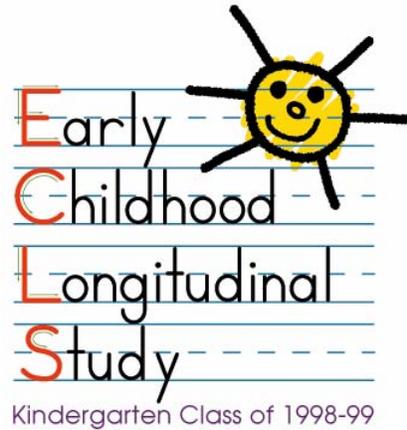

NATIONAL CENTER FOR EDUCATION STATISTICS



EARLY CHILDHOOD LONGITUDINAL STUDY, KINDERGARTEN CLASS OF 1998–99 (ECLS-K)

USER’S GUIDE TO THE LONGITUDINAL KINDERGARTEN–THIRD GRADE PUBLIC-USE DATA FILE NCES 2004-088

Prepared by

Westat

Karen Tourangeau
Christine Nord
Thanh Le
Siu Wan

**National Center for Education
Statistics**

Jerry West

This page intentionally left blank.

TABLE OF CONTENTS

<u>Section</u>		<u>Page</u>
1	Introduction.....	1
2	Individuals Included on the K–3 Longitudinal File.....	2
3	Content.....	4
4	K–3 Longitudinal Weights.....	6
5	Characteristics of Longitudinal Weights	10
6	Data File Structure	11
7	Electronic Code Book.....	13
8	K–3 Longitudinal Data File Limiting Fields	14

List of Tables

<u>Table</u>		<u>Page</u>
1	Characteristics of the child-level K-3 longitudinal weights: School years 1998–99, 1999–2000, and 2001–02.....	11

List of Exhibits

<u>Exhibit</u>		<u>Page</u>
1	Contents of the base year, first grade, and third grade user’s manuals: School years 1998–99, 1999–2000, and 2001–02	3
2	ECLS-K: K–3 longitudinal weights: School years 1998–99, 1999–2000, and 2001–02.....	7
3	Use of the K–3 Longitudinal weights: School years 1998–99, 1999–2000, and 2001–02.....	8
4	Organization of the K–3 longitudinal data file: School years 1998–99, 1999–2000, and 2001–02.....	12

USER'S GUIDE TO THE LONGITUDINAL KINDERGARTEN–THIRD GRADE PUBLIC-USE DATA FILE

1. Introduction

For the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), the longitudinal kindergarten–third grade public-use data file, referred to hereafter as the K–3 longitudinal data file, combines data from the base (kindergarten), first grade, and third grade years. It was created so that analysts can easily examine children's growth and development between kindergarten and third grade without having to go through the process of merging several different data files. Special K–3 longitudinal weights have been developed and are included on the file. All children included on the longitudinal K–1 data file released in 2002 are also included on this file to allow users to conduct K–1 as well as K–3 analyses. Thus this file can be used to study such topics as children's learning across school years, the extent of summer learning or loss between kindergarten and the fall of the following school year, and the school or classroom characteristics that are related to growth in reading and math skills between first and third grade.

There is no separate user's manual for the K–3 longitudinal data file. To use this file, users should refer to this user's guide and to the base year, first grade, and third grade user's manuals (*Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) Base Year Public-Use Data Files and Electronic Code Book: User's Manual* (NCES 2001–029: U.S. Department of Education, National Center for Education Statistics, 2001); *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) User's Manual for the ECLS-K First Grade Public-Use Data Files and Electronic Code Book* (NCES 2002–135: U.S. Department of Education, National Center for Education Statistics, 2002); and *Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K) User's Manual for the ECLS-K Third Grade Public-Use Data File and Electronic Code Book* (NCES 2004–001: U.S. Department of Education, National Center for Education Statistics, 2004)). These manuals are found in appendix C on the CD-ROM.

