NATIONAL YOUTH IN TRANSITION DATABASE (NYTD)

OUTCOMES FILE USER'S GUIDE

FY 2011 COHORT: WAVES 1 AND 2

NDACAN DATASET #182

December 1, 2014

NDACAN NATIONAL DATA ARCHIVE ON CHILD ABUSE AND NEGLECT

Cornell University
Ithaca, NY 14853
ndacan@cornell.edu

NATIONAL YOUTH IN TRANSITION DATABASE (NYTD)

OUTCOMES FILE 2011 COHORT WAVES 1 & 2

NDACAN Dataset #182

DATA PROVIDED BY

Children's Bureau
Administration on Children, Youth and Families
330 C Street, S.W.
Washington, DC 20204

FUNDED BY

Children's Bureau
Administration on Children, Youth and Families
Administration for Children and Families
United States Department of Health and Human Services

DISTRIBUTED BY

National Data Archive on Child Abuse and Neglect

USER'S GUIDE AND CODEBOOK WRITTEN BY

National Data Archive on Child Abuse and Neglect
in collaboration with
Tammy White, Ph.D.
Administration on Children, Youth and Families, DHHS

TABLE OF CONTENTS

ABSTRACT

The John H. Chafee Foster Care Independence Program (CFCIP) was initiated in an effort to improve outcomes for youth in foster care who are likely to reach their 18th birthday without having found a permanent home. The program provides funding to states to develop and administer programs designed to help ease the transition from foster care to independence.

The law that created CFCIP also required states to develop 1) a system for tracking the services provided through CFCIP, and 2) a method for collecting outcome measures so that the effectiveness of the program can be assessed.

These two components together form the National Youth in Transition Database (NYTD). The files contain data from all 50 states, as well as from the District of Columbia and Puerto Rico.

The Services component contains cross-sectional information on the services provided by states under CFCIP and the youth who receive those services. Data are submitted by the states every 6 months on a continuing basis. The Services file is available separately and can be ordered from NDACAN (www.NDACAN.Cornell.edu).

The Outcomes component contains the results of surveys conducted with youth to examine certain well-being, financial, and educational outcomes as they get older. Data from a survey on outcomes for a cohort of youth is provided every other year beginning with federal fiscal year 2011 (Oct 1, 2010 through Sep 30, 2011). There are three phases of outcome data for a given cohort -- a baseline survey during the year of the youth's 17 birthday, and two follow-ups: one at age 19, one at age 21. New cohorts will be established every three years (2014, 2017, 2020…) after the initial one in FY2011.

ACKNOWLEDGEMENT OF SOURCE

Authors should acknowledge the National Data Archive on Child Abuse and Neglect and the Children's Bureau when they publish manuscripts that use data provided by the Archive. Users of these data are urged to follow some adaptation of the statement below.

> *The data used in this publication were made available by the National Data Archive on Child Abuse and Neglect. Data from the National Youth in Transition Database were originally collected by the states and provided to the Children's Bureau. Funding was provided by the Children's Bureau, U.S. Department of Health and Human Services. The collector of the original data, the funder, the Archive, Cornell University and their agents or employees bear no responsibility for the analyses or interpretations presented here.*

PUBLICATION CITATION REQUIREMENT

In accordance with the *Terms of Use Agreement* for these datasets, users of these data are required to provide citations for any published work or report based wholly or in part on these data with the Archive. To obtain the Terms of Use Agreement for the NYTD Outcomes File, go to:

http://www.ndacan.cornell.edu/datasets/request-dataset.cfm

OVERVIEW OF NYTD

This section applies to both the Services and the Outcomes files

***NOTE**: The terms "variable" and "element" are used interchangeably in this document. They refer to the same entity.*

PURPOSE

The optimal outcome for children in foster care is permanency – a permanent home, either by reunification with the parents or adoption to a loving family. Some children, especially those who enter foster care when they are older, never find a permanent home. Each year, about 10% of children who exit foster care are released due to having reached the state's age limit for eligibility. These children are at particular risk for negative outcomes. Emancipated foster care youth encounter homelessness, incarceration, poor educational outcomes, lack of health insurance, unemployment, unplanned parenthood, and poverty at much higher rates than other children.

