ICPSR 29462

Head Start Impact Study (HSIS), 2002-2006 [United States]

United States Department of Health and Human Services. Administration for Children and Families. Office of Planning, Research and Evaluation
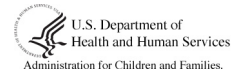
Covariates and Subgroup Variables Data Codebook

About *Research Connections*

These data are made available by the Child Care and Early Education *Research Connections* project. *Research Connections* promotes high quality research in child care and early education and the use of that research in policymaking.

Research Connections is operated by the National Center for Children in Poverty at the Mailman School of Public Health, Columbia University and the Inter-university Consortium for Political and Social Research at the Institute for Social Research, University of Michigan, through a cooperative agreement with the Child Care Bureau, Office of Family Assistance and the Office of Planning, Research, and Evaluation, Administration for Children and Families in the U.S. Department of Health and Human Services.

Terms of Use

Information about Copyrighted Content

<div align="center">

Head Start Impact Study (HSIS)
Crosswalk Document

</div>

To conform with ICPSR filenaming conventions, data and documentation files have been assigned new file names that differ from those assigned by the original data producer, Westat, Inc. Throughout the documentation, you may notice that the data files are referenced by their original names. Please refer to the table below to link the old and new file names.

<div align="center">

Head Start Impact Study (HSIS)

</div>

Data File Description	Original Data File prefix	ICPSR File prefix
Child Experiences Data	ChildExperiences	da29462-0001
Covariates and Subgroup Variables Data	covariates	da29462-0002
Fall 2002 Child Assessment Data	ChildAssessment_fall02	da29462-0003
Fall 2002 Parent Interview Data	ParentInterview_fall02	da29462-0004
Spring 2003 Child Assessment Data	ChildAssessment_spring03	da29462-0005
Spring 2003 Parent Interview Data	ParentInterview_spring03	da29462-0006
Spring 2003 Teacher Survey Data	Teacher_spring03	da29462-0007
Spring 2003 Teacher's Child Report Data	TCR_spring03	da29462-0008
Spring 2003 Center Director Interview Data	CenterDirector_spring03	da29462-0009
Spring 2003 Classroom Observation Data	ClassroomObservation_spring03	da29462-0010
Spring 2003 Care Provider Interview Data	CareProvider_spring03	da29462-0011
Spring 2003 Family Child Care Observation Data	FamilyChildcareObservation_spring03	da29462-0012
Spring 2004 Child Assessment Data	ChildAssessment_spring04	da29462-0013
Spring 2004 Parent Interview Data	ParentInterview_spring04	da29462-0014
Spring 2004 Teacher Survey Data	Teacher_spring04	da29462-0015
Spring 2004 Teacher's Child Report Data	TCR_spring04	da29462-0016
Spring 2004 Center Director Interview Data	CenterDirector_spring04	da29462-0017
Spring 2004 Classroom Observation Data	ClassroomObservation_spring04	da29462-0018
Spring 2004 Care Provider Interview Data	CareProvider_spring04	da29462-0019
Spring 2004 Family Child Care Observation Data	FamilyChildcareObservation_spring04	da29462-0020
Spring 2005 Child Assessment Data	ChildAssessment_spring05	da29462-0021
Spring 2005 Parent Interview Data	ParentInterview_spring05	da29462-0022
Spring 2005 Teacher Survey Data	Teacher_spring05	da29462-0023
Spring 2005 Teacher's Child Report Data	TCR_spring05	da29462-0024
Spring 2006 Child Assessment Data	ChildAssessment_spring06	da29462-0025
Spring 2006 Parent Interview Data	ParentInterview_spring06	da29462-0026
Spring 2006 Teacher Survey Data	Teacher_spring06	da29462-0027
Spring 2006 Teacher's Child Report Data	TCR_spring06	da29462-0028
Weights	weights	da29462-0029
JKN Factors	JKNFac76	da29462-0030

For more detailed information about this data collection, please consult the User Guide.

