

## Analysis of Data from the ECLS-B Kindergarten Rounds of Data Collection Practice Example

This document outlines the basic steps in an analysis intended to answer the following research question:

Is participation in early care and education during the year prior to kindergarten entry related to children's early reading and math scores at kindergarten entry?

SAS, SPSS, and Stata syntax for the steps needed to prepare the data for analysis, including the creation of "kindergarten round" variables that use information from the round in which children were first in school, is provided at the end of this document).

### Step 1- Create your data file from the Child Catalog on the Electronic Codebook (ECB).

Variables needed in analysis:

Kindergarten round indicator variables

XKWHENK XK FIRST YEAR OF KINDERGARTEN  
XKKDATA XK WAVE FIRST IN KINDERGARTEN

Children's early reading and math scale scores (kindergarten 2006 and kindergarten 2007)

X4RSCR2 X4 READING SCALE SCORE CLB P-K07  
X5RSCR2 X5 READING SCALE SCORE CLB P-K07  
X4MSCR2 X4 MATH SCALE SCORE CLB P-K07  
X5MSCR2 X5 MATH SCALE SCORE CLB P-K07

Early care and education participation variable for the year prior to kindergarten entry (preschool and kindergarten 2006)

Y3PRIMNW X3 PRIM CARE ARRNGMNT WHERE MOST HRS/WK (REVISED)  
*(note that this is a revised composite variable so it begins with the prefix Y3)*  
X4PRIMNW X4 PRIM CARE ARRNGMNT WHERE MOST HRS/WK

Weights

WKR0 WK RSP1/RSP2/RSP3/(RSP4/RSP5)-FULL SMP WGT  
WKR1 - WKR90 (corresponding replicate weights)

ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry  
> Slide 39 of 44

**Step 2- As appropriate, set missing value codes to missing.**

- Run descriptive statistics (frequency distributions for categorical variables of interest and means, minimums, and maximums for continuous variables) to determine which variables have missing data values (-7, -8, -9, and -1). Review hard copy instruments to see if there are any skip patterns and if any -1s need to be recoded to something other than missing (e.g., -1 to “no”).
- Note that you do not need to worry about the missing value codes in the kindergarten status indicator variables (XKWHENK and XKKDATA).
- Set missing value codes to missing.

**Step 3- As appropriate, collapse categories of categorical variables.**

- For this example, we want to combine the categories of the primary care arrangement variable into six categories.
  - No nonparental care
  - Relative care
  - Nonrelative care
  - Center care (not Head Start)
  - Head Start
  - Multiple arrangements for same number of hours
- Remember that we need to do this for both the preschool and kindergarten 2006 versions of the primary care arrangement variable.
- Create new value labels (formats) for your new variable if you choose to.

**Step 4- Create new “kindergarten round” variables.**

- Use XKKDATA to determine from which round the data should come.
- Note that XKKDATA is valid for all children who were enrolled in school or homeschooled in either 2006 or 2007 (including children who went straight to first grade or were in ungraded or multi-graded classes). If you want to limit your analysis, for example, to children who were enrolled only in traditional kindergarten in 2006 or 2007, you will need to use XKWHENK rather than XKKDATA.
- Cases are missing on XKKDATA if they were not yet enrolled in school by the time of the 2007 collection or if they were not yet enrolled in school in 2006 and were nonrespondents at 2007. These nonrespondents are accounted for in the weight.
- Note that if cases have missing data on the variable of interest (e.g. X\*RSCR2) for the round from which you are pulling the data, then the new variable (e.g.

## ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry > Slide 39 of 44

XKRSCR2) will have missing data for those cases, regardless of whether they have data for the other round.

### Step 5—Review data (e.g., run unweighted estimates).

- Run frequencies of the variables of interest. This will help determine whether there are enough cases in each category to support your analysis.

### Step 6—Produce weighted estimates.

- Weight data by WKRO.
- This is the appropriate weight because it is intended for analysis of parent interview data at all rounds and child assessment data at kindergarten entry.
- Limit your analysis to children who were enrolled only in traditional kindergarten in 2006 or 2007 using XKWHENK if this is important for your research question.
- Run estimates using a “weight” statement (in SPSS, you also could apply the sampling weight WKRO to the file).
- Run weighted means of XKRSCR2 and XKMSCR2 overall and for the different categories of XKPRIMNW.