In 1999, the John H. Chafee Foster Care Independence Act (42 USC § 677) was enacted. The goal of the Chafee Act is to support such youth by providing states with flexible funding for the design and administration of programs that will:

> *(1) identify children who are likely to remain in foster care until 18 years of age and to help these children make the transition to self-sufficiency by providing services such as assistance in obtaining a high school diploma, career exploration, vocational training, job placement and retention, training in daily living skills, training in budgeting and financial management skills, substance abuse prevention, and preventive health activities (including smoking avoidance, nutrition education, and pregnancy prevention);*
>
> *(2) help children who are likely to remain in foster care until 18 years of age receive the education, training, and services necessary to obtain employment;*
>
> *(3) help children who are likely to remain in foster care until 18 years of age prepare for and enter postsecondary training and education institutions;*
>
> *(4) provide personal and emotional support to children aging out of foster care, through mentors and the promotion of interactions with dedicated adults;*

(5) provide financial, housing, counseling, employment, education, and other appropriate support and services to former foster care recipients between 18 and 21 years of age to complement their own efforts to achieve self-sufficiency and to assure that program participants recognize and accept their personal responsibility for preparing for and then making the transition from adolescence to adulthood;

(6) make available vouchers for education and training, including postsecondary training and education, to youths who have aged out of foster care; and

(7) provide the services referred to in this subsection to children who, after attaining 16 years of age, have left foster care for kinship guardianship or adoption.

The act also requires the U.S. Department of Health and Human Services (DHHS) to collect two kinds of data:

1) Information about services and those who receive them, including "the number and characteristics of children receiving services", and "the type and quantity of services being provided."

2) Outcome data, including "measures of educational attainment, high school diploma, employment, avoidance of dependency, homelessness, nonmarital childbirth, incarceration, and high-risk behaviors."

These two datasets together constitute the National Youth in Transition Database.

GEOGRAPHIC AREA

NYTD data come from all 50 states, the District of Columbia, and Puerto Rico. However, due to confidentiality issues, Connecticut data are not currently available to the research community.

NYTD FILES

There are two NYTD datasets, one of all youth who receive independent living services using funds provided through the Chaffee Act, and the other for the results of a periodic survey of youth who turn 17 in certain years, along with follow-up surveys at ages 19 and 21.

Data from both populations are collected by the states and submitted to the Children's Bureau. Although there is some overlap, the two sets of data differ in the number and identity of the cases they cover, and the variables they contain. The number of youth who receive services is much larger than the number eligible to take the Outcomes Survey. Only 5% of those who received services are in the baseline outcomes survey population. However, most of the youth in the 2011 cohort who responded to the age 19 follow-up survey have received Chaffee services at some point.

Figure 1 (page 7) indicates the relative sizes of the Services and Outcomes files.

CONFIDENTIALITY PROTECTION

Because the same youth may appear in NYTD and the AFCARS Foster Care file, and the files can be linked, the same confidentiality protections used in AFCARS are applied to the NYTD files. These are:

- The child's day of birth (DOB) is recoded to the 15th of the month. This adjustment applies to NYTD element #4, *DOB*.

THE OUTCOMES FILE

COHORTS

Starting with the 2011 federal fiscal year, and every three years thereafter, all youth who reach their 17th birthday in the fiscal year and are in foster care within the 45-day period following their birthday (73 FR 10342) will be eligible for the outcomes survey. Youth who complete the survey will be followed-up two and four years later – at ages 19 and 21 -- with the same survey. Below is the schedule for outcomes data collection through the year 2020.

Fiscal Year	Cohort	Data Collected
2011	2011	Baseline Outcomes Survey (Age 17 in FC)
2012	-	
2013	2011	Age 19 Follow-up Survey
2014	2014	Baseline Outcomes Survey (Age 17 in FC)
2015	2011	Age 21 Follow-up Survey
2016	2014	Age 19 Follow-up Survey
2017	2017	Baseline Outcomes Survey (Age 17 in FC)
2018	2014	Age 21 Follow-up Survey
2019	2017	Age 19 Follow-up Survey
2020	2020	Baseline Outcomes Survey (Age 17 in FC)

THE THREE WAVES

WAVE 1 BASELINE POPULATION AND COHORT:

Baseline Population. All foster care youth who turn 17 in the baseline year are in the baseline population. All youth in the baseline population are required to be contacted and asked to complete the NYTD outcomes survey. Demographic data for all baseline youth will be recorded in the Wave 1 file, regardless of whether they respond to the survey.

Cohort. The Cohort is a subset of the Baseline Population. To be in the Cohort, a youth must meet the following conditions:

a) Youth is in the Baseline Population
b) Youth is in foster care on the day of the survey
c) Youth participated in the survey
d) Youth completed the survey within 45 days of her/his 17th birthday

e) At least one answer to Elements 37-58 is a valid answer other than "declined" or "not applicable" or all values are missing

The survey responses are recorded in the Wave 1 file along with demographic info for all youth in the baseline population.

