded from the computation as the percentage of literacy-related nonrespondents is reported as a separate statistic. In defining a valid skip, sampled persons who failed to provide a valid answer to a skip-controlling question (i.e., an item that determines the skip pattern) were excluded in item response rate calculations for the remaining questions in that skip. Suppose Q3 is a skip-controlling question for Q4 and Q5. If a sampled person failed to provide a valid response to Q3 (and therefore was not asked Q4 and Q5), then that sampled person should be excluded when computing item response rates for Q4 and Q5. Another example for defining a valid skip is when a valid response to Q3 signifies that Q4 and Q5 be skipped. Such cases were removed from the denominator for Q4 and Q5 item response rates.

7.4.3 Response Rates for all Background Questionnaire Items, Household Survey

The response rates for all BQ items for the Main Study and National Supplement for households are found in Appendix C, table C-1. Two sets of weighted item response rates are presented: the first treats "Don't Know" (DK) as a nonresponse; the second treats DK as a valid response. Thirty-one of the BQ items had response rates below 85 percent; 20 of these items had sample sizes below 100 and were disregarded for NRBA analysis (consistent with PIAAC Consortium guidelines). Nine of the items had adequate sample size and reported a low response; however, this was due to an error in coding, which disregarded the use of specific variables used in routing to the specific item. One item (additional payments received last year as part of current job) had a response rate of 70 percent when DK was not considered a valid response, but the rate rose to 93 percent when DKs were included as valid. Another item (total earnings from self-employment after deducting business expenses) had a response rate of 84 percent when DK was not considered a valid response. However, the rate rose to 90 percent when DKs were included as valid. DK seems a reasonable and valid response to the question about additional payments as some people may have lost count of these. Similarly, some respondents may validly not have known the answer to the question about total earnings after business expenses.

Both PIAAC and NCES standards require an NRBA to be conducted for any BQ items with a response rate lower than 85 percent. Since all items had greater than an 85 percent response rate (treating DK as a valid response), the potential for bias due to item nonresponse was considered negligible.

7.4.4 Response Rates for all Background Questionnaire Items, Prison Study

The response rates for all BQ items for the Prison Survey are found in Appendix C, table C-2. Two sets of weighted item response rates are presented: the first treats "Don't Know" as nonresponse; the second treats DK as a valid response. Five of the BQ items had response rates below 85 percent, but 2 of these items had sample sizes below 20 and were disregarded for NRBA analysis (consistent with PIAAC Consortium guidelines). Three items had adequate sample size and a low response rate. However, this was due to an error in coding, which disregarded the use of specific variables used in routing to the item. One item (number of firms or organizations in the last five years) had a response rate of 84 percent when DK was not considered a valid response, but the rate rose to 88 percent when DKs were included as valid. DK was considered valid. This seemed reasonable because prisoners may have lost count of different temporary jobs in the last five years. Two items (both on father's education) that had a response rate of 81 percent when DK was not considered a valid response rose to a 100 percent response rate when DK was considered valid. Again, it seems reasonable that respondents might not know father's education level if he had not been present in the home.

Both PIAAC and NCES standards require an NRBA to be conducted for any BQ items with a response rate lower than 85 percent. Since all items had greater than an 85 percent response rate (treating DK as a valid response), the potential for bias due to item nonresponse was considered negligible.

7.5 Summary of NRBA Results

For the Main Study, the basic NRBA for the BQ found seven variables that were significant at the $\alpha = 0.05$ level: census region; percentage of the population age 25 and older with at least a high school education; percentage of the population below 150 percent of poverty; percentage of the population age 18–64 that is employed; age category; indicator for children under age 16 in household; and gender. As shown in table A-6a in appendix A, the multivariate analysis identified the lowest response rate for a combination of the following characteristics:

- Hispanics age 26 and older;
- with no children under age 16 in the household;
- not living in the Northeastern United States;
- living in segments with unemployment exceeding 4.8 percent; and
- living in areas (census tracts) with less than 5.1 percent of the population being linguistically isolated.

The presence of children under age 16 in the household was the dominant variable in distinguishing response rate groups. In general, younger persons were found to be more likely to participate in an in-person household survey, as were those with children ages 16 and younger, and women.

For the area sample in the National Supplement, the basic NRBA for the BQ found nine variables that were significant at the $\alpha = 0.05$ level: census region; percentage of the population age 25 and older with at least a high school education; percentage of the population that is Hispanic or Non-Hispanic Black; percentage of the population below 150 percent of poverty; percentage of the population age 18–64 that is employed; race/ethnicity; age category; indicator for children under age 16 in household; and sampling

domain. As shown in table A-6b in appendix A, the multivariate analysis identified the lowest response rate for a combination of the following characteristics:

- with no children under age 16 in the household;
- not unemployed (age 16 to 34) or older (age 66 to 74);
- living in census tracts in which the employment rate exceeds 64.53 percent;
- living in the Northeastern United States;
- living in census tracts in which more than 2.42 percent of the population is foreign born;
- persons age 25 to 34 or older than 55; and
- living in census tracts in which the unemployment rate is 4.48 percent or less.

The presence of children under age 16 in the household was the dominant variable in distinguishing response rate groups. Four additional variables were significant at predicting the BQ response propensity: sampling domain; percentage of the population age 25 and older with at least a high school education; census region; and percentage of households that are linguistically isolated.

For the list sample in the National Supplement, the basic NRBA for the BQ found two variables that were significant at the $\alpha = 0.05$ level: indicator of whether the PSU is part of a Metropolitan Statistical Area and percentage of the population below 150 percent of poverty. As shown in table A-6c in appendix A, the multivariate analysis identified the lowest response rate for a combination of the following characteristics:

- living in a Metropolitan Statistical Area;
- female;
- living in the Western and Northeastern United States;
- living in census tracts in which less than 28.57 percent of the population has a high school education; and
- with no children under age 16 in the household.

The indicator of whether a sampled person resided in a Metropolitan Statistical Area was the dominant variable in distinguishing response rate groups. Census region was also found to be a significant predictor of the BQ response propensity.

The conclusions from the extended NRBA for the household sample are as follows:

- An examination of the impact of weighting adjustments indicates the potential for bias is at a low level.
- The comparison with alternate external data yields no important indications of potential bias in outcome statistics due to nonresponse.

- Collectively, the BQ nonresponse adjustment cells had a correlation of $r = 0.30$ with the first plausible value for literacy. The correlations across the seven raking dimensions ranged from $r = 0.29$ to 0.58. When also considering the BQ nonresponse adjustment cells with the seven raking dimensions, the correlation was $r = 0.65$. This indicates that the weighting variables were moderately correlated with the survey's outcome and likely were effective in reducing bias due to nonresponse.

- When more finely detailed age groupings were used to recalibrate the weights (using the same external data source for the control totals as during the formal weighting process), no statistically significant differences were found between the final-weighted estimates and the reweighted estimates.

- Further analysis shows that low-effort respondents (respondents obtained after a small number of contact attempts) scored lower, and that the high-effort respondents (more difficult to locate and/or to respond) scored higher. One plausible assumption is that nonrespondents would score at about the same level as high-effort respondents. The extra level of effort provided different proportions of certain characteristics of respondents.

According to PIAAC and NCES standards, an NRBA is to be conducted for any BQ items with a response rate lower than 85 percent. Since all items for the household sample had greater than an 85 percent response rate (treating DK as a valid response), the potential for bias due to item nonresponse was considered negligible.

No NRBA was needed for the Prison Study because the weighted response rates for all data collection stages and all BQ items are above the 85 percent response rate requirement (treating DK as a valid response).

Finally, the overall conclusion from the PIAAC study on nonresponse bias is that some minimal potential for nonresponse bias exists in the PIAAC estimates; however, the analysis shows that the bias is negligible.

This page intentionally blank.

8. WEIGHTING AND VARIANCE ESTIMATION

The PIAAC Consortium was responsible for deriving sampling weights for the Main Study for all participating countries. However, some countries, such as the United States, were permitted to create their own sampling weights provided that they submitted weighting and nonresponse bias analysis plans several months before data collection for Consortium approval and that they adhered to particular standards and guidelines specified in the *Technical Standards and Guidelines*. These included several quality control measures described earlier in chapter 2.

Sample weights for the U.S. PIAAC were produced for all sample persons in the Main Study and National Supplement samples who completed the BQ as well as those who did not complete the BQ because they were not sufficiently fluent in English or Spanish or had a learning or mental disability that prevented them from completing it. Although the combined Main Study and National Supplement household and prison sample is not part of the international comparison analysis, the PIAAC weighting standards were followed to be consistent with the weighting process for the Main Study. Those who did not complete the BQ for a literacy-related reason[41] received a final weight despite the lack of BQ or assessment data because they are considered part of the PIAAC target population and cannot be represented by survey respondents. Assigning final weights to such cases allows them to be addressed in estimation. This can be done by reporting the percentage of literacy-related nonresponse or by imputing low scores for the BQ literacy-related nonrespondents.

The main purpose of calculating sample weights was to permit inferences from sample persons to the population from which they were drawn. In addition, the sample weighting process was designed to accomplish the following objectives:

1. Permit unbiased estimates, taking into account the fact that all persons in the population did not have the same probability of selection.

2. Minimize the potential bias arising from differences between respondents and nonrespondents.

3. Use auxiliary data on known population characteristics in such a way as to reduce sampling errors and to bring data up to the dimensions of the population totals.

4. Reduce the variation of the weights and prevent a small number of observations from dominating domain estimates.

5. Facilitate sampling error estimation under complex sample designs.

Objective 1 was accomplished by computing base weights for the households selected for screening and, subsequently, for persons selected for the BQ and assessment from the eligible participating households in the household sample. It was also accomplished by computing base weights for the sampled prisons and then inmates sampled in the participating prisons for the BQ and assessment in the prison sample. The details of the base weight calculations for the Screener and the BQ are presented in sections 8.1.2.1.1 and 8.1.2.2.1 for the household sample and 8.2.2.1 and 8.2.2.2 for the prison sample, respectively.

[41] Literacy-related nonrespondents refer to individuals with a reason for nonresponse that was related to proficiency score for the assessment. The reasons include being unable to communicate in English or Spanish, lacking sufficient reading or writing skills to respond to a questionnaire, and learning or mental disability.

Objective 2 was accomplished through nonresponse weighting adjustments that accounted for Screener nonresponse and BQ nonresponse. Sections 8.1.2.1.2 and 8.1.2.2.2, as well as section 8.2.2.3 discuss the nonresponse adjustments for the Screener and BQ for the household sample and the prison sample, respectively. Some reduction in potential bias was also achieved while meeting Objective 3 by calibrating the weights. This was accomplished by using weighting variables that were not used for nonresponse adjustment because data were available only for respondents.

To meet Objective 3, for the household sample the weights were calibrated to known totals from the 2012 American Community Survey (ACS).[42] For the prison sample the weights were calibrated to known totals provided by the Bureau of Justice Statistics. The weights were calibrated using a raking procedure (i.e., iterative poststratification) so that numerous totals calculated with the resulting full-sample weights would agree with the ACS totals. The calibration procedure is described in sections 8.1.4 and section 8.2.2.4 for the household and prison sample, respectively.

Objective 4 was addressed by trimming the weights. A small number of weights were reduced using an inspection approach (referred to as the k x median rule) as required by PIAAC weighting guidelines. After the trimming procedure, the weights were again calibrated to ACS totals. The trimming procedure is described in section 8.1.4 for the household sample. No trimming was conducted for the prison sample.

Finally, Objective 5 was accomplished by creating 45 replicate weights using the stratified jackknife method. Full-sample and replicate weights were calculated for each record to facilitate the computation of unbiased estimates and their standard errors. The weighting procedures were repeated for 45 strategically constructed subsets of the sample to create a set of replicate weights for variance estimation using the jackknife method. The replication scheme was designed to produce stable estimates of standard errors. The replication design and the significance of the number of replicates are discussed further in sections 8.3 and 8.4.

Weighting was performed separately for the household and prison samples. The household sample weighting is described in section 8.1. The weighting process for the prison sample is detailed in section 8.2. Replicate weights and variance estimation for the household and prison samples are discussed in sections 8.3 and 8.4, respectively.

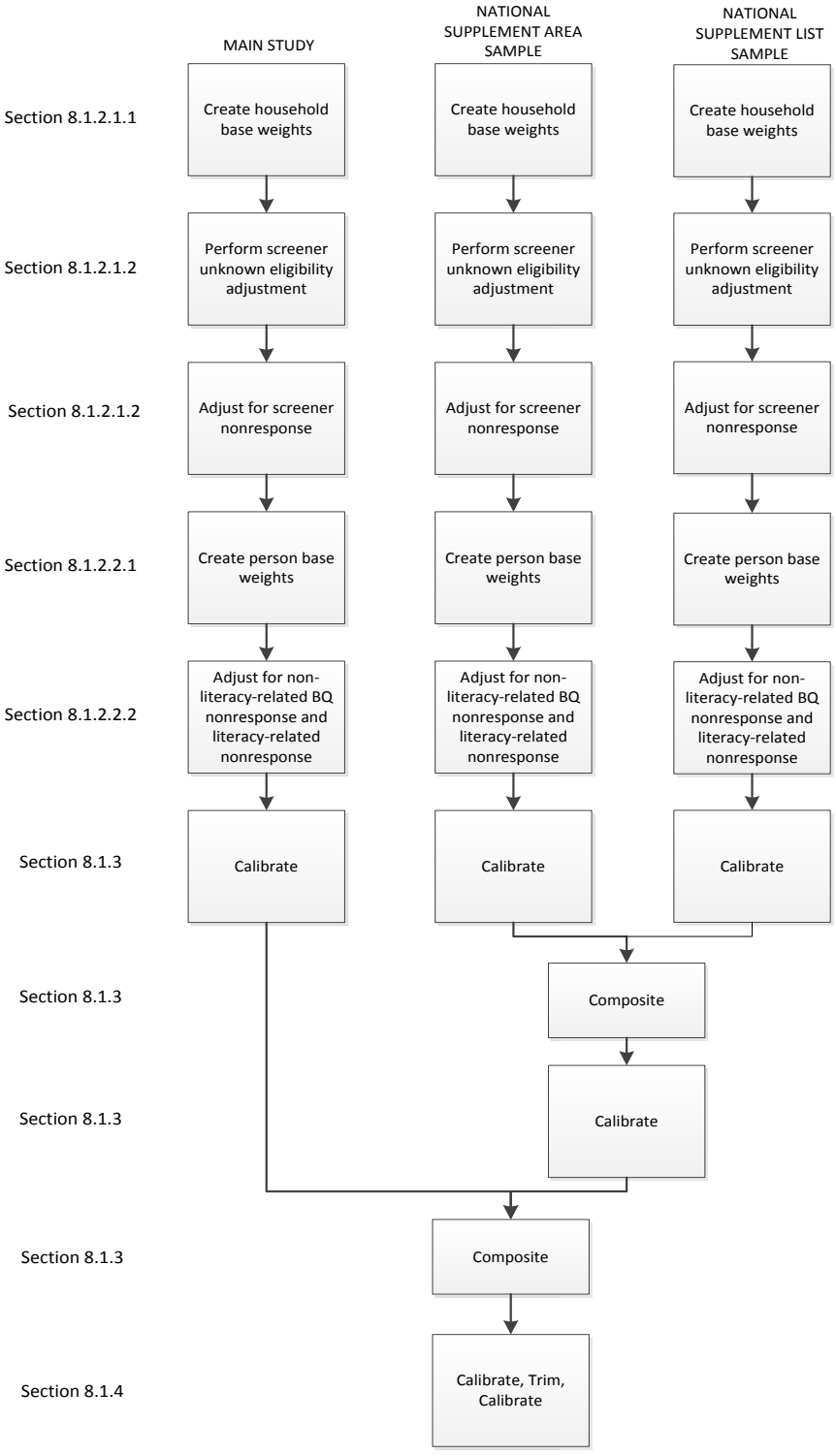
8.1 Household Sample Weighting

For the household sample, an additional goal of the weighting process was to improve the precision of estimates for unemployed persons and two groups of young adults (ages 16–24 and 25–34) by combining the Main Study and National Supplement. Composite weights were produced so that national estimates could be generated for the combined sample. Some discussion of combining samples is provided in the 2003 National Assessment of Adult Literacy (NAAL) Technical Report in the context of combining several independent state samples with the national sample. Other discussions as it relates to dual-frame estimation can be found in Lohr (2011).

Figure 8-1 illustrates the weighting process for the PIAAC household sample. The Main Study sample, National Supplement area sample, and National Supplement list sample were weighted separately to account for nonresponse, calibrated, composited, and then recalibrated.

[42] The ACS is conducted every year as a supplement to the U.S. decennial census, which provides an official count of the entire U.S. population to Congress.

Figure 8-1. Weighting process for the combined Main Study and National Supplement household samples



SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Several steps were completed prior to the weighting process. These preliminary steps are discussed in section 8.1.1. Section 8.1.2 describes the precompositing weighting steps, and section 8.1.3 describes the compositing process. Finally, section 8.1.4 provides a description of the final weighting adjustments after compositing.

8.1.1 Preliminary Steps in Weighting

Prior to weighting, the variables considered for the weighting adjustments were evaluated. Only variables of high quality, available for all eligible units, and related to literacy and response propensity were considered for the nonresponse adjustment. These included some variables created for sampling or collected through PIAAC survey instruments: an indicator of whether the primary sampling unit (PSU) is part of a Metropolitan Statistical Area (MeSA), census region, age category, race/ethnicity, language used at screening, and whether anyone younger than 16 resided in the household. They also included several tract-level estimates derived from the ACS, such as the percentage of the population in the tract that was foreign born. For the National Supplement samples, the sampling domain (unemployed age 16 to 65, not unemployed age 16 to 34 or unknown employment status age 16 to 65, and age 66 to 74) was also considered. For the final calibration adjustment, variables were required to have reliable control totals available from an external source and have less than 5 percent missing data for PIAAC respondents. Again, these included variables created for sampling or collected through the survey instruments: census region, age category, race/ethnicity, gender, level of education, and country of birth. In addition, employment data was used to create composite weights, facilitating the combining of the Main Study and National Supplement household samples.

8.1.1.1 Imputation for Weighting Variables

Missing data values among weighting variables were imputed for weighting purposes. The imputed variables were used only for weighting and, in keeping with the PIAAC *Technical Standards and Guidelines*, were not available for data analysis. Imputation was performed separately for the Main Study and National Supplement but followed the same general procedure.

Although age, race/ethnicity, and gender were collected in both the Screener and the BQ, the BQ measure was preferred for all items as these demographic data were self-reported. For the few cases in which the BQ measure was missing, the Screener value was used as a direct substitute. After this procedure, there were no missing values for gender. Missing values of age category (10 cases) were imputed using the broad age range collected in the Screener. Race/ethnicity for cases missing this item (175 cases) was created by imputing ethnicity (Hispanic/not Hispanic) first, and then race. To obtain values for ethnicity, cells were formed by PSU, segment, and language spoken at the Screener. Then a hotdeck procedure[43] was used to assign the value from a random donor within the cell to the missing case. To obtain values for race, cells were formed by PSU and segment and values imputed using the hotdeck procedure.

For level of education and country of birth, which were not collected through the Screener, a limited amount of imputation was performed to fill in the data for respondents so that the variables could be used in the raking process. Since the raking variables were needed for all cases receiving final weights (i.e.,

[43] Hotdeck is an imputation procedure that uses data from the same sample survey.

BQ respondents and literacy-related nonrespondents), missing values of country of birth and education were imputed using separate procedures for the two groups.

BQ respondents missing values of country of birth and education were imputed using the same hotdeck procedure. To impute the three missing values of country of birth, cells were formed by PSU and segment. To impute the 10 cases missing education category, cells were formed by age category (16–18, 19–24, 25–29, and 30+), PSU, and imputed race/ethnicity.

For persons who were unable to complete the BQ due to a language barrier, country of birth was imputed as "non-U.S." Education was imputed by taking a random draw from the 2008–2010 ACS education distribution for those who speak English not at all plus a portion of those who do not speak English well, in the appropriate age group. For persons who were unable to complete the BQ due to a learning or mental disability, country of birth was assigned as "U.S." with a probability of 0.904 and "nonU.S." otherwise. The probability of 0.904 is based on the percentage of persons with disabilities who were born in the United States according to the 2008–2010 ACS. Education was imputed by taking a random draw from the education distribution for the appropriate age group, where the education distribution is based on the 2008–2010 ACS distribution for persons with disabilities adjusted by the proportion of the disabled population within each education level that had a mental disability according to the 2008–2010 ACS.

Employment status was used to define the domains for compositing, as described in section 8.1.3, and nonmissing values were required for all BQ respondents and literacy-related nonrespondents. For the National Supplement, these data were collected in both the Screener and BQ. The BQ value was taken when available and the Screener value otherwise. This resulted in no missing values of employment status for respondents. For the two literacy-related nonrespondents missing employment status, a value was imputed using the same hotdeck procedure as above, where the cells were formed by PSU and segment, and the donors were limited to other literacy-related nonrespondents.

No employment status information was collected in the Screener in the Main Study, so a different imputation approach was needed. For the two respondents with missing values, cells were formed by PSU and segment and values imputed using the hotdeck procedure. Imputation for the literacy-related nonrespondents was done by taking a random draw from the employment distributions from the 2012 ACS. For language problems,[44] this was based on the distribution of employment for those that speak English not well or not at all. For learning/mental disabilities, imputation used the distribution of employment for persons with cognitive difficulty.

In the sections that follow, the weighting process refers to nonliteracy-related adjustments and literacy-related adjustments. Particular attention was given to classify nonresponse into literacy-related and nonliteracy-related nonresponse categories. All nonliteracy-related nonrespondents were considered to be similar to respondents with respect to proficiency scores; however, the literacy-related nonrespondents were not considered to be similar to the respondents since they likely would have scored lower than average if they attempted the assessment. For the Screener, literacy-related nonresponse occurred if the household representative could not speak the language of the interview. For the BQ and assessment, literacy-related nonresponse occurred if the sample person could not speak the language of the interview or could not complete the interview and assessment due to reading or writing difficulty or a learning/mental disability that precluded answering the questions.

[44] Language problems: Individuals who were unable to comprehend, speak, or read English or Spanish.

8.1.1.2 Selection of Nonresponse Adjustment Variables

The weighting nonresponse adjustments can be effective in reducing nonresponse bias in the PIAAC estimates if the variables used in the adjustments are both related to the survey outcome (proficiency) and to response propensity (Little 1986). In accordance with the PIAAC *Technical Standards and Guidelines*, an initial nonresponse bias analysis was conducted to evaluate the relationship between the potential nonresponse adjustment variables and response propensity. The Main Study, National Supplement area sample, and National Supplement list sample were analyzed separately to inform the separate nonresponse weighting adjustments for each sample.

The classification software package Search,[45] which employs a hierarchical tree algorithm, was used for the initial nonresponse bias analyses at both the Screener and BQ levels. The software identifies the auxiliary variables that best define subgroups with differential response rates. The analyses were performed using base weights (equal to the inverse of the selection probability at each stage) to reflect any differential sampling rates. To the extent that the auxiliary variables are related to proficiency, the differential response rates among the subgroups (or differences in the distributions of respondents and nonrespondents) may indicate potential nonresponse bias in the unadjusted PIAAC estimates. Using these variables in the weighting adjustments should help alleviate this bias.

8.1.1.2.1 Screener Analysis

A Screener-level nonliteracy-related nonresponse status variable was created and used as the dependent variable for the analysis. Because very little was known about the households that did not respond to the Screener, information used to form weighting classes had to come from a different source. The frame contained only aggregate demographic information, such as census region and MeSA status. However, because the sampling was performed using census geography, the sampled SSUs were merged to ACS tract-level[46] data files to create approximate SSU-level weighting variables. ACS 2005–2009 was used for the Main Study ACS 2008–2012 for the National Supplement. All the characteristics related to proficiency considered as predictors are shown in table 8-1. The MeSA indicator and census region come from the PSU sampling frame. The other variables are estimated quartiles of tract-level data from the ACS 2005–2009 for the Main Study and ACS 2008–2012 for the National Supplement.

Table 8-1 also indicates which variables were found to be significant predictors of Screener response and were used to form the Screener nonresponse adjustment cells. Twenty-six final cells were formed for the Main Study, 26 for the National Supplement area sample, and 20 for the National Supplement list sample. These cells were used for both the Screener-level nonliteracy-related nonresponse adjustment and the unknown eligibility adjustment.

[45] Details on the history and development of the software and also some references to early evaluations may be found at http://www.isr.umich.edu/src/smp/search/. Details of some computations in Search are provided in Sonquist, Baker, and Morgan (1974).

[46] Blocks are very fine partitions of the United States, formed using visible semipermanent features such as roads, railroad tracks, mountain ridges, bodies of water, and power lines. The only invisible boundaries used are county, state, and national boundaries. Minor civil division boundaries and property lines are ignored. A block group is a small group of contiguous blocks. A tract is a collection of contiguous block groups all within the same county.

Table 8-1. Screener nonresponse adjustment variables

Variables	Main Study	National Supplement	
		Area sample	List sample
Indicator of whether the PSU is part of a Metropolitan Statistical Area	X	X	
Census region	X	X	X
Percentage of housing units occupied by owner	X	X	X
Percentage of the population age 25 and older with at least a high school education		X	X
Percentage of the population that is Hispanic or Non-Hispanic Black	X	X	X
Percentage of the population that is Hispanic	X	X	X
Percentage of the population age 18–64 that is unemployed	X	X	
Percentage of the population below 150 percent of poverty	X	X	X
Percentage of the population that is foreign born	X	X	X
Percentage of households that are linguistically isolated	X	X	
Categorized average household size	X	X	X
Percentage of the population age 18–64 that is employed		X	
Percentage of the population age 25 and older with a high school education	X	X	
Percentage of the population age 25 and older with some college education	X	X	X

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

8.1.1.2.2 BQ Analysis

A BQ-level nonliteracy-related nonresponse status variable was created and used as the dependent variable for the analysis. In addition to the items used for the Screener analysis, characteristics from the Screener that were also considered predictors and are shown below.

From the Screener enumeration:

- indicator for children under age 16 in household.

From the Screener:

- age category (after imputation);
- gender;
- race/ethnicity (after imputation);
- language used at screening; and
- sampling domain (for the National Supplement only).

Table 8-2 shows the items that were found to be the most significant predictors of nonliteracy-related response to the BQ and were used to form the adjustment cells. Twenty-six final cells were formed for the

Main Study, 26 for the National Supplement area sample, and 15 for the National Supplement list sample. These cells were then used for the BQ-level nonliteracy-related nonresponse adjustment.

Table 8-2. Background Questionnaire nonresponse adjustment variables

Variables	Main Study	National Supplement: Area sample	National Supplement: List sample
Indicator of whether the PSU is part of a Metropolitan Statistical Area			X
Census region	X	X	X
Percentage of housing units occupied by owner	X	X	
Percentage of the population age 25 and older with at least a high school education	X	X	
Percentage of the population that is Hispanic or Non-Hispanic Black	X		
Percentage of the population that is Hispanic		X	
Percentage of the population age 18–64 that is unemployed	X	X	
Percentage of the population below 150 percent of poverty		X	X
Percentage of the population that is foreign born	X	X	
Percentage of households that are linguistically isolated	X		X
Categorized average household size	X		
Percentage of the population age 18–64 that is employed	X	X	
Percentage of the population age 25 and older with a high school education	X	X	X
Percentage of the population age 25 and older with some college education	X	X	
Age category (after imputation)	X	X	X
Gender			X
Race/ethnicity (after imputation)			
Language used at screening			
Indicator for children under age 16 in household	X	X	X
Sampling domain (National Supplement area sample only)		X	

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

8.1.1.3 Selection of Post-Compositing Calibration Adjustment Variables

The PIAAC Consortium stipulated that at minimum the weights must be benchmarked to control totals for age and gender. Other key variables of interest were race/ethnicity, educational attainment, country of birth, and census region. Two-way crossing of these variables formed raking dimensions for the 16- to 65-year-olds and 66- to 74-year-olds separately, as shown in table 8-3. Since the Main Study sample included only 16- to 65-year-olds, creating raking dimensions separately for the two age groups helped maintain consistency of estimates between the Main Study and National Supplement, and helped improve the precision of estimates for the two age groups. Seven raking dimensions were created for the 16- to 65-year-olds, while three raking dimensions were created for the 66- to 74-year-olds given the smaller sample size.

Table 8-3. Variables involved in the calibration process

16- to 65-year-olds

Categories:	Education[1] (4 levels): < HS, HS, Some college, College degree +	Race/ethnicity[2] (3 levels): Hispanic, Non-Hispanic Black, Other	Age (5 levels): 16–25, 26–35, 36–45, 46–55, 56–65	Gender (2 levels): Male, Female	Country of birth (2 levels): U.S., not U.S.	Region (4 levels): Northeast, Midwest, South, West
Raking dimension						
1 (12 levels)	✓	✓				
2 (20 levels)	✓		✓			
3 (8 levels)	✓			✓		
4 (15 levels)		✓	✓			
5 (6 levels)		✓		✓		
6 (10 levels)			✓		✓	
7 (8 levels)					✓	✓

66- to 74-year-olds

Raking dimension	Categories
1 (8 levels)	Education (4 levels same as above) by gender (2 levels)
2 (4 levels)	Race/ethnicity by gender: Hispanic, Non-Hispanic Black, Other male, Other female
3 (6 levels)	Country of birth by region: born in U.S. by 4-level region, and not born in U.S. by 2-level region (Northeast and Midwest, South and West)

[1] <HS: Less than high school; HS: High school diploma or equivalent; Some college: Some college, no degree received; College degree +: A college degree or higher.
[2] All adults of Hispanic origin are classified as Hispanic regardless of race. Those classified as Black are non-Hispanic Black only. Those classified as other include non-Hispanics of all other races, including multiracial.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

The 2012 ACS 1-year Public Use Micro Sample (PUMS) data were used to construct control totals for the raking step in the weighting process. The 2012 PUMS data contain 3,113,030 person records (all ages). Even though data from the ACS are subject to sampling error, the ACS is the largest survey other than the Decennial Census that the Census Bureau administers and, as such, results in more accurate estimates of the raking dimensions than is possible with the smaller PIAAC sample.

The ACS totals include residents in institutional group quarters and military barracks that are not part of the PIAAC target population. Therefore, they were removed from the ACS PUMS data before constructing control totals.

8.1.2 Pre-compositing

8.1.2.1 Screener Weighting Adjustments

The weighting process began with the creation of household-level base weights. The household-level weights reflect the household selection probability (section 8.1.2.1.1) and were adjusted for nonresponse to the Screener (section 8.1.2.1.2). The Main Study sample, National Supplement area sample, and National Supplement list sample were processed separately.

8.1.2.1.1 Screener Base Weights

The Screener base weight was assigned to all sampled households and is equal to the reciprocal of the household selection probability. Since PIAAC has a multistage design, with households selected within primary and secondary sampling units (PSUs and segments/tracts), the selection probability is equal the product of the conditional selection probabilities at each stage:

$$W_k = \frac{1}{P_k} = \frac{1}{P_i P_{j|i} P_{k|j}},$$

where, P_i is the probability of selecting PSU i, $P_{j|i}$ is the conditional probability of selecting segment j within PSU i (or tract j within PSU i for the National Supplement list sample), and $P_{k|j}$ is the conditional probability of selecting household k within segment j. The selection probability also reflects any changes to the subsampling procedures. Since the total number of dwelling units (DUs) selected included a reserve sample, this selection probability also includes the proportion of the larger sample that was released to the field.

Since the Main Study had an equal probability design, W_k was equal to 13,483, the inverse of the sampling rate, for all cases initially. However, two DUs were originally listed erroneously as multifamily structures, with seven units each on the DU sampling frame. The actual household selection probability for these two DUs is seven times that for other DUs on the frame, requiring a further adjustment to their Screener base weights.

The National Supplement area sample also had an equal probability design, with W_k equal to 13,269 initially. Three DUs were originally listed as multifamily structures with either two or three units. The weights were adjusted accordingly by a factor of one-half or one-third.

The National Supplement list sample did not have an equal probability design. An equal probability sample of DUs would have resulted in large variation in the number of completed cases across tracts. Any variation in the weights is dampened after compositing with the Main Study sample and area sample. Therefore, the weights were allowed to vary to achieve a more balanced workload. In addition, a further adjustment was needed for one DU erroneously shown as a three-unit structure on the U.S. Postal Service (USPS) address list. The Screener base weights W_k ranged from 24 to 28,835.

8.1.2.1.2 Screener Weighting Adjustments

The next step in the weighting process was to adjust the base weights to reduce potential bias as a result of nonresponse to the Screener. Sampled DUs were classified into one of the following categories: Screener respondent, Screener literacy-related nonrespondent, Screener nonliteracy-related nonrespondent, ineligible unit, and unit with unknown eligibility. The adjustment for units with unknown eligibility is described below. This adjustment differed for the Main Study and National Supplement given the different eligibility definitions. This section also addresses the treatment of literacy-related nonrespondents and describes the Screener nonresponse adjustment. All eligible households that responded to the Screener were assigned a nonzero adjusted Screener weight.

Unknown Eligibility in the Main Study

Before any nonresponse adjustment was processed, an adjustment for unknown eligibility was performed. In this step, a portion of the weights of the households with unknown eligibility status (i.e., whether they contain a person age 16 to 65) were distributed to the ineligible cases. The down-weighted unknown eligibility cases were then treated as eligible nonrespondents.

The adjustment was done within the weighting cells defined for the nonresponse adjustment (see section 8.1.1.2.1). Weighting classes were combined if a cell size was less than 30 or an adjustment factor for the unknown eligibility cases was less than 0.55. Within each weighting cell, the Screener unknown eligibility adjustment factor was computed as follows:

$$F_{1k} = \begin{cases} \dfrac{S_L + S_R + S_{NR} + S_I + S_U}{S_L + S_R + S_{NR} + S_I} & \text{if } k \in \text{I} \\ \dfrac{S_L + S_R + S_{NR}}{S_L + S_R + S_{NR} + S_I} & \text{if } k \in \text{U} \\ 1 & \text{if } k \in \text{L, R, NR,} \end{cases}$$

where S represents the sum of the Screener base weights over records in the same adjustment cell as household k. The factors and weights shown here are for a household k and households are classified as R: respondent, L: literacy-related nonrespondent, NR: nonliteracy-related nonrespondent, I: ineligible, or U: unknown eligibility. The resulting factors for the unknown eligibility cases ranged from 0.56 to 0.78. The factors for the eligible cases were set to 1.00. The Screener unknown eligibility adjusted weight was calculated as $W_k F_{1k}$.

Unknown Eligibility in the National Supplement

For the National Supplement, a step was needed to adjust the weights of the Screener literacy-related nonrespondents and Screener nonliteracy-related nonrespondents (i.e., disabilities and partial completes) for unknown eligibility. This step is not part of the standard PIAAC weighting process but was necessary due to the different eligibility definition for the National Supplement. To be consistent with the Main Study, it was assumed that age eligibility of the Screener nonrespondents was known (i.e., there was someone in the household between the ages of 16 and 74 for the area sample or between the ages of 16 and 65 for the list sample). However, the employment status eligibility was unknown.

The adjustment was done within the weighting cells defined for the nonresponse adjustment. Within each cell, the adjustment factor was calculated as

$$F_{0k} = \begin{cases} \dfrac{S_I + (1 - p_{unemp})(S_L + S_{NR})}{S_I} & \text{if } k \in \text{I} \\ p_{unemp} & \text{if } k \in \text{L, NR} \\ 1 & \text{if } k \in \text{R, U} \end{cases},$$

Where S represents the sum of the Screener base weights over records in the same adjustment cell as household k, and $p_{unemp}$is the estimated proportion of households that have a least one person in the target population given that there is at least one age-eligible person, calculated as

$$p_{unemp} = \frac{p_{elg}}{p_{age}},$$

where p_{elg}= the overall Screener eligibility rate calculated from the sample, and p_{age}= the proportion of households with at least one person age 16 to 74 for the area sample or the proportion of households with at least one person age 16 to 65 for the list sample, based on ACS 2012. The factors and weights shown above are for a household k and households are classified as R: respondent, L: literacy-related nonrespondent, NR: nonliteracy-related nonrespondent, I: ineligible, or U: unknown eligibility. The resulting adjustment factor F_{0k} for Screener literacy-related nonrespondents and nonliteracy-related nonrespondents was equal to 0.48 in the area sample and 0.11 in the list sample.

Next, the standard unknown eligibility adjustment was performed to distribute a portion of the weights of the remaining households with unknown eligibility status (i.e., whether they contained a member of the target population) to the ineligible cases. For these households, both age and employment eligibility were unknown. The down-weighted unknown eligibility cases were then treated as eligible nonrespondents.

The adjustment was done within the same weighting cells as the previous adjustment. Within each weighting cell, the Screener unknown eligibility adjustment factor $F_{1k}$is computed as shown above for the Main Study, except using the weights after the initial adjustment. The resulting factors for the unknown eligibility cases ranged from 0.29 to 0.50 for the area sample and from 0.03 to 0.22 for the list sample. The factors for the eligible cases were set to 1.00. The Screener unknown eligibility adjusted weight is calculated as $W_k F_{0k} F_{1k}$.

Treatment of Literacy-Related Nonresponse

As mentioned above, for the Screener nonresponse adjustment, the nonrespondents were divided into two categories. The first consisted of cases involving nonliteracy-related nonresponse. Nonliteracy-related nonrespondents were considered to be similar to respondents with respect to proficiency scores. The second consisted of cases involving literacy-related nonresponse. Language problems were the only type of literacy-related nonresponse at the Screener level. Households with this type of nonresponse were presumed to differ from responding households with respect to literacy. Therefore, the weighting procedures adjusted the weights of the respondents to represent the nonliteracy-related nonresponse only. The weights of the literacy-related nonresponse cases were not adjusted during the Screener-level nonresponse adjustment because their literacy was expected to differ from that of respondents. The contribution of the Screener-level literacy-related nonresponse to the total population is accounted for by the literacy-related nonresponse adjustment carried out for the BQ sample (refer to section 8.1.2.2.2).

Nonresponse Adjustment

An adjustment was made to distribute the Screener base weights (adjusted for unknown-eligibility status) of the nonliteracy-related nonrespondents to the Screener respondents. The nonresponse adjustment was performed within cells that were defined based on the analysis described in section 8.1.1.2.1. The cells were required to adhere to certain criteria. First, since it is generally preferable to redistribute the weights of nonrespondents over a relatively large pool of cases, the minimum cell size allowed was 30 cases. Second, variation in nonresponse adjustment factors typically increases the variation in the weights, which in turn generally increases the variance of the survey estimates. Since the amount of variation in the nonresponse adjustment factors was a concern, the maximum allowable nonresponse adjustment factor was 2.0.

Within each weighting cell, the nonliteracy-related Screener nonresponse adjustment factor is computed as follows:

$$F_{2k} = \begin{cases} 1 & \text{if } k \in \text{L, I} \\ \dfrac{S_R + S_{NR} + S_U}{S_R} & \text{if } k \in \text{R} \\ 0 & \text{if } k \in \text{NR, U,} \end{cases}$$

where S represents the sum of the unknown eligibility-adjusted Screener weights over records in the same adjustment cell as household k. The factors and weights shown here are for a household k and households are classified as R: respondent, L: literacy-related nonrespondent, NR: nonliteracy-related nonrespondent, I: ineligible, or U: unknown eligibility. For the Main Study, the resulting factors ranged from 1.04 to 1.33 with the average adjustment factor being 1.14. The Screener nonresponse adjusted weight is calculated as $W_k F_{1k} F_{2k}$. For the National Supplement, the resulting factors ranged from 1.08 to 1.51 with a mean of 1.17 for the area sample and from 1.01 to 1.26 with a mean of 1.11 for the list sample. The Screener nonresponse adjusted weight is calculated as $W_k F_{0k} F_{1k} F_{2k}$.

8.1.2.2 BQ Weighting Adjustments

The person-level weights reflect the nonresponse to the Screener and the within-household selection probabilities (section 8.1.2.2.1) and were adjusted for nonresponse to the BQ (section 8.1.2.2.2). The three samples (Main Study, National Supplement area sample, and National Supplement list sample) were processed separately.

8.1.2.2.1 BQ Base Weights

The derivation of base weights was necessary to prevent potentially serious biases in the outcome statistics. For the Main Study, the study specifications called for the selection of one person in households with fewer than four eligible members and two persons in households with four or more eligible members. Therefore, the probabilities of selection varied by the number of eligible persons in the household. For example, members of households with only one eligible member had twice the chance of selection as those in households with two (or four) eligible members. For the National Supplement, the sampling rules differed depending on the age and employment status of the household member and the sample release group of the household, as described in sections 3.1.2.1.3 and 3.1.2.2.3. To produce unbiased estimates, different weights had to be used to account for the within-household selection rate.

The BQ base weights account for both nonresponse to the Screener and different within-household selection rates. The BQ base weights were computed as the product of the Screener nonresponse-adjusted weight and the reciprocal of the within-household probability of section for person *l* within household *k* of PSU *i* and segment/tract *j*:

$$W_l = \begin{cases} W_k F_{1k} F_{2k} \dfrac{1}{P_{l|k}} \ for\ the\ Main\ Study \\ W_k F_{0k} F_{1k} F_{2k} \dfrac{1}{P_{l|k}} \ for\ the\ National\ Supplement. \end{cases}$$

The within-household person selection probability, $P_{l|k}$, was defined in chapter 3 for each sample.

8.1.2.2.2 BQ Nonresponse Adjustments

The BQ base weights were then adjusted for nonresponse using the same procedure as described in section 8.1.2.1.2 but with an additional step. Before the BQ weights were calibrated, the weights of the literacy-related respondents to the BQ and assessment were adjusted to account for the literacy-related Screener nonrespondents. This adjustment was necessary primarily to allow the literacy-related respondents to represent the literacy-related Screener nonrespondents in the calibration procedure. This adjustment assumes that the literacy-related nonrespondents to the Screener, BQ, and the assessment are similar in literacy.

Nonliteracy-Related Nonresponse Adjustment

An adjustment was made to distribute the BQ base weights of the nonliteracy-related nonrespondents to the nonliteracy-related respondents. The nonresponse adjustment was performed within cells that were defined based on the analysis described in section 8.1.1.2.2. Again, the cells were required to contain at least 30 cases and to have nonresponse adjustment factors no larger than 2.0.

Within each weighting cell, the nonliteracy-related BQ nonresponse adjustment factor is computed as follows:

$$F_{3l} = \begin{cases} 1 & \text{if } l \in \text{L, I} \\ \dfrac{S_R + S_{NR}}{S_R} & \text{if } l \in \text{R} \\ 0 & \text{if } l \in \text{NR,} \end{cases}$$

where S represents the sum of the BQ base weights over records in the same adjustment cell as person l. The persons are classified as R: respondent, L: literacy-related nonrespondent, NR: nonliteracy-related nonrespondent, or I: ineligible. For the Main Study, the resulting factors ranged from 1.07 to 1.69, with the average adjustment factor being 1.23. For the National Supplement, the resulting factors ranged from 1.09 to 1.95 with a mean of 1.31 for the area sample and from 1.00 to 1.27 with a mean of 1.08 for the list sample. For all samples, the BQ nonliteracy-related nonresponse adjusted weight is calculated as $W_l F_{3l}$.

Literacy-Related Nonresponse Adjustment

The next step was to distribute the weights of the Screener cases that did not respond due to language problems to the literacy-related BQ and assessment nonrespondents (those with language problems or learning or mental disabilities that prevented them from fully completing the questionnaire or assessment). At the Screener level, there were 106 literacy-related nonrespondents in the Main Study, 130 in the National Supplement area sample, and 88 in the National Supplement list sample. At the BQ and assessment level, there were 120 literacy-related nonrespondents in the Main Study, 61 in the National Supplement area sample, and 13 in the National Supplement list sample. The weights of both the BQ and assessment literacy-related nonrespondents were adjusted to account for the Screener literacy-related nonrespondents. The Screener weights associated with the cases with literacy-related nonresponse received a Screener nonresponse adjustment factor equal to one (see section 8.1.2.1.2) under the presumption that the Screener literacy-related nonrespondents were unlike other respondents and, therefore, were excluded from the Screener nonresponse adjustment process.

The literacy-related nonresponse adjustment at the BQ stage was necessary to allow the Screener literacy-related nonrespondents to be accounted for in the raking procedure as part of the total population. To do so, we allowed the BQ and assessment literacy-related nonrespondents to represent the Screener literacy-related nonrespondents. We assumed that the literacy-related nonrespondents to the Screener, the BQ, and the assessment were similar in proficiency.

Within each weighting cell, the adjustment is the following:

$$F_{4l} = \begin{cases} 1 & \text{if } l \notin \text{L} \\ \dfrac{S_L^{screener} + S_L^{BQ} + S_L^{assessment}}{S_L^{BQ} + S_L^{assessment}} & \text{if } l \in \text{L}, \end{cases}$$

where

$S_L^{Screener}$ = sum of Screener base weights for DUs with a Screener disposition of a literacy-related nonresponse category;

S_L^{BQ} = sum of Screener base weights for DUs with a BQ disposition of a literacy-related nonresponse; and

$S_L^{assessment}$ = sum of Screener base weights for DUs with an assessment disposition of a literacy-related nonresponse.

The Consortium recommended using the literacy-related disposition codes (language problem or learning or mental disability) as weighting classes for this step. However, given the small number of learning or mental disability cases, only one weighting class was used for all literacy-related cases. The adjustment factor for all the literacy-related BQ and assessment respondents was equal to 1.88 for the Main Study, 2.09 for the National Supplement area sample, and 2.37 for the National Supplement list sample. The final BQ nonresponse adjusted weight (adjusted for literacy and nonliteracy nonresponse) is calculated as $W_l F_{3l} F_{4l}$.

8.1.3 Compositing

After adjusting each sample for nonresponse to the Screener and BQ, the weights from the three samples were composited. This was done in two steps. First, the area and list samples from the National Supplement were combined, as described in section 8.1.3.1. Then the resulting sample was combined with the Main Study, as described in section 8.1.3.2.

8.1.3.1 Compositing the National Supplement Area and List Samples

As indicated in chapter 3, the National Supplement list sample was selected from five purposively chosen tracts within each PSU (400 tracts total in the 80 PSUs). It was assumed that the five tracts represented only themselves and formed a stratum within each PSU. The composite weighting process respected the tract strata such that the list sample was combined with the National Supplement area sample cases that were sampled as unemployed and resided in one of the 400 tracts. A geocoding operation was conducted to identify DUs in the area sample that fell into one of the tracts in the list sample frame.

8.1.3.1.1 Pre-Compositing Calibration

Prior to compositing, the weights were calibrated (poststratified) to population totals for area sample and list sample, respectively. To be consistent with the compositing domains, the area sample was calibrated separately by tract strata. The precompositing calibration was necessary because of the "poststrata" that defined the stratum for the tracts, as mentioned above. The area sample had a design that did not pay attention to that stratum boundary, and therefore had segments that represented other segments both in and out of the stratum. In fact, some of the segments were split across the two strata because the segments were defined on Census 2000 geography definitions while the tracts were defined on the Census 2010 definitions. The calibration process brought the sum of weights to match the population totals within each stratum, helped determine the compositing factor in the subsequent compositing process, and reduced potential bias in survey estimates.

Poststratification adjusts survey weights of respondents so that the weighted sample distribution is the same as some known population distribution. In other words, the sums of the adjusted weights of the respondents are equal to known population totals for certain subgroups of the population. For the list sample and each tract stratum in the area sample, the weights were poststratified to population totals by sampling domain crossed with age. For each tract stratum in the area sample, there were four poststratification cells: unemployed age 16 to 34, unemployed age 35 to 65, not unemployed age 16 to 34, and age 66 to 74. For the list sample, there were two cells: unemployed age 16 to 34 and unemployed age 35 to 65. Screener responses for employment status and age were used for the adjustment.

The control totals were based on tract-level data from the most recent ACS 5-year (2008–2012) summary file (SF). Separate totals were obtained for the 400 tracts within the sampled PSUs and the other tracts in the sampled PSUs, and then these totals were weighted using the PSU weights (inverse of the PSU selection probability) to create national population totals. The ACS SF includes institutionalized population, which is not part of PIAAC target population. Therefore, a reduction to the ACS totals was required in order to remove them from the ACS control totals. In addition, the age groups in the ACS tables did not match PIAAC target population precisely and thus needed an adjustment as well.

Employment status is not defined exactly the same in ACS[47] and PIAAC, and it is not possible to derive the ACS estimate from the set of PIAAC BQ items. (Refer to figure 8-2 for a side-by-side comparison of the ACS, Current Population Survey (CPS), and PIAAC definition of employment status.) Table 8-4 shows the unemployment rates for the PIAAC Main Study, ACS, and CPS. As can be seen, the unemployment rates from the PIAAC Main Study and ACS are fairly close for a similar time period. Thus the ACS employment control totals were used, given the rates are similar, and employment status was only used in the precompositing calibration.

[47] The ACS definition of employment status is available at:
https://www2.census.gov/programs-surveys/acs/tech_docs/subject_definitions/2014_ACSSubjectDefinitions.pdf.

Table 8-4. Unemployment rates

Survey	Unemployment rate
PIAAC Main Study (8/2011–4/2012)	9.7 percent (noninstitutional 16–65)
2012 ACS	9.4 percent (civilian 16+)
2008–2012 ACS	9.3 percent (civilian 16+)
2012 CPS	8.2 percent (civilian noninstitutional population age 16+)

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012; U.S. Department of Commerce, Census Bureau, 2012 and 2008–2012 ACS; U.S. Department of Commerce, Census Bureau, 2012 Current Population Survey (CPS).

Figure 8-2. Comparison of the ACS, CPS, and PIAAC definitions of unemployment status: 2012

Employed persons are those who…			Unemployed persons are those who…		
ACS	CPS	PIAAC	ACS	CPS	PIAAC
Did **any** work for pay or profit during the survey reference week.	Did **any** work for pay or profit during the survey reference week.	Did any paid work **for at least one hour** during the survey reference week.	Did not have a job at all during the survey reference week, were **actively** looking for work during the prior 4 weeks, and were available for work last week.	Did not have a job at all during the survey reference week, **made at least one specific active effort** to find a job during the prior 4 weeks (**unless temporarily ill**), and were available for work **that week**. *["that week" inferred from the questionnaire.]*	Did not meet any of the criteria listed under "employed", were looking for work and **did something** to find a job in the last 4 weeks, and could start work **within two weeks** if a job had been available last week.
Did **at least 15 hours** of unpaid work on a family farm or in a family business.	Did **at least 15 hours** of unpaid work in a family-owned enterprise operated by someone in their household.	Did **any** unpaid work for at least one hour for a business they or a family member owns.	Did not work at all during the reference week, were waiting to be called back to a job from which they had been laid off, and were available for work except for temporarily ill.	Were not working and were waiting to be called back to a job from which they had been laid off (they need not be looking for work to be classified as unemployed).	Waiting to start a job within 3 months but could start work within 2 weeks if a job had been available last week.

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014; U.S. Department of Commerce, Census Bureau, 2012 ACS; U.S. Department of Commerce, Census Bureau, 2012 Current Population Survey (CPS).

Within each adjustment cell, the precompositing calibration adjustment is as follows:

$$F_{5l} = \frac{S_1^*}{S_R + S_L},$$

where S represents the sum of the BQ nonresponse adjusted weights over records in the same adjustment cell as person l. The persons are classified as R: respondent and L: literacy-related nonrespondent. S_1^* is the control total for the adjustment cell. The resulting factors ranged from 1.07 to 1.51 for the area sample within the 400 tracts, from 0.83 to 1.22 for the area sample outside the 400 tracts, and from 1.19 to 1.46 for the list sample. The precompositing calibration weight is calculated as $W_l F_{3l} F_{4l} F_{5l}$.

8.1.3.1.2 Compositing Procedure

The compositing was performed within the following domains: adults sampled as unemployed age 16 to 34 within the 400 tract stratum, adults sampled as unemployed age 35 to 65 within the 400 tract stratum, and the remaining sample. Let W_{lc}^{AS} denote the precompositing calibration weight for the area sample and W_{lc}^{LS} denote the precompositing calibration weight for the list sample for person l in domain c. Composite weights ($\widetilde{W}_{lc}^{NS}$) were created from the two sets of weights as follows:

$$\widetilde{W}_{lc}^{NS} = \alpha_c^{NS} W_{lc}^{AS} \mathrm{I}_{AS}(l) + (1 - \alpha_c^{NS}) W_{lc}^{LS} \mathrm{I}_{LS}(l),$$

where the term α_c^{NS} is the compositing factor for domain c, AS stands for National Supplement area sample, LS stands for National Supplement list sample, NS stands for National Supplement, and

$$\mathrm{I}_{AS}(l) = \begin{cases} 1 \; if \; l \text{ in the National Supplement area sample} \\ 0 \; otherwise \end{cases}$$

$$\mathrm{I}_{LS}(l) = \begin{cases} 1 \; if \; l \; in \; the \; National \; Supplement \; list \; sample \\ 0 \; otherwise \end{cases}.$$

This method produces unbiased estimates for any value of the compositing factor. The optimum value is the one that results in the lowest variance. For a particular estimate Y, the optimum value would be calculated as:

$$\alpha_{Yc}^{NS} = \frac{V(\hat{Y}_c^{LS})}{V(\hat{Y}_c^{AS}) + V(\hat{Y}_c^{LS})},$$

where $V(\hat{Y}_c^{AS})$ represents the variance of the estimate of Y in domain c for the area sample and $V(\hat{Y}_c^{LS})$ represents the variance of the estimate of Y in domain c for the list sample. Given that proficiency is the main estimate of interest, and the scores for the National Supplement were not available at the time of weighting, the value of α_{Yc}^{A} was unknown.

One main quantity that greatly affects the variance is the sample size (n). In the context of simple random sampling (SRS), the variance is inversely proportional to the sample size. This property leads to the following compositing factor:

$$\alpha_c^{NS} = \frac{n_c^{AS}}{n_c^{AS} + n_c^{LS}}.$$

To reflect the complex sample design of PIAAC, effective sample size was used in place of sample size. Effective sample size is the sample size divided by the design effect. Under the assumption that the design effect due to clustering was similar for the different samples, the design effect was estimated as the design effect due to unequal weighting, which is calculated as $1 + CV(W_{lc}^{AS})^2$ for the area sample and $1 + CV(W_{lc}^{LS})^2$ for the list sample. The resulting compositing factor was 0.38 for unemployed age 16 to 34 within the 400 tract stratum, 0.35 for unemployed age 35 to 65 within the 400 tract stratum, and 1 otherwise. The composited National Supplement weight $\widetilde{W}_{lc}^{NS}$ is calculated as defined above.

8.1.3.2 Compositing the National Supplement and Main Study Samples

After combining the area and list samples, the next step in the weighting process was to composite the resulting sample with the Main Study.

8.1.3.2.1 Pre-Compositing Calibration

Prior to compositing, the weights for the Main Study and National Supplement were each calibrated (poststratified) to population totals. Twelve poststratification cells were formed for the Main Study by crossing the final age-employment status classification (unemployed age 16 to 34, unemployed age 35 to 65, not unemployed age 16 to 34, not unemployed age 35 to 65) with race/ethnicity (Hispanic, not Hispanic Black, and other). For the National Supplement, 15 cells were formed by crossing the final age-employment status classification (unemployed age 16 to 34, unemployed age 35 to 65, not unemployed age 16 to 34, not unemployed age 35 to 65 and age 66 to 74) with race/ethnicity (Hispanic, not Hispanic Black, and other). The control totals were based on data from the 2012 ACS 1-year Public Use Micro Sample (PUMS), after removing institutional group quarters and active duty military in noninstitutional group quarters. BQ data were used for this adjustment because employment status was not collected in the Main Study Screener.

Within each adjustment cell, the calibration adjustment prior to Main Study and National Supplement compositing was as follows:

$$F_{6l} = \frac{S_2^*}{S_R + S_L},$$

where S represents the sum of the Main Study BQ nonresponse adjusted weights or the National Supplement composited weights over records in the same adjustment cell as person l. The persons are classified as R: respondent and L: literacy-related nonrespondent. S_2^* is the control total for the adjustment cell. The resulting factors ranged from 0.89 to 1.44 for the Main Study and from 0.93 to 48.98 for the National Supplement. The large adjustment factors in the National Supplement were for the three cells

corresponding to not unemployed age 35 to 65. The sample sizes for these groups were small in the National Supplement given that they were not part of the target population but came in through misclassification. The large calibration factors were offset by small compositing factors in the next step. The calibration weight prior to Main Study and National Supplement compositing was calculated as $W_l F_{3l} F_{4l} F_{6l}$ for the Main Study and $\widetilde{W}_{lc}^{NS} F_{6l}$ for the National Supplement.

8.1.3.2.2 Compositing Procedure

The compositing of the Main Study and National Supplement samples was performed within the following domains: unemployed age 16 to 65, not unemployed age 16 to 34, not unemployed age 35 to 65, and age 66 to 74. Age and employment status were based on the final values from the BQ, with missing values imputed as described in section 8.1.1.1 Let W_{ld}^{MS} denote the precompositing calibration weight for the Main Study and W_{ld}^{NS} denote the precompositing calibration weight for the National Supplement for person *l* in domain *d*. Composite weights ($\widetilde{W}_{ld}$) were created from the two sets of weights as follows:

$$\widetilde{W}_{ld} = \alpha_d \, W_{ld}^{MS} I_{MS}(l) + \left(1 - \alpha_d\right) W_{ld}^{NS} \mathrm{I}_{NS}(l) \,,$$

where the term α_d is the compositing factor, MS stands for Main Study, NS stands for National Supplement, and

$$\mathrm{I}_{MS}(l) = \begin{cases} 1 \; if \; l \text{ in the Main Study sample} \\ 0 \; otherwise \end{cases}$$

$$\mathrm{I}_{NS}(l) = \begin{cases} 1 \; if \; l \; in \; the \; National \; Supplement \; sample. \\ 0 \; otherwise \end{cases}$$

The compositing factor is

$$\alpha_d = \frac{n_d^{MS}}{n_d^{MS} + n_d^{NS}},$$

where *n* represents the effective sample size as described in section 8.1.3.1.2. The resulting compositing factor was 0.46 for unemployed age 16 to 65, 0.59 for not unemployed age 16 to 34, 0.97 for not unemployed age 35 to 65, and 0 for age 66 to 74. The composited weight $\widetilde{W}_{ld}$ is calculated as defined above.

8.1.4 Post-Compositing

After compositing, steps relating to trimming of extreme weights and calibration to external population controls were performed. Careful planning of the sample design can limit the need for weight reduction procedures. However, even a carefully designed sample cannot fully prevent the need for reducing extreme weights. The use of nonresponse and calibration adjustments also introduces variation in sampling weights.

In general, trimming procedures introduce some bias into the sampling weights (Lee 1995). However, as Lee discusses, the trimming adjustment in most cases will reduce the sampling error component of the overall mean square error more than it increases the bias when the adjustment is applied to only a very small number of weights. To limit the number of weights affected by this adjustment across the international PIAAC sample, the PIAAC Consortium developed a trimming threshold that incorporates the variation of the weights so that the amount of bias is not only limited but also kept uniform across the participating country samples. This threshold and details of the trimming adjustment are presented in section 8.1.4.2.

Undercoverage of the target population occurs when some population units are not included in the sampling frame and have no chance of being selected into the sample. Almost all surveys are subject to some amount of undercoverage, and PIAAC is no exception. Calibration is commonly used in sample surveys to reduce the mean square error of estimates and to create consistency with statistics from other studies. Respondents who completed the BQ were included in the calibration. Literacy-related nonrespondents were also included because they are part of the target population from which the control totals were derived and were part of the analysis file. Variables critical to the weighting were recoded and imputed, as necessary, before the calculation of base weights as described in section 8.1.1.

A raking procedure (i.e., iterative poststratification) was used for the calibration of the composited weights. In raking, categories are formed from certain variables, and the weights are calibrated to control totals for each category. In some instances, such cross-tabulations may contain sparse cells, or population distributions may be known for the marginal but not the joint distributions for variables used to define the weighting classes. Typically, raking is conducted when the control totals for interior cells of a cross-tabulation are unknown or sample sizes in some cells are too small for efficient estimation. Raking is related to poststratification in that it poststratifies (or calibrates) to marginal population totals of several variables (or raking dimensions) in an iterative manner. Oh and Scheuren (1987) provide a concise description of the raking procedure and its properties.

Raking and trimming were performed in an iterative process. Raking brings weights into alignment with independent control totals by adjusting the weights in an iterative process to each dimension's set of control totals. In this raking process, some of these weights may become especially distorted as each raking iteration readjusts the weights to meet the control constraints. Combining raking with trimming reduces the larger weights generated in the raking process. After trimming, though, the trimmed weights no longer match the control constraints, and another raking process may be necessary. The resultant weights meet the control constraints (add to the control totals), and at the same time no weights exceed bounds.

8.1.4.1 Pre-Trimming Calibration

A pretrimming raked weight was calculated for each respondent as follows. Let N_γ denote the population count in the raking dimension category γ as obtained from the 2012 ACS, as discussed in section 8.1.1.3. Let $\widehat{N}_{1\gamma}$ be the corresponding survey estimate obtained by using the survey weights prior to raking (as calculated below):

$$\widehat{N}_{1\gamma} = \sum_{i=L,R(\gamma)} \widetilde{W}_i,$$

where

$\widetilde{W}_l$ = the composite weight for person *l*, reflecting all weighting adjustments prior to raking or trimming; and

$L,R_{(\gamma)}$ = the set of literacy-related BQ nonrespondents and BQ respondents in raking dimension category γ.

The initial adjustment factor for raking dimension category γ is given by $F_{7\gamma} = N_\gamma / \widehat{N}_{1\gamma}$. The same process was applied for each raking dimension, each time using the adjusted weights from the previous dimension. This was done iteratively until the sums of the adjusted weights equal all control totals. The raking processes all converged in fewer than 60 iterations.

For simplicity, the raking factor can be denoted as $F_{7\gamma}$, where γ can denote each of the interior cells defined by the raking dimensions shown in table 8-3. The resulting factors ranged from 0.40 to 3.06. The initial calibrated composite weight is calculated as $\widetilde{W}_l F_{7\gamma}$.

8.1.4.2 Trimming

The *Technical Standards and Guidelines* included a guideline permitting the trimming of extreme weights. To account for the different sample designs, trimming was performed separately within domains defined by the sample and sampling domain. The following five trimming domains were formed: the Main Study sample; unemployed age 16 to 65 in the National Supplement area sample; not unemployed age 16 to 34 in the area sample; age 66 to 74 in the area sample; and the National Supplement list sample.

The trimming method implemented and recommended by Consortium statisticians was to trim weights that were over 3.5 times the median weight within a domain. The Consortium further recommended that the trimming threshold incorporate the variation of the weights so that the amount of trimming conducted across the international sample is limited. The impact of a weight trimming process is an increase to bias due to trimming weights, and a decrease to the variation in the weights. As such, after the initial calibration, the trimming threshold in each domain was calculated as follows in order to result in a limited number of cases trimmed and at the same time achieve comparability across countries:

$$3.5\sqrt{1 + CV(\widetilde{W}_l F_{7\gamma})^2}\ ,$$

where the CV is the coefficient of variation. The standard threshold would have resulted in an excessive number of cases trimmed for the list sample. Therefore, the list sample threshold was increased so that less than 5 percent of the sum of weights was trimmed. The resulting values for the thresholds were 4.03 for the Main Study, 4.48 for unemployed age 16 to 65 in the area sample, 4.31 for not unemployed age 16 to 34 in the area sample, 3.73 for age 66 to 74 in the area sample, and 34.73 for the list sample. The cutoff value for domain *t* (c_t) was then calculated as the threshold multiplied by the median weight (the 50th percentile of the weights in the domain). Weights over this value were trimmed to the cutoff, and the trimmed portion of the weights redistributed to the untrimmed cases within the domain, so that the overall

sum of weights for the domain remained unchanged. The extra redistribution step was necessary since more trimming was needed for the unemployed domains, and employment status was not used in the final calibration. The trimming factor was calculated as

$$F_{8l} = \begin{cases} \dfrac{c_t}{\widetilde{W}_l F_{7\gamma}} \; if \; \widetilde{W}_l F_{7\gamma} > c_t \\ \dfrac{\sum_{u \in T'} \widetilde{W}_u F_{7\gamma} + \sum_{u \in T} (\widetilde{W}_u F_{7\gamma} - c_t)}{\sum_{u \in T'} \widetilde{W}_u F_{7\gamma}} \; if \; \widetilde{W}_l F_{7\gamma} \leq c_t, \end{cases}$$

where u represents a case in the same domain as person l, T is the set of cases in domain t for which $\widetilde{W}_l F_{7\gamma} > c_t$ and T' is the set of cases in domain t for which $\widetilde{W}_l F_{7\gamma} \leq c_t$.

Seventy-one cases were trimmed, with their trimming factors ranging from 0.39 to just below 1.00. The trimmed composite weight is calculated as $\widetilde{W}_l F_{7\gamma} F_{8l}$.

8.1.4.3 Post-Trimming Calibration

After trimming, the survey estimate obtained by using the trimmed survey weights was calculated as:

$$\widehat{N}_{2\gamma} = \sum_{i=L,R(\gamma)} \widetilde{W}_l F_{7\gamma} F_{8l},$$

where

$\widetilde{W}_l F_{7\gamma} F_{8l}$ = the composite weight for person l, reflecting all weighting adjustments prior to the second iteration of raking; and

$L,R(\gamma)$ = the set of BQ respondents and literacy-related BQ nonrespondents in raking dimension category γ.

The initial adjustment factor for raking dimension category γ is then given by $F_{9\gamma} = N_\gamma / \widehat{N}_{2\gamma}$. The same process was applied for each raking dimension, each time using the adjusted weights from the previous dimension. This was done iteratively until the sums of the adjusted weights equal all control totals. The raking processes all converged in fewer than 40 iterations.

For simplicity, the raking factor can be denoted as $F_{9\gamma}$, where γ can denote each of the interior cells defined by the raking dimensions shown in table 8-3. The resulting factors ranged from 0.98 to 1.06. The final composite weight is calculated as $\widetilde{W}_l F_{7\gamma} F_{8l} F_{9\gamma}$.

8.2 Prison Sample Weighting

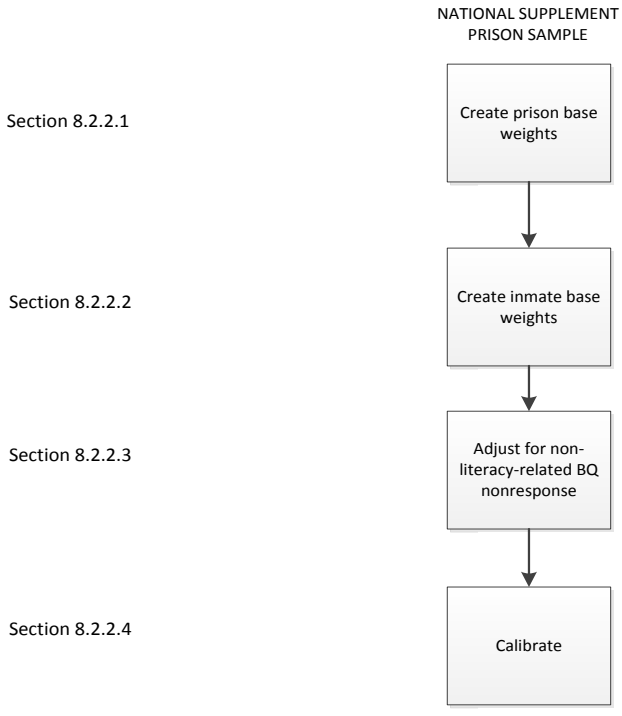
Sample weights were produced for all sample inmates who completed the BQ as well as those who did not complete the BQ due to language problems or learning/mental disabilities. As discussed in Kalton and Flores-Cervantes (2003), the main purpose of calculating sample weights is to permit inferences from

sample persons to the population from which they were drawn. The sample weighting process is designed to accomplish the following objectives:

1. permit unbiased estimates, taking into account the fact that all inmates in the population did not have the same probability of selection;
2. minimize the potential bias arising from differences between respondents and nonrespondents;
3. use auxiliary data on known population characteristics in such a way as to reduce sampling errors and to calibrate weights to the dimensions of the population totals;
4. reduce the variation of the weights and prevent a small number of observations from dominating domain estimates; and
5. facilitate sampling error estimation under the complex sample design.

To make valid inferences from the responding inmates to the target population, the sample must be weighted to account for the special sample design features as well as other complexities arising from nonresponse. First, prison base weights were assigned. Then, a nonresponse adjustment was made to the prison base weights to account for nonparticipating prisons. Next, inmate base weights were computed using the prison nonresponse-adjusted weight and the within-prison sampling rate. Then, an adjustment for nonresponse to the BQ was conducted. Finally, the nonresponse-adjusted inmate weights were raked to control totals based on data provided by the Bureau of Justice Statistics to account for inmate nonresponse and noncoverage. Figure 8-3 illustrates the weighting process for the PIAAC prison sample.

Figure 8-3. Weighting process for the National Supplement prison sample



SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Several steps were completed prior to the weighting process. These preliminary steps are discussed in section 8.2.1. Section 8.2.2 provides a description of the weighting adjustments.

Simple formulas that assume simple random sampling for variance estimation were not appropriate under the Prison Study sample design. Even if sampling weights were used to construct the survey estimates, inferences would not be valid unless the corresponding variance estimator appropriately reflected all the complex features of the sample design. Estimates of variance were facilitated through the use of the jackknife replication method (discussed in section 8.3) and can also be estimated through Taylor series (linearization).

A standardized weighting system was used to perform the tasks of nonresponse adjustment, raking, and the creation of replicate weights for variance estimation. The system has been used on numerous surveys for NCES and has also been adapted to conform to PIAAC quality control checks according to the PIAAC *Technical Standards and Guidelines*.[48] Quality control checks were embedded in the weighting system, as discussed in section 8.2.5. For example, after each weighting step, the weighted totals and percent distribution for several key domains were compared to the weighted totals and percent distribution prior to the weighting step. This allowed the statistician to identify any large changes in distribution and investigate the reasons for the changes.

