Education Longitudinal Study of 2002 (ELS:2002) Postsecondary Education Transcript Study: Addendum to the Public-Use File

Read-Me Documentation

Content Contact
Elise M. Christopher
(202) 502-7899
elise.christopher@ed.gov

U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION STATISTICS

Institute of Education Sciences

This page intentionally left blank

Contents

List of Tables

This page intentionally left blank

1. Addendum Purposes and Overview

The Education Longitudinal Study of 2002 is sponsored by the National Center for Education Statistics, and follows nationally representative samples of both a sophomore cohort (starting in 2002) and a senior cohort (starting in 2004) to a final interview at about age 26 in 2012. The Education Longitudinal Study of 2002 Postsecondary Education Transcript Study (ELS:2002 PETS) gathered institutional data concerning the post-high school coursetaking, credit accrual, and grades of the subset of the sample that had pursued postsecondary education by the time of the final interview.

The Public-Use Addendum to ELS:2002 PETS is designed to serve two related purposes. In particular, the Addendum supplies information about the contents of the public- (NCES 2015-314) and restricted-use (NCES 2015-035) files that researchers can consult in order to plan their research. This information will enable potential analysts to assess whether they will need to apply for a license to obtain the restricted data. Indeed, the contents of this Addendum should help analysts to determine whether their intended analyses can be conducted with public-use data, with restricted-use data, or with neither or both. Because the data file documentation (DFD) (Ingels et al. 2015, NCES 2015-033) is based on the restricted-use files, the documentation itself is restricted. Like the data, the DFD can only be accessed by those who have obtained a license from the National Center for Education Statistics (NCES) for doing so.

The public-use variables are, of course, a subset of the restricted-use variables. The public-use variables reflect, in many cases, the fact of their suppression, but also, even where the select data do appear on the public files, it is often in modified form (for example, a continuous variable might be categorized, or a categorical variable might be recoded into grosser categories, among other possibilities). This constraint applies to the composite variables as well as other measures: postsecondary transcript composite variables may be suppressed in the public version, or they may be modified, or they may be available in precisely the same form as the restricted variables. One way of looking at the utility of the public-use version is that it may afford opportunities to do preliminary or exploratory work in an area through the public files, and move on at a later time to tap the greater specificity and richness of the restricted files, which contain data on specific courses, credits, and grades that researchers can utilize in in-depth follow-up analysis. For some research questions, however, the difference between the public and restricted files is categorical rather

than just a matter of degree (say, coarsening). Whole subfiles, in some cases, have been excluded from the public-use data—for example, federal loan and grant information is restricted only, as are transcript-institutional microdata, GED testing program data, and CPS (Central Processing System data, which report information from the Free Application for Federal Student Aid).

This Addendum provides basic information about ELS:2002 PETS data collection and weights. In addition to addressing the purposes of the Addendum (section 1) it briefly discusses data collection procedures and results (section 2), composite variables (section 3), use of weights in analysis of the public-use file (section 4), and the content and structure of the public-use data files (section 5), as well as the tools available for analyzing ELS:2002 PETS data (section 6).

An appendix displays the Table of Contents for the restricted Data File Documentation, so that potential users can see what further information is available to licensees. A second appendix supplies an extended discussion of issues in the analytical use of case weights. A brief References section is also included.

2. Background: Data Collection Procedures and Results

The third follow-up (2012) of ELS:2002 provides a picture of outcomes at about age 26, a time when most sample members had completed the transition to adulthood. As of 2012, for the sophomore cohort, some 16 percent had no postsecondary enrollment and hence fall outside of PETS, while 33 percent had a bachelor's degree or higher and 51 percent had at least some postsecondary enrollment or an associate's degree or undergraduate certification (Lauff and Ingels 2014). Postsecondary transcripts were requested for each of the ELS:2002 sample members who reported in the student interview that they had attended an Integrated Postsecondary Education Data System (IPEDS) institution. Some 12,549 sample members were so identified. (Non-IPEDS institutions, such as postsecondary institutions outside the United States, were not included in the transcript collection.) At the student level, a transcript was received from at least one institution for 11,623 sample members. The weighted response rate was 77 percent.

The fielded institution sample for the postsecondary transcript collection included 3,666 distinct IPEDS institutions. Of the 3,666 institutions, 2 percent were

determined to be ineligible because the institution had closed or because a sample member had reported a school that was not a postsecondary institution. Of the remaining 3,598 institutions, 3,309 (92 percent) provided transcripts. For further information on postsecondary transcript data collection results, and the coding and processing of the transcript data, see the restricted-use *ELS:2002 Postsecondary Education Transcript Study Data File Documentation (DFD)* (Ingels et al. 2015, NCES 2015-033).

Sample sizes for eligible (that is, all nondeceased) cohort members and response rate data from the base year through third follow-up (including PETS) are summarized in table 1 below.

