



Considerations for Analysis of ECLS-B Data

Module Objectives

Describe the analysis considerations for using information from

- The Household Roster
 - Identifying Respondents, Parents, and Other Household Members
- Fathers
- Twin Pairs
- Kindergarten Rounds

The Household Roster

Includes information about every individual living in the household with the study child at each round



The Household Roster

- Relationship to child
- Specific type of mother, father, sibling, or nonrelative
- Sex
- Race/ethnicity (may identify an individual as any combination of races, along with whether the individual is Hispanic)
- Age
- Whether currently residing in household
- Reason no longer in household, if applicable

The Household Roster, 9 Months

At 9 months, parent respondents were asked to identify everyone living in the household with the child



The Household Roster

Now I have a few questions about you and your household.

We have listed that you and {CHILD}{and{TWIN}} currently live in this household. Please tell me the names and ages of all the other people who normally live here. Please do not include anyone staying here temporarily who usually lives somewhere else.

The Household Roster, After 9 Months

In the 2-year, preschool, kindergarten 2006, and kindergarten 2007 rounds, household information was updated in order to add new household members and identify members who were no longer in the household



The Household Roster

I am going to read a list of the people who lived in this household at the time of our last interview. As I read each person's name, please tell me if he or she still lives in this household.

Does {FULL NAME} still live in this household?

[If No] Why is {FIRST NAME} no longer living in this household?

Anyone else (living in this household)? {[People sometimes join households as a result of marriage, or a marriage-like partner or a relative or boarder moving in.]}

Structure of the Household Roster

- Each household member is assigned a "position number" (1-21) in the household roster
- Position/roster numbers do not change across rounds
- Individuals remain in the roster whether or not they remained in the household
- Roster variables indicate whether each individual was living in the household at any given round
- Roster variable names begin with "Q" and a number indicating the round of data collection and end with a double-digit roster position number (e.g., Q1REL_01, Q4AGE_15, etc.)



The Household Roster

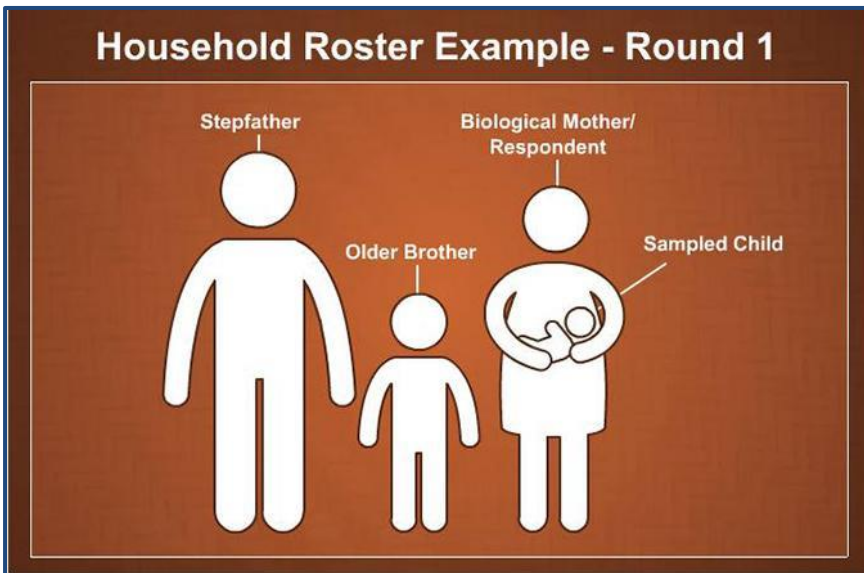
Structure of the Household Roster (Continued)