8.2.1 Preliminary Steps in Weighting

Prior to weighting, the variables considered for the weighting adjustments were evaluated. Only variables of high quality, available for all eligible units, and related to literacy and response propensity were considered for the nonresponse adjustment. For the final calibration adjustment, variables were required to have reliable control totals available from an external source and have less than 5 percent missing data for PIAAC respondents.

8.2.1.1 Imputation for Weighting Variables

Imputation was conducted for variables used in the weighting process. Age, race/ethnicity, and gender were collected in the Screener and the BQ. Gender had no missing values. Age had a small number of missing values, but they were deduced with almost certainty using administrative information from the prisons. The race/ethnicity variable was coded as Hispanic, non-Hispanic Black, or Non-Hispanic Other. For the seven cases with missing race/ethnicity, ethnicity (Hispanic/not Hispanic) was imputed first, and then race for non-Hispanics. To obtain values for ethnicity, cells were formed by prison ID and language spoken as determined by the Screener questionnaire. Then a hotdeck procedure[49] was used to assign the value from a random donor within the cell to the missing case. To obtain values for race, cells were formed by prison ID, language spoken, and imputed ethnicity.

[48] The PIAAC *Technical Standards and Guidelines* are available at http://www.oecd.org/site/piaac/PIAAC-NPM(2014_06)PIAAC_Technical_Standards_and_Guidelines.pdf. The guidelines for the weighting standards are in chapter 14.
[49] Hotdeck is an imputation procedure that uses data from the same sample survey.

8.2.1.2 Selection of Weight Adjustment Variables

In accordance with the PIAAC *Technical Standards and Guidelines*, an initial nonresponse bias analysis was conducted to evaluate the relationship between the potential nonresponse adjustment variables and response propensity. The analysis was conducted at the inmate level, using Search, which employs a hierarchical tree algorithm that was introduced in section 8.1. The software identifies the auxiliary variables that best define subgroups with differential response rates. The analysis was performed using inmate base weights to reflect any differential sampling rates. To the extent that the auxiliary variables are related to literacy, the differential response rates among the subgroups (or differences in the distributions of respondents and nonrespondents) may indicate potential nonresponse bias in the unadjusted PIAAC estimates. Using these variables in the weighting adjustments should help alleviate the presence of bias.

A BQ-level nonliteracy-related nonresponse status variable was created and used as the dependent variable for the analysis. The following prison-level variables were available as predictors from the sampling frame:

- region (Northeast, Midwest, South, West);
- type (federal/state/other);
- security level (supermax, maximum/close/high, medium, minimum/low); and
- gender (male, coed, female).

The following person-level weighting variables come from the prison administrative data and the self-reported BQ data for respondents:

- age category (<30 year; 30–49 years; >= 50 years);
- gender (male; female); and
- race/ethnicity (after imputation).

The control totals for age, race/ethnicity, and gender were provided by the Bureau of Justice Statistics based on their estimates of prison population (section 8.2.2.4).

8.2.2 Weighting Adjustments

The PIAAC prison-level weighting process involved the prison base weight creation (section 8.2.2.1), and the subsequent inmate weighting adjustments (sections 8.2.2.2 through 8.2.2.4). During the inmate nonresponse adjustment, particular attention was given to classifying nonresponse into literacy-related and nonliteracy-related nonresponse categories. All nonliteracy-related nonrespondents were considered similar to respondents with respect to literacy scores; however, the literacy-related nonrespondents were not considered similar to the respondents since they likely would have scored lower than average if they attempted the assessment. For the BQ, literacy-related nonresponse occurred if the sample person could not speak the language of the interview or could not complete the interview due to reading or writing difficulty or learning/mental disability. The weights of the respondents were adjusted to account for the nonliteracy-related nonrespondents, while the literacy-related nonresponse cases retained their weights

during the nonresponse adjustment process. In the final calibration process, the weights of the literacy-related nonrespondents and the respondents were adjusted to align with the control totals.

8.2.2.1 Prison Base Weights

Prison base weights were assigned to all sampled prisons as the reciprocal of the prison selection probability. Prisons were selected with probability proportional to size, and therefore the prison base weights varied with the size of the prison. The prisons were also selected within strata defined by the gender of the inmates in the facility (male/coed, female). The prison base weight was assigned as follows:

$$W_{hi} = \frac{1}{P_{hi}},$$

where P_{hi} is the probability of selecting facility *i* from stratum *h*.

The prison base weights W_{hi} ranged from 1.005 to 64.812. The prison-level response rate was 100 percent after the two nonresponding prisons, out of the 98 eligible prisons, were replaced by their substitutes (discussed in section 3.2). Therefore, a prison level nonresponse adjustment was not needed.

8.2.2.2 Inmate Base Weights

The next step in the weighting process was to create the inmate-level base weight. The inmate base weights account for the prison selection probability and for the different inmate selection rates across prisons. The full-sample inmate base weight for inmate j of prison i was computed as the product of the prison nonresponse-adjusted base weight and the reciprocal of the inmate sampling rate, as given by

$$W_{hij} = W_{hi}\frac{1}{P_{j|hi}},$$

where $P_{j|hi}$ is the probability of selecting inmate *j* from facility *i* in stratum *h*. The inmate base weights W_{hij} ranged from 190.3 to 307.7 for the all-female prison stratum and from 229.6 to 1231.0 for the male/coed prison stratum. Variation in the inmate base weights was due to oversampling all-female prisons, and due to assigning a minimum (12) and maximum (19) number of inmates to sample per prison. Exceeding the bounds on inmate sample size occurred when differences existed between the sampling frame inmate counts, updated inmate counts at time of negotiations with prisons, and the inmate counts determined through the within-prison sampling procedure conducted during data collection.

8.2.2.3 Nonresponse Adjustment

An adjustment was made to distribute the inmate base weights of the nonliteracy-related nonrespondents to the respondents. The nonresponse adjustment was performed within cells that were defined based on the analysis described in section 8.2.1.2. The cells were required to adhere to certain criteria. First, since it is generally preferable to redistribute the weights of nonrespondents over a relatively large pool of cases, the minimum cell size allowed was 30 cases. Second, variation in nonresponse adjustment factors

typically increases the variation in the weights, which generally increases the variance of the survey estimates. Since the amount of variation in the nonresponse adjustment factors was a concern, the maximum allowable nonresponse adjustment factor was 2.0.

Within each weighting cell, the nonliteracy-related nonresponse adjustment factor was computed as follows:

$$F_{2j} = \begin{cases} 1 & \text{if } l \in L \\ \dfrac{S_R + S_{NR}}{S_R} & \text{if } l \in R \\ 0 & \text{if } l \in NR \end{cases},$$

where *S* represents the sum of weights over records in the same adjustment cell as inmate *hij*. The factors and weights shown here are for an inmate *hij* classified as R: respondent, L: literacy-related nonrespondent, I: ineligibles or NR: nonliteracy-related nonrespondent. The resulting factors ranged from 1.02 to 1.54, with the average adjustment factor being 1.17. The BQ nonresponse adjusted weight is calculated as $W_{hij}F_{2j}$.

8.2.2.4 Calibration

In the weighting process for the prison sample, the last step was to calibrate the inmate nonresponse adjusted weights to external control totals. Raking is a special case of the broader set of calibration procedures, and it is also known as iterative poststratification. It assumes the existence of a set of marginal totals for selected raking dimensions. After raking, the sample estimates for the selected variables will equal the known totals to reduce bias due to noncoverage.

Sampled inmates completing the BQ were included in the raking process. Literacy-related nonparticipants (those with a language problem, reading and writing difficulty, or learning/mental disability) were also included to account for nonsampled inmates who were not able to complete the BQ due to literacy-related reasons. The control totals (the prison population estimates form 2013) were provided by the Bureau of Justice Statistics (BJS). The totals were estimated from multiple sources of administrative data and survey data. Table 8-5 shows the prison population estimates for 2013 from BJS and the sample totals before calibration. The total prison population provided by BJS was about 15 percent more than the sample total before calibration. Since the prison population provided by BJS might have included inmates in facilities that were outside of the scope of the prison study, it was decided to use BJS percentages in the calibration process. More specifically, the control totals were computed by multiplying the sample total sum of weights before calibration, 1,226,009, by the percentages by age, sex, and race from BJS. After calibration, the sample total remained the same, but the distributions by age, sex, and race matched the percentages provided by BJS.

Table 8-5. Comparison of 2013 prison population estimates and sample estimates before calibration

	2013 population estimate		Sample totals before calibration	
	Count	Percent	Count	Percent
Total prison population[1]	1,420,600		1,226,009	
Age				
<30 years	417,100	29.36	366,204	29.87
30–49 years	765,000	53.85	650,643	53.07
≥ 50 years	238,400	16.78	209,163	17.06
Sex				
Male	1,325,400	93.30	1,153,783	94.11
Female	95,200	6.70	72,227	5.89
Race				
Black, non-Hispanic	518,000	36.47	436,256	35.58
Hispanic, any race	312,200	21.98	249,557	20.36
Other, non-Hispanic	590,300	41.55	540,196	44.06

[1] Number of prisoners in the custody of state or federal facilities or privately operated facilities that primarily hold state or federal inmates; excludes prisoners in community-based facilities.
NOTE: Prison population estimates for 2013 were rounded to the nearest 100.
SOURCE: 2013 Bureau of Justice Statistics prison population estimates.

A raking procedure (i.e., iterative poststratification) was used for the calibration of the nonresponse adjusted weights. Let $\widetilde{N}_\gamma$ denote the population control total in the raking dimension category γ as computed by multiplying the sample total sum of weights before calibration by the percentage of the raking dimension category γ from BJS. Let $\widehat{N}_\gamma$ be the corresponding survey estimate obtained by using the survey weights prior to raking (as calculated below):

$$\widehat{N}_\gamma = \sum_{hij=L,R(\gamma)} W_{hij} F_{2j} \ ,$$

where

$L,R_{(\gamma)}$ = the set of literacy-related BQ nonrespondents and BQ respondents in raking dimension category γ.

The initial adjustment factor for raking dimension category γ was given by $F_{3\gamma} = \widetilde{N}_\gamma / \widehat{N}_\gamma$. The same process was applied for each raking dimension, each time using the adjusted weights from the previous dimension. This was done iteratively until the sums of the adjusted weights equaled all control totals. The raking processes all converged in 4 iterations.

For simplicity, the raking factor can be denoted as $F_{3\gamma}$ where γ denotes each of the interior cells defined by the raking dimensions. The calibrated inmate weight is calculated as $W_{hij} F_{2j} F_{3\gamma}$. The resulting factors ranged from 0.93 to 1.28. The resulting calibrated weights ranged from 490 to 1643 for male inmates and from 238 to 441 for female inmates. Trimming was not done since the variation in the calibrated weights was not large.

8.3 Replicate Weights

There are two commonly used approaches for estimating variances for complex surveys: replication and Taylor series (linearization). The replication approach was used for PIAAC due to the need to accommodate the complexities of the sample design, the generation of plausible values, and the impact of the weighting adjustments. The software developed for the analysis of PIAAC data, such as the PIAAC Data Explorer and the IDB Analyzer, uses the replicate weights for generating valid variance estimates.

Replication methods were applied to surveys by dividing the sample into specially designed replicate subsamples that mirrored the design of the full sample. To form the replicate subsamples, variance strata and variance units were defined. Each subsample was reweighted to account for the subsampling that occurred. For each subsequent stage of weighting, adjustments made to the full-sample weights were also made to the replicate weights. As a result of these adjustments, the effects of the weighting procedures were properly reflected in variance estimates. As described below, 45 replicate weights were formed for the household sample (section 8.3.1) and the prison sample (8.3.2). Since the PIAAC Data Explorer required 80 replicate weights, the remaining 35 replicates were created by simply setting each to the set of full-sample weights.

An estimate was then calculated for the full sample and each of the replicate subsamples. The variance of the full sample estimate was computed as the sum of squared deviations between each replicate subsample estimate and the full sample estimate. We used the paired jackknife approach (commonly referred to as JK2 and described in Rust and Rao [1996]) with two variance units per stratum for the prison sample. The sampling error variance estimator with this approach takes the following form:

$$Var(\hat{\theta}) = \sum_{i} (\hat{\theta}_g - \hat{\theta}_0)^2,$$

where

$\hat{\theta}_0$ = full sample estimate; and

$\hat{\theta}_r$ = estimate for replicate r.

8.3.1 Household Sample Replicate Weight Formation

As mentioned in section 3.1.2, the 80 PSUs selected for the PIAAC Main Study were used for both the area sample component and list sample component of the National Supplement. The same segments selected for the Main Study were also used in the area sample of the National Supplement. However, the list sample consisted of 80 new segments, with one selected from each of the 80 PSUs.

Since PSUs selected with certainty were self-representing, variance strata and units were formed differently for these PSUs than for the nonself-representing PSUs. Each self-representing PSU formed its own variance stratum. To reflect the degrees of freedom in variance estimation appropriately, each large self-representing PSU was split into two or more variance strata to reflect its measure of size (MOS). The number of variance strata formed by each PSU was based on the size of the PSU relative to the other PSUs on the sampling frame, or the number of times the PSU was "hit" during sample selection. As a result, one PSU comprised three variance strata, another comprised two, and the two remaining self-representing PSUs comprised only one variance stratum each.

The variance units were formed differently for the self-representing PSUs in the Main Study and area sample than for the list sample. For the Main Study and area sample, segments were paired in the order in which they were selected and assigned alternately to the assigned variance strata. Within the variance strata assigned to the self-representing PSUs, variance units were the segments. Segments were assigned to variance units 1 and 2 alternately according to their selection order within each variance stratum. For the list sample, since there was only one segment per PSU, the DUs in each segment were assigned to a number of groups alternatively in the order in which they were selected. The number of groups depended upon the number of variance strata the PSU comprised. The groups were then randomly assigned to variance units 1 and 2 within each variance stratum.

The remaining 76 PSUs (nonself-representing) were assigned to variance strata by pairing the PSUs according to the nested stratification scheme to form 38 variance strata, for a total of 45.

Once the variance strata and variance units were assigned, the replicate weights were created. With JK2, one variance unit is dropped from each stratum to form each replicate. That is, the first set of replicate weights is formed by taking the Screener base weights and setting all the weights in a randomly selected variance unit (1 or 2) of variance stratum 1 to zero, doubling the weights of the cases in the other variance unit in the variance stratum and keeping all other weights unchanged. For PIAAC, this process continued over the 45 variance strata for a total of 45 replicate weights.

8.3.2 Prison Sample Replicate Weight Formation

To create the replicates, the 98 eligible prisons (two sampled prisons were closed) were sorted in their order of selection. Prisons were paired consecutively and assigned to 49 variance strata. To be consistent with the household sample, some variance strata were combined to obtain a total of 45 replicate weights. Once the variance strata and variance units were assigned, the replicate weights were created. With JK2, one variance unit is dropped from each stratum to form each replicate. Within each variance stratum, one prison was randomly assigned to variance unit 1 and the other to variance unit 2. Replicates were then created using the paired jackknife approach as described in the previous section.

8.3.3 Accounting for the Imputation Error Variance Component

In addition to the sampling variance component as described in sections 8.3.1 and 8.3.2, another procedure that affects the variances, is the generation of plausible values using item response theory (IRT) models (Birnbaum 1968; Lord 1980). Because different respondents take different sets of items that could be different in level of difficulty, it would be inappropriate to base the competency estimates simply on the number of correct answers obtained. Therefore, large-scale assessments using matrix sampling rely on IRT models. The IRT model uses the item responses for each individual and regards the latent literacy score as random. Several predicted values, referred to as plausible values, are generated from the IRT model and the variation among them captures the measurement error.

For estimation using plausible values (PVs), calculations must account for both the sampling error component and the variance due to imputation of proficiency scores. The estimator of the population mean is the average of the *M* PV means,

$$\hat{\bar{Y}}^* = \sum_{m=1}^{M} \hat{\bar{Y}}_m / M.$$

The variance of the estimated mean $\hat{\bar{Y}}^*$ is computed using formulas specific to PVs as follows:

$$v\left(\hat{\bar{Y}}^*\right) = U^* + B\left(1 + \frac{1}{M}\right)$$

where, the "within" variance component is computed as the average of the sampling variance for each of the M plausible values, computed as

$$U^* = \left(\sum_{m=1}^{M} U_m\right) / \mathrm{M}$$

where the sampling variance of the estimated mean $\hat{\bar{Y}}_m$ for plausible value m is U_m, and where, the "between" component is calculated as

$$B = \left[\sum_{m=1}^{M}\left(\hat{\bar{Y}}_m - \hat{\bar{Y}}^*\right)^2\right] / \text{(M-1)}$$

where, the mean of each of the *M* PVs $y_{l1}, y_{l2}, \ldots, y_{lM}$. for sample unit l is computed as

$$\hat{\bar{Y}}_m = \sum_{l \epsilon s} w_l y_{lm} / \sum_{l \epsilon s} w_l \; ; m = 1, \ldots, M,$$

where s denotes the set of sample units.

The standard error is computed as the square root of the total variance, $\sqrt{v\left(\hat{\bar{Y}}^*\right)}$.

This page intentionally blank.

9. DATA PREPARATION AND PROCESSING

9.1 Introduction

The Consortium required that Main Study data preparation and processing be performed in a uniform way within and across countries and with an acceptable quality level. Key data preparation tasks that ensured this uniformity consisted of manual data entry of scoring sheets, generation and review of edits on computer generated data files, management of coding, scoring of related files, validation of the structural consistency of the database, and delivery of the national database to the Consortium.

As outlined in chapter 2, sections 2.8 and 2.9, the Consortium-provided Data Management Expert (DME) software was used to perform many of these data preparation and processing activities. The Consortium provided each country with the DME software, which was used to assemble, manage, verify, and edit each country's national database. The national DME database consisted of two parts: (1) data collected by the virtual machine's processing of the Background Questionnaire (BQ) and the computer-based assessment items or tests administered on the interviewer laptops, and (2) scoring data entered manually and generated as the result of scoring the paper-based assessment booklets.

The same data preparation and processing procedures were followed for the National Supplement as those used for the Main Study. An updated version of the Consortium-provided DME was used for all data preparation and processing activities.

9.2 Overview of Data Preparation and Processing Activities

The initial phase of PIAAC data preparation in the U.S. Main Study included the customization of the DME software to accommodate the Consortium-approved U.S. adaptations that were made to the international BQ. Each country was required to map its national dataset into a highly structured, standardized record layout required by the Consortium to facilitate comparability during data analysis. The international record layout specified the position, format, and length of each field, and included a description of each variable, indicating the valid response categories and ranges to be collected for the variable.

Once the adaptations to the DME were completed and approved by the Consortium, the U.S. version of the DME software was deployed to a separate computer loaded with the DME software and secured in a separate, locked area. Files transmitted from the interviewers were processed by the DME software and added to the DME database each morning, using the TAO (Testing Assisté par Ordinateur) data import capabilities. Edit reports contained within the DME were run on a regular basis; issues and discrepancies such as differences in age reported by the Screener respondent and sampled person were reviewed and resolved throughout the data collection period.

Coding of the appropriate verbatim fields was accomplished following the PIAAC *Technical Standards and Guidelines* throughout the data collection period, and coded responses were imported into the DME on a regular basis as coding and verification was completed.

After the completion of the data collection period, hard-copy assessment and reading component booklets were scored by Pearson. The results of the scoring activities were entered into the scoring part of the DME software. The DME scoring database was used by Pearson to generate inter-rater reliability reports

during the paper assessment scoring period and to deliver these reports along with the scoring data to Westat for inclusion in the final U.S. DME database. The DME scoring database was imported into the DME database at Westat after all scoring was completed.

Finally, DME edit reports were run and reconciled and frequencies of the deliverable datasets were run and reviewed to ensure accuracy and completeness of the required delivery products.

The same data preparation and processing activities were followed for the National Supplement.

9.3 Reconciliation of Demographic Information with Screener

Demographic information collected from the household respondent during the Screener was confirmed and updated by the sample person at the start of the BQ. Comparison edits between the demographic information obtained in the BQ and demographic information obtained during screening were conducted as completed interviews were received. When age discrepancies of less than 5 years were observed, it was assumed that the Screener respondent reported the sample person's age in error. The age reported by the sample person completing the BQ was considered to be the accurate response. In cases where the age varied by more than 5 years or there was gender disagreement, the field supervisor validated the case and provided an explanation or clarification of what had occurred. When it was determined that a mistake had been made by the respondent or interviewer (e.g., a keying error), the BQ was updated with the correct date of birth or gender and subsequent fields that would have been affected were also reviewed and updated as appropriate. Edits were performed in accordance with the Consortium requirements and were discussed and agreed upon with field management staff and statisticians prior to their implementation.

The same demographic information reconciliation process was followed for the National Supplement.

9.4 Coding Operations of Required Fields

Throughout the data collection period and continuing 2 weeks after the end of data collection, selected BQ variables were coded following the coding schemes specified by the Consortium. Experienced Westat coding staff were trained to understand the selected international coding schemes' principles and the required coding structure. Training consisted of an item-focused training session, which included written materials and training exercises for each required coding scheme. All coding was 100 percent verified or double coded to ensure accuracy and consistency across coding staff. Coding reports were generated weekly, and error rates exceeding 6 percent triggered review by the coding supervisor and retraining if necessary. As coding was completed for a group of cases, verified codes were imported to the PIAAC DME for delivery to the Consortium along with the other study data.

Coding schemes followed for U.S. PIAAC included the following:

- 1997 International Standard Classification of Education (ISCED), allowing for nonformal education and excluding foreign certification, was used to code the level of education variables for foreign degrees obtained by the respondent. In addition, based on the updated 2011 ISCED categorization of education, additional derived variables were added to the combined Main Study and National Supplement datasets.

- CIP 2010[50] coding scheme, at the four-digit level when possible, was used to code the U.S. adaptations added to capture the area of study for each of the education question series.
- ISO 639-2 alpha-3[51] was the coding scheme used for languages.
- UN M.49[52] coding scheme was used to code the country of earned education and the country of birth.
- 2008 International Standard Classification of Occupations (ISCO-08) was used to code occupations reported in the BQ. Occupational coding was done to the four-digit level when enough information was available. Note that the Consortium analysis (Main Study dataset) was completed at the three-digit level.
- International Standard Industrial Classification of All Economic Activities (ISIC), Revision 4 was followed to assign industry codes. Industry coding was done to the four-digit level when enough information was available. Note that the Consortium analysis (Main Study dataset) was completed at the three-digit level.

A list of the coded variables and their definitions is presented below:

- ISCED_H (Level of highest qualification – Respondent, only coded for the foreign degree level);
- ISCO08_C (Current job – Respondent);
- ISCO08_L (Last job – Respondent);
- ISIC4_C (Industry of current job – Respondent);
- ISIC4_L (Industry of last job – Respondent);
- LNG_L1 (First language learned at home in childhood, still understood – Respondent);
- LNG_L2 (Second language learned at home in childhood, still understood – Respondent);
- LNG_HOME (Language most often spoken at home – Respondent);
- CNT_H (Country in which highest qualification was gained – Respondent);
- CNT_BIRTH (Country of birth – Respondent);
- USCIP_H (Broad field of highest qualification, CIP codes – Respondent);

[50] The Classification of Instructional Programs (CIP) provides a taxonomic scheme that supports the accurate tracking and reporting of fields of study and program completions activity. To see the CIP codes, go to: http://nces.ed.gov/ipeds/cipcode/Default.aspx?y=55.
[51] The ISO 639-2 alpha-3 provides a three-digit alphabetic coding scheme that supports the consistent reporting of languages. To see the ISO 639-2 alpha-3 codes, go to: http://www.loc.gov/standards/iso639-2/langhome.html.
[52] The UN M.49 coding scheme is a standard for area codes used for statistical purposes. The scheme is developed and maintained by the United Nations Statistics Division. To see the UN M.49 codes, go to http://unstats.un.org/unsd/methods/m49/m49alpha.htm.

- USCIP_C (Broad field of currently studied for qualification, CIP codes – Respondent);
- USCIP_L (Broad field of last year qualification, CIP codes – Respondent);
- ISCO08_US (Apprentice – Respondent); and
- REGION_US (U.S. census region of the respondent's residence).

Table 9-1 presents a crosswalk for each of the coded variables listed above and identifies the variable or variables within the BQ and the coding scheme used.

Table 9-1. Coded variables, input variables from Background Questionnaire, and coding scheme used

Coded variables	Variables from Background Questionnaire	Coding scheme used
ISCED_H	B_Q01aUS, B_S01a1, B_Q01a3US	ISCED 1997 (1–14, incl nonformal, excl foreign)
ISCO08_C	D_Q01a, D_Q01b, D_Q08a, D_Q04	ISCO 2008
ISCO08_L	E_Q01a, E_Q01b, E_Q04	ISCO 2008
ISIC4_C	D_Q02a, D_Q02b, D_Q03US	ISIC Rev 4
ISIC4_L	E_Q02a, E_Q02b, E_Q03US	ISIC Rev 4
LNG_L1	J_Q05a1US, J_S05a1	ISO 639-2 alpha-3
LNG_L2	J_Q05a2US, J_S05a2	ISO 639-2 alpha-3
LNG_HOME	J_Q05bUS, J_S05b	ISO 639-2 alpha-3
CNT_H	B_Q01a2US, B_S01a2	UN M.49
CNT_BIRTH	J_Q04a, J_Q04bUS, J_S04b	UN M.49
USCIP_H	B_Q01bUSX, B_S01a1	CIP 2010
USCIP_C	B_Q02cUSX, B_S02b1	CIP 2010
USCIP_L	B_Q05bUSX, B_S05a1	CIP 2010
ISCO08_US	B_Q29bUSX	ISCO 2008
REGION_US	SDIF	U.S. census regions

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

9.5 Processing and Scoring of Hard-Copy Assessment Materials

Approximately 15 percent of the sample persons in the Main Study and 23 percent of the sample persons in the National Supplement household sample completed one or more paper test booklets during their assessment. 36 percent of the sample persons in the National Supplement prison sample completed one or more paper test booklets during their assessment. Pearson, the scoring subcontractor used successfully for the PIAAC Field Test, also scored the paper assessment materials for the Main Study and the National Supplement. The hard-copy assessment materials were maintained securely at Westat and shipped to Pearson for data capture of the responses in the reading component booklets and scoring of the Core, Literacy, and Numeracy items in the exercise booklets. Pearson used the DME software for data capture following the scoring guidelines outlined in the PIAAC *Technical Standards and Guidelines*.

Exercise booklets were shipped to Pearson on a monthly basis. Upon receipt of each shipment Pearson inventoried the received booklets and verified that expected booklets listed on the transmittal sheet were received. Pearson notified Westat that the shipment was received and noted any discrepancies, if found.

Pearson treated the assessment booklets as confidential material and stored them in a secured location at all times.

The reading component booklets were handled separately from exercise booklets containing the Core, Literacy, and Numeracy items. Reading component booklets did not require scoring but rather accurate data capture of selected answers. Pearson completed the data capture each month upon receipt of the reading component booklets. Scorers were trained and then scored the reading component booklets before they were trained, and subsequently scored, the exercise booklets. (This was an approved deviation from Consortium scoring procedures.) Following the approved PIAAC reconciliation procedures, the data-entry function of the DME was used. After data entry, Pearson used the DME software to verify that 100 percent double-data entry was performed and used the DME to identify any inconsistencies between the two sets of data. Any inconsistencies identified were reconciled with the original scoring sheets and the values were corrected in the DME database. Pearson returned electronic files of these data monthly to Westat.

Scoring of the three types of exercise booklets was not undertaken until the end of the data collection period so that all booklets could be scored together. Pearson inventoried and stored the completed exercise booklets received each month until they were ready to score the booklets. Pearson trained the scoring staff to score items in one domain (Core, Literacy, or Numeracy) at a time and then scored the items for that domain immediately following the training. Based on experience during the Field Test, Pearson felt this method resulted in higher scoring efficiency and quality.

After careful review of the PIAAC standards Pearson decided to use the default three-scorer design as outlined in the PIAAC *Technical Standards and Guidelines*. Pearson felt the three-scorer design would result in greater efficiencies during the scoring process. The U.S. also decided to check the consistency of scoring by having a second scorer rescore 100 percent of the items even though this was above and beyond the Consortium requirement. It was determined that because of the small number of exercise booklets to be scored it was more efficient to second score 100 percent than to implement the complicated random selection process required by the Consortium if less than 100 percent of the booklets had a second scorer.

Once Pearson completed the scoring process, the company delivered the scoring data and reports to Westat. The six delivered reports were generated by the DME software, which consisted of two reports for each exercise booklet type (Core, Literacy, and Numeracy). There was an inter-rater log and an inter-rater results report for each exercise booklet type.

The Main Study overall inter-rater reliability for each of the three PIAAC exercise booklet types was as follows:

- Core—Total for all items = 99.3 percent;
- Literacy—Total for all items = 99.4 percent; and
- Numeracy—Total for all items = 98.9 percent.

The National Supplement overall inter-rater reliability for each of the three PIAAC exercise booklet types was as follows:

- Core—Total for all items = 99.6 percent;
- Literacy—Total for all items = 99.7 percent; and

- Numeracy—Total for all items = 99.9 percent.

At the end of the scoring process all booklets were returned to Westat for secure storage and archival. Final scoring data files were delivered to Westat and were imported into a consolidated DME database. Frequencies were created and reviewed as a final quality control of these data.

The same hard-copy assessment scoring process was followed for the National Supplement.

9.6 Reconciliation with Data Management Expert (DME)

Once all data processing, coding, and scoring activities were completed, the data from the individual processes were combined to create the final DME database. During each step of this process, rows were matched and reconciled to ensure consistency between each step. The final database included all data from the virtual machine (BQ and computer-based assessment), coding, and scoring. To ensure that the delivered data were free of errors, the consolidated DME database was verified and edited. Examples of completed checks were ensuring that no variables or case records were dropped during the loading process, looking for data that were loaded into the wrong variable, and ensuring that scoring data were received for every completed paper assessment booklet.

Aggregated data cases were reconciled with the Study Management System (SMS). The reconciled state of the end of each case was compared to the data available for each case and, if needed, editing adjustments were made. Some technical problem cases were discovered by recognizing disparity between the reconciled disposition and the lack of or incompleteness of the available data. As a result, cases with missing data were reviewed closely and in some cases this review yielded the recovery of the missing data.

The first steps in the verification process were to run the consistency edit checks provided by the DME and to review reported errors. Next, additional edits not provided by the DME software were run, including a reconciliation of case statuses and a general frequency review. The SMS reconciliation matched SMS case status codes with the DME status codes by case ID to ensure that all completed cases in the SMS were loaded into the DME. Basic frequencies were generated and reviewed to ensure that reported values fit defined variable ranges or specified response categories for each variable.

In some instances, the review required a manual comparison of the original source data with the data contained in the DME database to guarantee that the data were loaded accurately. In other instances, groups of data were reviewed to ensure that the loading process correctly imported whole blocks of data. If needed, audit trails of individual interviews were checked to determine how the discrepancies might have been generated. Data contained within the SMS, including the electronic records of calls and case comments, were reviewed for additional information to help resolve questions.

Throughout the process, possible errors were investigated, documented, and resolved prior to the delivery of the final dataset to NCES. Remaining discrepancies were documented in the final delivery notes that were delivered along with the final data files.

The same DME reconciliation process was followed for the National Supplement. However, it should be noted that a separate DME model was used for the prison sample to accommodate the BQ changes implemented for the prison sample. The DME for the prison sample was built upon the existing household sample DME, incorporating the additional prison BQ variables. There were no changes to the coding frame or the types of edit checks executed.

9.7 Reconciliation with Study Management System (SMS)

At the end of the data collection period, cases still being worked were reviewed and finalized, or closed by the supervisor. Finalization included a review of the case history and the assignment of a final status code for each incomplete task. Generally, the SMS was the best source for final case statuses, and these results were merged back into the aggregated data in the DME repository.

As noted in section 9.6, the SMS data were reconciled with the aggregated DME database to ensure that all completed cases were loaded accurately. The SMS case history and comments were reviewed thoroughly for problem cases in order to understand the operational circumstances, arrive at a resolution and set the final status within the DME. For example, a case in which the BQ was completed but the assessment was never completed might be listed as "incomplete" in the virtual machine operational data, while the SMS would reflect a different final disposition assigned by the supervisor (for example, Language Problem or Refusal). Due to the complexity of Consortium-implemented workflow and the number of disposition codes available for the various paths of the assessment, accurate reconciliation was complex and required careful scrutiny during the review.

For most completed cases, reconciliation amounted to ensuring that all data were present in the dataset. However, for exceptions, partial cases, or sample persons identified in screening but not interviewed, a more detailed comparison was performed to ensure that the delivered data accurately reflected the final status. The reconciled, finalized status was compared to the data available for each case. Potential problem cases were discovered by recognizing disparity between dispositions and the incomplete data within the DME. When possible, some or all data for problem cases were retrieved from backup files.

The same SMS reconciliation process was followed for the National Supplement.

9.8 Delivery of Data to the Consortium

Following the end of the Main Study field period, a single, complete, and validated database reflecting the U.S. national DME layout was submitted to the Consortium, along with supplemental documentation and log data. Data files were delivered in accord with the expected format required for processing and analysis at the international level.

Before delivery of the final dataset, a confidentiality operation assigned replacement randomized identifiers for each delivered case. Since operational case and person identifiers contained some substantively identifying information, these were overwritten with entirely random identifiers for delivery. This confounded revealing any substantive meaning for cases and prevented matching-back to any remaining operations information. Westat securely retains a cross-walk of this identifier replacement should a legitimate need arise later.

In addition, a few other techniques were implemented to ensure confidentiality within the delivered U.S. dataset beyond those discussed in sections 6.3 and 6.4. Data items that directly identify sample persons (e.g., name, address, and phone number) were suppressed. Also, BQ log data (DME BQL table) was not delivered because these raw data files contain literal recording of all values, including confidential data that cannot be suppressed. As the Consortium supplied no solution for data value suppression within the BQ log data format, and since BQ log data are not fundamental to international data handling or analysis,

Westat excluded these files from the international delivery. This was done with prior Consortium notice and agreement.

The products aggregated for the delivery included the following:

- a single integrated, verified, confidentialized, and exported database in the DME's format and using the adapted national codebooks;
- a single zip archive including all raw TAO XML output files for the computer-based exercise;
- a free-format documentation on "Other" questions in the BQ, summarizing the National Project Manager's review of the response made to the open-ended questions in the BQ that did not require formalized coding;
- a comprehensive and detailed free-format documentation of any implemented confidentiality edits, if any;
- a comprehensive and detailed free-format documentation of any other issues or notes that required attention from the Consortium during data processing and analysis; and
- a free-format documentation with double-coding reliability evidence and explanations for Quality Control purposes. This included both documentation of the achieved double-coding accuracy at the four-digit level for ISCO and ISIC instances and tables in which the PIAAC data were compared to the most recent Labor Force Survey on the distributions of highest level of education, labor force status, current occupation, and current industry.

All data products were posted on the NCES secure FTP site. These were forwarded by NCES staff to the Consortium via the IEA secure server.

During its further review of the Main Study data, IEA identified a BQ flow variation among U.S. adaptations/extensions. A U.S.-adapted data path caused the U.S. BQ to collect more responses in the B section (education) than the international BQ flow intended. After reviewing these issues with NCES, it was decided that the U.S. path was preferred and was therefore implemented in the BQ. IEA removed the extra data points to standardize the U.S. dataset with the international version and Westat delivered a supplemental national delivery file (Main Study) to NCES that provided these data points for analysis. Westat also provided a supplemental national delivery file to NCES with these additional data points for analysis from the National Supplement.

The same data delivery process to NCES and the Consortium was followed for the National Supplement; however, separate deliveries were made for the household sample and the prison sample. IEA completed the same review of the delivered household data as was completed for the Main Study. IEA completed a similar review of the prison sample data; however, revisions to IEA's process were required in order to ensure that the BQ changes were accurately reflected in the DME. These revisions were made in order to ensure that the prison BQ changes that impacted question routing were implemented accurately. Throughout the process IEA worked closely with ETS and Westat to verify the data and identify any possible discrepancies in order to ensure that valid skip patterns and data cleaning routines were implemented correctly.

10. DATA QUALITY, IRT ANALYSES, AND POPULATION MODELING

10.1 Introduction

The aim of this chapter is to give an overview of the analyses that were conducted to examine the quality of the data, the IRT scaling approach that was used to link the National Supplement scales to the PIAAC scales, and the population modeling that was conducted to generate plausible values.

The following steps were taken to assure comparability of the National Supplement results to those in the PIAAC Main Study in terms of *instrumentation* and *survey operations*. The National Supplement's household and prison samples were tested using the same instruments used in the Main Study. These instruments assessed the PIAAC domains of literacy (including reading components), numeracy, and problem solving for technology rich environments (PSTRE). As in the PIAAC 2012 Main Study, items were administered through either a paper-based assessment (PBA) or computer-based assessment (CBA). The systems of test administration, scoring and the evaluation of scoring accuracies employed for cognitive items were the same as those in the Main Study (see 10.2 for more information). The same applies regarding the analysis methods and procedures, which are based on the same psychometric principles.

As in the Main Study, the test design was based on a variant of sampling design most common to the major large-scale surveys. This design, called *matrix sampling*, entails administering a subset of items from a larger item pool, with different groups of respondents answering different sets of items. This design allowed a reduction in the response burden for an individual while expanding the item pool to represent the framework as completely as possible.

The use of this design makes it inappropriate to use any statistic based on the number of correct responses to describe or compare the skills of respondents, but this limitation can be overcome through item response theory (IRT) scaling. When a set of items requires a given skill, the response patterns should show regularities that can be modeled using the underlying commonalities among the items. These regularities can be used to characterize respondents (by estimating so-called person or ability parameters through IRT models) and items (by estimating certain item parameters through IRT models; e.g., item difficulty) in terms of a common scale, even if not all respondents take identical sets of items. In other words, if an item pool is used to measure a certain skill unidimensionally (i.e., only one skill is necessary to solve the items), respondents can be compared even if they responded to different sets of items in the pool (as long as the item pool is scaled using a certain IRT model and showed appropriate model fit). This makes it possible to describe distributions of performance in a population or subpopulation and to estimate the relationships between proficiency and background variables.

Before data can be used for analyses, the quality of the data must be evaluated. This was done by reviewing the item responses to determine whether each respondent received the items and booklets as planned in the design (completion), reviewing item analyses (percent of correct responses per item) within and across groups to detect potential errors in translation or scoring, and reviewing scorer agreement to evaluate consistency in scoring (reliability). Quality checks were also performed to evaluate the handling and pattern of missing values (i.e., missing by design, omitted by the respondent).

In order to link the PIAAC National Supplement and the Main Study in terms of a common scale, the item parameters in the IRT scaling for the supplemental study were fixed to those item parameters obtained from the Main Study (fixed item parameter linking). Using common IRT item parameters

assures that the scale linkage can be established and the inference structures remain intact from PIAAC. To achieve this, the majority of item parameters needed to be the same as in the PIAAC Main Study (common item parameters), while a few items needed unique item parameters in certain groups (newly estimated item parameters in case they showed no fit to the common item parameters obtained in the Main Study).

In the following sections, the data evaluation process and the scaling model used for the PIAAC National Supplement are described.

10.2 Data Handling and Evaluation: Missing Values, Completion, Item Analysis, and Scoring Reliability

Assurance of data quality is an important step prior to IRT scaling and population modeling. Only if the analyses are based on correct data can reasonable and meaningful results be provided. Procedures for evaluating scoring and handling of missing data, data completion, and item analyses are illustrated below.

10.2.1 Scoring and Handling of Missing Data

The National Supplement followed the same scoring guidelines and procedures as those applied in the Main Study. The literacy and numeracy items were dichotomously scored: correct responses were scored as 1 and incorrect responses as 0. The problem-solving items received polytomous scores: partly correct or fully correct responses were scored with 1 and 2, or 1, 2, and 3 depending on the item; incorrect responses were scored as 0. The two kinds of missing values were coded differently for IRT calibration purpose: items that were administered but omitted by the respondent were coded as 8, and items that were not administered by design were coded as 9.

Regarding the handling of missing data, the design of the National Supplement followed the same procedure to that used in the Main Study in order to maintain comparability. The data have a characteristic structure of missing responses that are derived from the matrix sampling design and the instituted accommodation for respondents with very low proficiency skills through core items. This structure is characterized by data missing completely at random (coded as 9) due to the test design (random assignment of booklets) and data missing due to omitted responses (8). More specifically, there were different types of missing values within the cognitive part of the assessment:

1. Missing by design: items that were not presented to each respondent due to the matrix sampling design. Accordingly, these structural missing data, unrelated to respondents' literacy, numeracy, and problem-solving skills, were not included in the estimation of respondent proficiencies.

2. Omitted responses: missing responses that occurred when respondents chose not to perform one or more presented items, either because they were unable to do so or some other reason. Any missing response followed by a valid response (whether correct or incorrect) was defined as an omitted response. Omitted responses were treated as wrong because a random response to an open-ended item would almost certainly result in a wrong answer.

3. Not reached or not attempted responses: missing responses at the end of a booklet were treated as if they were not presented (i.e., missing by design) due to the difficulty

of determining if the respondent was unable to finish these items or simply abandoned them.

4. Cases where respondents did not answer a sufficient number of Background Questionnaire (BQ) questions (< 5 items) were considered as incomplete cases and not used in the latent regression; they were also not included in computing plausible values.

Some respondents who answered a sufficient number of BQ questions may not have been able to respond to the cognitive items or were unwilling to respond to the cognitive items. The treatment of these cases is described in the next section. However, the nonresponse rate in the U.S. National Supplement sample was low (19.9 percent for the household supplemental sample, and 14.7 percent for the prison sample).

10.2.2 Data Completion—Treatment of Respondents with Fewer than Five Cognitive Item Responses

This section addresses the issue of respondents who provided background information but did not completely respond to the cognitive items. A minimum of five completed items per domain was necessary to assure sufficient information about the proficiency of respondents.

In some cases a sampled person decided to stop the assessment. The reasons for stopping may be classified into two groups: those unable to respond to the cognitive items (i.e., for cognitive-related reasons), and those unwilling to respond (i.e., for noncognitive-related reasons).

The National Supplement followed the Main Study procedure with respect to cases with responses to fewer than five cognitive items per domain. All consecutively missing responses at the end of a block of items were treated as incorrect if the reason for not responding to the cognitive items was related to literacy, numeracy, or problem-solving skills (i.e., cognitive-related reasons). Otherwise, all consecutively missing responses were treated as "not reached" and coded as 9.

This scoring method is important with regard to the latent regression population model. The treatment of nonresponding examinees due to noncognitive-related reasons has no impact on the likelihood function of proficiency, but the treatment due to cognitive-related reasons does. With this scoring procedure, summary statistics can be produced for the entire population, including those who respond to cognitive items correctly as well as those who were not able to respond to cognitive items.

10.2.3 Item Analyses

Once the data were prepared, item analyses were conducted separately for each domain (literacy, numeracy, problem solving, and reading components) and the following two groups: a combined group including the Main Study sample with adjusted weights and the National Supplement sample. The purpose of the item analyses was to identify outliers or unexpected patterns that might signify issues with translations of items or scoring guides, or issues related to a misinterpretation of scoring guides during scoring training. The item analysis report includes the following statistics for each item:

- summary statistics;
- response statistics;

- item response categories within blocks; and
- item statistics.

Summary Statistics:

- Statistics for the computation of the alpha reliability coefficient and standard error of measurement for the test;
- Summary statistics for the literacy and numeracy block scores across all respondents. The block score is the sum of correct responses (across all subjects) per domain for each group (defined sample);
- Frequency distribution of block scores of all subjects, including summary statistics;
- Summary statistics for the zlogit score across all respondents;
- Score category analysis per domain and block; and
- Item-level statistics gathered from the response analysis, including the biserial correlation (R-BIS), proportion correct (P-PLUS), percent not reached (% NOTRCH), percent off-task reached (% OFFTSK) and percent omitted (% OMIT). The ninth column contains percent of missing responses (% MISS), calculated as the sum of the percent omitted and percent not reached. Flag symbols for each item according to whether these statistics exceeded certain criterion values are included as well.

Response Statistics:

The response statistics are illustrated in Table 10-1.

Table 10-1. Response statistics

N	The number of respondents in each response category for the question
Percent	The percent of respondents in each response category for the question
Mean Score	The average zlogit scores for the respondents in each response category for the question
Std. Dev.	The standard deviation of the zlogit scores for the respondents in each response category for the question
Resp Wt	Scoring weight applied to item responses: 1 = correct and 0 = incorrect

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Item Response Categories Within Blocks:

The response categories of each item within a block are illustrated in table 10-2.

Table 10-2. Response categories of each item within the block

NOT RCH	Respondents who did not respond to or omitted the question and did not respond to any subsequent questions
OFF TSK	Respondents who did not see or start the block
OMIT	Respondents who did not respond to or omitted the question but did respond to at least one subsequent question in the block
1*	Respondents who responded correctly to dichotomously scored items
1, 2	Respondents who responded partially correct (1) or fully correct (2) to polytomously scored items
7	Respondents who responded incorrectly
TOTAL	The aggregation of respondents who either omitted the item or had valid response codes These statistics do not include the respondents who did not reach the item

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Item Statistics:

The item statistics are illustrated in table 10-3.

Table 10-3. Item statistics

R BIS	The Rbis (R-biserial) indicates the correlation between respondents' performance on an individual question and their performance on the criterion score. It is a measure of a question's power to discriminate among respondents of different abilities. A relatively high R-biserial indicates that respondents who scored higher on the criterion score were more likely than respondents who scored lower to get that individual question correct. The R-biserial estimates the product moment correlation that would be obtained from two continuous distributions if the dichotomized variable were normally distributed. In special cases it can take on a value greater than 1, and it is actually unbounded in both directions.
PT BIS	The point biserial is the Pearson product moment correlation coefficient between the dichotomous item score (0, 1) and the continuous criterion score. Its range is (-1, 1).
P+	P+ is the percent of respondents who reached the question and selected the correct answer (calculated as: sum of the scores/sum of all possible scores)
Delta	Delta index is the inverse-normal transformation of proportions correct to describe item difficulty with the mean of 13.0 and the standard deviation of 4. Smaller delta index indicates easier item and larger number indicates difficult item. The index can vary often between 1 and 25.

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

10.3 IRT Scaling: Estimation of Item Parameters

The IRT scaling provides the estimations of item parameters and the proficiency distribution of the population. It was carried out separately for the domains of literacy, numeracy, and problem solving (no IRT analyses were conducted for reading components). Similar to the Main Study, the National Supplement used the two-parameter logistic model (2PL; Birnbaum, 1968) for dichotomously scored

responses (literacy, numeracy) and the general partial credit model (GPCM; Muraki, 1992) for polytomous data (problem solving). Incorrect responses were coded as 0; correct responses in the 2PL model as 1, and 1, 2, or 3 in the GPCM; omitted responses were coded as incorrect responses (0); and missing responses by design as missing values (9) so that these responses did not contribute to the IRT estimations.

The IRT scaling was conducted for 166 items from the cognitive assessment (as in the Main Study): literacy (76 items), numeracy (76 items), and problem solving (14 items). No item had to be excluded from the analyses.

The *2PL model* is a mathematical model for the probability that an individual will respond correctly to a particular item from a single domain of items. The probability of solving an item (i) depends only on the ability or proficiency (θj) of the respondent (j) and two item parameters characterizing the properties of the item (item difficulty β_i, which is the location parameter of item i characterizing its difficulty, and item discrimination α_i, which is the slope parameter of item I characterizing its sensitivity to proficiency).

The *GPCM*, like the 2PL model, is a mathematical model for the probability that an individual will respond in a certain response category on a particular item. While the 2PL model is suitable for dichotomous responses only, the GPCM can be used with polytomous and dichotomous responses considering m_i+1 ordered response categories for an item *i*. The GPCM reduces to the 2PL model when applied to dichotomous responses.

For more details about the IRT models and the model equations see the PIAAC Main Study technical report (Yamamoto, Khorramdel, and von Davier, 2013).

A central assumption of IRT is conditional independence (sometimes also called local independence). In other words, item response probabilities depend only on the respondent's ability and the specified item parameters – there is no dependence on any demographic characteristics of the examinees, or responses to any other items presented in a test, or the survey administration conditions. Moreover, the 2PL model and the GPCM assume unidimensionality; that is, a single latent variable, the ability or proficiency θ, accounts for performance on a set of items.

For more details about the models and IRT scaling process, see the technical report for the PIAAC Main Study (Yamamoto, Khorramdel, and von Davier, 2013).

As the National Supplement uses the items from the Main Study, the invariance of item parameters was evaluated on the supplemental data. For this, the item parameters in the IRT scaling of the National Supplement data were fixed to the values of the item parameters obtained in the Main Study. The original household sample from the Main Study was included in the IRT analyses together with the two samples (household and prison) from the National Supplement. The original sample received updated standardized sample weights reflecting the Main Study data and the National Supplement household sample data. The standardized sample weights were calculated to sum up to 5000 for the Main Study sample (2012 data) together with the National Supplement household sample, and up to 5000 for the prison sample. Cases with more than 5 cognitive responses were used to evaluate the fit of the Main Study IRT parameters for the PBA and CBA items (note that all cases received plausible values). While the Main Study sample was used together with the National Supplement household sample as one group in the population model (see 10.5), they were treated as separate groups in the IRT scaling because first it had to be evaluated whether the same (common) item parameter can be used for these samples. The three different samples (Main Study, household, prison) were further divided by mode (CBA and PBA) resulting in six different groups. Tables 10-4 and 10-5 give an overview of the sample sizes for the six groups used in the IRT scaling (note that PSTRE exists only in the CBA).

Table 10-4. Sample Sizes for the literacy and numeracy domains

	Group used in IRT Analyses	Mode	N	Standardized weighted n
1	Main Study household sample	PBA	615	449.71
2	Main Study household sample	CBA	4101	2885.00
3	National Supplement household sample	PBA	671	252.63
4	National Supplement household sample	CBA	2777	988.45
5	National Supplement prison sample	PBA	397	1628.97
6	National Supplement prison sample	CBA	843	3029.44

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Table 10-5. Sample sizes for the problem solving in technology-rich environments (PSTRE) domain

	Group used in IRT Analyses	Mode	N	Standardized weighted n
1	Main Study household sample	CBA	4716	3334.71
2	National Supplement household sample	CBA	3448	1241.07
3	National Supplement prison sample	CBA	1240	4658.41

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

The item fit statistics used are the mean deviation (MD) and the root mean square deviation (RMSD). Both measures quantify the magnitude and direction of deviations in the observed data from the estimated item characteristic curve (ICC) for each single item. While the MD is most sensitive to the deviations of observed item difficulty parameters from the estimated ICC, the RMSD is sensitive to the deviations of both the observed item difficulty parameters and item slope parameters.

The scaling procedure needed to take into account the possibility of any systematic interaction between the national subsamples and the items that were used to produce estimates of the item parameters and sample distributions. For this reason, the 2PL model and the GPCM were estimated as a multiple-group IRT model using a mixture of normal population distributions (one for each subsample) where item parameters were generally constrained to be equal across groups with a unique mean and variance for each group (see the above described concurrent calibration). The moments of these distributions were updated at each iteration during IRT calibration.

Items that showed deviations from the common Main Study item parameters were assumed to work differently in the National Supplement than in the Main Study and would harm the link. Poorly fitting items or ICCs were identified using an RMSD > 0.15, and a MD > 0.15 and < -0.15 criterion where a value of 0 indicates no discrepancy (in other words, a perfect fit of the model). As in the Main Study, it was assumed for such items that the common item parameters were not appropriate and group-specific unique item parameters were estimated in a second step.

In this subsequent step, unique item parameters were estimated to account for group-based deviations for a small subset of items. This involved a close monitoring of the IRT scaling for item-by-group interactions and allowing group-specific item parameters only in instances where substantial deviations were identified. This procedure takes into account that some items work differently in certain groups. The

common and unique item parameters were estimated using a mathematical algorithm[53] that still allows us to estimate all item parameters in relation to one another and, thus, common and unique item parameters were on the same latent scale. A large number of common item parameters support the comparability of the scales across the countries and assessments, and a few unique item parameters only reduce the measurement error further while not impacting the comparability of scales.

In the analyses for the National Supplement in most cases, the linking item responses across groups were accurately described by the common PIAAC item parameters. Only three literacy and three numeracy items needed unique item parameters in the subsample of incarcerated respondents (literacy: item M313412S for the PBA group, item E322003S and item E322004S for CBA; numeracy: item M600C04S for PBA, item C606509S and item C611517S for CBA). For PSTRE, no unique item parameters needed to be estimated. Table 10-6 provides the number of unique item parameters per group.

Table 10-6. Number of unique item parameters for each group and domain

		Number of group-specific item parameters		
Group	Mode	Literacy (76 items)	Numeracy (76 items)	PSTRE (14 items)
Main Study household sample	PBA	0	0	---
Main Study household sample	CBA	0	0	0
National Supplement household sample	PBA	0	0	---
National Supplement household sample	CBA	0	0	0
National Supplement prison sample	PBA	1	1	---
National Supplement prison sample	CBA	2	2	---

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Results show that the establishment of comparability of the Main Study and the National Supplement for the USA was successful. The data from both studies can be described on a common scale. Because no unique item parameter had to be estimated for the National Supplement household sample, the Main Study and the National Supplement household sample could be treated as one group in the subsequent population model (see 10.5). The National Supplement prison sample was treated as a separate group in the population model because unique item parameters had to be estimated for this sample (see table 10-6).

10.4 Reading Components

Reading components were administered to respondents who did not pass the literacy core items in the CBA (respondents with fewer than 3 correct core items out of 6) as well as to respondents in the PBA. The reading components consisted of three different blocks resulting in total of 100 items: vocabulary (34 items), sentence processing (22 items), and basic passage comprehension (44 items).

For the reading components, both response time and proportions correct had predictable relationships with the literacy proficiencies (see table 10-7). Results, similar to those of the PIAAC 2012 Main Study, show a high proportion of correct responses, as expected, even among least able respondents with scores lower than 175 (vocabulary: P+ = 0.85; sentence processing: P+ = 0.60; basic passage comprehension:

[53] The software *mdltm* (von Davier, 2005) was used for the IRT calibration, which provides marginal maximum likelihood (MML) estimates obtained using customary expectation-maximization (EM) methods, with optional acceleration.

P+ = 0.55). While high response accuracy was even among less able respondents, response fluency represented by the average response time indicates that less able respondents (respondents who are below level 1) took 2 to 3 times longer to answer reading component items compared to most able respondents (respondents who are in level 3 and above). (Please note that the reading components were not administered with any time limit.)

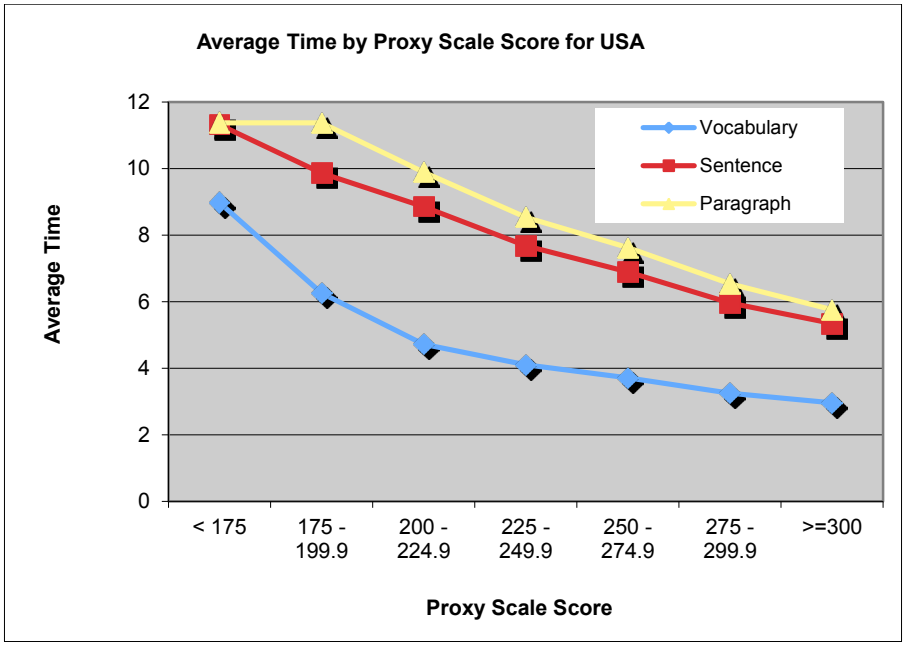
Table 10-7. Reading components average proportions correct and average response time by literacy posterior means

		Literacy Posterior Means						
		< 175	175 – 199.9	200 – 224.9	225 – 249.9	250 – 274.9	275 – 299.9	>= 300
Vocabulary	Average proportions correct P+	0.85	0.92	0.96	0.98	1.00	1.00	1.00
	Average response time per item (sec)	8.98	6.26	4.72	4.10	3.71	3.25	2.96
Sentence	Average proportions correct P+	0.60	0.75	0.84	0.91	0.95	0.97	0.98
	Average response time per item (sec)	11.32	9.87	8.86	7.68	6.90	5.97	5.34
Passage	Average proportions correct P+	0.55	0.76	0.85	0.93	0.97	0.99	1.00
	Average response time per item (sec)	11.38	11.38	9.90	8.55	7.63	6.54	5.76

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Figure 10-1 shows the fluency represented by the average response time for the reading components scale projected onto the literacy proficiency scale.

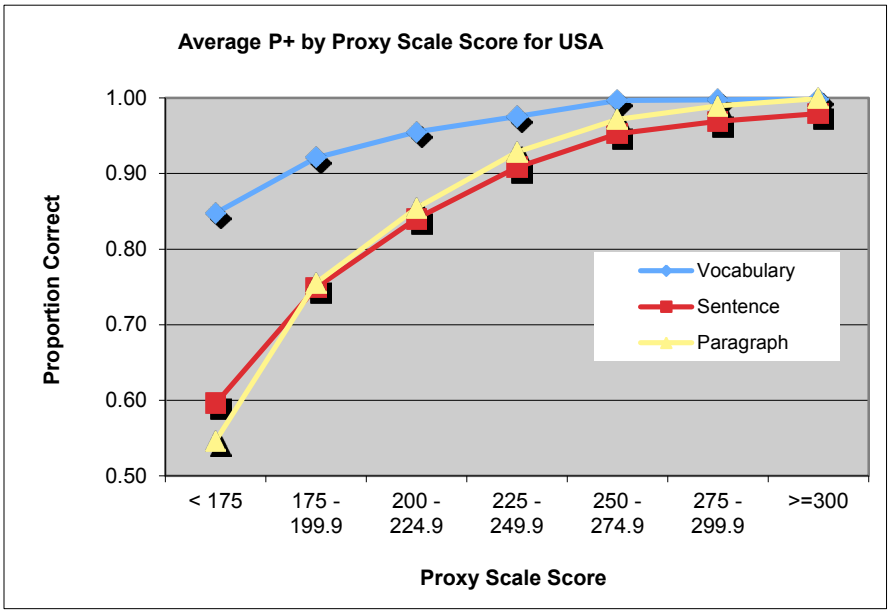
Figure 10-1. Fluency (average response time) of the reading components scale projected onto the literacy proficiency scale



SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

In terms of accuracy of the reading components, results are similar to the Main Study results as well, showing a high proportion of correct responses as expected (vocabulary: P+ = 0.96; sentence processing: P+ = 0.86; basic passage comprehension: P+ = 0.87). This means that the reading components were easy for every respondent. Figure 10-2 shows the proportion of correct responses for the reading components scale projected onto the literacy proficiency scale.

Figure 10-2. Accuracy (discrimination by means of conditional P+) of the reading components scale projected onto the literacy proficiency scale



SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

10.5 Population Modeling

The population or conditioning model used for the Main Study and the National Supplement is a combination of the IRT models described above and a latent regression model. In the latent regression model, the distribution of the proficiency variable (θ) is assumed to depend not only on the cognitive item responses X but also on a number of predictors Y, which are variables obtained from the BQ (e.g., gender, country of birth, education, occupation, employment status, reading practices, etc.). Both the item parameters from the IRT calibration stage (see section 10.3) and the estimates from the regression analysis are needed to generate plausible values.

Usually, a considerable number of background variables (predictors) are collected in international large-scale assessments, with a principal component analysis extracting the components that explain 90 percent of the variation for further analysis. For the PIAAC analysis, the background variables are contrast coded and 80 percent of explained variance are used to avoid overparameterization. The use of principal components also serves to retain information for examinees with missing responses to one or more background variables.

The latent regression parameters Γ and $\sum$ are estimated conditional on the previously determined item parameter estimates (from the item calibration stage) and plausible values for each respondent j are drawn from the conditional distribution.

All consecutively missing responses at the end of a block of items were treated as incorrect if the reason for not responding to the cognitive items was related to the cognitive skills. Otherwise, all consecutively missing responses were treated as not reached (i.e., not presented). Furthermore, examinees with responses to fewer than five cognitive items per domain were not included in a first run of the population

modeling (with regard to the regression model) to obtain unbiased Γ and Σ. In a second round of analysis, the regression parameters were treated as fixed to obtain plausible values for all cases, including those with fewer than five responses to cognitive items.

For more details about the population model and the computational procedure, please see the technical report for the PIAAC 2012 Main Study (Yamamoto, Khorramdel, and von Davier, 2013).

10.6 Generating Plausible Values

Plausible values are multiple imputed proficiency values based on information from the cognitive items (IRT scaling) and the BQs (principal components). Plausible values are used to obtain more accurate estimates of group proficiency than would be obtained through an aggregation of point estimates. A more detailed description is given in Mislevy (1991), Thomas (2002), and von Davier, Sinharay, Oranje, and Beaton (2006). For more details about the population model and the computational procedure, please see the technical report for the PIAAC 2012 Main Study (Yamamoto, Khorramdel, and von Davier 2013).

In PIAAC, the computation of group-level reporting statistics involving scores in the cognitive domains is based on 10 independently drawn plausible values for each domain assigned to each respondent. Each set of plausible values is equally well designed to estimate population statistics (such as group means and standard deviations); however, multiple plausible values (in PIAAC: 10) are required for each respondent to represent the uncertainty in the domain measures appropriately (von Davier, Gonzalez, and Mislevy 2009). The statistics based on scores are always computed at population or subpopulation levels and should never be used to draw inferences at the individual level.

In the analyses of PIAAC, a normal multivariate distribution was assumed for $P(\theta_j|x_j, y_j, \Gamma, \Sigma)$, with a common variance, Σ, and with a mean given by a linear model with slope parameters, Γ, based on the principal components of several hundred selected main effects from the vector of background variables. The item parameters for the cognitive items were obtained from the IRT calibration. All background variables were contrast coded before they were processed further in the population model. Contrast coding allows the inclusion of codes for refused responses as well as codes for responses that were not collected by means of routing and avoiding the necessity of linear coding. The increased number of variables obtained through contrast coding is substantial. To capture most of the common variance in the contrast-coded background questions with a reduced set of variables, a principal component analysis was conducted. Because each population can have unique associations among the background variables, the extraction of principal components was carried out separately for the following two groups: the prison sample, and a group combining the Main Study and National Supplement household sample. In PIAAC, each set of principal components y^c (or conditioning variables) was selected to include 80 percent of the variance with the aim of explaining as much variance as possible while at the same time avoiding overparameterization.

References

Bell, B., Mohadjer, L., Montaquila, J., and Rizzo, L. (1999). Creating a Frame of Newly Constructed Units for Household Surveys. *JSM Proceedings*, Survey Research Methods Section: 306–310.

Birnbaum, A. (1968). Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. In F.M. Lord and M.R. Novick (Eds.), *Statistical Theories of Mental Test Scores* (pp.397–479). Reading, MA: Addison-Wesley.

Carsten, R., Daniel, T., and Gonzalez, E. (2013). International Database and Data Analysis Tools. In *Technical Report of the Survey of Adult Skills (PIAAC)* [Pre-publication copy] (Chapter 23, pp. 539–544). Paris: Organisation for Economic Cooperation and Development. Retrieved from: https://www.oecd.org/site/piaac/_Technical%20Report_17OCT13.pdf.

Cochran, W.G. (1977). *Sampling Techniques* (3rd ed.). New York: John Wiley & Sons.

Denis, V. (2014). User's Guide for a SAS Macro enabling the computation of design based standard errors in the Survey of Adult Skills (PIAAC). Organization for Economic Co-operation and Development (OECD), Version 1.2. Retrieved from the OECD website: http://www.oecd.org/skills/piaac/PIAAC%20SAS%20Macro%20User%20Guide%20-%202.2.pdf

Dohrmann, S., Harding, L., and Li, L. (2008). Using Digital Imagery to Update the Measures of Size of Area Segments. *JSM Proceedings*, Survey Research Methods Section: 1718–1725.

Dohrmann, S., Li, L., and Mohadjer, L. (2011). Updating the Measures of Size of Local Areas Late in the Decade Using USPS Address Lists. *JSM Proceedings*, Survey Research Methods Section: 2891–2901.

Kalton, G., and Flores-Cervantes, I. (2003). Weighting Methods. *Journal of Official Statistics 19*(2): 81–97.

Kalton, G., Kali, J., and Sigman, R. (2014). Handling Frame Problems When Address-Based Sampling is Used for In-Person Household Surveys. *Journal of Survey Statistics and Methodology 2*(3): 283–304.

Krenzke, T., and Haung, W. (2009). Revisiting Nested Stratification of Primary Sampling Units. *Proceedings of the Federal Committee on Statistical Methodology*. Retrieved from http://fcsm.sites.usa.gov/files/2014/05/2009FCSM_Krenzke_IX-C.pdf.

Lee, H. (1995). Outliers in Business Surveys. In Brenda G. Cox et al. (Eds.), *Business Survey Methods* (pp. 503–526). New York: John Wiley & Sons.

Lessler, J., and Kalsbeek, W. (1992). *Nonsampling Error in Surveys*. New York: John Wiley & Sons.

Little, R.J.A. (1986). Survey Nonresponse Adjustments for Estimates of Means. *International Statistical Review 54*(2): 139–157.

Lohr, S.L. (2011). Alternative Survey Sample Designs: Sampling with Multiple Overlapping Frames. *Survey Methodology 37*(2): 197–213.

Lord, F.M. (1980). *Applications of Item Response Theory to Practical Testing Problems.* Hillsdale, NJ: Erlbaum Associates.

Mislevy, R.J. (1991). Randomization-based Inference About Latent Variables From Complex Samples. *Psychometrika, 56*(2): 177–196.

Mohadjer, L., Kalton, G., Krenzke, T., Liu, B., Van de Kerkhove, W., Li, L., Sherman, D., Dillman, J., Rao, J.N.K., and White, S. (2009). *National Assessment of Adult Literacy: Indirect County and State Estimates of the Percentage of Adults at the Lowest Level of Literacy for 1992 and 2003* (NCES 2009–482). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Mohadjer, L., Montaquila, J.M., and Sherris, E. (2002). Evaluation of a Two-phase Approach to Segment Selection for Area Probability Samples Late in the Decade. *JSM Proceedings*, Survey Research Methods Section: 2395–2399.

Montaquila, J.M., Bell, B., Mohadjer, L., and Rizzo, L. (1999). A Methodology for Sampling Households Late in a Decade. *JSM Proceedings*, Survey Research Methods: 311–315.

Montaquila, J.M., Hsu, V., and Brick, J.M. (2011). Using a "Match Rate" Model to Predict Areas Where USPS-Based Address Lists May Be Used in Place of Traditional Listing. *Public Opinion Quarterly 75*(2): 317–335.

Muraki, E. (1992). A generalized partial credit model: application of an EM algorithm. *Applied Psychological Measurement, 16*(2): 159–177.

PIAAC Consortium. (2014). *Technical Standards and Guidelines.* Paris: OECD. Retrieved http://www.oecd.org/site/piaac/PIAAC-NPM(2014_06)PIAAC_Technical_Standards_and_Guidelines.pdf.

Pokropek, A., and Jakubowski, M. (2013). PIAACTOOLS: Stata® programs for statistical computing using PIAAC data. Organization for Economic Co-operation and Development (OECD). Retrieved from: http://www.oecd.org/site/piaac/PIACTOOLS_16OCT_for_web.pdf.

Oh, H.L., and Scheuren, F.J. (1987). Modified Raking Ratio Estimation. *Survey Methodology 13*(2): 209–219.

Rao, J.N.K., and Scott, A.J. (1984). On Chi-Squared Tests for Multiway Contingency Tables with Cell Proportions Estimated From Survey Data. *Annals of Statistics 12*(1): 46–60.

Rao, J.N.K., and Thomas, D.R. (1989). Chi-Squared Tests for Contingency Tables. In D.Holt, C.J. Skinner, and T.M.F. Smith (Eds.), *Analysis of Complex Surveys* (pp. 89–114). New York: John Wiley & Sons.

Rao, J.N.K., and Thomas, D.R. (2003). Analysis of Categorical Response Data from Complete Survey: An Appraisal and Update. In R.L. Chambers and C.J. Skinner (Eds.), *Analysis of Survey Data* (pp. 85–108). West Sussex, England: John Wiley & Sons.

Rust, K., and Rao, J.N.K. (1996). Variance Estimation for Complex Surveys Using Replication Techniques. *Statistical Methods in Medical Research 5*: 283–310.

Sonquist, J., Baker, E., and Morgan, J. (1974). *Searching for Structure: An Approach to Analysis of Substantial Bodies of Micro-Data and Documentation for a Computer Program* (Rev. ed.). Ann Arbor, MI: Institute for Social Research, University of Michigan.

Thomas, N. (2002). The Role of Secondary Covariates When Estimating Latent Trait Population Distributions. *Psychometrika, 67*(1), 33–48.

Thomas, D.R., and Rao, J.N.K. (1987). Small-Sample Comparisons of Level and Power for Simple Goodness-of-Fit Statistics Under Cluster Sampling. *Journal of the American Statistical Association*, *82*(398): 630–636.

von Davier, M. (2005). *A General Diagnostic Model Applied to Language Testing Data* (Research Report RR-05-16). Princeton, NJ: Educational Testing Service.

von Davier, M., Gonzalez, E., and Mislevy, R. (2009). What Are Plausible Values and Why Are They Useful? In *IERI Monograph Series: Issues and Methodologies in Large Scale Assessments, Vol. 2,* (pp.9–36). Retrieved from IERI website: http://www.ierinstitute.org/fileadmin/Documents/IERI_Monograph/IERI_Monograph_Volume_02.pdf.

von Davier, M. Sinharay, S., Oranje, A., and Beaton, A. (2006). Statistical Procedures Used in the National Assessment of Educational Progress (NAEP): Recent Developments and Future Directions. In C.R. Rao and S. Sinharay (Eds.), *Handbook of Statistics, Vol. 26: Psychometrics* (pp.1039–1055). Amsterdam: Elsevier.