Head Start Impact Study

Covariates and Subgroup Variables

Codebook

Number	Name	Label
1	HSIS_RAPROGID	RANDOM ASSIGNMENT PROGRAM ID
2	HSIS_RACNTRID	RANDOM ASSIGNMENT CENTER ID
3	HSIS_CHILDID	CHILD ID
4	CHILDCOHORT	CHILD COHORT
5	CHILDRESULTGROUP	TREATMENT/CONTROL
6	CROSSOVER	FLAG FOR SAMPLED CONTROL WHO RECEIVED HS
7	NOSHOW	FLAG FOR SAMPLED TREATMENT WHO DID NOT RECEIVE HS
8	D_1CGVAGE	BASELINE CAREGIVER AGE (AS OF SEPT 1, 2002)
9	D_BBIO_WITHCH	BOTH BIOLOGICAL PARENTS LIVE WITH CHILD
10	D_CDS	CES-D DEPRESION SCALE (4 CATEGORIES)
11	D_CF_CHSEX01	CHILD GENDER
12	D_CHRACE3	CHILD RACE (3 CATEGORIES)
13	D_FALLSPR_TESTLANG	LANGUAGE CHILD TEST IN FALL 2002/SPRING 2003
14	D_LANG_HM	HOME LANGUAGE
15	D_LQUART	CHILD'S PRE-ACADEMIC SKILLS (LOWEST QUARTILE)
16	D_LS1	LANGUAGE SAMPLE DUMMY VARIABLE INDICATOR
17	D_MBIO_RECIMM	BIOLOGICAL MOTHER RECENT IMMIGRANT STATUS
18	D_MORACE3	BIOLOGICAL MOTHER/CAREGIVER RACE/ETHNICITY (3 CATEGORIES)
19	D_MOTHHIGHED3	MOTHER'S EDUCATION STATUS (3 CATEGORIES)
20	D_MOTH_MARSTAT	MOTHER'S MARITAL STATUS
21	D_PI_SPECNEEDS_F02	CHILD HAS SPECIAL NEEDS - FALL 2002
22	D_RISK3	HOUSEHOLD RISK INDEX (3 CATEGORIES)
23	D_TEENMOM	TEEN MOM
24	D_URBAN	URBANICITY

Name: HSIS_RAPROGID

Label: RANDOM ASSIGNMENT PROGRAM ID

Number: 1

Type: Numeric

Width: 3

Decimals: 0
Location: 1-3

Explanation
ID for program of random assignment.

THE FOLLOWING COVARIATES ARE LOCATED IN THE PARENT INTERVIEW OR CHILD ASSESSMENT FILES:

CHILD ASSESSMENT FILES D_CA_AGEATTESTING (CHILD AGE AT TIME OF SPRING ASSESSMENT) D_CA_WEEKOFTESTING (NUMBER OF WEEKS ELAPSED BETWEEN 9/1/2002 AND FALL TESTING)

PARENT INTERVIEW FILES D_PI_WEEKOFINTERVIEW (NUMBER OF WEEKS ELAPSED BETWEEN 9/1/2002 AND PARENT INTERVIEW)

Name: HSIS_RACNTRID
Label: RANDOM ASSIGNMENT CENTER ID
Number: 2
Type: Numeric
Width: 4
Decimals: 0
Location: 4-7

Explanation
ID for center of random assignment.

Name: HSIS_CHILDID
Label: CHILD ID
Number: 3
Type: Numeric
Width: 6
Decimals: 0
Location: 8-13

Name: CHILDCOHORT
Label: CHILD COHORT
Number: 4
Type: Character
Width: 1
Location: 14

Explanation

Designates whether the child is in the 3-year old or 4-year old
cohort.

Frequency Code Description

Frequency	Code	Description
2449	3	Cohort 3
1993	4	Cohort 4

Name: CHILDRESULTGROUP
Label: TREATMENT/CONTROL
Number: 5
Type: Numeric
Width: 1
Decimals: 0
Location: 15

Explanation

Designates whether the child is in the treatment or control group.