### Step 7—Calculate standard errors for the complex sample design to do hypothesis testing

- Use software packages such as SAS, SPSS, Stata, R, WesVar, SUDAAN, or AM to determine appropriate standard errors to use for significance testing.
- Use your standard errors to compute t-tests. The formula used to calculate the t statistic is as follows:

$$t = \frac{\text{Estimate 2} - \text{Estimate 1}}{\sqrt{(\text{Standard Error 2})^2 + (\text{Standard Error 1})^2}}$$

ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry  
> Slide 39 of 44

**Example of complete SAS syntax for analysis of early reading and math scores at kindergarten entry by participation in early care and education during the year prior to kindergarten entry**

```
Data practice;
set orig_data_set;

/*Recode missing values (-9) to system missing*/

Array original(6) X4RSCR2 X5RSCR2 X4MSCR2 X5MSCR2 Y3PRIMNW X4PRIMNW;
Array newvar(6) X4RSCR2_r X5RSCR2_r X4MSCR2_r X5MSCR2_r Y3PRIMNW_r X4PRIMNW_r;

Do i = 1 to 6;
    If original(i) < 0 then newvar(i) = .;
    Else newvar(i) = original(i);
End;

/*Collapse categories of X/Y*PRIMNW into 6 categories*/

If Y3PRIMNW_r = 0 then Y3PRIMNW_r = 0; /*No primary care*/
Else if Y3PRIMNW_r in (1,2,3) then Y3PRIMNW_r = 1; /*Relative care*/
Else if Y3PRIMNW_r in (4,5,6) then Y3PRIMNW_r = 2; /*Non-relative care*/
Else if Y3PRIMNW_r = 7 then Y3PRIMNW_r = 3; /*Center (not Head Start) care*/
Else if Y3PRIMNW_r = 8 then Y3PRIMNW_r = 4; /*Head Start*/
Else if Y3PRIMNW_r = 9 then Y3PRIMNW_r = 5; /*Multiple care arrangements*/

If X4PRIMNW_r = 0 then X4PRIMNW_r = 0; /*No primary care*/
Else if X4PRIMNW_r in (1,2,3) then X4PRIMNW_r = 1; /*Relative care*/
Else if X4PRIMNW_r in (4,5,6) then X4PRIMNW_r = 2; /*Non-relative care*/
Else if X4PRIMNW_r = 7 then X4PRIMNW_r = 3; /*Center (not Head Start) care*/
Else if X4PRIMNW_r = 8 then X4PRIMNW_r = 4; /*Head Start*/
Else if X4PRIMNW_r = 9 then X4PRIMNW_r = 5; /*Multiple care arrangements*/

/*Create "XK" version of reading, math, and primary care variables for analysis of children at kindergarten
entry*/

/*Reading scale score*/
XKRSCR2 = .;
If XKKDATA = 1 then XKRSCR2 = X4RSCR2_r;
Else if XKKDATA = 2 then XKRSCR2 = X5RSCR2_r;

/*Math scale score*/
XKMSCR2 = .;
If XKKDATA = 1 then XKMSCR2 = X4MSCR2_r;
Else if XKKDATA = 2 then XKMSCR2 = X5MSCR2_r;

/*Primary care variable (year prior to kindergarten entry)*/
XKPRIMNW = .;
If XKKDATA = 1 then XKPRIMNW = Y3PRIMNW_r2;
Else if XKKDATA = 2 then XKPRIMNW = X4PRIMNW_r2;
Label XKRSCR2 = "Reading scale score at first kindergarten entry";
Label XKMSCR2 = "Math scale score at first kindergarten entry";
Label XKPRIMNW = "Primary care arrangement, year prior to kind.";

/*Create format (value labels) for new primary care variable*/
```

**ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry  
> Slide 39 of 44**

```
proc format;
value primcr6c
    0 = "0: No regular care"
    1 = "1: Relative care"
    2 = "2: Nonrelative care"
    3 = "3: Center care (not HS)"
    4 = "4: Head Start"
    5 = "5: Multiple care arrangements";
```

```
/*Run unweighted frequencies of primary care variable to check cell counts (limit analysis to cases
enrolled only in traditional kindergarten programs)*/
```

```
Proc freq data=practice;
tables XKPRIMNW/missprint;
where XKWHENK in (2,3);
format XKPRIMNW primcr5c.;
run;
```

```
/*Run weighted means of reading and math scale scores*/
```

```
Proc means data=practice n mean min max;
var XKRSCR2 XKMSCR2;
weight WKR0;
where XKWHENK in (2,3);
run;
```

```
/*Run weighted means of reading and math scale scores by categories of primary care*/
```

```
Proc means data=practice n mean min max;
var XKRSCR2 XKMSCR2;
class XKPRIMNW;
weight WKR0;
where XKWHENK in (2,3);
format XKPRIMNW primcr5c.;
run;
```

```
/*Use SAS survey procedures to obtain the correct standard errors (syntax not presented here)*/
```

ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry  
> Slide 39 of 44

**Example of complete SPSS syntax for analysis of early reading and math scores at kindergarten entry by participation in early care and education during the year prior to kindergarten entry**

\*Recode missing values (-9) to system missing\*.

```
RECODE X4RSCR2 X4MSCR2 X5RSCR2 X5MSCR2 X4PRIMNW Y3PRIMNW (SYSMIS=SYSMIS) (-9  
thru -1=SYSMIS)  
(ELSE=Copy) INTO X4RSCR2_r X4MSCR2_r X5RSCR2_r X5MSCR2_r X4PRIMNW_r Y3PRIMNW_r.  
EXECUTE.
```

\*Collapse categories of X/Y\*PRIMNW into 6 categories\*.

```
RECODE Y3PRIMNW_r X4PRIMNW_r (0=0) (7=3) (8=4) (9=5) (1 thru 3=1) (4 thru 6=2)  
(ELSE=SYSMIS).  
EXECUTE.
```

\*Create "XK" version of reading, math, and primary care variables for analysis of children at kindergarten entry\*

\*Reading scale score\*.

```
DO IF XKKDATA=1.  
compute XKRSCR2=X4RSCR2_r.  
ELSE IF XKKDATA=2.  
compute XKRSCR2=X5RSCR2_r.  
End if.  
EXECUTE.
```

\*Math scale score\*.

```
DO IF XKKDATA=1.  
compute XKMSCR2=X4MSCR2_r.  
ELSE IF XKKDATA=2.  
compute XKMSCR2=X5MSCR2_r.  
End if.  
EXECUTE.
```

\*Primary care arrangement (year prior to kindergarten entry)\*.

```
DO IF XKKDATA=1.  
compute XKPRIMNW=Y3PRIMNW_r.  
ELSE IF XKKDATA=2.  
compute XKPRIMNW=X4PRIMNW_r.  
End if.  
EXECUTE.
```

\*Assign variable labels and value labels to new variables\*.

```
VARIABLE LABELS  
XKRSCR2 'Reading scale score at kindergarten entry'  
XKMSCR2 'Math scale score at kindergarten entry'  
XKPRIMNW 'Primary care arrangement in year prior to K'.
```

**ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry  
> Slide 39 of 44**

## VALUE LABELS

/XKPRIMNW

0 'No regular care'  
1 'Relative'  
2 'Non-relative'  
3 'Center'  
4 'Head Start'  
5 'Multiple'

.

\*Run unweighted frequencies of primary care variable to check cell counts\*.

TEMPORARY.

SELECT IF (WKWHENK = 2 | WKWHENK = 3).

## FREQUENCIES

VARIABLES= XKPRIMNW  
/ORDER= ANALYSIS .

\*Run weighted means of reading and math scale scores\*.

WEIGHT BY WKR0.

TEMPORARY.

SELECT IF (WKWHENK = 2 | WKWHENK = 3).

MEANS TABLES=XKRSCR2 XKMSCR2

/CELLS MEAN COUNT.

\*Run weighted means of reading and math scale scores by categories of primary care\*.

TEMPORARY.

SELECT IF (WKWHENK = 2 | WKWHENK = 3).