Yamamoto, K., Khorramdel, L., and von Davier, M. (2013). Scaling PIAAC Cognitive Data. In *Technical Report of the Survey of Adult Skills (PIAAC)* [Pre-publication copy] (Chapter 17, pp. 406–438). Paris: Organisation for Economic Cooperation and Development. Retrieved from: https://www.oecd.org/site/piaac/_Technical%20Report_17OCT13.pdf.

Westat. (2007). *WesVar 4.3 User's Guide*. Rockville, MD: Westat.

This page intentionally blank.

Appendix A

Basic Analysis Tables

This page intentionally blank.

Appendix A. Basic Analysis Tables

Table A-1. PIAAC weighted Screener and Background Questionnaire response rates: 2012 Main Study (MS) and 2014 List and Area Samples

	Weighted Screener response rate (SCR) (percent)			Weighted BQ response rate (BQ) (percent)			Weighted cumulative response rate (SCR x BQ) (percent)		
Analysis variable	MS	Area	List	MS	Area	List	MS	Area	List
Overall	86.5	81.4	84.8	82.1	77.9	93.0	71.0	63.4	78.8
Region									
Northeast	83.7	75.0	84.9	77.7	71.4	84.7	65.1	53.6	71.9
Midwest	89.0	87.7	86.0	84.3	79.5	95.9	75.1	69.8	82.4
South	87.6	84.5	84.3	82.9	79.1	92.8	72.6	66.8	78.3
West	83.8	74.7	75.4	82.9	80.9	95.1	69.5	60.4	71.7
Indicator of whether the PSU is part of a Metropolitan Statistical Area									
Nonmetro. area	90.7	89.4	86.3	85.1	81.7	88.7	77.2	73.0	76.5
Metro. area	85.6	79.8	84.1	81.6	77.3	95.4	69.8	61.7	80.2
Categorized average household size[1]									
1st quartile or less	87.9	82.4	85.8	82.0	74.5	90.1	72.1	61.4	77.3
(1st quartile – median]	88.3	85.6	88.7	83.2	78.0	96.2	73.4	66.8	85.3
(median – 3rd quartile]	86.1	81.9	83.3	80.9	79.1	92.5	69.6	64.8	77.0
More than 3rd quartile	83.5	76.2	76.1	82.6	79.5	90.3	68.9	60.5	68.8
Percentage of the population below 150 percent of poverty[1]									
1st quartile or less	83.1	79.9	87.1[2]	78.1	70.5	92.8[2]	64.9	56.3	80.8[2]
(1st quartile – median]	86.4	82.3		81.0	76.9		70.0	63.3	
(median – 3rd quartile]	89.2	83.5	85.6	84.2	79.1	89.0	75.1	66.0	76.2
More than 3rd quartile	86.8	80.2	83.1	85.4	84.1	95.9	74.1	67.4	79.7
Percentage of the population that is foreign born[1]									
1st quartile or less	92.0	89.2	88.1	85.2	80.5	90.6	78.4	71.9	79.9
(1st quartile – median]	87.3	85.5	83.7	80.6	76.3	94.7	70.3	65.2	79.2
(median – 3rd quartile]	83.3	78.9	83.6	80.1	76.4	95.2	66.8	60.3	79.6
More than 3rd quartile	82.2	71.3	67.4	82.3	78.4	91.2	67.7	55.9	61.5
Percentage of the population age 25 and older with at least a high school education[1]									
1st quartile or less	87.0	80.8	86.3	85.7	85.1	94.2	74.6	68.7	81.3
(1st quartile – median]	89.0	83.0	82.6	83.4	80.0	89.4	74.0	66.4	73.8
(median – 3rd quartile]	84.4	82.1	85.4	80.0	73.1	97.0	67.5	60.0	82.9
More than 3rd quartile	85.4	80.3	66.2	79.7	72.4	88.1	67.8	58.1	58.4
Percentage of households that are linguistically isolated[1]									
1st quartile or less	90.2	88.1	88.5	84.2	79.7	92.1	75.9	70.2	81.5
(1st quartile – median]	87.9	84.2	84.8	80.3	76.9	93.7	70.5	64.7	79.5
(median – 3rd quartile]	85.6	79.4	83.7	80.1	75.0	92.1	68.6	59.6	77.1
More than 3rd quartile	81.8	73.5	77.3	84.1	79.9	95.0	68.7	58.8	73.4

See notes at end of table.

Table A-1. PIAAC weighted Screener and Background Questionnaire response rates: 2012 Main Study (MS) and 2014 List and Area Samples—Continued

	Weighted Screener response rate (SCR) (percent)			Weighted BQ response rate (BQ) (percent)			Weighted cumulative response rate (SCR x BQ) (percent)		
Analysis variable	MS	Area	List	MS	Area	List	MS	Area	List
Percentage of housing units occupied by owner[1]									
1st quartile or less	84.2	73.5	81.8	82.4	81.1	94.5	69.3	59.6	77.3
(1st quartile – median]	87.5	82.3	85.0	82.9	78.5	94.6	72.6	64.6	80.4
(median – 3rd quartile]	87.7	84.8	87.4	83.8	77.9	92.2	73.5	66.1	80.6
More than 3rd quartile	86.1	84.0	82.5	79.7	74.3	89.0	68.6	62.4	73.4
Percentage of the population age 18–64 that is unemployed[1]									
1st quartile or less	86.5	82.2	-	81.9	75.6	-	70.9	62.1	-
(1st quartile – median]	85.8	82.1	83.5[2]	81.7	76.7	91.9[2]	70.1	63.0	76.7[2]
(median – 3rd quartile]	88.0	82.1	86.3	82.7	79.5	93.2	72.8	65.3	80.4
More than 3rd quartile	85.6	78.7	84.2	82.3	80.2	93.3	70.4	63.1	78.6
Race/ethnicity (after imputation)									
Hispanic	-	-	-	84.5	84.2	87.3	-	-	-
Non-Hispanic Black	-	-	-	82.3	82.8	95.6	-	-	-
Other	-	-	-	81.7	75.6	92.2	-	-	-
Age category for Main Study (after imputation)[1]									
16–25	-	-	-	85.8	-	-	-	-	-
26–35	-	-	-	83.9	-	-	-	-	-
36–45	-	-	-	80.2	-	-	-	-	-
46–55	-	-	-	80.8	-	-	-	-	-
56–65	-	-	-	79.9	-	-	-	-	-
Age category for National Supplement (after imputation)[1]									
16–24	-	-	-	-	80.3	92.6	-	-	-
25–34	-	-	-	-	77.6	96.1	-	-	-
35–45	-	-	-	-	94.5	87.4	-	-	-
46–55	-	-	-	-	80.7	92.7	-	-	-
56–65	-	-	-	-	87.4	96.7	-	-	-
Older than 65	-	-	-	-	70.2	-	-	-	-
Gender									
Male	-	-	-	80.6	78.3	89.9	-	-	-
Female	-	-	-	83.6	77.5	94.9	-	-	-
Sampling domain									
Unemployed age 16–65	-	-	-	-	88.9	-	-	-	-
Not unemployed age 16–34	-	-	-	-	77.6	-	-	-	-
Age 66–74	-	-	-	-	70.5	-	-	-	-

[1] Age categories were altered for the National Supplement to reflect the analytic domains of interest.
[2] The first and second quartiles were collapsed due to small sample sizes.
- The item was not available for both respondents and nonrespondents, and therefore the response rate could not be computed.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Table A-1a. Values of the quartiles for the Census tract-level quartile variables

	2012 Main Study	2014 Area and List Samples
Categorized average household size		
1st quartile or less	2.296 or less	2.33 or less
(1st quartile – median]	2.297 – 2.554	2.34 – 2.56
(median – 3rd quartile]	2.555 – 2.8218	2.57 – 2.85
More than 3rd quartile	2.8219 or more	2.86 or more
Percentage of the population below 150 percent of poverty		
1st quartile or less	10.280 or less	11.92 or less
(1st quartile – median]	10.290 – 19.390	11.93 – 22.95
(median – 3rd quartile]	19.391 – 31.050	22.96 – 34.99
More than 3rd quartile	31.060 or more	35.00 or more
Percentage of the population that is foreign born		
1st quartile or less	2.620 or less	2.42 or less
(1st quartile – median]	2.630 – 7.270	2.43 – 7.46
(median – 3rd quartile]	7.271 – 17.740	7.47 – 19.45
More than 3rd quartile	17.750 or more	19.46 or more
Percentage of the population age 25 and older with at least a high school education		
1st quartile or less	77.80 or less	78.07 or less
(1st quartile – median]	77.81 – 87.42	78.08 – 87.88
(median – 3rd quartile]	87.43 – 93.45	87.89 – 93.78
More than 3rd quartile	93.46 or more	93.79 or more
Percentage of households that are linguistically isolated		
1st quartile or less	0.260 or less	0.14 or less
(1st quartile – median]	0.270 – 1.690	0.15 – 1.87
(median – 3rd quartile]	1.700 – 5.103	1.88 – 5.70
More than 3rd quartile	5.104 or more	5.71 or more
Percentage of housing units occupied by owner		
1st quartile or less	52.33 or less	48.67 or less
(1st quartile – median]	52.34 – 71.71	48.68 – 68.95
(median – 3rd quartile]	71.72 – 83.50	68.96 – 81.34
More than 3rd quartile	83.60 or more	81.35 or more
Percentage of the population age 18–64 that is unemployed		
1st quartile or less	3.0900 or less	4.48 or less
(1st quartile – median]	3.1000 – 4.7780	4.49 – 6.40
(median – 3rd quartile]	4.7790 – 6.6652	6.41 – 9.34
More than 3rd quartile	6.6653 or more	9.35 or more
Percentage of the population that is Hispanic		
1st quartile or less	1.5422 or less	1.99 or less
(1st quartile – median]	1.5423 – 4.4652	2.00 – 5.33
(median – 3rd quartile]	4.4653 – 13.230	5.34 – 15.72
More than 3rd quartile	13.240 or more	15.73 or more

See notes at end of table.

Table A-1a. Values of the quartiles for the Census tract-level quartile variables—Continued

	2012 Main Study	2014 Area and List Samples
Percentage of the population that is Hispanic or Non-Hispanic Black		
1st quartile or less	4.62 or less	5.44 or less
(1st quartile – median]	4.63 – 13.60	5.45 – 16.38
(median – 3rd quartile]	13.61 – 37.20	16.39 – 45.56
More than 3rd quartile	37.30 or more	45.57 or more
Percentage of the population age 25 and older with some college education		
1st quartile or less	21.53 or less	22.40 or less
(1st quartile – median]	21.54 – 26.22	22.41 – 27.72
(median – 3rd quartile]	23.23 – 31.17	27.73 – 32.60
More than 3rd quartile	31.18 or more	32.61 or more
Percentage of the population age 25 and older with a high school education		
1st quartile or less	20.20 or less	20.35 or less
(1st quartile – median]	20.30 – 28.94	20.36 – 28.56
(median – 3rd quartile]	28.95 – 36.74	28.57 – 35.64
More than 3rd quartile	36.75 or more	35.65 or more
Percentage of the population age 18–64 that is employed		
1st quartile or less	67.59 or less	64.53 or less
(1st quartile – median]	67.60 – 74.01	64.54 – 71.66
(median – 3rd quartile]	74.02 – 78.48	71.67 – 77.00
More than 3rd quartile	78.49 or more	77.01 or more

NOTE: For the Main Study, the quartiles for the tract-level variables are based on the distribution of the Main Study tracts using ACS 2005–2009 data. For the National Supplement, the quartiles are based on the distribution of the tracts in the National Supplement using ACS 2008–2012 data.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Table A-2a. Results of Screener bivariate analysis: 2012 Main Study

Analysis variable	Respondents		Nonrespondents		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	*p* value
Indicator of whether the PSU is part of a Metropolitan Statistical Area					2.7189	0.0992
Nonmetro. area	15.6	1.50	11.1	2.37		
Metro. area	84.4	1.50	88.9	2.37		
Region					7.7991	0.0368
Northeast	20.6	3.67	25.5	4.84		
Midwest	22.2	2.89	18.5	2.87		
South	40.7	3.68	36.1	3.74		
West	16.5	3.08	20.0	3.90		
Percentage of housing units occupied by owner					2.5402	0.3808
52.33 or less	22.4	2.32	22.3	2.68		
52.34 – 71.71	25.1	1.74	23.5	2.10		
71.72 – 83.50	27.4	1.94	26.0	2.30		
83.60 or more	25.2	1.64	28.2	2.74		
Percentage of the population age 25 and older with at least a high school education					17.0706	0.0006
77.80 or less	24.8	2.57	19.1	2.57		
77.81 – 87.42	25.4	2.56	21.9	2.64		
87.43 – 93.45	25.2	2.21	30.4	2.60		
93.46 or more	24.5	2.76	28.6	3.16		
Percentage of the population that is Hispanic or Non-Hispanic Black					6.0971	0.0847
4.62 or less	27.1	2.47	21.1	2.78		
4.63 – 13.60	25.2	1.95	29.0	2.52		
13.61 – 37.20	24.7	2.32	26.4	2.98		
37.30 or more	23.1	2.06	23.6	2.68		
Percentage of the population that is Hispanic					10.4419	0.0112
1.5422 or less	25.0	2.47	18.7	2.23		
1.5423 – 4.4652	26.7	2.03	25.7	2.46		
4.4653 – 13.230	25.3	2.03	29.6	2.56		
13.240 or more	23.0	2.00	26.0	3.08		
Percentage of the population age 18–64 that is unemployed					4.9813	0.1312
3.0900 or less	24.9	1.73	27.6	2.71		
3.1000 – 4.7780	26.5	1.68	28.1	2.27		
4.7790 – 6.6652	25.7	1.43	22.4	1.80		
6.6653 or more	22.9	1.86	22.0	2.62		

See note at end of table.

Table A-2a. Results of Screener bivariate analysis: 2012 Main Study—Continued

Analysis variable	Respondents		Nonrespondents		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	*p* value
Percentage of the population below 150 percent of poverty					25.9633	<.0001
10.280 or less	23.7	2.60	32.7	3.43		
10.290 – 19.390	26.0	1.94	28.0	2.75		
19.391 – 31.050	25.8	2.21	20.1	2.47		
31.060 or more	24.5	2.13	19.2	2.87		
Percentage of the population that is foreign born					31.9742	<.0001
2.620 or less	27.9	3.32	17.0	2.38		
2.630 – 7.270	25.5	2.67	25.1	2.93		
7.271 – 17.740	24.2	2.45	31.5	2.83		
17.750 or more	22.3	2.52	26.5	2.88		
Percentage of households that are linguistically isolated					10.6839	0.0110
0.260 or less	26.4	2.41	20.7	1.72		
0.270 – 1.690	25.8	1.85	23.8	2.51		
1.700 – 5.103	24.7	2.11	28.6	2.80		
5.104 or more	23.2	2.09	27.0	3.23		
Categorized average household size					7.3363	0.0551
2.296 or less	22.3	2.37	21.9	2.60		
2.297 – 2.554	26.1	1.88	22.1	2.24		
2.555 – 2.8218	26.3	1.86	26.7	2.62		
2.8219 or more	25.4	2.51	29.3	3.07		
Percentage of the population age 18–64 that is employed					6.1197	0.0833
67.59 or less	24.7	2.48	20.3	2.93		
67.60 – 74.01	25.1	2.02	24.2	2.71		
74.02 – 78.48	24.8	1.50	26.9	2.49		
78.49 or more	25.4	2.97	28.6	3.64		
Percentage of the population age 25 and older with a high school education					28.1600	<.0001
20.20 or less	23.7	2.86	30.9	3.12		
20.30 – 28.94	23.9	1.79	25.5	2.42		
28.95 – 36.74	25.9	2.37	25.4	2.48		
36.75 or more	26.5	3.29	18.1	2.25		
Percentage of the population age 25 and older with some college education					5.7646	0.1131
21.53 or less	24.2	3.25	26.3	3.83		
21.54 – 26.22	24.1	1.91	19.6	2.34		
23.23 – 31.17	24.7	2.03	26.0	2.32		
31.18 or more	27.0	2.60	28.1	3.15		

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012.

Table A-2b. Results of Screener bivariate analysis: 2014 Area Sample

Analysis variable	Respondents Percent	Respondents Standard error	Nonrespondents Percent	Nonrespondents Standard error	Chi-square Statistic	Chi-square *p* value
Indicator of whether the PSU is part of a Metropolitan Statistical Area					5.1834	0.0228
Nonmetro. area	16.25	1.79	11.24	2.51		
Metro. area	83.75	1.79	88.76	2.51		
Region					10.6675	0.0089
Northeast	20.26	3.99	28.20	4.80		
Midwest	21.45	2.57	16.39	2.37		
South	41.13	3.83	37.67	3.95		
West	17.15	3.08	17.73	3.76		
Percentage of housing units occupied by owner					4.1056	0.1587
48.67 or less	21.42	2.04	25.10	3.49		
48.68 – 68.95	26.06	1.97	22.81	2.39		
68.96 – 81.34	27.00	2.24	25.04	2.36		
81.35 or more	25.52	1.65	27.06	2.45		
Percentage of the population age 25 and older with at least a high school education					19.7483	<.0001
78.07 or less	26.39	2.55	17.45	2.50		
78.08 – 87.88	26.01	2.66	23.64	2.82		
87.89 – 93.78	25.52	2.23	28.43	2.57		
93.79 or more	22.07	2.55	30.48	3.43		
Percentage of the population that is Hispanic or Non-Hispanic Black					15.0998	0.0016
5.44 or less	27.40	2.47	21.35	2.43		
5.45 – 16.38	23.61	1.87	31.72	2.59		
16.39 – 45.56	25.69	2.39	26.60	2.90		
45.57 or more	23.30	2.03	20.33	2.59		
Percentage of the population that is Hispanic					8.7083	0.0287
1.99 or less	27.06	2.31	21.52	2.00		
2.00 – 5.33	22.31	1.82	27.62	2.97		
5.34 – 15.72	24.25	2.27	26.56	2.89		
15.73 or more	26.38	2.31	24.30	3.08		
Percentage of the population age 18–64 that is unemployed					2.4912	0.4624
4.48 or less	25.46	2.29	27.94	2.99		
4.49 – 6.40	25.59	1.98	26.57	2.03		
6.41 – 9.34	27.68	2.22	24.39	2.67		
9.35 or more	21.27	1.99	21.09	2.90		

See note at end of table.

Table A-2b. Results of Screener bivariate analysis: 2014 Area Sample—Continued

Analysis variable	Respondents		Nonrespondents		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	*p* value
Percentage of the population below 150 percent of poverty					20.1966	<.0001
11.92 or less	21.87	2.71	31.13	3.21		
11.93 – 22.95	27.00	2.39	25.49	2.58		
22.96 – 34.99	25.62	2.09	23.43	2.58		
35.00 or more	25.51	2.35	19.96	2.79		
Percentage of the population that is foreign born					15.5671	0.0009
2.42 or less	26.93	3.47	19.46	2.75		
2.43 – 7.46	25.93	2.97	23.12	2.77		
7.47 – 19.45	23.78	2.05	31.12	3.29		
19.46 or more	23.36	2.72	26.30	3.27		
Percentage of households that are linguistically isolated					11.1055	0.0093
0.14 or less	27.80	2.35	21.81	2.35		
0.15 – 1.87	23.71	1.75	25.43	2.41		
1.88 – 5.70	25.15	2.40	30.62	2.60		
5.71 or more	23.33	2.13	22.14	2.35		
Categorized average household size					3.9854	0.1916
2.33 or less	22.44	2.57	26.62	3.53		
2.34 – 2.56	25.52	2.38	22.88	2.18		
2.57 – 2.85	25.24	2.05	25.95	2.11		
2.86 or more	26.80	2.73	24.54	3.04		
Percentage of the population age 18–64 that is employed					4.2980	0.1922
64.53 or less	25.76	2.64	21.35	3.37		
64.54 – 71.66	25.69	2.10	26.83	2.68		
71.67 – 77.00	25.84	1.85	26.57	2.71		
77.01 or more	22.71	2.72	25.25	3.10		
Percentage of the population age 25 and older with a high school education					7.6578	0.0277
20.35 or less	22.31	2.50	29.57	3.96		
20.36 – 28.56	23.71	2.16	22.82	2.00		
28.57 – 35.64	27.08	2.15	25.75	2.80		
35.65 or more	26.90	3.19	21.86	2.42		
Percentage of the population age 25 and older with some college education					1.5125	0.5168
22.40 or less	24.99	3.29	27.87	4.21		
22.41 – 27.72	23.36	1.53	23.58	1.62		
27.73 – 32.60	25.36	2.13	24.82	2.78		
32.61 or more	26.29	2.55	23.73	3.26		

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Table A-2c. Results of Screener bivariate analysis: 2014 List sample

Analysis variable	Respondents		Nonrespondents		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	*p* value
Indicator of whether the PSU is part of a Metropolitan Statistical Area					1.3094	0.2525
Nonmetro. area	37.66	8.71	49.60	10.54		
Metro. area	62.34	8.71	50.40	10.54		
Region					0.6893	0.6606
Northeast	10.37	4.58	9.28	5.55		
Midwest	27.03	8.05	33.26	9.46		
South	57.47	10.04	51.40	12.56		
West	5.13	2.74	6.07	2.90		
Percentage of housing units occupied by owner					5.7312	0.0855
48.67 or less	16.35	5.40	9.82	3.59		
48.68 – 68.95	27.90	13.23	10.01	3.97		
68.96 – 81.34	45.87	10.62	60.59	9.32		
81.35 or more	9.87	5.19	19.57	8.16		
Percentage of the population age 25 and older with at least a high school education					8.3196	0.0341
78.07 or less	53.63	10.24	30.30	10.53		
78.08 – 87.88	31.27	8.29	46.04	10.33		
87.89 – 93.78	13.65	5.72	19.14	8.10		
93.79 or more	1.45	0.73	4.52	2.53		
Percentage of the population that is Hispanic or Non-Hispanic Black					10.2276	0.0118
5.44 or less	26.34	9.43	41.49	12.41		
5.45 – 16.38	17.89	6.47	37.46	13.33		
16.39 – 45.56	26.55	9.95	13.25	6.22		
45.57 or more	29.22	11.20	7.80	2.04		
Percentage of the population that is Hispanic					2.1939	0.3278
1.99 or less	35.86	9.87	31.04	9.43		
2.00 – 5.33	15.64	5.25	29.30	10.54		
5.34 – 15.72	37.00	13.22	29.66	13.82		
15.73 or more	11.49	5.45	10.00	5.29		
Percentage of the population age 18–64 that is unemployed					4.5167	0.0954
6.40 or less	20.71	6.71	40.89	13.25		
6.41 – 9.34	36.11	11.63	19.82	9.05		
9.35 or more	43.18	9.48	39.30	10.60		
Percentage of the population below 150 percent of poverty					1.4691	0.409
22.95 or less	13.72	5.58	21.40	8.82		
22.96 – 34.99	36.65	10.53	39.57	9.77		
35.00 or more	49.63	11.80	39.03	13.46		

See note at end of table.

Table A-2c. Results of Screener bivariate analysis: 2014 List sample—Continued

Analysis variable	Respondents		Nonrespondents		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	*p* value
Percentage of the population that is foreign born					1.2092	0.4756
2.42 or less	41.15	10.61	48.06	12.92		
2.43 – 7.46	35.46	13.06	22.82	12.52		
7.47 – 19.45	19.04	7.26	24.89	10.01		
19.46 or more	4.35	2.48	4.23	1.75		
Percentage of households that are linguistically isolated					1.8531	0.4589
0.14 or less	32.86	9.01	33.43	9.02		
0.15 – 1.87	32.14	13.74	20.03	10.48		
1.88 – 5.70	23.87	8.06	35.76	13.60		
5.71 or more	11.13	5.40	10.78	5.42		
Categorized average household size					1.169	0.6319
2.33 or less	16.99	6.60	18.69	9.30		
2.34 – 2.56	28.35	9.56	30.96	11.67		
2.57 – 2.85	46.57	13.15	37.12	13.19		
2.86 or more	8.09	3.49	13.22	5.61		
Percentage of the population age 18–64 that is employed					10.0378	0.0061
64.53 or less	44.84	11.91	14.42	5.28		
64.54 – 71.66	32.88	9.58	45.52	10.41		
71.67 or more	22.28	6.27	40.06	11.20		
Percentage of the population age 25 and older with a high school education					2.3236	0.2902
20.35 or less	1.84	1.03	1.89	1.27		
20.36 – 28.56	15.43	6.19	23.64	9.53		
28.57 – 35.64	34.26	11.40	19.75	7.65		
35.65 or more	48.47	12.61	54.71	10.94		
Percentage of the population age 25 and older with some college education					2.0064	0.4898
22.40 or less	29.63	13.51	15.05	7.58		
22.41 – 27.72	26.34	9.98	34.81	14.16		
27.73 – 32.60	24.96	10.15	28.76	11.31		
32.61 or more	19.07	5.32	21.39	7.54		

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Table A-3a. PIAAC multivariate analysis of Screener response indicators, by response cell: 2012 Main Study

Response cell	Weighted response rate (percent)
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Percentage of the population below 150 percent of poverty in 1st and 2nd quartiles & Non-MSA	75.7
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Percentage of the population below 150 percent of poverty in 1st and 2nd quartiles & MSA & Percentage of the population age 18–64 that is unemployed in 4th quartile	80.1
Percentage of the population that is foreign born in 1st quartile & Categorized average household size in 4th quartile	86.0
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Percentage of the population below 150 percent of poverty in 3rd and 4th quartiles & Census Region West, Midwest, and Northeast & Non-MSA	96.0
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Percentage of the population below 150 percent of poverty in 3rd and 4th quartiles & Census Region West, Midwest, and Northeast & MSA & Percentage of the population age 18–64 that is unemployed in 1st and 2nd quartiles	81.4
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & MSA & Percentage of the population age 18–64 that is unemployed in 1st, 2nd, and 3rd quartiles & Percentage of the population below 150 percent of poverty in 2nd quartile & Census Region West	90.1
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & MSA & Percentage of the population age 18–64 that is unemployed in 1st, 2nd, and 3rd quartiles & Percentage of the population below 150 percent of poverty in 1st quartile & Percentage of the population that is Hispanic in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st quartile	84.9
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Percentage of the population below 150 percent of poverty in 3rd and 4th quartiles & Census Region South & Percentage of the population that is Hispanic or Non-Hispanic Black in 1st, 2nd, and 3rd Quartiles	93.6
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & MSA & Percentage of the population age 18–64 that is unemployed in 1st, 2nd, and 3rd quartiles & Percentage of the population below 150 percent of poverty in 2nd quartile & Census Region Northeast, Midwest, and South & Percentage of the population that is Hispanic or Non-Hispanic Black in 1st and 2nd quartile	82.8

See note at end of table.

Table A-3a. PIAAC multivariate analysis of Screener response indicators, by response cell: 2012 Main Study—Continued

Response cell	Weighted response rate (percent)
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & MSA & Percentage of the population age 18–64 that is unemployed in 1st, 2nd, and 3rd quartiles & Percentage of the population below 150 percent of poverty in 2nd quartile & Census Region Northeast, Midwest, and South & Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles	87.6
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & MSA & Percentage of the population age 18–64 that is unemployed in 1st, 2nd, and 3rd quartiles & Percentage of the population below 150 percent of poverty in 1st quartile & Percentage of the population that is Hispanic in 1st and 2nd quartiles & Categorized average household size in 1st and 2nd quartiles	90.8
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Percentage of the population below 150 percent of poverty in 3rd and 4th quartiles & Census Region West, Midwest, and Northeast & MSA & Percentage of the population age 18–64 that is unemployed in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 4th quartile	83.7
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & MSA & Percentage of the population age 18–64 that is unemployed in 1st, 2nd, and 3rd quartiles & Percentage of the population below 150 percent of poverty in 1st quartile & Percentage of the population that is Hispanic in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles & Percentage of households that are linguistically isolated in 1st quartile	90.3
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & MSA & Percentage of the population age 18–64 that is unemployed in 1st, 2nd, and 3rd quartiles & Percentage of the population below 150 percent of poverty in 1st quartile & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles & Percentage of households that are linguistically isolated in 2nd, 3rd, and 4th quartiles & Percentage of the population that is Hispanic in 4th quartile	76.6
Percentage of the population that is foreign born in 1st quartile & Census Region South and Northeast & Categorized average household size in 1st and 2nd quartiles	93.2
Percentage of the population that is foreign born in 1st quartile & Census Region South and Northeast & Categorized average household size in 3rd quartile	88.5

See note at end of table.

Table A-3a. PIAAC multivariate analysis of Screener response indicators, by response cell: 2012 Main Study—Continued

Response cell	Weighted response rate (percent)
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Percentage of the population below 150 percent of poverty in 3rd and 4th quartiles & Census Region South & Percentage of the population that is Hispanic or Non-Hispanic Black in 4th quartile & Percentage of the population age 25 and older with a high school education in 1st and 2nd quartiles	91.3
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Percentage of the population below 150 percent of poverty in 3rd and 4th quartiles & Census Region South & Percentage of the population that is Hispanic or Non-Hispanic Black in 4th quartile & Percentage of the population age 25 and older with a high school education in 3rd and 4th quartiles	86.9
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & MSA & Percentage of the population age 18–64 that is unemployed in 1st, 2nd, and 3rd quartiles & Percentage of the population below 150 percent of poverty in 1st quartile & Percentage of the population that is Hispanic in 1st and 2nd quartiles & Categorized average household size in 3rd and 4th quartiles & Census Region South	89.1
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & MSA & Percentage of the population age 18–64 that is unemployed in 1st, 2nd, and 3rd quartiles & Percentage of the population below 150 percent of poverty in 1st quartile & Percentage of the population that is Hispanic in 1st and 2nd quartiles & Categorized average household size in 3rd and 4th quartiles & Census Region West, Midwest, and Northeast	82.1
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & MSA & Percentage of the population age 18–64 that is unemployed in 1st, 2nd, and 3rd quartiles & Percentage of the population below 150 percent of poverty in 1st quartile & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles & Percentage of households that are linguistically isolated in 2nd, 3rd, and 4th quartiles & Percentage of the population that is Hispanic in 3rd quartile & Percentage of the population that is Hispanic or Non-Hispanic Black in 2nd quartile	77.8

See note at end of table.

Table A-3a. PIAAC multivariate analysis of Screener response indicators, by response cell: 2012 Main Study—Continued

Response cell	Weighted response rate (percent)
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & MSA & Percentage of the population age 18–64 that is unemployed in 1st, 2nd, and 3rd quartiles & Percentage of the population below 150 percent of poverty in 1st quartile & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles & Percentage of households that are linguistically isolated in 2nd, 3rd, and 4th quartiles & Percentage of the population that is Hispanic in 3rd quartile & Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles	85.0
Percentage of the population that is foreign born in 1st quartile & Categorized average household size in 1st, 2nd, and 3rd quartiles & Census Region West and Midwest & Percentage of the population that is Hispanic in 1st and 2nd quartiles	94.7
Percentage of the population that is foreign born in 1st quartile & Categorized average household size in 1st, 2nd, and 3rd quartiles & Census Region West and Midwest & Percentage of the population that is Hispanic in 3rd quartile	82.1
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Percentage of the population below 150 percent of poverty in 3rd and 4th quartiles & Census Region West, Midwest, and Northeast & MSA & Percentage of the population age 18–64 that is unemployed in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st, 2nd, and 3rd quartiles & Percentage of households that are linguistically isolated in 2nd quartile	97.0
Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Percentage of the population below 150 percent of poverty in 3rd and 4th quartiles & Census Region West, Midwest, and Northeast & MSA & Percentage of the population age 18–64 that is unemployed in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st, 2nd, and 3rd quartiles & Percentage of households that are linguistically isolated in 3rd and 4th quartiles	88.6

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012.

Table A-3b. PIAAC multivariate analysis of Screener response indicators, by response cell: 2014 Area Sample

Response cell	Weighted response rate (percent)
Percentage of the population age 25 and older with at least a high school education in 1st quartile & Census Region Northeast	84.3
Percentage of the population age 25 and older with at least a high school education in 1st quartile & Census Region Midwest, West, and South & Percentage of the population that is Hispanic or Non-Hispanic Black in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is unemployed in 1st and 2nd quartiles	83.4
Percentage of the population age 25 and older with at least a high school education in 1st quartile & Census Region Midwest, West, and South & Percentage of the population that is Hispanic or Non-Hispanic Black in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is unemployed in 3rd and 4th quartiles	90.3
Percentage of the population age 25 and older with at least a high school education in 1st quartile & Census Region Midwest, West, and South & Percentage of the population that is Hispanic or Non-Hispanic Black in 4th quartile & Percentage of the population age 25 and older with some college education in 1st quartile	93.3
Percentage of the population age 25 and older with at least a high school education in 1st quartile & Census Region Midwest, West, and South & Percentage of the population that is Hispanic or Non-Hispanic Black in 4th quartile & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles	90.6
Percentage of the population age 25 and older with at least a high school education in 2nd quartile & Census Region Midwest and West & Percentage of the population that is Hispanic in 1st quartile	86.7
Percentage of the population age 25 and older with at least a high school education in 2nd quartile & Census Region Midwest and West & Percentage of the population that is Hispanic in 2nd, 3rd, and 4th quartiles	90.7
Percentage of the population age 25 and older with at least a high school education in 2nd quartile & Census Region Northeast and South & Non-MSA	89.9

See note at end of table.

Table A-3b. PIAAC multivariate analysis of Screener response indicators, by response cell: 2014 Area Sample—Continued

Response cell	Weighted response rate (percent)
Percentage of the population age 25 and older with at least a high school education in 2nd quartile & Census Region Northeast and South & MSA & Percentage of the population age 25 and older with some college education in 1st and 2nd quartiles & Percentage of the population below 150 percent of poverty in 1st and 2nd quartiles	82.8
Percentage of the population age 25 and older with at least a high school education in 2nd quartile & Census Region Northeast and South & MSA & Percentage of the population age 25 and older with some college education in 1st and 2nd quartiles & Percentage of the population below 150 percent of poverty in 3rd and 4th quartiles	70.9
Percentage of the population age 25 and older with at least a high school education in 2nd quartile & Census Region Northeast and South & MSA & Percentage of the population age 25 and older with some college education in 3rd and 4th quartiles & Percentage of the population age 25 and older with a high school education in 1st and 2nd quartiles	89.2
Percentage of the population age 25 and older with at least a high school education in 2nd quartile & Census Region Northeast and South & MSA & Percentage of the population age 25 and older with some college education in 3rd and 4th quartiles & Percentage of the population age 25 and older with a high school education in 3rd and 4th quartiles	81.6
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartiles & Percentage of the population that is foreign born in 1st and 2nd quartiles & Census Region Midwest and South & Percentage of households that are linguistically isolated in 1st quartile & Categorized average household size in 1st and 2nd quartiles	91.0
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartiles & Percentage of the population that is foreign born in 1st and 2nd quartiles & Census Region Midwest and South & Percentage of households that are linguistically isolated in 1st quartile & Categorized average household size in 3rd and 4th quartiles	87.7

See note at end of table.

Table A-3b. PIAAC multivariate analysis of Screener response indicators, by response cell: 2014 Area Sample—Continued

Response cell	Weighted response rate (percent)
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartiles & Percentage of the population that is foreign born in 1st and 2nd quartiles & Census Region Midwest and South & Percentage of households that are linguistically isolated in 2nd, 3rd, and 4th quartiles & Percentage of the population age 18–64 that is employed in 1st, 2nd, and 3rd quartiles	82.8
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartile & Percentage of the population that is foreign born in 1st and 2nd quartiles & Census Region Midwest and South & Percentage of households that are linguistically isolated in 2nd, 3rd, and 4th quartiles & Percentage of the population age 18–64 that is employed in 4th quartile	89.4
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartile & Percentage of the population that is foreign born in 1st and 2nd quartiles & Census Region West and Northeast & Percentage of the population that is Hispanic or Non-Hispanic Black in 1st quartile	85.3
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartile & Percentage of the population that is foreign born in 1st and 2nd quartiles & Census Region West and Northeast & Percentage of the population that is Hispanic or Non-Hispanic Black in 2nd and 3rd quartiles	76.9
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartile & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is employed in 1st and 2nd quartiles	66.0
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartile & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is employed in 3rd and 4th quartiles & Census Region South and Midwest & Percent of housing units occupied by owner in 1st quartile	78.0

See note at end of table.

Table A-3b. PIAAC multivariate analysis of Screener response indicators, by response cell: 2014 Area Sample—Continued

Response cell	Weighted response rate (percent)
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartile & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is employed in 3rd and 4th quartiles & Census Region South and Midwest & Percent of housing units occupied by owner in 2nd, 3rd, and 4th quartiles	84.1
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartile & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is employed in 3rd and 4th quartiles & Census Region West and Northeast & Percent of housing units occupied by owner in 1st quartile	61.0
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartile & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is employed in 3rd and 4th quartiles & Census Region West and Northeast & Percent of housing units occupied by owner in 2nd quartile	84.7
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartile & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is employed in 3rd and 4th quartiles & Census Region West and Northeast & Percent of housing units occupied by owner in 3rd and 4th quartiles	71.9
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartile & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 4th quartile & Percentage of the population that is Hispanic in 1st and 2nd quartiles	72.6

See note at end of table.

Table A-3b. PIAAC multivariate analysis of Screener response indicators, by response cell: 2014 Area Sample—Continued

Response cell	Weighted response rate (percent)
Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartile &	
Percentage of the population that is foreign born in 3rd and 4th quartiles &	85.3
Percentage of the population age 25 and older with some college education in 4th quartile &	
Percentage of the population that is Hispanic in 3rd and 4th quartiles	

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Table A-3c. PIAAC multivariate analysis of Screener response indicators, by response cell: 2014 List Sample

Response cell	Weighted response rate (percent)
Percentage of the population that is Hispanic or Non-Hispanic Black in 1st and 2nd quartiles & Percentage of the population that is Hispanic in 1st quartile & Census Region Northeast and South	94.5
Percentage of the population that is Hispanic or Non-Hispanic Black in 1st and 2nd quartiles & Percentage of the population that is Hispanic in 1st quartile & Census Region and West and Midwest & Percent of housing units occupied by owner in 1st and 2nd quartiles	84.5
Percentage of the population that is Hispanic or Non-Hispanic Black in 1st and 2nd quartiles & Percentage of the population that is Hispanic in 1st quartile & Census Region and West and Midwest & Percent of housing units occupied by owner in 3rd and 4th quartiles	84.2
Percentage of the population that is Hispanic or Non-Hispanic Black in 1st and 2nd quartiles & Percentage of the population that is Hispanic in 2nd and 3rd quartiles & Categorized average household size in 1st and 2nd quartiles	82.6
Percentage of the population that is Hispanic or Non-Hispanic Black in 1st and 2nd quartiles & Percentage of the population that is Hispanic in 2nd and 3rd quartiles & Categorized average household size in 3rd and 4th quartiles & Percentage of the population age 25 and older with at least a high school education in 1st and 2nd quartiles	61.4
Percentage of the population that is Hispanic or Non-Hispanic Black in 1st and 2nd quartiles & Percentage of the population that is Hispanic in 2nd and 3rd quartiles & Categorized average household size in 3rd and 4th quartiles & Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartiles	76.2
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 1st and 2nd quartiles & Census Region and West and Northeast	88.8
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 1st and 2nd quartiles & Census Region and Midwest and South & Categorized average household size in 1st quartile	95.0

See note at end of table.

Table A-3c. PIAAC multivariate analysis of Screener response indicators, by response cell: 2014 List Sample—Continued

Response cell	Weighted response rate (percent)
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 1st and 2nd quartiles & Census Region and Midwest and South & Categorized average household size in 2nd, 3rd, 4th quartiles & Percentage of the population that is Hispanic in 1st and 2nd quartiles & Categorized average household size in 2nd and 3rd quartiles	98.3
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 1st and 2nd quartiles & Census Region and Midwest and South & Categorized average household size in 2nd, 3rd, 4th quartiles & Percentage of the population that is Hispanic in 1st and 2nd quartiles & Categorized average household size in 4th quartile	94.7
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 1st and 2nd quartiles & Census Region and Midwest and South & Categorized average household size in 2nd, 3rd, 4th quartiles & Percentage of the population that is Hispanic in 3rd and 4th quartiles	99.4
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st quartile & Categorized average household size in 1st and 2nd quartiles	80.3
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st quartile & Categorized average household size in 3rd and 4th quartiles & Percent of housing units occupied by owner in 1st quartile	85.3
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st quartile & Categorized average household size in 3rd and 4th quartiles & Percent of housing units occupied by owner in 2nd, 3rd, and 4th quartiles	87.2

See note at end of table.

Table A-3c. PIAAC multivariate analysis of Screener response indicators, by response cell: 2014 List Sample—Continued

Response cell	Weighted response rate (percent)
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles & Census Region and South and West & Percentage of the population below 150 percent of poverty in 2nd and 3rd quartiles	90.9
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles & Census Region and South and West & Percentage of the population below 150 percent of poverty in 4th quartile & Census Region and West	93.5
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles & Census Region and South and West & Percentage of the population below 150 percent of poverty in 4th quartile & Census Region and South	97.4
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles & Census Region Northeast and Midwest & Percentage of the population that is foreign born in 3rd quartile & Percentage of the population age 25 and older with some college education in 2nd and 3rd quartiles	92.7
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles & Census Region Northeast and Midwest & Percentage of the population that is foreign born in 3rd quartile & Percentage of the population age 25 and older with some college education in 4th quartile	89.0

See note at end of table.

Table A-3c. PIAAC multivariate analysis of Screener response indicators, by response cell: 2014 List Sample—Continued

Response cell	Weighted response rate (percent)
Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles & Percentage of the population that is foreign born in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles & Census Region Northeast and Midwest & Percentage of the population that is foreign born in 4th quartile	80.6

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Table A-4a. PIAAC multivariate analysis of Screener response indicators: 2012 Main Study

Analysis variables	*F* statistic	Degrees of freedom		*p* value
		Numerator	Denominator	
Overall fit	2.705	16	30	0.009
Indicator of whether the PSU is part of a Metropolitan Statistical Area	0.054	1	45	0.818
Region	1.328	3	43	0.278
Categorized average household size	0.006	1	45	0.939
Percentage of the population below 150 percent of poverty	4.593	1	45	0.038
Percentage of the population age 18–64 that is unemployed	0.157	1	45	0.693
Percentage of the population that is foreign born	1.982	1	45	0.166
Percentage of the population age 25 and older with at least a high school education	0.690	1	45	0.411
Percentage of households that are linguistically isolated	0.396	1	45	0.532
Percentage of housing units occupied by owner	0.688	1	45	0.411
Percentage of the population that is Hispanic	0.493	1	45	0.486
Percentage of the population that is Hispanic or Non-Hispanic Black	1.245	1	45	0.270
Percentage of the population age 25 and older with some college education	0.271	1	45	0.605
Percentage of the population age 25 and older with a high school education	1.039	1	45	0.314
Percentage of the population age 18–64 that is unemployed	0.019	1	45	0.890

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012.

Table A-4b. PIAAC multivariate analysis of Screener response indicators: 2014 Area Sample

Analysis variables	Degrees of freedom	Wald Chi-Square	*p* value
Percentage of the population age 25 and older with at least a high school education	1	15.0453	0.0001
Percentage of the population that is foreign born	1	7.6817	0.0056
Census region	3	7.6873	0.0529
Percentage of the population age 18–64 that is employed	1	1.3921	0.2381

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Table A-4c. PIAAC multivariate analysis of Screener response indicators: 2014 List Sample

Analysis variables	Degrees of freedom	Wald Chi-Square	*p* value
Percentage of the population that is Hispanic or Non-Hispanic Black	1	43.3689	<0.0001
Percentage of the population that is Hispanic	1	6.7184	0.0095
Percentage of the population age 25 and older with at least a high school education	1	8.0477	0.0046
Census region	3	5.9823	0.1125
Percentage of the population that is foreign born	1	2.5846	0.1079

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Table A-5a. PIAAC results of Background Questionnaire bivariate analysis: 2012 Main Study

Analysis variable	Respondents Percent	Respondents Standard error	Nonrespondents Percent	Nonrespondents Standard error	Chi-square Statistic	Chi-square *p* value
Indicator of whether the PSU is part of a Metropolitan Statistical Area					1.5453	0.2138
Nonmetro. Area	15.4	1.42	12.8	2.20		
Metro. Area	84.6	1.42	87.2	2.20		
Region					8.9051	0.0259
Northeast	19.4	3.70	26.3	4.69		
Midwest	22.8	2.77	19.2	3.63		
South	39.8	3.81	37.3	3.76		
West	17.9	3.29	17.2	3.47		
Percentage of housing units occupied by owner					4.7618	0.1456
52.33 or less	20.9	2.17	20.9	2.63		
52.34 – 71.71	24.3	1.82	23.4	1.90		
71.72 – 83.50	27.9	1.96	24.9	2.28		
83.60 or more	26.9	1.87	30.8	2.56		
Percentage of the population age 25 and older with at least a high school education					11.9699	0.0055
77.80 or less	25.0	2.74	20.1	2.56		
77.81 – 87.42	25.6	2.49	23.0	3.12		
87.43 – 93.45	24.7	2.28	28.3	2.65		
93.46 or more	24.7	2.65	28.8	2.93		
Percentage of the population that is Hispanic or Non-Hispanic Black					5.8024	0.1082
4.62 or less	27.5	2.38	25.2	2.81		
4.63 – 13.60	25.1	2.01	27.8	2.45		
13.61 – 37.20	24.1	2.23	27.2	2.68		
37.30 or more	23.2	2.11	19.8	2.40		
Percentage of the population that is Hispanic					5.6801	0.1229
1.5422 or less	24.8	2.40	23.3	3.10		
1.5423 – 4.4652	26.4	2.01	25.4	2.97		
4.4653 – 13.230	25.2	2.08	30.4	2.77		
13.240 or more	23.6	2.27	20.9	2.08		
Percentage of the population age 18–64 that is unemployed					0.4485	0.9212
3.0900 or less	25.3	1.83	25.3	2.54		
3.1000 – 4.7780	26.9	1.83	27.7	2.39		
4.7790 – 6.6652	25.7	1.40	24.7	2.22		
6.6653 or more	22.1	1.82	22.3	2.20		

See note at end of table.

Table A-5a. PIAAC results of Background Questionnaire bivariate analysis: 2012 Main Study—Continued

Analysis variable	Respondents		Nonrespondents		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	*p* value
Percentage of the population below 150 percent of poverty					13.5705	0.0027
10.280 or less	24.8	2.80	30.9	3.11		
10.290 – 19.390	25.8	2.10	27.5	2.27		
19.391 – 31.050	25.5	2.29	22.5	2.40		
31.060 or more	23.9	2.22	19.1	2.16		
Percentage of the population that is foreign born					7.4088	0.0556
2.620 or less	27.4	3.27	21.7	3.80		
2.630 – 7.270	26.0	2.79	28.3	3.39		
7.271 – 17.740	23.9	2.35	26.8	3.33		
17.750 or more	22.8	2.60	23.2	2.83		
Percentage of households that are linguistically isolated					7.0699	0.0660
0.260 or less	26.5	2.47	22.7	2.57		
0.270 – 1.690	25.5	1.84	28.5	2.81		
1.700 – 5.103	24.5	2.10	27.7	2.56		
5.104 or more	23.6	2.26	21.1	2.52		
Categorized average household size					0.7632	0.7765
2.296 or less	20.4	2.35	20.2	2.31		
2.297 – 2.554	25.1	1.91	24.0	2.51		
2.555 – 2.8218	26.5	1.87	28.3	2.45		
2.8219 or more	28.0	2.88	27.6	2.97		
Percentage of the population age 18–64 that is employed					7.4806	0.0448
67.59 or less	24.6	2.50	19.7	2.30		
67.60 – 74.01	25.2	2.08	26.1	2.67		
74.02 – 78.48	25.2	1.63	27.0	2.17		
78.49 or more	25.1	2.90	27.2	3.46		
Percentage of the population age 25 and older with a high school education					3. 2291	0.3158
20.20 or less	23.9	2.87	27.7	3.21		
20.30 – 28.94	24.7	1.91	23.3	2.17		
28.95 – 36.74	25.6	2.43	25.0	2.87		
36.75 or more	25.8	3.24	24.0	3.63		
Percentage of the population age 25 and older with some college education					2.9739	0.3398
21.53 or less	23.4	3.16	24.7	3.67		
21.54 – 26.22	24.0	1.97	21.3	2.31		
23.23 – 31.17	24.3	2.17	27.0	2.48		
31.18 or more	28.3	2.68	27.1	3.27		

See note at end of table.

Table A-5a. PIAAC results of Background Questionnaire bivariate analysis: 2012 Main Study—Continued

Analysis variable	Respondents		Nonrespondents		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	*p* value
Race/ethnicity (after imputation)					0.6561	0.6725
Hispanic	12.2	1.71	11.2	1.26		
Non-Hispanic Black	12.5	1.50	12.0	1.42		
Other	75.3	1.73	76.9	1.81		
Age category (after imputation)					19.8287	0.0003
16–25	21.5	0.94	16.3	1.47		
26–35	20.2	0.69	17.7	1.43		
36–45	19.7	0.62	22.1	1.14		
46–55	21.6	0.61	23.4	1.34		
56–65	17.0	0.67	20.6	1.53		
Language used at screening					1.4654	0.2261
English	97.6	0.71	98.4	0.52		
Spanish	2.4	0.71	1.6	0.52		
Indicator for children under age 16 in household					18.1768	<.0001
No	59.5	1.12	68.6	1.77		
Yes	40.5	1.12	31.4	1.77		
Gender					6.7792	0.0092
Male	47.0	0.70	52.2	1.75		
Female	53.0	0.70	47.8	1.75		

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012.

Table A-5b. PIAAC results of Background Questionnaire bivariate analysis: 2014 Area Sample

	Respondents		Nonrespondents		Chi-square	
Analysis variable	Percent	Standard error	Percent	Standard error	Statistic	*p* value
Indicator of whether the PSU is part of a Metropolitan Statistical Area					1.7921	0.1807
Nonmetro. Area	15.96	2.13	12.76	1.80		
Metro. Area	84.04	2.13	87.24	1.80		
Region					11.0199	0.0100
Northeast	18.40	3.93	26.41	4.51		
Midwest	21.62	2.43	19.77	3.52		
South	42.02	4.00	38.47	4.03		
West	17.95	3.17	15.35	3.44		
Percentage of housing units occupied by owner					5.3800	0.1261
48.67 or less	23.48	2.23	19.26	2.68		
48.68 – 68.95	25.96	2.05	24.98	3.14		
68.96 – 81.34	27.06	2.57	27.13	2.35		
81.35 or more	23.50	1.64	28.62	2.96		
Percentage of the population age 25 and older with at least a high school education					24.4627	<.0001
78.07 or less	28.19	2.66	17.64	2.63		
78.08 – 87.88	26.67	2.70	23.75	3.46		
87.89 – 93.78	24.08	2.45	31.24	2.93		
93.79 or more	21.06	2.66	27.38	3.52		
Percentage of the population that is Hispanic or Non-Hispanic Black					11.0922	0.0089
5.44 or less	25.89	2.62	28.45	2.83		
5.45 – 16.38	22.87	2.04	27.76	2.86		
16.39 – 45.56	26.01	2.61	26.13	3.09		
45.57 or more	25.24	2.20	17.66	2.67		
Percentage of the population that is Hispanic					3.6683	0.2335
1.99 or less	25.83	2.45	26.46	2.98		
2.00 – 5.33	22.26	1.99	22.79	2.16		
5.34 – 15.72	23.75	2.21	26.89	2.95		
15.73 or more	28.15	2.76	23.85	2.54		

See note at end of table.

Table A-5b. PIAAC results of Background Questionnaire bivariate analysis: 2014 Area Sample—Continued

Analysis variable	Respondents		Nonrespondents		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	*p* value
Percentage of the population age 18–64 that is unemployed					2.7580	0.3801
4.48 or less	25.14	2.42	28.03	3.07		
4.49–6.40	25.88	2.25	27.60	2.56		
6.41–9.34	27.20	2.34	25.42	3.11		
9.35 or more	21.78	2.08	18.95	3.01		
Percentage of the population below 150 percent of poverty					24.9767	<.0001
11.92 or less	19.88	2.63	28.94	3.64		
11.93 – 22.95	26.80	2.34	28.57	2.91		
22.96 – 34.99	25.71	2.36	23.99	2.21		
35.00 or more	27.62	2.36	18.50	2.78		
Percentage of the population that is foreign born					2.7829	0.4170
2.42 or less	26.71	3.53	22.97	3.77		
2.43 – 7.46	24.65	2.98	26.82	3.42		
7.47 – 19.45	24.26	2.44	26.27	2.14		
19.46 or more	24.38	3.02	23.94	3.28		
Percentage of households that are linguistically isolated					4.1422	0.2039
0.14 or less	27.82	2.52	24.56	2.92		
0.15 – 1.87	23.16	1.78	24.87	2.45		
1.88 – 5.70	24.31	2.54	28.41	3.51		
5.71 or more	24.71	2.40	22.16	3.15		
Categorized average household size					3.8189	0.2730
2.33 or less	21.47	2.50	25.36	3.36		
2.34 – 2.56	25.50	2.62	25.46	2.84		
2.57 – 2.85	25.16	2.27	23.68	2.20		
2.86 or more	27.87	3.14	25.49	2.80		
Percentage of the population age 18–64 that is employed					20.6847	<.0001
64.53 or less	27.51	2.74	16.82	2.70		
64.54 – 71.66	24.58	2.16	29.53	2.67		
71.67 – 77.00	25.62	2.16	28.53	2.61		
77.01 or more	22.29	2.87	25.11	3.26		

See note at end of table.

Table A-5b. PIAAC results of Background Questionnaire bivariate analysis: 2014 Area Sample—Continued

Analysis variable	Respondents		Nonrespondents		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	*p* value
Percentage of the population age 25 and older with a high school education					6.0343	0.1038
20.35 or less	21.03	2.63	25.94	3.09		
20.36 – 28.56	24.31	2.31	25.25	3.17		
28.57 – 35.64	27.12	2.21	25.85	2.92		
35.65 or more	27.55	3.45	22.97	3.29		
Percentage of the population age 25 and older with some college education					1.3784	0.6221
22.40 or less	25.08	3.47	25.34	3.53		
22.41 – 27.72	23.06	1.57	23.72	1.70		
27.73 – 32.60	24.94	2.05	26.62	3.11		
32.61 or more	26.92	2.90	24.32	3.05		
Race/ethnicity (after imputation)					11.9217	0.0022
Hispanic	15.23	2.04	10.57	1.92		
Non-Hispanic Black	14.97	2.00	10.62	1.80		
Other	69.80	1.80	78.81	2.36		
Age category (after imputation)					59.8079	<.0001
16–24	38.43	1.37	31.77	2.13		
25–34	37.53	1.29	37.44	2.11		
35–45	3.48	0.31	0.77	0.21		
46–55	2.25	0.32	1.83	0.51		
56–65	1.78	0.23	0.88	0.23		
Older than 65	16.54	0.70	27.32	1.70		
Language used at screening					1.6663	0.1968
English	97.43	0.92	98.57	0.68		
Spanish	2.57	0.92	1.43	0.68		
Indicator for children under age 16 in household					35.3441	<.0001
No	60.93	1.34	75.15	1.98		
Yes	39.07	1.34	24.85	1.98		
Gender					0.3801	0.5375
Male	49.51	0.88	48.41	1.73		
Female	50.49	0.88	51.59	1.73		

See note at end of table.

Table A-5b. PIAAC results of Background Questionnaire bivariate analysis: 2014 Area Sample—Continued

Analysis variable	Respondents		Nonrespondents		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	*p* value
Sampling domain					65.6865	<.0001
Unemployed age 16–65	16.99	1.13	7.27	0.73		
Not unemployed age 16–34	66.31	1.32	65.41	1.76		
Age 66–74	16.70	0.70	27.32	1.70		

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Table A-5c. PIAAC results of Background Questionnaire bivariate analysis: 2014 List Sample

Analysis variable	Respondents Percent	Respondents Standard error	Nonrespondents Percent	Nonrespondents Standard error	Chi-square Statistic	Chi-square *p* value
Indicator of whether the PSU is part of a Metropolitan Statistical Area					8.2278	0.0041
Nonmetro. Area	34.89	9.15	58.77	9.52		
Metro. Area	65.11	9.15	41.23	9.52		
Region					6.1571	0.0839
Northeast	9.05	3.88	21.59	9.31		
Midwest	28.10	8.76	15.95	8.23		
South	57.82	10.44	58.96	11.77		
West	5.03	2.75	3.50	1.85		
Percentage of housing units occupied by owner					1.3745	0.5862
48.67 or less	16.62	5.80	12.75	4.32		
48.68 – 68.95	29.86	14.70	22.55	10.36		
68.96 – 81.34	44.33	11.16	49.51	10.55		
81.35 or more	9.19	5.14	15.19	9.23		
Percentage of the population age 25 and older with at least a high school education					6.6159	0.0713
78.07 or less	55.46	10.69	45.37	11.05		
78.08 – 87.88	29.83	8.79	46.79	10.06		
87.89 – 93.78	13.40	5.76	5.49	3.56		
93.79 or more	1.32	0.68	2.35	1.74		
Percentage of the population that is Hispanic or Non-Hispanic Black					1.4840	0.6411
5.44 or less	25.25	9.55	33.63	13.04		
5.45 – 16.38	16.49	6.05	14.71	8.59		
16.39 – 45.56	26.61	10.14	29.24	12.77		
45.57 or more	31.64	12.41	22.42	7.78		
Percentage of the population that is Hispanic					0.3102	0.9331
1.99 or less	35.27	10.18	39.11	11.88		
2.00 – 5.33	14.95	5.00	16.92	10.01		
5.34 – 15.72	38.24	14.29	34.53	13.37		
15.73 or more	11.55	6.10	9.44	3.95		
Percentage of the population age 18–64 that is unemployed					0.0893	0.9143
6.40 or less	18.65	6.51	21.78	11.79		
6.41 – 9.34	37.97	12.45	36.69	12.88		
9.35 or more	43.38	10.26	41.53	10.80		

See note at end of table

Table A-5c. PIAAC results of Background Questionnaire bivariate analysis: 2014 List Sample—Continued

Analysis variable	Respondents		Nonrespondents		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	*p* value
Percentage of the population below 150 percent of poverty					6.5109	0.0264
22.95 or less	12.70	5.35	13.07	7.52		
22.96 – 34.99	34.99	10.85	57.04	9.61		
35.00 or more	52.31	12.18	29.89	11.02		
Percentage of the population that is foreign born					2.6401	0.3050
2.42 or less	39.40	10.71	53.91	13.57		
2.43 – 7.46	37.44	14.18	27.78	12.71		
7.47 – 19.45	18.84	7.49	12.68	5.35		
19.46 or more	4.33	2.82	5.63	2.35		
Percentage of households that are linguistically isolated					0.5765	0.8849
0.14 or less	31.65	9.12	35.83	13.24		
0.15 – 1.87	34.49	14.95	30.69	12.31		
1.88 – 5.70	22.68	7.77	25.60	11.16		
5.71 or more	11.18	6.05	7.88	2.83		
Categorized average household size					3.2416	0.2368
2.33 or less	15.79	6.25	22.87	10.23		
2.34 – 2.56	29.24	10.05	15.24	8.26		
2.57 – 2.85	47.14	14.00	50.75	14.04		
2.86 or more	7.83	3.75	11.14	4.64		
Percentage of the population age 18–64 that is employed					0.0632	0.9517
64.53 or less	46.96	13.00	44.74	11.66		
64.54 – 71.66	32.99	10.60	33.08	10.27		
71.67 or more	20.05	6.23	22.18	9.69		
Percentage of the population age 25 and older with a high school education					0.6996	0.7867
20.35 or less	1.80	1.05	1.75	1.06		
20.36 – 28.56	14.43	6.03	9.93	3.86		
28.57 – 35.64	33.48	11.86	37.24	12.94		
35.65 or more	50.29	13.43	51.09	12.45		
Percentage of the population age 25 and older with some college education					1.2755	0.6752
22.40 or less	31.49	14.85	27.04	11.75		
22.41 – 27.72	25.48	10.03	27.70	12.30		
27.73 – 32.60	24.77	10.36	32.36	13.38		
32.61 or more	18.26	5.35	12.89	5.13		

See note at end of table.

Table A-5c. PIAAC results of Background Questionnaire bivariate analysis: 2014 List Sample—Continued

Analysis variable	Respondents		Nonrespondents		Chi-square	
	Percent	Standard error	Percent	Standard error	Statistic	*p* value
Race/ethnicity (after imputation)					2.0241	0.2871
Hispanic	6.58	3.38	12.74	7.17		
Non-Hispanic Black	34.11	13.00	20.89	8.99		
Other	59.31	10.38	66.38	10.70		
Age category (after imputation)					2.9674	0.2694
16–24	30.63	3.65	32.57	9.33		
25–34	30.75	2.17	16.39	4.27		
35–45	16.88	2.05	32.18	11.70		
46–55	15.44	2.59	15.97	7.45		
56 or older	6.31	1.90	2.90	1.80		
Language used at screening					0.4097	0.5221
English	97.54	1.89	98.36	1.50		
Spanish	2.46	1.89	1.64	1.50		
Indicator for children under age 16 in household					0.9894	0.3199
No	50.78	5.05	63.02	11.97		
Yes	49.22	5.05	36.98	11.97		
Gender					2.7152	0.0994
Male	36.99	3.10	55.06	9.52		
Female	63.01	3.10	44.94	9.52		

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Table A-6a. PIAAC multivariate analysis of Background Questionnaire response indicators, by response cell: 2012 Main Study

Response cell	Weighted response rate (percent)
No indicator for children under age 16 in household & Census Region West, South, and Midwest & Age categories (after imputation): 26–35, 36–45, 46–55, 56–65 & Race/ethnicity (after imputation): Non-Hispanic Black	74.2
Indicator for children under age 16 in household & Percentage of the population age 18–64 that is employed in 1st quartile	91.3
No indicator for children under age 16 in household & Census Region West, South, and Midwest & Age categories (after imputation): 26–35, 36–45, 46–55, 56–65 & Race/ethnicity (after imputation): Hispanic and Other & Percentage of households that are linguistically isolated in 4th quartile	84.5
No indicator for children under age 16 in household & Census Region Northeast & Percentage of households that are linguistically isolated in 1st quartile	83.2
No indicator for children under age 16 in household & Age categories (after imputation): 26–35, 36–45, 46–55, 56–65 & Race/ethnicity (after imputation): Hispanic and Other & Percentage of households that are linguistically isolated in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is unemployed in 1st and 2nd quartiles & Census Region West	74.0
Indicator for children under age 16 in household & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Male & Age category (after imputation): 16–25	88.5
No indicator for children under age 16 in household & Census Region Northeast & Percentage of households that are linguistically isolated in 2nd, 3rd, and 4th quartiles & Race/ethnicity (after imputation): Non-Hispanic Black	81.1
Indicator for children under age 16 in household & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Female & Categorized average household size in 1st and 2nd quartiles	88.4
No indicator for children under age 16 in household & Census Region Northeast & Percentage of households that are linguistically isolated in 2nd, 3rd, and 4th quartiles & Race/ethnicity (after imputation): Hispanic and Other & Percentage of the population that is Hispanic or Non-Hispanic Black in 3rd and 4th quartiles	63.7

See note at end of table.

Table A-6a. PIAAC multivariate analysis of Background Questionnaire response indicators, by response cell: 2012 Main Study—Continued

Response cell	Weighted response rate (percent)
No indicator for children under age 16 in household & Age categories (after imputation): 26–35, 36–45, 46–55, 56–65 & Race/ethnicity (after imputation): Hispanic and Other & Percentage of households that are linguistically isolated in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is unemployed in 1st and 2nd quartiles & Census Region South and Midwest & Percentage of the population age 18–64 that is employed in 1st, 2nd, and 3rd quartiles	79.5
No indicator for children under age 16 in household & Age categories (after imputation): 26–35, 36–45, 46–55, 56–65 & Race/ethnicity (after imputation): Hispanic and Other & Percentage of households that are linguistically isolated in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is unemployed in 1st and 2nd quartiles & Census Region South and Midwest & Percentage of the population age 18–64 that is employed in 4th quartile	86.5
No indicator for children under age 16 in household & Census Region West, South, and Midwest & Age categories (after imputation): 26–35, 36–45, 46–55, 56–65 & Percentage of households that are linguistically isolated in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is unemployed in 3rd and 4th quartiles & Race/ethnicity (after imputation): Hispanic	58.5
Indicator for children under age 16 in household & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Male & Age categories (after imputation): 26–35, 36–45, 46–55, 56–65 & Percentage of the population that is Hispanic or Non-Hispanic Black in 1st quartile	81.7
No indicator for children under age 16 in household & Census Region West, South, and Midwest & Age categories (after imputation): 26–35, 36–45, 46–55, 56–65 & Percentage of households that are linguistically isolated in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is unemployed in 3rd and 4th quartiles & Race/ethnicity (after imputation): Other & Percentage of the population that is Hispanic or Non-Hispanic Black in 4th quartile	68.4
No indicator for children under age 16 in household & Census Region West, South, and Midwest & Age categories (after imputation): 26–35, 36–45, 46–55, 56–65 & Percentage of households that are linguistically isolated in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is unemployed in 3rd and 4th quartiles & Race/ethnicity (after imputation): Other & Percentage of the population that is Hispanic or Non-Hispanic Black in 1st, 2nd, and 3rd quartiles & Percentage of housing units occupied by owner in 1st, 2nd, and 3rd quartiles	81.3

See note at end of table.

Table A-6a. PIAAC multivariate analysis of Background Questionnaire response indicators, by response cell: 2012 Main Study—Continued

Response cell	Weighted response rate (percent)
No indicator for children under age 16 in household & Census Region West, South, and Midwest & Age categories (after imputation): 26–35, 36–45, 46–55, 56–65 & Percentage of households that are linguistically isolated in 1st, 2nd, and 3rd quartiles & Percentage of the population age 18–64 that is unemployed in 3rd and 4th quartiles & Race/ethnicity (after imputation): Other & Percentage of the population that is Hispanic or Non-Hispanic Black in 1st, 2nd, and 3rd quartiles & Percentage of housing units occupied by owner in 4th quartile	73.3
No indicator for children under age 16 in household & Census Region West, South, and Midwest & Age category (after imputation): 16–25 & Percentage of the population that is foreign born in 1st quartile	93.1
Indicator for children under age 16 in household & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Female & Categorized average household size in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st quartile	91.2
Indicator for children under age 16 in household & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Female & Categorized average household size in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles	82.1
Indicator for children under age 16 in household & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Male & Age categories (after imputation): 26–35, 36–45, 46–55, 56–65 & Percentage of the population that is Hispanic or Non-Hispanic Black in 2nd, 3rd, and 4th quartiles & Percentage of the population age 25 and older with at least a high school education in 1st, 2nd, and 3rd quartiles	77.8
Indicator for children under age 16 in household & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Male & Age categories (after imputation): 26–35, 36–45, 46–55, 56–65 & Percentage of the population that is Hispanic or Non-Hispanic Black in 2nd, 3rd, and 4th quartiles & Percentage of the population age 25 and older with at least a high school education in 4th quartile	67.1

See note at end of table.

Table A-6a. PIAAC multivariate analysis of Background Questionnaire response indicators, by response cell: 2012 Main Study—Continued

Response cell	Weighted response rate (percent)
No indicator for children under age 16 in household & Census Region Northeast & Race/ethnicity (after imputation): Hispanic and Other & Percentage of the population that is Hispanic or Non-Hispanic Black in 1st and 2nd quartiles & Percentage of households that are linguistically isolated in 4th quartile	82.8
No indicator for children under age 16 in household & Census Region Northeast & Race/ethnicity (after imputation): Hispanic and Other & Percentage of the population that is Hispanic or Non-Hispanic Black in 1st and 2nd quartiles & Percentage of households that are linguistically isolated in 2nd and 3rd quartiles & Percentage of the population age 25 and older with at least a high school education in 1st, 2nd, and 3rd quartiles	67.5
No indicator for children under age 16 in household & Census Region Northeast & Race/ethnicity (after imputation): Hispanic and Other & Percentage of the population that is Hispanic or Non-Hispanic Black in 1st and 2nd quartiles & Percentage of households that are linguistically isolated in 2nd and 3rd quartiles & Percentage of the population age 25 and older with at least a high school education in 4th quartile	75.6
No indicator for children under age 16 in household & Age category (after imputation): 16–25 & Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Census Region West	77.5
No indicator for children under age 16 in household & Age category (after imputation): 16–25 & Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Census Region South and Midwest	84.1

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012.

Table A-6b. PIAAC multivariate analysis of Background Questionnaire indicators, by response cell: 2014 Area Sample

Response cell	Weighted response rate (percent)
No indicator for children under age 16 in household & Sampling domain: unemployed age 16–65	87.0
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 1st quartile & Age categories (after imputation): 16–24	86.2
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 1st quartile & Age categories (after imputation): 25–34, 35–45, 46–55, 56–65, older than 65 & Percentage of the population below 150 percent of poverty in 1st, 2nd, and 3rd quartiles	67.3
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 1st quartile & Age categories (after imputation): 25–34, 35–45, 46–55, 56–65, older than 65 & Percentage of the population below 150 percent of poverty in 4th quartile	78.7
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region Northeast & Percentage of the population that is foreign born in 1st quartile	80.4
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region Northeast & Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Age categories (after imputation): 16–24	67.5
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region Northeast & Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Age categories (after imputation): 25–34, 56–65, older than 65 & Percentage of the population age 18–64 that is unemployed in 1st quartile	37.7
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region Northeast & Percentage of the population that is foreign born in 2nd, 3rd, and 4th quartiles & Age categories (after imputation): 25–34, 56–65, older than 65 & Percentage of the population age 18–64 that is unemployed in 2nd, 3rd, and 4th quartiles	58.4

See note at end of table.