	Round 1	Round 2	Round 3	Round 4	Round 5
Relationship to child	Q1REL_01- Q1REL_21	Q2REL_01- Q2REL_21	Q3REL_01- Q3REL_21	Q4REL_01- Q4REL_21	Q5REL_01- Q5REL_21
Mother type	Q1MOM_01- Q1MOM_21	Q2MOM_01- Q2MOM_21	Q3MOM_01- Q3MOM_21	Q4MOM_01- Q4MOM_21	Q5MOM_01- Q5MOM_21
Father type	Q1DAD_01- Q1DAD_21	Q2DAD_01- Q2DAD_21	Q3DAD_01- Q3DAD_21	Q4DAD_01- Q4DAD_21	Q5DAD_01- Q5DAD_21
Sister type	Q1SIS_01- Q1SIS_21	Q2SIS_01- Q2SIS_21	Q3SIS_01- Q3SIS_21	Q4SIS_01- Q4SIS_21	Q5SIS_01- Q5SIS_21
Brother type	Q1BRO_01- Q1BRO_21	Q2BRO_01- Q2BRO_21	Q3BRO_01- Q3BRO_21	Q4BRO_01- Q4BRO_21	Q5BRO_01- Q5BRO_21
Age	Q1AGE_01- Q1AGE_21	Q2AGE_01- Q2AGE_21	Q3AGE_01- Q3AGE_21	Q4AGE_01- Q4AGE_21	Q5AGE_01- Q5AGE_21
Sex	Q_SEX_01- Q_SEX_21	N/A	N/A	N/A	N/A
Race (white)	Q_WH_01- Q_WH_21	N/A	N/A	N/A	N/A
Race (black)	Q_BLK_01- Q_BLK_21	N/A	N/A	N/A	N/A
Currently in household?	Q1CURR01- Q1CURR21	Q2CURR01- Q2CURR21	Q3CURR01- Q3CURR21	Q4CURR01- Q4CURR21	Q5CURR01- Q5CURR21



The Household Roster

Household Roster Example, Round 1



The Household Roster

Household Roster Example, Round 1 (Continued)

	Person 1: _01	Person 2: _02	Person 3: _03	Person 4: _04	Person 5: _05
Q1REL_##	1 = Mother	-1 = Not applicable	2 = Father	4 = Brother	-1 = Not applicable
Q1MOM_##	1 = Biological mother	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable
Q1DAD_##	-1 = Not applicable	-1 = Not applicable	3 = Stepfather	-1 = Not applicable	-1 = Not applicable
Q1SIS_##	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable
Q1BRO_##	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable	3 = Stepbrother	-1 = Not applicable
Q1AGE_##	28	1	25	2	-1 = Not applicable
Q_SEX_##	2 = Female	2 = Female	1 = Male	1 = Male	-1 = Not applicable
Q_WH_##	1 = Yes	1 = Yes	1 = Yes	-9 = Not ascertained	-1 = Not applicable
Q_BLK_##	2 = No	2 = No	1 = Yes	-9 = Not ascertained	-1 = Not applicable
Q1CURR##	1 = Yes	1 = Yes	1 = Yes	1 = Yes	2 = No



The Household Roster

Identification Variables - Example, Round 1

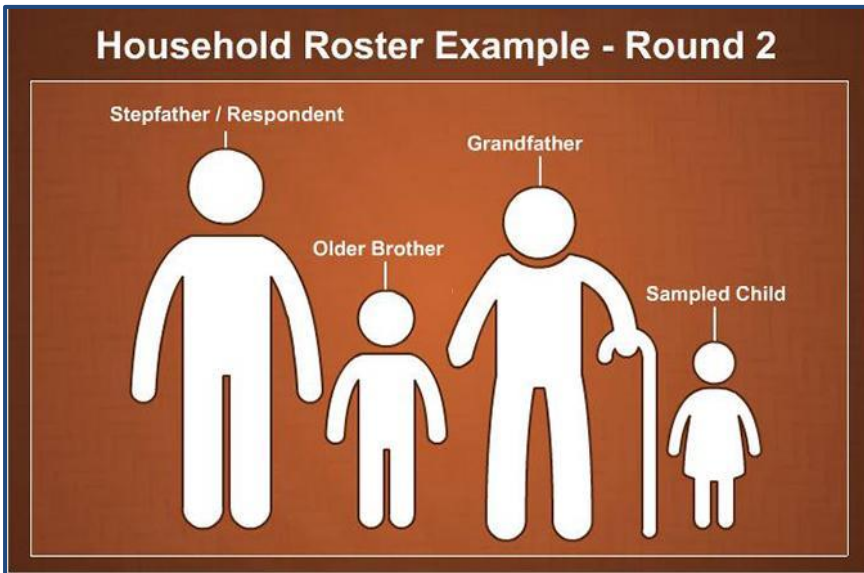
ID variables indicate certain individual's position number for the household roster