Table 1. Eligible sample and weighted response rates for ELS:2002		
Study round	Eligible[1]	Weighted response rate
Base-year	17,754	87.4
First follow-up	16,733	88.7
Second follow-up	16,700	81.9
Third follow-up	16,562	77.8
Postsecondary transcripts	12,549	77.2

[1] Eligibility is based on membership in the sophomore or freshened senior cohort. Detected sampling errors (based on erroneous school report of grade level)—and deceased individuals—have been removed. In addition, sample members who did not report postsecondary education participation have been removed from the postsecondary transcript response rate denominator.
NOTE: The numbers of eligible sample members reported here, as well as the weighted response rates reported here, may differ somewhat from statistics reported in previous ELS:2002 documentation, owing to changing status of questionnaire-incapable sample members over time, and factors such as the selective use of rounding (in the second follow-up, data were initially released in restricted form, which employs rounding as a data perturbation technique helping to protect against disclosure of respondent identity). Note also that while table 1 reports response rates, participation rates based on the fielded sample only (thus excluding temporary out-of-scopes from the response rate denominator) are reported separately, using unweighted response rates, and appear in the *ELS:2002 Postsecondary Education Transcript Study Data File Documentation* (DFD) (Ingels et al. 2015).
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

3. Composite Variables

Composite variables—also called constructed, derived, or created variables—are usually generated using responses from two or more questionnaire items, from logical or stochastic imputation, or from recoding of a variable (often for disclosure avoidance reasons). Some are copied from another source (e.g., a variable supplied in sampling or imported from an external database). Examples of composite variables include school variables (school sector, urbanicity, region of the country); high school completion status; math assessment scores (achievement quartile [quarter] in math); demographic variables (socioeconomic status [SES], sex, race, Hispanic

ethnicity, and month and year of birth); and employment histories. Most of the composite variables can be used as classification variables or independent variables in data analysis. The ELS:2002 PETS data file includes many composite variables for the convenience of data users.

Some examples of postsecondary transcript variables that can be found on the ELS:2002 PETS public-use file include F3TZPS1SEC (sector of first known postsecondary institution attended as reported to IPEDS); F3TZPS1SLC (selectivity of the first known postsecondary institution attended); F3TZPS1START (month/year the student first attended the postsecondary institution); F3TZANYDEGRE (any known degree attained as of June 2013); F3TZHIGHDEG (highest known degree attained as of June 2013); F3TZDEG1SEC (sector of institution where first known degree attained; F3TZDEG1SLC (institutional selectivity where first degree attained); F3TZCERTLTDT (date of most recent known certificate); F3TZASCLCIP2 (most recent known associate's degree major/field of study as 2-digit CIP code); F3TZASCIERN (all science courses known credits earned); F3TZASCIGPA (all science courses known GPA); and F3TZCALCERN (known credits earned in calculus/advanced math).

On the other hand, examples of postsecondary transcript variables NOT included on the public-use file, include F3TZCRTLCIP6 (most recent known certificate major/field of study, 6-digit CIP code); F3FEDLNCUM (cumulative federal loan amount borrowed for undergraduate and graduate education); and F3FEDLNDUE (total amount owed, principal and interest, on all Federal direct loans).

4. Using Case Weights for ELS:2002 PETS Public-Use File Analysis

The general purpose of weighting data is to compensate for unequal probabilities of selection and to adjust for the effects of nonresponse. Weights are typically calculated in two main steps. In the first step, unadjusted weights (called base weights or design weights—these are not available on the data files) are calculated as the inverse of the probabilities of selection, taking into account all stages of the sample selection process. In the second step, these initial weights are adjusted to compensate for nonresponse. Weights are critical for population estimation. Analytic weights should be used in combination with software that accounts for the ELS:2002 complex survey design to produce estimates for the target population, with

appropriate standard errors. The ELS:2002 PETS weights, further, were specifically designed to support analyses that incorporate information gleaned from postsecondary transcripts in such a way as to retain the generalizability of results to the two ELS:2002 cohorts, the 10th- and 12th-grade student populations. Students who attended postsecondary institutions may be in either cohort or both. Both Taylor-series weights and balanced repeated replicate (BRR) weights are available on the ELS:2002 PETS public-use data files.

Eight student weights were computed for ELS:2002 PETS (for detailed information about these weights, and the weighting process, see chapter 5 of the ELS:2002 PETS DFD [Ingels et al. 2015]). The eight new weights and the types of analyses that may be conducted using these weights are summarized below.