Table A-6b. PIAAC multivariate analysis of Background Questionnaire indicators, by response cell: 2014 Area Sample—Continued

Response cell	Weighted response rate (percent)
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region and Midwest, West, and South & Percent of housing units occupied by owner in 1st and 2nd quartiles & Age categories (after imputation): 16–24, 25–34, 35–45, 56–65 & Percentage of the population age 25 and older with some college education in 1st quartile	69.9
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region and Midwest, West, and South & Percent of housing units occupied by owner in 1st and 2nd quartiles & Age categories (after imputation): 16–24, 25–34, 35–45, 56–65 & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles & Percentage of the population below 150 percent of poverty in 1st, 2nd, and 3rd quartiles	84.4
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region and Midwest, West, and South & Percent of housing units occupied by owner in 1st and 2nd quartiles & Age categories (after imputation): 16–24, 25–34, 35–45, 56–65 & Percentage of the population age 25 and older with some college education in 2nd, 3rd, and 4th quartiles & Percentage of the population below 150 percent of poverty in 4th quartile	71.0
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region and Midwest, West, and South & Percent of housing units occupied by owner in 1st and 2nd quartiles & Age categories (after imputation): older than 65	65.5
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region and Midwest, West, and South & Percent of housing units occupied by owner in 3rd and 4th quartiles & Percentage of the population age 18–64 that is employed in 2nd and 3rd quartiles & Percentage of the population age 25 and older with a high school education in 1st and 2nd quartiles & Percentage of the population age 25 and older with at least a high school education in 1st and 2nd quartiles	69.0

See note at end of table.

Table A-6b. PIAAC multivariate analysis of Background Questionnaire indicators, by response cell: 2014 Area Sample—Continued

Response cell	Weighted response rate (percent)
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region and Midwest, West, and South & Percent of housing units occupied by owner in 3rd and 4th quartiles & Percentage of the population age 18–64 that is employed in 2nd and 3rd quartiles & Percentage of the population age 25 and older with a high school education in 1st and 2nd quartiles & Percentage of the population age 25 and older with at least a high school education in 3rd and 4th quartiles	51.9
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region and Midwest, West, and South & Percent of housing units occupied by owner in 3rd and 4th quartiles & Percentage of the population age 18–64 that is employed in 2nd and 3rd quartiles & Percentage of the population age 25 and older with a high school education in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 1st, 2nd, and 3rd quartiles	67.4
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region and Midwest, West, and South & Percent of housing units occupied by owner in 3rd and 4th quartiles & Percentage of the population age 18–64 that is employed in 2nd and 3rd quartiles & Percentage of the population age 25 and older with a high school education in 3rd and 4th quartiles & Percentage of the population age 25 and older with some college education in 4th quartile	77.2
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region and Midwest, West, and South & Percent of housing units occupied by owner in 3rd and 4th quartiles & Percentage of the population age 18–64 that is employed in 4th quartile & Age categories (after imputation): 16–24, 25–34, 35–45, 56–65	79.9

See note at end of table.

Table A-6b. PIAAC multivariate analysis of Background Questionnaire indicators, by response cell: 2014 Area Sample—Continued

Response cell	Weighted response rate (percent)
No indicator for children under age 16 in household & Sampling domain: not unemployed age 16–34, age 66–74 & Percentage of the population age 18–64 that is employed in 2nd, 3rd, and 4th quartiles & Census Region and Midwest, West, and South & Percent of housing units occupied by owner in 3rd and 4th quartiles & Percentage of the population age 18–64 that is employed in 4th quartile & Age categories (after imputation): older than 65	67.1
Indicator for children under age 16 in household & Percentage of the population age 25 and older with at least a high school education in 1st quartile	91.2
Indicator for children under age 16 in household & Percentage of the population age 25 and older with at least a high school education in 2nd, 3rd, and 4th quartiles & Age categories (after imputation): 16–24, 25–34, 35–45, 46–55 & Percentage of the population age 18–64 that is unemployed in 1st and 2nd quartiles & Percentage of the population that is Hispanic in 1st quartile	91.8
Indicator for children under age 16 in household & Percentage of the population age 25 and older with at least a high school education in 2nd, 3rd, and 4th quartiles & Age categories (after imputation): 16–24, 25–34, 35–45, 46–55 & Percentage of the population age 18–64 that is unemployed in 1st and 2nd quartiles & Percentage of the population that is Hispanic in 2nd, 3rd, and 4th quartiles & Percentage of the population age 25 and older with at least a high school education in 2nd and 3rd quartiles	86.1
Indicator for children under age 16 in household & Percentage of the population age 25 and older with at least a high school education in 2nd, 3rd, and 4th quartiles & Age categories (after imputation): 16–24, 25–34, 35–45, 46–55 & Percentage of the population age 18–64 that is unemployed in 1st and 2nd quartiles & Percentage of the population that is Hispanic in 2nd, 3rd, and 4th quartiles & Percentage of the population age 25 and older with at least a high school education in 4th quartile	72.3
Indicator for children under age 16 in household & Percentage of the population age 25 and older with at least a high school education in 2nd, 3rd, and 4th quartiles & Age categories (after imputation): 16–24, 25–34, 35–45, 46–55 & Percentage of the population age 18–64 that is unemployed in 3rd and 4th quartiles & Census Region and Midwest and West	84.5

See note at end of table.

Table A-6b. PIAAC multivariate analysis of Background Questionnaire indicators, by response cell: 2014 Area Sample—Continued

Response cell	Weighted response rate (percent)
Indicator for children under age 16 in household & Percentage of the population age 25 and older with at least a high school education in 2nd, 3rd, and 4th quartiles & Age categories (after imputation): 16–24, 25–34, 35–45, 46–55 & Percentage of the population age 18–64 that is unemployed in 3rd and 4th quartiles & Census Region Northeast and South & Sampling domain: unemployed age 16–65	86.6
Indicator for children under age 16 in household & Percentage of the population age 25 and older with at least a high school education in 2nd, 3rd, and 4th quartiles & Age categories (after imputation): 16–24, 25–34, 35–45, 46–55 & Percentage of the population age 18–64 that is unemployed in 3rd and 4th quartiles & Census Region Northeast and South & Sampling domain: not unemployed age 16–34	68.7
Indicator for children under age 16 in household & Percentage of the population age 25 and older with at least a high school education in 2nd, 3rd, and 4th quartiles & Age categories (after imputation): 56–65, older than 65	57.5

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Table A-6c. PIAAC multivariate analysis of Background Questionnaire response indicators, by response cell: 2014 List Sample

Response cell	Weighted response rate (percent)
Non-MSA & Percentage of households that are linguistically isolated in 1st quartile	87.1
Non-MSA & Percentage of households that are linguistically isolated in 2nd and 3rd quartiles	89.9
MSA & Male & Census Region West	98.5
MSA & Male & Census Region Northeast and South & Percentage of the population below 150 percent of poverty in 1st, 2nd, and 3rd quartiles	82.4
MSA & Male & Census Region Northeast and South & Percentage of the population below 150 percent of poverty in 4th quartile & No indicator for children under age 16 in household & Census Region Northeast	89.1
MSA & Male & Census Region Northeast and South & Percentage of the population below 150 percent of poverty in 4th quartile & No indicator for children under age 16 in household & Census Region South	87.4
MSA & Male & Census Region Northeast and South & Percentage of the population below 150 percent of poverty in 4th quartile & Indicator for children under age 16 in household	98.0
MSA & Male & Census Region Midwest	97.3
MSA & Female & Census Region West and Northeast & Percentage of the population age 25 and older with a high school education in 1st and 2nd quartiles & No indicator for children under age 16 in household	79.0

See note at end of table.

Table A-6c. PIAAC multivariate analysis of Background Questionnaire response indicators, by response cell: 2014 List Sample—Continued

Response cell	Weighted response rate (percent)
MSA & Female & Census Region West and Northeast & Percentage of the population age 25 and older with a high school education in 1st and 2nd quartiles & Indicator for children under age 16 in household	92.3
MSA & Female & Census Region West and Northeast & Percentage of the population age 25 and older with a high school education in 3rd and 4th quartiles	96.2
MSA & Female & Census Region Midwest and South & Percentage of households that are linguistically isolated in 1st and 2nd quartiles & No indicator for children under age 16 in household	98.7
MSA & Female & Census Region Midwest and South & Percentage of households that are linguistically isolated in 1st and 2nd quartiles & Indicator for children under age 16 in household	100.0
MSA & Female & Census Region Midwest and South & Percentage of households that are linguistically isolated in 3rd and 4th quartiles & Age categories (after imputation): 16–24, 25–34	93.9
MSA & Female & Census Region Midwest and South & Percentage of households that are linguistically isolated in 3rd and 4th quartiles & Age categories (after imputation): 35–45, 46–55, 56–65, older than 65	99.3

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Table A-7a. PIAAC multivariate analysis of Background Questionnaire response indicators: 2012 Main Study

Analysis variables	F statistic	Degrees of freedom		*p* value
		Numerator	Denominator	
Overall fit	2.622	23	23	0.012
Age category (without imputation)	4.035	4	42	0.007
Indicator for children under age 16 in household	13.421	1	45	0.001
Language used at screening	0.890	1	45	0.350
Percentage of the population below 150 percent of poverty	3.557	1	45	0.066
Percentage of the population age 18–64 that is employed	0.001	1	45	0.970
Percentage of the population that is foreign born	0.641	1	45	0.428
Percentage of households that are linguistically isolated	0.258	1	45	0.614
Percentage of housing units occupied by owner	0.415	1	45	0.523
Percentage of the population that is Hispanic	0.041	1	45	0.841
Percentage of the population that is Hispanic or Non-Hispanic Black	0.100	1	45	0.753
Percentage of the population age 25 and older with a high school education	0.380	1	45	0.541
Percentage of the population age 25 and older with some college education	0.022	1	45	0.883
Indicator of whether the PSU is part of a Metropolitan Statistical Area	0.282	1	45	0.598
Region	1.346	3	43	0.272
Categorized household size	0.002	1	45	0.962
Gender	6.473	1	45	0.014
Race/ethnicity (after imputation)	0.268	2	44	0.766

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012.

Table A-7b. PIAAC multivariate analysis of Background Questionnaire response indicators: 2014 Area Sample

Analysis variables	Degrees of freedom	Wald Chi-Square	*p* value
Sampling domain	2	51.7030	0.0000
Percentage of the population age 25 and older with at least a high school education	1	15.2660	0.0001
Indicator for children under age 16 in household	1	13.3558	0.0003
Census region	3	11.0129	0.0117
Percentage of households that are linguistically isolated	1	4.0777	0.0435
Percent of housing units occupied by owner	1	2.6431	0.1040

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Table A-7c. PIAAC multivariate analysis of Background Questionnaire response indicators: 2014 List Sample

Analysis variables	Degrees of freedom	Wald Chi-Square	*p* value
Indicator of whether the PSU is part of a Metropolitan Statistical Area	1	9.0026	0.0027
Gender	1	3.0835	0.0791
Census Region	3	10.2467	0.0166

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2014.

Appendix B

Extended Analysis Tables

This page intentionally blank.

Appendix B. Extended Analysis Tables

Table B-1. Estimated percentages, standard errors, and relative differences for selected weighting steps, by subgroup, for Main Study, National Supplement area sample and list sample separately

Subgroup	Base weights: eligible sample		Base weights: respondents			Nonresponse adjusted weights		
	Percent	Standard error	Percent	Standard error	*Relative difference*	Percent	Standard error	*Relative difference*
Total	100.0	†	100.0	†	†	100.0	†	†
Main Study:								
Percentage of the population age 18–64 that is unemployed[1]								
1st quartile	25.1	1.85	25.1	1.80	-0.027	25.2	1.78	0.054
2nd quartile	27.1	1.81	26.9	1.81	-0.072	26.9	1.82	-0.110
3rd quartile	25.6	1.41	25.8	1.40	0.156	26.2	1.42	0.426
4th quartile	22.2	1.77	22.1	1.77	-0.017	21.8	1.74	-0.226
Race/ethnicity								
Hispanic	13.3	1.83	13.0	1.63	-0.175	13.8	1.72	0.273
Non-Hispanic								
Black	12.2	1.48	12.1	1.39	-0.041	12.1	1.41	-0.068
Other	74.5	1.81	74.9	1.62	0.210	74.1	1.72	-0.221
Age (in years) [2]								
16–25	23.6	0.98	20.6	0.87	-3.092	22.8	0.92	-0.816
26–35	19.7	0.65	19.7	0.60	0.031	19.7	0.61	0.000
36–45	20.3	0.57	20.2	0.50	-0.175	20.3	0.57	0.000
46–55	21.3	0.61	21.9	0.56	0.951	21.7	0.60	0.656
56–65	15.1	0.67	17.6	0.72	3.776	15.5	0.67	0.597

See notes at end of table.

Table B-1. Estimated percentages, standard errors, and relative differences for selected weighting steps, by subgroup, for Main Study, National Supplement area sample and list sample separately—Continued

Subgroup	Base weights: eligible sample		Base weights: respondents			Nonresponse adjusted weights		
	Percent	Standard error	Percent	Standard error	*Relative difference*	Percent	Standard error	*Relative difference*
Gender								
Male	47.3	0.68	48.1	0.57	1.206	48.0	0.66	1.029
Female	52.8	0.68	51.9	0.57	-1.353	52.0	0.66	-1.176
National Supplement area sample:								
Percentage of the population age 18–64 that is unemployed[1]								
1st quartile	26.0	2.40	25.3	2.36	-0.279	25.7	2.39	-0.092
2nd quartile	26.2	1.99	25.8	2.09	-0.193	26.1	2.09	-0.012
3rd quartile	26.7	2.36	27.2	2.38	0.205	27.1	2.38	0.166
4th quartile	21.2	2.08	21.7	2.05	0.274	21.0	2.06	-0.070
Race/ethnicity								
Hispanic	14.9	1.74	16.1	1.91	0.711	15.2	1.80	0.219
Non-Hispanic								
Black	13.9	1.69	14.8	1.93	0.526	14.2	1.81	0.205
Other	71.3	1.61	69.1	1.70	-1.324	70.5	1.61	-0.454
Age (in years)[2]								
16–24	36.5	1.31	37.9	1.39	1.083	36.5	1.40	0.042
25–34	38.0	1.22	38.0	1.31	-0.005	37.1	1.29	-0.708
35–45	2.8	0.25	3.4	0.31	2.446	3.1	0.30	1.230
46–55	2.2	0.28	2.2	0.31	0.329	2.0	0.27	-0.562
56–65	1.6	0.18	1.8	0.23	1.247	1.7	0.25	0.917
66–74	19.0	0.69	16.7	0.68	-3.374	19.5	0.75	0.711

See notes at end of table.

Table B-1. Estimated percentages, standard errors, and relative differences for selected weighting steps, by subgroup, for Main Study, National Supplement area sample and list sample separately—Continued

Subgroup	Base weights: eligible sample		Base weights: respondents			Nonresponse adjusted weights		
	Percent	Standard error	Percent	Standard error	*Relative difference*	Percent	Standard error	*Relative difference*
Gender								
Male	49.4	0.86	49.7	0.90	0.338	50.1	0.97	0.740
Female	50.6	0.86	50.3	0.90	-0.338	49.9	0.97	-0.740
National Supplement list sample:								
Percentage of the population age 18–64 that is unemployed[1]								
1st and 2nd quartile	19.4	6.78	19.1	6.87	-0.039	19.0	6.65	-0.059
3rd quartile	36.5	13.24	36.6	13.62	0.003	36.0	13.14	-0.043
4th quartile	44.1	10.58	44.3	10.89	0.021	45.0	10.61	0.091
Race/ethnicity								
Hispanic	7.3	3.78	6.8	3.78	-0.125	6.9	3.64	-0.120
Non-Hispanic								
Black	32.1	14.09	33.1	14.54	0.066	32.3	13.97	0.016
Other	60.6	11.06	60.1	11.42	-0.042	60.8	11.04	0.021
Age (in years)[2]								
16–24	30.8	3.99	30.6	3.91	-0.033	31.0	3.99	0.062
25–34	30.0	2.18	31.0	2.22	0.452	30.9	2.17	0.414
35–45	17.8	2.24	16.7	2.09	-0.477	16.5	2.01	-0.577
46–55	15.3	2.45	15.2	2.81	-0.013	15.2	2.79	-0.030
56–65	6.2	1.78	6.4	1.88	0.136	6.4	1.80	0.120

See notes at end of table.

Table B-1. Estimated percentages, standard errors, and relative differences for selected weighting steps, by subgroup, for Main Study, National Supplement area sample and list sample separately—Continued

Subgroup	Base weights: eligible sample		Base weights: respondents			Nonresponse adjusted weights		
	Percent	Standard error	Percent	Standard error	*Relative difference*	Percent	Standard error	*Relative difference*
Gender								
Male	38.3	3.05	37.1	3.23	-0.406	37.6	3.23	-0.225
Female	61.7	3.05	62.9	3.23	0.406	62.4	3.23	0.225

† Not applicable.
[1] Quantiles among tract-level percentages. Cutpoint values are provided in table A-5.
[2] The Main Study and National Supplement list sample did not sample people age 66 and older.
NOTE: Relative differences are relative to the standard error of the eligible sample estimate.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC) 2012/2014.

Table B-2. Estimated percentages and standard errors from PIAAC and Current Population Survey (CPS), by subgroup

Subgroup	PIAAC		CPS	
	Percent	Standard error	Percent	Standard error[1]
Age (in years)				
16–18	5.5	0.17	5.8	0.06
19–25	13.7	0.17	13.3	0.08
25–30	9.2	0.22	9.3	0.07
31–35	8.9	0.22	9.0	0.07
36–40	9.1	0.27	8.5	0.07
41–45	8.7	0.27	9.3	0.07
46–50	10.0	0.32	9.9	0.07
51–55	9.4	0.32	9.8	0.07
56–60	8.8	0.32	8.7	0.07
61–65	7.8	0.32	7.6	0.06
66–74	8.9	0.00	8.9	0.07
Gender				
Male	48.9	0.00	49.0	0.12
Female	51.1	0.00	51.0	0.12
Race/Ethnicity				
Hispanic	15.6	0.00	15.7	0.09
Non-Hispanic Black Alone	12.0	0.00	11.9	0.08
Non-Hispanic Other	72.4	0.00	72.4	0.11
Education Attainment				
Less than high school	15.6	0.00	15.7	0.08
High school	26.5	0.00	28.4	0.10
Some college	31.4	0.00	28.3	0.10
College degree or higher	26.6	0.00	27.7	0.10

See notes at end of table.

Table B-2. Estimated percentages and standard errors from PIAAC and Current Population Survey (CPS), by subgroup—Continued

	PIAAC		CPS	
Subgroup	Percent	Standard error	Percent	Standard error[1]
Region				
Northeast	18.1	0.05	18.0	n/a[2]
Midwest	21.3	0.05	21.3	n/a
South	37.2	0.09	37.2	n/a
West	23.4	0.09	23.5	n/a

[1] Standard errors for the CPS estimates were derived from the generalized variance function formula using the parameters in Table 4 of Source and Accuracy of Estimates for Income, Poverty, and Health Insurance Coverage in the United States: 2012

[2] Region totals are fixed as part of the final weighting step in the ASEC survey of the CPS, so they do not have associated standard errors. For details see Technical Paper 63RV *The Current Population Survey: Design and Methodology*, Issued March 2002

SOURCE: National Center for Education Statistics' PIAAC 2012 and 2014 and the Bureau of Labor Statistics' Current Population Survey, Annual Social and Economic Supplement, 2012.

Table B-3. Correlations between literacy score and subgroup

Subgroup	Correlation
Census region	0.09
Metropolitan statistical region	0.07
Categorized household size	0.10
Percentage of the population below 150 percent of poverty[1]	0.31
Percentage of the population that is foreign born[1]	0.18
Percentage of the population with at least a high school education[1]	0.36
Percentage of the population that are linguistically isolated[1]	0.19
Percentage of housing units occupied by owner[1]	0.13
Percentage of the population age 18–64 that is unemployed[1]	0.15
Race/ethnicity	0.35
Age (in years)	0.20
Gender	0.02
Employment status (C_D05)	0.19
Education attainment (B_Q01aUS)	0.54
Nonresponse adjustment cells (NRA cells)	0.30
Calibration dimension (race/ethnicity by education attainment by age (16–65;66–74))	0.58
Calibration dimension (age by education attainment)	0.56
Calibration dimension (gender by education attainment by age (16–65;66–74))	0.52
Calibration dimension (age by race/ethnicity)	0.42
Calibration dimension (gender by race/ethnicity by age (16–65;66–74))	0.38
Calibration dimension (age by country of birth)	0.29
Calibration dimension (census region by country of birth by age (16–65;66–74))	0.29
Calibration dimension (race/ethnicity by education attainment by age (16–65;66–74)) + NRA cells	0.60
Calibration dimension (age by education attainment) + NRA cells	0.59
Calibration dimension (gender by education attainment by age (16–65;66–74)) + NRA cells	0.57
Calibration dimension (age by race/ethnicity) + NRA cells	0.46
Calibration dimension (gender by race/ethnicity by age (16–65;66–74)) + NRA cells	0.43
Calibration dimension (age by country of birth) + NRA cells	0.41
Calibration dimension (census region by country of birth by age (16–65;66–74)) + NRA cells	0.40
All calibration dimensions + NRA cells	0.65

[1] Quantiles among tract-level percentages.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC) 2012/2014.

Table B-4. Mean literacy score (1st plausible value) and standard errors from the final weights and recalibrated weights, by subgroup

Subgroup	Final weight		Re-weighted	
	Mean	Standard error	Mean	Standard error
Gender				
Male	271.7	0.95	271.7	0.94
Female	269.0	1.15	269.0	1.15
Age (in years)				
16–25	274.6	1.49	274.3	1.50
26–35	281.7	1.51	281.7	1.51
36–45	273.1	1.60	273.0	1.62
46–55	267.0	1.58	267.1	1.57
56–65	263.3	1.31	263.3	1.30
66–74	253.2	2.37	253.2	2.37
Education attainment				
1 Preprimary or no schooling	216.8	18.00	216.5	17.74
2 Grades 1-6	176.3	5.61	176.4	5.64
3 Grades 7-9	233.2	2.09	233.9	1.98
4 High school diploma	253.8	1.13	253.7	1.14
5 Preassociate education. Attended trade school, college, or university; no certificate or degree received	275.4	1.86	274.5	1.85
7 A certificate from a college or trade school for completion of a program prior to the associate/bachelor's degree	267.9	1.93	269.7	1.92
8 Associate degree	284.2	2.06	284.1	2.06
9 Bachelor's degree (e.g. BA, AB, BS)	300.3	1.29	300.3	1.30
10 Master's degree (e.g. MA, MS, MEng, MEd, MSW, MBA)	309.5	1.80	309.5	1.81
11 Professional degree (e.g. MD, DDS, DVM, LLB, JD)	315.1	4.85	315.1	4.83
12 Doctorate degree (e.g. PhD, EdD)	306.6	6.91	306.3	7.03
13 Foreign degree	252.0	15.43	251.5	15.33

See notes at end of table.

Table B-4. Mean literacy score (1st plausible value) and standard errors from the final weights and recalibrated weights, by subgroup—Continued

Subgroup	Final weight		Re-weighted	
	Mean	Standard error	Mean	Standard error
Employment status				
1 Currently working	276.8	1.11	276.8	1.11
2 Finding a job	260.4	1.66	260.4	1.66
3 Not finding	254.3	1.37	254.3	1.38
4 Else	204.6	79.70	202.5	77.56

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC) 2012/2014.

This page intentionally blank.

Appendix C

PIAAC 2012 and 2014
Weighted Item Response Rates

This page intentionally blank.

Appendix C. PIAAC 2012 and 2014 Weighted Item Response Rates

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
A_D01A1	General – Interview month (DERIVED BY CAPI)	8,488	100	100
A_D01A2	General – Year before interview (DERIVED BY CAPI)	8,488	100	100
A_D01A3	General – Interview year (DERIVED BY CAPI)	8,488	100	100
A_N01	General – Gender of respondent	8,488	100	100
A_Q01A	General – Year of birth	8,488	100	100
A_Q01B	General – Month of birth	8,488	100	100
BQLANG	Language for Background Questionnaire	8,488	100	100
B_D01D[1]	Education – Highest qualification – Months elapsed since finished (DERIVED BY CAPI)	8,459	10	10
B_D03D[1]	Education – Uncompleted qualification – Months elapsed since dropout (DERIVED BY CAPI)	6,679	2	2
B_D12H	Activities – Last year – Number of learning activities (DERIVED BY CAPI)	7,986	100	100
B_Q01A	Education – Highest qualification – Level	8,488	100	100
B_Q01A2US	Education – Highest qualification – Country of foreign qualification	24	63	63
B_Q01A3	Education – Highest qualification – Level of foreign qualification	24	59	63
B_Q01A3US	Education – Highest qualification – Level of foreign qualification	24	59	63
B_Q01AUS	Education – Highest qualification – Level	8,488	100	100
B_Q01B[2]	Education – Highest qualification – Area of study	7,073	69	69
B_Q01BUSX	Education – Highest qualification – Area of study verbatim	4,462	100	100
B_Q01C1	Education – Highest qualification – Age of finish	8,459	99	100
B_Q01C2	Education – Highest qualification – Year of finish	8,459	99	100
B_Q01D	Education – Highest qualification – Month of finish	1,055	98	99
B_Q02A	Education – Current qualification	8,488	100	100

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
B_Q02B	Education – Current qualification – Level	1,816	99	99
B_Q02BUS	Education – Current qualification – Level	1,816	99	99
B_Q02C[2]	Education – Current qualification – Area of study	1,805	72	72
B_Q02CUSX	Education – Current Qualification – Area of Study Verbatim	1,243	100	100
B_Q03A	Education – Uncompleted qualification	6,679	100	100
B_Q03B	Education – Uncompleted qualification – Level	2,082	100	99
B_Q03BUS	Education – Uncompleted qualification – Level	2,111	100	100
B_Q03C1	Education – Uncompleted qualification – Age of dropout	2,082	99	99
B_Q03C2	Education – Uncompleted qualification – Year of dropout	2,082	100	99
B_Q03D	Education – Uncompleted qualification – Month of dropout	184	92	94
B_Q04A	Education – Formal qualification	6,327	100	100
B_Q04B	Education – Formal qualification – How many qualifications	258	96	96
B_Q05A	Education – Formal qualification – Level	258	92	96
B_Q05AUS	Education – Formal qualification – Level	635	97	100
B_Q05B	Education – Formal qualification – Area of study	321	91	92
B_Q05BUSX	Education – Formal Qualification – Area of Study Verbatim	302	100	99
B_Q05C	Education – Formal qualification – Reason job related	1,917	98	98
B_Q05CUSX1	Education – Formal qualification – Degree personal interest	1,554	100	100
B_Q05CUSX2	Education – Formal qualification – Degree personal interest or work related	743	100	100
B_Q10A	Education – Formal qualification – Employed	1,917	98	98
B_Q10B	Education – Formal qualification – Employed – Working hours	1,319	99	99
B_Q10C	Education – Formal qualification – Employed – Useful for job	1,319	99	99
B_Q11	Education – Formal qualification – Grant from employer	1,917	98	98
B_Q12A	Activities – Last year – Open or distance edu	7,986	100	100
B_Q12B	Activities – Last year – Open or distance edu – How many	1,252	99	99
B_Q12C	Activities – Last year – On the job training	7,986	100	100
B_Q12D	Activities – Last year – On the job training – How many	2,727	100	100

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
B_Q12E	Activities – Last year – Seminars or workshops	7,986	100	100
B_Q12F	Activities – Last year – Seminars or workshops – How many	2,101	100	100
B_Q12G	Activities – Last year – Private lessons	7,986	100	100
B_Q12H	Activities – Last year – Private lessons – How many	646	99	99
B_Q13	Activities – Last year – Activity specified	3,237	100	100
B_Q14A	Activities – Last year – Job related	2,158	100	100
B_Q14B[2]	Activities – Last year – Reason for participating	3,286	58	58
B_Q14BUSX1	Activities – Last year – Activity Participation for personal interest	1,415	100	100
B_Q14BUSX2	Activities – Last year – Activity Participation mainly personal reasons	895	100	100
B_Q15A	Activities – Last year – Employed	2,158	100	100
B_Q15B	Activities – Last year – During working hours	3,441	100	100
B_Q15C	Activities – Last year – Useful for job	3,441	100	100
B_Q16	Activities – Last year – Grant from employer	4,073	100	100
B_Q17	Activities – Last year – Time spend – Unit	4,073	100	100
B_Q18A	Activities – Last year – Time spend for activities – Weeks	1,069	99	99
B_Q19A	Activities – Last year – Time spend for activities – Days	1,100	99	99
B_Q20A	Activities – Last year – Time spend for activities – Hours	1,917	99	100
B_Q20B	Activities – Last year – Time spend for activities – Proportion of job-related hours	3,227	100	100
B_Q26A	Activities – Last year – Wanted but didn't start	7,986	100	100
B_Q26B	Activities – Last year – Wanted but didn't start – Reason	2,831	100	100
B_Q27AUSX	Activities – Class – Class/tutor basic skills	2,590	100	100
B_Q27BUSX	Activities – Class – Class/tutor GED	2,590	100	100
B_Q27CUSX	Activities – Class – Class/tutor other equivalency	2,590	100	100
B_Q27DUSX	Activities – Class – Class/tutor main reason	375	98	99
B_Q27EUSXA	Activities – Class – Class attendance, amount	375	96	99
B_Q27EUSXB	Activities – Class – Class attendance, unit	375	97	99
B_Q29AUSX	Activities – Apprentice – Was apprentice	8,488	100	100

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
B_Q29BUSX	Activities – Apprentice – Which trade or craft	144	100	100
B_S01A1	Education – Highest qualification – Name of foreign qualification	24	63	63
B_S01A2	Education – Highest qualification – Country of foreign qualification (other)	24	39	39
B_S27EUSX	Activities – Class – Class attendance, other specify	7	75	100
C_D04	Current status/work history – Last month – Active steps to find job (DERIVED BY CAPI)	1,551	98	98
C_D05	Current status/work history – Employment status (DERIVED BY CAPI)	8,488	100	100
C_D06	Current status/work history – Current – Paid job or family business (DERIVED BY CAPI)	8,488	100	100
C_D08C	Current status/work history – Left work in past 5 years (DERIVED BY CAPI)	1,845	99	99
C_D09	Current status/work history – Work experience (DERIVED BY CAPI)	8,488	100	100
C_Q01A	Current status/work history – Last week – Paid work	8,488	100	100
C_Q01B	Current status/work history – Last week – Away from job but will return	3,917	100	100
C_Q01C	Current status/work history – Last week – Unpaid work for own business	3,665	100	100
C_Q02A	Current status/work history – Last month – Looking for paid work	3,496	100	100
C_Q02B	Current status/work history – Last month – Waiting to start job	1,952	100	100
C_Q02C	Current status/work history – Last month – Waiting to start job – Next 3 months	22	51	51
C_Q03_01	Current status/work history – Last month – Reason not looking for work – Waiting for result of application	1,937	99	99
C_Q03_02	Current status/work history – Last month – Reason not looking for work – Being a student	1,937	99	99
C_Q03_03	Current status/work history – Last month – Reason not looking for work – Looking after the family	1,937	99	99
C_Q03_04	Current status/work history – Last month – Reason not looking for work – Temp sick	1,937	99	99
C_Q03_05	Current status/work history – Last month – Reason not looking for work – Long–term illness	1,937	99	99

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
C_Q03_06	Current status/work history – Last month – Reason not looking for work – Nothing available	1,937	99	99
C_Q03_07	Current status/work history – Last month – Reason not looking for work – Did not get around	1,937	99	99
C_Q03_08	Current status/work history – Last month – Reason not looking for work – No need	1,937	99	99
C_Q03_09	Current status/work history – Last month – Reason not looking for work – Retired	1,937	99	99
C_Q03_10	Current status/work history – Last month – Reason not looking for work – Other	1,937	99	99
C_Q04A	Current status/work history – Last month – Ways of looking for work – Contact public employment	1,551	98	98
C_Q04B	Current status/work history – Last month – Ways of looking for work – Contact private agency	1,551	98	98
C_Q04C	Current status/work history – Last month – Ways of looking for work – Apply to employers	1,551	98	98
C_Q04D	Current status/work history – Last month – Ways of looking for work – Ask family/friends	1,551	98	98
C_Q04E	Current status/work history – Last month – Ways of looking for work – Place/answer advertisements	1,551	98	98
C_Q04F	Current status/work history – Last month – Ways of looking for work – Study advertisements	1,551	98	98
C_Q04G	Current status/work history – Last month – Ways of looking for work – Recruitment test	1,551	98	98
C_Q04H	Current status/work history – Last month – Ways of looking for work – Premises	1,551	98	98
C_Q04I	Current status/work history – Last month – Ways of looking for work – Licenses/financial resources	1,551	98	98
C_Q04J	Current status/work history – Last month – Ways of looking for work – Other	1,551	98	98
C_Q05	Current status/work history – Ability to start job within 2 weeks	1,563	98	98
C_Q06	Current status/work history – Last week – Number of jobs	4,830	100	100

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
C_Q07	Current status/work history – Subjective status	8,488	100	100
C_Q08A	Current status/work history – Ever paid work	3,665	100	100
C_Q08B	Current status/work history – Last year – Paid work	3,130	100	100
C_Q08C1	Current status/work history – Age when stopped working in last job	1,845	98	99
C_Q08C2	Current status/work history – Year when stopped working in last job	1,845	99	99
C_Q09	Current status/work history – Years of paid work during lifetime	7,953	100	100
C_Q10A	Current status/work history – Last 5 years – How many diff firms or organisations	7,183	100	100
C_S03	Current status/work history – Last month – Months looking for paid work	1,566	98	98
C_S04J	Current status/work history – Last month – Ways of looking for work – Other specified	274	92	92
D_D16A	Current work – Earnings – Salary interval per hour (DERIVED BY CAPI)	4,032	100	100
D_Q01A	Current work – Job title	5,000	100	100
D_Q01B	Current work – Responsibilities	5,000	100	100
D_Q02A	Current work – Kind of business, industry or service	5,000	100	100
D_Q02B	Current work – Main product of firm or organisation	5,000	100	100
D_Q03	Current work – Economic sector	5,000	99	100
D_Q03US	Current work – Economic sector	5,000	99	100
D_Q04	Current work – Employee or self–employed	4,831	100	100
D_Q05A1	Current work – Start of work for employer – Age	4,300	99	100
D_Q05A2	Current work – Start of work for employer – Year	4,300	99	100
D_Q05A3	Current work – Start of work for employer – Month	1,238	99	99
D_Q05B1	Current work – Start of work for business – Age	708	98	99
D_Q05B2	Current work – Start of work for business – Year	708	98	99
D_Q05B3	Current work – Start of work for business – Month	128	91	91
D_Q06A	Current work – Amount of people working for employer	4,300	99	100
D_Q06B	Current work – Amount of people working for employer increased	4,300	99	100
D_Q06C	Current work – Part of a larger organization	4,300	99	100

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
D_Q07A	Current work – Employees working for you	708	99	99
D_Q07B	Current work – Employees working for you – Amount	215	96	96
D_Q08A	Current work – Managing other employees	4,300	100	100
D_Q08B	Current work – Managing other employees – Amount	1,288	99	99
D_Q09	Current work – Type of contract	4,131	100	100
D_Q10	Current work – Hours/week	5,000	99	100
D_Q11A	Current work – Work flexibility – Sequence of tasks	5,000	100	100
D_Q11B	Current work – Work flexibility – How to do the work	5,000	100	100
D_Q11C	Current work – Work flexibility – Speed of work	5,000	100	100
D_Q11D	Current work – Work flexibility – Working hours	5,000	100	100
D_Q12A	Current work – Requirements – Education level	4,300	99	100
D_Q12AUS	Current work – Requirements – Education level	4,300	99	100
D_Q12B	Current work – Requirements – To do the job satisfactorily	4,112	100	100
D_Q12C	Current work – Requirements – Related work experience	4,300	99	100
D_Q13A	Current work – Learning – Learning from coworkers/supervisors	4,507	100	100
D_Q13B	Current work – Learning – Learning-by-doing	5,000	100	100
D_Q13C	Current work – Learning – Keeping up to date	5,000	100	100
D_Q14	Current work – Job satisfaction	5,000	100	100
D_Q16A	Current work – Earnings – Salary interval	4,131	98	98
D_Q16B	Current work – Earnings – Gross pay	4,032	94	95
D_Q16C	Current work – Earnings – Gross pay in broad categories	184	85	85
D_Q16D1	Current work – Earnings – Broad categories – Gross pay per hour	32	63	67
D_Q16D2	Current work – Earnings – Broad categories – Gross pay per day	10	9	9
D_Q16D3	Current work – Earnings – Broad categories – Gross pay per week	19	37	43
D_Q16D4	Current work – Earnings – Broad categories – Gross pay per 2 weeks	19	47	50
D_Q16D5	Current work – Earnings – Broad categories – Gross pay per month	15	52	52
D_Q16D6	Current work – Earnings – Broad categories – Gross pay per year	67	81	82

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
D_Q17A	Current work – Earnings – Additional payments	4,032	99	99
D_Q17B	Current work – Earnings – Additional payments amount last year	1,657	88	95
D_Q17C	Current work – Earnings – Additional payments in broad categories	68	82	82
D_Q17D	Current work – Earnings – Additional payments – Broad – Last year	157	70	93
D_Q18A	Current work – Earnings – Total earnings last year	708	84	90
D_Q18B	Current work – Earnings – Total earnings broad categories	62	80	80
D_Q18C1	Current work – Earnings – Broad categories – Total earnings last month	9	3	3
D_Q18C2	Current work – Earnings – Broad categories – Total earnings last year	78	79	83
D_S09	Current work – Other type of contract specified	90	87	91
D_S16A	Current work – Earnings – Hours per piece	23	58	59
EDLEVEL3	Educational level of the respondent (DERIVED BY CAPI)	8,488	100	100
E_Q01A	Last job – Job title	2,191	99	99
E_Q01B	Last job – Responsibilities	2,191	99	99
E_Q02A	Last job – Kind of business, industry or service	2,191	99	99
E_Q02B	Last job – Main product of firm or organisation	2,191	99	99
E_Q03	Last job – Economic sector	2,191	98	99
E_Q03US	Last job – Economic sector	2,191	98	99
E_Q04	Last job – Employee or self–employed	2,191	99	99
E_Q05A1	Last job – Start of work for employer – Age	1,987	99	99
E_Q05A2	Last job – Start of work for employer – Year	1,987	99	99
E_Q05B1	Last job – Start of work for business – Age	212	92	92
E_Q05B2	Last job – Start of work for business – Year	212	93	93
E_Q06	Last job – Amount of people working for employer	1,987	99	99
E_Q07A	Last job – Employees working for you	212	93	93
E_Q07B	Last job – Employees working for you – Amount	42	72	72
E_Q08	Last job – Type of contract	1,987	99	99
E_Q09	Last job – Hours/week	2,191	99	99

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
E_Q10	Last job – Reason for end of job	1,987	99	99
E_S08	Last job – Other type of contract specified	48	73	75
F_Q01B	Skill use work – Time cooperating with coworkers	5,599	100	100
F_Q02A	Skill use work – How often – Sharing work–related info	6,202	100	100
F_Q02B	Skill use work – How often – Teaching people	6,202	100	100
F_Q02C	Skill use work – How often – Presentations	6,202	100	100
F_Q02D	Skill use work – How often – Selling	6,202	100	100
F_Q02E	Skill use work – How often – Advising people	6,202	100	100
F_Q03A	Skill use work – How often – Planning own activities	6,202	100	100
F_Q03B	Skill use work – How often – Planning others activities	6,202	100	100
F_Q03C	Skill use work – How often – Organising own time	6,202	100	100
F_Q04A	Skill use work – How often – Influencing people	6,202	100	100
F_Q04B	Skill use work – How often – Negotiating with people	6,202	100	100
F_Q05A	Skill use work – Problem solving – Simple problems	6,202	100	100
F_Q05B	Skill use work – Problem solving – Complex problems	6,202	100	100
F_Q06B	Skill use work – How often – Working physically for long	6,202	100	100
F_Q06C	Skill use work – How often – Using hands or fingers	6,202	100	100
F_Q07A	Skill use work – Not challenged enough	5,000	100	100
F_Q07B	Skill use work – Need more training	5,000	100	100
G_Q01A	Skill use work – Literacy – Read directions or instructions	6,202	100	100
G_Q01B	Skill use work – Literacy – Read letters memos or mails	6,202	100	100
G_Q01C	Skill use work – Literacy – Read newspapers or magazines	6,202	100	100
G_Q01D	Skill use work – Literacy – Read professional journals or publications	6,202	100	100
G_Q01E	Skill use work – Literacy – Read books	6,202	100	100
G_Q01F	Skill use work – Literacy – Read manuals or reference materials	6,202	100	100
G_Q01G	Skill use work – Literacy – Read financial statements	6,202	100	100
G_Q01H	Skill use work – Literacy – Read diagrams maps or schematics	6,202	100	100

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
G_Q02A	Skill use work – Literacy – Write letters memos or mails	6,202	100	100
G_Q02B	Skill use work – Literacy – Write articles	6,202	100	100
G_Q02C	Skill use work – Literacy – Write reports	6,202	100	100
G_Q02D	Skill use work – Literacy – Fill in forms	6,202	100	100
G_Q03B	Skill use work – Numeracy – How often – Calculating costs or budgets	6,202	100	100
G_Q03C	Skill use work – Numeracy – How often – Use or calculate fractions or percentages	6,202	100	100
G_Q03D	Skill use work – Numeracy – How often – Use a calculator	6,202	100	100
G_Q03F	Skill use work – Numeracy – How often – Prepare charts graphs or tables	6,202	100	100
G_Q03G	Skill use work – Numeracy – How often – Use simple algebra or formulas	6,202	100	100
G_Q03H	Skill use work – Numeracy – How often – Use advanced math or statistics	6,202	100	100
G_Q04	Skill use work – ICT – Experience with computer in job	6,202	100	100
G_Q05A	Skill use work – ICT – Internet – How often – For mail	4,403	100	100
G_Q05C	Skill use work – ICT – Internet – How often – Work related info	4,403	100	100
G_Q05D	Skill use work – ICT – Internet – How often – Conduct transactions	4,403	100	100
G_Q05E	Skill use work – ICT – Computer – How often – Spreadsheets	4,403	100	100
G_Q05F	Skill use work – ICT – Computer – How often – Word	4,403	100	100
G_Q05G	Skill use work – ICT – Computer – How often – Programming language	4,403	100	100
G_Q05H	Skill use work – ICT – Computer – How often – Real-time discussions	4,403	100	100
G_Q06	Skill use work – ICT – Computer – Level of computer use	4,403	99	100
G_Q07	Skill use work – ICT – Computer – Got the skills needed	4,403	100	100
G_Q08	Skill use work – ICT – Computer – Lack of skills affect career	4,403	100	100
H_Q01A	Skill use everyday life – Literacy – Read directions or instructions	8,488	100	100
H_Q01B	Skill use everyday life – Literacy – Read letters memos or mails	8,488	100	100
H_Q01C	Skill use everyday life – Literacy – Read newspapers or magazines	8,488	100	100
H_Q01D	Skill use everyday life – Literacy – Read professional journals or publications	8,488	100	100
H_Q01E	Skill use everyday life – Literacy – Read books	8,488	100	100
H_Q01F	Skill use everyday life – Literacy – Read manuals or reference materials	8,488	100	100

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
H_Q01G	Skill use everyday life – Literacy – Read financial statements	8,488	100	100
H_Q01H	Skill use everyday life – Literacy – Read diagrams maps or schematics	8,488	100	100
H_Q02A	Skill use everyday life – Literacy – Write letters memos or mails	8,488	100	100
H_Q02B	Skill use everyday life – Literacy – Write articles	8,488	100	100
H_Q02C	Skill use everyday life – Literacy – Write reports	8,488	100	100
H_Q02D	Skill use everyday life – Literacy – Fill in forms	8,488	100	100
H_Q03B	Skill use everyday life – Numeracy – How often – Calculating costs or budgets	8,488	100	100
H_Q03C	Skill use everyday life – Numeracy – How often – Use or calculate fractions or percentages	8,488	100	100
H_Q03D	Skill use everyday life – Numeracy – How often – Use a calculator	8,488	100	100
H_Q03F	Skill use everyday life – Numeracy – How often – Prepare charts graphs or tables	8,488	100	100
H_Q03G	Skill use everyday life – Numeracy – How often – Use simple algebra or formulas	8,488	100	100
H_Q03H	Skill use everyday life – Numeracy – How often – Use advanced math or statistics	8,488	100	100
H_Q04A	Skill use everyday life – ICT – Ever used computer	4,093	100	100
H_Q04B	Skill use everyday life – ICT – Experience with computer everyday life	8,038	100	100
H_Q05A	Skill use everyday life – ICT – Internet – How often – For mail	6,794	100	100
H_Q05C	Skill use everyday life – ICT – Internet – How often – In order to better understand various issues	6,794	100	100
H_Q05D	Skill use everyday life – ICT – Internet – How often – Conduct transactions	6,794	100	100
H_Q05E	Skill use everyday life – ICT – Computer – How often – Spreadsheets	6,794	100	100
H_Q05F	Skill use everyday life – ICT – Computer – How often – Word	6,794	100	100
H_Q05G	Skill use everyday life – ICT – Computer – How often – Programming language	6,794	100	100
H_Q05H	Skill use everyday life – ICT – Computer – How often – Real–time discussions	6,794	100	100
I_Q04B	About yourself – Learning strategies – Relate new ideas into real life	8,488	100	100
I_Q04D	About yourself – Learning strategies – Like learning new things	8,488	100	100
I_Q04H	About yourself – Learning strategies – Attribute something new	8,488	100	100
I_Q04J	About yourself – Learning strategies – Get to the bottom of difficult things	8,488	100	100

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
I_Q04L	About yourself – Learning strategies – Figure out how different ideas fit together	8,488	100	100
I_Q04M	About yourself – Learning strategies – Looking for additional info	8,488	100	100
I_Q05F	About yourself – Cultural engagement – Voluntary work for non–profit organisations	8,488	100	100
I_Q06A	About yourself – Political efficacy – No influence on the government	8,488	100	100
I_Q06DUSX1A	About yourself – Political efficacy – Information from newspapers	8,488	100	100
I_Q06DUSX1B	About yourself – Political efficacy – Information from magazines	8,488	100	100
I_Q06DUSX1C	About yourself – Political efficacy – Information from internet	8,488	100	100
I_Q06DUSX1D	About yourself – Political efficacy – Information from radio	8,488	100	100
I_Q06DUSX1E	About yourself – Political efficacy – Information from television	8,488	100	100
I_Q06DUSX1F	About yourself – Political efficacy – Information from books or brochures	8,488	100	100
I_Q06DUSX1G	About yourself – Political efficacy – Information from family members, friends, or coworkers	8,488	100	100
I_Q07A	About yourself – Social trust – Trust only few people	8,488	100	100
I_Q07B	About yourself – Social trust – Other people take advantage of you	8,488	100	100
I_Q08	About yourself – Health – State	8,488	100	100
I_Q08USX1	About yourself – Health – Difficulty seeing print	8,488	100	100
I_Q08USX2	About yourself – Health – Difficulty hearing conversation	8,488	100	100
I_Q08USX3	About yourself – Health – Diagnosed learning disabled	8,488	100	100
I_Q10BUSX1	About yourself – Health – Have medical insurance	8,488	100	100
I_Q10BUSX2A	About yourself – Health – Health information from newspapers	8,488	100	100
I_Q10BUSX2B	About yourself – Health – Health information from magazines	8,488	100	100
I_Q10BUSX2C	About yourself – Health – Health information from internet	8,488	100	100
I_Q10BUSX2D	About yourself – Health – Health information from radio	8,488	100	100
I_Q10BUSX2E	About yourself – Health – Health information from television	8,488	100	100
I_Q10BUSX2F	About yourself – Health – Health information from books or brochures	8,488	100	100

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
I_Q10BUSX2G	About yourself – Health – Health information from family members, friends, or coworkers	8,488	100	100
I_Q10BUSX2H	About yourself – Health – Health information from health professional	8,488	100	100
I_Q10BUSX3A	About yourself – Health – Flu shot in past year	8,488	100	100
I_Q10BUSX3B	About yourself – Health – Mammogram in past year	2,087	100	100
I_Q10BUSX3C	About yourself – Health – Pap smear in past year	4,332	100	100
I_Q10BUSX3D	About yourself – Health – Screen for colon cancer in past year	2,570	100	100
I_Q10BUSX3E	About yourself – Health – Vision check in past year	8,488	100	100
I_Q10BUSX3F	About yourself – Health – Screen for prostate cancer in past year	1,174	100	100
I_Q10BUSX3G	About yourself – Health – Screen for osteoporosis in past year	2,570	99	100
I_Q10BUSX3H	About yourself – Health – Seen dentist in past year	8,488	100	100
J_N05A2	Background – More than one language mentioned	8,488	100	100
J_Q01	Background – People in household	8,488	100	100
J_Q02A	Background – Living with spouse or partner	7,115	100	100
J_Q02C	Background – Work situation of spouse or partner	4,077	100	100
J_Q03A	Background – Children	8,488	100	100
J_Q03B	Background – Number of children	5,127	100	100
J_Q03C	Background – Age of the child	1,285	99	99
J_Q03D1	Background – Age of the youngest child	3,850	100	100
J_Q03D2	Background – Age of the oldest child	3,850	99	100
J_Q04A	Background – Born in country	8,488	100	100
J_Q04BUS	Background – Country of birth	1,079	100	100
J_Q04C1	Background – Age of immigration	1,087	99	99
J_Q04C2	Background – Year of immigration	1,087	99	99
J_Q04DUSX1A	Background – Hispanic	8,488	100	100
J_Q04DUSX1B_01	Background – Hispanic origin – Mexican	1,101	100	100
J_Q04DUSX1B_02	Background – Hispanic origin – Puerto Rican	1,101	100	100

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
J_Q04DUSX1B_03	Background – Hispanic origin – Cuban	1,101	100	100
J_Q04DUSX1B_04	Background – Hispanic origin – Central/South America	1,101	100	100
J_Q04DUSX1B_05	Background – Hispanic origin – Other	1,101	100	100
J_Q04DUSX2_01	Background – Race – White	8,488	99	99
J_Q04DUSX2_02	Background – Race – Black	8,488	99	99
J_Q04DUSX2_03	Background – Race – Asian	8,488	99	99
J_Q04DUSX2_04	Background – Race – American Indian	8,488	99	99
J_Q04DUSX2_05	Background – Race – Native Hawaiian	8,488	99	99
J_Q05A1US	Background – First learned language	8,488	100	100
J_Q05A2US	Background – Second learned language	369	100	100
J_Q05A2USX2	Background – Age learned English	1,401	99	99
J_Q05BUS	Background – Language spoken at home	1,401	99	99
J_Q05CUSX1	Background – Language spoken most	1,401	99	99
J_Q05CUSX2	Background – English outside home	1,401	99	99
J_Q05CUSX3A	Background – Ability to understand spoken English	8,488	100	100
J_Q05CUSX3B	Background – Ability to speak English	8,488	100	100
J_Q05CUSX3D	Background – Ability to read English	8,488	100	100
J_Q05CUSX3E	Background – Ability to write English	8,488	100	100
J_Q05CUSX4	Background – ESL class/tutor in past year	1,290	99	99
J_Q05CUSX5	Background – Reason for ESL class/tutor	77	100	100
J_Q05CUSX6	Background – Class/tutor learn English as adult	1,290	99	99
J_Q06A	Background – Mother/female guardian – Whether born in country	8,488	100	100
J_Q06B	Background – Mother/female guardian – Highest level of education	8,488	98	100
J_Q06BUS	Background – Mother/female guardian – Highest level of education	8,488	98	100
J_Q07A	Background – Father/male guardian – Whether born in country	8,488	99	100
J_Q07B	Background – Father/male guardian – Highest level of education	8,488	95	100
J_Q07BUS	Background – Father/male guardian – Highest level of education	8,488	95	100

See notes at end of table.

Table C-1. PIAAC 2012 and 2014 household survey weighted item response rates (Items with response rates below 85 percent asked of at least 100 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
J_Q08	Background – Number of books at home	8,488	100	100
J_S04B[2]	Background – Country of birth (other)	1,087	53	53
J_S05A1[2]	Background – First learned language (other)	8,488	5	5
J_S05A2[2]	Background – Second learned language (other)	377	14	14
J_S05B[2]	Background – Language spoken at home (other)	8,488	3	3

[1] Response rate is below 85 percent for this CAPI derived variable because the derivation did not distinguish between a valid skip and nonresponse. As a result, valid skips were counted as nonresponse.
[2] Response rate is below 85 percent since valid skips were not coded as such in the US–specific variable used for routing to this item.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC) 2012/2014.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
A_D01A1	General – Interview month (DERIVED BY CAPI)	1,315	100	100
A_D01A2	General – Year before interview (DERIVED BY CAPI)	1,315	100	100
A_D01A3	General – Interview year (DERIVED BY CAPI)	1,315	100	100
A_N01	General – Gender of respondent	1,315	100	100
A_Q01A	General – Year of birth	1,315	100	100
A_Q01B	General – Month of birth	1,315	100	100
B_D01D[1]	Education – Highest qualification – Months elapsed since finished (DERIVED BY CAPI)	1,306	5	5
B_Q01A	Education – Highest qualification – Level	1,315	100	100
B_Q01AUS	Education – Highest qualification – Level	1,315	100	100
B_Q01B	Education – Highest qualification – Area of study	303	100	100
B_Q01BUSX	Education – Highest qualification – Area of study verbatim	303	100	100
B_Q01C1US	Education – Highest qualification – Age of finish	1,306	99	100
B_Q01C2	Education – Highest qualification – Year of finish	1,306	99	100
B_Q01D	Education – Highest qualification – Month of finish	77	96	100
B_Q02A	Education – Current qualification	1,315	100	100
B_Q02B	Education – Current qualification – Level	293	100	100
B_Q02BUS	Education – Current qualification – Level	293	100	100
B_Q02C	Education – Current qualification – Area of study	134	100	100
B_Q02CUSX	Education – Current Qualification – Area of Study Verbatim	134	100	100
B_Q03A	Education – Uncompleted qualification	1,022	100	100
B_Q03B	Education – Uncompleted qualification – Level	331	100	100
B_Q03BUS	Education – Uncompleted qualification – Level	331	100	100
B_Q27aUSP	Activities – Class – Class/tutor basic skills during current incarceration	1,114	100	100
B_Q27bUSP	Activities – Class – Class/tutor GED during current incarceration	1,114	100	100
B_Q27cUSP	Activities – Class – Class/tutor other equivalency during current incarceration	745	100	100

See notes at end of table.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
B_Q27eUSPa	Activities – Class – Class attendance, amount during current incarceration	461	98	100
B_Q27eUSPb	Activities – Class – Class attendance, unit during current incarceration	461	98	100
B_S27eUSP	Activities – Class – Class attendance, other specify during current incarceration	15	74	100
BQLANG	Language for Background Questionnaire	1,315	100	100
C_Q07USP	Status/work history (prior to incarceration) – Subjective status	1,315	100	100
C_Q08AUSP	Current status/work history – Ever paid work	481	100	100
C_Q09	Current status/work history – Years of paid work during lifetime	1,315	91	91
C_Q10AUSP	Current status/work history – Last 5 years – How many diff firms or organisations	525	84	88
COMPUTEREXPERIENCEUSP	Respondent experience with computer (including current prison job) (DERIVED BY CAPI)	1,315	100	100
D_Q01AUSP	Prison job – Job title	1,137	100	100
D_Q01BUSP	Prison job – Responsibilities	1,137	100	100
D_Q10USP	Prison job – Hours/week	1,137	99	100
E_Q01AUSP	Last job (prior to incarceration) – Job title	1,315	91	91
E_Q01BUSP	Last job (prior to incarceration) – Responsibilities	1,315	91	91
E_Q02AUSP	Last job (prior to incarceration) – Kind of business, industry or service	1,315	90	91
E_Q02BUSP	Last job (prior to incarceration) – Main product of firm or organisation	1,315	90	91
E_Q03US_P	Last job (prior to incarceration) – Economic sector	1,208	98	99
E_Q03USP	Last job (prior to incarceration) – Economic sector	1,208	98	99
E_Q04USP	Last job (prior to incarceration) – Employee or self-employed	1,208	99	99
E_Q05A1USP	Last job (prior to incarceration) – Start of work for employer – Age	1,022	98	99
E_Q05A2USP	Last job (prior to incarceration) – Start of work for employer – Year	1,022	98	99
E_Q05B1USP	Last job (prior to incarceration) – Start of work for business – Age	186	99	100
E_Q05B2USP	Last job (prior to incarceration) – Start of work for business – Year	186	99	100
E_Q06USP	Last job (prior to incarceration) – Amount of people working for employer	754	99	100
E_Q07AUSP	Last job (prior to incarceration) – Employees working for you	72	97	97

See notes at end of table.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
E_Q07BUSP	Last job (prior to incarceration) – Employees working for you – Count	10	81	81
E_Q08USP	Last job (prior to incarceration) – Type of contract	1,022	97	99
E_Q09USP	Last job (prior to incarceration) – Hours/week	186	100	100
E_Q10USP	Last job (prior to incarceration) – Reason for end of job	96	100	100
E_S08USP	Last job (prior to incarceration) – Other type of contract specified	1,022	99	99
F_Q01BUSP	Skill use work (prior to incarceration) – Time cooperating with co-workers	1,208	98	99
F_Q02AUSP	Skill use work (prior to incarceration) – How often – Sharing work-related info	1,022	99	100
F_Q02BUSP	Skill use work (prior to incarceration) – How often – Teaching people	16	100	100
F_Q02CUSP	Skill use work (prior to incarceration) – How often – Presentations	487	96	96
F_Q02DUSP	Skill use work (prior to incarceration) – How often – Selling	487	96	96
F_Q02EUSP	Skill use work (prior to incarceration) – How often – Advising people	487	96	96
F_Q03AUSP	Skill use work (prior to incarceration) – How often – Planning own activities	487	96	96
F_Q03BUSP	Skill use work (prior to incarceration) – How often – Planning others activities	487	96	96
F_Q03CUSP	Skill use work (prior to incarceration) – How often – Organising own time	487	96	96
F_Q04AUSP	Skill use work (prior to incarceration) – How often – Influencing people	487	95	96
F_Q04BUSP	Skill use work (prior to incarceration) – How often – Negotiating with people	487	95	96
F_Q05AUSP	Skill use work (prior to incarceration) – Problem solving – Simple problems	487	94	96
F_Q05BUSP	Skill use work (prior to incarceration) – Problem solving – Complex problems	487	94	96
F_Q06BUSP	Skill use work (prior to incarceration) – How often – Working physically for long	487	96	96
F_Q06CUSP	Skill use work (prior to incarceration) – How often – Using hands or fingers	487	96	96
F_Q07AUSP	Skill use work (prior to incarceration) – Not challenged enough	487	95	96
F_Q07BUSP	Skill use work (prior to incarceration) – Need more training	487	96	96
G_Q01AUSP	Skill use work (prior to incarceration) – Literacy – Read directions or instructions	487	96	96
G_Q01BUSP	Skill use work (prior to incarceration) – Literacy – Read letters memos or mails	487	96	96
G_Q01CUSP	Skill use work (prior to incarceration) – Literacy – Read newspapers or magazines	487	96	96
G_Q01DUSP	Skill use work (prior to incarceration) – Literacy – Read professional journals or publications	487	96	96

See notes at end of table.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
G_Q01EUSP	Skill use work (prior to incarceration) – Literacy – Read books	487	96	96
G_Q01FUSP	Skill use work (prior to incarceration) – Literacy – Read manuals or reference materials	487	96	96
G_Q01GUSP	Skill use work (prior to incarceration) – Literacy – Read financial statements	487	95	96
G_Q01HUSP	Skill use work (prior to incarceration) – Literacy – Read diagrams maps or schematics	487	96	96
G_Q02AUSP	Skill use work (prior to incarceration) – Literacy – Write letters memos or mails	487	96	96
G_Q02BUSP	Skill use work (prior to incarceration) – Literacy – Write articles	487	96	96
G_Q02CUSP	Skill use work (prior to incarceration) – Literacy – Write reports	487	96	96
G_Q02DUSP	Skill use work (prior to incarceration) – Literacy – Fill in forms	487	96	96
G_Q03BUSP	Skill use work (prior to incarceration) – Numeracy – How often – Calculating costs or budgets	487	96	96
G_Q03CUSP	Skill use work (prior to incarceration) – Numeracy – How often – Use or calculate fractions or percentages	487	96	96
G_Q03DUSP	Skill use work (prior to incarceration) – Numeracy – How often – Use a calculator	487	95	96
G_Q03FUSP	Skill use work (prior to incarceration) – Numeracy – How often – Prepare charts graphs or tables	487	96	96
G_Q03GUSP	Skill use work (prior to incarceration) – Numeracy – How often – Use simple algebra or formulas	487	96	96
G_Q03HUSP	Skill use work (prior to incarceration) – Numeracy – How often – Use advanced math or statistics	487	96	96
G_Q04USP	Skill use work (prior to incarceration) – ICT – Experience with computer in job	487	96	96
G_Q05AUSP	Skill use work (prior to incarceration) – ICT – Internet – How often – For mail	199	100	100
G_Q05CUSP	Skill use work (prior to incarceration) – ICT – Internet – How often – Work related info	199	100	100
G_Q05DUSP	Skill use work (prior to incarceration) – ICT – Internet – How often – Conduct transactions	199	100	100

See notes at end of table.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
G_Q05EUSP	Skill use work (prior to incarceration) – ICT – Computer – How often – Spreadsheets	199	100	100
G_Q05FUSP	Skill use work (prior to incarceration) – ICT – Computer – How often – Word	199	100	100
G_Q05GUSP	Skill use work (prior to incarceration) – ICT – Computer – How often – Programming language	199	100	100
G_Q05HUSP	Skill use work (prior to incarceration) – ICT – Computer – How often – Real-time discussions	199	100	100
G_Q06USP	Skill use work (prior to incarceration) – ICT – Computer – Level of computer use	201	97	100
G_Q07USP	Skill use work (prior to incarceration) – ICT – Computer – Got the skills needed	199	100	100
G_Q08USP	Skill use work (prior to incarceration) – ICT – Computer – Lack of skills affect career	199	100	100
H_Q01AUSP	Skill use everyday life (in prison) – Literacy – Read directions or instructions	1,315	100	100
H_Q01BUSP	Skill use everyday life (in prison) – Literacy – Read letters or memos	1,315	100	100
H_Q01CUSP	Skill use everyday life (in prison) – Literacy – Read newspapers or magazines	1,315	100	100
H_Q01DUSP	Skill use everyday life (in prison) – Literacy – Read professional journals or publications	1,315	100	100
H_Q01EUSP	Skill use everyday life (in prison) – Literacy – Read books	1,315	100	100
H_Q01FUSP	Skill use everyday life (in prison) – Literacy – Read manuals or reference materials	1,315	100	100
H_Q01GUSP	Skill use everyday life (in prison) – Literacy – Read financial statements	1,315	100	100
H_Q01HUSP	Skill use everyday life (in prison) – Literacy – Read diagrams maps or schematics	1,315	100	100
H_Q02AUSP	Skill use everyday life (in prison) – Literacy – Write letters or memos	1,315	100	100
H_Q02BUSP	Skill use everyday life (in prison) – Literacy – Write articles	1,315	100	100
H_Q02CUSP	Skill use everyday life (in prison) – Literacy – Write reports	1,315	100	100
H_Q02DUSP	Skill use everyday life (in prison) – Literacy – Fill in forms	1,315	100	100
H_Q03BUSP	Skill use everyday life (in prison) – Numeracy – How often – Calculating costs or budgets	1,315	100	100

See notes at end of table.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
H_Q03CUSP	Skill use everyday life (in prison) – Numeracy – How often – Use or calculate fractions or percentages	1,315	100	100
H_Q03DUSP	Skill use everyday life (in prison) – Numeracy – How often – Use a calculator	1,315	100	100
H_Q03FUSP	Skill use everyday life (in prison) – Numeracy – How often – Prepare charts graphs or tables	1,315	100	100
H_Q03GUSP	Skill use everyday life (in prison) – Numeracy – How often – Use simple algebra or formulas	1,315	100	100
H_Q03HUSP	Skill use everyday life (in prison) – Numeracy – How often – Use advanced math or statistics	1,315	100	100
H_Q04AUSP	Skill use everyday life – ICT – Ever used computer	1,114	100	100
H_Q04BUSP	Skill use everyday life (in prison) – ICT – Experience with computer everyday life	1,174	100	100
H_Q05EUSP	Skill use everyday life (in prison) – ICT – Computer – How often – Spreadsheets	236	100	100
H_Q05FUSP	Skill use everyday life (in prison) – ICT – Computer – How often – Word	236	100	100
H_Q05GUSP	Skill use everyday life (in prison) – ICT – Computer – How often – Programming language	236	100	100
I_Q04B	About yourself – Learning strategies – Relate new ideas into real life	1,315	97	100
I_Q04D	About yourself – Learning strategies – Like learning new things	1,315	98	100
I_Q04H	About yourself – Learning strategies – Attribute something new	1,315	98	100
I_Q04J	About yourself – Learning strategies – Get to the bottom of difficult things	1,315	98	100
I_Q04L	About yourself – Learning strategies – Figure out how different ideas fit together	1,315	98	100
I_Q04M	About yourself – Learning strategies – Looking for additional info	1,315	98	100
I_Q06A	About yourself – Political efficacy – No influence on the government	1,315	97	100
I_Q06DUSP1G	About yourself – Political efficacy – Information from family members, friends, or inmates	1,315	98	100
I_Q06DUSX1A	About yourself – Political efficacy – Information from newspapers	1,315	98	100
I_Q06DUSX1B	About yourself – Political efficacy – Information from magazines	1,315	98	100
I_Q06DUSX1D	About yourself – Political efficacy – Information from radio	1,315	98	100

See notes at end of table.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
I_Q06DUSX1E	About yourself – Political efficacy – Information from television	1,315	98	100
I_Q06DUSX1F	About yourself – Political efficacy – Information from books or brochures	1,315	98	100
I_Q07A	About yourself – Social trust – Trust only few people	1,315	97	100
I_Q07B	About yourself – Social trust – Other people take advantage of you	1,315	97	100
I_Q08	About yourself – Health – State	1,315	98	100
I_Q08USX1	About yourself – Health – Difficulty seeing print	1,315	98	100
I_Q08USX2	About yourself – Health – Difficulty hearing conversation	1,315	98	100
I_Q08USX3	About yourself – Health – Diagnosed learning disabled	1,315	98	100
I_Q10BUSP2G	About yourself – Health – Health information from family members, friends, or inmates	1,315	98	100
I_Q10BUSX2A	About yourself – Health – Health information from newspapers	1,315	98	100
I_Q10BUSX2B	About yourself – Health – Health information from magazines	1,315	98	100
I_Q10BUSX2D	About yourself – Health – Health information from radio	1,315	98	100
I_Q10BUSX2E	About yourself – Health – Health information from television	1,315	98	100
I_Q10BUSX2F	About yourself – Health – Health information from books or brochures	1,315	98	100
I_Q10BUSX2H	About yourself – Health – Health information from health professional	1,315	98	100
J_N05A2	Background – More than one language mentioned	1,315	100	100
J_Q02C	Background – Work situation of spouse or partner	217	89	100
J_Q03A	Background – Children	1,315	100	100
J_Q03B	Background – Number of children	942	100	100
J_Q03C	Background – Age of the child	261	100	100
J_Q03D1	Background – Age of the youngest child	680	100	100
J_Q03D2	Background – Age of the oldest child	680	100	100
J_Q04A	Background – Born in country	1,315	100	100
J_Q04BUS	Background – Country of birth	87	100	100
J_Q04C1	Background – Age of immigration	87	100	100
J_Q04C2	Background – Year of immigration	87	100	100

See notes at end of table.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as Nonresponse	Item response rate (percent) "Don't Know" as valid response
J_Q04DUSX1A	Background – Hispanic	1,315	100	100
J_Q04DUSX1B_01	Background – Hispanic origin – Mexican	267	100	100
J_Q04DUSX1B_02	Background – Hispanic origin – Puerto Rican	267	100	100
J_Q04DUSX1B_03	Background – Hispanic origin – Cuban	267	100	100
J_Q04DUSX1B_04	Background – Hispanic origin – Central/South America	267	100	100
J_Q04DUSX1B_05	Background – Hispanic origin – Other	267	100	100
J_Q04DUSX2_01	Background – Race – White	1,315	98	99
J_Q04DUSX2_02	Background – Race – Black	1,315	98	99
J_Q04DUSX2_03	Background – Race – Asian	1,315	98	99
J_Q04DUSX2_04	Background – Race – American Indian	1,315	98	99
J_Q04DUSX2_05	Background – Race – Native Hawaiian	1,315	98	99
J_Q05A1US	Background – First learned language	1,315	100	100
J_Q05A2US	Background – Second learned language	61	100	100
J_Q05A2USX2	Background – Age learned English	196	99	100
J_Q05BUS	Background – Language spoken at home	196	99	100
J_Q05CUSP4	Background – ESL class/tutor during current incarceration	173	100	100
J_Q05CUSP5	Background – Reason for ESL class/tutor during current incarceration	27	95	100
J_Q05CUSX1	Background – Language spoken most	196	100	100
J_Q05CUSX2	Background – English outside home	196	100	100
J_Q05CUSX3A	Background – Ability to understand spoken English	1,315	100	100
J_Q05CUSX3B	Background – Ability to speak English	1,315	100	100
J_Q05CUSX3D	Background – Ability to read English	1,315	100	100
J_Q05CUSX3E	Background – Ability to write English	1,315	100	100
J_Q05CUSX6	Background – Class/tutor learn English as adult	173	100	100
J_Q06A	Background – Mother/female guardian – Whether born in country	1,315	100	100
J_Q06B	Background – Mother/female guardian – Highest level of education	1,315	91	100
J_Q06BUS	Background – Mother/female guardian – Highest level of education	1,315	91	100

See notes at end of table.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
J_Q07A	Background – Father/male guardian – Whether born in countrty	1,315	97	100
J_Q07B	Background – Father/male guardian – Highest level of education	1,315	81	100
J_Q07BUS	Background – Father/male guardian – Highest level of education	1,315	81	100
J_Q08	Background – Number of books at home	1,315	98	100
J_S04B[2]	Background – Country of birth (other)	87	36	36
J_S05A1	Background – First learned language (other)	21	100	100
J_S05A2[2]	Background – Second learned language (other)	61	10	10
J_S05B	Background – Language spoken at home (other)	8	100	100
P_G_Q01a	Prison – Skill use work (prison job) – Literacy – Read directions or instructions	797	100	100
P_G_Q01b	Prison – Skill use work (prison job) – Literacy – Read letters or memos	797	100	100
P_G_Q01c	Prison – Skill use work (prison job) – Literacy – Read newspapers or magazines	797	100	100
P_G_Q01d	Prison – Skill use work (prison job) – Literacy – Read professional journals or publications	797	100	100
P_G_Q01e	Prison – Skill use work (prison job) – Literacy – Read books	797	100	100
P_G_Q01f	Prison – Skill use work (prison job) – Literacy – Read manuals or reference materials	797	100	100
P_G_Q01g	Prison – Skill use work (prison job) – Literacy – Read financial statements	797	100	100
P_G_Q01h	Prison – Skill use work (prison job) – Literacy – Read diagrams maps or schematics	797	100	100
P_G_Q02a	Prison – Skill use work (prison job) – Literacy – Write letters or memos	797	100	100
P_G_Q02b	Prison – Skill use work (prison job) – Literacy – Write articles	797	100	100
P_G_Q02c	Prison – Skill use work (prison job) – Literacy – Write reports	797	100	100
P_G_Q02d	Prison – Skill use work (prison job) – Literacy – Fill in forms	797	100	100
P_G_Q03b	Prison – Skill use work (prison job) – Numeracy – How often – Calculating costs or budgets	797	100	100
P_G_Q03c	Prison – Skill use work (prison job) – Numeracy – How often – Use or calculate fractions or percentages	797	100	100

See notes at end of table.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
P_G_Q03d	Prison – Skill use work (prison job) – Numeracy – How often – Use a calculator	797	100	100
P_G_Q03f	Prison – Skill use work (prison job) – Numeracy – How often – Prepare charts graphs or tables	797	100	100
P_G_Q03g	Prison – Skill use work (prison job) – Numeracy – How often – Use simple algebra or formulas	797	100	100
P_G_Q03h	Prison – Skill use work (prison job) – Numeracy – How often – Use advanced math or statistics	797	100	100
P_G_Q04	Prison – Skill use work (prison job) – ICT – Experience with computer in job	797	100	100
P_G_Q05e	Prison – Skill use work (prison job) – ICT – Computer – How often – Spreadsheets	81	100	100
P_G_Q05f	Prison – Skill use work (prison job) – ICT – Computer – How often – Word	81	100	100
P_G_Q05g	Prison – Skill use work (prison job) – ICT – Computer – How often – Programming language	81	100	100
P_G_Q06	Prison – Skill use work (prison job) – ICT – Computer – Level of computer use	81	95	95
P_G_Q07	Prison – Skill use work (prison job) – ICT – Computer – Got the skills needed	81	99	99
P_G_Q08	Prison – Skill use work (prison job) – ICT – Computer – Lack of skills affect career	797	100	100
P_Q010a	Prison – Education – Last 12 months – Time in class verbatim	290	99	99
P_Q010b	Prison – Education – Last 12 months – Time, unit	290	100	100
P_Q020	Prison – Education – Last 12 months – Degree offering	290	99	100
P_Q030	Prison – Education – Basic skills – Location	534	86	86
P_Q040	Prison – Education – Basic skills – Reason	536	86	86
P_Q040_S	Prison – Education – Basic skills – Reason other	14	100	100
P_Q050	Prison – Education – Prior – Reason stopped	1,293	100	100
P_Q050_S	Prison – Education – Basic skills – Reason stopped other	160	97	97
P_Q060	Prison – Education – Enrollment	1,315	91	92
P_Q070	Prison – Education – Waiting list	830	99	100

See notes at end of table.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
P_Q080	Prison – Education – Enrollment – Degree	830	99	100
P_Q090	Prison – Education – Enrollment – Reason	830	100	100
P_Q090_S	Prison – Education – Enrollment – Reason other	22	100	100
P_Q100	Prison – Education – Enrollment – Reason not	367	99	100
P_Q100_S	Prison – Education – Enrollment – Reason not other	191	99	100
P_Q120	Prison – Education – Current – Highest level	1,315	100	100
P_Q130	Prison – Education – Current – Course of study	529	99	100
P_Q130_S	Prison – Education – Current – Course offered other	42	99	100
P_Q140a	Prison – Education – Current – Time in class verbatim	529	99	100
P_Q140b	Prison – Education – Current – Time, unit	529	99	100
P_Q150	Prison – Education – Current – Reason	529	100	100
P_Q150_S	Prison – Education – Current – Reason other	40	100	100
P_Q160_1	Prison – Incarceration – Most recent – Month	1,315	98	100
P_Q160_2	Prison – Incarceration – Most recent – Year	1,315	100	100
P_Q170	Prison – Incarceration – Prior	1,315	100	100
P_Q180	Prison – Incarceration – Release	1,315	99	100
P_Q190a	Prison – Education – Current – Readiness class	1,315	100	100
P_Q190a_1	Prison – Education – Readiness class – Hours	330	96	100
P_Q190a_2	Prison – Education – Readiness class – Reason	330	100	100
P_Q190b	Prison – Education – Current – Parenting class	1,315	100	100
P_Q190b_1	Prison – Education – Parenting class – Hours	225	97	100
P_Q190b_2	Prison – Education – Parenting class – Reason	225	100	100
P_Q190c	Prison – Education – Current – Life skills class	1,315	100	100
P_Q190c_1	Prison – Education – Life skills class – Hours	432	98	100
P_Q190c_2	Prison – Education – Life skills class – Reason	432	100	100
P_Q190d	Prison – Education – Current – Adjustment class	1,315	100	100
P_Q190d_1	Prison – Education – Adjustment class – Hours	505	97	100

See notes at end of table.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
P_Q190d_2	Prison – Education – Adjustment class – Reason	505	100	100
P_Q190e	Prison – Education – Current – Addiction group	1,315	100	100
P_Q190e_1	Prison – Education – Addiction group – Hours	546	98	100
P_Q190e_2	Prison – Education – Addiction group – Reason	546	100	100
P_Q190f	Prison – Education – Current – Mental health group	1,315	100	100
P_Q190f_1	Prison – Education – Mental health group – Hours	223	96	100
P_Q190f_2	Prison – Education – Mental health group – Reason	223	100	100
P_Q190g	Prison – Education – Current – Inmate assistance group	1,315	100	100
P_Q190g_1	Prison – Education – Inmate assistance group – Hours	117	99	100
P_Q190g_2	Prison – Education – Inmate assistance group – Reason	117	100	100
P_Q190h	Prison – Education – Current – Religious study group	1,315	100	100
P_Q190h_1	Prison – Education – Religious study group – Hours	567	97	100
P_Q190h_2	Prison – Education – Religious study group – Reason	567	100	100
P_Q190i	Prison – Education – Current – Ethnic/racial organization	1,315	100	100
P_Q190i_1	Prison – Education – Ethnic/racial organization – Hours	51	96	100
P_Q190i_2	Prison – Education – Ethnic/racial organization – Reason	51	100	100
P_Q190j	Prison – Education – Current – Other	1,315	100	100
P_Q190j_1	Prison – Education – Current – Other specified	208	100	100
P_Q190j_2	Prison – Education – Other – Hours	208	97	100
P_Q190j_3	Prison – Education – Other – Reason	208	100	100
P_Q220	Prison – Education – Job training	1,315	100	100
P_Q230	Prison – Education – Waiting list job training	1,011	99	100
P_Q240_01	Prison – Reason job skill program, 1st response	451	100	100
P_Q240_02	Prison – Reason job skill program, 2nd response	451	100	100
P_Q240_03	Prison – Reason job skill program, 3rd response	451	100	100
P_Q240_04	Prison – Reason job skill program, 4th response	451	100	100
P_Q240_05	Prison – Reason job skill program, 5th response	451	100	100

See notes at end of table.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
P_Q240_06	Prison – Reason job skill program, 6th response	451	100	100
P_Q240_S	Prison – Education – Job training – Reason other	17	100	100
P_Q250	Prison – Education – Job training – Reason not	864	99	100
P_Q250_S	Prison – Education – Job training – Reason not other	320	100	100
P_Q260	Prison – Education – Job training – Intent	147	94	100
P_Q270	Prison – Education – Job training – Hours spent	451	100	100
P_Q280	Prison – Education – Job training – Hours scheduled	451	99	100
P_Q290	Prison – Education – IT certification	1,315	100	100
P_Q300	Prison – Education – IT test for certification	61	100	100
P_Q310	Prison – Education – IT test prepared	57	100	100
P_Q320	Prison – Education – Certificate other than IT	1,315	100	100
P_Q330	Prison – Education – Non-IT test for certification	258	100	100
P_Q340	Prison – Education – Non-IT test prepared	237	99	100
P_Q340_S	Prison – Education – Non-IT test prepared, other	3	100	100
P_Q350a	Prison – Education – Current – Time spent, amount	1,315	98	100
P_Q350b	Prison – Education – Current – Time unit	1,315	98	100
P_Q360	Prison – Education – Access to library	1,315	99	100
P_Q370	Prison – Education – Library, often use	1,169	100	100
P_Q380	Prison – Education – Library, length to access	1,169	98	100
P_Q390	Prison – Jobs – Current incarceration	1,315	100	100
P_Q400	Prison – Jobs – Waiting list	518	98	100
P_Q410	Prison – Jobs – Ever in prison	518	100	100
P_Q420	Prison – Jobs – Site	1,137	100	100
P_Q430_01	Prison – Income sources (prior to incarceration), 1st response	1,315	100	100
P_Q430_02	Prison – Income sources (prior to incarceration), 2nd response	1,315	100	100
P_Q430_03	Prison – Income sources (prior to incarceration), 3rd response	1,315	100	100
P_Q430_04	Prison – Income sources (prior to incarceration), 4th response	1,315	100	100

See notes at end of table.