The Household Roster

Variable	Description	Value (for example)
Y1RESPID	Respondent ID: Person identified in X1RSPREL and referenced in Respondent Information (RI) section of parent interview	1
Y1MOMID	Mother ID: Person identified in Y1MOMTYP	1
Y1FTHID	Father ID: Person identified in Y1FTHTYP	3
P1SPON_R	Spouse ID: Person referenced in Spouse Information (SI) section of parent interview	3
X1CHID	Sampled child ID: Equals 2 except for twin cases where it can equal 2 or 3	2

Household Roster Example, Round 2



The Household Roster

Household Roster Example, Round 2 (Continued)

	Person 1_01	Person 2_02	Person 3_03	Person 4_04	Person 5_05
Q2REL_##	1 = Mother	-1 = Not applicable	2 = Father	4 = Brother	8 = Grandfather
Q2MOM_##	1 = Biological mother	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable
Q2DAD_##	-1 = Not applicable	-1 = Not applicable	3 = Stepfather	-1 = Not applicable	-1 = Not applicable
Q2SIS_##	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable
Q2BRO_##	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable	3 = Stepbrother	-1 = Not applicable
Q2AGE_##	-1 = Not applicable	2	26	3	57
Q_SEX_##	2 = Female	2 = Female	1 = Male	1 = Male	1 = Male
Q_WH_##	1 = Yes	1 = Yes	1 = Yes	-9 = Not ascertained	-9 = Not ascertained
Q_BLK_##	2 = No	2 = No	1 = Yes	-9 = Not ascertained	-9 = Not ascertained
Q2CURR##	2 = No	1 = Yes	1 = Yes	1 = Yes	1 = Yes
Q2REAS##	4 = Living elsewhere for employment	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable	-1 = Not applicable



The Household Roster

Identification Variables - Example, Round 2

Variable	Description	Value (for example)
Y2RESPID	Respondent ID: Person identified in Y2RSPREL and referenced in Respondent Information (RI) section of parent interview	3
Y2MOMID	Mother ID: Person identified in Y2MOMTYP	-1
Y2FTHID	Father ID: Person identified in Y2FHTTYP	3
P2SPON_R	Spouse ID: Person referenced in Spouse Information (SI) section of parent interview	-1
X2CHID	Sampled child ID: Equals 2 except for twin cases where it can equal 2 or 3	2



The Household Roster

Using Identification Variables

Roster variables and identification variables can help track changes to the household composition longitudinally



The Household Roster

		Round 1	Round 2
Respondent	X/Y*RESPID	1	3
Mother	X/Y*MOMID	1	-1
Father	X/Y*FTHID	3	3
Spouse	P*SPON_R (P*SPONUM)	3	-1
Sampled child	X*CHID	2	2

Relationship Between Identification and Composite Variables

- Parent characteristic composite variables (X/Y*MOMED, X/Y*HMEMP, X/Y*MOMOCC, X/Y*HMRACE, X/Y*HMAG_B, etc.) refer to the people identified by X/Y*MOMID and X/Y*FTHID
- The major sources of data for the composite variables are the Respondent and Spouse Information sections of the parent interview
 - If the respondent and his/her spouse are not identified as a mother/mother figure and/or father/father figure, composite data are either missing or imputed



The Household Roster

Identification of and Collection of Information from Resident Fathers

- How were resident fathers identified?
 - Identified in the parent interview
 - Household member who was the spouse or partner of the parent respondent
- Resident father asked to complete and return a self-administered questionnaire in the 9-month, 2-year, and preschool data collections
- Some fathers were parent interview respondents
 - Completed a section just for fathers in the parent interview, rather than a self-administered questionnaire for fathers



Fathers

Identification and Eligibility of Nonresident Fathers

- Identified in the parent interview
- Limited to biological fathers who did not reside with the child
- The child had to be living with his or her biological mother, who had to be the respondent to the parent interview
- Father must have had minimum level of contact with child or child's mother to be included in study
 - Saw the child at least once in the last month
 - Saw the child at least 7 days in the last 3 months
 - Was in contact with the child's birth mother at least once a month in the 3 months preceding the parent interview