Table 2. PETS analysis weights	
Weight Variable Name	Potential use
PSWT PSTSCWT	Analyses based on postsecondary transcript data – alone (PSWT) or in conjunction with high school transcript data (PSTSCWT)
F3QPSWT F3QTSCPSWT	Analyses based on postsecondary transcript data and third follow-up data – alone (F3QPSWT) or in conjunction with high school transcript data (F3QTSCPSWT).
F3BYPNLPSWT F3BYTSCPSWT	Analyses based on postsecondary transcript data and third follow-up data in combination with base-year data (or other prior rounds of data) – alone (F3BYPNLPSWT) or in conjunction with high school transcript data (F3BYTSCPSWT).
F3F1PNLPSWT F3F1TSCPSWT	Analyses based on postsecondary transcript data and third follow-up data in combination with first follow-up data (or second follow-up data) – alone (F3F1PNLPSWT) or in conjunction with high school transcript data (F3F1TSCPSWT).

Further detail on weighting issues in analysis may be found in appendix B of this Addendum.

5. ELS:2002 PETS Public Use Data Files: Content and Structure

To protect the confidentiality of NCES data that contain information about specific individuals, ELS:2002 data were subject to various procedures to minimize disclosure risk. These procedures are described in detail in the two third follow-up DFDs (NCES 2014-364 and NCES 2015-033). As a result of disclosure analysis and procedures, the public-use version of the data reflects a reduced set of the data

gathered, reflecting suppression of some variables and alteration (for example, top- or bottom coding) of others. As a final step, the ELS:2002 data were analyzed to confirm that the disclosure limitation techniques did not compromise the analytic utility of the data.

This section describes the general structure of the ELS:2002 postsecondary education transcript study (PETS) data files and the individual file components.

The longitudinal postsecondary transcript study data file combines the full complement of base-year to third follow-up data (including high school transcript data) with postsecondary transcript data and related composite variables. A summary of the comprehensive set of files that comprises the PETS study data file release is provided in table 3. Nonpublic files are noted in the table.

Table 3.	Summary of files included within the base-year to third follow-up data file			
File name	Accessibility	Level	Record count	Longitudinal composition
Student file (BYF3PSTSTU)	Public and restricted use	Student level	16,197	BY, F1, F2, and F3 student interview, BY parent interview, BY teacher interview, school-level (BY and F1) administrator and staff interviews, school-level composites, HS transcript composites (F1), postsecondary educational transcript composites (F3), extant sources
School file (BYF1TSCH)	Public and restricted use	School level	1,954	BY and F1 administrator interview, BY facilities checklist, HS transcript (F1), extant sources
HS student transcript file (HSTRNSTU)	Restricted use	Student-course level	638,967	HS student transcripts (F1)
HS school transcript file (HSTRNSCH)	Restricted use	School-course level	117,151	HS course catalog (F1)
PS transcript school file (PSTINST)	Restricted use	Institution level	3,554	PS student transcripts (F3)
PS transcript student-school file (PSTSTUINST)	Restricted use	Student-school level	23,702	PS student transcripts (F3)
PS transcript term file (PSTTERM)	Restricted use	Student-term level	140,285	PS student transcripts (F3)
PS transcript course file (PSTCOURSE)	Restricted use	Student-course level	558,257	PS student transcripts (F3)
PS transcript degree/major file (PSTDEGMJR)	Restricted use	Student-degree/major level	25,386	PS student transcripts (F3)
PS transcript entrance exam/test file (PSTTEST)	Restricted use	Student level	4,650	PS student transcripts (F3)
F2 student-institution file (F2INST)	Public and restricted use	Student-institution level	33,495	F2 interview, IPEDS
F3 student-institution file (F3INST)	Public and restricted use	Student-institution level	20,951	F2 and F3 interview, IPEDS

See notes at end of table.

Table 3.   Summary of files included within the base-year to third follow-up data file—Continued

File name	Accessibility	Level	Record count	Longitudinal composition
CPS 2004-05 file (CPS0405)	Restricted use	Student level	6,484	CPS, academic year 2004–05
CPS 2005-06 file (CPS0506)	Restricted use	Student level	5,261	CPS, academic year 2005–06
CPS 2006-07 file (CPS0607)	Restricted use	Student level	4,343	CPS, academic year 2006–07
CPS 2009-10 file (CPS0910)	Restricted use	Student level	3,296	CPS, academic year 2009–10
CPS 2010-11 file (CPS1011)	Restricted use	Student level	3,143	CPS, academic year 2010–11
CPS 2011-12 file (CPS1112)	Restricted use	Student level	2,408	CPS, academic year 2011–12
CPS 2012-13 file (CPS1213)	Restricted use	Student level	2,721	CPS, academic year 2012–13
CPS 2013-14 file (CPS1314)	Restricted use	Student level	1,074	CPS, academic year 2013–14
NSLDS Pell file[1] (NSLDSPL)	Restricted use	Student level	18,277	NSLDS
NSLDS Loan file[2] (NSLDSLN)	Restricted use	Student level	52,665	NSLDS
GED file (GED)	Restricted use	Student level	496	GED
Student BRR weight file (BRRSTU)	Public and restricted use	Student level	16,197	BY, F1, HS transcript, F2, F3 and PS transcript replicate weights
School BRR weight file (BRRSCH)	Public and restricted use	School level	751	BY and F1 replicate weights