Table C-2. PIAAC 2014 National Supplement Prison survey weighted item response rates (Items with response rates below 85 percent asked of at least 20 sampled persons are highlighted)—Continued

Background Questionnaire item	Description	Item sample size	Item response rate (percent) "Don't Know" as nonresponse	Item response rate (percent) "Don't Know" as valid response
P_Q430_05	Prison – Income sources (prior to incarceration), 5th response	1,315	100	100
P_Q430_06	Prison – Income sources (prior to incarceration), 6th response	1,315	100	100
P_Q430_07	Prison – Income sources (prior to incarceration), 7th response	1,315	100	100
P_Q430_08	Prison – Income sources (prior to incarceration), 8th response	1,315	100	100
P_Q430_S	Prison – Income sources (prior to incarceration), other	78	99	99
P_Q440	Prison – Background – Marital status	1,315	100	100

[1]Response rate is below 85 percent for this CAPI derived variable because the derivation did not distinguish between a valid skip and nonresponse. As a result, valid skips were counted as nonresponse.
[2]Response rate is below 85 percent since valid skips were not coded as such in the US–specific variable used for routing to this item.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC) 2014.

Appendix D

NRBA Technical Notes

This page intentionally blank.

Appendix D. NRBA Technical Notes

D.1 Total Survey Error

There are two major components of total survey error: sampling error and nonsampling error. Sampling error is the error that occurs because population estimates are based on a sample rather than a census. Due to clustering effects typical of household surveys, the sample size can be misleading to users when judging the magnitude of sampling error. Therefore, precise measurement of sampling error is necessary and is facilitated in PIAAC through the replication method known as the stratified jackknife technique. A description of the PIAAC 2012 and 2014 sample design, weighting procedures, and variance estimation methods can be found in chapters 3 and 8.

Nonsampling error contains all sources of error besides sampling error. According to Lessler and Kalsbeek (1992), there are three subcomponents of nonsampling error: (1) frame error, (2) nonresponse error, and (3) measurement error. This report is focused on nonresponse error, which is the error arising from failure to obtain a response, whether it is unit nonresponse or item nonresponse. A key measure of the impact of nonresponse on total survey error is nonresponse bias. Nonresponse bias can be substantial when two conditions hold: (1) when response rate is relatively low, and (2) when the difference between the characteristics of respondents and nonrespondents is relatively large. An estimate for nonresponse bias can be expressed as follows:

$$\text{Bias}(\bar{y}_R) = (1 - W_R)(\bar{Y}_R - \bar{Y}_N),$$

where W_R is the response rate and $\bar{Y}_R$ and $\bar{Y}_N$ are the mean values of the survey items estimated among the respondents and nonrespondents, respectively. Because survey values for nonrespondents are not available, nonresponse bias is not known and can only be estimated by using data available for both respondents and nonrespondents.

An alternative model of nonresponse assumes each sampled person has a certain propensity to respond, and NRB in a characteristic is a function of the covariance between the response propensity and the characteristic:

$$\text{Bias}(\bar{y}_R) \frac{\sigma_{yp}}{\bar{p}}$$

where σ_{yp} is the covariance between the outcome variable and response propensity, and $\bar{p}$ is the mean response propensity.

D.2 Response Rate Computations

The response rates shown in table D-1 were computed for the three components given in the PIAAC standards. The disposition codes (DISP) in table D-1 are described more fully in table D-2 (Screener), table D-3 (BQ), and table D-4 (assessment).

Table D-1. Actual response rates

Stage	Actual response rate			Description
Screener	COMPLETE/ELIGIBLE			
	COMPLETE	=	C^s	Completed screeners
	ELIGIBLE	=	$HH^s - I^s - U^s * (I^s/K^s)$	All sampled households
	C^s	=	DISP_SCR(01,02)	HHs known to be ineligible
	HH^s	=	All sampled households[1]	
	I^s	=	DISP_SCR(19,22,26,28)	HHs with unknown eligibility status
	U^s	=	DISP_SCR(04,05,17,20,21,24)	
	K^s	=	DISP_SCR(01,02,03,07,09, 12-16,19,22,26,28)	HHs with known eligibility status
Background questionnaire	COMPLETE/ELIGIBLE			
	COMPLETE	=	$C^b + LR^b$	Completed BQ cases
	ELIGIBLE	=	$SP^b – D^b – I^b$	Literacy-related nonrespondents
	C^b	=	DISP_CIBQ(01,90[2])	All sampled persons
	LR^b	=	DISP_CIBQ(07,08,09) and	SPs with a disability
	QCFLAG_LR	=	1	SPs known to be ineligible
	SP^b	=	All sampled persons[3]	
	D^b	=	DISP_CIBQ(12,13,15,16)	
	I^b	=	DISP_CIBQ(18,25)	
Assessment[4]	COMPLETE/ELIGIBLE			Completed assessments
	COMPLETE	=	$C^a + LR^a$	Literacy-related nonrespondents
	ELIGIBLE	=	$C^b – D^a – I^a$	Completed BQ cases
	C^a	=	DISP_MAIN(01,90[2])	SPs with a disability
	LR^a	=	DISP_MAIN(07,08,09,14)	SPs known to be ineligible
	C^b	=	DISP_CIBQ(01,90[2])	
	D^a	=	DISP_MAIN(12,13,15,16)	
	I^a	=	DISP_MAIN(18)	

[1] Includes the original sample of dwelling units plus the reserve sample released as a random subset of the population.
[2] A code of 90 represents technical problems during the BQ or assessment. For the U.S. and U.S. reports, technical problems are treated as nonresponses. For PIAAC international reports, they are treated as completes to be consistent with the PIAAC Standards and Guidelines.
[3] Includes the original sample of persons plus the reserve sample released as a random subset of the population.
[4] The assessment response rate with reading components can be computed by replacing DISP_MAIN with DISP_MAINWRC. The reading components conditional response rate is the assessment response rate with reading components divided by the assessment response rate without reading components.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Table D-2. Disposition codes for PIAAC Screener

Description	Code
Complete – 1 sample person selected	01
Complete – 2 sample persons selected	02
Partial complete/break-off	03
Refusal – household member	04
Refusal – gatekeeper	05
Language problem (i.e., unable to communicate in English or Spanish)	07
Learning/mental disability	09
Hearing impairment	12
Blindness/visual impairment	13
Speech impairment	14
Physical disability	15
Other disability	16
Other (unspecified), such as sickness, falsification or unusual circumstances	17
Complete – no eligible sample persons	19
Unable to locate dwelling unit	20
Maximum number of calls	21
Dwelling unit under construction	22
Temporarily absent/unavailable during field period	24
Vacant dwelling unit (e.g., holiday or temporary residence only, such as cabins on a lake)	26
Duplication – already interviewed	27
Address not a dwelling unit (for example, nonresidential units such as businesses, government offices, and other organizations and residential units such as institutions (e.g., prisons or sanitariums) and military barracks)	28

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Table D-3. Disposition codes for PIAAC Background Questionnaire

Description	Code
Complete	01
Partial complete/break-off (e.g., a paused interview that eventually cannot be completed)	03
Refusal – sample person (e.g., for refusal to participate due to time constraints or lack of interest)	04
Refusal – other	05
Language problem (i.e., unable to communicate in English or Spanish)	07
Reading and writing difficulty	08
Learning/mental disability	09
Hearing impairment	12
Blindness/visual impairment	13
Speech impairment	14
Physical disability	15
Other disability	16
Other (unspecified), such as sickness, falsification or unusual circumstances)	17
Death	18
Maximum number of calls (e.g., respondent not successfully contacted)	21
Temporarily absent/unavailable during field period (e.g., travelling and will not be back during the field period or moved within the same community but not successfully located by the interviewers)	24
Ineligible (e.g., moved outside the country)	25
Duplication – already interviewed	27
Technical problem	90

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Table D-4. Disposition codes for PIAAC direct assessments

Assessment	Description	Code
Main	Complete	01
Main	Partial complete/break-off (e.g., a paused interview that eventually cannot be terminated)	03
Main	Refusal – sample person (e.g., for refusal to participate due to time constraints)	04
Main	Refusal – other	05
Main	Language problem (i.e., did not read or write in English)	07
Main	Reading and writing difficulty	08
Main	Learning/mental disability	09
Main	Hearing impairment	12
Main	Blindness/visual impairment	13
Main	Speech impairment	14
Main	Physical disability	15
Main	Other disability	16
Main	Other (unspecified), such as sickness, falsification or unusual circumstances	17
Main	Death	18
Main	Maximum number of calls (e.g., respondent not successfully contacted)	21
Main	Temporarily absent/unavailable during field period	24
Main	Duplication – already interviewed	27
Main	Technical problem	90
Main	Missing paper booklet	91

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

D.3 Tests of Significance

The comparisons discussed in this report have been tested for statistical significance. For example, when comparing results obtained from the full sample with those obtained from only the responding sample units, tests of statistical significance were used to establish whether or not the observed differences are statistically significant. The estimation of the standard errors that was required in order to undertake the tests of significance was complicated by the complex sample and assessment designs, both of which generated error variance. Together they mandate a set of statistically complex procedures in order to estimate the correct standard errors. As a consequence, the estimated standard errors contain a sampling variance component estimated by the stratified jackknife method. Details on the procedures used can be found in the WesVar 4.3 User's Guide (Westat 2007).

Two kinds of statistical tests are included in the report: t tests and chi-square tests.

D.4 *T* Tests

t tests were used for testing for the hypothesis that no difference exists between the estimated percentages of variables for two groups. In particular, comparisons were made between the final PIAAC estimates and estimates based on alternative weighting adjustments. Suppose that $\bar{x}_A$ and $\bar{x}_B$ are the percentages for two groups that are being compared and $se(\bar{x}_A - \bar{x}_B)$ is the standard error of the difference between the percentages which accounts for the complex survey design. Then the t test is defined as

$$t = \frac{|\bar{x}_A - \bar{x}_B|}{se(\bar{x}_A - \bar{x}_B)}$$

This statistic is then compared to the critical values of the appropriate Student t-distribution, to determine whether the difference is statistically significant. The appropriate number of degrees of freedom for the distribution is given by the number of primary sampling units in the design minus the number of sampling strata. Note that this procedure recognizes that the two samples in question were not independent samples. This effect was accounted for in calculating the standard error of the difference.

D.5 Chi-Square Tests

Chi-square tests are used for testing whether two distributions of a given categorical variable are different. These tests are conducted in a way that reflects the impact of the complex sample design on sampling variance. In this instance one distribution is for the responding sample and one for nonrespondents. Suppose that the categorical variable in question has c levels, cross-tabulated to produce weighted proportions p. The Pearson chi-square statistic is calculated as

$$X^2 = n \sum_{i=1}^{2} \sum_{j=1}^{c} \left(p_{ij} - p_{i\cdot} p_{\cdot j}\right)^2 / p_{i\cdot} p_{\cdot j}$$

where j denotes the categories of the categorical variable, i indexes the samples (respondents and nonrespondents), and n indicates the overall sample size. This statistic is not suitable for use with these data because it does not account for the complex sample design used to collect the data.

Therefore, the Pearson chi-square statistic had to be modified to account for the complex design. The resulting test statistic is referred to as the Rao-Scott adjusted chi-square statistic. The Rao-Scott adjustment was based on the Satterthwaite approximation and referred to here as the Satterthwaite-adjusted chi-square statistic.[1] The number of degrees of freedom for the chi-square test, normally given as $(c-1)$, where c is the number of categories of the categorical variable for each distribution, is also modified to account for the complex design. The modified test statistic is then compared to the chi-square distribution with the appropriate number of degrees of freedom, to determine whether the difference in the two distributions is statistically significant. A detailed description of the technique is provided in the following paragraphs (also see Rao and Thomas 2003).

The first step in the calculation of the Satterthwaite-adjusted chi-square statistic is to form the following vector:

$$Y = \sqrt{n}\begin{pmatrix} p_{11} - p_{1\cdot}p_{\cdot 1} \\ p_{12} - p_{1\cdot}p_{\cdot 2} \\ \vdots \\ p_{rc} - p_{r\cdot}p_{\cdot c} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{rc} \end{pmatrix}$$

An rc x 1 vector made up of the products of the marginal proportions is defined as

$$p = \begin{pmatrix} p_{1\cdot}p_{\cdot 1} \\ p_{1\cdot}p_{\cdot 2} \\ \vdots \\ p_{r\cdot}p_{\cdot c} \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{rc} \end{pmatrix}$$

For each replicate, an rc x rc matrix is calculated whose ij-th element is made up of

$$(y_{ig} - y_i)(y_{jg} - y_j)$$

where y_{ig} and y_{jg} are the i-th and j-th elements of **Y** calculated for the g-th replicate and y_i and y_j are the corresponding full-sample values. The ij-th element of the estimated covariance matrix for **Y, B**=cov(**Y**), is calculated using the following formula:

$$B_{ij} = \sum_{g=1}^{G}(y_{ig} - y_i)(y_{jg} - y_j)$$

The Satterthwaite's approximation to degrees of freedom for the chi-square statistic to be calculated is

$$v = \frac{\left(\sum_{i=1}^{rc} B_{ii} p_i\right)^2}{\sum_{i=1}^{rc}\sum_{j=1}^{rc}\frac{B_{ij}^2}{p_i p_j}}$$

[1]Compared to the other type of Rao-Scott adjusted chi-square statistic, the Satterthwaite-adjusted chi-square statistic can provide some performance advantages for Type I error and power, particularly when the design effects are variable (Thomas and Rao 1987; Rao and Thomas 1989).

Finally, the adjusted chi-square statistic is defined as

$$\mathrm{RS3} = \frac{\mathrm{X}^2}{\sum_{\mathrm{i}=1}^{\mathrm{rc}} \frac{\mathrm{B_{ii}}}{\mathrm{p_i}}}$$

D.6 Logistic Regression Models

A linear model for investigating the relationship between binary (dichotomous) outcomes and a set of explanatory variables is referred to as a logistic regression model. The data are assumed to follow a binomial distribution, with probabilities that depend on the independent variables. In this instance the binary outcome of interest was whether or not the sampled unit completed the PIAAC Background Questionnaire.

Let p_i denote the probability that the i-th sampled person will participate. Under the logistic regression model, the log odds of response propensity (expressed in terms of the logarithm of $p_i/(1-p_i)$), is assumed to have the following linear form:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \quad + \beta_p X_{pi}$$

where $X_{1i}, X_{2i}, \ldots, X_{pi}$ are p auxiliary variables associated with the i-th sampled person, and $\beta_0, \beta_1, \ldots, \beta_p$ are coefficients to be estimated. Asymptotic assumptions are used to develop statistical tests to determine which, if any, of the coefficients are significantly different from zero. In the analyses in this report the standard procedures for carrying out logistic regression analyses have been modified both to incorporate the sampling weights in the estimation of the coefficients, and to reflect the effect of the complex sample design on the variance-covariance matrix of the coefficients.

The Newton-Raphson algorithm is used to iteratively solve for parameter solutions in the logistic regression. Let $q(\boldsymbol{\beta}) = \partial L_n(\boldsymbol{\beta})/\partial \boldsymbol{\beta}$ be the vector of first partial derivatives of the sample log-likelihood with respect to $\boldsymbol{\beta}$. Let $H(\boldsymbol{\beta})$ be the matrix of second partial derivatives (or Hessian) of the sample log-likelihood having entries $\frac{\partial^2 \mathrm{L}}{\partial \beta_{\mathrm{a}} \partial \beta_{\mathrm{b}}}$, where β_{a} and β_{b} are two separate components of $\boldsymbol{\beta}$. Denote by $\mathbf{q}^t$ and $\mathbf{H}^t$ the values of $\mathbf{q}(\boldsymbol{\beta})$ and $\mathrm{H}(\boldsymbol{\beta})$ evaluated at $\mathbf{b}^t$, the value of the estimate $\mathbf{b}$ at step t.

The general approach is to approximate the sample log-likelihood at the desired estimate, $L_n(\mathrm{b})$, at step t in the iterative process near the point $\mathbf{b}^t$ by a second-order Taylor series expansion:

$$L_n^t(\mathbf{b}) \cong L_n(\mathbf{b}^t) + \mathbf{q}^{t\prime}(\mathbf{b} - \mathbf{b}^t) + \tfrac{1}{2}(\mathbf{b} - \mathbf{b}^t)' \mathbf{H}^t(\mathbf{b} - \mathbf{b}^t).$$

Solving $\frac{\partial L^t}{\partial \mathbf{b}} = \mathbf{q}^t + \mathbf{H}^t(\mathbf{b} - \mathbf{b}^t) = 0$ for $\mathbf{b}$ yields the iteration equations

$$\mathbf{b}^{t+1} = \mathbf{b}^t - [\mathbf{H}^t]^{-1}\mathbf{q}^t$$

assuming $\mathbf{H}^t$ has an inverse. Given an initial value for $t = 0$, the set of iteration equations is solved for $\mathbf{b}^1$, $\mathbf{b}^1$ is used to solve for $\mathbf{b}^2$, and so on, until the convergence criterion is satisfied. The standard error of each estimated β coefficient is calculated using the stratified jackknife method and repeating the procedure for each replicate.

Appendix E

Data User Guidance

This page intentionally blank.

Appendix E. Data User Guidance

This appendix provides information to help support analyses of the combined 2012 and 2014 U.S. PIAAC data, which were released following the data collection for the U.S. PIAAC National Supplement in 2013–14. U.S. PIAAC microdata are in the following forms:

- U.S. national public-use file (PUF) for the combined household sample (16- to 74-year-olds in the combined Main Study and National Supplement household sample);
- U.S. national PUF for the prison sample (16- to 74-year-olds);
- U.S. national restricted-use file (RUF) for the combined household sample (16- to 74-year-olds in the combined Main Study and National Supplement household sample);
- U.S. national RUF for the prison sample (16- to 74-year-olds);
- Main Study household sample files to verify published results from the 2012 sample; and
- International PUF for the combined household sample (16- to 65-year-olds in the combined Main Study and National Supplement household sample).

The U.S. national PUF for the combined household sample includes respondents (ages 16–74) from both the Main Study and National Supplement household samples. It contains variables that are common across all countries that participated in PIAAC, variables that were administered to U.S. respondents only, and U.S.-only derived variables. The U.S. national PUF for the prison sample contains only part of the variables that are common across all PIAAC countries[1] and other variables that were administered to the U.S. prison respondents only. In addition, the PUF for the prison sample contains some U.S.-only variables that are also in the U.S. household sample.

For both the combined household sample and the prison sample, the U.S. national RUFs contain more variables than the PUFs. In addition to the variables in the PUFs, the RUFs contain detailed versions of variables and additional data collected through U.S. specific questionnaire routing (see the section "RUF variables related to U.S. specific routing" below for more information). The RUFs can be accessed through a restricted-use license agreement with the National Center for Education Statistics (NCES). To review the NCES protocols for a restricted-use license arrangement, please see https://nces.ed.gov/statprog/rudman/ (accessed February 2, 2016). To apply for a restricted-use data license granting access to the PIAAC RUFs, follow the instructions at https://nces.ed.gov/statprog/instruct.asp.

The combined sample is to be used for analyses of PIAAC data that are intended for publication. The combined sample is a larger sample, and supports more detailed analysis, as well as more accurate estimates. Besides the larger combined sample, the improved accuracy of estimates are due in part to the revised population estimates based on the Census 2010 data, which were unavailable when PIAAC 2012 went into the field. The PUF and RUF files released for the 2012 Main Study household sample remain available upon request for verification of results published from that sample. The international PUF

[1] Note that although the prison sample data files contain some of the international variables, the prison file estimates are not comparable to the international PIAAC file estimates since the international PIAAC sample does not contain people in prison.

includes respondents from both the Main Study sample and National Supplement household samples, but is limited to 16- to 65-year-olds only per the age range for which the data were collected across PIAAC-participating countries in 2011–12. It contains variables that are common across all countries that participated in PIAAC and following the internationally imposed routing. Some variables in the international PUF were suppressed for U.S. respondents due to confidentiality concerns. The international PUF will be available on the OECD website[2] in the summer of 2016.

The U.S. national PUFs and RUFs are in the format of SPSS and SAS files, and have associated codebooks in PDF format. The SPSS files are fully labeled and considered self-documenting, and the SAS format scripts can be used to assign value labels. The codebook for the variables that reside on the PUF can be accessed at the following website: https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2016667. (The codebook for the variables that reside on the RUF is provided along with the RUF dataset.)

The PUF and RUF codebooks contain the following details for each variable in the file:

- NAME—displays the unique identifier for the variable in the data file;
- LABEL—displays a short description associated with the variable;
- QUESTION—displays the survey question wording associated with the variable;
- NUMBER—displays the order of the variable in the data file;
- TYPE—displays the variable type (integer, string/character, or numeric/floating point);
- WIDTH/DECIMALS—displays the number of digits or characters in the variable and the number of digits to the right of the decimal point;
- LOCATION—displays the position of the variable within the data recorded, indicated with the start and end columns;
- COUNT—displays the number of records for each level of categorical variables;
- PERCENT—displays the percent of records for each level of categorical variables;
- VALUE—displays a list of the possible coded response options for categorical variables;
- DESCRIPTION—displays a description of each coded response option for categorical variables;
- VALID N—displays the number of records with nonmissing values for continuous variables;
- MINIMUM—displays the minimum value for continuous variables;
- MAXIMUM—displays the maximum value for continuous variables;
- MEAN—displays the average value for continuous variables;

[2] The OECD website is http://www.oecd.org/site/piaac/publicdataandanalysis.htm (accessed February 2, 2016).

- STANDARD DEVIATION—displays the standard deviation for continuous variables; and
- MISSING N—displays the number of records with missing values for continuous variables.

Each respondent is uniquely identified by the variable SEQID.

Most variables use a standard scheme for missing values. Codes are used to indicate item nonresponse and legitimate skips. The specific coded values for each item are detailed in the codebook. The coding scheme in Table E-1 was used for the SAS and SPSS datasets.

Table E-1. Missing value codes in SAS and SPSS datasets

Missing value code description	SAS missing value code	SPSS missing value code[1]
Valid skip	.V for numeric variables, 9-fill ending with 6 (for e.g., 6, 96, 996, …) for character variables	9-fill ending with 6 (for e.g., 6, 96, 996, …)
Don't know	.D for numeric variables, 9-fill ending with 7 (for e.g., 7, 97, 997, …) for character variables	9-fill ending with 7 (for e.g., 7, 97, 997, …)
Refused	.R for numeric variables, 9-fill ending with 8 (for e.g., 8, 98, 998, …) for character variables	9-fill ending with 8 (for e.g., 8, 98, 998, …)
Not stated or inferred	.N for numeric variables, 9-fill ending with 9 (for e.g., 9, 99, 999, …) for character variables	9-fill ending with 9 (for e.g., 9, 99, 999, …)
Don't Know/Refused	.M	9-fill ending with 8 (for e.g., 8, 98, 998, …)
Unknown	.U	9-fill ending with 5 (for e.g., 5, 95, 995, …)
<16 years old[2]	.A	93
>65 years old[2]	.B	94
Student in regular cycle of studies[3]	.A	4
Native born[4]	.A	94
Has not worked more than 5 years[5]	.A for numeric variables, 9995 for character variables	94 for numeric variables, 9995 for string variables.
Still in education[6]	.A	94
Citizen by birth[7]	.A	5

See note at end of table.

Table E-1. Missing value codes in SAS and SPSS datasets—Continued

Missing value code description	SAS missing value code	SPSS missing value code[1]
Adults older than 24[8]	.A	4
Did not participate[9]	.A	9994, 9994.00

[1] The missing value code depends on the length of the field for each data item. All SPSS missing value codes are classified as missing except for string variables in which the missing value codes are defined in the value labels.
[2] This description is for variable AGEG10LFS_T.
[3] This description is for variables: FAET12, FAET12JR, FAET12NJR, FNFAET12, FNFAET12JR, and FNFAET12NJR.
[4] This description is for variables: IMYRS and IMYRS_C.
[5] This description is for all industry or occupation related variables: ISCOSKIL4, ISCO1C, ISCO1L, ISCO2C, ISCO2L, ISIC1C, ISIC1L, ISIC2C, ISIC2L, ISIC4_C, ISIC4_L, ISIC4_CUS_C, and ISIC4_LUS_C.
[6] This description is for variables: LEAVEDU and LEAVEDUUS_C.
[7] This description is for variable J_Q04c2_T.
[8] This description is for variable LEAVER1624.
[9] This description is for variables: NFEHRS, NFEHRSJR, and NFEHRSNJR.
SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

The data item types in the U.S. national PUF and RUF include identification variables, background questionnaire item variables, derived variables, weighting variables, and assessment item scores. Users can find more information on the background questionnaire item variables through the background questionnaire available on the NCES PIAAC website.[3] A list of all the variables in the data file can be found in the codebooks. Side-by-side lists of the variables that are in the PUF and RUF are given in tables E-5 (for the household files) and E-6 (for the prison files) at the end of this appendix. The variables that are common across all countries that participated in PIAAC were provided by OECD. The variable names for these items were generally maintained as provided by OECD. U.S.-specific variables were then added to the end of the data file, and these variables have "US" at or near the end of the names. The prison files also include prison-only variables which have either the prefix "P_" or the suffixes[4] "USP" and "_P" in the names. The prison-only variables are not directly comparable[5] with the variables in household files and hence named differently. The variables that are comparable between the household and prison samples were named exactly the same in the household and prison files.

Coding of selected variables captured in the electronic Background Questionnaire (BQ) used the coding schemes specified by the PIAAC Consortium. Detailed responses to questions regarding the respondent's current and past employer, industry, occupation, and duties were used to assign industry and occupation codes. Occupation and industry coding was done to the full four digits when enough information for the four digit level was available. Codes from the 2008 International Standard Classification of Occupations (ISCO-08) were used to code the occupation of the respondent. Industry codes were done using the International Standard Industrial Classification of All Economic Activities (ISIC), Revision 4. The coded responses for industry and occupation included in the RUFs are as follows:

ISCO08_C Respondent's current ISCO-08 occupation code

ISCO08_L Respondent's last ISCO-08 occupation code

ISIC4_C Respondent's ISIC rev4 code of industry of current job (household sample only)

[3] https://nces.ed.gov/surveys/piaac/english_q.asp, accessed February 2, 2016.
[4] Note the suffix "USP" precedes most other suffixes such as "_C" or "_WLE_CA".
[5] The variables were not deemed comparable because of routing, question wording or context reasons.

ISIC4_L Respondent's ISIC rev4 code of industry of last job

ISCO08_US Respondent's apprentice ISCO-08 occupation code (household sample only)

Disclosure analyses and data masking were conducted to provide reasonable assurance that the data disseminated for PIAAC would not lead to the identification of individuals. Even though analysts will need to sign a license to use the restricted-use data for analysis purposes only, the data are masked to further protect the identities of individuals, and to provide consistency between results from both the RUF and PUF. The microdata were masked through techniques of data coarsening and data swapping so that one cannot be certain of the identity of individuals, nor that the data are attributable to the individual. The masking procedures were done in a manner that has a negligible impact on data analyses.

RUF Variables Related to U.S. Specific Routing

The U.S. adaptation of the Background Questionnaire introduced a new routing path that caused collection of additional responses that would not occur in the international routing. This occurred after question B_Q2bUS (in both household and prison questionnaire) and routed high school graduates to follow-up questions about starting, but not finishing, an advanced degree and further down-path questions. In the household questionnaire, this had the potential to route about 550 respondents through the group of questions B_Q03a – B_Q05b in variation of the international routing. Altogether thirty-eight questions (B_Q03a - B_Q26b) were affected. In the prison questionnaire, this had the potential to route about 290 respondents through the questions B_Q03a and B_Q03b in variation of the international routing and affected those two questions. In both the national and international PUFs, the additional data collected through the U.S. specific routing were removed to be consistent with the other countries. However, these data are included in the RUF and can be identified by the "USR" suffix in variable names.

Variable Changes from the PIAAC Main Study File

The combined household sample RUF and national PUF not only include more respondents than the previously-released Main Study files, but also contain new variables that were not on the Main Study files. These variables are shown in Table E-2. All of these variables were added to the RUF and some of them were added to the national PUF to enhance the analytic value of PIAAC data and to facilitate comparison with the PIAAC prison sample and international files. In addition, the variables AGEG5LFS and AGEG10LFS were replaced with AGEG5LFSEXT and AGEG10LFSEXT with slightly different age categories to incorporate the 66-74 years old that were added to the National Supplement sample as a category on its own, maintaining the age breakdowns on the Main Study ages 16-65 sample. Also note that I_Q010bUSX1 was renamed I_Q10bUSX1 to be consistent with the Background Questionnaire.

Table E-2. New variables in the household sample files

Variable name	Description
AGE1634	Age 16-34 flag (derived)
AGE6674	Age 66-74 flag (derived)
B_Q01A_ISCED11	Education - Highest qualification - Level, ISCED 2011
B_Q01A3_ISCED11	Education - Highest qualification - Level of foreign qualification, ISCED 2011
B_Q02B_ISCED11	Education - Current qualification - Level, ISCED 2011
B_Q03B_ISCED11	Education - Uncompleted qualification - Level, ISCED 2011
B_Q05A_ISCED11	Education - Formal qualification - Level, ISCED 2011
D_Q12A_ISCED11	Current work - Requirements - Education level, ISCED 2011
EMP_6CAT	Employment status (derived, 6 categories)
EMPSTAT	Employment and student status (derived, 5 categories)

See note at end of table.

Table E-2. New variables in the household sample files—Continued

Variable name	Description
INTLFLAG	Flag for international comparison of U.S. combined Main Study and National Supplement
J_Q06B_ISCED11	Background - Mother/female guardian - Highest level of education, ISCED 2011
J_Q07B_ISCED11	Background - Father/male guardian - Highest level of education, ISCED 2011
SAMPFLAG	Flag for U.S. combined Main Study and National Supplement (Household and Prison samples)
UNEMPFLAG	Unemployment flag
URBAN_12CAT	Urbanicity (derived, 12 categories)
URBAN_4CAT	Urbanicity (derived, 4 categories)

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

In addition to the above variables, CNTRYID has been changed in the U.S. national household and prison PUF/RUF files. CNTRYID for the national combined household files is 84091 and labeled as "United states (16-74 y.o. sample)." CNTRYID for the prison files is 84092 and labeled as "United states (prison sample)." Note that in the U.S. international PUF CNTRYID remains as previously identified, 840, with a label "United States."

Some Formulas for Weighted Estimates

The sampling weights can be used to estimate several types of statistics, including means, totals, proportions, ratios of person characteristics, regression coefficients, etc. For instance, for person l, an estimate of the total number of persons in domain d is simply the sum of the weights (W) in domain d:

$$\widehat{N}_d = \sum_{l \in d} W_l \, .$$

To estimate the proportion of persons in domain d, compute the following:

$$\hat{P}_d = \frac{\sum_{l \in d} W_l}{\sum_{l} W_l} .$$

To estimate a weighted mean, compute the following:

$$\hat{\bar{Y}} = \frac{\sum_{l} W_l y_l}{\sum_{l} W_l}$$

where y_i = reported value of survey item y for person l.

To estimate the weighted mean for literacy score from the $M = 10$ plausible values (PVs), use the following formula:

$$\hat{\bar{Y}}^* = \frac{\sum_m \hat{\bar{Y}}_m}{M}$$

where

$$\hat{\bar{Y}}_m = \frac{\sum_l W_l y_{lm}}{\sum_l W_l};\ m = 1, \ldots, M$$

When estimating standard errors associated with PIAAC estimates, such as those computed using the above formulas, special approaches are needed to capture the features of the PIAAC design. These features include effects related to sampling (i.e., stratification, systematic sampling from a sorted list, clustering, and effects of probability proportionate to size sampling), weighting (i.e., nonresponse adjustments, trimming, and raking), and multiple imputation of scores (plausible values). The variance of the estimated mean is computed to account for the imputation error variance component, due to the use of PVs, as follows:

$$v\left(\hat{\bar{Y}}^*\right) = U^* + B\left(1 + \frac{1}{M}\right)$$

where the "within" variance component is computed as the average of the sampling variance for each of the M plausible values, computed as

$$U^* = \left(\sum_{m=1}^{M} U_m\right)/M$$

where the sampling variance of the estimated mean $\hat{\bar{Y}}_m$ for plausible value m is U_m, and where the "between" component is calculated as

$$B = \left[\sum_{m=1}^{M} \left(\hat{\bar{Y}}_m - \hat{\bar{Y}}^*\right)^2\right]/(M-1)$$

The standard error is computed as the square root of the total variance, $\sqrt{v\left(\hat{\bar{Y}}^*\right)}$

Software Tools

There are a variety of specialized tools available to support valid analyses of the U.S. PIAAC data, such as:

- Web-based tools:
  - NCES International Data Explorer (IDE)
  - OECD IDE (household sample with international variables only)

- Statistical software tools:
  - International Database (IDB) Analyzer (used with SPSS)
  - SAS macro
  - Stata macro

Information on how to use the NCES IDE is available in the PIAAC IDE help guide: http://nces.ed.gov/surveys/international/ide/. Descriptions of the OECD IDE and statistical software tools and how to use them with PIAAC data can be found on the OECD website for public data and analysis at http://www.oecd.org/site/piaac/publicdataandanalysis.htm. Both the NCES and OECD IDE allow the user to create statistical tables and charts to explore adults' (16–65 years of age) performance in literacy, numeracy, and problem solving in technology-rich environments. Data can be analyzed by country as well as in relation to demographic characteristics, social and linguistic background, education, use of skills, and employment status. In addition, the NCES IDE allows analyses on the 66- to 74-year-olds and the prison sample. The NCES IDE also has some U.S.-only and prison-only variables that are not available in the OECD IDE.

The IDB Analyzer is an application developed by the IEA Data Processing and Research Center (DPC) to facilitate the analysis of data from IEA's large-scale assessments. It creates SPSS code that can be used with SPSS to conduct statistical analyses, taking into account the complex sample and assessment structures of the databases.

More information on the SAS macro can be found in Denis (2014). Pokropek and Jakubowski (2013) discuss how to use PIAAC data with Stata. Among the statistical analysis tools mentioned above, each can compute descriptive statistics for multiple countries, although only the IDB Analyzer enables the user to conduct statistical hypothesis testing among groups in the population without having to write any programming code. On the OECD website for public data and analysis, users of the U.S. RUF and PUF are encouraged to review the technical descriptions of the analysis tools in Chapter 23 of the *Technical Report for the Survey of Adult Skills* (Carsten, Daniel, and Gonzalez 2013).

PUF Compendia

The PUF compendia are sets of tables that report weighted percentage statistics for the background items in the U.S. national PUFs. The purpose of the compendia is to support PUF users so that they can gain knowledge of the contents of the PUF, and can use the compendia results to be sure that they are performing PUF analyses correctly. The compendia are available at https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2016667.

Using the IDB Analyzer with U.S. National PUF/RUF Data and with International PUF Data from all PIAAC Countries

If any of the national files is used on its own, no preparatory or additional steps are needed for working with it in the IDB Analyzer. Simply open the U.S. national file in the IDB Analyzer Analysis module and proceed as prompted.

Two types of combinations are possible with the currently released files: (1) U.S. international PUF with international PUFs for other countries; (2) U.S. national household PUF/RUF with U.S. national prison PUF/RUF.

For combining the U.S. international PUF with the PUFs for other countries, access the U.S. international PUF from the NCES website and files for other PIAAC participating countries form the OECD website (the NCES website contains a link to location of files on the OECD website). Note that the international PUF for the U.S. contains only international variables for the sample of 16- to 65-year-olds from the combined household sample, which is comparable to the other PIAAC participating countries. The combined U.S. national PUF contains additional U.S.-only variables as well as the sample of 66- to 74-year-olds. Thus, the U.S. PUF is not comparable to data from other PIAAC countries.

The steps to combine U.S. international PUF with other countries' PUFs are as follows:

1. Download the international PUF SPSS file for the U.S. from the NCES PIAAC website, and download the international PUF SPSS files for all the other countries from the OECD PIAAC website. Note all of the files need to be downloaded into the same directory.

2. Follow the steps in the IDB Analyzer Merge module to merge all international PUF SPSS files into one file.

3. Save the merged file as "prgALLms_PUF.sav." There should be 152,514 records (22 countries) and 1,329 variables in the file. You can then proceed to use the IDB Analyzer Analysis module to conduct analysis on the merged file.

For combining the U.S. national prison PUF/RUF with household PUF/RUF, it is necessary to first combine the two files outside of the IDB Analyzer, in SPSS. In order to do that, follow the steps below:

1. In SPSS open either the combined household or prison U.S. national files. In subsequent steps we assume prison file is opened first.

2. Use the SPSS menu options to add cases from the combined household U.S. national file: DATA -> Merge Files -> Add Cases. You will have a number of unpaired variables. These are prison-only variables and one could add them to have a complete set of both prison and household variables, knowing that the data will be missing for prison-only variables for household cases. Note it is not necessary to create a flag variable to indicate which file a case is from, since the variables CNTRYID and SAMPFLAG already have different values for the household file and prison file.

3. Save the merged file. There will be 9,989 observations in the file. The number of variables will depend on whether PUFs or RUFs are merged, as well as whether the prison-only variables are added to the list of household-/prison-comparable variables in the step 2 above. You can then proceed to use the IDB Analyzer Analysis module to conduct an analysis on the merged file. The IDB Analyzer will use CNTRYID as the default grouping variable for analysis, thus, as recommended, computing results for household sample and prison sample separately.

IDB Analyzer Examples

Here are two IDB Analyzer examples that use the international merged file created in the previous section. It is important to note that the IDB Analyzer Analysis module performs all analyses in accordance with the technical guidelines for analyzing PIAAC data that are provided in the PIAAC Technical Report. The OECD PIAAC Technical Report is available from http://www.oecd.org/site/piaac/publications.htm. The IDB Analyzer Help Manual is available by pressing F1 within the application, or clicking the help button on the interface.

Using as input the international U.S. and other countries' merged SPSS data file (prgALLms_PUF.sav), the following IDB Analyzer instructions produce specified cross-country comparisons. If interested in conducting *t* tests, Appendix C of the IDB Analyzer Help Manual provides instructions.

Example 1: Produce an Excel table that contains cross-country results by gender (as defined by the variable GENDER_R) for literacy scale means, percentages, and standard errors.

1. Open the IDB Analyzer and choose "Analysis Module";
2. Select the "prgALLms_PUF.sav" data file from the designated file path. If the file was not created using the IDB Analyzer a dialog box will pop up asking for the study type, choose "PIAAC";
3. Under "Analysis Type," choose "PIAAC (using final full sample weight)." Under "Statistic Type," choose "Percentages and Means." Under the "Plausible Value Option," choose "Use PVs." Change "Number of Decimals" to your preference;
4. Select the variable "GENDER_R," and move it under the "Grouping Variables";
5. Click on the "Plausible Values" section. Three types of scores will become available in the "Select Variables" window. Choose the Literacy Scale score;
6. Click "Define," choose a folder to store the output and a file name;
7. Click "Start SPSS," and run the SPSS code generated by the IDB Analyzer.

Example 2: Produce an Excel table that contains cross-country percentages, along with their respective standard errors, for literacy scale levels by gender (using the variables GENDER_R and PVLIT1-PVLIT10) for those respondents who were administered the literacy assessment (i.e., only those cases that have plausible values).

1. Open the IDB Analyzer and choose "Analysis Module";
2. Select the "prgALLms_PUF.sav" data file from designated file path. If the file was not created using the IDB Analyzer a dialog box will pop up asking for study type, choose "PIAAC";
3. Under "Analysis Type," choose "PIAAC (using final full sample weight)." Under "Statistic Type," choose "Benchmarks." Under "Benchmark Option," choose "Discrete." Change "Number of Decimals" to your preference;
4. Select the variable "GENDER_R," and move it under the "Grouping Variables";
5. Click on the "Plausible Values" section. Three types of scores will become available in the "Select Variables" window, choose the Literacy Scale score. If you want to include cases

without PVs in your report, you will need to check "Report cases with no plausible values" above the Plausible Values box;

6. Enter the cut points under "Achievement Benchmarks," separated by spaces in the ascending order (for example, for Literacy you would enter: 176 226 276 326 376 for level 1 through level 5). The cut points for the levels are in the appendix of the Help Manual of the IDB Analyzer, accessible by pressing the F1 key;

7. Click "Define," choose a folder to store the output and a file name;

8. Click "Start SPSS," run the SPSS code generated by the IDB Analyzer.

Output from the Two Examples Produced by the IDB Analyzer

Tables E-3 and E-4 provide selected summary statistics results for the two examples presented above, respectively. Due to the size of the spreadsheet that is produced, only results for Literacy Level 1 (from 176 to Below 226) were extracted from the Example 2 spreadsheet.

Table E-3. Example 1: Means and percentages overall for PIAAC Literacy Scale, by country and gender

CNTRYID	GENDER_R	*DVAR*	*N*	*SPFWT0*	*SUMW_SE*	*PCT*	*PCT_SE*	*MNPV*	*MNPV_SE*	*SDPV*	*SDPV_SE*
Austria	Male	PVLIT	2,479	2,764,088.46	7,207.03	49.86	0.10	271.53	1.04	44.64	0.86
Austria	Female	PVLIT	2,546	2,780,051.37	6,922.30	50.14	0.10	267.39	0.93	43.17	0.75
Belgium	Male	PVLIT	2,467	1,984,965.42	6,758.70	50.58	0.13	278.09	0.97	47.91	0.86
Belgium	Female	PVLIT	2,517	1,939,810.58	7,650.74	49.42	0.13	272.81	1.08	46.07	0.89
Canada	Male	PVLIT	12,442	11,684,548.30	0.00	49.97	0.00	274.49	0.86	50.99	0.75
Canada	Female	PVLIT	14,241	11,696,518.54	0.00	50.03	0.00	272.19	0.78	49.84	0.64
Czech Republic	Male	PVLIT	2,756	3,706,815.34	9,665.44	50.44	0.10	275.68	1.26	40.83	1.06
Czech Republic	Female	PVLIT	3,325	3,642,379.54	10,916.46	49.56	0.10	272.32	1.30	40.67	1.10
Denmark	Male	PVLIT	3,590	1,819,081.82	2,144.36	50.32	0.04	270.58	1.03	49.70	0.81
Denmark	Female	PVLIT	3,696	1,796,079.90	1,167.73	49.68	0.04	271.00	0.80	45.63	0.86
Estonia	Male	PVLIT	3,432	427,035.49	419.81	47.83	0.03	275.06	1.09	45.44	0.68
Estonia	Female	PVLIT	4,154	465,698.67	339.19	52.17	0.03	276.64	0.81	43.42	0.63
Finland	Male	PVLIT	2,757	1,758,421.71	1,147.43	50.29	0.03	285.96	1.21	51.99	1.10
Finland	Female	PVLIT	2,707	1,738,487.29	1,147.43	49.71	0.03	289.15	0.99	49.25	1.19
France	Male	PVLIT	3,382	19,392,790.91	78,456.80	48.83	0.19	262.05	0.87	49.32	0.58
France	Female	PVLIT	3,525	20,318,410.94	77,142.10	51.17	0.19	262.23	0.69	48.73	0.64
Germany	Male	PVLIT	2,641	26,701,427.44	76,883.02	50.51	0.12	272.35	1.17	47.69	0.84
Germany	Female	PVLIT	2,738	26,163,761.05	79,086.02	49.49	0.12	267.21	1.19	46.96	0.87
Ireland	Male	PVLIT	2,728	1,458,047.52	4,775.05	48.92	0.14	267.71	1.17	49.17	1.05
Ireland	Female	PVLIT	3,235	1,522,255.80	4,108.07	51.08	0.14	265.43	1.10	45.19	1.07
Italy	Male	PVLIT	2,220	19,557,641.99	45,174.50	50.00	0.05	250.36	1.50	46.41	1.08
Italy	Female	PVLIT	2,369	19,555,182.14	31,803.85	50.00	0.05	250.61	1.32	42.90	0.87
Japan	Male	PVLIT	2,468	40,246,126.33	74,110.55	50.27	0.06	297.78	0.88	40.35	0.78
Japan	Female	PVLIT	2,705	39,811,288.52	57,228.06	49.73	0.06	294.69	1.01	38.99	0.80
Korea	Male	PVLIT	3,092	17,166,875.15	20,962.89	49.75	0.04	275.72	0.75	41.57	0.76
Korea	Female	PVLIT	3,559	17,342,502.71	16,397.74	50.25	0.04	269.43	0.87	41.56	0.65
Netherlands	Male	PVLIT	2,501	5,475,351.56	17,524.78	50.20	0.13	287.06	1.08	48.73	0.95
Netherlands	Female	PVLIT	2,581	5,431,153.00	16,173.16	49.80	0.13	280.92	0.94	47.85	0.88
Norway	Male	PVLIT	2,557	1,639,476.09	3,792.47	51.09	0.08	280.34	0.97	47.76	0.80

See note at end of table.

Table E-3. Example 1: Means and percentages overall for PIAAC Literacy Scale, by country and gender—Continued

CNTRYID	GENDER_R	*DVAR*	*N*	*SPFWT0*	*SUMW_SE*	*PCT*	*PCT_SE*	*MNPV*	*MNPV_SE*	*SDPV*	*SDPV_SE*
Norway	Female	PVLIT	2,390	1,569,515.99	3,610.70	48.91	0.08	276.43	0.91	46.14	0.97
Poland	Male	PVLIT	4,733	13,229,030.39	3,849.71	49.47	0.01	263.66	0.97	49.62	0.90
Poland	Female	PVLIT	4,633	13,512,956.61	3,849.71	50.53	0.01	270.08	0.86	46.09	0.88
Russian Federation	Male	PVLIT	1,344	41,758,488.41	79,572.18	47.77	0.09	272.90	2.98	44.45	1.63
Russian Federation	Female	PVLIT	2,547	45,654,647.98	79,454.22	52.23	0.09	277.37	2.88	41.26	1.76
Slovak Republic	Male	PVLIT	2,697	1,929,712.55	1,800.22	49.99	0.03	273.47	0.86	40.58	0.73
Slovak Republic	Female	PVLIT	3,005	1,930,792.85	1,720.19	50.01	0.03	274.22	0.82	39.55	0.81
Spain	Male	PVLIT	2,929	15,500,659.30	24,189.68	50.24	0.05	254.11	1.00	49.94	0.80
Spain	Female	PVLIT	3,042	15,353,142.28	24,822.11	49.76	0.05	249.45	1.04	47.98	0.81
Sweden	Male	PVLIT	2,253	3,036,908.24	4,280.14	50.73	0.07	280.88	1.08	49.97	0.95
Sweden	Female	PVLIT	2,216	2,949,015.04	4,280.14	49.27	0.07	277.54	1.10	51.11	1.19
United Kingdom	Male	PVLIT	3,693	17,398,114.39	42,304.58	49.82	0.08	273.90	1.37	50.08	1.09
United Kingdom	Female	PVLIT	5,113	17,526,525.88	37,135.73	50.18	0.08	271.03	1.29	47.79	1.04
United States	Male	PVLIT	3,566	96,800,620.62	635,462.93	48.91	0.19	272.26	1.14	51.03	0.95
United States	Female	PVLIT	4,193	101,120,189.15	625,092.24	51.09	0.19	271.21	1.16	49.61	1.08

SOURCE: Organization for Economic Cooperation and Development, Survey of Adult Skills (PIAAC) 2012; U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Table E-4. Example 2: Percentages for PIAAC Literacy Scale levels, by country and gender

CNTRYID	GENDER_R	*DVAR*	*CUTVAR*	*N*	*SPFWT0*	*SUMW_SE*	*PCT*	*PCT_SE*
Austria	Male	PVLIT	2.From 176 to Below 226	280	343,832	25,670.46	12.44	0.93
Austria	Female	PVLIT	2.From 176 to Below 226	309	381,141	24,949.74	13.71	0.90
Belgium	Male	PVLIT	2.From 176 to Below 226	268	215,033	17,150.06	10.83	0.86
Belgium	Female	PVLIT	2.From 176 to Below 226	321	251,672	16,109.27	12.97	0.83
Canada	Male	PVLIT	2.From 176 to Below 226	1,876	1,461,361	84,999.64	12.51	0.73
Canada	Female	PVLIT	2.From 176 to Below 226	2,076	1,514,212	60,136.94	12.95	0.51
Czech Republic	Male	PVLIT	2.From 176 to Below 226	261	372,413	45,079.77	10.05	1.22
Czech Republic	Female	PVLIT	2.From 176 to Below 226	331	386,481	41,001.52	10.61	1.13
Denmark	Male	PVLIT	2.From 176 to Below 226	481	233,769	13,778.76	12.85	0.76
Denmark	Female	PVLIT	2.From 176 to Below 226	469	197,711	14,265.85	11.01	0.79
Estonia	Male	PVLIT	2.From 176 to Below 226	389	48,079	3,109.28	11.26	0.73
Estonia	Female	PVLIT	2.From 176 to Below 226	446	50,543	2,835.94	10.85	0.61
Finland	Male	PVLIT	2.From 176 to Below 226	222	149,478	12,763.68	8.50	0.73
Finland	Female	PVLIT	2.From 176 to Below 226	189	128,607	11,481.55	7.40	0.66
France	Male	PVLIT	2.From 176 to Below 226	556	3,341,452	138,577.64	17.23	0.72
France	Female	PVLIT	2.From 176 to Below 226	517	3,158,453	148,385.55	15.54	0.73
Germany	Male	PVLIT	2.From 176 to Below 226	323	3,656,914	258,522.80	13.70	0.97
Germany	Female	PVLIT	2.From 176 to Below 226	369	3,983,781	256,581.91	15.23	0.98
Ireland	Male	PVLIT	2.From 176 to Below 226	343	191,149	14,789.91	13.11	1.01
Ireland	Female	PVLIT	2.From 176 to Below 226	426	202,840	15,362.05	13.32	1.01
Italy	Male	PVLIT	2.From 176 to Below 226	461	4,530,923	280,099.69	23.17	1.43
Italy	Female	PVLIT	2.From 176 to Below 226	463	4,193,733	253,744.19	21.45	1.30
Japan	Male	PVLIT	2.From 176 to Below 226	105	1,817,447	208,784.72	4.52	0.52
Japan	Female	PVLIT	2.From 176 to Below 226	105	1,671,113	229,480.67	4.20	0.58
Korea	Male	PVLIT	2.From 176 to Below 226	302	1,631,208	111,979.26	9.50	0.65
Korea	Female	PVLIT	2.From 176 to Below 226	433	2,051,018	132,417.31	11.83	0.76
Netherlands	Male	PVLIT	2.From 176 to Below 226	199	465,750	38,840.74	8.51	0.71
Netherlands	Female	PVLIT	2.From 176 to Below 226	240	551,576	40,729.18	10.16	0.75
Norway	Male	PVLIT	2.From 176 to Below 226	213	153,724	12,506.19	9.38	0.76

See note at end of table.