Fathers

Collection of Information from Nonresident Fathers

- A self-administered questionnaire was mailed to the nonresident father to complete and return
- Data from nonresident fathers were collected in the 9-month and 2-year data collections



Fathers

How to Conceptualize Data on Fathers in the ECLS-B

- The unit of analysis is always the child
 - Data are not representative of fathers
- Data describe characteristics of fathers who have regular contact with the child
 - Resident fathers at a given round
 - Biological nonresident fathers to whom mothers granted access, who had recent contact with the child, and who completed the self-administered questionnaire



Fathers

Key Variables for Identifying Fathers

Resident and nonresident father questionnaire status variables	Data Collection		
	9-Month	2-Year	Preschool
Resident father self-administered questionnaire data available	X1STTRFQ = 1	X2STTRFQ = 1	X3STTRFQ = 1
Nonresident father questionnaire data available	X1STTNRQ = 1	X2STTNRQ = 1	N/A
The person who completed the parent interview	X1RSPREL	Y2RSPREL	Y3RSPREL
Type of resident father in the household	Y1FTHTYP	Y2FTHTYP	Y3FTHTYP
Type of father that completed the resident father questionnaire.	F1RELCH	F2RELCH	F3RELCH
Same or different resident father at 9-month, 2-year, and preschool rounds	N/A	Y2SAMERF	Y3SAMERF



Fathers

Additional Data and Analysis Considerations

- Fathers may change across rounds: new fathers in the household, fathers leaving the household
- Some children have data for multiple fathers (i.e., children with a resident father and a nonresident biological father)
- Resident father composite variables are available



Fathers

The ECLS-B Twin Sample

The ECLS-B twin oversample allows for better estimates of twin characteristics than would be possible if twins were sampled at the rate they occur in the population



Twin Pairs

- Round 1 (9 months): ~1,600 twins, 800 twin pairs
- Round 2 (2 years): ~1,500 twins, 750 twin pairs
- Round 3 (Preschool): ~1,400 twins, 700 twin pairs
- Round 4 (Kindergarten 2006): ~1,200 twins, 600 twin pairs
- Round 5 (Kindergarten 2007*): ~400 twins, 200 twin pairs

*Twins of children eligible for round 5 were included in the round 5 sample regardless of their own eligibility

Data Collection for Twins

- Questions in parent interview sections applicable to both twins and for which responses were not expected to vary across twins were asked only once during the parent interview (e.g., Family structure (FS), Community and social support (CS), Respondent information (RI), Household income and assets (HI))
- Parent interview sections specific to the individual child were asked twice, once for each twin (e.g., Child development, literacy and school readiness (CD), Child health (CH), School experiences (SE))
- Other study components handled twin-specific questions differently (see the Data File User's Manual and survey instruments for details)



Twin Pairs

Identifying Twins

- Twins are included as separate child cases (with unique I_IDs) on the ECLS-B data file
- For each child, the composite variable X*TWIN indicates the presence of a sampled twin case on the data file at each round
- I_TWID is the I_ID of a child's sampled twin
- I_TWIDPR is a unique ID for every twin pair
 - I_TWIDPR equals the I_ID of one of the twins in each pair
- X*CHID is the roster position of the child within the household



Twin Pairs

Identifying Twins (Continued)

I_ID	I_TWINID	I_TWINPR	X1TWIN	X1CHID
398474	427038	398474	1 = Twin in HH	2
427038	398474	398474	1 = Twin in HH	3
018843	.	.	0 = No twin in HH	2
<u>Symbol</u>	<u>Label</u>			
.	Sysmis. System missing.			



Twin Pairs

Twin Zygosity

- Zygosity is the degree of identity in the genome of twins
 - Fraternal twins are “dizygotic”
 - Two eggs are independently fertilized, forming two zygotes and two embryos
 - Identical twins are “monozygotic”
 - A single egg is fertilized to form one zygote which then divides into two separate embryos
- Definitive evidence of zygosity is a genetic test
 - The ECLS-B did not collect biological samples from children (i.e., no DNA)



Twin Pairs

Twin Zygosity (Continued)

The ECLS-B included a series of questions in the 2-year parent interview intended to help determine the zygosity of same-sex twin pairs, e.g.