[1] The NSLDS Pell file contains all Pell records available as of the third follow-up. Additional notes related to the cumulative nature of this file are provided in section 6.1.7.
[2] The NSLDS loan file contains all loan records available as of the third follow-up. Additional notes related to the cumulative nature of this file are provided in section6.1.7.
NOTE: BRR = balanced repeated replicate. BY = base year. CPS = Central Processing System. F1 = first follow-up. F2 = second follow-up. F3 = third follow-up. GED = General Educational Development credential. HS = high school. IPEDS = Integrated Postsecondary Education Data System. NSLDS = National Student Loan Data System.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

As table 3 indicates, only a subset of the various files is available in a public-use version; other files are categorically excluded and have a status of restricted-only. More specifically, the **student file**, **school file**, **student institutional files for the second follow-up** and **student institutional files for the third follow-up** contain public-use variables (although even within these files, specific questionnaire items or composites may be altered or suppressed). In addition, BRR weights are available at both the student and school level. Files containing exclusively restricted data include the eight transcript files from HSTRNSTU through PSTTEST, the eight CPS (Central Processing System—financial aid application) files, the two National Student Loan Data System (NSLDS) files (Pell grant and loan) and the GED (General Educational Development credential) files.

Sections of the student file (BYF3PSTSTU) are as follows:

- ID and Universe Variables;
- Base-year (BY) Weights and Composites;
- First follow-up (F1) Weights and Composites;
- High School Transcript Weights and Composites;
- Second follow-up (F2) Weights and Composites;
- F2 Extant Data Source Composites;
- Third follow-up (F3) Weights and Composites;
- Postsecondary Transcript Weights and Composites;
- F3 Extant Data Source Composites;
- BY Student Questionnaire;
- F1 Student Questionnaire;
- F1 Dropout Questionnaire;
- F1 Transfer Questionnaire;
- F1 Early Graduate Questionnaire;
- F1 New Participant Supplement;
- F2 Student survey;
- F3 Student survey;
- BY Parent Questionnaire;
- BY Teacher Questionnaire (English);
- BY Teacher Questionnaire (Math);
- BY School Composites;
- F1 School Composites;
- BY Administrator Questionnaire;
- F1 Administrator Questionnaire;
- BY Library Questionnaire; and
- BY Facilities Checklist.

The *high school* file (BYF1TSCH) reflects data for the base-year survey, first follow-up survey, and first follow-up high school transcript (course offerings) data composites. The *second follow-up postsecondary student-institution* file (F2INST) is a student-level file that links students to postsecondary institutions applied to or attended as of the second follow-up interview. The *third follow-up postsecondary student-institution attendance* file (F3INST) is a student-level file that links students to postsecondary institutions for which attendance was reported as of the third follow-up interview.

6. Analysis Tools

ELS:2002 data have been made available in public-use form via the web-based EDAT and PowerStats applications and (for licensed users) in restricted-use form[1] (NCES 2015-035) via ECB format. The restricted-use ECB is installed from a DVD and is designed to be run in a Windows environment. ECB software packages are available at no cost from NCES. A summary of all data access and analysis tools is provided in table 4.

Table 4. Summary of ELS:2002 data access and analysis tools		
ELS:2002 Base-year to Third Follow-up		
Detail	Public use	Restricted use
Longitudinal composition	BY, F1, F2, F3, postsecondary transcripts	BY, F1, F2, F3, high school transcripts, postsecondary transcripts
Documentation publication number	NCES 2015-314	NCES 2015-035
Access point	EDAT, PowerStats	Request DVD under license
Data file format	EDAT: One of six programming languages (SAS, Stata, SPSS, Sudaan, R, S-Plus) (or ASCII or CSV) PowerStats: Does not allow extraction of raw data	One of three programming languages (SAS, SPSS, and Stata)
Data file variable selection	EDAT: Tag Files PowerStats: Does not provide access to microdata	Electronic codebook
Analysis capability	EDAT: Does not support interactive analysis; provides access to microdata PowerStats: General statistics (percentages, means, etc.), data tables, table regressions	Does not support interactive analysis; provides access to micro data

ELS:2002 Base-year to Third Follow-up Data Availability						
	Base year	First follow-up	High school transcripts	Second follow-up	Third follow-up	Postsecondary transcripts
Year conducted	2002	2004	2005	2006	2012	2013–2014
Date available	Now	Now	Now	Now	Now	Spring 2015
Restricted-use (DVD) NCES 2015-035	✔	✔	✔	✔	✔	Released as separate file
Public-use (EDAT) NCES 2015-314	✔	✔		✔	✔	

NOTE: BY = base year. EDAT = Educational Data Analysis Tool. F1 = first follow-up. F2 = second follow-up. F3 = third follow-up. SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

There are three basic analysis tools available for ELS:2002 data: the restricted-use Electronic Codebook, public-use Educational Data Analysis Tool, and PowerStats.