Table E-4. Example 2: Percentages for PIAAC Literacy Scale levels, by country and gender—Continued

CNTRYID	GENDER_R	*DVAR*	*CUTVAR*	*N*	*SPFWT0*	*SUMW_SE*	*PCT*	*PCT_SE*
Norway	Female	PVLIT	2.From 176 to Below 226	203	150,812	13,382.21	9.61	0.85
Poland	Male	PVLIT	2.From 176 to Below 226	632	2,146,269	139,588.48	16.22	1.06
Poland	Female	PVLIT	2.From 176 to Below 226	488	1,819,016	111,957.46	13.46	0.83
Russian Federation	Male	PVLIT	2.From 176 to Below 226	140	5,570,977	667,579.54	13.34	1.60
Russian Federation	Female	PVLIT	2.From 176 to Below 226	226	4,466,872	655,180.36	9.78	1.43
Slovak Republic	Male	PVLIT	2.From 176 to Below 226	277	192,759	14,222.78	9.99	0.74
Slovak Republic	Female	PVLIT	2.From 176 to Below 226	299	184,352	14,595.88	9.55	0.76
Spain	Male	PVLIT	2.From 176 to Below 226	575	3,011,395	152,188.42	19.43	0.98
Spain	Female	PVLIT	2.From 176 to Below 226	650	3,294,139	173,245.70	21.46	1.13
Sweden	Male	PVLIT	2.From 176 to Below 226	181	284,544	25,927.75	9.37	0.85
Sweden	Female	PVLIT	2.From 176 to Below 226	179	289,186	23,217.92	9.81	0.79
United Kingdom	Male	PVLIT	2.From 176 to Below 226	521	2,371,682	185,807.47	13.63	1.06
United Kingdom	Female	PVLIT	2.From 176 to Below 226	710	2,280,218	155,513.94	13.01	0.89
United States	Male	PVLIT	2.From 176 to Below 226	569	13,491,223	968,710.87	13.94	0.99
United States	Female	PVLIT	2.From 176 to Below 226	605	13,001,407	749,238.75	12.86	0.73

SOURCE: Organization for Economic Cooperation and Development, Survey of Adult Skills (PIAAC) 2012; U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Tables E-5 and E-6 present side-by-side lists of U.S. public-use file and restricted-use file variables for the household sample and for the prison sample, respectively.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
A_D01a1		x	General - Interview month (DERIVED BY CAPI)
A_D01a2		x	General - Year before interview (DERIVED BY CAPI)
A_D01a3		x	General - Interview year (DERIVED BY CAPI)
A_N01		x	General - Gender of respondent
A_N01_T	x	x	Gender (Trend-IALS/ALL)
A_Q01a		x	General - Year of birth
A_Q01b		x	General - Month of birth
ACTIVE_SECTION		x	Active section (final state on export)
AETPOP	x	x	Adult education/training population (AET) – excludes youths 16-24 in initial cycle of studies(derived)
AGE_R		x	Person resolved age from BQ and QC check (derived)
AGE_R_ORG		x	Person resolved age from BQ and QC check (derived, original before trimming)
AGE1634	x	x	Age 16-34 flag (derived)
AGE6674	x	x	Age 66-74 flag (derived)
AGEG10LFS_T	x	x	Age in 10 year bands (Trend-IALS/ALL)
AGEG10LFSEXT	x	x	Age in 10 year bands extended to include ages over 65 (derived)
AGEG5LFSEXT	x	x	Age in 5 year bands extended to include ages over 65 (derived)
B_D01d		x	Education - Highest qualification - Months elapsed since finished (DERIVED BY CAPI)
B_D01d_C	x	x	Education - Time elapsed since finished highest qualification (categorised, 5 categories)
B_D03d		x	Education - Uncompleted qualification - Months elapsed since dropout (DERIVED BY CAPI)
B_D03d_C	x	x	Education - Derived Months elapsed since leaving education without completing program (categorized, 3 categories)
B_D03d_USR		x	Education - Uncompleted qualification - Months elapsed since dropout (DERIVED BY CAPI), corrected for U.S. routing
B_D12h	x	x	Activities - Last year - Number of learning activities (DERIVED BY CAPI)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
B_D12h_USR		x	Activities - Last year - Number of learning activities (DERIVED BY CAPI), corrected for U.S. routing
B_Q01a	x	x	Education - Highest qualification - Level
B_Q01A_C	x	x	Education - Highest qualification - Level (3 categories) (derived from B_Q01A)
B_Q01A_ISCED11	x	x	Education - Highest qualification - Level, ISCED 2011
B_Q01a_T	x	x	Highest level of schooling (Trend-IALS/ALL)
B_Q01a2US		x	Education - Highest qualification - Country of foreign qualification
B_Q01a3		x	Education - Highest qualification - Level of foreign qualification
B_Q01a3_C		x	Education - Highest Qualification - Level of foreign qualification (collapsed, 14 categories)
B_Q01A3_ISCED11		x	Education - Highest qualification - Level of foreign qualification, ISCED 2011
B_Q01a3US		x	Education - Highest qualification - Level of foreign qualification
B_Q01aUS		x	Education - Highest qualification - Level
B_Q01AUS_C	x	x	Education - Highest qualification - Level (3 categories) (derived from B_Q01AUS)
B_Q01b	x	x	Education - Highest qualification - Area of study
B_Q01bUSX		x	Education - Highest qualification - Area of study verbatim
B_Q01c1		x	Education - Highest qualification - Age of finish
B_Q01c1_C	x	x	Education - Highest qualification - Age of finish (categorised, 6 categories)
B_Q01c1_T		x	Age at completion of highest level of schooling (Trend-IALS/ALL)
B_Q01c2		x	Education - Highest qualification - Year of finish
B_Q01d		x	Education - Highest qualification - Month of finish
B_Q02a	x	x	Education - Current qualification
B_Q02a_T1	x	x	Education or training in last 12 months (Trend-IALS/ALL)
B_Q02a_T2	x	x	Courses toward certificate, diploma, or degree in program of studies in last 12 months (Trend-IALS/ALL)
B_Q02b		x	Education - Current qualification - Level

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
B_Q02b_C		x	Education - Current Qualification (collapsed, 10 categories)
B_Q02B_ISCED11		x	Education - Current qualification - Level, ISCED 2011
B_Q02bUS		x	Education - Current qualification - Level
B_Q02bUS_C	x	x	Education - Current qualification - Level (6 categories) (derived from B_Q02b)
B_Q02c		x	Education - Current qualification - Area of study
B_Q02cUS_C	x	x	Education - Current qualification - Area of study (8 categories) (derived from B_Q02c)
B_Q02cUSX		x	Education - Current Qualification - Area of Study Verbatim
B_Q03a	x	x	Education - Uncompleted qualification
B_Q03a_USR		x	Education - Uncompleted qualification, corrected for U.S. routing
B_Q03b		x	Education - Uncompleted qualification - Level
B_Q03b_C		x	Education - Uncompleted qualification - Level (collapsed, 10 categories)
B_Q03B_ISCED11		x	Education - Uncompleted qualification - Level, ISCED 2011
B_Q03b_USR		x	Education - Uncompleted qualification - Level, corrected for U.S. routing
B_Q03bUS		x	Education - Uncompleted qualification - Level
B_Q03bUS_C	x	x	Education - Uncompleted qualification - Level (6 categories) (derived from B_Q03b)
B_Q03c1		x	Education - Uncompleted qualification - Age of dropout
B_Q03c1_C	x	x	Education - Uncompleted qualification - Age of dropout (categorised, 6 categories)
B_Q03c1_USR		x	Education - Uncompleted qualification - Age of dropout, corrected for U.S. routing
B_Q03c2		x	Education - Uncompleted qualification - Year of dropout
B_Q03c2_USR		x	Education - Uncompleted qualification - Year of dropout, corrected for U.S. routing
B_Q03d		x	Education - Uncompleted qualification - Month of dropout
B_Q03d_USR		x	Education - Uncompleted qualification - Month of dropout, corrected for U.S. routing
B_Q04a	x	x	Education - Formal qualification
B_Q04a_USR		x	Education - Formal qualification, corrected for U.S. routing

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
B_Q04b		x	Education - Formal qualification - Count
B_Q04b_C	x	x	Education - Formal qualification - Count (top-coded at 2)
B_Q04b_USR		x	Education - Formal qualification - Count, corrected for U.S. routing
B_Q05a		x	Education - Formal qualification - Level
B_Q05A_ISCED11		x	Education - Formal qualification - Level, ISCED 2011
B_Q05a_USR		x	Education - Formal qualification - Level, corrected for U.S. routing
B_Q05aUS		x	Education - Formal qualification - Level
B_Q05b		x	Education - Formal qualification - Area of study
B_Q05b_USR		x	Education - Formal qualification - Area of study, corrected for U.S. routing
B_Q05bUSX		x	Education - Formal Qualification - Area of Study Verbatim
B_Q05c	x	x	Education - Formal qualification - Reason job related
B_Q05c_T	x	x	Main reason for program of studies (Trend-IALS/ALL)
B_Q05c_USR		x	Education - Formal qualification - Reason job related, corrected for U.S. routing
B_Q05cUSX1	x	x	Education - Formal qualification - Degree personal interest
B_Q05cUSX2	x	x	Education - Formal qualification - Degree personal interest or work related
B_Q10a	x	x	Education - Formal qualification - Employed
B_Q10a_USR		x	Education - Formal qualification - Employed, corrected for U.S. routing
B_Q10b	x	x	Education - Formal qualification - Employed - Working hours
B_Q10b_USR		x	Education - Formal qualification - Employed - Working hours, corrected for U.S. routing
B_Q10c	x	x	Education - Formal qualification - Employed - Useful for job
B_Q10c_USR		x	Education - Formal qualification - Employed - Useful for job, corrected for U.S. routing
B_Q11	x	x	Education - Formal qualification - Grant from employer
B_Q11_USR		x	Education - Formal qualification - Grant from employer, corrected for U.S. routing

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
B_Q12a	x	x	Activities - Last year - Open or distance education
B_Q12a_T	x	x	Courses outside of program of studies in last 12 months (Trend-IALS/ALL)
B_Q12a_USR		x	Activities - Last year - Open or distance education, corrected for U.S. routing
B_Q12b		x	Activities - Last year - Open or distance education - Count
B_Q12b_USR		x	Activities - Last year - Open or distance education - Count, corrected for U.S. routing
B_Q12bUS_C	x	x	Activities - Last year - Open or distance edu - How many (topcode at 7) (derived from B_Q12b)
B_Q12c	x	x	Activities - Last year - On the job training
B_Q12c_USR		x	Activities - Last year - On the job training, corrected for U.S. routing
B_Q12d		x	Activities - Last year - On the job training - Count
B_Q12d_C	x	x	Activities - Last year - On the job training - Count (top-coded at 5)
B_Q12d_USR		x	Activities - Last year - On the job training - Count, corrected for U.S. routing
B_Q12e	x	x	Activities - Last year - Seminars or workshops
B_Q12e_USR		x	Activities - Last year - Seminars or workshops, corrected for U.S. routing
B_Q12f		x	Activities - Last year - Seminars or workshops - Count
B_Q12f_C	x	x	Activities - Last year - Seminars or workshops - Count (top-coded at 5)
B_Q12f_USR		x	Activities - Last year - Seminars or workshops - Count, corrected for U.S. routing
B_Q12g	x	x	Activities - Last year - Private lessons
B_Q12g_USR		x	Activities - Last year - Private lessons, corrected for U.S. routing
B_Q12h		x	Activities - Last year - Private lessons - Count
B_Q12h_C	x	x	Activities - Last year - Private lessons - Count (top-coded at 5)
B_Q12h_USR		x	Activities - Last year - Private lessons - Count, corrected for U.S. routing
B_Q13	x	x	Activities - Last year - Activity specified
B_Q13_USR		x	Activities - Last year - Activity specified, corrected for U.S. routing
B_Q14a	x	x	Activities - Last year - Job related

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
B_Q14a_USR		x	Activities - Last year - Job related, corrected for U.S. routing
B_Q14b	x	x	Activities - Last year - Reason for participating
B_Q14b_USR		x	Activities - Last year - Reason for participating, corrected for U.S. routing
B_Q14bUSX1	x	x	Activities - Last year - Activity Participation for personal interest
B_Q14bUSX2	x	x	Activities - Last year - Activity Participation mainly personal reasons
B_Q15a	x	x	Activities - Last year - Employed
B_Q15a_USR		x	Activities - Last year - Employed, corrected for U.S. routing
B_Q15b	x	x	Activities - Last year - During working hours
B_Q15b_USR		x	Activities - Last year - During working hours, corrected for U.S. routing
B_Q15c	x	x	Activities - Last year - Useful for job
B_Q15c_USR		x	Activities - Last year - Useful for job, corrected for U.S. routing
B_Q16	x	x	Activities - Last year - Grant from employer
B_Q16_USR		x	Activities - Last year - Grant from employer, corrected for U.S. routing
B_Q17	x	x	Activities - Last year - Time spend - Unit
B_Q17_USR		x	Activities - Last year - Time spend - Unit, corrected for U.S. routing
B_Q18a	x	x	Activities - Last year - Time spend for activities - Weeks
B_Q18a_USR		x	Activities - Last year - Time spend for activities - Weeks, corrected for U.S. routing
B_Q19a	x	x	Activities - Last year - Time spend for activities - Days
B_Q19a_USR		x	Activities - Last year - Time spend for activities - Days, corrected for U.S. routing
B_Q20a	x	x	Activities - Last year - Time spend for activities - Hours
B_Q20a_USR		x	Activities - Last year - Time spend for activities - Hours, corrected for U.S. routing
B_Q20b	x	x	Activities - Last year - Time spend for activities - Proportion of job-related hours
B_Q20b_USR		x	Activities - Last year - Time spend for activities - Proportion of job-related hours, corrected for U.S. routing
B_Q26a	x	x	Activities - Last year - Wanted but didn't start

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
B_Q26a_T	x	x	Training or education for career or job wanted but not taken in last 12 months (Trend-IALS/ALL)
B_Q26a_USR		x	Activities - Last year - Wanted but didn't start, corrected for U.S. routing
B_Q26b	x	x	Activities - Last year - Wanted but didn't start – Reason
B_Q26b_USR		x	Activities - Last year - Wanted but didn't start - Reason, corrected for U.S. routing
B_Q27aUSX	x	x	Activities - Class - Class/tutor basic skills
B_Q27bUSX	x	x	Activities - Class - Class/tutor GED
B_Q27cUSX	x	x	Activities - Class - Class/tutor other equivalency
B_Q27dUSX	x	x	Activities - Class - Class/tutor main reason
B_Q27eUSXa		x	Activities - Class - Class attendance, amount
B_Q27eUSXb		x	Activities - Class - Class attendance, unit
B_Q29aUSX	x	x	Activities - Apprentice - Was apprentice
B_Q29bUSX		x	Activities - Apprentice - Which trade or craft
B_S01a1		x	Education - Highest qualification - Name of foreign qualification
B_S01a2		x	Education - Highest qualification - Country of foreign qualification (other)
B_S27eUSX		x	Activities - Class - Class attendance, other specify
BIRTHRGN		x	Country of birth (9 regions - derived)
BIRTHRGNUS_C	x	x	Country of birth (9 regions) (3 categories) (derived from BIRTHRGN)
BORNLANG	x	x	Interactions between place of birth and language status (derived)
BQLANG		x	Language for background questionnaire
C_D04	x	x	Current status/work history - Last month - Active steps to find job (DERIVED BY CAPI)
C_D05	x	x	Current status/work history - Employment status (DERIVED BY CAPI)
C_D06	x	x	Current status/work history - Current - Paid job or family business (DERIVED BY CAPI)
C_D08c	x	x	Current status/work history - Left work in past 5 years (DERIVED BY CAPI)
C_D09	x	x	Current status/work history - Work experience (DERIVED BY CAPI)
C_D09_T	x	x	Worked at job or business in last 12 months (any number of hours) (Trend-IALS/ALL)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C_Q01a	x	x	Current status/work history - Last week - Paid work
C_Q01b	x	x	Current status/work history - Last week - Away from job but will return
C_Q01c	x	x	Current status/work history - Last week - Unpaid work for own business
C_Q02a	x	x	Current status/work history - Last month - Looking for paid work
C_Q02b	x	x	Current status/work history - Last month - Waiting to start job
C_Q02c		x	Current status/work history - Last month - Waiting to start job - Next 3 months
C_Q03_01	x	x	Current status/work history - Last month - Reason not looking for work - Waiting for result of application
C_Q03_02	x	x	Current status/work history - Last month - Reason not looking for work - Being a student
C_Q03_03	x	x	Current status/work history - Last month - Reason not looking for work - Looking after the family
C_Q03_04	x	x	Current status/work history - Last month - Reason not looking for work - Temp sick
C_Q03_05	x	x	Current status/work history - Last month - Reason not looking for work - Long-term illness
C_Q03_06	x	x	Current status/work history - Last month - Reason not looking for work - Nothing available
C_Q03_07	x	x	Current status/work history - Last month - Reason not looking for work - Did not get around
C_Q03_08	x	x	Current status/work history - Last month - Reason not looking for work - No need
C_Q03_09	x	x	Current status/work history - Last month - Reason not looking for work - Retired
C_Q03_10	x	x	Current status/work history - Last month - Reason not looking for work - Other
C_Q04a	x	x	Current status/work history - Last month - Ways of looking for work - Contact public employment
C_Q04b	x	x	Current status/work history - Last month - Ways of looking for work - Contact private agency

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C_Q04c	x	x	Current status/work history - Last month - Ways of looking for work - Apply to employers
C_Q04d	x	x	Current status/work history - Last month - Ways of looking for work - Ask family/friends
C_Q04e	x	x	Current status/work history - Last month - Ways of looking for work - Place/answer advertisements
C_Q04f	x	x	Current status/work history - Last month - Ways of looking for work - Study advertisements
C_Q04g	x	x	Current status/work history - Last month - Ways of looking for work - Recruitment test
C_Q04h	x	x	Current status/work history - Last month - Ways of looking for work - Premises
C_Q04i		x	Current status/work history - Last month - Ways of looking for work - Licenses/financial resources
C_Q04j	x	x	Current status/work history - Last month - Ways of looking for work - Other
C_Q05		x	Current status/work history - Ability to start job within 2 weeks
C_Q06	x	x	Current status/work history - Last week - Number of jobs
C_Q07	x	x	Current status/work history - Subjective status
C_Q07_T	x	x	Current work situation (Trend-IALS/ALL)
C_Q08a	x	x	Current status/work history - Ever paid work
C_Q08b	x	x	Current status/work history - Last year - Paid work
C_Q08c1		x	Current status/work history - Age when stopped working in last job
C_Q08c1_C	x	x	Current status/work history - Age when stopped working in last job (categorised, 10 categories)
C_Q08c2		x	Current status/work history - Year when stopped working in last job
C_Q08C2US_C	x	x	Current status/work history - Year when stopped working in last job (4 categories) (derived from C_Q08C2)
C_Q09		x	Current status/work history - Years of paid work during lifetime
C_Q09_C	x	x	Current status/work history - Years of paid work during lifetime (top-coded at 47)
C_Q10a		x	Current status/work history - Last 5 years - How many different firms or organisations

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C_Q10a_C	x	x	Current status/work history - Last 5 years - How many different firms or organisations (top-coded at 7)
C_S03		x	Current status/work history - Last month - Months looking for paid work
C_S03US_C	x	x	Current status/work history - Last month - Months looking for paid work (topcode at 13) (derived from C_S03)
C_S04j		x	Current status/work history - Last month - Ways of looking for work - Other specified
C300C02A	x	x	CLC / 300 - Employment Advertisement (Number of Actions)
C300C02F	x	x	CLC / 300 - Employment Advertisement (Timing First Action)
C300C02S	x	x	CLC / 300 - Employment Advertisement (Scored Response)
C300C02T	x	x	CLC / 300 - Employment Advertisement (Timing)
C301C05A	x	x	CLC / 301 - SGIH (Number of Actions)
C301C05F	x	x	CLC / 301 - SGIH (Timing First Action)
C301C05S	x	x	CLC / 301 - SGIH (Scored Response)
C301C05T	x	x	CLC / 301 - SGIH (Timing)
C305215A	x	x	CL / 305 - TMN AntiTheft (Number of Actions)
C305215F	x	x	CL / 305 - TMN AntiTheft (Timing First Action)
C305215S	x	x	CL / 305 - TMN AntiTheft (Scored Response)
C305215T	x	x	CL / 305 - TMN AntiTheft (Timing)
C305218A	x	x	CL / 305 - TMN AntiTheft (Number of Actions)
C305218F	x	x	CL / 305 - TMN AntiTheft (Timing First Action)
C305218S	x	x	CL / 305 - TMN AntiTheft (Scored Response)
C305218T	x	x	CL / 305 - TMN AntiTheft (Timing)
C308116A	x	x	CL / 308 - Baltic Stock Market (Number of Actions)
C308116F	x	x	CL / 308 - Baltic Stock Market (Timing First Action)
C308116S	x	x	CL / 308 - Baltic Stock Market (Scored Response)
C308116T	x	x	CL / 308 - Baltic Stock Market (Timing)
C308117A	x	x	CL / 308 - Baltic Stock Market (Number of Actions)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C308117F	x	x	CL / 308 - Baltic Stock Market (Timing First Action)
C308117S	x	x	CL / 308 - Baltic Stock Market (Scored Response)
C308117T	x	x	CL / 308 - Baltic Stock Market (Timing)
C308118A	x	x	CL / 308 - Baltic Stock Market (Number of Actions)
C308118F	x	x	CL / 308 - Baltic Stock Market (Timing First Action)
C308118S	x	x	CL / 308 - Baltic Stock Market (Scored Response)
C308118T	x	x	CL / 308 - Baltic Stock Market (Timing)
C308119A	x	x	CL / 308 - Baltic Stock Market (Number of Actions)
C308119F	x	x	CL / 308 - Baltic Stock Market (Timing First Action)
C308119S	x	x	CL / 308 - Baltic Stock Market (Scored Response)
C308119T	x	x	CL / 308 - Baltic Stock Market (Timing)
C308120A	x	x	CL / 308 - Baltic Stock Market (Number of Actions)
C308120F	x	x	CL / 308 - Baltic Stock Market (Timing First Action)
C308120S	x	x	CL / 308 - Baltic Stock Market (Scored Response)
C308120T	x	x	CL / 308 - Baltic Stock Market (Timing)
C308121A	x	x	CL / 308 - Baltic Stock Market (Number of Actions)
C308121F	x	x	CL / 308 - Baltic Stock Market (Timing First Action)
C308121S	x	x	CL / 308 - Baltic Stock Market (Scored Response)
C308121T	x	x	CL / 308 - Baltic Stock Market (Timing)
C309319A	x	x	CL / 309 - Generic Medicines (Number of Actions)
C309319F	x	x	CL / 309 - Generic Medicines (Timing First Action)
C309319S	x	x	CL / 309 - Generic Medicines (Scored Response)
C309319T	x	x	CL / 309 - Generic Medicines (Timing)
C309320A	x	x	CL / 309 - Generic Medicines (Number of Actions)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C309320F	x	x	CL / 309 - Generic Medicines (Timing First Action)
C309320S	x	x	CL / 309 - Generic Medicines (Scored Response)
C309320T	x	x	CL / 309 - Generic Medicines (Timing)
C309321A	x	x	CL / 309 - Generic Medicines (Number of Actions)
C309321F	x	x	CL / 309 - Generic Medicines (Timing First Action)
C309321S	x	x	CL / 309 - Generic Medicines (Scored Response)
C309321T	x	x	CL / 309 - Generic Medicines (Timing)
C309322A	x	x	CL / 309 - Generic Medicines (Number of Actions)
C309322F	x	x	CL / 309 - Generic Medicines (Timing First Action)
C309322S	x	x	CL / 309 - Generic Medicines (Scored Response)
C309322T	x	x	CL / 309 - Generic Medicines (Timing)
C310406A	x	x	CL / 310 - Memory Training (Number of Actions)
C310406F	x	x	CL / 310 - Memory Training (Timing First Action)
C310406S	x	x	CL / 310 - Memory Training (Scored Response)
C310406T	x	x	CL / 310 - Memory Training (Timing)
C310407A	x	x	CL / 310 - Memory Training (Number of Actions)
C310407F	x	x	CL / 310 - Memory Training (Timing First Action)
C310407S	x	x	CL / 310 - Memory Training (Scored Response)
C310407T	x	x	CL / 310 - Memory Training (Timing)
C313410A	x	x	CL / 313 - International Calls (Number of Actions)
C313410F	x	x	CL / 313 - International Calls (Timing First Action)
C313410S	x	x	CL / 313 - International Calls (Scored Response)
C313410T	x	x	CL / 313 - International Calls (Timing)
C313411A	x	x	CL / 313 - International Calls (Number of Actions)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C313411F	x	x	CL / 313 - International Calls (Timing First Action)
C313411S	x	x	CL / 313 - International Calls (Scored Response)
C313411T	x	x	CL / 313 - International Calls (Timing)
C313412A	x	x	CL / 313 - International Calls (Number of Actions)
C313412F	x	x	CL / 313 - International Calls (Timing First Action)
C313412S	x	x	CL / 313 - International Calls (Scored Response)
C313412T	x	x	CL / 313 - International Calls (Timing)
C313413A	x	x	CL / 313 - International Calls (Number of Actions)
C313413F	x	x	CL / 313 - International Calls (Timing First Action)
C313413S	x	x	CL / 313 - International Calls (Scored Response)
C313413T	x	x	CL / 313 - International Calls (Timing)
C313414A	x	x	CL / 313 - International Calls (Number of Actions)
C313414F	x	x	CL / 313 - International Calls (Timing First Action)
C313414S	x	x	CL / 313 - International Calls (Scored Response)
C313414T	x	x	CL / 313 - International Calls (Timing)
C600C04A	x	x	CNC / 600 - Election results (Number of Actions)
C600C04F	x	x	CNC / 600 - Election results (Timing First Action)
C600C04S	x	x	CNC / 600 - Election results (Scored Response)
C600C04T	x	x	CNC / 600 - Election results (Timing)
C601C06A	x	x	CNC / 601 - Bottles (Number of Actions)
C601C06F	x	x	CNC / 601 - Bottles (Timing First Action)
C601C06S	x	x	CNC / 601 - Bottles (Scored Response)
C601C06T	x	x	CNC / 601 - Bottles (Timing)
C602501A	x	x	CN / 602 - Price Tags (Number of Actions)
C602501F	x	x	CN / 602 - Price Tags (Timing First Action)
C602501S	x	x	CN / 602 - Price Tags (Scored Response)
C602501T	x	x	CN / 602 - Price Tags (Timing)
C602502A	x	x	CN / 602 - Price Tags (Number of Actions)
C602502F	x	x	CN / 602 - Price Tags (Timing First Action)
C602502S	x	x	CN / 602 - Price Tags (Scored Response)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C602502T	x	x	CN / 602 - Price Tags (Timing)
C602503A	x	x	CN / 602 - Price Tags (Number of Actions)
C602503F	x	x	CN / 602 - Price Tags (Timing First Action)
C602503S	x	x	CN / 602 - Price Tags (Scored Response)
C602503T	x	x	CN / 602 - Price Tags (Timing)
C604505A	x	x	CN / 604 - Gas Gauge (Number of Actions)
C604505F	x	x	CN / 604 - Gas Gauge (Timing First Action)
C604505S	x	x	CN / 604 - Gas Gauge (Scored Response)
C604505T	x	x	CN / 604 - Gas Gauge (Timing)
C605506A	x	x	CN / 605 - Photo (Number of Actions)
C605506F	x	x	CN / 605 - Photo (Timing First Action)
C605506S	x	x	CN / 605 - Photo (Scored Response)
C605506T	x	x	CN / 605 - Photo (Timing)
C605507A	x	x	CN / 605 - Photo (Number of Actions)
C605507F	x	x	CN / 605 - Photo (Timing First Action)
C605507S	x	x	CN / 605 - Photo (Scored Response)
C605507T	x	x	CN / 605 - Photo (Timing)
C605508A	x	x	CN / 605 - Photo (Number of Actions)
C605508F	x	x	CN / 605 - Photo (Timing First Action)
C605508S	x	x	CN / 605 - Photo (Scored Response)
C605508T	x	x	CN / 605 - Photo (Timing)
C606509A	x	x	CN / 606 - Solution (Number of Actions)
C606509F	x	x	CN / 606 - Solution (Timing First Action)
C606509S	x	x	CN / 606 - Solution (Scored Response)
C606509T	x	x	CN / 606 - Solution (Timing)
C607510A	x	x	CN / 607 - TV (Number of Actions)
C607510F	x	x	CN / 607 - TV (Timing First Action)
C607510S	x	x	CN / 607 - TV (Scored Response)
C607510T	x	x	CN / 607 - TV (Timing)
C608513A	x	x	CN / 608 - Tree (Number of Actions)
C608513F	x	x	CN / 608 - Tree (Timing First Action)
C608513S	x	x	CN / 608 - Tree (Scored Response)
C608513T	x	x	CN / 608 - Tree (Timing)
C611516A	x	x	CN / 611 - Temp Scale (Number of Actions)
C611516F	x	x	CN / 611 - Temp Scale (Timing First Action)
C611516S	x	x	CN / 611 - Temp Scale (Scored Response)
C611516T	x	x	CN / 611 - Temp Scale (Timing)
C611517A	x	x	CN / 611 - Temp Scale (Number of Actions)
C611517F	x	x	CN / 611 - Temp Scale (Timing First Action)
C611517S	x	x	CN / 611 - Temp Scale (Scored Response)
C611517T	x	x	CN / 611 - Temp Scale (Timing)
C612518A	x	x	CN / 612 - Dioxin (Number of Actions)
C612518F	x	x	CN / 612 - Dioxin (Timing First Action)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C612518S	x	x	CN / 612 - Dioxin (Scored Response)
C612518T	x	x	CN / 612 - Dioxin (Timing)
C613520A	x	x	CN / 613 - Logbook (Number of Actions)
C613520F	x	x	CN / 613 - Logbook (Timing First Action)
C613520S	x	x	CN / 613 - Logbook (Scored Response)
C613520T	x	x	CN / 613 - Logbook (Timing)
C614601A	x	x	CN / 614 - Watch (Number of Actions)
C614601F	x	x	CN / 614 - Watch (Timing First Action)
C614601S	x	x	CN / 614 - Watch (Scored Response)
C614601T	x	x	CN / 614 - Watch (Timing)
C615602A	x	x	CN / 615 - Candles (Number of Actions)
C615602F	x	x	CN / 615 - Candles (Timing First Action)
C615602S	x	x	CN / 615 - Candles (Scored Response)
C615602T	x	x	CN / 615 - Candles (Timing)
C615603A	x	x	CN / 615 - Candles (Number of Actions)
C615603F	x	x	CN / 615 - Candles (Timing First Action)
C615603S	x	x	CN / 615 - Candles (Scored Response)
C615603T	x	x	CN / 615 - Candles (Timing)
C617605A	x	x	CN / 617 - Map (Number of Actions)
C617605F	x	x	CN / 617 - Map (Timing First Action)
C617605S	x	x	CN / 617 - Map (Scored Response)
C617605T	x	x	CN / 617 - Map (Timing)
C617606A	x	x	CN / 617 - Map (Number of Actions)
C617606F	x	x	CN / 617 - Map (Timing First Action)
C617606S	x	x	CN / 617 - Map (Scored Response)
C617606T	x	x	CN / 617 - Map (Timing)
C618607A	x	x	CN / 618 - Six Pack (Number of Actions)
C618607F	x	x	CN / 618 - Six Pack (Timing First Action)
C618607S	x	x	CN / 618 - Six Pack (Scored Response)
C618607T	x	x	CN / 618 - Six Pack (Timing)
C618608A	x	x	CN / 618 - Six Pack (Number of Actions)
C618608F	x	x	CN / 618 - Six Pack (Timing First Action)
C618608S	x	x	CN / 618 - Six Pack (Scored Response)
C618608T	x	x	CN / 618 - Six Pack (Timing)
C619609A	x	x	CN / 619 - Tiles (Number of Actions)
C619609F	x	x	CN / 619 - Tiles (Timing First Action)
C619609S	x	x	CN / 619 - Tiles (Scored Response)
C619609T	x	x	CN / 619 - Tiles (Timing)
C620610A	x	x	CN / 620 - Inflation (Number of Actions)
C620610F	x	x	CN / 620 - Inflation (Timing First Action)
C620610S	x	x	CN / 620 - Inflation (Scored Response)
C620610T	x	x	CN / 620 - Inflation (Timing)
C620612A	x	x	CN / 620 - Inflation (Number of Actions)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C620612F	x	x	CN / 620 - Inflation (Timing First Action)
C620612S	x	x	CN / 620 - Inflation (Scored Response)
C620612T	x	x	CN / 620 - Inflation (Timing)
C622615A	x	x	CN / 622 - Classified (Number of Actions)
C622615F	x	x	CN / 622 - Classified (Timing First Action)
C622615S	x	x	CN / 622 - Classified (Scored Response)
C622615T	x	x	CN / 622 - Classified (Timing)
C623616A	x	x	CN / 623 - Wine (Number of Actions)
C623616F	x	x	CN / 623 - Wine (Timing First Action)
C623616S	x	x	CN / 623 - Wine (Scored Response)
C623616T	x	x	CN / 623 - Wine (Timing)
C623617A	x	x	CN / 623 - Wine (Number of Actions)
C623617F	x	x	CN / 623 - Wine (Timing First Action)
C623617S	x	x	CN / 623 - Wine (Scored Response)
C623617T	x	x	CN / 623 - Wine (Timing)
C624619A	x	x	CN / 624 - BMI (Number of Actions)
C624619F	x	x	CN / 624 - BMI (Timing First Action)
C624619S	x	x	CN / 624 - BMI (Scored Response)
C624619T	x	x	CN / 624 - BMI (Timing)
C624620A	x	x	CN / 624 - BMI (Number of Actions)
C624620F	x	x	CN / 624 - BMI (Timing First Action)
C624620S	x	x	CN / 624 - BMI (Scored Response)
C624620T	x	x	CN / 624 - BMI (Timing)
CASEID		x	Household operational ID
CBA_CORE_STAGE1_SCORE	x	x	CBA Core score for stage 1
CBA_CORE_STAGE2_SCORE	x	x	CBA Core score for stage 2
CBA_START	x	x	Computer-based exercise agreement
CBAMOD1	x	x	CBA module 1 branch (derived)
CBAMOD1STG1	x	x	CBA module 1, stage 1 branch (derived)
CBAMOD1STG2	x	x	CBA module 1, stage 2 branch (derived)
CBAMOD2	x	x	CBA module 2 branch (derived)
CBAMOD2ALT	x	x	CBA module 1 and 2 branch (derived)
CBAMOD2STG1	x	x	CBA module 2, stage 1 branch (derived)
CBAMOD2STG2	x	x	CBA module 2, stage 2 branch (derived)
CILANG	x	x	Language for exercise
CNT_BRTH		x	Country of birth - Respondent (UN M49 numerical) (coded)
CNT_BRTHUS_C	x	x	Country of birth - Respondent (UN M49 numerical) (2 categories) (derived from CNT_BRTH)
CNT_H		x	Country in which highest qualification was gained - Respondent (UN M49 numerical) (coded)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
CNTRY	x	x	Country ID and sub-national entity sample code (string)
CNTRY_E	x	x	Participating country or sub-national entity code (string)
CNTRYID	x	x	Country ID (ISO 3166, numeric)
CNTRYID_E	x	x	Participating country or sub-national entity code (numeric)
COMPUTEREXPERIENCE	x	x	Respondent experience with computer (DERIVED BY CAPI)
CORESTAGE1_PASS	x	x	Core Stage 1 pass status
CORESTAGE2_PASS	x	x	Core Stage 2 pass status
CTRYQUAL		x	Country where highest qualification obtained (9 regions - derived)
CTRYRGN	x	x	Country region (9 regions)
D_D16a	x	x	Current work - Earnings - Salary interval per hour (DERIVED BY CAPI)
D_Q01a		x	Current work - Job title
D_Q01b		x	Current work - Responsibilities
D_Q02a		x	Current work - Kind of business, industry or service
D_Q02b		x	Current work - Main product of firm or organisation
D_Q03	x	x	Current work - Economic sector
D_Q03US	x	x	Current work - Economic sector
D_Q04	x	x	Current work - Employee or self-employed
D_Q04_T	x	x	Status at this job or business - six levels (Trend-IALS/ALL)
D_Q04_T1	x	x	Status at this job or business - four levels (Trend-IALS/ALL)
D_Q05a1		x	Current work - Start of work for employer - Age
D_Q05a1_C	x	x	Current work - Start of work for employer - Age (categorised, 9 categories)
D_Q05a2		x	Current work - Start of work for employer - Year
D_Q05A2US_C	x	x	Current work - Start of work for employer - Year (4 categories) (derived from D_Q05A2)
D_Q05a3		x	Current work - Start of work for employer - Month
D_Q05b1		x	Current work - Start of work for business - Age
D_Q05b1_C	x	x	Current work - Start of work for business - Age (categorised, 9 categories)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
D_Q05b2		x	Current work - Start of work for business - Year
D_Q05B2US_C	x	x	Current work - Start of work for business - Year (4 categories) (derived from D_Q05B2)
D_Q05b3		x	Current work - Start of work for business - Month
D_Q06a	x	x	Current work - Amount of people working for employer
D_Q06b	x	x	Current work - Amount of people working for employer increased
D_Q06c	x	x	Current work - Part of a larger organisation
D_Q07a	x	x	Current work - Employees working for you
D_Q07b		x	Current work - Employees working for you - Count
D_Q07b_C	x	x	Current work - Employees working for you - Count (collapsed, 2 categories)
D_Q08a	x	x	Current work - Managing other employees
D_Q08b	x	x	Current work - Managing other employees - Count
D_Q09		x	Current work - Type of contract
D_Q09US_C	x	x	Current work - Type of contract (5 categories) (derived from D_Q09)
D_Q10	x	x	Current work - Hours/week
D_Q10_C	x	x	Current work - Hours/week (top-coded at 60)
D_Q10_T	x	x	Hours per week at this job or business - number of hours (top coded at 97, Trend-IALS/ALL)
D_Q10_T1	x	x	Hours per week at this job or business - range of hours (Trend-IALS/ALL)
D_Q11a	x	x	Current work - Work flexibility - Sequence of tasks
D_Q11b	x	x	Current work - Work flexibility - How to do the work
D_Q11c	x	x	Current work - Work flexibility - Speed of work
D_Q11d	x	x	Current work - Work flexibility - Working hours
D_Q12a	x	x	Current work - Requirements - Education level
D_Q12A_C	x	x	Current work - Requirements - Education level (3 categories) (derived from D_Q12A)
D_Q12A_ISCED11	x	x	Current work - Requirements - Education level, ISCED 2011
D_Q12aUS		x	Current work - Requirements - Education level

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
D_Q12AUS_C	x	x	Current work - Requirements - Education level (3 categories) (derived from D_Q12AUS)
D_Q12b	x	x	Current work - Requirements - To do the job satisfactorily
D_Q12c	x	x	Current work - Requirements - Related work experience
D_Q13a	x	x	Current work - Learning - Learning from coworkers/supervisors
D_Q13b	x	x	Current work - Learning - Learning-by-doing
D_Q13c	x	x	Current work - Learning - Keeping up to date
D_Q14	x	x	Current work - Job satisfaction
D_Q16a	x	x	Current work - Earnings - Salary interval
D_Q16b		x	Current work - Earnings - Gross pay
D_Q16bUS_C	x	x	Current work - Earnings - Gross pay (topcode at 120000) (derived from D_Q16b)
D_Q16c	x	x	Current work - Earnings - Gross pay in broad categories
D_Q16d1		x	Current work - Earnings - Broad categories - Gross pay per hour
D_Q16d2		x	Current work - Earnings - Broad categories - Gross pay per day
D_Q16d3		x	Current work - Earnings - Broad categories - Gross pay per week
D_Q16d4		x	Current work - Earnings - Broad categories - Gross pay per 2 weeks
D_Q16d5		x	Current work - Earnings - Broad categories - Gross pay per month
D_Q16d6		x	Current work - Earnings - Broad categories - Gross pay per year
D_Q17a	x	x	Current work - Earnings - Additional payments
D_Q17b		x	Current work - Earnings - Additional payments amount last year
D_Q17bUS_C	x	x	Current work - Earnings - Additional payments amount last year (11 categories) (derived from D_Q17b)
D_Q17c	x	x	Current work - Earnings - Additional payments in broad categories
D_Q17d		x	Current work - Earnings - Additional payments - Broad - Last year
D_Q18a		x	Current work - Earnings - Total earnings last year
D_Q18a_T	x	x	Annual net income before taxes and deductions (Trend-IALS/ALL)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
D_Q18b	x	x	Current work - Earnings - Total earnings broad categories
D_Q18c1		x	Current work - Earnings - Broad categories - Total earnings last month
D_Q18c2		x	Current work - Earnings - Broad categories - Total earnings last year
D_S09		x	Current work - Other type of contract specified
D_S16a		x	Current work - Earnings - Hours per piece
D302C02A	x	x	CLC / 302 - Election Results (Number of Actions)
D302C02F	x	x	CLC / 302 - Election Results (Timing First Action)
D302C02S	x	x	CLC / 302 - Election Results (Scored Response)
D302C02T	x	x	CLC / 302 - Election Results (Timing)
D304710A	x	x	CL / 304 - Contact Employer (Number of Actions)
D304710F	x	x	CL / 304 - Contact Employer (Timing First Action)
D304710S	x	x	CL / 304 - Contact Employer (Scored Response)
D304710T	x	x	CL / 304 - Contact Employer (Timing)
D304711A	x	x	CL / 304 - Contact Employer (Number of Actions)
D304711F	x	x	CL / 304 - Contact Employer (Timing First Action)
D304711S	x	x	CL / 304 - Contact Employer (Scored Response)
D304711T	x	x	CL / 304 - Contact Employer (Timing)
D306110A	x	x	CL / 306 - Canco (Number of Actions)
D306110F	x	x	CL / 306 - Canco (Timing First Action)
D306110S	x	x	CL / 306 - Canco (Scored Response)
D306110T	x	x	CL / 306 - Canco (Timing)
D306111A	x	x	CL / 306 - Canco (Number of Actions)
D306111F	x	x	CL / 306 - Canco (Timing First Action)
D306111S	x	x	CL / 306 - Canco (Scored Response)
D306111T	x	x	CL / 306 - Canco (Timing)
D307401A	x	x	CL / 307 - MEDCO Aspirin (Number of Actions)
D307401F	x	x	CL / 307 - MEDCO Aspirin (Timing First Action)
D307401S	x	x	CL / 307 - MEDCO Aspirin (Scored Response)
D307401T	x	x	CL / 307 - MEDCO Aspirin (Timing)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
D307402A	x	x	CL / 307 - MEDCO Aspirin (Number of Actions)
D307402F	x	x	CL / 307 - MEDCO Aspirin (Timing First Action)
D307402S	x	x	CL / 307 - MEDCO Aspirin (Scored Response)
D307402T	x	x	CL / 307 - MEDCO Aspirin (Timing)
D311701A	x	x	CL / 311 - Dutch Women (Number of Actions)
D311701F	x	x	CL / 311 - Dutch Women (Timing First Action)
D311701S	x	x	CL / 311 - Dutch Women (Scored Response)
D311701T	x	x	CL / 311 - Dutch Women (Timing)
D315512A	x	x	CL / 315 - Distances-Mexican Cities (Number of Actions)
D315512F	x	x	CL / 315 - Distances-Mexican Cities (Timing First Action)
D315512S	x	x	CL / 315 - Distances-Mexican Cities (Scored Response)
D315512T	x	x	CL / 315 - Distances-Mexican Cities (Timing)
DISP_BQ	x	x	Final disposition code for BQ/JRA
DISP_CIBQ	x	x	Final disposition code for person - combining CI and BQ/JRA (derived)
DOBMM		x	Date of birth (derived from BQ)
DOBYY		x	Date of birth year (derived from BQ)
E_Q01a		x	Last job - Job title
E_Q01b		x	Last job - Responsibilities
E_Q02a		x	Last job - Kind of business, industry or service
E_Q02b		x	Last job - Main product of firm or organisation
E_Q03	x	x	Last job - Economic sector
E_Q03US	x	x	Last job - Economic sector
E_Q04	x	x	Last job - Employee or self-employed
E_Q05a1		x	Last job - Start of work for employer - Age
E_Q05a1_C	x	x	Last job - Start of work for employer - Age (categorised, 9 categories)
E_Q05a2		x	Last job - Start of work for employer - Year
E_Q05A2US_C	x	x	Last job - Start of work for employer - Year (4 categories) (derived from E_Q05A2)
E_Q05b1		x	Last job - Start of work for business - Age
E_Q05b1_C	x	x	Last job - Start of work for business - Age (categorised, 9 categories)
E_Q05b2		x	Last job - Start of work for business - Year
E_Q05B2US_C	x	x	Last job - Start of work for business - Year (2 categories) (derived from E_Q05B2)
E_Q06	x	x	Last job - Amount of people working for employer

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
E_Q07a		x	Last job - Employees working for you
E_Q07b		x	Last job - Employees working for you - Count
E_Q08		x	Last job - Type of contract
E_Q08US_C	x	x	Last job - Type of contract (5 categories) (derived from E_Q08)
E_Q09	x	x	Last job - Hours/week
E_Q09_C	x	x	Last work - Hours/week (top-coded at 60)
E_Q10	x	x	Last job - Reason for end of job
E_S08		x	Last job - Other type of contract specified
E318001A	x	x	CL / 318 - Civil Engineering (Number of Actions)
E318001F	x	x	CL / 318 - Civil Engineering (Timing First Action)
E318001S	x	x	CL / 318 - Civil Engineering (Scored Response)
E318001T	x	x	CL / 318 - Civil Engineering (Timing)
E318003A	x	x	CL / 318 - Civil Engineering (Number of Actions)
E318003F	x	x	CL / 318 - Civil Engineering (Timing First Action)
E318003S	x	x	CL / 318 - Civil Engineering (Scored Response)
E318003T	x	x	CL / 318 - Civil Engineering (Timing)
E320001A	x	x	CL / 320 - Discussion forum (Number of Actions)
E320001F	x	x	CL / 320 - Discussion forum (Timing First Action)
E320001S	x	x	CL / 320 - Discussion forum (Scored Response)
E320001T	x	x	CL / 320 - Discussion forum (Timing)
E320003A	x	x	CL / 320 - Discussion forum (Number of Actions)
E320003F	x	x	CL / 320 - Discussion forum (Timing First Action)
E320003S	x	x	CL / 320 - Discussion forum (Scored Response)
E320003T	x	x	CL / 320 - Discussion forum (Timing)
E320004A	x	x	CL / 320 - Discussion forum (Number of Actions)
E320004F	x	x	CL / 320 - Discussion forum (Timing First Action)
E320004S	x	x	CL / 320 - Discussion forum (Scored Response)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
E320004T	x	x	CL / 320 - Discussion forum (Timing)
E321001A	x	x	CL / 321 - Internet Poll (Number of Actions)
E321001F	x	x	CL / 321 - Internet Poll (Timing First Action)
E321001S	x	x	CL / 321 - Internet Poll (Scored Response)
E321001T	x	x	CL / 321 - Internet Poll (Timing)
E321002A	x	x	CL / 321 - Internet Poll (Number of Actions)
E321002F	x	x	CL / 321 - Internet Poll (Timing First Action)
E321002S	x	x	CL / 321 - Internet Poll (Scored Response)
E321002T	x	x	CL / 321 - Internet Poll (Timing)
E322001A	x	x	CL / 322 - Lakeside Fun Run (Number of Actions)
E322001F	x	x	CL / 322 - Lakeside Fun Run (Timing First Action)
E322001S	x	x	CL / 322 - Lakeside Fun Run (Scored Response)
E322001T	x	x	CL / 322 - Lakeside Fun Run (Timing)
E322002A	x	x	CL / 322 - Lakeside Fun Run (Number of Actions)
E322002F	x	x	CL / 322 - Lakeside Fun Run (Timing First Action)
E322002S	x	x	CL / 322 - Lakeside Fun Run (Scored Response)
E322002T	x	x	CL / 322 - Lakeside Fun Run (Timing)
E322003A	x	x	CL / 322 - Lakeside Fun Run (Number of Actions)
E322003F	x	x	CL / 322 - Lakeside Fun Run (Timing First Action)
E322003S	x	x	CL / 322 - Lakeside Fun Run (Scored Response)
E322003T	x	x	CL / 322 - Lakeside Fun Run (Timing)
E322004A	x	x	CL / 322 - Lakeside Fun Run (Number of Actions)
E322004F	x	x	CL / 322 - Lakeside Fun Run (Timing First Action)
E322004S	x	x	CL / 322 - Lakeside Fun Run (Scored Response)
E322004T	x	x	CL / 322 - Lakeside Fun Run (Timing)
E322005A	x	x	CL / 322 - Lakeside Fun Run (Number of Actions)
E322005F	x	x	CL / 322 - Lakeside Fun Run (Timing First Action)
E322005S	x	x	CL / 322 - Lakeside Fun Run (Scored Response)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
E322005T	x	x	CL / 322 - Lakeside Fun Run (Timing)
E323002A	x	x	CL / 323 - Library Search (Number of Actions)
E323002F	x	x	CL / 323 - Library Search (Timing First Action)
E323002S	x	x	CL / 323 - Library Search (Scored Response)
E323002T	x	x	CL / 323 - Library Search (Timing)
E323003A	x	x	CL / 323 - Library Search (Number of Actions)
E323003F	x	x	CL / 323 - Library Search (Timing First Action)
E323003S	x	x	CL / 323 - Library Search (Scored Response)
E323003T	x	x	CL / 323 - Library Search (Timing)
E323004A	x	x	CL / 323 - Library Search (Number of Actions)
E323004F	x	x	CL / 323 - Library Search (Timing First Action)
E323004S	x	x	CL / 323 - Library Search (Scored Response)
E323004T	x	x	CL / 323 - Library Search (Timing)
E323005A	x	x	CL / 323 - Library Search (Number of Actions)
E323005F	x	x	CL / 323 - Library Search (Timing First Action)
E323005S	x	x	CL / 323 - Library Search (Scored Response)
E323005T	x	x	CL / 323 - Library Search (Timing)
E327001A	x	x	CL / 327 - Summer Streets (Number of Actions)
E327001F	x	x	CL / 327 - Summer Streets (Timing First Action)
E327001S	x	x	CL / 327 - Summer Streets (Scored Response)
E327001T	x	x	CL / 327 - Summer Streets (Timing)
E327002A	x	x	CL / 327 - Summer Streets (Number of Actions)
E327002F	x	x	CL / 327 - Summer Streets (Timing First Action)
E327002S	x	x	CL / 327 - Summer Streets (Scored Response)
E327002T	x	x	CL / 327 - Summer Streets (Timing)
E327003A	x	x	CL / 327 - Summer Streets (Number of Actions)
E327003F	x	x	CL / 327 - Summer Streets (Timing First Action)
E327003S	x	x	CL / 327 - Summer Streets (Scored Response)
E327003T	x	x	CL / 327 - Summer Streets (Timing)
E327004A	x	x	CL / 327 - Summer Streets (Number of Actions)
E327004F	x	x	CL / 327 - Summer Streets (Timing First Action)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
E327004S	x	x	CL / 327 - Summer Streets (Scored Response)
E327004T	x	x	CL / 327 - Summer Streets (Timing)
E329002A	x	x	CL / 329 - Work-related Stress (Number of Actions)
E329002F	x	x	CL / 329 - Work-related Stress (Timing First Action)
E329002S	x	x	CL / 329 - Work-related Stress (Scored Response)
E329002T	x	x	CL / 329 - Work-related Stress (Timing)
E329003A	x	x	CL / 329 - Work-related Stress (Number of Actions)
E329003F	x	x	CL / 329 - Work-related Stress (Timing First Action)
E329003S	x	x	CL / 329 - Work-related Stress (Scored Response)
E329003T	x	x	CL / 329 - Work-related Stress (Timing)
E632001A	x	x	CN / 632 - Educational level (Number of Actions)
E632001F	x	x	CN / 632 - Educational level (Timing First Action)
E632001S	x	x	CN / 632 - Educational level (Scored Response)
E632001T	x	x	CN / 632 - Educational level (Timing)
E632002A	x	x	CN / 632 - Educational level (Number of Actions)
E632002F	x	x	CN / 632 - Educational level (Timing First Action)
E632002S	x	x	CN / 632 - Educational level (Scored Response)
E632002T	x	x	CN / 632 - Educational level (Timing)
E634001A	x	x	CN / 634 - Peanuts (Number of Actions)
E634001F	x	x	CN / 634 - Peanuts (Timing First Action)
E634001S	x	x	CN / 634 - Peanuts (Scored Response)
E634001T	x	x	CN / 634 - Peanuts (Timing)
E634002A	x	x	CN / 634 - Peanuts (Number of Actions)
E634002F	x	x	CN / 634 - Peanuts (Timing First Action)
E634002S	x	x	CN / 634 - Peanuts (Scored Response)
E634002T	x	x	CN / 634 - Peanuts (Timing)
E635001A	x	x	CN / 635 - Parking Map (Number of Actions)
E635001F	x	x	CN / 635 - Parking Map (Timing First Action)
E635001S	x	x	CN / 635 - Parking Map (Scored Response)
E635001T	x	x	CN / 635 - Parking Map (Timing)
E636001A	x	x	CN / 636 - Lab Report (Number of Actions)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
E636001F	x	x	CN / 636 - Lab Report (Timing First Action)
E636001S	x	x	CN / 636 - Lab Report (Scored Response)
E636001T	x	x	CN / 636 - Lab Report (Timing)
E641001A	x	x	CN / 641 - Amoeba (Number of Actions)
E641001F	x	x	CN / 641 - Amoeba (Timing First Action)
E641001S	x	x	CN / 641 - Amoeba (Scored Response)
E641001T	x	x	CN / 641 - Amoeba (Timing)
E644002A	x	x	CN / 644 - NZExports (Number of Actions)
E644002F	x	x	CN / 644 - NZExports (Timing First Action)
E644002S	x	x	CN / 644 - NZExports (Scored Response)
E644002T	x	x	CN / 644 - NZExports (Timing)
E645001A	x	x	CNC / 645 - Airport Timetable (Number of Actions)
E645001F	x	x	CNC / 645 - Airport Timetable (Timing First Action)
E645001S	x	x	CNC / 645 - Airport Timetable (Scored Response)
E645001T	x	x	CNC / 645 - Airport Timetable (Timing)
E646002A	x	x	CN / 646 - Rug Production (Number of Actions)
E646002F	x	x	CN / 646 - Rug Production (Timing First Action)
E646002S	x	x	CN / 646 - Rug Production (Scored Response)
E646002T	x	x	CN / 646 - Rug Production (Timing)
E650001A	x	x	CN / 650 - Urban Population (Number of Actions)
E650001F	x	x	CN / 650 - Urban Population (Timing First Action)
E650001S	x	x	CN / 650 - Urban Population (Scored Response)
E650001T	x	x	CN / 650 - Urban Population (Timing)
E651002A	x	x	CN / 651 - Fertilizer (Number of Actions)
E651002F	x	x	CN / 651 - Fertilizer (Timing First Action)
E651002S	x	x	CN / 651 - Fertilizer (Scored Response)
E651002T	x	x	CN / 651 - Fertilizer (Timing)
E655001A	x	x	CN / 655 - Path (Number of Actions)
E655001F	x	x	CN / 655 - Path (Timing First Action)
E655001S	x	x	CN / 655 - Path (Scored Response)
E655001T	x	x	CN / 655 - Path (Timing)
E657001A	x	x	CN / 657 - Package (Number of Actions)
E657001F	x	x	CN / 657 - Package (Timing First Action)
E657001S	x	x	CN / 657 - Package (Scored Response)
E657001T	x	x	CN / 657 - Package (Timing)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
E660003A	x	x	CN / 660 - Weight history (Number of Actions)
E660003F	x	x	CN / 660 - Weight history (Timing First Action)
E660003S	x	x	CN / 660 - Weight history (Scored Response)
E660003T	x	x	CN / 660 - Weight history (Timing)
E660004A	x	x	CN / 660 - Weight history (Number of Actions)
E660004F	x	x	CN / 660 - Weight history (Timing First Action)
E660004S	x	x	CN / 660 - Weight history (Scored Response)
E660004T	x	x	CN / 660 - Weight history (Timing)
E661001A	x	x	CN / 661 - Study fees (Number of Actions)
E661001F	x	x	CN / 661 - Study fees (Timing First Action)
E661001S	x	x	CN / 661 - Study fees (Scored Response)
E661001T	x	x	CN / 661 - Study fees (Timing)
E661002A	x	x	CN / 661 - Study fees (Number of Actions)
E661002F	x	x	CN / 661 - Study fees (Timing First Action)
E661002S	x	x	CN / 661 - Study fees (Scored Response)
E661002T	x	x	CN / 661 - Study fees (Timing)
E664001A	x	x	CN / 664 - Orchestra tickets (Number of Actions)
E664001F	x	x	CN / 664 - Orchestra tickets (Timing First Action)
E664001S	x	x	CN / 664 - Orchestra tickets (Scored Response)
E664001T	x	x	CN / 664 - Orchestra tickets (Timing)
E665001A	x	x	CN / 665 - Cooper test (Number of Actions)
E665001F	x	x	CN / 665 - Cooper test (Timing First Action)
E665001S	x	x	CN / 665 - Cooper test (Scored Response)
E665001T	x	x	CN / 665 - Cooper test (Timing)
E665002A	x	x	CN / 665 - Cooper test (Number of Actions)
E665002F	x	x	CN / 665 - Cooper test (Timing First Action)
E665002S	x	x	CN / 665 - Cooper test (Scored Response)
E665002T	x	x	CN / 665 - Cooper test (Timing)
EARNFLAG	x	x	Earnings including bonuses reporting method (derived)
EARNHR		x	Hourly earnings excluding bonuses for wage and salary earners (derived)
EARNHRBONUS		x	Hourly earnings including bonuses for wage and salary earners (derived)
EARNHRBONUSDCL	x	x	Hourly earnings including bonuses for wage and salary earners, in deciles (derived)
EARNHRBONUSPPP		x	Hourly earnings including bonuses for wage and salary earners, PPP corrected $US (derived)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
EARNHRBONUSPPPUS_C	x	x	Hourly earnings including bonuses for wage and salary earners, PPP corrected $US (topcoded) (derived from EARNHRBONUSPPP)
EARNHRBONUSUS_C	x	x	Hourly earnings including bonuses for wage and salary earners (topcoded) (derived from EARNHRBONUS)
EARNHRDCL	x	x	Hourly earnings excluding bonuses for wage and salary earners, in deciles (derived)
EARNHRPPP		x	Hourly earnings excluding bonuses for wage and salary earners, PPP corrected $US (derived)
EARNHRPPPUS_C	x	x	Hourly earnings excluding bonuses for wage and salary earners, PPP corrected $US (topcoded) (derived from EARNHRPPP)
EARNHRUS_C	x	x	Hourly earnings excluding bonuses for wage and salary earners (topcoded) (derived from EARNHR)
EARNMTH		x	Monthly earnings excluding bonuses for wage and salary earners (derived)
EARNMTHALL		x	Monthly earnings including bonuses for wage and salary earners and self-employed (derived)
EARNMTHALLDCL	x	x	Monthly earnings including bonuses for wage and salary earners and self-employed, in deciles (derived)
EARNMTHALLPPP		x	Monthly earnings including bonuses for wage and salary earners and self-employed, PPP corrected $US (derived)
EARNMTHALLPPPUS_C	x	x	Monthly earnings including bonuses for wage and salary earners and self-employed, PPP corrected $US (topcoded) (derived from EARNMTHALLPPP)
EARNMTHALLUS_C	x	x	Monthly earnings including bonuses for wage and salary earners and self-employed (topcoded) (derived from EARNMTHALL)
EARNMTHBONUS		x	Monthly earnings including bonuses for wage and salary earners (derived)
EARNMTHBONUSPPP		x	Monthly earnings including bonuses for wage and salary earners, PPP corrected $US (derived)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
EARNMTHBONUSPPPUS_C	x	x	Monthly earnings including bonuses for wage and salary earners, PPP corrected $US (topcoded) (derived from EARNMTHBONUSPPP)
EARNMTHBONUSUS_C	x	x	Monthly earnings including bonuses for wage and salary earners (topcoded) (derived from EARNMTHBONUS)
EARNMTHPPP		x	Monthly earnings excluding bonuses for wage and salary earners, PPP corrected $US (derived)
EARNMTHPPPUS_C	x	x	Monthly earnings excluding bonuses for wage and salary earners, PPP corrected $US (topcoded) (derived from EARNMTHPPP)
EARNMTHSELFPPP		x	Monthly earnings for self-employed, PPP corrected $US (derived)
EARNMTHSELFPPPUS_C	x	x	Monthly earnings for self-employed, PPP corrected $US (tocoded) (derived from EARNMTHSELFPPP)
EARNMTHUS_C	x	x	Monthly earnings excluding bonuses for wage and salary earners (topcoded) (derived from EARNMTH)
EDCAT6	x	x	Highest level of formal education obtained (6 categories - derived)
EDCAT7	x	x	Highest level of formal education obtained (7 categories - derived)
EDCAT8	x	x	Highest level of formal education obtained (8 categories - derived)
EDLEVEL3	x	x	Educational level of the respondent (DERIVED BY CAPI)
EDWORK	x	x	Interaction between adults' work and education status (derived)
EMP_6CAT	x	x	Employment status (derived, 6 categories)
EMPSTAT	x	x	Employment and student status (derived, 5 categories)
EXCFRM_PROP		x	Proportion in target population who are excluded from the sampling frame
F_Q01b	x	x	Skill use work - Time cooperating with coworkers
F_Q02a	x	x	Skill use work - How often - Sharing work-related info
F_Q02b	x	x	Skill use work - How often - Teaching people
F_Q02c	x	x	Skill use work - How often - Presentations
F_Q02d	x	x	Skill use work - How often - Selling

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
F_Q02e	x	x	Skill use work - How often - Advising people
F_Q03a	x	x	Skill use work - How often - Planning own activities
F_Q03b	x	x	Skill use work - How often - Planning others activities
F_Q03c	x	x	Skill use work - How often - Organising own time
F_Q04a	x	x	Skill use work - How often - Influencing people
F_Q04b	x	x	Skill use work - How often - Negotiating with people
F_Q05a	x	x	Skill use work - Problem solving - Simple problems
F_Q05b	x	x	Skill use work - Problem solving - Complex problems
F_Q06b	x	x	Skill use work - How often - Working physically for long
F_Q06c	x	x	Skill use work - How often - Using hands or fingers
F_Q07a	x	x	Skill use work - Not challenged enough
F_Q07b	x	x	Skill use work - Need more training
FAET12	x	x	Participated in formal AET in 12 months preceding survey (see AETPOP - derived)
FAET12JR	x	x	Participated in formal AET for job-related reasons in 12 months preceding survey (see AETPOP - derived)
FAET12NJR	x	x	Participated in formal AET for non job-related reasons in 12 months preceding survey (see AETPOP - derived)
FE12	x	x	Participated in formal education in 12 months preceding survey (derived)
FIRLGRGN		x	Source region of first language learned at home in childhood and still understand (9 regions - derived)
FIRLGRGNUS_C	x	x	Source region of first language learned at home in childhood and still understand (9 regions) (2 categories) (derived from FIRLGRGN)
FNFAET12	x	x	Participated in formal or nonformal AET in 12 months preceding survey (see AETPOP - derived)
FNFAET12JR	x	x	Participated in formal or nonformal AET for job-related reasons in 12 months preceding survey (see AETPOP - derived)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
FNFAET12NJR	x	x	Participated in formal or nonformal AET for non job-related reasons in 12 mon. preceding survey (see AETPOP - derived)
FNFE12JR	x	x	Participated in formal or nonformal education for job-related reasons in 12 months preceding the survey (derived)
FORBILANG	x	x	Has learned as a child and still understands at least two languages not including test language (derived)
FORBORNLANG	x	x	Interactions between foreign-born and language status (2 categories - derived)
G_Q01a	x	x	Skill use work - Literacy - Read directions or instructions
G_Q01a_T	x	x	As part of job, read or use directions or instructions (Trend-IALS/ALL)
G_Q01a_T1	x	x	As part of job, read or use directions or instructions - levels collapsed (Trend-IALS/ALL)
G_Q01b	x	x	Skill use work - Literacy - Read letters memos or mails
G_Q01b_T	x	x	As part of job, read or use letters, memos, e-mails (Trend-IALS/ALL)
G_Q01b_T1	x	x	As part of job, read or use letters, memos, e-mails - levels collapsed (Trend-IALS/ALL)
G_Q01c	x	x	Skill use work - Literacy - Read newspapers or magazines
G_Q01c_T	x	x	As part of job, read or use reports, articles, magazines, journals (Trend-IALS/ALL)
G_Q01c_T1	x	x	As part of job, read or use reports, articles, magazines, journals - levels collapsed (Trend-IALS/ALL)
G_Q01d	x	x	Skill use work - Literacy - Read professional journals or publications
G_Q01e	x	x	Skill use work - Literacy - Read books
G_Q01f	x	x	Skill use work - Literacy - Read manuals or reference materials
G_Q01f_T	x	x	As part of job, read or use manuals, reference books, catalogues (Trend-IALS/ALL)
G_Q01f_T1	x	x	As part of job, read or use manuals, reference books, catalogues - levels collapsed (Trend-IALS/ALL)
G_Q01g	x	x	Skill use work - Literacy - Read financial statements

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
G_Q01g_T	x	x	As part of job, read or use bills, invoices, spreadsheets, budget tables (Trend-IALS/ALL)
G_Q01g_T1	x	x	As part of job, read or use bills, invoices, spreadsheets, budget tables - levels collapsed (Trend-IALS/ALL)
G_Q01h	x	x	Skill use work - Literacy - Read diagrams maps or schematics
G_Q01h_T	x	x	As part of job, read or use diagrams or schematics (Trend-IALS/ALL)
G_Q01h_T1	x	x	As part of job, read or use diagrams or schematics - levels collapsed (Trend-IALS/ALL)
G_Q02a	x	x	Skill use work - Literacy - Write letters memos or mails
G_Q02b	x	x	Skill use work - Literacy - Write articles
G_Q02c	x	x	Skill use work - Literacy - Write reports
G_Q02d	x	x	Skill use work - Literacy - Fill in forms
G_Q03b	x	x	Skill use work - Numeracy - How often - Calculating costs or budgets
G_Q03c	x	x	Skill use work - Numeracy - How often - Use or calculate fractions or percentages
G_Q03d	x	x	Skill use work - Numeracy - How often - Use a calculator
G_Q03f	x	x	Skill use work - Numeracy - How often - Prepare charts graphs or tables
G_Q03g	x	x	Skill use work - Numeracy - How often - Use simple algebra or formulas
G_Q03h	x	x	Skill use work - Numeracy - How often - Use advanced math or statistics
G_Q04	x	x	Skill use work - ICT - Experience with computer in job
G_Q04_T	x	x	Ever used computer (Trend-IALS/ALL)
G_Q05a	x	x	Skill use work - ICT - Internet - How often - For mail
G_Q05c	x	x	Skill use work - ICT - Internet - How often - Work related info
G_Q05d	x	x	Skill use work - ICT - Internet - How often - Conduct transactions
G_Q05e	x	x	Skill use work - ICT - Computer - How often - Spreadsheets
G_Q05f	x	x	Skill use work - ICT - Computer - How often - Word

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
G_Q05g	x	x	Skill use work - ICT - Computer - How often - Programming language
G_Q05h	x	x	Skill use work - ICT - Computer - How often - Real-time discussions
G_Q06	x	x	Skill use work - ICT - Computer - Level of computer use
G_Q07	x	x	Skill use work - ICT - Computer - Got the skills needed
G_Q08	x	x	Skill use work - ICT - Computer - Lack of skills affect career
GENDER		x	Person gender (derived from BQ)
GENDER_R	x	x	Person resolved gender from BQ and QC check (derived)
H_Q01a	x	x	Skill use everyday life - Literacy - Read directions or instructions
H_Q01b	x	x	Skill use everyday life - Literacy - Read letters memos or mails
H_Q01b_T	x	x	In daily life, read or use letters, notes, e-mails (Trend-IALS/ALL)
H_Q01c	x	x	Skill use everyday life - Literacy - Read newspapers or magazines
H_Q01c_T	x	x	In daily life, read or use newspapers, magazines, articles (Trend-IALS/ALL)
H_Q01d	x	x	Skill use everyday life - Literacy - Read professional journals or publications
H_Q01e	x	x	Skill use everyday life - Literacy - Read books
H_Q01e_T	x	x	In daily life, read or use books (fiction or nonfiction; not for job or school) (Trend-IALS/ALL)
H_Q01f	x	x	Skill use everyday life - Literacy - Read manuals or reference materials
H_Q01g	x	x	Skill use everyday life - Literacy - Read financial statements
H_Q01h	x	x	Skill use everyday life - Literacy - Read diagrams maps or schematics
H_Q02a	x	x	Skill use everyday life - Literacy - Write letters memos or mails
H_Q02b	x	x	Skill use everyday life - Literacy - Write articles
H_Q02c	x	x	Skill use everyday life - Literacy - Write reports
H_Q02d	x	x	Skill use everyday life - Literacy - Fill in forms

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
H_Q03b	x	x	Skill use everyday life - Numeracy - How often - Calculating costs or budgets
H_Q03c	x	x	Skill use everyday life - Numeracy - How often - Use or calculate fractions or percentages
H_Q03d	x	x	Skill use everyday life - Numeracy - How often - Use a calculator
H_Q03f	x	x	Skill use everyday life - Numeracy - How often - Prepare charts graphs or tables
H_Q03g	x	x	Skill use everyday life - Numeracy - How often - Use simple algebra or formulas
H_Q03h	x	x	Skill use everyday life - Numeracy - How often - Use advanced math or statistics
H_Q04a	x	x	Skill use everyday life - ICT - Ever used computer
H_Q04b	x	x	Skill use everyday life - ICT - Experience with computer everyday life
H_Q05a	x	x	Skill use everyday life - ICT - Internet - How often - For mail
H_Q05c	x	x	Skill use everyday life - ICT - Internet - How often - In order to better understand various issues
H_Q05d	x	x	Skill use everyday life - ICT - Internet - How often - Conduct transactions
H_Q05e	x	x	Skill use everyday life - ICT - Computer - How often - Spreadsheets
H_Q05f	x	x	Skill use everyday life - ICT - Computer - How often - Word
H_Q05g	x	x	Skill use everyday life - ICT - Computer - How often - Programming language
H_Q05h	x	x	Skill use everyday life - ICT - Computer - How often - Real-time discussions
HIDD_DU		x	Hidden dwelling unit (DU)
HOMLANG	x	x	Test language same as language spoken most often at home (derived)
HOMLGRGN		x	Source region of language spoken most often at home (9 regions - derived)
HOMLGRGNUS_C	x	x	Source region of language spoken most often at home (9 regions) (2 categories) (derived from HOMLGRGN)
I_Q04b	x	x	About yourself - Learning strategies - Relate new ideas into real life
I_Q04d	x	x	About yourself - Learning strategies - Like learning new things

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
I_Q04h	x	x	About yourself - Learning strategies - Attribute something new
I_Q04j	x	x	About yourself - Learning strategies - Get to the bottom of difficult things
I_Q04l	x	x	About yourself - Learning strategies - Figure out how different ideas fit together
I_Q04m	x	x	About yourself - Learning strategies - Looking for additional info
I_Q05f	x	x	About yourself - Cultural engagement - Voluntary work for nonprofit organisations
I_Q06a	x	x	About yourself - Political efficacy - No influence on the government
I_Q06dUSX1a	x	x	About yourself - Political efficacy - Information from newspapers
I_Q06dUSX1b	x	x	About yourself - Political efficacy - Information from magazines
I_Q06dUSX1c	x	x	About yourself - Political efficacy - Information from internet
I_Q06dUSX1d	x	x	About yourself - Political efficacy - Information from radio
I_Q06dUSX1e	x	x	About yourself - Political efficacy - Information from television
I_Q06dUSX1f	x	x	About yourself - Political efficacy - Information from books or brochures
I_Q06dUSX1g	x	x	About yourself - Political efficacy - Information from family members, friends, or coworkers
I_Q07a	x	x	About yourself - Social trust - Trust only few people
I_Q07b	x	x	About yourself - Social trust - Other people take advantage of you
I_Q08	x	x	About yourself - Health - State
I_Q08_T	x	x	General health (Trend-IALS/ALL)
I_Q08USX1	x	x	About yourself - Health - Difficulty seeing print
I_Q08USX2	x	x	About yourself - Health - Difficulty hearing conversation
I_Q08USX3	x	x	About yourself - Health - Diagnosed learning disabled
I_Q10bUSX1	x	x	About yourself - Health - Have medical insurance
I_Q10bUSX2a	x	x	About yourself - Health - Health information from newspapers

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
I_Q10bUSX2b	x	x	About yourself - Health - Health information from magazines
I_Q10bUSX2c	x	x	About yourself - Health - Health information from internet
I_Q10bUSX2d	x	x	About yourself - Health - Health information from radio
I_Q10bUSX2e	x	x	About yourself - Health - Health information from television
I_Q10bUSX2f	x	x	About yourself - Health - Health information from books or brochures
I_Q10bUSX2g	x	x	About yourself - Health - Health information from family members, friends, or coworkers
I_Q10bUSX2h	x	x	About yourself - Health - Health information from health professional
I_Q10bUSX3a	x	x	About yourself - Health - Flu shot in past year
I_Q10bUSX3b	x	x	About yourself - Health - Mammogram in past year
I_Q10bUSX3c	x	x	About yourself - Health - Pap smear in past year
I_Q10bUSX3d	x	x	About yourself - Health - Screen for colon cancer in past year
I_Q10bUSX3e	x	x	About yourself - Health - Vision check in past year
I_Q10bUSX3f	x	x	About yourself - Health - Screen for prostate cancer in past year
I_Q10bUSX3g	x	x	About yourself - Health - Screen for osteoporosis in past year
I_Q10bUSX3h	x	x	About yourself - Health - Seen dentist in past year
ICTHOME	x	x	Index of use of ICT skills at home (derived)
ICTHOME_WLE_CA	x	x	Index of use of ICT skills at home, categorised WLE (derived)
ICTWORK	x	x	Index of use of ICT skills at work (derived)
ICTWORK_WLE_CA	x	x	Index of use of ICT skills at work, categorised WLE (derived)
ID_PSU		x	Sampling ID: Primary sampling unit (PSU) identification number
ID_SSU		x	Sampling ID: Second-stage sampling unit (SSU) identification number
IMGEN	x	x	First and second generation immigrants (derived)
IMPAR	x	x	Parents' immigration status (derived)
IMYRCAT	x	x	Years in country (2-category - derived)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
IMYRS		x	Years in country (derived)
IMYRS_C	x	x	Years in country (categorised, 4 categories)
INFLUENCE	x	x	Index of use of influencing skills at work (derived)
INFLUENCE_WLE_CA	x	x	Index of use of influencing skills at work, categorised WLE (derived)
INPIAAC	x	x	PIAAC Sample Indicator
INTLFLAG	x	x	Flag for international comparison of U.S. combined Main Study and National Supplement
ISCED_HF		x	Level of Highest Qualification (Foreign) - Respondent (ISCED) (coded)
ISCED_HF_C		x	Level of Highest Qualification (collapsed, 14 categories)
ISCO08_C		x	Current Job Occupation - Respondent (ISCO 2008) (coded)
ISCO08_CUS_C	x	x	Current Job Occupation - Respondent (ISCO 2008) (combined into 3-digit categories) (derived from ISCO08_C)
ISCO08_L		x	Last Job Occupation - Respondent (ISCO 2008) (coded)
ISCO08_LUS_C	x	x	Last Job Occupation - Respondent (ISCO 2008) (combined into 3-digit categories) (derived from ISCO08_L)
ISCO08_US		x	Trade or Craft - Respondent (ISCO 2008)
ISCO08_US_C	x	x	Trade or Craft – (ISCO 2008) (combined into 3-digit categories) (derived from ISCO08_US)
ISCO1C	x	x	Occupational classification of respondent's job at 1-digit level (ISCO 2008), current job (derived)
ISCO1L	x	x	Occupational classification of respondent's job at 1-digit level (ISCO 2008), last job (derived)
ISCO2C	x	x	Occupational classification of respondent's job at 2-digit level (ISCO 2008), current job (derived)
ISCO2L	x	x	Occupational classification of respondent's job at 2-digit level (ISCO 2008), last job (derived)
ISCOSKIL4	x	x	Occupational classification of respondent's job (4 skill based categories), last or current (derived)
ISIC1C	x	x	Industry classification of respondent's job at 1-digit level (ISIC rev 4), current job (derived)
ISIC1L	x	x	Industry classification of respondent's job at 1-digit level(ISIC rev 4), last job (derived)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
ISIC2C	x	x	Industry classification of respondent's job at 2-digit level (ISIC rev 4), current job (derived)
ISIC2L	x	x	Industry classification of respondent's job at 2-digit level (ISIC rev 4), last job (derived)
ISIC4_C		x	Current Job Industry - Respondent (ISIC rev 4) (coded)
ISIC4_CUS_C	x	x	Current Job Industry - Respondent (ISIC rev 4) (combined into 3-digit categories) (derived from ISIC4_C)
ISIC4_L		x	Last Job Industry - Respondent (ISIC rev 4) (coded)
ISIC4_LUS_C	x	x	Last Job Industry - Respondent (ISIC rev 4) (combined into 3-digit categories) (derived from ISIC4_L)
J_N05a2	x	x	Background - More than one language mentioned
J_Q01		x	Background - People in household
J_Q01_C	x	x	Background - People in household (top-coded at 6)
J_Q01_T		x	Number living in household (Trend-IALS/ALL)
J_Q01_T1	x	x	Number living in household (from 1 to 7) (Trend-IALS/ALL)
J_Q02a	x	x	Background - Living with spouse or partner
J_Q02c	x	x	Background - Work situation of spouse or partner
J_Q03a	x	x	Background – Children
J_Q03b		x	Background - Number of children
J_Q03b_C	x	x	Background - Number of children (top-coded at 4)
J_Q03c		x	Background - Age of the child
J_Q03c_C	x	x	Background - Age of the child (categorised, 4 categories)
J_Q03CUS_C	x	x	Background - Age of the child (5 categories) (derived from J_Q03C)
J_Q03d1		x	Background - Age of the youngest child
J_Q03d1_C	x	x	Background - Age of the youngest child (categorised, 4 categories)
J_Q03D1US_C	x	x	Background - Age of the youngest child (5 categories) (derived from J_Q03D1)
J_Q03d2		x	Background - Age of the oldest child
J_Q03d2_C	x	x	Background - Age of the oldest child (categorised, 4 categories)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
J_Q03D2US_C	x	x	Background - Age of the oldest child (5 categories) (derived from J_Q03D2)
J_Q04a	x	x	Background - Born in country
J_Q04a_T	x	x	Born in country (Trend-IALS/ALL)
J_Q04bUS		x	Background - Country of birth
J_Q04c1		x	Background - Age of immigration
J_Q04c1_C	x	x	Background - Age of immigration (categorised, 9 categories)
J_Q04c2		x	Background - Year of immigration
J_Q04c2_T		x	Year of immigration to country (Trend-IALS/ALL)
J_Q04c2_T1	x	x	Year of immigration to country - range of years (Trend-IALS/ALL)
J_Q04C2US_C	x	x	Background - Year of immigration (4 categories) (derived from J_Q04C2)
J_Q04dUSX1a	x	x	Background - Hispanic
J_Q04dUSX1b_01		x	Background - Hispanic origin - Mexican
J_Q04dUSX1b_02		x	Background - Hispanic origin - Puerto Rican
J_Q04dUSX1b_03		x	Background - Hispanic origin - Cuban
J_Q04dUSX1b_04		x	Background - Hispanic origin - Central/South America
J_Q04dUSX1b_05		x	Background - Hispanic origin - Other
J_Q04dUSX2_01		x	Background - Race - White
J_Q04dUSX2_02		x	Background - Race - Black
J_Q04dUSX2_03		x	Background - Race - Asian
J_Q04dUSX2_04		x	Background - Race - American Indian
J_Q04dUSX2_05		x	Background - Race - Native Hawaiian
J_Q05a1US		x	Background - First learned language
J_Q05a2US		x	Background - Second learned language
J_Q05a2USX2	x	x	Background - Age learned English
J_Q05bUS		x	Background - Language spoken at home
J_Q05cUSX1		x	Background - Language spoken most
J_Q05cUSX2	x	x	Background - English outside home
J_Q05cUSX3a	x	x	Background - Ability to understand spoken English
J_Q05cUSX3b	x	x	Background - Ability to speak English
J_Q05cUSX3d	x	x	Background - Ability to read English
J_Q05cUSX3e	x	x	Background - Ability to write English
J_Q05cUSX4	x	x	Background - ESL class/tutor in past year
J_Q05cUSX5		x	Background - Reason for ESL class/tutor
J_Q05cUSX6	x	x	Background - Class/tutor learn English as adult
J_Q06a	x	x	Background - Mother/female guardian - Whether born in country