Frequency Code Description

Frequency	Code	Description
2646	2	Treatment
1796	3	Control

Name: CROSSOVER
Label: FLAG FOR SAMPLED CONTROL WHO RECEIVED HS
Number: 6
Type: Numeric
Width: 1
Decimals: 0
Location: 16

Explanation

Designates a child who was assigned as a control but participated in
Head Start

Frequency Code Description

Frequency	Code	Description
4225	0	No
217	1	Yes

Valid: 4442
Minimum: 0.00
Maximum: 1.00
Mean: 0.05
Standard Deviation: 0.22

Name: NOSHOW

Label: FLAG FOR SAMPLED TREATMENT WHO DID NOT RECEIVE HS

Number: 7

Type: Numeric

Width: 1

Decimals: 0

Location: 17

Explanation

Designates a child who was assigned to Head Start but did not participate in Head Start

Frequency	Code	Description
3938	0	No
504	1	Yes

Valid: 4442

Minimum: 0.00

Maximum: 1.00

Mean: 0.11

Standard Deviation: 0.32

Name: D_1CGVAGE

Label: BASELINE CAREGIVER AGE (AS OF SEPT 1, 2002)

Number: 8

Type: Numeric

Width: 2

Decimals: 0

Location: 18-19

Explanation

Assigned as the mother's age if the respondent's relationship to the child is the mother both in Fall 2002 and Spring 2003. Assigned as the father's age if the respondent's relationship to the child is the father in both Fall 2002 and Spring 2003. For all other cases where the relationship is neither the father nor the mother, the respondent's age is assigned. This variable also includes imputed values.

Valid: 4442

Minimum: 16.00

Maximum: 78.00

Mean: 28.91

Standard Deviation: 7.34

Name: D_BBIO_WITHCH

Label: BOTH BIOLOGICAL PARENTS LIVE WITH CHILD

Number: 9

Type: Numeric

Width: 1

Decimals: 0

Location: 20

Explanation

This variable is based on the Fall 2002 parent interview. It is based
on the responses to P1BIOMO, P4MALIVE, P1MINHH, P1RFATH, P4FALIVE,
P1FINHH and includes imputed values.

DERIVATION:
if p1biomo = 2 and p4malive = 1 and p1minhh = 2
then d_mbio_withch = 0;
else if p1biomo = 2 and p4malive = 1 and p1minhh = 1
then d_mbio_withch = 1;
else if p1biomo = 2 and p4malive = 2 and p1minhh = 1
then d_mbio_withch = 99;
else if p1biomo = 2 and p4malive = 2 and p1minhh = 2
then d_mbio_withch = 0;
else if p1biomo = 1 and p4malive = 1 and p1minhh = 1
then d_mbio_withch = 1;
else if p1biomo = 1 and p4malive = 1 and p1minhh = 2
then d_mbio_withch = 0;
else if p1biomo = 1 and p4malive = 2 and p1minhh = 1
then d_mbio_withch = 1;
else if p1biomo = 2 and p4malive = 9 and p1minhh = 1
then d_mbio_withch = 1;
else if p1biomo = 2 and p4malive = 9 and p1minhh = 2
then d_mbio_withch = 0;
else if p1biomo = 2 and p4malive = . and p1minhh = 1
then d_mbio_withch = 1;
else if p1biomo = 1 and p4malive = . and p1minhh = .
then d_mbio_withch = 1;
else d_mbio_withch = .;

if p1rfath = 2 and p4falive = 1 and p1finhh = 2
then d_fbio_withch = 0;
else if p1rfath = 2 and p4falive = 1 and p1finhh = 1
then d_fbio_withch = 1;
else if p1rfath = 2 and p4falive = 2 and p1finhh = 1
then d_fbio_withch = 99;
else if p1rfath = 2 and p4falive = 2 and p1finhh = 2
then d_fbio_withch = 0;
else if p1rfath = 1 and p4falive = 1 and p1finhh = 1
then d_fbio_withch = 1;
else if p1rfath = 1 and p4falive = 1 and p1finhh = 2
then d_fbio_withch= 0;
else if p1rfath = 1 and p4falive = 2 and p1finhh = 1