Users will obtain basic information about the K–3 longitudinal data file within this guide. It begins with a description of the individuals included on the file. It then provides an overview of the content of the data file, the K–3 longitudinal weights, and the data file structure. The last section provides an overview of the electronic code book (ECB). For information on the background of the ECLS-K, the data collection instruments, and the assessment and rating scale scores used in the ECLS-K, users should

refer to the base year, first grade, and third grade user's manuals. These manuals also describe the sample design and implementation and provide detailed information about the data file content and composite variables. For detailed information about the ECB and how to use it, users should refer to the online help found in the root directory of the Longitudinal Kindergarten–Third Grade Public-Use ECB. The same information can also be found in chapter 8 of the base year, first grade, or third grade user's manuals. Exhibit 1 provides an overview of the content of these manuals (they all follow the same format). This exhibit will help users locate specific information about the ECLS-K.

2. Individuals Included on the K–3 Longitudinal File

The K–3 longitudinal file contains all children included on the K–1 longitudinal file released in 2002 as well as most of the children for whom either a parent interview or a child assessment was completed in third grade. To be included on the original K–1 longitudinal file, a parent interview or child assessment must have been conducted in the spring of kindergarten (spring 1999) and in at least one point in time during the first grade year (fall 1999 or spring 2000). Thus, children were included if there was either a parent or child assessment conducted during the spring of kindergarten **and** a parent interview or child assessment was completed in any of the following data collections:

- Fall-first grade *or*
- Fall-kindergarten and fall-first grade *or*
- Spring-first grade *or*
- Fall-kindergarten and spring-first grade *or*
- Fall-kindergarten and fall-first grade and spring-first grade.

Not all children on the original K–1 longitudinal file completed a parent interview or a child assessment in the third grade.¹

¹ Of children on the K–1 longitudinal file, 2,402 (14.0 percent) had neither a parent interview nor a child assessment completed in third grade; 792 (4.6 percent) had a child assessment completed in third grade, but no parent interview; and 1,693 (9.8 percent) had a parent interview completed in third grade, but no child assessment. All of these children are included on the K–3 longitudinal file so that analysts can conduct longitudinal analyses involving just the kindergarten and first grade data.

Exhibit 1. Contents of the base year, first grade, and third grade user’s manuals: School years 1998–99, 1999–2000, and 2001–02

Chapter	Content
Chapter 1	Introduction Provides information about the background of the ECLS-K, the conceptual model underlying the study, the study components, and the ECLS-K data files.
Chapter 2	Description of Data Collection Instruments Describes the different data collection instruments, including the direct child assessments, the parent interview, the teacher questionnaires, and the school questionnaires.
Chapter 3	Assessment and Rating Scale Scores Used in the ECLS-K Provides detailed information on the assessment and rating scale scores developed for the ECLS-K and how to use them.
Chapter 4	Sample Design and Implementation Provides information on sampling procedures, calculation and use of sample weights, variance estimation, and design effects. Note: Information specific to the creation of cross-round (i.e., any combination of kindergarten-first grade-third grade) weights included in the K–3 longitudinal file is <u>not</u> presented in this chapter. Details about the cross-round weights are described in chapter 9 of the first grade and third grade user’s manuals.
Chapter 5	Data Collection Methods and Response Rates Provides information about the data collection methods, field staff training, response rates, and data collection quality control.
Chapter 6	Data Preparation Provides information on data preparation, including coding and editing specifications for computer-assisted interviewing and hard-copy questionnaires.
Chapter 7	Data File Content and Composite Variables Provides detailed information on the identification variables, missing values, and composite variables that have been created. Note: Section 7.7 in the first grade user’s manual contains details on how to merge base year school-level data with the first grade child-level data. This is not relevant to the K–3 longitudinal data file. These data have already been merged for users of the K–3 longitudinal data file.
Chapter 8	Electronic Code Book Provides information on the hardware and software requirements of the electronic code book (ECB), its features, instructions for installing, starting, and exiting the ECB, working with taglists (i.e., lists of commonly used variables), and extracting data from the ECB to create an analysis data file.
Chapter 9	Creating a Longitudinal File This chapter was included in the first and third grade public-use and restricted-use cross-sectional user’s manuals for users who wanted to create their own longitudinal file. It provides information on the types of research questions that can be answered with a longitudinal file. Chapter 9 in the third grade manuals provides detailed information about the development and use of the K–3 longitudinal (i.e., cross-round) weights. Note: The instructions on merging the files in section 9.3 are not relevant to the K–3 longitudinal data file as the data have already been merged.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 base year, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