Twin Pairs

- Are {CHILD} and {TWIN} identical or fraternal (non-identical) twins?*
- Are there any differences between your twins in the following physical characteristics... (texture of your twins' hair, shades of your twins' hair colors, eye color, complexion, facial appearance, shapes of your twins' ear lobes)?
- Did the twins' first teeth begin to come in at about the same time?
- Are twins often mistaken for each other... (by respondent, by relatives, by strangers)?
- Would you say that child and twin are as physically alike as "two peas in a pod"...?

*This question was asked in both the 9-month and 2-year data collections

Twin Zygosity (Continued)

- Similar questions were also included in an interviewer remarks questionnaire, which was completed at the end of the home visit
- Responses from these questions were used to create a derived zygosity variable. Guided by Goldsmith (1991),¹ this variable used data from the parent and field interviewer to determine the twins' most likely zygosity
- This twin zygosity variable, along with a description of the coding and a record layout, is available in a supplemental data file included on the ECLSB 9-month through Kindergarten 2007 Restricted-Use DVD in a folder titled "Zygosity"



Twin Pairs

¹Goldsmith, H.H. (1991) A Zygosity Questionnaire for Young Twins: A Research Note. Behavior Genetics, 21: 257-269

Implications of the Kindergarten Round Study Design for Research Questions

- Kindergarten data are not representative of all kindergartners
 - Cohort of children born in the U.S. in 2001
- Data are representative of the characteristics and experiences of children in the cohort at kindergarten entry
 - The data for most cases in the kindergarten 2006 and 2007 collections were collected in the fall



Kindergarten
Rounds

Kindergarten Status - Review

- Round 4 (kindergarten 2006 round): Approximately 75 percent of the ECLS-B children were in kindergarten or higher for the first time in 2006
- Approximately 25 percent of the ECLS-B children were in kindergarten or higher for the first time in 2007 (kindergarten 2007 round, round 5)
- Approximately 4 percent of the ECLS-B sample who were in kindergarten in 2006 repeated kindergarten in 2007



Kindergarten
Rounds

Selection of Data Round

- Which round(s) of data should you use?
 - It depends on your research question
- Use round 4 (kindergarten 2006) data to answer questions
 - about children born in 2001 when they were approximately 5 years old
 - when school enrollment (i.e., kindergarten status) is not likely to be related to the outcome of interest
 - What child and family characteristics are related to the likelihood that a child born in 2001 was overweight when he or she was about 5 years old?
 - What percentage of children born in the U.S. in 2001 had been diagnosed with certain illnesses, such as asthma or repeated ear infections, by the time they were about 5 years old?



Kindergarten
Rounds

Selection of Data Round (Continued)

- Which round(s) of data should you use?
 - It depends on your research question
- Use round 5 (kindergarten 2007) data to answer questions
 - about children born in 2001 who first attended kindergarten in 2006 and repeated kindergarten in 2007
 - What percentage of children born in 2001 attended kindergarten in 2006 and repeated kindergarten in 2007?
 - How is the likelihood of repeating kindergarten related to child, family, and school characteristics? Who is most likely to repeat kindergarten?



Kindergarten
Rounds

Selection of Data Round (Continued)

- Which round(s) of data should you use?
 - It depends on your research question
- Combine round 4 and round 5 data to answer questions
 - about children born in 2001 when they were first in kindergarten
 - when school enrollment (kindergarten status) is likely to be related to the outcome of interest
 - What child and family characteristics relate to children's early reading and math ability in the beginning of kindergarten?
 - Is participation in early care and education prior to kindergarten entry related to children's cognitive and/or socioemotional development as they start kindergarten?