then d_fbio_withch = 1;
else if p1rfath = 2 and p4falive in (7, 8, 9) and p1finhh = 1
then d_fbio_withch = 1;
else if p1rfath = 2 and p4falive in (7, 8, 9) and p1finhh = 2
then d_fbio_withch = 0;
else if p1rfath = 1 and p4falive = . and p1finhh = .
then d_fbio_withch = 1;
else if p1rfath = 2 and p4falive = . and p1finhh = 1
then d_fbio_withch = 1;
else d_fbio_withch = .;

if d_mbio_withch = 0 or d_fbio_withch = 0
then d_bbio_withch = 0;
else if d_mbio_withch = 1 and d_fbio_withch = 1
then d_bbio_withch = 1;
else if d_mbio_withch = 99 or d_fbio_withch = 99
then d_bbio_withch = .;
else d_bbio_withch = .;

Frequency	Code	Description
2239	0	No
2203	1	Yes

Valid:	4442
Minimum:	0.00
Maximum:	1.00
Mean:	0.50
Standard Deviation:	0.50

Name:	D_CDS
Label:	CES-D DEPRESION SCALE (4 CATEGORIES)
Number:	10
Type:	Numeric
Width:	1
Decimals:	0
Location:	21

Explanation

This variable is based on the Fall 2002 parent interview. It is based
on the responses to P1BOTHR P1NOEAT P1BLUES P1MINDON P1DEPR P1EFFRT
P1FFUL P1SLRES P1TKLES P1LONLY P1SAD P1NOTGO.

DERIVATION:
The responses are first rescaled so that 7,8,9 are assigned missing
values.

1->0, 2->1, 3->2, 4->3.
The scale is the sum of these values.
If the parent did not respond to at least 3 elements this variable is

set to missing.

Frequency	Code	Description
1871	1	No depressive symptoms
909	2	Mild depressive symptoms
456	3	Moderate depressive symptoms
341	4	Severe depressive symptoms
865	BLANK	Missing

Missing Data Code: BLANK

Name: D_CF_CHSEX01
Label: CHILD GENDER
Number: 11
Type: Numeric
Width: 1
Decimals: 0
Location: 22

Explanation
This variable is based on the response to P1CHSEX in the Fall 2002
parent interview. When this response is missing, information collected
during recruitment were used. This variable includes imputed values.

Frequency	Code	Description
2239	0	Male
2203	1	Female

Name: D_CHRACE3
Label: CHILD RACE (3 CATEGORIES)
Number: 12
Type: Numeric
Width: 1
Decimals: 0
Location: 23

Explanation
This variable is primarily based on responses from the Fall 2002
parent interview. When this data are not available, information
collected during recruitment are used to construct the variable.
This variable includes imputed values.

Frequency	Code	Description
1418	1	White/other

| 1354 | 2 Black |
| 1670 | 3 Hispanic |

Name: D_FALLSPR_TESTLANG
Label: LANGUAGE CHILD TEST IN FALL 2002/SPRING 2003
Number: 13
Type: Numeric
Width: 1
Decimals: 0
Location: 24

Explanation

This variable is created by comparing the language of assessment in the Fall 2002 and Spring 2003. It is also used to subset the analysis sample.

Frequency	Code	Description
3301	0	English - English
1141	1	Spanish - English
0	2	Spanish - Spanish

Name: D_LANG_HM
Label: HOME LANGUAGE
Number: 14
Type: Numeric
Width: 1
Decimals: 0
Location: 25

Explanation

This variable is based on the response to P1LANHO in the Fall 2002 parent interview. When the response is not available, information obtained during recruitment are used. This variable includes imputed values.

Frequency	Code	Description
1324	0	English
3118	1	Spanish

Name: D_LQUART
Label: CHILD'S PRE-ACADEMIC SKILLS (LOWEST QUARTILE)
Number: 15

Type: Numeric
Width: 1
Decimals: 0
Location: 26

Explanation

This variable is created using the following cutoffs from the Fall
2002 Pre-academic W-score:

Cohort 3
(English-English) 320; (Spanish-English) 347.7;

Cohort 4
(English-English) 342; (Spanish-English) 363.7.