Electronic Codebook (ECB). The ECB is available only in restricted-use versions of the data files. The ECB system serves as an electronic version of a fully documented

---

[1] A license is required to access the restricted-use ECB (http://nces.ed.gov/statprog/confid.asp).

survey codebook. It offers full access to all available microdata and composites. It allows the data user to browse through the ELS:2002 variables contained on the data files; search variable and value names for keywords related to particular research questions; review the wording of these items along with notes and other pertinent information related to them; examine the definitions and programs used to develop composite and classification variables; and export SAS, SPSS, or Stata syntax programs for statistical analysis. The ECB also provides an electronic display of the distribution of counts and percentages for each variable in the dataset. Analysts can use the ECB to select or "tag" variables of interest; export codebooks that display the distributions of the tagged variables; and generate SAS, SPSS, and Stata program code (including variable and value labels) that can be used with the analyst's own statistical software. A guide for using the ECB is accessible via the "View Help" option of the "Help" tab in the ECB's main menu.

Education Data Analysis Tool (EDAT). The EDAT web application, available via the NCES website (http://nces.ed.gov/EDAT), allows users to download public-use data files in formats compatible with one of six statistical programming languages (SAS, SPSS, S-Plus, Stata, R, and SUDAAN). The contents of downloaded files can be customized via the application's "tag" functionality by which users can select variables pertaining to their specific research interests. Additional information regarding the EDAT application is available via the NCES website. The ELS:2002 PETS public-use files are identified by the NCES publication number, 2015-314.

PowerStats. The PowerStats web application, available via the NCES website, allows users to generate simple statistics such as data tables and regressions via an interactive interface. The application also allows users to store work associated with their analyses or import (via external files) previously generated analyses. Additional information regarding the PowerStats application is available via the NCES website.

References

Ingels, S.J., Pratt, D.J., Alexander, C.P., Jewell, D.M., Lauff, E., Mattox, T.L., Wilson, D. (2014). *Education Longitudinal Study of 2002 (ELS:2002) Base-Year to Third Follow-Up Data File Documentation* (NCES 2014-364). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC.

Ingels, S.J., Pratt, D.J., Alexander, C.P., Bryan, M., Jewell, D.M. Lauff, E., Mattox, T.L., and Wilson, D. (2015). *Education Longitudinal Study of 2002 (ELS:2002) Postsecondary Education Transcript Study Data File Documentation* (NCES 2015-033). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Note: *restricted* document available only to licensees.

Lauff, E., and Ingels, S.J. (2014). *Education Longitudinal Study of 2002 (ELS:2002): A First Look at 2002 High School Sophomores 10 Years Later* (NCES 2014-363). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Lauff, E., and Ingels, S.J. (2015). *Education Longitudinal Study of 2002 (ELS:2002): A First Look at the Postsecondary Transcripts of 2002 High School Sophomores.* (NCES 2015-034). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, DC.

This page intentionally left blank

Appendix A: Table of Contents, ELS:2002 PETS DFD (NCES 2015-033)

This page intentionally left blank

Appendix B: Detailed Information for Use of Weights

PSWT and PSTSCWT analytic weights. The weights PSWT and PSTSCWT were constructed to support analysis of postsecondary transcript data alone, using PSWT or in conjunction with high school (first follow-up) transcript data using PSTSCWT. For an example of transcript data analysis using the sophomore cohort in conjunction with PSWT—the weight developed to support analyses of the postsecondary transcript data either alone or in conjunction with high school transcript data)—see Lauff and Ingels, 2015.

PSTSCWT is unique among the set of PETS weights in that it is non-zero for any second or third follow-up respondent for whom a postsecondary transcript was collected from an IPEDS institution. Because of its construction, some sample members with a non-zero value of this weight will not have interview data from one or more rounds or high school transcript data. Care should be taken to consider the set of data to be used with the analysis weight PSTSCWT to avoid excluding too many records with a nonmissing weight value. As a general rule of thumb, PSTSCWT will most likely be the best available weight to use when incorporating postsecondary transcript data in analyses.