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
J_Q06a_T	x	x	Mother or female guardian born in country (Trend-IALS/ALL)
J_Q06b	x	x	Background - Mother/female guardian - Highest level of education
J_Q06B_ISCED11	x	x	Background - Mother/female guardian - Highest level of education, ISCED 2011
J_Q06b_T	x	x	Highest level of education - mother or female guardian (Trend-IALS/ALL)
J_Q06bUS	x	x	Background - Mother/female guardian - Highest level of education
J_Q07a	x	x	Background - Father/male guardian - Whether born in country
J_Q07a_T	x	x	Father or male guardian born in country (Trend-IALS/ALL)
J_Q07b	x	x	Background - Father/male guardian - Highest level of education
J_Q07B_ISCED11	x	x	Background - Father/male guardian - Highest level of education, ISCED 2011
J_Q07b_T	x	x	Highest level of education - father or male guardian (Trend-IALS/ALL)
J_Q07bUS	x	x	Background - Father/male guardian - Highest level of education
J_Q08	x	x	Background - Number of books at home
J_S04b		x	Background - Country of birth (other)
J_S05a1		x	Background - First learned language (other)
J_S05a2		x	Background - Second learned language (other)
J_S05b		x	Background - Language spoken at home (other)
LANGUAGE	x	x	Background - English language status (derived)
LEARNATWORK	x	x	Index of learning at work (derived)
LEARNATWORK_WLE_CA	x	x	Index of learning at work, categorised WLE (derived)
LEAVEDU		x	Respondent's age when leaving formal education (derived)
LEAVEDUUS_C	x	x	Respondent's age when leaving formal education (10 categories) (derived from LEAVEDU)
LEAVER1624	x	x	Youth aged 16 to 24 who have left education without completing ISCED 3 or higher (derived)
LITSTATUS	x	x	Literacy - PV Status
LNG_BQ	x	x	Language for background questionnaire (derived, ISO 639-2/T)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
LNG_CI	x	x	Language for exercise (derived, ISO 639-2/T)
LNG_HOME		x	Language most often spoken at home - Respondent (ISO 639-2/T) (coded)
LNG_HOMEUS_C	x	x	Language most often spoken at home - Respondent (ISO 639-2/T) (3 categories) (derived from LNG_HOME)
LNG_L1		x	First language learned at home in childhood and still understood - Respondent (IS0 639-2/T) (coded)
LNG_L1US_C	x	x	First language learned at home in childhood and still understood - Respondent (IS0 639-2/T) (3 categories) (derived from LNG_L1)
LNG_L2		x	Second language learned at home in childhood and still understood - Respondent (IS0 639-2/T) (coded)
LNG_L2US_C	x	x	Second language learned at home in childhood and still understood - Respondent (IS0 639-2/T) (3 categories) (derived from LNG_L2)
M300C02S	x	x	PLC / 300 - Q5 - Employment Advertisement - Employees (Scored Response)
M301C05S	x	x	PLC / 301 - Q1 - SGIH - Tel. number (Scored Response)
M305215S	x	x	PL / 305 - Q10 - TMN AntiTheft - Documents (Scored Response)
M305218S	x	x	PL / 305 - Q11 - TMN AntiTheft - Block/Unblock (Scored Response)
M309319S	x	x	PL / 309 - Q17- Generic Medicines - Limited Use (Scored Response)
M309320S	x	x	PL / 309 - Q18 - Generic Medicines - Market share (Scored Response)
M309321S	x	x	PL / 309 - Q19 - Generic Medicines – 10% or more (Scored Response)
M309322S	x	x	PL / 309 - Q20 - Generic Medicines - Reasons (Scored Response)
M310406S	x	x	PL / 310 - Q15 - Memory Training - Brain Part (Scored Response)
M310407S	x	x	PL / 310 - Q16 - Memory Training - Discovery (Scored Response)
M313410S	x	x	PL / 313 - Q3 - International Calls - Dial 098 (Scored Response)
M313411S	x	x	PL / 313 - Q4 - International Calls - Full number (Scored Response)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
M313412S	x	x	PL / 313 - Q5 - International Calls - Country code (Scored Response)
M313413S	x	x	PL / 313 - Q6 - International Calls - Information (Scored Response)
M313414S	x	x	PL / 313 - Q7 - International Calls - Call Canada (Scored Response)
M600C04S	x	x	PNC / 600 - Q4 - Election results - Votes (Scored Response)
M602501S	x	x	PN / 602 - Q17 - Price Tag - Packed first (Scored Response)
M602502S	x	x	PN / 602 - Q18 - Price Tag - Change (Scored Response)
M602503S	x	x	PN / 602 - Q19 - Price Tag - Quarter (Scored Response)
M604505S	x	x	PN / 604 - Q14 - Gas Gauge - Gas remaining (Scored Response)
M610515S	x	x	PN / 610 - Q15 - Compound Interest - Invested (Scored Response)
M615602S	x	x	PN / 615 - Q1 - Candles - Layers (Scored Response)
M615603S	x	x	PN / 615 - Q2 - Candles - Total Weight (Scored Response)
M618607S	x	x	PN / 618 - Q12 - Six Pack - Price per can (Scored Response)
M618608S	x	x	PN / 618 - Q13 - Six Pack - Discount (Scored Response)
M620610S	x	x	PN / 620 - Q4 - Inflation - Prediction (Scored Response)
M620612S	x	x	PN / 620 - Q5 - Inflation - Actual rate (Scored Response)
M623616S	x	x	PN / 623 - Q7 - Wine - Bottles drink (Scored Response)
M623617S	x	x	PN / 623 - Q8 - Wine - Gallon (Scored Response)
M623618S	x	x	PN / 623 - Q9 - Wine - Spain / US (Scored Response)
M624619S	x	x	PN / 624 - Q10 - BMI - Weight zone (Scored Response)
M624620S	x	x	PN / 624 - Q11 - BMI - What is BMI (Scored Response)
MONTHLYINCPR	x	x	Monthly income percentile rank category (derived)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
N302C02S	x	x	PLC / 302 - Q3 - Election Results - Candidate (Scored Response)
N306110S	x	x	PL / 306 - Q1 - CANCO - Information (Scored Response)
N306111S	x	x	PL / 306 - Q2 - CANCO - Two ways (Scored Response)
NATBILANG	x	x	Has learned as a child and still understands at least two languages including test language (derived)
NATIVELANG	x	x	Test language same as native language (derived)
NATIVESPEAKER	x	x	Respondent is a native speaker (DERIVED BY CAPI)
NEET	x	x	Adults not employed at time of survey and not in education or training in 12 months preceding the survey (derived)
NFE12	x	x	Participated in nonformal education in 12 months preceding survey (derived)
NFE12JR	x	x	Participated in nonformal education for job-related reasons in 12 months preceding survey (derived)
NFE12NJR	x	x	Participated in nonformal education for non job-related reasons in 12 months preceding survey (derived)
NFEHRS	x	x	Number of hours of participation in nonformal education (derived)
NFEHRSJR	x	x	Number of hours of participation in nonformal education for job-related reasons (derived)
NFEHRSNJR	x	x	Number of hours of participation in nonformal education for nonjob-related reasons (derived)
NOPAIDWORKEVER	x	x	Adults who never had paid work including self-employment in past (derived)
NUM_ELG		x	Number of eligible persons in the household from screener
NUM_SEL		x	Number of selected persons in the household from screener
NUMHOME	x	x	Index of use of numeracy skills at home (basic and advanced - derived)
NUMHOME_WLE_CA	x	x	Index of use of numeracy skills at home (basic and advanced), categorised WLE (derived)
NUMSTATUS	x	x	Numeracy - PV Status
NUMWORK	x	x	Index of use of numeracy skills at work (basic and advanced - derived)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
NUMWORK_WLE_CA	x	x	Index of use of numeracy skills at work (basic and advanced), categorised WLE (derived)
P317001S	x	x	PL / 317 - Q12 - Apples - Evidence (Scored Response)
P317002S	x	x	PL / 317 - Q13 - Apples - Composition (Scored Response)
P317003S	x	x	PL / 317 - Q14 - Apples - Occupation (Scored Response)
P324002S	x	x	PL / 324 - Q8 - Milk Label - Safe (Scored Response)
P324003S	x	x	PL / 324 - Q9 - Milk Label - Calcium (Scored Response)
P330001S	x	x	PLC / 330 - Q2 - Guadeloupe - Falls (Scored Response)
P601C06S	x	x	PNC / 601 - Q6 - Bottles - Bottles (Scored Response)
P614601S	x	x	PNC / 614 - Q7 - Watch - Price (Scored Response)
P640001S	x	x	PN / 640 - Q3 - Odometer - Trip Miles (Scored Response)
P645001S	x	x	PNC / 645 - Q8 - AirportTimetable - Departure (Scored Response)
P655001S	x	x	PN / 655 - Q20 - Path - Length (Scored Response)
P664001S	x	x	PN / 664 - Q16 - Orchestra tickets - Student ticket (Scored Response)
P666001S	x	x	PN / 666 - Q6 - Rope - Measure (Scored Response)
P901002R	x	x	PV Q1 square (Actual Response)
P901002S	x	x	PV Q1 square (Scored Response)
P901003R	x	x	PV Q3 hand (Actual Response)
P901003S	x	x	PV Q3 hand (Scored Response)
P901004R	x	x	PV Q7 moon (Actual Response)
P901004S	x	x	PV Q7 moon (Scored Response)
P901005R	x	x	PV Q4 baby (Actual Response)
P901005S	x	x	PV Q4 baby (Scored Response)
P901006R	x	x	PV Q6 bird (Actual Response)
P901006S	x	x	PV Q6 bird (Scored Response)
P901011R	x	x	PV Q13 piano (Actual Response)
P901011S	x	x	PV Q13 piano (Scored Response)
P901013R	x	x	PV Q8 bread (Actual Response)
P901013S	x	x	PV Q8 bread (Scored Response)
P901015R	x	x	PV Q33 wrist (Actual Response)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P901015S	x	x	PV Q33 wrist (Scored Response)
P901017R	x	x	PV Q15 elephant (Actual Response)
P901017S	x	x	PV Q15 elephant (Scored Response)
P901018R	x	x	PV Q16 saw (Actual Response)
P901018S	x	x	PV Q16 saw (Scored Response)
P901019R	x	x	PV Q17 bus (Actual Response)
P901019S	x	x	PV Q17 bus (Scored Response)
P901020R	x	x	PV Q14 computer (Actual Response)
P901020S	x	x	PV Q14 computer (Scored Response)
P901021R	x	x	PV Q22 zipper (Actual Response)
P901021S	x	x	PV Q22 zipper (Scored Response)
P901024R	x	x	PV Q26 candle (Actual Response)
P901024S	x	x	PV Q26 candle (Scored Response)
P901025R	x	x	PV Q27 fountain (Actual Response)
P901025S	x	x	PV Q27 fountain (Scored Response)
P902014R	x	x	PV Q28 microscope (Actual Response)
P902014S	x	x	PV Q28 microscope (Scored Response)
P902021R	x	x	PV Q19 umbrella (Actual Response)
P902021S	x	x	PV Q19 umbrella (Scored Response)
P902022R	x	x	PV Q25 door (Actual Response)
P902022S	x	x	PV Q25 door (Scored Response)
P902024R	x	x	PV Q12 crown (Actual Response)
P902024S	x	x	PV Q12 crown (Scored Response)
P903007R	x	x	PV Q9 shovel (Actual Response)
P903007S	x	x	PV Q9 shovel (Scored Response)
P903012R	x	x	PV Q10 chess (Actual Response)
P903012S	x	x	PV Q10 chess (Scored Response)
P903017R	x	x	PV Q20 lion (Actual Response)
P903017S	x	x	PV Q20 lion (Scored Response)
P903021R	x	x	PV Q32 saddle (Actual Response)
P903021S	x	x	PV Q32 saddle (Scored Response)
P903024R	x	x	PV Q29 envelope (Actual Response)
P903024S	x	x	PV Q29 envelope (Scored Response)
P904009R	x	x	PV Q5 ruler (Actual Response)
P904009S	x	x	PV Q5 ruler (Scored Response)
P904010R	x	x	PV Q11 pants (Actual Response)
P904010S	x	x	PV Q11 pants (Scored Response)
P904012R	x	x	PV Q2 butterfly (Actual Response)
P904012S	x	x	PV Q2 butterfly (Scored Response)
P904014R	x	x	PV Q21 bottle (Actual Response)
P904014S	x	x	PV Q21 bottle (Scored Response)
P904015R	x	x	PV Q23 neck (Actual Response)
P904015S	x	x	PV Q23 neck (Scored Response)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P904020R	x	x	PV Q24 television (Actual Response)
P904020S	x	x	PV Q24 television (Scored Response)
P904021R	x	x	PV Q34 arrow (Actual Response)
P904021S	x	x	PV Q34 arrow (Scored Response)
P904022R	x	x	PV Q18 stool (Actual Response)
P904022S	x	x	PV Q18 stool (Scored Response)
P904024R	x	x	PV Q30 bell (Actual Response)
P904024S	x	x	PV Q30 bell (Scored Response)
P904025R	x	x	PV Q31 axe (Actual Response)
P904025S	x	x	PV Q31 axe (Scored Response)
P911001R	x	x	SP S1 (Actual Response)
P911001S	x	x	SP S1 (Scored Response)
P911003R	x	x	SP S3 (Actual Response)
P911003S	x	x	SP S3 (Scored Response)
P911004R	x	x	SP S4 (Actual Response)
P911004S	x	x	SP S4 (Scored Response)
P911005R	x	x	SP S5 (Actual Response)
P911005S	x	x	SP S5 (Scored Response)
P911006R	x	x	SP S6 (Actual Response)
P911006S	x	x	SP S6 (Scored Response)
P911007R	x	x	SP S7 (Actual Response)
P911007S	x	x	SP S7 (Scored Response)
P911008R	x	x	SP S8 (Actual Response)
P911008S	x	x	SP S8 (Scored Response)
P911009R	x	x	SP S9 (Actual Response)
P911009S	x	x	SP S9 (Scored Response)
P911010R	x	x	SP S10 (Actual Response)
P911010S	x	x	SP S10 (Scored Response)
P911014R	x	x	SP S15 (Actual Response)
P911014S	x	x	SP S15 (Scored Response)
P911017R	x	x	SP S19 (Actual Response)
P911017S	x	x	SP S19 (Scored Response)
P911020R	x	x	SP S22 (Actual Response)
P911020S	x	x	SP S22 (Scored Response)
P912002R	x	x	SP S2 (Actual Response)
P912002S	x	x	SP S2 (Scored Response)
P912011R	x	x	SP S12 (Actual Response)
P912011S	x	x	SP S12 (Scored Response)
P912013R	x	x	SP S14 (Actual Response)
P912013S	x	x	SP S14 (Scored Response)
P912019R	x	x	SP S21 (Actual Response)
P912019S	x	x	SP S21 (Scored Response)
P913013R	x	x	SP S11 (Actual Response)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P913013S	x	x	SP S11 (Scored Response)
P914012R	x	x	SP S13 (Actual Response)
P914012S	x	x	SP S13 (Scored Response)
P914015R	x	x	SP S16 (Actual Response)
P914015S	x	x	SP S16 (Scored Response)
P914016R	x	x	SP S18 (Actual Response)
P914016S	x	x	SP S18 (Scored Response)
P914018R	x	x	SP S20 (Actual Response)
P914018S	x	x	SP S20 (Scored Response)
P914019R	x	x	SP S17 (Actual Response)
P914019S	x	x	SP S17 (Scored Response)
P921002R	x	x	PC P1 S2 The Birthday Party (Actual Response)
P921002S	x	x	PC P1 S2 The Birthday Party (Scored Response)
P921003R	x	x	PC P1 S3 The Birthday Party (Actual Response)
P921003S	x	x	PC P1 S3 The Birthday Party (Scored Response)
P921004R	x	x	PC P1 S4 The Birthday Party (Actual Response)
P921004S	x	x	PC P1 S4 The Birthday Party (Scored Response)
P921005R	x	x	PC P1 S5 The Birthday Party (Actual Response)
P921005S	x	x	PC P1 S5 The Birthday Party (Scored Response)
P921007R	x	x	PC P1 S7 The Birthday Party (Actual Response)
P921007S	x	x	PC P1 S7 The Birthday Party (Scored Response)
P921008R	x	x	PC P1 S8 The Birthday Party (Actual Response)
P921008S	x	x	PC P1 S8 The Birthday Party (Scored Response)
P921009R	x	x	PC P1 S9 The Birthday Party (Actual Response)
P921009S	x	x	PC P1 S9 The Birthday Party (Scored Response)
P921010R	x	x	PC P1 S10 The Birthday Party (Actual Response)
P921010S	x	x	PC P1 S10 The Birthday Party (Scored Response)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P921011R	x	x	PC P1 S11 The Birthday Party (Actual Response)
P921011S	x	x	PC P1 S11 The Birthday Party (Scored Response)
P921013R	x	x	PC P1 S13 The Birthday Party (Actual Response)
P921013S	x	x	PC P1 S13 The Birthday Party (Scored Response)
P921014R	x	x	PC P1 S14 The Birthday Party (Actual Response)
P921014S	x	x	PC P1 S14 The Birthday Party (Scored Response)
P921015R	x	x	PC P1 S15 The Birthday Party (Actual Response)
P921015S	x	x	PC P1 S15 The Birthday Party (Scored Response)
P921016R	x	x	PC P1 S16 The Birthday Party (Actual Response)
P921016S	x	x	PC P1 S16 The Birthday Party (Scored Response)
P921017R	x	x	PC P1 S17 The Birthday Party (Actual Response)
P921017S	x	x	PC P1 S17 The Birthday Party (Scored Response)
P921018R	x	x	PC P1 S18 The Birthday Party (Actual Response)
P921018S	x	x	PC P1 S18 The Birthday Party (Scored Response)
P921019R	x	x	PC P1 S19 The Birthday Party (Actual Response)
P921019S	x	x	PC P1 S19 The Birthday Party (Scored Response)
P921020R	x	x	PC P1 S20 The Birthday Party (Actual Response)
P921020S	x	x	PC P1 S20 The Birthday Party (Scored Response)
P921021R	x	x	PC P1 S21 The Birthday Party (Actual Response)
P921021S	x	x	PC P1 S21 The Birthday Party (Scored Response)
P921035R	x	x	PC P3 S2 World Sports (Actual Response)
P921035S	x	x	PC P3 S2 World Sports (Scored Response)
P921036R	x	x	PC P3 S3 World Sports (Actual Response)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P921036S	x	x	PC P3 S3 World Sports (Scored Response)
P921037R	x	x	PC P3 S4 World Sports (Actual Response)
P921037S	x	x	PC P3 S4 World Sports (Scored Response)
P921038R	x	x	PC P3 S5 World Sports (Actual Response)
P921038S	x	x	PC P3 S5 World Sports (Scored Response)
P921040R	x	x	PC P3 S7 World Sports (Actual Response)
P921040S	x	x	PC P3 S7 World Sports (Scored Response)
P921041R	x	x	PC P3 S8 World Sports (Actual Response)
P921041S	x	x	PC P3 S8 World Sports (Scored Response)
P921042R	x	x	PC P3 S9 World Sports (Actual Response)
P921042S	x	x	PC P3 S9 World Sports (Scored Response)
P921043R	x	x	PC P3 S10 World Sports (Actual Response)
P921043S	x	x	PC P3 S10 World Sports (Scored Response)
P922023R	x	x	PC P2 S2 A Letter to the Editor - Clinic (Actual Response)
P922023S	x	x	PC P2 S2 A Letter to the Editor - Clinic (Scored Response)
P922024R	x	x	PC P2 S3 A Letter to the Editor - Clinic (Actual Response)
P922024S	x	x	PC P2 S3 A Letter to the Editor - Clinic (Scored Response)
P922025R	x	x	PC P2 S4 A Letter to the Editor - Clinic (Actual Response)
P922025S	x	x	PC P2 S4 A Letter to the Editor - Clinic (Scored Response)
P922026R	x	x	PC P2 S5 A Letter to the Editor - Clinic (Actual Response)
P922026S	x	x	PC P2 S5 A Letter to the Editor - Clinic (Scored Response)
P922027R	x	x	PC P2 S6 A Letter to the Editor - Clinic (Actual Response)
P922027S	x	x	PC P2 S6 A Letter to the Editor - Clinic (Scored Response)
P922028R	x	x	PC P2 S7 A Letter to the Editor - Clinic (Actual Response)
P922028S	x	x	PC P2 S7 A Letter to the Editor - Clinic (Scored Response)
P922030R	x	x	PC P2 S9 A Letter to the Editor - Clinic (Actual Response)
P922030S	x	x	PC P2 S9 A Letter to the Editor - Clinic (Scored Response)
P922031R	x	x	PC P2 S10 A Letter to the Editor - Clinic (Actual Response)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P922031S	x	x	PC P2 S10 A Letter to the Editor - Clinic (Scored Response)
P922032R	x	x	PC P2 S11 A Letter to the Editor - Clinic (Actual Response)
P922032S	x	x	PC P2 S11 A Letter to the Editor - Clinic (Scored Response)
P922033R	x	x	PC P2 S12 A Letter to the Editor - Clinic (Actual Response)
P922033S	x	x	PC P2 S12 A Letter to the Editor - Clinic (Scored Response)
P924035R	x	x	PC P4 S2 Going to the Movies (Actual Response)
P924035S	x	x	PC P4 S2 Going to the Movies (Scored Response)
P924036R	x	x	PC P4 S3 Going to the Movies (Actual Response)
P924036S	x	x	PC P4 S3 Going to the Movies (Scored Response)
P924037R	x	x	PC P4 S4 Going to the Movies (Actual Response)
P924037S	x	x	PC P4 S4 Going to the Movies (Scored Response)
P924038R	x	x	PC P4 S5 Going to the Movies (Actual Response)
P924038S	x	x	PC P4 S5 Going to the Movies (Scored Response)
P924040R	x	x	PC P4 S7 Going to the Movies (Actual Response)
P924040S	x	x	PC P4 S7 Going to the Movies (Scored Response)
P924041R	x	x	PC P4 S8 Going to the Movies (Actual Response)
P924041S	x	x	PC P4 S8 Going to the Movies (Scored Response)
P924042R	x	x	PC P4 S9 Going to the Movies (Actual Response)
P924042S	x	x	PC P4 S9 Going to the Movies (Scored Response)
P924043R	x	x	PC P4 S10 Going to the Movies (Actual Response)
P924043S	x	x	PC P4 S10 Going to the Movies (Scored Response)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
PAIDWORK12	x	x	Adults who have had paid work during the 12 months preceding the survey (derived)
PAIDWORK5	x	x	Adults who have had paid work in last 5 years (derived)
PAPER	x	x	Paper branch (derived)
PARED	x	x	Highest of mother or father's level of education (derived)
PBROUTE	x	x	Paper-based routing code (derived)
PLANNING	x	x	Index of use of planning skills at work (derived)
PLANNING_WLE_CA	x	x	Index of use of planning skills at work, categorised WLE (derived)
PPC_SCORE	x	x	Final score for the paper core assessment
PRC_PC_SCR	x	x	Total Score for Reading Components Section - Passage Comprehension (derived)
PRC_PF_Q1	x	x	Sentence Timer for Passage Comprehension items - passage 1
PRC_PF_Q2	x	x	Sentence Timer for Passage Comprehension items - passage 2
PRC_PF_Q3	x	x	Sentence Timer for Passage Comprehension items - passage 3 and 4
PRC_PV_Q1	x	x	Sentence Timer for Print Vocabulary items
PRC_PV_SCR	x	x	Total Score for Reading Components Section - Print Vocabulary (derived)
PRC_SP_Q1	x	x	Sentence Timer for Sentence Processing items
PRC_SP_SCR	x	x	Total Score for Reading Components Section - Sentence Processing (derived)
PROB_PERS		x	Person probability of selection (within HHs, if applicable)
PSLSTATUS	x	x	Problem Solving - PV Status
PVLIT1	x	x	Literacy scale score - Plausible value 1
PVLIT10	x	x	Literacy scale score - Plausible value 10
PVLIT2	x	x	Literacy scale score - Plausible value 2
PVLIT3	x	x	Literacy scale score - Plausible value 3
PVLIT4	x	x	Literacy scale score - Plausible value 4
PVLIT5	x	x	Literacy scale score - Plausible value 5
PVLIT6	x	x	Literacy scale score - Plausible value 6
PVLIT7	x	x	Literacy scale score - Plausible value 7
PVLIT8	x	x	Literacy scale score - Plausible value 8
PVLIT9	x	x	Literacy scale score - Plausible value 9
PVNUM1	x	x	Numeracy scale score - Plausible value 1
PVNUM10	x	x	Numeracy scale score - Plausible value 10
PVNUM2	x	x	Numeracy scale score - Plausible value 2

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
PVNUM3	x	x	Numeracy scale score - Plausible value 3
PVNUM4	x	x	Numeracy scale score - Plausible value 4
PVNUM5	x	x	Numeracy scale score - Plausible value 5
PVNUM6	x	x	Numeracy scale score - Plausible value 6
PVNUM7	x	x	Numeracy scale score - Plausible value 7
PVNUM8	x	x	Numeracy scale score - Plausible value 8
PVNUM9	x	x	Numeracy scale score - Plausible value 9
PVPSL1	x	x	Problem-solving scale score - Plausible value 1
PVPSL10	x	x	Problem-solving scale score - Plausible value 10
PVPSL2	x	x	Problem-solving scale score - Plausible value 2
PVPSL3	x	x	Problem-solving scale score - Plausible value 3
PVPSL4	x	x	Problem-solving scale score - Plausible value 4
PVPSL5	x	x	Problem-solving scale score - Plausible value 5
PVPSL6	x	x	Problem-solving scale score - Plausible value 6
PVPSL7	x	x	Problem-solving scale score - Plausible value 7
PVPSL8	x	x	Problem-solving scale score - Plausible value 8
PVPSL9	x	x	Problem-solving scale score - Plausible value 9
RACETHN_4CAT	x	x	Background - Race/ethnicity (derived, 4 categories)
RACETHN_5CAT	x	x	Background - Race/ethnicity (derived, 5 categories)
RACETHN_6CAT		x	Background - Race/ethnicity (derived, 6 categories)
READHOME	x	x	Index of use of reading skills at home (prose and document texts - derived)
READHOME_WLE_CA	x	x	Index of use of reading skills at home (prose and document texts), categorised WLE (derived)
READWORK	x	x	Index of use of reading skills at work (prose and document texts - derived)
READWORK_WLE_CA	x	x	Index of use of reading skills at work (prose and document texts), categorised WLE (derived)
READYTOLEARN	x	x	Index of readiness to learn (derived)
READYTOLEARN_WLE_CA	x	x	Index of readiness to learn, categorised WLE (derived)
REGION_US	x	x	Geographical region – Respondent (US Census regions)
SAMPFLAG	x	x	Flag for U.S. combined Main Study and National Supplement (Household and Prison samples)
SAMPTYPE		x	Flag for oversample

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
SECLGRGN		x	Source region of second language learned at home in childhood and still understand (9 regions - derived)
SECLGRGNUS_C	x	x	Source region of second language learned at home in childhood and still understand (9 regions) (2 categories) (derived from SECLGRGN)
SEQID	x	x	Sequential ID (randomly derived)
SORT_PSU		x	Sort order for PSU selection (or persons if one-stage design)
SORT_SSU		x	Sort order for SSU selection
SPFWT0	x	x	Final full sample weight
SPFWT1	x	x	Final replicate weight (1)
SPFWT10	x	x	Final replicate weight (10)
SPFWT11	x	x	Final replicate weight (11)
SPFWT12	x	x	Final replicate weight (12)
SPFWT13	x	x	Final replicate weight (13)
SPFWT14	x	x	Final replicate weight (14)
SPFWT15	x	x	Final replicate weight (15)
SPFWT16	x	x	Final replicate weight (16)
SPFWT17	x	x	Final replicate weight (17)
SPFWT18	x	x	Final replicate weight (18)
SPFWT19	x	x	Final replicate weight (19)
SPFWT2	x	x	Final replicate weight (2)
SPFWT20	x	x	Final replicate weight (20)
SPFWT21	x	x	Final replicate weight (21)
SPFWT22	x	x	Final replicate weight (22)
SPFWT23	x	x	Final replicate weight (23)
SPFWT24	x	x	Final replicate weight (24)
SPFWT25	x	x	Final replicate weight (25)
SPFWT26	x	x	Final replicate weight (26)
SPFWT27	x	x	Final replicate weight (27)
SPFWT28	x	x	Final replicate weight (28)
SPFWT29	x	x	Final replicate weight (29)
SPFWT3	x	x	Final replicate weight (3)
SPFWT30	x	x	Final replicate weight (30)
SPFWT31	x	x	Final replicate weight (31)
SPFWT32	x	x	Final replicate weight (32)
SPFWT33	x	x	Final replicate weight (33)
SPFWT34	x	x	Final replicate weight (34)
SPFWT35	x	x	Final replicate weight (35)
SPFWT36	x	x	Final replicate weight (36)
SPFWT37	x	x	Final replicate weight (37)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
SPFWT38	x	x	Final replicate weight (38)
SPFWT39	x	x	Final replicate weight (39)
SPFWT4	x	x	Final replicate weight (4)
SPFWT40	x	x	Final replicate weight (40)
SPFWT41	x	x	Final replicate weight (41)
SPFWT42	x	x	Final replicate weight (42)
SPFWT43	x	x	Final replicate weight (43)
SPFWT44	x	x	Final replicate weight (44)
SPFWT45	x	x	Final replicate weight (45)
SPFWT46	x	x	Final replicate weight (46)
SPFWT47	x	x	Final replicate weight (47)
SPFWT48	x	x	Final replicate weight (48)
SPFWT49	x	x	Final replicate weight (49)
SPFWT5	x	x	Final replicate weight (5)
SPFWT50	x	x	Final replicate weight (50)
SPFWT51	x	x	Final replicate weight (51)
SPFWT52	x	x	Final replicate weight (52)
SPFWT53	x	x	Final replicate weight (53)
SPFWT54	x	x	Final replicate weight (54)
SPFWT55	x	x	Final replicate weight (55)
SPFWT56	x	x	Final replicate weight (56)
SPFWT57	x	x	Final replicate weight (57)
SPFWT58	x	x	Final replicate weight (58)
SPFWT59	x	x	Final replicate weight (59)
SPFWT6	x	x	Final replicate weight (6)
SPFWT60	x	x	Final replicate weight (60)
SPFWT61	x	x	Final replicate weight (61)
SPFWT62	x	x	Final replicate weight (62)
SPFWT63	x	x	Final replicate weight (63)
SPFWT64	x	x	Final replicate weight (64)
SPFWT65	x	x	Final replicate weight (65)
SPFWT66	x	x	Final replicate weight (66)
SPFWT67	x	x	Final replicate weight (67)
SPFWT68	x	x	Final replicate weight (68)
SPFWT69	x	x	Final replicate weight (69)
SPFWT7	x	x	Final replicate weight (7)
SPFWT70	x	x	Final replicate weight (70)
SPFWT71	x	x	Final replicate weight (71)
SPFWT72	x	x	Final replicate weight (72)
SPFWT73	x	x	Final replicate weight (73)
SPFWT74	x	x	Final replicate weight (74)
SPFWT75	x	x	Final replicate weight (75)
SPFWT76	x	x	Final replicate weight (76)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
SPFWT77	x	x	Final replicate weight (77)
SPFWT78	x	x	Final replicate weight (78)
SPFWT79	x	x	Final replicate weight (79)
SPFWT8	x	x	Final replicate weight (8)
SPFWT80	x	x	Final replicate weight (80)
SPFWT9	x	x	Final replicate weight (9)
STRAT_HH		x	Explicit strata used for stratifying HHs
STRAT_PSU		x	Explicit strata used for stratifying PSUs (or persons if one-stage design)
STRAT_SSU		x	Explicit strata used for stratifying SSUs
SUBSAMP		x	Sub-sample flag
TASKDISC	x	x	Index of use of task discretion at work (derived)
TASKDISC_WLE_CA	x	x	Index of use of task discretion at work, categorised WLE (derived)
TECHPROB		x	Technical problem flag
U01a000A	x	x	Unit01a Number of Actions
U01a000F	x	x	Unit01a Time to First Action
U01a000S	x	x	Problem-solving Unit 01a (Polytomous scored response - derived)
U01a000T	x	x	Unit01a Total Time
U01b000A	x	x	Unit01b Number of Actions
U01b000F	x	x	Unit01b Time to First Action
U01b000S	x	x	Problem-solving Unit 01b (Dichotomous scored response - derived)
U01b000T	x	x	Unit01b Total Time
U02x000A	x	x	Unit02 Number of Actions
U02x000F	x	x	Unit02 Time to First Action
U02x000S	x	x	Problem-solving Unit 02x (Polytomous scored response - derived)
U02x000T	x	x	Unit02 Total Time
U03a000A	x	x	Unit03a Number of Actions
U03a000F	x	x	Unit03a Time to First Action
U03a000S	x	x	Problem-solving Unit 03a (Dichotomous scored response - derived)
U03a000T	x	x	Unit03a Total Time
U04a000A	x	x	Unit04a Number of Actions
U04a000F	x	x	Unit04a Time to First Action
U04a000S	x	x	Problem-solving Unit 04a (Polytomous scored response - derived)
U04a000T	x	x	Unit04a Total Time
U06a000A	x	x	Unit06a Number of Actions
U06a000F	x	x	Unit06a Time to First Action

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
U06a000S	x	x	Problem-solving Unit 06a (Dichotomous scored response - derived)
U06a000T	x	x	Unit06a Total Time
U06b000A	x	x	Unit06b Number of Actions
U06b000F	x	x	Unit06b Time to First Action
U06b000S	x	x	Problem-solving Unit 06b (Dichotomous scored response - derived)
U06b000T	x	x	Unit06b Total Time
U07x000A	x	x	Unit07 Number of Actions
U07x000F	x	x	Unit07 Time to First Action
U07x000S	x	x	Problem-solving Unit 07x (Dichotomous scored response - derived)
U07x000T	x	x	Unit07 Total Time
U11b000A	x	x	Unit11b Number of Actions
U11b000F	x	x	Unit11b Time to First Action
U11b000S	x	x	Problem-solving Unit 11b (Polytomous scored response - derived)
U11b000T	x	x	Unit11b Total Time
U16x000A	x	x	Unit16 Number of Actions
U16x000F	x	x	Unit16 Time to First Action
U16x000S	x	x	Problem-solving Unit 16x (Dichotomous scored response - derived)
U16x000T	x	x	Unit16 Total Time
U19a000A	x	x	Unit19a Number of Actions
U19a000F	x	x	Unit19a Time to First Action
U19a000S	x	x	Problem-solving Unit 19a (Dichotomous scored response - derived)
U19a000T	x	x	Unit19a Total Time
U19b000A	x	x	Unit19b Number of Actions
U19b000F	x	x	Unit19b Time to First Action
U19b000S	x	x	Problem-solving Unit 19b (Polytomous scored response - derived)
U19b000T	x	x	Unit19b Total Time
U21x000A	x	x	Unit21 Number of Actions
U21x000F	x	x	Unit21 Time to First Action
U21x000S	x	x	Problem-solving Unit 21x (Dichotomous scored response - derived)
U21x000T	x	x	Unit21 Total Time
U23x000A	x	x	Unit23 Number of Actions
U23x000F	x	x	Unit23 Time to First Action
U23x000S	x	x	Problem-solving Unit 23x (Polytomous scored response - derived)
U23x000T	x	x	Unit23 Total Time

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
UNEMPFLAG	x	x	Unemployment flag
URBAN_12CAT		x	Urbanicity (derived, 12 categories)
URBAN_4CAT	x	x	Urbanicity (derived, 4 categories)
USCIP_C		x	Education - Current Qualification - Area of Study (coded)
USCIP_C_C	x	x	Education - Current Qualification - Area of Study (combined into 4-digit categories) (derived from USCIP_C)
USCIP_H		x	Education - Highest qualification - Area of study (coded)
USCIP_H_C	x	x	Education - Highest qualification - Area of study (combined into 4-digit categories) (derived from USCIP_H)
USCIP_L		x	Education - Formal qualification in last 12 months- Area of study (coded)
USCIP_L_C	x	x	Education - Formal qualification in last 12 months- Area of study (combined into 4-digit categories) (derived from USCIP_L)
VARSTRAT	x	x	Variance stratum
VARUNIT	x	x	Variance unit
VEFAYFAC	x	x	Fay's K factor used in creating replicate weights (BRR only)
VEMETHOD	x	x	Replication approach (string)
VEMETHODN	x	x	Replication approach (numeric)
VENREPS	x	x	Number of replicate weights used
VET	x	x	Respondent's highest level of education obtained is vocationally oriented (derived, ISCED3 and 4 only)
WRITHOME	x	x	Index of use of writing skills at home (derived)
WRITHOME_WLE_CA	x	x	Index of use of writing skills at home, categorised WLE (derived)
WRITWORK	x	x	Index of use of writing skills at work (derived)
WRITWORK_WLE_CA	x	x	Index of use of writing skills at work, categorised WLE (derived)
YEARLYINCPR	x	x	Yearly income percentile rank category (derived)
YRSGET	x	x	Imputed years of formal education needed to get the job (self-reported - derived)
YRSQUAL	x	x	Highest level of education obtained imputed into years of education (derived)
YRSQUAL_T	x	x	Derived variable on total years of schooling during lifetime - top coded at 24 (Trend-IALS/ALL)

See note at end of table.

Table E-5. Side-by-side list of U.S. public-use file and restricted-use file variables for the household sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
ZZ1a	x	x	Observation module: Presence of additional person
ZZ1b_01	x	x	Observation module: Assistance in background questionnaire
ZZ1b_02	x	x	Observation module: Assistance in skills assessment
ZZ1cUSX		x	Second SP present
ZZ2	x	x	Observation module: Respondent understood the questions
ZZ3	x	x	Observation module: Clarification necessary
ZZ4_01	x	x	Observation module: Respondent held a conversation with someone else
ZZ4_02	x	x	Observation module: Respondent answered a phone call, text message or e-mail
ZZ4_03	x	x	Observation module: Respondent was looking after children
ZZ4_04	x	x	Observation module: Respondent was undertaking domestic tasks
ZZ4_05	x	x	Observation module: Television, radio, game console or music player was in use in the vicinity of respondent
ZZ4_06	x	x	Observation module: Respondent was interrupted by some other activity, task or event
ZZ5	x	x	Observation module: Assessment taking too long
ZZ6	x	x	Observation module: Room of assessment
ZZ7USX		x	Observed income

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample [x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
A_D01a1		x	General - Interview month (DERIVED BY CAPI)
A_D01a2		x	General - Year before interview (DERIVED BY CAPI)
A_D01a3		x	General - Interview year (DERIVED BY CAPI)
A_N01		x	General - Gender of respondent
A_Q01a		x	General - Year of birth
A_Q01b		x	General - Month of birth
ACTIVE_SECTION		x	Active section (final state on export)
AGE_R		x	Person resolved age from BQ and QC check (derived)
AGE_R_ORG		x	Person resolved age from BQ and QC check (derived, original before trimming)
AGE1634	x	x	Age 16-34 flag (derived)
AGE6674	x	x	Age 66-74 flag (derived)
AGEG10LFSEXT	x	x	Age in 10 year bands extended to include ages over 65 (derived)
AGEG5LFSEXT	x	x	Age in 5 year bands extended to include ages over 65 (derived)
B_D01d		x	Education - Highest qualification - Months elapsed since finished (DERIVED BY CAPI)
B_D01d_C	x	x	Education - Time elapsed since finished highest qualification (categorised, 5 categories)
B_Q01a	x	x	Education - Highest qualification - Level
B_Q01A_C	x	x	Education - Highest qualification – Level (3 categories) (derived from B_Q01A)
B_Q01A_ISCED11	x	x	Education - Highest qualification - Level, ISCED 2011
B_Q01a2US		x	Education - Highest qualification - Country of foreign qualification
B_Q01a3		x	Education - Highest qualification - Level of foreign qualification
B_Q01a3_C		x	Education - Highest Qualification - Level of foreign qualification (collapsed, 14 categories)
B_Q01A3_ISCED11		x	Education - Highest qualification - Level of foreign qualification, ISCED 2011
B_Q01a3US		x	Education - Highest qualification - Level of foreign qualification
B_Q01AUS_C	x	x	Education - Highest qualification – Level (3 categories) (derived from B_Q01AUS)
B_Q01aUS		x	Education - Highest qualification - Level

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
B_Q01b	x	x	Education - Highest qualification - Area of study
B_Q01bUSX		x	Education - Highest qualification - Area of study verbatim
B_Q01c1_C	x	x	Education - Highest qualification - Age of finish (categorised, 6 categories)
B_Q01c1US		x	Education - Highest qualification - Age of finish
B_Q01c2		x	Education - Highest qualification - Year of finish
B_Q01d		x	Education - Highest qualification - Month of finish
B_Q02a	x	x	Education - Current qualification
B_Q02b		x	Education - Current qualification - Level
B_Q02b_C		x	Education - Current Qualification (collapsed, 10 categories)
B_Q02B_ISCED11		x	Education - Current qualification - Level, ISCED 2011
B_Q02bUS_C	x	x	Education - Current qualification - Level (6 categories) (derived from B_Q02b)
B_Q02bUS		x	Education - Current qualification - Level
B_Q02c		x	Education - Current qualification - Area of study
B_Q02cUS_C	x	x	Education - Current qualification - Area of study (8 categories) (derived from B_Q02c)
B_Q02cUSX		x	Education - Current Qualification - Area of Study Verbatim
B_Q03a	x	x	Education - Uncompleted qualification
B_Q03A_USR		x	Education - Uncompleted qualification, corrected for U.S. routing
B_Q03b		x	Education - Uncompleted qualification - Level
B_Q03b_C		x	Education - Uncompleted qualification - Level (collapsed, 10 categories)
B_Q03B_ISCED11		x	Education - Uncompleted qualification - Level, ISCED 2011
B_Q03B_USR		x	Education - Uncompleted qualification - Level, corrected for U.S. routing
B_Q03bUS_C	x	x	Education - Uncompleted qualification - Level (6 categories) (derived from B_Q03b)
B_Q03bUS		x	Education - Uncompleted qualification - Level

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
B_Q27aUSP	x	x	Activities - Class - Class/tutor basic skills during current incarceration
B_Q27bUSP	x	x	Activities - Class - Class/tutor GED during current incarceration
B_Q27cUSP	x	x	Activities - Class - Class/tutor other equivalency during current incarceration
B_Q27eUSPa		x	Activities - Class - Class attendance, amount during current incarceration
B_Q27eUSPb		x	Activities - Class - Class attendance, unit during current incarceration
B_S01a1		x	Education - Highest qualification - Name of foreign qualification
B_S01a2		x	Education - Highest qualification - Country of foreign qualification (other)
B_S27eUSP		x	Activities - Class - Class attendance, other specify during current incarceration
BIRTHRGN		x	Country of birth (9 regions - derived)
BIRTHRGNUS_C	x	x	Country of birth (9 regions) (3 categories) (derived from BIRTHRGN)
BORNLANG	x	x	Interactions between place of birth and language status (derived)
BQLANG		x	Language for background questionnaire
C_Q07USP	x	x	Status/work history (prior to incarceration) - Subjective status
C_Q08aUSP	x	x	Current status/work history - Ever paid work
C_Q09		x	Current status/work history - Years of paid work during lifetime
C_Q09_C	x	x	Current status/work history - Years of paid work during lifetime (top-coded at 47)
C_Q10aUSP		x	Current status/work history - Last 5 years - How many different firms or organisations
C_Q10aUSP_C	x	x	Current status/work history - Last 5 years - How many different firms or organisations (top-coded at 7)
C300C02A	x	x	CLC / 300 - Employment Advertisement (Number of Actions)
C300C02F	x	x	CLC / 300 - Employment Advertisement (Timing First Action)
C300C02S	x	x	CLC / 300 - Employment Advertisement (Scored Response)
C300C02T	x	x	CLC / 300 - Employment Advertisement (Timing)
C301C05A	x	x	CLC / 301 - SGIH (Number of Actions)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C301C05F	x	x	CLC / 301 - SGIH (Timing First Action)
C301C05S	x	x	CLC / 301 - SGIH (Scored Response)
C301C05T	x	x	CLC / 301 - SGIH (Timing)
C305215A	x	x	CL / 305 - TMN AntiTheft (Number of Actions)
C305215F	x	x	CL / 305 - TMN AntiTheft (Timing First Action)
C305215S	x	x	CL / 305 - TMN AntiTheft (Scored Response)
C305215T	x	x	CL / 305 - TMN AntiTheft (Timing)
C305218A	x	x	CL / 305 - TMN AntiTheft (Number of Actions)
C305218F	x	x	CL / 305 - TMN AntiTheft (Timing First Action)
C305218S	x	x	CL / 305 - TMN AntiTheft (Scored Response)
C305218T	x	x	CL / 305 - TMN AntiTheft (Timing)
C308116A	x	x	CL / 308 - Baltic Stock Market (Number of Actions)
C308116F	x	x	CL / 308 - Baltic Stock Market (Timing First Action)
C308116S	x	x	CL / 308 - Baltic Stock Market (Scored Response)
C308116T	x	x	CL / 308 - Baltic Stock Market (Timing)
C308117A	x	x	CL / 308 - Baltic Stock Market (Number of Actions)
C308117F	x	x	CL / 308 - Baltic Stock Market (Timing First Action)
C308117S	x	x	CL / 308 - Baltic Stock Market (Scored Response)
C308117T	x	x	CL / 308 - Baltic Stock Market (Timing)
C308118A	x	x	CL / 308 - Baltic Stock Market (Number of Actions)
C308118F	x	x	CL / 308 - Baltic Stock Market (Timing First Action)
C308118S	x	x	CL / 308 - Baltic Stock Market (Scored Response)
C308118T	x	x	CL / 308 - Baltic Stock Market (Timing)
C308119A	x	x	CL / 308 - Baltic Stock Market (Number of Actions)
C308119F	x	x	CL / 308 - Baltic Stock Market (Timing First Action)
C308119S	x	x	CL / 308 - Baltic Stock Market (Scored Response)
C308119T	x	x	CL / 308 - Baltic Stock Market (Timing)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C308120A	x	x	CL / 308 - Baltic Stock Market (Number of Actions)
C308120F	x	x	CL / 308 - Baltic Stock Market (Timing First Action)
C308120S	x	x	CL / 308 - Baltic Stock Market (Scored Response)
C308120T	x	x	CL / 308 - Baltic Stock Market (Timing)
C308121A	x	x	CL / 308 - Baltic Stock Market (Number of Actions)
C308121F	x	x	CL / 308 - Baltic Stock Market (Timing First Action)
C308121S	x	x	CL / 308 - Baltic Stock Market (Scored Response)
C308121T	x	x	CL / 308 - Baltic Stock Market (Timing)
C309319A	x	x	CL / 309 - Generic Medicines (Number of Actions)
C309319F	x	x	CL / 309 - Generic Medicines (Timing First Action)
C309319S	x	x	CL / 309 - Generic Medicines (Scored Response)
C309319T	x	x	CL / 309 - Generic Medicines (Timing)
C309320A	x	x	CL / 309 - Generic Medicines (Number of Actions)
C309320F	x	x	CL / 309 - Generic Medicines (Timing First Action)
C309320S	x	x	CL / 309 - Generic Medicines (Scored Response)
C309320T	x	x	CL / 309 - Generic Medicines (Timing)
C309321A	x	x	CL / 309 - Generic Medicines (Number of Actions)
C309321F	x	x	CL / 309 - Generic Medicines (Timing First Action)
C309321S	x	x	CL / 309 - Generic Medicines (Scored Response)
C309321T	x	x	CL / 309 - Generic Medicines (Timing)
C309322A	x	x	CL / 309 - Generic Medicines (Number of Actions)
C309322F	x	x	CL / 309 - Generic Medicines (Timing First Action)
C309322S	x	x	CL / 309 - Generic Medicines (Scored Response)
C309322T	x	x	CL / 309 - Generic Medicines (Timing)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C310406A	x	x	CL / 310 - Memory Training (Number of Actions)
C310406F	x	x	CL / 310 - Memory Training (Timing First Action)
C310406S	x	x	CL / 310 - Memory Training (Scored Response)
C310406T	x	x	CL / 310 - Memory Training (Timing)
C310407A	x	x	CL / 310 - Memory Training (Number of Actions)
C310407F	x	x	CL / 310 - Memory Training (Timing First Action)
C310407S	x	x	CL / 310 - Memory Training (Scored Response)
C310407T	x	x	CL / 310 - Memory Training (Timing)
C313410A	x	x	CL / 313 - International Calls (Number of Actions)
C313410F	x	x	CL / 313 - International Calls (Timing First Action)
C313410S	x	x	CL / 313 - International Calls (Scored Response)
C313410T	x	x	CL / 313 - International Calls (Timing)
C313411A	x	x	CL / 313 - International Calls (Number of Actions)
C313411F	x	x	CL / 313 - International Calls (Timing First Action)
C313411S	x	x	CL / 313 - International Calls (Scored Response)
C313411T	x	x	CL / 313 - International Calls (Timing)
C313412A	x	x	CL / 313 - International Calls (Number of Actions)
C313412F	x	x	CL / 313 - International Calls (Timing First Action)
C313412S	x	x	CL / 313 - International Calls (Scored Response)
C313412T	x	x	CL / 313 - International Calls (Timing)
C313413A	x	x	CL / 313 - International Calls (Number of Actions)
C313413F	x	x	CL / 313 - International Calls (Timing First Action)
C313413S	x	x	CL / 313 - International Calls (Scored Response)
C313413T	x	x	CL / 313 - International Calls (Timing)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C313414A	x	x	CL / 313 - International Calls (Number of Actions)
C313414F	x	x	CL / 313 - International Calls (Timing First Action)
C313414S	x	x	CL / 313 - International Calls (Scored Response)
C313414T	x	x	CL / 313 - International Calls (Timing)
C600C04A	x	x	CNC / 600 - Election results (Number of Actions)
C600C04F	x	x	CNC / 600 - Election results (Timing First Action)
C600C04S	x	x	CNC / 600 - Election results (Scored Response)
C600C04T	x	x	CNC / 600 - Election results (Timing)
C601C06A	x	x	CNC / 601 - Bottles (Number of Actions)
C601C06F	x	x	CNC / 601 - Bottles (Timing First Action)
C601C06S	x	x	CNC / 601 - Bottles (Scored Response)
C601C06T	x	x	CNC / 601 - Bottles (Timing)
C602501A	x	x	CN / 602 - Price Tags (Number of Actions)
C602501F	x	x	CN / 602 - Price Tags (Timing First Action)
C602501S	x	x	CN / 602 - Price Tags (Scored Response)
C602501T	x	x	CN / 602 - Price Tags (Timing)
C602502A	x	x	CN / 602 - Price Tags (Number of Actions)
C602502F	x	x	CN / 602 - Price Tags (Timing First Action)
C602502S	x	x	CN / 602 - Price Tags (Scored Response)
C602502T	x	x	CN / 602 - Price Tags (Timing)
C602503A	x	x	CN / 602 - Price Tags (Number of Actions)
C602503F	x	x	CN / 602 - Price Tags (Timing First Action)
C602503S	x	x	CN / 602 - Price Tags (Scored Response)
C602503T	x	x	CN / 602 - Price Tags (Timing)
C604505A	x	x	CN / 604 - Gas Gauge (Number of Actions)
C604505F	x	x	CN / 604 - Gas Gauge (Timing First Action)
C604505S	x	x	CN / 604 - Gas Gauge (Scored Response)
C604505T	x	x	CN / 604 - Gas Gauge (Timing)
C605506A	x	x	CN / 605 - Photo (Number of Actions)
C605506F	x	x	CN / 605 - Photo (Timing First Action)
C605506S	x	x	CN / 605 - Photo (Scored Response)
C605506T	x	x	CN / 605 - Photo (Timing)
C605507A	x	x	CN / 605 - Photo (Number of Actions)
C605507F	x	x	CN / 605 - Photo (Timing First Action)
C605507S	x	x	CN / 605 - Photo (Scored Response)
C605507T	x	x	CN / 605 - Photo (Timing)
C605508A	x	x	CN / 605 - Photo (Number of Actions)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C605508F	x	x	CN / 605 - Photo (Timing First Action)
C605508S	x	x	CN / 605 - Photo (Scored Response)
C605508T	x	x	CN / 605 - Photo (Timing)
C606509A	x	x	CN / 606 - Solution (Number of Actions)
C606509F	x	x	CN / 606 - Solution (Timing First Action)
C606509S	x	x	CN / 606 - Solution (Scored Response)
C606509T	x	x	CN / 606 - Solution (Timing)
C607510A	x	x	CN / 607 - TV (Number of Actions)
C607510F	x	x	CN / 607 - TV (Timing First Action)
C607510S	x	x	CN / 607 - TV (Scored Response)
C607510T	x	x	CN / 607 - TV (Timing)
C608513A	x	x	CN / 608 - Tree (Number of Actions)
C608513F	x	x	CN / 608 - Tree (Timing First Action)
C608513S	x	x	CN / 608 - Tree (Scored Response)
C608513T	x	x	CN / 608 - Tree (Timing)
C611516A	x	x	CN / 611 - Temp Scale (Number of Actions)
C611516F	x	x	CN / 611 - Temp Scale (Timing First Action)
C611516S	x	x	CN / 611 - Temp Scale (Scored Response)
C611516T	x	x	CN / 611 - Temp Scale (Timing)
C611517A	x	x	CN / 611 - Temp Scale (Number of Actions)
C611517F	x	x	CN / 611 - Temp Scale (Timing First Action)
C611517S	x	x	CN / 611 - Temp Scale (Scored Response)
C611517T	x	x	CN / 611 - Temp Scale (Timing)
C612518A	x	x	CN / 612 - Dioxin (Number of Actions)
C612518F	x	x	CN / 612 - Dioxin (Timing First Action)
C612518S	x	x	CN / 612 - Dioxin (Scored Response)
C612518T	x	x	CN / 612 - Dioxin (Timing)
C613520A	x	x	CN / 613 - Logbook (Number of Actions)
C613520F	x	x	CN / 613 - Logbook (Timing First Action)
C613520S	x	x	CN / 613 - Logbook (Scored Response)
C613520T	x	x	CN / 613 - Logbook (Timing)
C614601A	x	x	CN / 614 - Watch (Number of Actions)
C614601F	x	x	CN / 614 - Watch (Timing First Action)
C614601S	x	x	CN / 614 - Watch (Scored Response)
C614601T	x	x	CN / 614 - Watch (Timing)
C615602A	x	x	CN / 615 - Candles (Number of Actions)
C615602F	x	x	CN / 615 - Candles (Timing First Action)
C615602S	x	x	CN / 615 - Candles (Scored Response)
C615602T	x	x	CN / 615 - Candles (Timing)
C615603A	x	x	CN / 615 - Candles (Number of Actions)
C615603F	x	x	CN / 615 - Candles (Timing First Action)
C615603S	x	x	CN / 615 - Candles (Scored Response)
C615603T	x	x	CN / 615 - Candles (Timing)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C617605A	x	x	CN / 617 - Map (Number of Actions)
C617605F	x	x	CN / 617 - Map (Timing First Action)
C617605S	x	x	CN / 617 - Map (Scored Response)
C617605T	x	x	CN / 617 - Map (Timing)
C617606A	x	x	CN / 617 - Map (Number of Actions)
C617606F	x	x	CN / 617 - Map (Timing First Action)
C617606S	x	x	CN / 617 - Map (Scored Response)
C617606T	x	x	CN / 617 - Map (Timing)
C618607A	x	x	CN / 618 - Six Pack (Number of Actions)
C618607F	x	x	CN / 618 - Six Pack (Timing First Action)
C618607S	x	x	CN / 618 - Six Pack (Scored Response)
C618607T	x	x	CN / 618 - Six Pack (Timing)
C618608A	x	x	CN / 618 - Six Pack (Number of Actions)
C618608F	x	x	CN / 618 - Six Pack (Timing First Action)
C618608S	x	x	CN / 618 - Six Pack (Scored Response)
C618608T	x	x	CN / 618 - Six Pack (Timing)
C619609A	x	x	CN / 619 - Tiles (Number of Actions)
C619609F	x	x	CN / 619 - Tiles (Timing First Action)
C619609S	x	x	CN / 619 - Tiles (Scored Response)
C619609T	x	x	CN / 619 - Tiles (Timing)
C620610A	x	x	CN / 620 - Inflation (Number of Actions)
C620610F	x	x	CN / 620 - Inflation (Timing First Action)
C620610S	x	x	CN / 620 - Inflation (Scored Response)
C620610T	x	x	CN / 620 - Inflation (Timing)
C620612A	x	x	CN / 620 - Inflation (Number of Actions)
C620612F	x	x	CN / 620 - Inflation (Timing First Action)
C620612S	x	x	CN / 620 - Inflation (Scored Response)
C620612T	x	x	CN / 620 - Inflation (Timing)
C622615A	x	x	CN / 622 - Classified (Number of Actions)
C622615F	x	x	CN / 622 - Classified (Timing First Action)
C622615S	x	x	CN / 622 - Classified (Scored Response)
C622615T	x	x	CN / 622 - Classified (Timing)
C623616A	x	x	CN / 623 - Wine (Number of Actions)
C623616F	x	x	CN / 623 - Wine (Timing First Action)
C623616S	x	x	CN / 623 - Wine (Scored Response)
C623616T	x	x	CN / 623 - Wine (Timing)
C623617A	x	x	CN / 623 - Wine (Number of Actions)
C623617F	x	x	CN / 623 - Wine (Timing First Action)
C623617S	x	x	CN / 623 - Wine (Scored Response)
C623617T	x	x	CN / 623 - Wine (Timing)
C624619A	x	x	CN / 624 - BMI (Number of Actions)
C624619F	x	x	CN / 624 - BMI (Timing First Action)
C624619S	x	x	CN / 624 - BMI (Scored Response)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
C624619T	x	x	CN / 624 - BMI (Timing)
C624620A	x	x	CN / 624 - BMI (Number of Actions)
C624620F	x	x	CN / 624 - BMI (Timing First Action)
C624620S	x	x	CN / 624 - BMI (Scored Response)
C624620T	x	x	CN / 624 - BMI (Timing)
CASEID		x	Household operational ID
CBA_CORE_STAGE1_SCORE	x	x	CBA Core score for stage 1
CBA_CORE_STAGE2_SCORE	x	x	CBA Core score for stage 2
CBA_START	x	x	Computer-based exercise agreement
CBAMOD1	x	x	CBA module 1 branch (derived)
CBAMOD1STG1	x	x	CBA module 1, stage 1 branch (derived)
CBAMOD1STG2	x	x	CBA module 1, stage 2 branch (derived)
CBAMOD2	x	x	CBA module 2 branch (derived)
CBAMOD2ALT	x	x	CBA module 1 and 2 branch (derived)
CBAMOD2STG1	x	x	CBA module 2, stage 1 branch (derived)
CBAMOD2STG2	x	x	CBA module 2, stage 2 branch (derived)
CILANG	x	x	Language for exercise
CNT_BRTH		x	Country of birth - Respondent (UN M49 numerical) (coded)
CNT_BRTHUS_C	x	x	Country of birth - Respondent (UN M49 numerical) (2 categories) (derived from CNT_BRTH)
CNT_H		x	Country in which highest qualification was gained - Respondent (UN M49 numerical) (coded)
CNTRY	x	x	Country ID and sub-national entity sample code (string)
CNTRY_E	x	x	Participating country or sub-national entity code (string)
CNTRYID	x	x	Country ID (ISO 3166, numeric)
CNTRYID_E	x	x	Participating country or sub-national entity code (numeric)
COMPUTEREXPERIENCEUSP	x	x	Respondent experience with computer (including current prison job) (DERIVED BY CAPI)
CORESTAGE1_PASS	x	x	Core Stage 1 pass status
CORESTAGE2_PASS	x	x	Core Stage 2 pass status
CTRYQUAL		x	Country where highest qualification obtained (9 regions - derived)
CTRYRGN	x	x	Country region (9 regions)
D_Q01aUSP		x	Prison job - Job title
D_Q01bUSP		x	Prison job - Responsibilities
D_Q10USP	x	x	Prison job - Hours/week