Different cutoffs for the language groups were used because the
underlying tests are different.

Frequency Code Description

Frequency	Code	Description
3345	0	Not lowest quartile
1097	1	Lowest quartile

Valid: 4442
Minimum: 0.00
Maximum: 1.00
Mean: 0.25
Standard Deviation: 0.43

Name: D_LS1
Label: LANGUAGE SAMPLE DUMMY VARIABLE INDICATOR
Number: 16
Type: Numeric
Width: 1
Decimals: 0
Location: 27

Explanation

This variable is derived from D_FALLSPR_TESTLANG. The value is 1 if
the the child is in the English-English subsample.

Frequency Code Description

Frequency	Code	Description
1208	0	Dual language learners
3234	1	English speaking

Valid: 4442
Minimum: 0.00

Maximum: 1.00
Mean: 0.73
Standard Deviation: 0.45

Name: D_MBIO_RECIMM
Label: BIOLOGICAL MOTHER RECENT IMMIGRANT STATUS
Number: 17
Type: Numeric
Width: 1
Decimals: 0
Location: 28

Explanation

This variable is based on the responses to P1MBORN and P1MYRS in the
Fall 2020 parent interview. Both responses must be non-missing
otherwise this variable is considered missing. A mother is considered
a recent immigrant if she was foreign born and was in the United
States for less than 10 years.

Frequency Code Description

Frequency	Code	Description
3587	0	No
855	1	Yes

Valid: 4442
Minimum: 0.00
Maximum: 1.00
Mean: 0.19
Standard Deviation: 0.39

Name: D_MORACE3
Label: BIOLOGICAL MOTHER/CAREGIVER RACE/ETHNICITY (3 CATEGORIES)
Number: 18
Type: Numeric
Width: 1
Decimals: 0
Location: 29

Explanation

This variable is based on responses from the Fall 2002 parent
interview. When data were not available, information collected
during recruitment was used to construct the variable. This variable
includes imputed values.

Frequency Code Description

1496	1 White
1348	2 Black
1598	3 Hispanic

Name: D_MOTHHIGHED3
Label: MOTHER'S EDUCATION STATUS (3 CATEGORIES)
Number: 19
Type: Numeric
Width: 1
Decimals: 0
Location: 30

Explanation

This variable is based on the response to P1GRMO in the Fall 2002
parent interview. It includes imputed values.

DERIVATION:
if p1grmo in (00, 01) then d_moth_highed = 1;
else if p1grmo in (02, 03) then d_moth_highed = 2;
else if p1grmo = 04 then d_moth_highed = 3;
else if p1grmo = 05 then d_moth_highed = 4;
else if p1grmo in (06, 07) then d_moth_highed = 5;
else if p1grmo = 08 then d_moth_highed = 6;
else if p1grmo = 09 then d_moth_highed = 7;
else if p1grmo = 10 then d_moth_highed = 8;
else if p1grmo in (11, 12, 13, 14) then d_moth_highed = 9;
else if p1grmo in (97,98,99) then d_moth_highed = .;
else d_moth_highed = .;

if d_moth_highed in (1,2) then d_mothhighed3 = 1;
else if d_moth_highed in (3,4) then d_mothhighed3 = 2;
else if d_moth_highed >= 5 then d_mothhighed3 = 3;

Frequency Code Description

1687	1 Less than High School
1481	2 High school diploma or GED
1274	3 Beyond High School

Name: D_MOTH_MARSTAT
Label: MOTHER'S MARITAL STATUS
Number: 20
Type: Numeric
Width: 1
Decimals: 0
Location: 31

Explanation

This variable is based on the response to P1MARMO in the Fall 2002
parent interview. It also includes imputed values.