In addition to the children from the K–1 longitudinal file, the K–3 longitudinal file also includes children if there were either a parent or child assessment conducted during the spring of third grade **and** a parent interview or child assessment was completed in any of the following data collections:

- Spring-first grade *or*
- Spring-kindergarten and spring-first grade *or*
- Fall-kindergarten, spring-kindergarten, and spring-first grade *or*
- Fall-kindergarten and spring-first grade *or*
- Fall-kindergarten, spring-kindergarten, fall-first grade, and spring-first grade.

There are 306 children with either a completed parent interview or a completed child assessment in third grade who are not included on the K–3 longitudinal file. These children were interviewed only in fall-first grade and spring-third grade or were missing both fall- and spring-first grade data.

Analyses using combinations of K–3 data other than those listed above can be conducted. Even though customized weights were not created for the other K–3 data combinations, existing weights can be used for analyses of these combinations. See the section on K–3 longitudinal weights for details.

The K–3 longitudinal data file is a child-level file. All parent, teacher, and school information collected for any particular child from each round of data collection has been attached to that child’s record (a more detailed description of the record layout follows). In all, the K–3 longitudinal data file has 17,401 child records. For detailed information about response rates in each round of data collection, see chapter 5 of the base year, first grade, and third grade user’s manuals.

3. Content

With a few exceptions, the K–3 longitudinal data file contains all the data that were collected from parents, children, teachers, or schools in the base year (fall and spring), first grade (fall and spring), and spring-third grade data collections. To streamline the file, however, the data from the household rosters that listed all household members, their relationship to the sampled child, and selected other characteristics are not included on the file. The composite variables describing critical household roster-based information, such as the children’s family structure and selected characteristics of the family

members, have been retained on the file (see chapter 7 of the base year, first grade, and third grade public-use user's manuals for a description of these and other composite variables).

In addition, cross-sectional weights (associated with a single wave of data collection) and within-grade longitudinal weights (for within-kindergarten or within-first grade longitudinal analysis) are not included on the K–3 longitudinal data file, which contains just the K–3 longitudinal weights (for analysis of kindergarten, first grade, and third grade data). There are no cross-sectional or within-grade longitudinal weights included on the file because the K–3 longitudinal data file should not be used to examine just the kindergarten data, just the first grade data, or just the third grade data. The reason it should not be used in this way is that not all children interviewed in kindergarten were interviewed in first grade. Similarly, not all children interviewed in first grade were interviewed in third grade. And, as noted previously, not all children interviewed in third grade are included on the K–3 longitudinal file. Thus, the population of base year respondents contained in the K–3 longitudinal data file is a subset of those who were interviewed during the base year. Similarly, the population of first grade and third grade children on the K–3 longitudinal data file is a subset of those who were interviewed during the first and third grade data collections.

Similar to the first grade files, the K–3 longitudinal data file contains a few base year variables that were not in the base year files. They fall into three categories: (1) base year recalibrated assessment scores, (2) base year recalibrated Academic Rating Scale (ARS) scores and (3) new and corrected base year composites. The direct child assessment scores were recalibrated to obtain gain scores that could be compared across five waves of data. The ARS scores were recalibrated because an error was identified in the base year ARS scores. Specifically, the fall and spring base year ARS scores used slightly different metrics. These scores were recalibrated using a combined calibration of fall- and spring-kindergarten ratings. Therefore, the unit for the corrected fall- and spring-kindergarten scores is the same, though comparisons between fall- and spring-kindergarten scores are not recommended. Although the item stems are similar across grades, the actual items include performance criteria that increase from one grade to the next. Moreover, the ARS score metric is different at each point. Therefore, change scores should *not* be used to compare third grade ratings with those from earlier rounds.