It is also important to keep cohort membership closely in mind when analyzing PETS data. As noted above, postsecondary transcript respondents may be in the population of Spring-term 2002 10th-grade students, or in the population of Spring-term 2004 12th-grade students, or in both populations. Analyses designed to assess characteristics of one of the populations must take care to restrict analyses to those postsecondary transcript respondents in the population of interest. To identify those postsecondary transcript respondents who are members of the two student populations, two flag variables, G10COHRT and G12COHRT, are available. Those PETS respondents with a value of 1 for G10COHRT are members of the population of Spring 2002 10th-grade students. Those PETS respondents with a value of 1 (determined in the first follow-up) or a value of 2 (determined in the second follow-up) for G12COHRT are members of the population of Spring 2004 12th-grade students. The combinations of these weights can be summarized as follows:

- PSWT when used in conjunction with G10COHRT permits estimates from postsecondary transcript data to be representative of students enrolled in 10th grade in Spring term 2002.

- PSWT when used in conjunction with G12COHRT permits estimates from postsecondary transcript data to be representative of students enrolled in 12th grade in Spring term 2004.
- PSTSCWT when used in conjunction with G10COHRT permits estimates from postsecondary transcript data in combination with high school transcript data to be representative of students enrolled in 10th grade in Spring term 2002.
- PSTSCWT when used in conjunction with G12COHRT permits estimates from postsecondary transcript data in combination with high school transcript data to be representative of students enrolled in 12th grade in Spring term 2004.

Table B-1 summarizes, from the base year to the third-follow-up PETS, the ELS:2002 analysis weights along with associated universe flags, populations, and respondents:

Table B-1. Relationship among student weights, universe flags, populations, and respondents: 2002–2014

Weight	Universe flag	Population	Respondent
BYSTUWT	G10COHRT	A—Spring 2002 10th-grader	Respondent to questionnaire in 2002
BYEXPWT	G10COHRT	A—Spring 2002 10th-grader	Respondent to questionnaire in 2002 or were incapable of completing a questionnaire
F1PNLWT	G10COHRT	A—Spring 2002 10th-grader	Respondent to questionnaire in 2002 and 2004 (base-year data may be from the new participant supplement[1] or imputed)
F1XPNLWT	G10COHRT	A—Spring 2002 10th-grader	Respondent to questionnaire in 2002 and 2004 (base-year data may be from the new participant supplement or imputed) or were incapable of completing a questionnaire in 2002 or 2004
F1QWT	G10COHRT[1] G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Respondent to questionnaire in 2004
F1EXPWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Respondent to questionnaire in 2004 or were incapable of completing a questionnaire in 2004
F1TRSCWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	High school transcript respondent and respondent to first follow-up or base-year questionnaire or were members of the expanded sample
F2QWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Respondent to questionnaire in 2006
F2QTSCWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Respondent to questionnaire in 2006 and fully or partially complete high school transcript data
F2F1WT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Respondent to questionnaire in 2004 and 2006 or were incapable of completing a questionnaire in 2004 and respondent to questionnaire in 2006
F2BYWT	G10COHRT	A—Spring 2002 10th-grader	Respondent to questionnaire in 2002 and 2006 or were incapable of completing a questionnaire in 2002 and respondent to questionnaire in 2006
F3QWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Respondent to questionnaire in 2012

See notes at end of table.

Table B-1. Relationship among student weights, universe flags, populations, and respondents: 2002–2014—Continued

Weight	Universe flag	Population	Respondent
F3QTSCWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Respondent to questionnaire in 2012 and high school transcript respondent
F3BYPNLWT	G10COHRT	A—Spring 2002 10th-grader	Respondent to questionnaire in 2002 and 2012 or were incapable of completing a questionnaire in 2002 and respondent to questionnaire in 2012
F3BYTSCWT	G10COHRT	A—Spring 2002 10th-grader	Respondent to questionnaire in 2002 and 2012 or were incapable of completing a questionnaire in 2002 and respondent to questionnaire in 2012 and high school transcript respondent
F3F1PNLWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Respondent to questionnaire in 2004 and 2012 or were incapable of completing a questionnaire in 2004 and respondent to questionnaire in 2012
F3F1TSCWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Respondent to questionnaire in 2004 and 2012 or were incapable of completing a questionnaire in 2004 and respondent to questionnaire in 2012 and high school transcript respondent
PSWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Postsecondary transcript respondent
PSTSCWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Postsecondary transcript respondent and high school transcript respondent
F3QPSWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Postsecondary transcript respondent and respondent to questionnaire in 2012
F3QTSCPSWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Postsecondary transcript respondent and respondent to questionnaire in 2012 and high school transcript respondent
F3BYPNLPSWT	G10COHRT	A—Spring 2002 10th-grader	Postsecondary transcript respondent and respondent to questionnaire in 2002 and 2012 or were incapable of completing a questionnaire in 2002 and respondent to questionnaire in 2012
F3BYTSCPSWT	G10COHRT	A—Spring 2002 10th-grader	Postsecondary transcript respondent and respondent to questionnaire in 2002 and 2012 or were incapable of completing a questionnaire in 2002 and respondent to questionnaire in 2012 and high school transcript respondent
F3F1PNLPSWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Postsecondary transcript respondent and respondent to questionnaire in 2004 and 2012 or were incapable of completing a questionnaire and respondent to questionnaire in 2012
F3F1TSCPSWT	G10COHRT G12COHRT	A—Spring 2002 10th-grader B—Spring 2004 12th-grader	Postsecondary transcript respondent and respondent to questionnaire in 2004 and 2012 or were incapable of completing a questionnaire in 2004 and respondent to questionnaire in 2012 and high school transcript respondent