DERIVATION:
if p1marmo = 5 then d_moth_marstat = 1;
else if p1marmo = 1 then d_moth_marstat = 2;
else if p1marmo in (2,3,4) then d_moth_marstat = 3;
else if p1marmo in (7,8,9) then d_moth_marstat = .;

Frequency	Code	Description
1742	1	Never married
1972	2	Currently married
724	3	Separated/divorced/widowed
0	97	Refused
0	98	Don't know
0	99	Not ascertained
4	BLANK	Missing

Missing Data Codes: 97,98,99,BLANK

Name: D_PI_SPECNEEDS_F02
Label: CHILD HAS SPECIAL NEEDS - FALL 2002
Number: 21
Type: Numeric
Width: 1
Decimals: 0
Location: 32

Explanation
This variable is based on the response to P1DOC in the Fall 2002
parent interview and includes imputed values.

DERIVATION:
if p1doc = 1 then d_pi_specneeds = 1;
else if p1doc = 2 then d_pi_specneeds = 0;
else d_pi_specneeds = .;

Frequency	Code	Description
3872	0	Not special needs
570	1	Special needs

Valid:	4442
Minimum:	0.00
Maximum:	1.00
Mean:	0.13
Standard Deviation:	0.33

Name: D_RISK3

Label: HOUSEHOLD RISK INDEX (3 CATEGORIES)

Number: 22

Type: Numeric

Width: 1

Decimals: 0

Location: 33

Explanation

This is a scale created based on the sum of five variables:
(1) whether the household received food stamps or TANF in Fall 2002,
(2) if neither parent was a high school graduate,
(3) if neither parent is working,
(4) if the mother was a teen mother,
(5) and if the mother is a single mother.

DERIVATION:
if p1fstmp = 2 then d_rec_fdstmp = 0;
else if p1fstmp = 1 then d_rec_fdstmp = 1;
else if p1fstmp in (7,8,9) then d_rec_fdstmp = .;
else d_rec_fdstmp = .;

if p1tanf = 2 then d_rec_tanf = 0;
else if p1tanf = 1 then d_rec_tanf = 1;
else if p1tanf in (7,8,9) then d_rec_tanf = .;
else d_rec_tanf = .;

if d_rec_fdstmp = 1 or d_rec_tanf = 1 then d_tanforfs = 1;
else if d_rec_fdstmp ne . and d_rec_tanf ne . then d_tanforfs = 0;

if p1grmo in (01 02 03) or p1grfa in (01 02 03) then d_nohighsch = 1;
else if d_fath_highed in (1,2) or d_moth_highed in (1,2)
then d_nohighsch = 1;
else if p1grmo ne . or p1grfa ne . or d_moth_highed ne .
or d_moth_highed ne . then d_nohighsch = 0;

if d_moth_empstat = 3 and d_fath_empstat = 3 then d_nowork = 1;
else if d_moth_empstat ne . or d_fath_empstat ne .
then d_nowork = 0;

if d_moth_marstat in (1,3) then d_singlemom = 1;
else if d_moth_marstat ne . then d_singlemom = 0;

Teen Mom: See below.

d_risk = sum(d_teenmom, d_tanforfs, d_nohighsch,
d_nowork, d_singlemom);

if d_risk in (0,1,2) then d_risk3 = 1 ;
else if d_risk = 3 then d_risk3 = 2 ;
else if d_risk in (4,5) d_then risk3 = 3 ;

Frequency	Code	Description
3383	1	Low/no risk
741	2	Moderate risk
318	3	High risk

Name: D_TEENMOM
Label: TEEN MOM
Number: 23
Type: Numeric
Width: 1
Decimals: 0
Location: 34

Explanation
This variable is computed based on the age of the mother. If the age is 18 or less the child has a teen mom.

Frequency	Code	Description
3690	0	No
752	1	Yes

Valid:	4442
Minimum:	0.00
Maximum:	1.00
Mean:	0.17
Standard Deviation:	0.38

Name: D_URBAN
Label: URBANICITY
Number: 24
Type: Numeric
Width: 1
Decimals: 0
Location: 35

Explanation
Whether the child lives in an urban or non-urban environment. This variable is based on geo-coding of the center of random assignment.

Frequency	Code	Description
696	0	Not urban
3746	1	Urban