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
D_Q10USP_C	x	x	Prison job - Hours/week (top-coded at 60)
D302C02A	x	x	CLC / 302 - Election Results (Number of Actions)
D302C02F	x	x	CLC / 302 - Election Results (Timing First Action)
D302C02S	x	x	CLC / 302 - Election Results (Scored Response)
D302C02T	x	x	CLC / 302 - Election Results (Timing)
D304710A	x	x	CL / 304 - Contact Employer (Number of Actions)
D304710F	x	x	CL / 304 - Contact Employer (Timing First Action)
D304710S	x	x	CL / 304 - Contact Employer (Scored Response)
D304710T	x	x	CL / 304 - Contact Employer (Timing)
D304711A	x	x	CL / 304 - Contact Employer (Number of Actions)
D304711F	x	x	CL / 304 - Contact Employer (Timing First Action)
D304711S	x	x	CL / 304 - Contact Employer (Scored Response)
D304711T	x	x	CL / 304 - Contact Employer (Timing)
D306110A	x	x	CL / 306 - Canco (Number of Actions)
D306110F	x	x	CL / 306 - Canco (Timing First Action)
D306110S	x	x	CL / 306 - Canco (Scored Response)
D306110T	x	x	CL / 306 - Canco (Timing)
D306111A	x	x	CL / 306 - Canco (Number of Actions)
D306111F	x	x	CL / 306 - Canco (Timing First Action)
D306111S	x	x	CL / 306 - Canco (Scored Response)
D306111T	x	x	CL / 306 - Canco (Timing)
D307401A	x	x	CL / 307 - MEDCO Aspirin (Number of Actions)
D307401F	x	x	CL / 307 - MEDCO Aspirin (Timing First Action)
D307401S	x	x	CL / 307 - MEDCO Aspirin (Scored Response)
D307401T	x	x	CL / 307 - MEDCO Aspirin (Timing)
D307402A	x	x	CL / 307 - MEDCO Aspirin (Number of Actions)
D307402F	x	x	CL / 307 - MEDCO Aspirin (Timing First Action)
D307402S	x	x	CL / 307 - MEDCO Aspirin (Scored Response)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
D307402T	x	x	CL / 307 - MEDCO Aspirin (Timing)
D311701A	x	x	CL / 311 - Dutch Women (Number of Actions)
D311701F	x	x	CL / 311 - Dutch Women (Timing First Action)
D311701S	x	x	CL / 311 - Dutch Women (Scored Response)
D311701T	x	x	CL / 311 - Dutch Women (Timing)
D315512A	x	x	CL / 315 - Distances-Mexican Cities (Number of Actions)
D315512F	x	x	CL / 315 - Distances-Mexican Cities (Timing First Action)
D315512S	x	x	CL / 315 - Distances-Mexican Cities (Scored Response)
D315512T	x	x	CL / 315 - Distances-Mexican Cities (Timing)
DISP_BQ	x	x	Final disposition code for BQ/JRA
DISP_CIBQ	x	x	Final disposition code for person - combining CI and BQ/JRA (derived)
DOBMM		x	Date of birth (derived from BQ)
DOBYY		x	Date of birth year (derived from BQ)
E_Q01aUSP		x	Last job (prior to incarceration) - Job title
E_Q01bUSP		x	Last job (prior to incarceration) - Responsibilities
E_Q02aUSP		x	Last job (prior to incarceration) - Kind of business, industry or service
E_Q02bUSP		x	Last job (prior to incarceration) - Main product of firm or organisation
E_Q03US_P	x	x	Last job (prior to incarceration) - Economic sector
E_Q03USP	x	x	Last job (prior to incarceration) - Economic sector
E_Q04USP	x	x	Last job (prior to incarceration) - Employee or self-employed
E_Q05a1USP		x	Last job (prior to incarceration) - Start of work for employer - Age
E_Q05a1USP_C	x	x	Last job (prior to incarceration) - Start of work for employer - Age (categorised, 9 categories)
E_Q05a2USP		x	Last job (prior to incarceration) - Start of work for employer - Year
E_Q05A2USP_C	x	x	Last job (prior to incarceration) - Start of work for employer - Year (4 categories) (derived from E_Q05A2USP)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
E_Q05b1USP		x	Last job (prior to incarceration) - Start of work for business - Age
E_Q05b1USP_C	x	x	Last job (prior to incarceration) - Start of work for business - Age (categorised, 9 categories)
E_Q05b2USP		x	Last job (prior to incarceration) - Start of work for business - Year
E_Q05B2USP_C	x	x	Last job (prior to incarceration) - Start of work for business - Year (2 categories) (derived from E_Q05B2USP)
E_Q06USP	x	x	Last job (prior to incarceration) - Amount of people working for employer
E_Q07aUSP		x	Last job (prior to incarceration) - Employees working for you
E_Q07bUSP		x	Last job (prior to incarceration) - Employees working for you - Count
E_Q08USP		x	Last job (prior to incarceration) - Type of contract
E_Q08USP_C	x	x	Last job (prior to incarceration) - Type of contract (5 categories) (derived from E_Q08USP)
E_Q09USP	x	x	Last job (prior to incarceration) - Hours/week
E_Q09USP_C	x	x	Last job (prior to incarceration) - Hours/week (top-coded at 60)
E_Q10USP	x	x	Last job (prior to incarceration) - Reason for end of job
E_S08USP		x	Last job (prior to incarceration) - Other type of contract specified
E318001A	x	x	CL / 318 - Civil Engineering (Number of Actions)
E318001F	x	x	CL / 318 - Civil Engineering (Timing First Action)
E318001S	x	x	CL / 318 - Civil Engineering (Scored Response)
E318001T	x	x	CL / 318 - Civil Engineering (Timing)
E318003A	x	x	CL / 318 - Civil Engineering (Number of Actions)
E318003F	x	x	CL / 318 - Civil Engineering (Timing First Action)
E318003S	x	x	CL / 318 - Civil Engineering (Scored Response)
E318003T	x	x	CL / 318 - Civil Engineering (Timing)
E320001A	x	x	CL / 320 - Discussion forum (Number of Actions)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
E320001F	x	x	CL / 320 - Discussion forum (Timing First Action)
E320001S	x	x	CL / 320 - Discussion forum (Scored Response)
E320001T	x	x	CL / 320 - Discussion forum (Timing)
E320003A	x	x	CL / 320 - Discussion forum (Number of Actions)
E320003F	x	x	CL / 320 - Discussion forum (Timing First Action)
E320003S	x	x	CL / 320 - Discussion forum (Scored Response)
E320003T	x	x	CL / 320 - Discussion forum (Timing)
E320004A	x	x	CL / 320 - Discussion forum (Number of Actions)
E320004F	x	x	CL / 320 - Discussion forum (Timing First Action)
E320004S	x	x	CL / 320 - Discussion forum (Scored Response)
E320004T	x	x	CL / 320 - Discussion forum (Timing)
E321001A	x	x	CL / 321 - Internet Poll (Number of Actions)
E321001F	x	x	CL / 321 - Internet Poll (Timing First Action)
E321001S	x	x	CL / 321 - Internet Poll (Scored Response)
E321001T	x	x	CL / 321 - Internet Poll (Timing)
E321002A	x	x	CL / 321 - Internet Poll (Number of Actions)
E321002F	x	x	CL / 321 - Internet Poll (Timing First Action)
E321002S	x	x	CL / 321 - Internet Poll (Scored Response)
E321002T	x	x	CL / 321 - Internet Poll (Timing)
E322001A	x	x	CL / 322 - Lakeside Fun Run (Number of Actions)
E322001F	x	x	CL / 322 - Lakeside Fun Run (Timing First Action)
E322001S	x	x	CL / 322 - Lakeside Fun Run (Scored Response)
E322001T	x	x	CL / 322 - Lakeside Fun Run (Timing)
E322002A	x	x	CL / 322 - Lakeside Fun Run (Number of Actions)
E322002F	x	x	CL / 322 - Lakeside Fun Run (Timing First Action)
E322002S	x	x	CL / 322 - Lakeside Fun Run (Scored Response)
E322002T	x	x	CL / 322 - Lakeside Fun Run (Timing)
E322003A	x	x	CL / 322 - Lakeside Fun Run (Number of Actions)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
E322003F	x	x	CL / 322 - Lakeside Fun Run (Timing First Action)
E322003S	x	x	CL / 322 - Lakeside Fun Run (Scored Response)
E322003T	x	x	CL / 322 - Lakeside Fun Run (Timing)
E322004A	x	x	CL / 322 - Lakeside Fun Run (Number of Actions)
E322004F	x	x	CL / 322 - Lakeside Fun Run (Timing First Action)
E322004S	x	x	CL / 322 - Lakeside Fun Run (Scored Response)
E322004T	x	x	CL / 322 - Lakeside Fun Run (Timing)
E322005A	x	x	CL / 322 - Lakeside Fun Run (Number of Actions)
E322005F	x	x	CL / 322 - Lakeside Fun Run (Timing First Action)
E322005S	x	x	CL / 322 - Lakeside Fun Run (Scored Response)
E322005T	x	x	CL / 322 - Lakeside Fun Run (Timing)
E323002A	x	x	CL / 323 - Library Search (Number of Actions)
E323002F	x	x	CL / 323 - Library Search (Timing First Action)
E323002S	x	x	CL / 323 - Library Search (Scored Response)
E323002T	x	x	CL / 323 - Library Search (Timing)
E323003A	x	x	CL / 323 - Library Search (Number of Actions)
E323003F	x	x	CL / 323 - Library Search (Timing First Action)
E323003S	x	x	CL / 323 - Library Search (Scored Response)
E323003T	x	x	CL / 323 - Library Search (Timing)
E323004A	x	x	CL / 323 - Library Search (Number of Actions)
E323004F	x	x	CL / 323 - Library Search (Timing First Action)
E323004S	x	x	CL / 323 - Library Search (Scored Response)
E323004T	x	x	CL / 323 - Library Search (Timing)
E323005A	x	x	CL / 323 - Library Search (Number of Actions)
E323005F	x	x	CL / 323 - Library Search (Timing First Action)
E323005S	x	x	CL / 323 - Library Search (Scored Response)
E323005T	x	x	CL / 323 - Library Search (Timing)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
E327001A	x	x	CL / 327 - Summer Streets (Number of Actions)
E327001F	x	x	CL / 327 - Summer Streets (Timing First Action)
E327001S	x	x	CL / 327 - Summer Streets (Scored Response)
E327001T	x	x	CL / 327 - Summer Streets (Timing)
E327002A	x	x	CL / 327 - Summer Streets (Number of Actions)
E327002F	x	x	CL / 327 - Summer Streets (Timing First Action)
E327002S	x	x	CL / 327 - Summer Streets (Scored Response)
E327002T	x	x	CL / 327 - Summer Streets (Timing)
E327003A	x	x	CL / 327 - Summer Streets (Number of Actions)
E327003F	x	x	CL / 327 - Summer Streets (Timing First Action)
E327003S	x	x	CL / 327 - Summer Streets (Scored Response)
E327003T	x	x	CL / 327 - Summer Streets (Timing)
E327004A	x	x	CL / 327 - Summer Streets (Number of Actions)
E327004F	x	x	CL / 327 - Summer Streets (Timing First Action)
E327004S	x	x	CL / 327 - Summer Streets (Scored Response)
E327004T	x	x	CL / 327 - Summer Streets (Timing)
E329002A	x	x	CL / 329 - Work-related Stress (Number of Actions)
E329002F	x	x	CL / 329 - Work-related Stress (Timing First Action)
E329002S	x	x	CL / 329 - Work-related Stress (Scored Response)
E329002T	x	x	CL / 329 - Work-related Stress (Timing)
E329003A	x	x	CL / 329 - Work-related Stress (Number of Actions)
E329003F	x	x	CL / 329 - Work-related Stress (Timing First Action)
E329003S	x	x	CL / 329 - Work-related Stress (Scored Response)
E329003T	x	x	CL / 329 - Work-related Stress (Timing)
E632001A	x	x	CN / 632 - Educational level (Number of Actions)
E632001F	x	x	CN / 632 - Educational level (Timing First Action)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
E632001S	x	x	CN / 632 - Educational level (Scored Response)
E632001T	x	x	CN / 632 - Educational level (Timing)
E632002A	x	x	CN / 632 - Educational level (Number of Actions)
E632002F	x	x	CN / 632 - Educational level (Timing First Action)
E632002S	x	x	CN / 632 - Educational level (Scored Response)
E632002T	x	x	CN / 632 - Educational level (Timing)
E634001A	x	x	CN / 634 - Peanuts (Number of Actions)
E634001F	x	x	CN / 634 - Peanuts (Timing First Action)
E634001S	x	x	CN / 634 - Peanuts (Scored Response)
E634001T	x	x	CN / 634 - Peanuts (Timing)
E634002A	x	x	CN / 634 - Peanuts (Number of Actions)
E634002F	x	x	CN / 634 - Peanuts (Timing First Action)
E634002S	x	x	CN / 634 - Peanuts (Scored Response)
E634002T	x	x	CN / 634 - Peanuts (Timing)
E635001A	x	x	CN / 635 - Parking Map (Number of Actions)
E635001F	x	x	CN / 635 - Parking Map (Timing First Action)
E635001S	x	x	CN / 635 - Parking Map (Scored Response)
E635001T	x	x	CN / 635 - Parking Map (Timing)
E636001A	x	x	CN / 636 - Lab Report (Number of Actions)
E636001F	x	x	CN / 636 - Lab Report (Timing First Action)
E636001S	x	x	CN / 636 - Lab Report (Scored Response)
E636001T	x	x	CN / 636 - Lab Report (Timing)
E641001A	x	x	CN / 641 - Amoeba (Number of Actions)
E641001F	x	x	CN / 641 - Amoeba (Timing First Action)
E641001S	x	x	CN / 641 - Amoeba (Scored Response)
E641001T	x	x	CN / 641 - Amoeba (Timing)
E644002A	x	x	CN / 644 - NZExports (Number of Actions)
E644002F	x	x	CN / 644 - NZExports (Timing First Action)
E644002S	x	x	CN / 644 - NZExports (Scored Response)
E644002T	x	x	CN / 644 - NZExports (Timing)
E645001A	x	x	CNC / 645 - Airport Timetable (Number of Actions)
E645001F	x	x	CNC / 645 - Airport Timetable (Timing First Action)
E645001S	x	x	CNC / 645 - Airport Timetable (Scored Response)
E645001T	x	x	CNC / 645 - Airport Timetable (Timing)
E646002A	x	x	CN / 646 - Rug Production (Number of Actions)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
E646002F	x	x	CN / 646 - Rug Production (Timing First Action)
E646002S	x	x	CN / 646 - Rug Production (Scored Response)
E646002T	x	x	CN / 646 - Rug Production (Timing)
E650001A	x	x	CN / 650 - Urban Population (Number of Actions)
E650001F	x	x	CN / 650 - Urban Population (Timing First Action)
E650001S	x	x	CN / 650 - Urban Population (Scored Response)
E650001T	x	x	CN / 650 - Urban Population (Timing)
E651002A	x	x	CN / 651 - Fertilizer (Number of Actions)
E651002F	x	x	CN / 651 - Fertilizer (Timing First Action)
E651002S	x	x	CN / 651 - Fertilizer (Scored Response)
E651002T	x	x	CN / 651 - Fertilizer (Timing)
E655001A	x	x	CN / 655 - Path (Number of Actions)
E655001F	x	x	CN / 655 - Path (Timing First Action)
E655001S	x	x	CN / 655 - Path (Scored Response)
E655001T	x	x	CN / 655 - Path (Timing)
E657001A	x	x	CN / 657 - Package (Number of Actions)
E657001F	x	x	CN / 657 - Package (Timing First Action)
E657001S	x	x	CN / 657 - Package (Scored Response)
E657001T	x	x	CN / 657 - Package (Timing)
E660003A	x	x	CN / 660 - Weight history (Number of Actions)
E660003F	x	x	CN / 660 - Weight history (Timing First Action)
E660003S	x	x	CN / 660 - Weight history (Scored Response)
E660003T	x	x	CN / 660 - Weight history (Timing)
E660004A	x	x	CN / 660 - Weight history (Number of Actions)
E660004F	x	x	CN / 660 - Weight history (Timing First Action)
E660004S	x	x	CN / 660 - Weight history (Scored Response)
E660004T	x	x	CN / 660 - Weight history (Timing)
E661001A	x	x	CN / 661 - Study fees (Number of Actions)
E661001F	x	x	CN / 661 - Study fees (Timing First Action)
E661001S	x	x	CN / 661 - Study fees (Scored Response)
E661001T	x	x	CN / 661 - Study fees (Timing)
E661002A	x	x	CN / 661 - Study fees (Number of Actions)
E661002F	x	x	CN / 661 - Study fees (Timing First Action)
E661002S	x	x	CN / 661 - Study fees (Scored Response)
E661002T	x	x	CN / 661 - Study fees (Timing)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
E664001A	x	x	CN / 664 - Orchestra tickets (Number of Actions)
E664001F	x	x	CN / 664 - Orchestra tickets (Timing First Action)
E664001S	x	x	CN / 664 - Orchestra tickets (Scored Response)
E664001T	x	x	CN / 664 - Orchestra tickets (Timing)
E665001A	x	x	CN / 665 - Cooper test (Number of Actions)
E665001F	x	x	CN / 665 - Cooper test (Timing First Action)
E665001S	x	x	CN / 665 - Cooper test (Scored Response)
E665001T	x	x	CN / 665 - Cooper test (Timing)
E665002A	x	x	CN / 665 - Cooper test (Number of Actions)
E665002F	x	x	CN / 665 - Cooper test (Timing First Action)
E665002S	x	x	CN / 665 - Cooper test (Scored Response)
E665002T	x	x	CN / 665 - Cooper test (Timing)
EDCAT6	x	x	Highest level of formal education obtained (6 categories - derived)
EDCAT7	x	x	Highest level of formal education obtained (7 categories - derived)
EDCAT8	x	x	Highest level of formal education obtained (8 categories - derived)
EDLEVEL3	x	x	Educational level of the respondent (DERIVED BY CAPI)
EXCFRM_PROP		x	Proportion in target population who are excluded from the sampling frame
F_Q01bUSP	x	x	Skill use work (prior to incarceration) - Time cooperating with coworkers
F_Q02aUSP	x	x	Skill use work (prior to incarceration) - How often - Sharing work-related info
F_Q02bUSP	x	x	Skill use work (prior to incarceration) - How often - Teaching people
F_Q02cUSP	x	x	Skill use work (prior to incarceration) - How often - Presentations
F_Q02dUSP	x	x	Skill use work (prior to incarceration) - How often - Selling
F_Q02eUSP	x	x	Skill use work (prior to incarceration) - How often - Advising people
F_Q03aUSP	x	x	Skill use work (prior to incarceration) - How often - Planning own activities
F_Q03bUSP	x	x	Skill use work (prior to incarceration) - How often - Planning others activities
F_Q03cUSP	x	x	Skill use work (prior to incarceration) - How often - Organising own time

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
F_Q04aUSP	x	x	Skill use work (prior to incarceration) - How often - Influencing people
F_Q04bUSP	x	x	Skill use work (prior to incarceration) - How often - Negotiating with people
F_Q05aUSP	x	x	Skill use work (prior to incarceration) - Problem solving - Simple problems
F_Q05bUSP	x	x	Skill use work (prior to incarceration) - Problem solving - Complex problems
F_Q06bUSP	x	x	Skill use work (prior to incarceration) - How often - Working physically for long
F_Q06cUSP	x	x	Skill use work (prior to incarceration) - How often - Using hands or fingers
F_Q07aUSP	x	x	Skill use work (prior to incarceration) - Not challenged enough
F_Q07bUSP	x	x	Skill use work (prior to incarceration) - Need more training
FIRLGRGN		x	Source region of first language learned at home in childhood and still understand (9 regions - derived)
FIRLGRGNUS_C	x	x	Source region of first language learned at home in childhood and still understand (9 regions) (2 categories) (derived from FIRLGRGN)
FORBILANG	x	x	Has learned as a child and still understands at least two languages not including test language (derived)
FORBORNLANG	x	x	Interactions between foreign-born and language status (2 categories - derived)
G_Q01aUSP	x	x	Skill use work (prior to incarceration) - Literacy - Read directions or instructions
G_Q01bUSP	x	x	Skill use work (prior to incarceration) - Literacy - Read letters memos or mails
G_Q01cUSP	x	x	Skill use work (prior to incarceration) - Literacy - Read newspapers or magazines
G_Q01dUSP	x	x	Skill use work (prior to incarceration) - Literacy - Read professional journals or publications
G_Q01eUSP	x	x	Skill use work (prior to incarceration) - Literacy - Read books
G_Q01fUSP	x	x	Skill use work (prior to incarceration) - Literacy - Read manuals or reference materials

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
G_Q01gUSP	x	x	Skill use work (prior to incarceration) - Literacy - Read financial statements
G_Q01hUSP	x	x	Skill use work (prior to incarceration) - Literacy - Read diagrams maps or schematics
G_Q02aUSP	x	x	Skill use work (prior to incarceration) - Literacy - Write letters memos or mails
G_Q02bUSP	x	x	Skill use work (prior to incarceration) - Literacy - Write articles
G_Q02cUSP	x	x	Skill use work (prior to incarceration) - Literacy - Write reports
G_Q02dUSP	x	x	Skill use work (prior to incarceration) - Literacy - Fill in forms
G_Q03bUSP	x	x	Skill use work (prior to incarceration) - Numeracy - How often - Calculating costs or budgets
G_Q03cUSP	x	x	Skill use work (prior to incarceration) - Numeracy - How often - Use or calculate fractions or percentages
G_Q03dUSP	x	x	Skill use work (prior to incarceration) - Numeracy - How often - Use a calculator
G_Q03fUSP	x	x	Skill use work (prior to incarceration) - Numeracy - How often - Prepare charts graphs or tables
G_Q03gUSP	x	x	Skill use work (prior to incarceration) - Numeracy - How often - Use simple algebra or formulas
G_Q03hUSP	x	x	Skill use work (prior to incarceration) - Numeracy - How often - Use advanced math or statistics
G_Q04USP	x	x	Skill use work (prior to incarceration) - ICT - Experience with computer in job
G_Q05aUSP	x	x	Skill use work (prior to incarceration) - ICT - Internet - How often - For mail
G_Q05cUSP	x	x	Skill use work (prior to incarceration) - ICT - Internet - How often - Work related info
G_Q05dUSP	x	x	Skill use work (prior to incarceration) - ICT - Internet - How often - Conduct transactions
G_Q05eUSP	x	x	Skill use work (prior to incarceration) - ICT - Computer - How often - Spreadsheets
G_Q05fUSP	x	x	Skill use work (prior to incarceration) - ICT - Computer - How often - Word

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
G_Q05gUSP	x	x	Skill use work (prior to incarceration) - ICT - Computer - How often - Programming language
G_Q05hUSP	x	x	Skill use work (prior to incarceration) - ICT - Computer - How often - Real-time discussions
G_Q06USP	x	x	Skill use work (prior to incarceration) - ICT - Computer - Level of computer use
G_Q07USP	x	x	Skill use work (prior to incarceration) - ICT - Computer - Got the skills needed
G_Q08USP	x	x	Skill use work (prior to incarceration) - ICT - Computer - Lack of skills affect career
GENDER		x	Person gender (derived from BQ)
GENDER_R	x	x	Person resolved gender from BQ and QC check (derived)
H_Q01aUSP	x	x	Skill use everyday life (in prison) - Literacy - Read directions or instructions
H_Q01bUSP	x	x	Skill use everyday life (in prison) - Literacy - Read letters or memos
H_Q01cUSP	x	x	Skill use everyday life (in prison) - Literacy - Read newspapers or magazines
H_Q01dUSP	x	x	Skill use everyday life (in prison) - Literacy - Read professional journals or publications
H_Q01eUSP	x	x	Skill use everyday life (in prison) - Literacy - Read books
H_Q01fUSP	x	x	Skill use everyday life (in prison) - Literacy - Read manuals or reference materials
H_Q01gUSP	x	x	Skill use everyday life (in prison) - Literacy - Read financial statements
H_Q01hUSP	x	x	Skill use everyday life (in prison) - Literacy - Read diagrams maps or schematics
H_Q02aUSP	x	x	Skill use everyday life (in prison) - Literacy - Write letters or memos
H_Q02bUSP	x	x	Skill use everyday life (in prison) - Literacy - Write articles
H_Q02cUSP	x	x	Skill use everyday life (in prison) - Literacy - Write reports
H_Q02dUSP	x	x	Skill use everyday life (in prison) - Literacy - Fill in forms
H_Q03bUSP	x	x	Skill use everyday life (in prison) - Numeracy - How often - Calculating costs or budgets
H_Q03cUSP	x	x	Skill use everyday life (in prison) - Numeracy - How often - Use or calculate fractions or percentages

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
H_Q03dUSP	x	x	Skill use everyday life (in prison) - Numeracy - How often - Use a calculator
H_Q03fUSP	x	x	Skill use everyday life (in prison) - Numeracy - How often - Prepare charts graphs or tables
H_Q03gUSP	x	x	Skill use everyday life (in prison) - Numeracy - How often - Use simple algebra or formulas
H_Q03hUSP	x	x	Skill use everyday life (in prison) - Numeracy - How often - Use advanced math or statistics
H_Q04aUSP	x	x	Skill use everyday life - ICT - Ever used computer
H_Q04bUSP	x	x	Skill use everyday life (in prison) - ICT - Experience with computer everyday life
H_Q05eUSP	x	x	Skill use everyday life (in prison) - ICT - Computer - How often - Spreadsheets
H_Q05fUSP	x	x	Skill use everyday life (in prison) - ICT - Computer - How often - Word
H_Q05gUSP	x	x	Skill use everyday life (in prison) - ICT - Computer - How often - Programming language
HOMLANG	x	x	Test language same as language spoken most often at home (derived)
HOMLGRGN		x	Source region of language spoken most often at home (9 regions - derived)
HOMLGRGNUS_C	x	x	Source region of language spoken most often at home (9 regions) (2 categories) (derived from HOMLGRGN)
I_Q04b	x	x	About yourself - Learning strategies - Relate new ideas into real life
I_Q04d	x	x	About yourself - Learning strategies - Like learning new things
I_Q04h	x	x	About yourself - Learning strategies - Attribute something new
I_Q04j	x	x	About yourself - Learning strategies - Get to the bottom of difficult things
I_Q04l	x	x	About yourself - Learning strategies - Figure out how different ideas fit together
I_Q04m	x	x	About yourself - Learning strategies - Looking for additional info
I_Q06a	x	x	About yourself - Political efficacy - No influence on the government
I_Q06dUSP1g	x	x	About yourself - Political efficacy - Information from family members, friends, or inmates

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
I_Q06dUSX1a	x	x	About yourself - Political efficacy - Information from newspapers
I_Q06dUSX1b	x	x	About yourself - Political efficacy - Information from magazines
I_Q06dUSX1d	x	x	About yourself - Political efficacy - Information from radio
I_Q06dUSX1e	x	x	About yourself - Political efficacy - Information from television
I_Q06dUSX1f	x	x	About yourself - Political efficacy - Information from books or brochures
I_Q07a	x	x	About yourself - Social trust - Trust only few people
I_Q07b	x	x	About yourself - Social trust - Other people take advantage of you
I_Q08	x	x	About yourself - Health - State
I_Q08USX1	x	x	About yourself - Health - Difficulty seeing print
I_Q08USX2	x	x	About yourself - Health - Difficulty hearing conversation
I_Q08USX3	x	x	About yourself - Health - Diagnosed learning disabled
I_Q10bUSP2g	x	x	About yourself - Health - Health information from family members, friends, or inmates
I_Q10bUSX2a	x	x	About yourself - Health - Health information from newspapers
I_Q10bUSX2b	x	x	About yourself - Health - Health information from magazines
I_Q10bUSX2d	x	x	About yourself - Health - Health information from radio
I_Q10bUSX2e	x	x	About yourself - Health - Health information from television
I_Q10bUSX2f	x	x	About yourself - Health - Health information from books or brochures
I_Q10bUSX2h	x	x	About yourself - Health - Health information from health professional
ICTWORK_LUSP	x	x	Index of use of ICT skills at work (derived), last job prior to incarceration
ICTWORK_LUSP_WLE_CA	x	x	Index of use of ICT skills at work, categorised WLE (derived), last job prior to incarceration
ID_HH		x	Sampling ID: Household (HH) identification number
ID_PSU		x	Sampling ID: Primary sampling unit (PSU) identification number

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
ID_SSU		x	Sampling ID: Second-stage sampling unit (SSU) identification number
IMGEN	x	x	First and second generation immigrants (derived)
IMPAR	x	x	Parents' immigration status (derived)
IMYRCAT	x	x	Years in country (2-category - derived)
IMYRS		x	Years in country (derived)
IMYRS_C	x	x	Years in country (categorised, 4 categories)
INFLUENCE_LUSP	x	x	Index of use of influencing skills at work (derived), last job prior to incarceration
INFLUENCE_LUSP_WLE_CA	x	x	Index of use of influencing skills at work, categorised WLE (derived), last job prior to incarceration
INPIAAC	x	x	PIAAC Sample Indicator
INTLFLAG	x	x	Flag for international comparison of U.S. combined Main Study and National Supplement
ISCED_HF		x	Level of Highest Qualification (Foreign) - Respondent (ISCED) (coded)
ISCED_HF_C		x	Level of Highest Qualification (collapsed, 14 categories)
ISCO08_CUSP		x	Prison Job Occupation - Respondent (ISCO 2008) (coded)
ISCO08_CUSP_C	x	x	Prison Job Occupation - Respondent (ISCO 2008) (combined into 3-digit categories) (derived from ISCO08_CUSP)
ISCO08_LUSP		x	Last Job Occupation (prior to incarceration) - Respondent (ISCO 2008) (coded)
ISCO08_LUSP_C	x	x	Last Job Occupation (prior to incarceration) - Respondent (ISCO 2008) (combined into 3-digit categories) (derived from ISCO08_LUSP)
ISIC4_LUSP		x	Last Job Industry (prior to incarceration) - Respondent (ISIC rev 4) (coded)
ISIC4_LUSP_C	x	x	Last Job Industry (prior to incarceration) - Respondent (ISIC rev 4) (combined into 3-digit categories) (derived from ISIC4_LUSP)
J_N05a2	x	x	Background - More than one language mentioned
J_Q02c	x	x	Background - Work situation of spouse or partner
J_Q03a	x	x	Background - Children
J_Q03b		x	Background - Number of children

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
J_Q03b_C	x	x	Background - Number of children (top-coded at 4)
J_Q03c		x	Background - Age of the child
J_Q03c_C	x	x	Background - Age of the child (categorised, 4 categories)
J_Q03CUS_C	x	x	Background - Age of the child (5 categories) (derived from J_Q03C)
J_Q03d1		x	Background - Age of the youngest child
J_Q03d1_C	x	x	Background - Age of the youngest child (categorised, 4 categories)
J_Q03D1US_C	x	x	Background - Age of the youngest child (5 categories) (derived from J_Q03D1)
J_Q03d2		x	Background - Age of the oldest child
J_Q03d2_C	x	x	Background - Age of the oldest child (categorised, 4 categories)
J_Q03D2US_C	x	x	Background - Age of the oldest child (5 categories) (derived from J_Q03D2)
J_Q04a	x	x	Background - Born in country
J_Q04bUS		x	Background - Country of birth
J_Q04c1		x	Background - Age of immigration
J_Q04c1_C	x	x	Background - Age of immigration (categorised, 9 categories)
J_Q04c2		x	Background - Year of immigration
J_Q04C2US_C	x	x	Background - Year of immigration (4 categories) (derived from J_Q04C2)
J_Q04dUSX1a	x	x	Background - Hispanic
J_Q04dUSX1b_01		x	Background - Hispanic origin - Mexican
J_Q04dUSX1b_02		x	Background - Hispanic origin - Puerto Rican
J_Q04dUSX1b_03		x	Background - Hispanic origin - Cuban
J_Q04dUSX1b_04		x	Background - Hispanic origin - Central/South America
J_Q04dUSX1b_05		x	Background - Hispanic origin - Other
J_Q04dUSX2_01		x	Background - Race - White
J_Q04dUSX2_02		x	Background - Race - Black
J_Q04dUSX2_03		x	Background - Race - Asian
J_Q04dUSX2_04		x	Background - Race - American Indian
J_Q04dUSX2_05		x	Background - Race - Native Hawaiian
J_Q05a1US		x	Background - First learned language
J_Q05a2US		x	Background - Second learned language
J_Q05a2USX2	x	x	Background - Age learned English
J_Q05bUS		x	Background - Language spoken at home
J_Q05cUSP4	x	x	Background - ESL class/tutor during current incarceration

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
J_Q05cUSP5		x	Background - Reason for ESL class/tutor during current incarceration
J_Q05cUSX1		x	Background - Language spoken most
J_Q05cUSX2	x	x	Background - English outside home
J_Q05cUSX3a	x	x	Background - Ability to understand spoken English
J_Q05cUSX3b	x	x	Background - Ability to speak English
J_Q05cUSX3d	x	x	Background - Ability to read English
J_Q05cUSX3e	x	x	Background - Ability to write English
J_Q05cUSX6	x	x	Background - Class/tutor learn English as adult
J_Q06a	x	x	Background - Mother/female guardian - Whether born in country
J_Q06b	x	x	Background - Mother/female guardian - Highest level of education
J_Q06b_ISCED11	x	x	Background - Mother/female guardian - Highest level of education, ISCED 2011
J_Q06bUS	x	x	Background - Mother/female guardian - Highest level of education
J_Q07a	x	x	Background - Father/male guardian - Whether born in country
J_Q07b	x	x	Background - Father/male guardian - Highest level of education
J_Q07b_ISCED11	x	x	Background - Father/male guardian - Highest level of education, ISCED 2011
J_Q07bUS	x	x	Background - Father/male guardian - Highest level of education
J_Q08	x	x	Background - Number of books at home
J_S04b		x	Background - Country of birth (other)
J_S05a1		x	Background - First learned language (other)
J_S05a2		x	Background - Second learned language (other)
J_S05b		x	Background - Language spoken at home (other)
LANGUAGE	x	x	Background - English language status (derived)
LEAVER1624	x	x	Youth aged 16 to 24 who have left education without completing ISCED 3 or higher (derived)
LITSTATUS	x	x	Literacy - PV Status
LNG_BQ	x	x	Language for background questionnaire (derived, ISO 639-2/T)
LNG_CI	x	x	Language for exercise (derived, ISO 639-2/T)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
LNG_HOME		x	Language most often spoken at home - Respondent (ISO 639-2/T) (coded)
LNG_HOMEUS_C	x	x	Language most often spoken at home - Respondent (ISO 639-2/T) (3 categories) (derived from LNG_HOME)
LNG_L1		x	First language learned at home in childhood and still understood - Respondent (IS0 639-2/T) (coded)
LNG_L1US_C	x	x	First language learned at home in childhood and still understood - Respondent (IS0 639-2/T) (3 categories) (derived from LNG_L1)
LNG_L2		x	Second language learned at home in childhood and still understood - Respondent (IS0 639-2/T) (coded)
LNG_L2US_C	x	x	Second language learned at home in childhood and still understood - Respondent (IS0 639-2/T) (3 categories) (derived from LNG_L2)
M300C02S	x	x	PLC / 300 - Q5 - Employment Advertisement - Employees (Scored Response)
M301C05S	x	x	PLC / 301 - Q1 - SGIH - Tel. number (Scored Response)
M305215S	x	x	PL / 305 - Q10 - TMN AntiTheft - Documents (Scored Response)
M305218S	x	x	PL / 305 - Q11 - TMN AntiTheft - Block/Unblock (Scored Response)
M309319S	x	x	PL / 309 - Q17- Generic Medicines - Limited Use (Scored Response)
M309320S	x	x	PL / 309 - Q18 - Generic Medicines - Market share (Scored Response)
M309321S	x	x	PL / 309 - Q19 - Generic Medicines – 10% or more (Scored Response)
M309322S	x	x	PL / 309 - Q20 - Generic Medicines - Reasons (Scored Response)
M310406S	x	x	PL / 310 - Q15 - Memory Training - Brain Part (Scored Response)
M310407S	x	x	PL / 310 - Q16 - Memory Training - Discovery (Scored Response)
M313410S	x	x	PL / 313 - Q3 - International Calls - Dial 098 (Scored Response)
M313411S	x	x	PL / 313 - Q4 - International Calls - Full number (Scored Response)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
M313412S	x	x	PL / 313 - Q5 - International Calls - Country code (Scored Response)
M313413S	x	x	PL / 313 - Q6 - International Calls - Information (Scored Response)
M313414S	x	x	PL / 313 - Q7 - International Calls - Call Canada (Scored Response)
M600C04S	x	x	PNC / 600 - Q4 - Election results - Votes (Scored Response)
M602501S	x	x	PN / 602 - Q17 - Price Tag - Packed first (Scored Response)
M602502S	x	x	PN / 602 - Q18 - Price Tag - Change (Scored Response)
M602503S	x	x	PN / 602 - Q19 - Price Tag - Quarter (Scored Response)
M604505S	x	x	PN / 604 - Q14 - Gas Gauge - Gas remaining (Scored Response)
M610515S	x	x	PN / 610 - Q15 - Compound Interest - Invested (Scored Response)
M615602S	x	x	PN / 615 - Q1 - Candles - Layers (Scored Response)
M615603S	x	x	PN / 615 - Q2 - Candles - Total Weight (Scored Response)
M618607S	x	x	PN / 618 - Q12 - Six Pack - Price per can (Scored Response)
M618608S	x	x	PN / 618 - Q13 - Six Pack - Discount (Scored Response)
M620610S	x	x	PN / 620 - Q4 - Inflation - Prediction (Scored Response)
M620612S	x	x	PN / 620 - Q5 - Inflation - Actual rate (Scored Response)
M623616S	x	x	PN / 623 - Q7 - Wine - Bottles drink (Scored Response)
M623617S	x	x	PN / 623 - Q8 - Wine - Gallon (Scored Response)
M623618S	x	x	PN / 623 - Q9 - Wine - Spain / US (Scored Response)
M624619S	x	x	PN / 624 - Q10 - BMI - Weight zone (Scored Response)
M624620S	x	x	PN / 624 - Q11 - BMI - What is BMI (Scored Response)
N302C02S	x	x	PLC / 302 - Q3 - Election Results - Candidate (Scored Response)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
N306110S	x	x	PL / 306 - Q1 - CANCO - Information (Scored Response)
N306111S	x	x	PL / 306 - Q2 - CANCO - Two ways (Scored Response)
NATBILANG	x	x	Has learned as a child and still understands at least two languages including test language (derived)
NATIVELANG	x	x	Test language same as native language (derived)
NATIVESPEAKER	x	x	Respondent is a native speaker (DERIVED BY CAPI)
NUMHOMEUSP	x	x	Index of use of numeracy skills in everyday life (in prison) (basic and advanced - derived)
NUMHOMEUSP_WLE_CA	x	x	Index of use of numeracy skills in everyday life (in prison) (basic and advanced), categorised WLE (derived)
NUMSTATUS	x	x	Numeracy - PV Status
NUMWORK_CUSP	x	x	Index of use of numeracy skills at work (basic and advanced) (derived), current prison job
NUMWORK_CUSP_WLE_CA	x	x	Index of use of numeracy skills at work (basic and advanced), categorized WLE (derived), current prison job
NUMWORK_LUSP	x	x	Index of use of numeracy skills at work (basic and advanced - derived), last job prior to incarceration
NUMWORK_LUSP_WLE_CA	x	x	Index of use of numeracy skills at work (basic and advanced), categorised WLE (derived), last job prior to incarceration
P_G_Q01a	x	x	Prison - Skill use work (prison job) - Literacy - Read directions or instructions
P_G_Q01b	x	x	Prison - Skill use work (prison job) - Literacy - Read letters or memos
P_G_Q01c	x	x	Prison - Skill use work (prison job) - Literacy - Read newspapers or magazines
P_G_Q01d	x	x	Prison - Skill use work (prison job) - Literacy - Read professional journals or publications
P_G_Q01e	x	x	Prison - Skill use work (prison job) - Literacy - Read books
P_G_Q01f	x	x	Prison - Skill use work (prison job) - Literacy - Read manuals or reference materials

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P_G_Q01g	x	x	Prison - Skill use work (prison job) - Literacy - Read financial statements
P_G_Q01h	x	x	Prison - Skill use work (prison job) - Literacy - Read diagrams maps or schematics
P_G_Q02a	x	x	Prison - Skill use work (prison job) - Literacy - Write letters or memos
P_G_Q02b	x	x	Prison - Skill use work (prison job) - Literacy - Write articles
P_G_Q02c	x	x	Prison - Skill use work (prison job) - Literacy - Write reports
P_G_Q02d	x	x	Prison - Skill use work (prison job) - Literacy - Fill in forms
P_G_Q03b	x	x	Prison - Skill use work (prison job) - Numeracy - How often - Calculating costs or budgets
P_G_Q03c	x	x	Prison - Skill use work (prison job) - Numeracy - How often - Use or calculate fractions or percentages
P_G_Q03d	x	x	Prison - Skill use work (prison job) - Numeracy - How often - Use a calculator
P_G_Q03f	x	x	Prison - Skill use work (prison job) - Numeracy - How often - Prepare charts graphs or tables
P_G_Q03g	x	x	Prison - Skill use work (prison job) - Numeracy - How often - Use simple algebra or formulas
P_G_Q03h	x	x	Prison - Skill use work (prison job) - Numeracy - How often - Use advanced math or statistics
P_G_Q04	x	x	Prison - Skill use work (prison job) - ICT - Experience with computer in job
P_G_Q05e	x	x	Prison - Skill use work (prison job) - ICT - Computer - How often - Spreadsheets
P_G_Q05f	x	x	Prison - Skill use work (prison job) - ICT - Computer - How often - Word
P_G_Q05g	x	x	Prison - Skill use work (prison job) - ICT - Computer - How often - Programming language
P_G_Q06	x	x	Prison - Skill use work (prison job) - ICT - Computer - Level of computer use
P_G_Q07	x	x	Prison - Skill use work (prison job) - ICT - Computer - Got the skills needed

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P_G_Q08	x	x	Prison - Skill use work (prison job) - ICT - Computer - Lack of skills affect career
P_Q010a	x	x	Prison - Education - Last 12 months - Time in class verbatim
P_Q010b	x	x	Prison - Education - Last 12 months - Time, unit
P_Q020	x	x	Prison - Education - Last 12 months - Degree offering
P_Q030	x	x	Prison - Education - Basic skills - Location
P_Q040	x	x	Prison - Education - Basic skills - Reason
P_Q040_S		x	Prison - Education - Basic skills - Reason other
P_Q050	x	x	Prison - Education - Prior - Reason stopped
P_Q050_S		x	Prison - Education - Basic skills - Reason stopped other
P_Q060	x	x	Prison - Education - Enrollment
P_Q070	x	x	Prison - Education - Waiting list
P_Q080	x	x	Prison - Education - Enrollment - Degree
P_Q080_ISCED11	x	x	Prison - Education - Enrollment - Degree, ISCED 2011
P_Q090	x	x	Prison - Education - Enrollment - Reason
P_Q090_S		x	Prison - Education - Enrollment - Reason other
P_Q100	x	x	Prison - Education - Enrollment - Reason not
P_Q100_S		x	Prison - Education - Enrollment - Reason not other
P_Q120	x	x	Prison - Education - Current - Highest level
P_Q120_ISCED11	x	x	Prison - Education - Current incarceration- Highest level, ISCED 2011
P_Q130	x	x	Prison - Education - Current - Course of study
P_Q130_S		x	Prison - Education - Current - Course offered other
P_Q140a	x	x	Prison - Education - Current- Time in class verbatim
P_Q140b	x	x	Prison - Education - Current - Time, unit
P_Q150	x	x	Prison - Education - Current - Reason
P_Q150_S		x	Prison - Education - Current - Reason other
P_Q160_1	x	x	Prison - Incarceration - Most recent - Month
P_Q160_2	x	x	Prison - Incarceration - Most recent - Year
P_Q170	x	x	Prison - Incarceration - Prior
P_Q180	x	x	Prison - Incarceration - Release
P_Q190a	x	x	Prison - Education - Current - Readiness Class
P_Q190a_1	x	x	Prison - Education - Readiness Class - Hours

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P_Q190a_2	x	x	Prison - Education - Readiness Class - Reason
P_Q190b	x	x	Prison - Education - Current - Parenting Class
P_Q190b_1	x	x	Prison - Education - Parenting Class - Hours
P_Q190b_2	x	x	Prison - Education - Parenting Class - Reason
P_Q190c	x	x	Prison - Education - Current - Life Skills Class
P_Q190c_1	x	x	Prison - Education - Life Skills Class - Hours
P_Q190c_2	x	x	Prison - Education - Life Skills Class - Reason
P_Q190d	x	x	Prison - Education - Current - Adjustment Class
P_Q190d_1	x	x	Prison - Education - Adjustment Class - Hours
P_Q190d_2	x	x	Prison - Education - Adjustment Class - Reason
P_Q190e	x	x	Prison - Education - Current - Addiction Group
P_Q190e_1	x	x	Prison - Education - Addiction Group - Hours
P_Q190e_2	x	x	Prison - Education - Addiction Group - Reason
P_Q190f	x	x	Prison - Education - Current - Mental Health Group
P_Q190f_1	x	x	Prison - Education - Mental Health Group - Hours
P_Q190f_2	x	x	Prison - Education - Mental Health Group - Reason
P_Q190g	x	x	Prison - Education - Current - Inmate Assistance Group
P_Q190g_1	x	x	Prison - Education - Inmate Assistance Group - Hours
P_Q190g_2	x	x	Prison - Education - Inmate Assistance Group - Reason
P_Q190h	x	x	Prison - Education - Current - Religious Study Group
P_Q190h_1	x	x	Prison - Education - Religious Study Group - Hours
P_Q190h_2	x	x	Prison - Education - Religious Study Group - Reason
P_Q190i	x	x	Prison - Education - Current - Ethnic/Racial Organization
P_Q190i_1	x	x	Prison - Education - Ethnic/Racial Organization - Hours
P_Q190i_2	x	x	Prison - Education - Ethnic/Racial Organization - Reason
P_Q190j	x	x	Prison - Education - Current - Other

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P_Q190j_1		x	Prison - Education - Current - Other specified
P_Q190j_2		x	Prison - Education - Other - Hours
P_Q190j_3		x	Prison - Education - Other - Reason
P_Q220	x	x	Prison - Education - Job Training
P_Q230	x	x	Prison - Education - Waiting List Job Training
P_Q240_01	x	x	Prison - Reason job skill program, 1st response
P_Q240_02	x	x	Prison - Reason job skill program, 2nd response
P_Q240_03	x	x	Prison - Reason job skill program, 3rd response
P_Q240_04	x	x	Prison - Reason job skill program, 4th response
P_Q240_05	x	x	Prison - Reason job skill program, 5th response
P_Q240_06	x	x	Prison - Reason job skill program, 6th response
P_Q240_S		x	Prison - Education - Job Training - Reason Other
P_Q250	x	x	Prison - Education - Job Training - Reason not
P_Q250_S		x	Prison - Education - Job Training - Reason not Other
P_Q260	x	x	Prison - Education - Job Training - Intent
P_Q270	x	x	Prison - Education - Job Training - Hours spent
P_Q280	x	x	Prison - Education - Job Training - Hours scheduled
P_Q290	x	x	Prison - Education - IT certification
P_Q300	x	x	Prison - Education - IT test for certification
P_Q310	x	x	Prison - Education - IT test prepared
P_Q310_S		x	Prison - Education - IT test prepared, other
P_Q320	x	x	Prison - Education - Certificate other than IT
P_Q330	x	x	Prison - Education - non-IT test for certification
P_Q340	x	x	Prison - Education - non-IT test prepared
P_Q340_S		x	Prison - Education - non-IT test prepared, other
P_Q350a	x	x	Prison - Education - Current - Time spent, amount
P_Q350b	x	x	Prison - Education - Current - Time unit
P_Q360	x	x	Prison - Education - Access to library
P_Q370	x	x	Prison - Education - Library, often use
P_Q380	x	x	Prison - Education - Library, length to access
P_Q390	x	x	Prison - Jobs - Current incarceration
P_Q400	x	x	Prison - Jobs - Waiting list
P_Q410	x	x	Prison - Jobs - Ever in prison
P_Q420	x	x	Prison - Jobs - Site
P_Q430_01	x	x	Prison - Income sources (prior to incarceration), 1st response

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P_Q430_02	x	x	Prison - Income sources (prior to incarceration), 2nd response
P_Q430_03	x	x	Prison - Income sources (prior to incarceration), 3rd response
P_Q430_04	x	x	Prison - Income sources (prior to incarceration), 4th response
P_Q430_05	x	x	Prison - Income sources (prior to incarceration), 5th response
P_Q430_06	x	x	Prison - Income sources (prior to incarceration), 6th response
P_Q430_07	x	x	Prison - Income sources (prior to incarceration), 7th response
P_Q430_08	x	x	Prison - Income sources (prior to incarceration), 8th response
P_Q430_S		x	Prison - Income sources (prior to incarceration), other
P_Q440	x	x	Prison - Background - Marital status
P317001S	x	x	PL / 317 - Q12 - Apples - Evidence (Scored Response)
P317002S	x	x	PL / 317 - Q13 - Apples - Composition (Scored Response)
P317003S	x	x	PL / 317 - Q14 - Apples - Occupation (Scored Response)
P324002S	x	x	PL / 324 - Q8 - Milk Label - Safe (Scored Response)
P324003S	x	x	PL / 324 - Q9 - Milk Label - Calcium (Scored Response)
P330001S	x	x	PLC / 330 - Q2 - Guadeloupe - Falls (Scored Response)
P601C06S	x	x	PNC / 601 - Q6 - Bottles - Bottles (Scored Response)
P614601S	x	x	PNC / 614 - Q7 - Watch - Price (Scored Response)
P640001S	x	x	PN / 640 - Q3 - Odometer - Trip Miles (Scored Response)
P645001S	x	x	PNC / 645 - Q8 - AirportTimetable - Departure (Scored Response)
P655001S	x	x	PN / 655 - Q20 - Path - Length (Scored Response)
P664001S	x	x	PN / 664 - Q16 - Orchestra tickets - Student ticket (Scored Response)
P666001S	x	x	PN / 666 - Q6 - Rope - Measure (Scored Response)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P901002R	x	x	PV Q1 square (Actual Response)
P901002S	x	x	PV Q1 square (Scored Response)
P901003R	x	x	PV Q3 hand (Actual Response)
P901003S	x	x	PV Q3 hand (Scored Response)
P901004R	x	x	PV Q7 moon (Actual Response)
P901004S	x	x	PV Q7 moon (Scored Response)
P901005R	x	x	PV Q4 baby (Actual Response)
P901005S	x	x	PV Q4 baby (Scored Response)
P901006R	x	x	PV Q6 bird (Actual Response)
P901006S	x	x	PV Q6 bird (Scored Response)
P901011R	x	x	PV Q13 piano (Actual Response)
P901011S	x	x	PV Q13 piano (Scored Response)
P901013R	x	x	PV Q8 bread (Actual Response)
P901013S	x	x	PV Q8 bread (Scored Response)
P901015R	x	x	PV Q33 wrist (Actual Response)
P901015S	x	x	PV Q33 wrist (Scored Response)
P901017R	x	x	PV Q15 elephant (Actual Response)
P901017S	x	x	PV Q15 elephant (Scored Response)
P901018R	x	x	PV Q16 saw (Actual Response)
P901018S	x	x	PV Q16 saw (Scored Response)
P901019R	x	x	PV Q17 bus (Actual Response)
P901019S	x	x	PV Q17 bus (Scored Response)
P901020R	x	x	PV Q14 computer (Actual Response)
P901020S	x	x	PV Q14 computer (Scored Response)
P901021R	x	x	PV Q22 zipper (Actual Response)
P901021S	x	x	PV Q22 zipper (Scored Response)
P901024R	x	x	PV Q26 candle (Actual Response)
P901024S	x	x	PV Q26 candle (Scored Response)
P901025R	x	x	PV Q27 fountain (Actual Response)
P901025S	x	x	PV Q27 fountain (Scored Response)
P902014R	x	x	PV Q28 microscope (Actual Response)
P902014S	x	x	PV Q28 microscope (Scored Response)
P902021R	x	x	PV Q19 umbrella (Actual Response)
P902021S	x	x	PV Q19 umbrella (Scored Response)
P902022R	x	x	PV Q25 door (Actual Response)
P902022S	x	x	PV Q25 door (Scored Response)
P902024R	x	x	PV Q12 crown (Actual Response)
P902024S	x	x	PV Q12 crown (Scored Response)
P903007R	x	x	PV Q9 shovel (Actual Response)
P903007S	x	x	PV Q9 shovel (Scored Response)
P903012R	x	x	PV Q10 chess (Actual Response)
P903012S	x	x	PV Q10 chess (Scored Response)
P903017R	x	x	PV Q20 lion (Actual Response)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P903017S	x	x	PV Q20 lion (Scored Response)
P903021R	x	x	PV Q32 saddle (Actual Response)
P903021S	x	x	PV Q32 saddle (Scored Response)
P903024R	x	x	PV Q29 envelope (Actual Response)
P903024S	x	x	PV Q29 envelope (Scored Response)
P904009R	x	x	PV Q5 ruler (Actual Response)
P904009S	x	x	PV Q5 ruler (Scored Response)
P904010R	x	x	PV Q11 pants (Actual Response)
P904010S	x	x	PV Q11 pants (Scored Response)
P904012R	x	x	PV Q2 butterfly (Actual Response)
P904012S	x	x	PV Q2 butterfly (Scored Response)
P904014R	x	x	PV Q21 bottle (Actual Response)
P904014S	x	x	PV Q21 bottle (Scored Response)
P904015R	x	x	PV Q23 neck (Actual Response)
P904015S	x	x	PV Q23 neck (Scored Response)
P904020R	x	x	PV Q24 television (Actual Response)
P904020S	x	x	PV Q24 television (Scored Response)
P904021R	x	x	PV Q34 arrow (Actual Response)
P904021S	x	x	PV Q34 arrow (Scored Response)
P904022R	x	x	PV Q18 stool (Actual Response)
P904022S	x	x	PV Q18 stool (Scored Response)
P904024R	x	x	PV Q30 bell (Actual Response)
P904024S	x	x	PV Q30 bell (Scored Response)
P904025R	x	x	PV Q31 axe (Actual Response)
P904025S	x	x	PV Q31 axe (Scored Response)
P911001R	x	x	SP S1 (Actual Response)
P911001S	x	x	SP S1 (Scored Response)
P911003R	x	x	SP S3 (Actual Response)
P911003S	x	x	SP S3 (Scored Response)
P911004R	x	x	SP S4 (Actual Response)
P911004S	x	x	SP S4 (Scored Response)
P911005R	x	x	SP S5 (Actual Response)
P911005S	x	x	SP S5 (Scored Response)
P911006R	x	x	SP S6 (Actual Response)
P911006S	x	x	SP S6 (Scored Response)
P911007R	x	x	SP S7 (Actual Response)
P911007S	x	x	SP S7 (Scored Response)
P911008R	x	x	SP S8 (Actual Response)
P911008S	x	x	SP S8 (Scored Response)
P911009R	x	x	SP S9 (Actual Response)
P911009S	x	x	SP S9 (Scored Response)
P911010R	x	x	SP S10 (Actual Response)
P911010S	x	x	SP S10 (Scored Response)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P911014R	x	x	SP S15 (Actual Response)
P911014S	x	x	SP S15 (Scored Response)
P911017R	x	x	SP S19 (Actual Response)
P911017S	x	x	SP S19 (Scored Response)
P911020R	x	x	SP S22 (Actual Response)
P911020S	x	x	SP S22 (Scored Response)
P912002R	x	x	SP S2 (Actual Response)
P912002S	x	x	SP S2 (Scored Response)
P912011R	x	x	SP S12 (Actual Response)
P912011S	x	x	SP S12 (Scored Response)
P912013R	x	x	SP S14 (Actual Response)
P912013S	x	x	SP S14 (Scored Response)
P912019R	x	x	SP S21 (Actual Response)
P912019S	x	x	SP S21 (Scored Response)
P913013R	x	x	SP S11 (Actual Response)
P913013S	x	x	SP S11 (Scored Response)
P914012R	x	x	SP S13 (Actual Response)
P914012S	x	x	SP S13 (Scored Response)
P914015R	x	x	SP S16 (Actual Response)
P914015S	x	x	SP S16 (Scored Response)
P914016R	x	x	SP S18 (Actual Response)
P914016S	x	x	SP S18 (Scored Response)
P914018R	x	x	SP S20 (Actual Response)
P914018S	x	x	SP S20 (Scored Response)
P914019R	x	x	SP S17 (Actual Response)
P914019S	x	x	SP S17 (Scored Response)
P921002R	x	x	PC P1 S2 The Birthday Party (Actual Response)
P921002S	x	x	PC P1 S2 The Birthday Party (Scored Response)
P921003R	x	x	PC P1 S3 The Birthday Party (Actual Response)
P921003S	x	x	PC P1 S3 The Birthday Party (Scored Response)
P921004R	x	x	PC P1 S4 The Birthday Party (Actual Response)
P921004S	x	x	PC P1 S4 The Birthday Party (Scored Response)
P921005R	x	x	PC P1 S5 The Birthday Party (Actual Response)
P921005S	x	x	PC P1 S5 The Birthday Party (Scored Response)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P921007R	x	x	PC P1 S7 The Birthday Party (Actual Response)
P921007S	x	x	PC P1 S7 The Birthday Party (Scored Response)
P921008R	x	x	PC P1 S8 The Birthday Party (Actual Response)
P921008S	x	x	PC P1 S8 The Birthday Party (Scored Response)
P921009R	x	x	PC P1 S9 The Birthday Party (Actual Response)
P921009S	x	x	PC P1 S9 The Birthday Party (Scored Response)
P921010R	x	x	PC P1 S10 The Birthday Party (Actual Response)
P921010S	x	x	PC P1 S10 The Birthday Party (Scored Response)
P921011R	x	x	PC P1 S11 The Birthday Party (Actual Response)
P921011S	x	x	PC P1 S11 The Birthday Party (Scored Response)
P921013R	x	x	PC P1 S13 The Birthday Party (Actual Response)
P921013S	x	x	PC P1 S13 The Birthday Party (Scored Response)
P921014R	x	x	PC P1 S14 The Birthday Party (Actual Response)
P921014S	x	x	PC P1 S14 The Birthday Party (Scored Response)
P921015R	x	x	PC P1 S15 The Birthday Party (Actual Response)
P921015S	x	x	PC P1 S15 The Birthday Party (Scored Response)
P921016R	x	x	PC P1 S16 The Birthday Party (Actual Response)
P921016S	x	x	PC P1 S16 The Birthday Party (Scored Response)
P921017R	x	x	PC P1 S17 The Birthday Party (Actual Response)
P921017S`	x	x	PC P1 S17 The Birthday Party (Scored Response)
P921018R	x	x	PC P1 S18 The Birthday Party (Actual Response)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P921018S	x	x	PC P1 S18 The Birthday Party (Scored Response)
P921019R	x	x	PC P1 S19 The Birthday Party (Actual Response)
P921019S	x	x	PC P1 S19 The Birthday Party (Scored Response)
P921020R	x	x	PC P1 S20 The Birthday Party (Actual Response)
P921020S	x	x	PC P1 S20 The Birthday Party (Scored Response)
P921021R	x	x	PC P1 S21 The Birthday Party (Actual Response)
P921021S	x	x	PC P1 S21 The Birthday Party (Scored Response)
P921035R	x	x	PC P3 S2 World Sports (Actual Response)
P921035S	x	x	PC P3 S2 World Sports (Scored Response)
P921036R	x	x	PC P3 S3 World Sports (Actual Response)
P921036S	x	x	PC P3 S3 World Sports (Scored Response)
P921037R	x	x	PC P3 S4 World Sports (Actual Response)
P921037S	x	x	PC P3 S4 World Sports (Scored Response)
P921038R	x	x	PC P3 S5 World Sports (Actual Response)
P921038S	x	x	PC P3 S5 World Sports (Scored Response)
P921040R	x	x	PC P3 S7 World Sports (Actual Response)
P921040S	x	x	PC P3 S7 World Sports (Scored Response)
P921041R	x	x	PC P3 S8 World Sports (Actual Response)
P921041S	x	x	PC P3 S8 World Sports (Scored Response)
P921042R	x	x	PC P3 S9 World Sports (Actual Response)
P921042S	x	x	PC P3 S9 World Sports (Scored Response)
P921043R	x	x	PC P3 S10 World Sports (Actual Response)
P921043S	x	x	PC P3 S10 World Sports (Scored Response)
P922023R	x	x	PC P2 S2 A Letter to the Editor - Clinic (Actual Response)
P922023S	x	x	PC P2 S2 A Letter to the Editor - Clinic (Scored Response)
P922024R	x	x	PC P2 S3 A Letter to the Editor - Clinic (Actual Response)
P922024S	x	x	PC P2 S3 A Letter to the Editor - Clinic (Scored Response)
P922025R	x	x	PC P2 S4 A Letter to the Editor - Clinic (Actual Response)
P922025S	x	x	PC P2 S4 A Letter to the Editor - Clinic (Scored Response)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P922026R	x	x	PC P2 S5 A Letter to the Editor - Clinic (Actual Response)
P922026S	x	x	PC P2 S5 A Letter to the Editor - Clinic (Scored Response)
P922027R	x	x	PC P2 S6 A Letter to the Editor - Clinic (Actual Response)
P922027S	x	x	PC P2 S6 A Letter to the Editor - Clinic (Scored Response)
P922028R	x	x	PC P2 S7 A Letter to the Editor - Clinic (Actual Response)
P922028S	x	x	PC P2 S7 A Letter to the Editor - Clinic (Scored Response)
P922030R	x	x	PC P2 S9 A Letter to the Editor - Clinic (Actual Response)
P922030S	x	x	PC P2 S9 A Letter to the Editor - Clinic (Scored Response)
P922031R	x	x	PC P2 S10 A Letter to the Editor - Clinic (Actual Response)
P922031S	x	x	PC P2 S10 A Letter to the Editor - Clinic (Scored Response)
P922032R	x	x	PC P2 S11 A Letter to the Editor - Clinic (Actual Response)
P922032S	x	x	PC P2 S11 A Letter to the Editor - Clinic (Scored Response)
P922033R	x	x	PC P2 S12 A Letter to the Editor - Clinic (Actual Response)
P922033S	x	x	PC P2 S12 A Letter to the Editor - Clinic (Scored Response)
P924035R	x	x	PC P4 S2 Going to the Movies (Actual Response)
P924035S	x	x	PC P4 S2 Going to the Movies (Scored Response)
P924036R	x	x	PC P4 S3 Going to the Movies (Actual Response)
P924036S	x	x	PC P4 S3 Going to the Movies (Scored Response)
P924037R	x	x	PC P4 S4 Going to the Movies (Actual Response)
P924037S	x	x	PC P4 S4 Going to the Movies (Scored Response)
P924038R	x	x	PC P4 S5 Going to the Movies (Actual Response)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
P924038S	x	x	PC P4 S5 Going to the Movies (Scored Response)
P924040R	x	x	PC P4 S7 Going to the Movies (Actual Response)
P924040S	x	x	PC P4 S7 Going to the Movies (Scored Response)
P924041R	x	x	PC P4 S8 Going to the Movies (Actual Response)
P924041S	x	x	PC P4 S8 Going to the Movies (Scored Response)
P924042R	x	x	PC P4 S9 Going to the Movies (Actual Response)
P924042S	x	x	PC P4 S9 Going to the Movies (Scored Response)
P924043R	x	x	PC P4 S10 Going to the Movies (Actual Response)
P924043S	x	x	PC P4 S10 Going to the Movies (Scored Response)
PAPER	x	x	Paper branch (derived)
PARED	x	x	Highest of mother or father's level of education (derived)
PBROUTE	x	x	Paper-based routing code (derived)
PLANNING_LUSP	x	x	Index of use of planning skills at work (derived), last job prior to incarceration
PLANNING_LUSP_WLE_CA	x	x	Index of use of planning skills at work, categorised WLE (derived), last job prior to incarceration
PPC_SCORE	x	x	Final score for the paper core assessment
PRC_PC_SCR	x	x	Total Score for Reading Components Section - Passage Comprehension (derived)
PRC_PF_Q1	x	x	Sentence Timer for Passage Comprehension items - passage 1
PRC_PF_Q2	x	x	Sentence Timer for Passage Comprehension items - passage 2
PRC_PF_Q3	x	x	Sentence Timer for Passage Comprehension items - passage 3 and 4
PRC_PV_Q1	x	x	Sentence Timer for Print Vocabulary items
PRC_PV_SCR	x	x	Total Score for Reading Components Section - Print Vocabulary (derived)
PRC_SP_Q1	x	x	Sentence Timer for Sentence Processing items
PRC_SP_SCR	x	x	Total Score for Reading Components Section - Sentence Processing (derived)
PSLSTATUS	x	x	Problem Solving - PV Status

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
PVLIT1	x	x	Literacy scale score - Plausible value 1
PVLIT10	x	x	Literacy scale score - Plausible value 10
PVLIT2	x	x	Literacy scale score - Plausible value 2
PVLIT3	x	x	Literacy scale score - Plausible value 3
PVLIT4	x	x	Literacy scale score - Plausible value 4
PVLIT5	x	x	Literacy scale score - Plausible value 5
PVLIT6	x	x	Literacy scale score - Plausible value 6
PVLIT7	x	x	Literacy scale score - Plausible value 7
PVLIT8	x	x	Literacy scale score - Plausible value 8
PVLIT9	x	x	Literacy scale score - Plausible value 9
PVNUM1	x	x	Numeracy scale score - Plausible value 1
PVNUM10	x	x	Numeracy scale score - Plausible value 10
PVNUM2	x	x	Numeracy scale score - Plausible value 2
PVNUM3	x	x	Numeracy scale score - Plausible value 3
PVNUM4	x	x	Numeracy scale score - Plausible value 4
PVNUM5	x	x	Numeracy scale score - Plausible value 5
PVNUM6	x	x	Numeracy scale score - Plausible value 6
PVNUM7	x	x	Numeracy scale score - Plausible value 7
PVNUM8	x	x	Numeracy scale score - Plausible value 8
PVNUM9	x	x	Numeracy scale score - Plausible value 9
PVPSL1	x	x	Problem-solving scale score - Plausible value 1
PVPSL10	x	x	Problem-solving scale score - Plausible value 10
PVPSL2	x	x	Problem-solving scale score - Plausible value 2
PVPSL3	x	x	Problem-solving scale score - Plausible value 3
PVPSL4	x	x	Problem-solving scale score - Plausible value 4
PVPSL5	x	x	Problem-solving scale score - Plausible value 5
PVPSL6	x	x	Problem-solving scale score - Plausible value 6
PVPSL7	x	x	Problem-solving scale score - Plausible value 7
PVPSL8	x	x	Problem-solving scale score - Plausible value 8
PVPSL9	x	x	Problem-solving scale score - Plausible value 9
RACETHN_4CAT	x	x	Background - Race/ethnicity (derived, 4 categories)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
RACETHN_5CAT		x	Background - Race/ethnicity (derived, 5 categories)
RACETHN_6CAT		x	Background - Race/ethnicity (derived, 6 categories)
READHOMEUSP	x	x	Index of use of reading skills in everyday life (in prison) (prose and document texts - derived)
READHOMEUSP_WLE_CA	x	x	Index of use of reading skills in everyday life (in prison) (prose and document texts), categorised WLE (derived)
READWORK_CUSP	x	x	Index of use of reading skills at work (prose and document texts) (derived), current prison job
READWORK_CUSP_WLE_CA	x	x	Index of use of reading skills at work (prose and document texts), categorized WLE (derived), current prison job
READWORK_LUSP	x	x	Index of use of reading skills at work (prose and document texts - derived), last job prior to incarceration
READWORK_LUSP_WLE_CA	x	x	Index of use of reading skills at work (prose and document texts), categorised WLE (derived), last job prior to incarceration
READYTOLEARN	x	x	Index of readiness to learn (derived)
READYTOLEARN_WLE_CA	x	x	Index of readiness to learn, categorised WLE (derived)
REG_TL2	x	x	Geographical region - Respondent (OECD TL2) (coded)
REGION_US	x	x	Geographical region - Respondent (US Census regions)
SAMPFLAG	x	x	Flag for U.S. combined Main Study and National Supplement (Household and Prison samples)
SAMPTYPE		x	Flag for oversample
SECLGRGN		x	Source region of second language learned at home in childhood and still understand (9 regions - derived)
SECLGRGNUS_C	x	x	Source region of second language learned at home in childhood and still understand (9 regions) (2 categories) (derived from SECLGRGN)
SEQID	x	x	Sequential ID (randomly derived)
SORT_HH		x	Sort order for HH selection
SPFWT0	x	x	Final full sample weight

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
SPFWT1	x	x	Final replicate weight (1)
SPFWT10	x	x	Final replicate weight (10)
SPFWT11	x	x	Final replicate weight (11)
SPFWT12/	x	x	Final replicate weight (12)
SPFWT13	x	x	Final replicate weight (13)
SPFWT14	x	x	Final replicate weight (14)
SPFWT15	x	x	Final replicate weight (15)
SPFWT16	x	x	Final replicate weight (16)
SPFWT17	x	x	Final replicate weight (17)
SPFWT18	x	x	Final replicate weight (18)
SPFWT19	x	x	Final replicate weight (19)
SPFWT2	x	x	Final replicate weight (2)
SPFWT20	x	x	Final replicate weight (20)
SPFWT21	x	x	Final replicate weight (21)
SPFWT22	x	x	Final replicate weight (22)
SPFWT23	x	x	Final replicate weight (23)
SPFWT24	x	x	Final replicate weight (24)
SPFWT25	x	x	Final replicate weight (25)
SPFWT26	x	x	Final replicate weight (26)
SPFWT27	x	x	Final replicate weight (27)
SPFWT28	x	x	Final replicate weight (28)
SPFWT29	x	x	Final replicate weight (29)
SPFWT3	x	x	Final replicate weight (3)
SPFWT30	x	x	Final replicate weight (30)
SPFWT31	x	x	Final replicate weight (31)
SPFWT32	x	x	Final replicate weight (32)
SPFWT33	x	x	Final replicate weight (33)
SPFWT34	x	x	Final replicate weight (34)
SPFWT35	x	x	Final replicate weight (35)
SPFWT36	x	x	Final replicate weight (36)
SPFWT37	x	x	Final replicate weight (37)
SPFWT38	x	x	Final replicate weight (38)
SPFWT39	x	x	Final replicate weight (39)
SPFWT4	x	x	Final replicate weight (4)
SPFWT40	x	x	Final replicate weight (40)
SPFWT41	x	x	Final replicate weight (41)
SPFWT42	x	x	Final replicate weight (42)
SPFWT43	x	x	Final replicate weight (43)
SPFWT44	x	x	Final replicate weight (44)
SPFWT45	x	x	Final replicate weight (45)
SPFWT46	x	x	Final replicate weight (46)
SPFWT47	x	x	Final replicate weight (47)
SPFWT48	x	x	Final replicate weight (48)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
SPFWT49	x	x	Final replicate weight (49)
SPFWT5	x	x	Final replicate weight (5)
SPFWT50	x	x	Final replicate weight (50)
SPFWT51	x	x	Final replicate weight (51)
SPFWT52	x	x	Final replicate weight (52)
SPFWT53	x	x	Final replicate weight (53)
SPFWT54	x	x	Final replicate weight (54)
SPFWT55	x	x	Final replicate weight (55)
SPFWT56	x	x	Final replicate weight (56)
SPFWT57	x	x	Final replicate weight (57)
SPFWT58	x	x	Final replicate weight (58)
SPFWT59	x	x	Final replicate weight (59)
SPFWT6	x	x	Final replicate weight (6)
SPFWT60	x	x	Final replicate weight (60)
SPFWT61	x	x	Final replicate weight (61)
SPFWT62	x	x	Final replicate weight (62)
SPFWT63	x	x	Final replicate weight (63)
SPFWT64	x	x	Final replicate weight (64)
SPFWT65	x	x	Final replicate weight (65)
SPFWT66	x	x	Final replicate weight (66)
SPFWT67	x	x	Final replicate weight (67)
SPFWT68	x	x	Final replicate weight (68)
SPFWT69	x	x	Final replicate weight (69)
SPFWT7	x	x	Final replicate weight (7)
SPFWT70	x	x	Final replicate weight (70)
SPFWT71	x	x	Final replicate weight (71)
SPFWT72	x	x	Final replicate weight (72)
SPFWT73	x	x	Final replicate weight (73)
SPFWT74	x	x	Final replicate weight (74)
SPFWT75	x	x	Final replicate weight (75)
SPFWT76	x	x	Final replicate weight (76)
SPFWT77	x	x	Final replicate weight (77)
SPFWT78	x	x	Final replicate weight (78)
SPFWT79	x	x	Final replicate weight (79)
SPFWT8	x	x	Final replicate weight (8)
SPFWT80	x	x	Final replicate weight (80)
SPFWT9	x	x	Final replicate weight (9)
SUBSAMP		x	Sub-sample flag
TECHPROB		x	Technical problem flag
U01a000A	x	x	Unit01a Number of Actions
U01a000F	x	x	Unit01a Time to First Action
U01a000S	x	x	Problem-solving Unit 01a (Polytomous scored response - derived)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
U01a000T	x	x	Unit01a Total Time
U01b000A	x	x	Unit01b Number of Actions
U01b000F	x	x	Unit01b Time to First Action
U01b000S	x	x	Problem-solving Unit 01b (Dichotomous scored response - derived)
U01b000T	x	x	Unit01b Total Time
U02x000A	x	x	Unit02 Number of Actions
U02x000F	x	x	Unit02 Time to First Action
U02x000S	x	x	Problem-solving Unit 02x (Polytomous scored response - derived)
U02x000T	x	x	Unit02 Total Time
U03a000A	x	x	Unit03a Number of Actions
U03a000F	x	x	Unit03a Time to First Action
U03a000S	x	x	Problem-solving Unit 03a (Dichotomous scored response - derived)
U03a000T	x	x	Unit03a Total Time
U04a000A	x	x	Unit04a Number of Actions
U04a000F	x	x	Unit04a Time to First Action
U04a000S	x	x	Problem-solving Unit 04a (Polytomous scored response - derived)
U04a000T	x	x	Unit04a Total Time
U06a000A	x	x	Unit06a Number of Actions
U06a000F	x	x	Unit06a Time to First Action
U06a000S	x	x	Problem-solving Unit 06a (Dichotomous scored response - derived)
U06a000T	x	x	Unit06a Total Time
U06b000A	x	x	Unit06b Number of Actions
U06b000F	x	x	Unit06b Time to First Action
U06b000S	x	x	Problem-solving Unit 06b (Dichotomous scored response - derived)
U06b000T	x	x	Unit06b Total Time
U07x000A	x	x	Unit07 Number of Actions
U07x000F	x	x	Unit07 Time to First Action
U07x000S	x	x	Problem-solving Unit 07x (Dichotomous scored response - derived)
U07x000T	x	x	Unit07 Total Time
U11b000A	x	x	Unit11b Number of Actions
U11b000F	x	x	Unit11b Time to First Action
U11b000S	x	x	Problem-solving Unit 11b (Polytomous scored response - derived)
U11b000T	x	x	Unit11b Total Time
U16x000A	x	x	Unit16 Number of Actions
U16x000F	x	x	Unit16 Time to First Action

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
U16x000S	x	x	Problem-solving Unit 16x (Dichotomous scored response - derived)
U16x000T	x	x	Unit16 Total Time
U19a000A	x	x	Unit19a Number of Actions
U19a000F	x	x	Unit19a Time to First Action
U19a000S	x	x	Problem-solving Unit 19a (Dichotomous scored response - derived)
U19a000T	x	x	Unit19a Total Time
U19b000A	x	x	Unit19b Number of Actions
U19b000F	x	x	Unit19b Time to First Action
U19b000S	x	x	Problem-solving Unit 19b (Polytomous scored response - derived)
U19b000T	x	x	Unit19b Total Time
U21x000A	x	x	Unit21 Number of Actions
U21x000F	x	x	Unit21 Time to First Action
U21x000S	x	x	Problem-solving Unit 21x (Dichotomous scored response - derived)
U21x000T	x	x	Unit21 Total Time
U23x000A	x	x	Unit23 Number of Actions
U23x000F	x	x	Unit23 Time to First Action
U23x000S	x	x	Problem-solving Unit 23x (Polytomous scored response - derived)
U23x000T	x	x	Unit23 Total Time
USCIP_C		x	Education - Current Qualification - Area of Study (coded)
USCIP_C_C	x	x	Education - Current Qualification - Area of Study (combined into 4-digit categories) (derived from USCIP_C)
USCIP_H		x	Education - Highest qualification - Area of study (coded)
USCIP_H_C	x	x	Education - Highest qualification - Area of study (combined into 4-digit categories) (derived from USCIP_H)
VARSTRAT	x	x	Variance stratum
VARUNIT	x	x	Variance unit
VEFAYFAC	x	x	Fay's K factor used in creating replicate weights (BRR only)
VEMETHOD	x	x	Replication approach (string)
VEMETHODN	x	x	Replication approach (numeric)
VENREPS	x	x	Number of replicate weights used
VET	x	x	Respondent's highest level of education obtained is vocationally oriented (derived, ISCED3 and 4 only)

See note at end of table.

Table E-6. Side-by-side list of U.S. public-use file and restricted-use file variables for the prison sample—Continued

[x = included in file]

Name	U.S. public-use file	U.S. restricted-use file	Label
WRITHOMEUSP	x	x	Index of use of writing skills in everyday life (in prison) (derived)
WRITHOMEUSP_WLE_CA	x	x	Index of use of writing skills in everyday life (in prison), categorised WLE (derived)
WRITWORK_CUSP	x	x	Index of use of writing skills at work (derived), current prison job
WRITWORK_CUSP_WLE_CA	x	x	Index of use of writing skills at work, categorized WLE (derived), current prison job
WRITWORK_LUSP	x	x	Index of use of writing skills at work (derived), last job prior to incarceration
WRITWORK_LUSP_WLE_CA	x	x	Index of use of writing skills at work, categorised WLE (derived), last job prior to incarceration
YRSQUAL	x	x	Highest level of education obtained imputed into years of education (derived)
ZZ1a	x	x	Observation module: Presence of additional person
ZZ2	x	x	Observation module: Respondent understood the questions
ZZ3	x	x	Observation module: Clarification necessary
ZZ4_01	x	x	Observation module: Respondent held a conversation with someone else
ZZ4_05	x	x	Observation module: Television, radio, game console or music player was in use in the vicinity of respondent
ZZ4_06	x	x	Observation module: Respondent was interrupted by some other activity, task or event
ZZ5	x	x	Observation module: Assessment taking too long
ZZ6USP	x	x	Observation module: Room of assessment (prison)

SOURCE: U.S. Department of Education, National Center for Education Statistics, U.S. Program for the International Assessment of Adult Competencies (PIAAC), 2012/2014.