Kindergarten
Rounds

Analysis of Children at Kindergarten Entry

1. Identify the round in which the child was first in kindergarten
2. Choose the variables of interest from both the kindergarten 2006 and kindergarten 2007 rounds
3. Create new "kindergarten round" variables using data from the round in which the child was first in kindergarten
4. Choose the appropriate "WK" weight
5. Analyze data using the new "kindergarten round" variables and weight



Kindergarten
Rounds

Analysis of Children at Kindergarten Entry (Continued)

Identify the round in which the child was first in kindergarten



Kindergarten
Rounds

- XKWHENK
 1. Not enrolled in school in 2006 or 2007
 2. First-time kindergartner in 2006
 3. First-time kindergartner in 2007
 4. Straight to 1st grade in 2006
 5. Straight to 1st grade in 2007
 6. First enrolled in school in 2006 (grade unknown or ungraded)
 7. First enrolled in school in 2007 (grade unknown or ungraded)
 8. Homeschooled (beginning in 2006)
 9. Homeschooled (beginning in 2007)
 10. Unknown
- XKKDATA
 1. First schooling, 2006
 2. First schooling, 2007

Analysis of Children at Kindergarten Entry (Continued)

Choose variables

- Outcome variables and covariates (e.g., demographic variables, family characteristics, independent variables) will likely come from the same round (i.e., the round the child first entered kindergarten)
- In some cases, covariates may come from an earlier round
- Variables only collected in one round (e.g., race, sex, birth weight)
- Variables from the round prior to kindergarten entry



Kindergarten
Rounds

[Choosing Variables Resource Document](#)

Analysis of Children at Kindergarten Entry (Continued)

Create “kindergarten round” variables

- Use XKKDATA and variable(s) of interest to create a new variable that combines kindergarten 2006 and kindergarten 2007 information
- If the child was first in kindergarten in round 4 (XKKDATA = 1), then the new variable will use data collected in 2006
- If the child was first in kindergarten in round 5 (XKKDATA = 2), then the new variable will use data collected in 2007



Kindergarten
Rounds

[Choosing Variables Resource Document](#)

Analysis of Children at Kindergarten Entry (Continued)

Choose a [weight](#)

- Weights that begin with the prefix “WK” are intended for the analysis of children at their entry into kindergarten
- Choose the most appropriate WK weight for your analysis



Kindergarten
Rounds

Analysis of Children at Kindergarten Entry (Continued)

Analyze data

- Analyze data using the new kindergarten round" variables
 - Report results in terms of the ECLS-B target population
 - For example, "The average early reading score for children born in the U.S. in 2001 as they first enter kindergarten is..."



Kindergarten
Rounds

[Example Analysis Resource Document](#)

Children Repeating Kindergarten in 2007

- A screener questionnaire was sent to parents of children who were in kindergarten in the 2006 round
 - Included children homeschooled for kindergarten and children enrolled in ungraded or multi-graded classes
- The kindergarten 2007 round included all children not enrolled in school or homeschooled in 2006, children identified as kindergarten repeaters in 2007, and twins of both of these groups of children



Kindergarten
Rounds

Children Repeating Kindergarten in 2007 (Continued)

- Use X5RPTR to identify repeaters in the data file
 - 1 = Yes, child repeating kindergarten
 - 2 = No, child not repeating kindergarten
 - 3 = Child in 2-year, transitional, or ungraded program
- X5RPTR is “not applicable” (-1) for children who were not enrolled in school in 2006 or who were homeschooled in either or both rounds
- X5RPTR is derived using both teacher and parent reports of grade level and enrollment status
- Limit analytic sample to repeaters (X5RPTR = 1)



Kindergarten
Rounds

Children Repeating Kindergarten in 2007 (Continued)

- Analysts do not need to combine data from multiple rounds to analyze data for repeaters
- Analysis of repeaters should use a “W5” [weight](#)
- The ECLS-B data set does not include data on whether children who entered kindergarten for the first time in 2007 repeated kindergarten in 2008



Kindergarten
Rounds

Module Summary

Described the analysis considerations for using information from

- The Household Roster
 - Identified Respondents, Parents, and Other Household Members
- Fathers
- Twin Pairs
- Kindergarten Rounds

Module Resources

- [Choosing Variables Resource Document](#)
- [ECLS-B Sample Design, Weights, Variance, and Missing Data](#)
- [Analysis Example Resource Document](#)