[1] As noted in the ELS:2002 base-year to first follow-up data file documentation, base-year nonrespondents who responded in the first follow-up are considered to be members of the Spring 2002 10th-grade population, but there is no base-year weight (BYSTUWT or BYEXPWT) for them. The new participant supplement employed in the first follow-up ensured that the standard classification variables collected in the base year were also available for this group. However, key variables were imputed for base-year nonrespondents who were first follow-up respondents, so that these students could be analyzed as part of the sophomore cohort using F1PNLWT or F1XPNLWT. These students who are third follow-up respondents may also be analyzed as part of the sophomore cohort using F3F1PNLWT.

NOTE: First follow-up sample freshening resulted in the inclusion of students who were members of the 12th-grade cohort but not the 10th-grade cohort so the G10COHRT flag is required to restrict analyses to those first follow-up respondents who are members of the 10th-grade cohort. Similarly, the G12COHRT flag is required to restrict analyses to those first follow-up respondents who are members of the 12th-grade cohort.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

F3QPSWT and F3QTSCPSWT Analytic Weights. The weights F3QPSWT and F3QTSCPSWT were constructed to support analysis of postsecondary transcript data and third follow-up data alone (F3QPSWT) or in conjunction with high school (first follow-up) transcript data (F3QTSCPSWT). To identify those postsecondary transcript respondents who are members of the two student populations, two flag variables, G10COHRT and G12COHRT, are available. Those postsecondary transcript respondents with a value of 1 for G10COHRT are members of the population of Spring-term 2002 10th-grade students. Those postsecondary transcript respondents with a value of 1 (determined in the first follow-up) or a value of 2 (determined in the second follow-up) for G12COHRT are members of the population of Spring-term 2004 12th-grade students. The combination of these weights can be summarized as follows:

- F3QPSWT when used in conjunction with G10COHRT permits estimates from postsecondary transcript data and third follow-up survey data to be representative of students enrolled in 10th grade in Spring term 2002.
- F3QPSWT when used in conjunction with G12COHRT permits estimates from postsecondary transcript data and third follow-up survey data to be representative of students enrolled in 12th grade in Spring term 2004.
- F3QTSCPSWT when used in conjunction with G10COHRT permits estimates from postsecondary transcript data and third follow-up survey data in combination with high school transcript data to be representative of students enrolled in 10th grade in Spring term 2002.
- F3QTSCPSWT when used in conjunction with G12COHRT permits estimates from postsecondary transcript data and third follow-up survey data in combination with high school transcript data to be representative of students enrolled in 12th grade in Spring term 2004.

The analytic weights *F3BYPNLPSWT* and *F3BYTSCPSWT* were constructed to support estimates of students enrolled in 10th grade in Spring term 2002. F3BYPNLPSWT was produced for all ELS:2002 sample members who responded[2] in the base year and in the third follow-up and have a postsecondary transcript. F3BYTSCPSWT was produced for all ELS:2002 sample members who responded in the base year and in the third follow-up and have a sufficient amount of postsecondary transcript and high school (first follow-up) transcript data. It is not

---

[2] Sample members who did not respond in the base year but did respond in the first follow-up were administered new participant supplemental items on the questionnaire to gather some of the same information that was collected on base-year respondents. Consequently, these base-year nonrespondents who responded in the first follow-up were treated as base-year respondents in the construction of first and second follow-up longitudinal weights. These sample members were treated as base-year respondents in the construction of PETS weights.

necessary to use the flag variable G10COHRT in conjunction with F3BYPNLPSWT or F3BYTSCPSWT; by definition, only postsecondary transcript and third follow-up respondents who are members of the Spring-term 2002 10th-grade population will have a non-zero value for these weights. However, it is possible that statistical software not designed for the analysis of sample survey data may fail to exclude records that have analysis weights of zero. The G10COHRT flag may be used to specifically restrict analyses to members of the 10th-grade cohort to avoid such a situation from arising.