MEANS TABLES=XKRSCR2 XKMSCR2 BY XKPRIMNW

/CELLS MEAN COUNT.

\*Use SPSS complex samples modules to obtain the correct standard errors (syntax not presented here)\*.

ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry  
> Slide 39 of 44

**Example of complete STATA syntax for analysis of early reading and math scores at kindergarten entry by participation in early care and education during the year prior to kindergarten entry**

\* This is a do-file for STATA programming. In order to use it correctly,  
\* You should type 'do "C:\ECLSB Seminar\Practice" nostop' in command line  
\* after you launch the STATA application.

```
clear
set memory 500m
set maxvar 10000
infile using "C:\ECLSB Seminar\Practice.dct"

/*Recode missing values (-9) to system missing*/
recode X4RSCR2 X5RSCR2 X4MSCR2 X5MSCR2 Y3PRIMNW X4PRIMNW (min/-1=.)

/*Collapse categories of X/Y*PRIMNW into 6 categories*/
recode Y3PRIMNW (0=0) /*No primary care*/
recode Y3PRIMNW (1/3=1) /*Relative care*/
recode Y3PRIMNW (4/6=2) /*Non-relative care*/
recode Y3PRIMNW (7=3) /*Center (not Head Start) care*/
recode Y3PRIMNW (8=4) /*Head Start*/
recode Y3PRIMNW (9=5) /*Multiple care arrangements*/

recode X4PRIMNW (0=0) /*No primary care*/
recode X4PRIMNW (1/3=1) /*Relative care*/
recode X4PRIMNW (4/6=2) /*Non-relative care*/
recode X4PRIMNW (7=3) /*Center (not Head Start) care*/
recode X4PRIMNW (8=4) /*Head Start*/
recode X4PRIMNW (9=5) /*Multiple care arrangements*/

/*Create "XK" version of reading, math, and primary care variables for analysis of children at kindergarten entry*/
/*Reading scale score*/
gen XKRSCR2 = .
replace XKRSCR2 = X4RSCR2 if (XKKDATA == 1)
replace XKRSCR2 = X5RSCR2 if (XKKDATA == 2)

/*Math scale score*/
gen XKMSCR2 = .
replace XKMSCR2 = X4MSCR2 if (XKKDATA == 1)
replace XKMSCR2 = X5MSCR2 if (XKKDATA == 2)

/*Primary care variable (year prior to kindergarten entry)*/
gen XKPRIMNW = .
replace XKPRIMNW = Y3PRIMNW if (XKKDATA == 1)
replace XKPRIMNW = X4PRIMNW if (XKKDATA == 2)

/*Create variable labels*/
label variable XKRSCR2 "Reading scale score at first kindergarten entry"
label variable XKMSCR2 "Math scale score at first kindergarten entry"
label variable XKPRIMNW "Primary care arrangement, year prior to kind."
```



**ECLS-B Considerations for Analysis of ECLS-B Data > Analysis of Children at Kindergarten Entry  
> Slide 39 of 44**

```
/*Create value labels for new primary care variable*/  
label define primcr6c 0 "0: No regular care" 1 "1: Relative care" 2 "2: Nonrelative care" 3 "3: Center  
care (not HS)" 4 "4: Head Start" 5 "5: Multiple care arrangements"  
label values XKPRIMNW primcr6c
```

```
save "C:\ECLSB Seminar\Practice.dta", replace
```

```
use "C:\ECLSB Seminar\Practice.dta"
```

```
/*Run unweighted frequencies of primary care variable to check cell counts*/  
tabulate XKPRIMNW if XKWHENK==2 | XKWHENK==3
```

```
/*Run weighted means of reading and math scale scores*/  
summarize XKRSCR2 XKMSCR2 if XKWHENK==2 | XKWHENK==3 [weight=WKR0]
```

```
/*Run weighted means of reading and math scale scores by categories of primary care*/  
sort XKPRIMNW  
by XKPRIMNW: summarize XKRSCR2 XKMSCR2 if XKWHENK==2 | XKWHENK==3 [weight=WKR0]
```

```
/*Use Stata survey procedures to obtain the correct standard errors (syntax not presented here)*/
```