The specifics of the ARS and composite problems are described in the first grade public-use user's manual in the section titled Base Year Errata and Composites. The other errors listed in that section have either been corrected (errata number 1 through 7) or are not pertinent to the K–3 longitudinal data file (erratum number 8). For example, the base year poverty and urbanicity composites were detected to have errors and were recreated and included with the first grade data file (appendix D) and in the K–3

longitudinal data file. Specifically, WKPOV_R replaces WKPOVRTY and KURBAN_R replaces KURBAN. Similarly, the imputation flag IF_INC_R replaces IF_INC. Errata numbers 3, 6, and 7 were corrected but did not require replacing existing variables.

4. K–3 Longitudinal Weights

There are several sets of K–3 longitudinal weights computed for children with complete data from different combinations of rounds. All K–3 longitudinal weights are child-level weights. There are no K–3 longitudinal weights at the school or teacher level since school- and teacher-level weights were not computed for the first grade year nor for the third grade year due to lack of representativeness.

The K–3 longitudinal weights, available on the K–3 longitudinal data file ECB are described in exhibit 2: the use of the K–3 longitudinal weights is described in exhibit 3. Exhibit 3 is designed to help users choose appropriate weights for their analyses.

First, decide which two or more points in time are the focus of the analysis. The analysis could pertain to two points in time (e.g., spring-kindergarten and fall-first grade, or spring-kindergarten and spring-first grade, or spring-first grade and spring-third grade), three points in time (e.g., spring-kindergarten, spring-first grade, and spring-third grade), four points in time (any four of fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, and spring-third grade), or five points in time (all five rounds of data). For example, if the analysis uses spring-kindergarten and fall-first grade data, then the appropriate weight would be one that begins with C23 (denoting child-level data from round 2 AND round 3). If the analysis uses data from spring-kindergarten, spring-first, and spring-third grade, then the appropriate weight begins with C245 (denoting data from rounds 2, 4, AND 5).

Second, consider the source of the data, which also affects the choice of the weight. In exhibit 3, details under “to be used in the analysis of ...” column provide guidance based on whether the data were collected through the child assessments, parent interviews, or teacher questionnaires A or B. If parent data from spring-kindergarten and fall-first grade are needed for the analysis, then C23PW0 should be used, otherwise C23CW0 can be used. Similarly, if an analyst wants to examine the influence of parent characteristics on gains in assessment scores between kindergarten and third grade, the appropriate weight would be C245PW0 indicating that parent interview data was included. However, if only child or teacher data were used in the analyses, then the appropriate weight to use is C245CW0.

Exhibit 2. ECLS-K: K–3 longitudinal weights: School years 1998–99, 1999–2000, and 2001–02