As longitudinal studies progress, the set of possible weights expands. Analytic weights produce estimates representative of specific populations and also adjust for patterns of potential differential nonresponse to survey components. F3BYPNLPSWT and F3BYTSCPSWT are intended to support analysis of postsecondary transcript data and third follow-up data in combination with base-year data (F3BYNLPSWT) and high school transcript data (F3BYTSCPSWT). For analysis with data from the aforementioned components and any other combination of first, second, and third follow-up data, selection of an appropriate analysis weight requires careful consideration of the analyses to be conducted, the data to be used, and the pattern of item nonresponse.

The weights *F3F1PNLPSWT* and *F3F1TSCPSWT* were constructed primarily to support estimates of Spring-term 2004 12th-grade students. However, these weights may also be used to support estimates of Spring-term 2002 10th-grade students, depending on the data used to generate estimates. F3F1PNLPSWT was produced for all ELS:2002 sample members who responded in the first follow-up and in the third follow-up and have a postsecondary transcript. F3F1TSCPSWT was produced for sample members who responded in the first follow-up and in the third follow-up and have a sufficient amount of postsecondary transcript and high school (first follow-up) transcript data. As with PSWT and PSTSCWT, these weights need to be used in combination with a cohort flag. Again, to identify those postsecondary transcript respondents who are members of the two student populations, two flag variables, G10COHRT and G12COHRT, are available. Postsecondary transcript respondents with a value of 1 for G10COHRT are members of the population of Spring-term 2002 10th-grade students. Postsecondary transcript respondents with a value of 1 (determined in the first follow-up) or a value of 2 (determined in the second follow-up) for G12COHRT are members of the population of Spring-term 2004 12th-grade students. In other words, postsecondary transcript weights for the 12th-grade cohort are constructed for all sample members where G12COHRT equals 1 or 2, whereas first follow-up weights are constructed for all sample members where G12COHRT equals 1. Consequently, when using postsecondary transcript weights, the 12th-grade cohort is identified by restricting sample members

to those individuals with a value of 1 or 2 for G12COHORT. When using first follow-up weights, the 12th-grade cohort is identified by restricting sample members to those individuals with a value of 1 for G12COHRT. The combinations can be summarized as follows:

- F3F1PNLPSWT when used in conjunction with G10COHRT permits estimates from the postsecondary transcript data and third follow-up survey in combination with first follow-up data to be representative of students enrolled in 10th grade in Spring term 2002.[3]
- F3F1PNLPSWT when used in conjunction with G12COHRT permits estimates from the postsecondary transcript data and third follow-up survey in combination with first follow-up data to be representative of students enrolled in 12th grade in Spring term 2004.
- F3F1TSCPSWT when used in conjunction with G10COHRT permits estimates from the postsecondary transcript data and third follow-up survey data in combination with first follow-up data and high school transcript data to be representative of students enrolled in 10th grade in Spring 2002.
- F3F1TSCPSWT when used in conjunction with G12COHRT permits estimates from the postsecondary transcript data and third follow-up survey data in combination with first follow-up data and high school transcript data to be representative of students enrolled in 12th grade in Spring term 2004.

Note that these two weights are designed to support analyses that incorporate data from the PETS, third follow-up, first follow-up, and high school transcript data collection. If an analysis also incorporates data from the base year or the second follow-up, then these weights may not be preferred. For example, F3BYPNLPSWT and F3BYTSCPSWT may be preferred when base-year data are included in an analysis.

Again, as longitudinal studies progress, the set of possible weights dramatically expands. Analytic weights produce estimates representative of specific populations and also adjust for patterns of potential differential nonresponse to survey components. F3F1PNLPSWT and F3F1TSCPSWT are intended to support analysis of postsecondary transcript data and third follow-up data in combination with first follow-up data (F3F1PNLPSWT) and high school transcript data (F3F1TSCPSWT). For analysis with data from the aforementioned components and in combination

---

[3] Note that the branching of various pathways offered in the student questionnaires in the first follow-up gives rise to component skips that could result in a particular analysis excluding members of the 10th-grade cohort. For example, analysis of dropouts as of the first follow-up necessarily excludes (except as a comparison group) other members of the 10th-grade cohort.

with second follow-up data or base-year data, selection of an appropriate analysis weight requires careful consideration of the analyses to be conducted, the data to be used, and the pattern of item nonresponse.

Postsecondary Transcript Weights and Prior-round Weights. In both the base year and first follow-up of ELS:2002, some sample members were not able to complete the sample member questionnaires because of limited English proficiency or because of physical or mental limitations. However, information was collected from individuals, such as school administrators and teachers, associated with these sample members. In the base year and first follow-up, the set of respondents in each round combined with the set of sample members who were questionnaire-incapable in that round was referred to as the expanded sample for that round. Analysis weights were created just for those sample members who were respondents and were created for the expanded sample. Expanded sample weights are only included in restricted-use files. In the second and third follow-up and PETS, however, questionnaire-incapable sample members were treated as nonrespondents in those rounds.