K–3 longitudinal (panel) weight	is nonzero if ...
C23CW0	assessment data are present for both spring-kindergarten and fall-first grade, or if the child was excluded from direct assessment in both of these rounds of data collection due to a disability.
C23PW0	parent interview data are present for both spring-kindergarten and fall-first grade.
C123CW0	assessment data are present for fall- and spring-kindergarten and fall-first grade, or if the child was excluded from direct assessment in all three of these rounds of data collection due to a disability.
C123PW0	parent interview data are present for fall- and spring-kindergarten and fall-first grade.
C24CW0	assessment data are present for both spring-kindergarten and spring-first grade, or if the child was excluded from direct assessment in both of these rounds of data collection due to a disability.
C24PW0	parent interview data are present for both spring-kindergarten and spring-first grade.
C124CW0	assessment data are present for fall-kindergarten and spring-kindergarten and spring-first grade, or if the child was excluded from direct assessment in all three of these rounds of data collection due to a disability.
C124PW0	parent interview data are present for fall-kindergarten and spring-kindergarten and spring-first grade.
C1_4CW0	assessment data are present for four rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, and spring-first grade), or if the child was excluded from direct assessment in all of these four rounds of data collection due to a disability.
C1_4PW0	parent interview data are present for four rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, and spring-first grade).
Y2COMW0	assessment data are present for fall-kindergarten and spring-kindergarten and spring-first grade, or if the child was excluded from direct assessment in all three of these rounds of data collection and parent and/or teacher data are present for one or more base year rounds and parent and/or teacher data are present for spring-first grade.
C45CW0	assessment data are present for both spring-first grade and spring-third grade, or if the child was excluded from direct assessment in both of these rounds of data collection due to a disability.
C45PW0	parent interview data are present for both spring-first grade and spring-third grade.
C245CW0	assessment data are present for spring-kindergarten and spring-first grade and spring-third grade, or if the child was excluded from direct assessment in all of these three rounds of data collection due to a disability.
C245PW0	parent interview data are present for spring-kindergarten and spring-first grade and spring-third grade.
C1_5FC0	assessment data are present for four rounds of data collection involving the full sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, and spring-third grade), or if the child was excluded from direct assessment in all four of these rounds of data collection due to a disability.
C1_5FP0	parent interview data are present for four rounds of data collection involving the full sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, and spring-third grade).
C1_5SC0	assessment data are present for all five rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, and spring-third grade), or if the child was excluded from direct assessment in all five rounds of data collection due to a disability.
C1_5SP0	parent interview data are present for all five rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, and spring-third grade).

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 base year, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

Exhibit 3. Use of the K–3 longitudinal weights: School years 1998–99, 1999–2000, and 2001–02

K–3 longitudinal (panel) weights	to be used for analysis of ...
C23CW0	child direct assessment data from BOTH spring-kindergarten and fall-first grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race/ethnicity).
C23PW0	parent interview data from BOTH spring-kindergarten and fall-first grade.
C123CW0	child direct assessment data from fall- AND spring-kindergarten AND fall-first grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race/ethnicity).
C123PW0	parent interview data from fall- AND spring-kindergarten AND fall-first grade.
C24CW0	child direct assessment data from BOTH spring-kindergarten and spring-first grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race/ethnicity).
C24PW0	parent interview data from BOTH spring-kindergarten and spring-first grade.
C124CW0	child direct assessment data from fall-kindergarten AND spring-kindergarten AND spring-first grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race/ethnicity).
C124PW0	parent interview data from fall-kindergarten AND spring-kindergarten AND spring-first grade.
C1_4CW0	child direct assessment data from all four rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, and spring-first grade), alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C1_4PW0	parent interview data from all four rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, and spring-first grade).
Y2COMW0	child direct assessment data from fall-kindergarten AND spring-kindergarten AND spring-first grade, in conjunction with parent and/or teacher data from spring-first grade, AND one or more base year rounds of parent and/or teacher data.
C45CW0	child direct assessment data from BOTH spring-first grade and spring-third grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race/ethnicity).
C45PW0	parent interview data from BOTH spring-first grade and spring-third grade.
C245CW0	child direct assessment data from spring-kindergarten AND spring-first grade AND spring-third grade, alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race/ethnicity).
C245PW0	parent interview data from spring-kindergarten AND spring-first grade AND spring-third grade.
C1_5FC0	child direct assessment data from FOUR rounds of data collections involving the full sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade), alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, race/ethnicity).
C1_5FP0	parent interview data from FOUR rounds of data collections involving the FULL sample of children (fall-kindergarten, spring-kindergarten, spring-first grade, spring-third grade).
C1_5SC0	child direct assessment data from ALL FIVE rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, and spring-third grade), alone or in conjunction with any combination of a limited set of child characteristics (e.g., age, sex, and race/ethnicity).
C1_5SP0	parent interview data from ALL FIVE rounds of data collection (fall-kindergarten, spring-kindergarten, fall-first grade, spring-first grade, and spring-third grade).

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 base year, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

K–3 longitudinal weights are used to produce estimates of differences between two or more rounds of data collection spanning across kindergarten, first grade, and third grade. Simple examples involving two rounds of data collection are as follows: (1) estimating the differences in children’s mean assessment scores between spring–first grade and spring–third grade using C45CW0 and (2) estimating the difference in Social Rating Scale scores as reported by parents in spring–kindergarten and spring–first grade using C24PW0 (Social Rating Scale scores as reported by teachers and parents are not available for fall–first grade). K–3 longitudinal weights are also used to study the characteristics of children who were assessed in two or more rounds of data collection. For example, one can study the characteristics of kindergarten children that are associated with the greatest gains in learning in first and third grades. If the analysis includes data collected from the parents in rounds 4 and 5, then C45PW0 can be used in the analysis. However, if the analysis involves just the key characteristics (e.g., race) available for most children and the child assessment data from rounds 4 and 5, then C45CW0 can be used to estimate changes in assessment scores between spring–first grade and spring–third grade. An example where data from more than two rounds are used is as follows: examining whether gains children make in their reading knowledge and skills during the kindergarten year and from the end of kindergarten to the end of first grade are related to parents’ and teachers’ beliefs about kindergarten readiness and parental educational expectations. In this case, the weight Y2COMW0 would be appropriate. As noted in the first grade and third grade user’s manuals, any longitudinal analysis that uses data from fall–first grade will be limited to a 27 percent subsample of children.²

There may be combinations of data for which no weights were developed. For example, there is no specific weight to study changes in children’s classroom environments as they move from kindergarten to third grade if child assessment or parent data are not used in the analysis. In this example, the data come from part A of the teacher’s questionnaire (TQA). The preferred weight for this analysis would be C245CW0, which is the weight for child direct assessment data from spring–kindergarten, spring–first grade, and spring–third grade. Of children on the longitudinal K–3 file with TQA data in all three of these rounds (e.g., spring–kindergarten, spring–first grade, and spring–third grade), 99 percent (10,016) have nonzero C245CW0, compared with 90 percent (9,135) with nonzero C1_5FC0 and 27 percent (2,489) with nonzero C1_5SC0, the other two longitudinal weights available for analyses of child data. The preferred weight is the one that will yield the largest number of records for analysis, which is C245CW0 in this case. Analytically, it can be argued that since the direct assessments are conducted in schools, this weight comes closest to capturing the children in participating schools and thus to capturing

² As described in the first grade user’s manual, fall–first grade was a design enhancement to enable researchers to study the extent of summer learning losses and gains and the factors associated with them. The fall data collection was limited to students in a 30 percent subsample of schools.

the children with relevant school environment data. Similarly, if data from the school administrator's questionnaire are used in the analysis of the K–3 longitudinal data, then the same arguments can be used to select the weight. In this case, 54 percent of children have school administrator questionnaire data from all kindergarten, first grade and third grade; of these, 99 percent have nonzero C245CW0 compared with 90 percent with nonzero C1_5FC0 and 27 percent with nonzero C1_5SC0. Therefore, the preferred weight is also C245CW0. For further advice on which weights to use when analyzing a complex combination of data, contact NCES at *ECLS@ed.gov*.

5. Characteristics of Longitudinal Weights

The statistical characteristics of the longitudinal weights are presented in table 1. For each weight, the number of cases with nonzero values is presented together with the mean weight, the standard deviation, the coefficient of variation (i.e., the standard deviation as a percentage of the mean weight), the minimum value of the weight, the maximum value of the weight, the skewness, the kurtosis, and the sum of weights.

The difference in the estimate of the population of students (sum of weights) between the different panels of students and types of weights is due to a combination of factors, among them: (1) the number of base year respondents who became ineligible (due to death, leaving the country, or being a nonsampled mover) after the base year; (2) the adjustment of the weights for the children of unknown eligibility; and (3) the difference in the number of records used to construct sample-based control totals. Of the eight longitudinal weights computed in third grade, only the first two (C45CW0 and C45PW0) involve children sampled in first grade. For these two weights, the child records included in the file used for computing the control totals are records of base year respondents and records of eligible children sampled in first grade. For all other longitudinal weights, records of children sampled in first grade were not included in the file, causing the sum of weights to be smaller.

Table 1. Characteristics of child-level K–3 longitudinal weights: School years 1998–99, 1999–2000, and 2001–02

Variable name	Number of cases	Mean	Standard deviation	CV ¹ (× 100)	Minimum	Maximum	Skewness	Kurtosis	Sum
C23CW0	5,216	739.84	587.55	79.42	68.23	7,182.37	3.98	21.56	3,858,997
C23PW0	4,861	793.83	515.75	64.97	84.26	5,853.21	2.97	13.04	3,858,805
C123CW0	4,729	815.99	646.25	79.20	76.08	7,696.79	3.89	21.55	3,858,824
C123PW0	4,295	898.37	597.89	66.55	95.35	6,421.30	3.05	14.20	3,858,492
C24CW0	16,371	234.81	200.69	85.47	1.78	3,272.40	4.22	31.65	3,844,009
C24PW0	14,938	257.25	198.94	77.34	1.93	2,580.41	3.30	19.64	3,842,784
C124CW0	15,001	256.28	228.52	89.17	1.54	3,877.43	3.71	24.60	3,844,472
C124PW0	13,413	286.40	214.80	75.00	2.06	3,275.79	3.84	26.53	3,841,463
C1_4CW0	4,542	847.78	639.83	75.47	77.56	7,528.68	3.49	18.68	3,850,619
C1_4PW0	4,012	959.07	617.93	64.43	108.75	6,780.92	2.86	13.48	3,847,785
Y2COMW0	13,983	274.83	241.55	87.89	2.03	3,803.82	4.26	29.97	3,842,961
C45CW0	13,964	281.86	273.52	97.04	1.68	3,897.42	3.37	19.90	3,935,960
C45PW0	12,652	310.98	266.89	85.82	1.68	3,718.34	3.11	17.32	3,934,550
C245CW0	13,694	280.68	277.47	98.86	1.65	4,119.55	3.55	22.53	3,843,642
C245PW0	12,204	314.92	267.05	84.80	1.78	3,121.66	2.87	14.51	3,843,272
C1_5FC0	12,558	306.07	303.52	99.17	1.68	4,264.25	3.59	22.83	3,843,607
C1_5FP0	10,998	349.42	299.17	85.62	1.92	3,754.91	3.18	17.88	3,842,954
C1_5SC0	4,032	952.67	875.12	91.86	64.97	7,174.65	3.28	13.78	3,841,183
C1_5SP0	3,522	1,090.37	816.79	74.91	104.68	6,801.61	2.56	9.19	3,840,278

¹ Coefficient of variation.

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 base year, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

For information about the development of the longitudinal weights, see chapter 9 of the first grade and third grade user’s manuals.

6. Data File Structure

The K–3 longitudinal data file is arranged in blocks of related variables. Exhibit 4 provides an overview of the order and content of these blocks. A more detailed listing of the block arrangements with a description of the topical variable grouping and associated field IDs can be found in the online help of the ECB.

Exhibit 4. Organization of the K–3 longitudinal data file: School years 1998–99, 1999–2000, and 2001–02

Block	Description of types of variables within the block
ID	Identification variables for child, parent, teachers, and schools (e.g., child ID, teacher ID)
Field Management System	Variables taken from the field management system (e.g., whether child received special education services)
Demographic	Demographic variables used in drawing the sample (e.g., the type of school, census region, round the child joined the study, the child’s age, and the child’s race).
Change flags	Indicators of whether the student changed teachers or schools between the different rounds
Full sample weights	The 19 longitudinal weights
Status flags	Flags indicating the presence or absence of a component for each round of data collection (e.g., presence of fall-kindergarten direct child assessment, parent interview, teacher questionnaire A, B, or C, school administrator questionnaire, school facilities checklist, or student record abstract).
Edits and error flags	Indicators of whether the household roster in the parent interview was edited or contained errors
Child assessment scores	Re-calibrated child assessment scores and Social Rating Scale scores
Composites	Composite variables constructed for users (e.g., socio-economic status, family structure)
Data from the instruments	Variables from all the instruments. The data are ordered as follows: direct child assessment data, parent interview data, teacher questionnaire data, school administrator data, and school facilities checklist data. Each segment begins with round 1 data (fall-kindergarten) and ends with round 5 data (spring-third grade).
Taylor series identifiers	Taylor series primary sampling unit and strata identifiers for the K–3 longitudinal weights
Replicate weights	Replicate weights (40 replicates for weights based on fall first grade data, 90 replicates in all other cases)

SOURCE: U.S. Department of Education, National Center for Education Statistics, Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 base year, first grade, and third grade data collections, school years 1998–99, 1999–2000, and 2001–02.

7. **Electronic Code Book**

The ECLS-K K–3 longitudinal data file CD-ROM contains an ECB that allows users to easily examine the variables in the K–3 longitudinal data file.

The ECLS-K–3 longitudinal data file ECB allows a user to perform the following:

- Search the names and labels of variables in the database (called the catalog) to select variables for analysis (see section 8.3, Variable List from the base year, first grade, or third grade user’s manuals).
- Examine the question wording, response categories, and response frequencies for variables the user selects (see section 8.4.9, Viewing Code Book and Variable Information from the base year, first grade, or third grade user’s manuals).
- Create a list of variables (called “a taglist”) to be extracted from the catalog, save the list for later use, print the list as a code book, or use a predefined list on the ECB (see section 8.4, Working Taglist, from the base year, first grade, or third grade user’s manuals).
- Automatically generate SAS, SPSS for Windows, or Stata programs to extract selected variables from the whole data set or for a subset of the cases that are defined by the user (see section 8.5, Extracting Data from the ECB from the base year, first grade or third grade user’s manuals).

The data user can create SAS, SPSS for Windows, and Stata programs that will generate an extract data file from the data file on the CD-ROM. The ECLS-K K–3 longitudinal data file ECB does not create a SAS, SPSS for Windows, or Stata data file. It will prepare the statements that can be used with the user’s own SAS, SPSS for Windows, or Stata software to create a file. As noted earlier, the CD-ROM contains both ASCII data sets that the ECB uses to extract specific subdata files. The CD-ROM must be in the drive for the data to be extracted.

For detailed information about using the ECB, see chapter 8 in the online help. Users can also refer to chapter 8 of the base year, first grade, or third grade user’s manuals for information about using ECBs. Chapters 8 in the user’s manuals use examples from the base year, the first grade or the third grade data files. The functionality of the ECB is essentially the same across data sets; only the examples used to illustrate the different ECB features change.

8. K–3 Longitudinal Data File Limiting Fields

The limiting fields for the K–3 longitudinal data file include the following: (1) child sampled in round 3 (R3SAMPLE); (2) child’s school type changed between rounds 2 and 4 (R4R2SCHG); (3) child’s school type changed between rounds 4 and 5 (R5R4SCHG); (4) first-time kindergartner status (P1FIRKDG); and (5) grade level of the child in the spring- third grade data collection (T5GLVL). These limiting fields allow cases belonging to different categories within each variable to be included or excluded from the extraction. The selection indicator will be either a “Yes” or a “No” to specify whether each value (category) for the variable should be included or excluded, respectively. The default setting for every category (value) for the limiting variables is “Yes,” meaning that all records will be present in the extract file. To exclude all records that have data for that particular category, the user can change the “Yes” to “No” by double-clicking on it. For example, the user can select “No” for grade levels other than first grade and “Yes” for first grade to create a data file that only includes children in first grade. At least one of the codes must be selected as “Yes” or no records will be extracted for analysis. See chapter 8, section 8.6 of the first grade or third grade user’s manuals or section 8.6 in chapter 8 in the online help for further details on limiting fields.