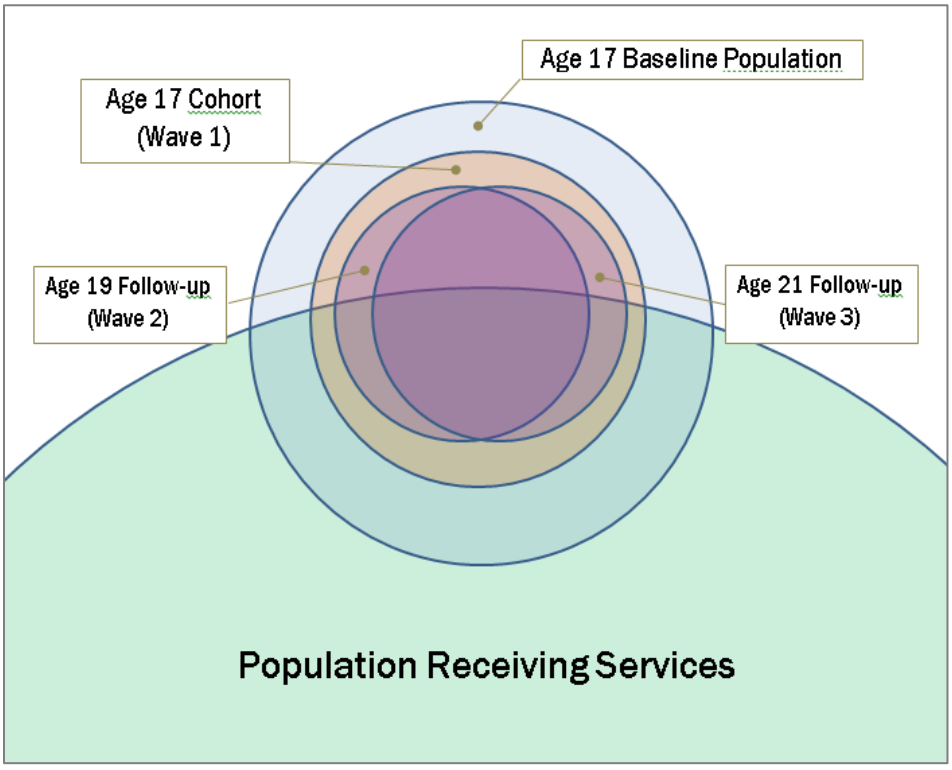
WAVE 2: TWO-YEAR FOLLOW-UP.

Youth in the cohort receive a follow-up survey during the 6-month reporting period that contains their 19th birthday. For the FY2011 cohort, this survey was conducted in FY2013, and the results are included in this version of the dataset.

WAVE 3: FOUR-YEAR FOLLOW-UP.

Youth in the cohort receive a second follow-up survey during the 6-month reporting period that contains their 21st birthday. The survey questions are the same for both follow-ups. The Age 19 follow-up survey for the 2011 cohort is scheduled to be conducted in FY 2015.

FIGURE 1: RELATIONSHIPS BETWEEN WAVES:



DATA COLLECTION

For the Baseline survey, results are reported to the Children's Bureau at 6-month intervals. The "A" period covers the first 6 months of the fiscal year (October through March); the "B" period covers the second 6 months (April through September). The Baseline Survey is administered during the 45-day period following the youth's 17th birthday. If a child's birthday occurs during the last 45 days of the "B" period, the state still has 45 days to collect the data. If the survey is administered after the end of the fiscal year, the data will be reported along with the "A" period data for the following year. This means that the full complement of baseline data will not be complete until after the May 15 deadline for submission of the "A" period file for the following fiscal year.

For the two follow-up surveys, responses are collected any time in the six-month semi-annual period that includes the youths' 19th and 21st birthdays. States are encouraged to collect the data early in the period to avoid performing a survey in one period but reporting it in the next.

SAMPLING

For the Baseline population, no sampling is done. The Baseline population consists of *all* youth in foster care at any point during the 45-day period beginning on their 17th birthday.

The Cohort is a self-selected, non-probabilistic sample of youth in the baseline population. Because youth are not selected randomly, there is no guarantee that the Cohort is representative of the Baseline population (i.e. 17-year-old kids in foster care).

Once the Cohort is selected, probabilistic sampling *may* be used to determine the two follow-up populations (at 19 and 21 years of age). Sampling is done once, and the same sample is used for both follow-up surveys. For the 2011 Cohort, twelve states opted to use sampling for their follow-up outcome surveys (Georgia, Iowa, Illinois, Indiana, Kentucky, Louisiana, Massachusetts, Ohio, Pennsylvania, Tennessee, Texas, and Washington).

The NYTD regulations specify the following regarding the sampling frame, sampling method, and sample size (73 FR 10371 §1356.84):

(b) The State agency must select the follow-up sample using simple random sampling procedures based on random numbers generated by a computer program, unless ACF approves another sampling procedure. The sampling universe consists of youth in the baseline population consistent with 45 CFR 1356.81(b) who participated in the State agency's data collection at age 17.

(c) The sample size is based on the number of youth in the baseline population who participated in the State agency's data collection at age 17.

(1) If the number of youth in the baseline population who participated in the outcome data collection at age 17 is 5,000 or less, the State agency must calculate the sample size using the formula in appendix C of this part, with the Finite Population Correction (FPC). The State agency must increase the resulting number by 30 percent to allow for attrition, but the sample size may not be larger than the number of youth who participated in data collection at age 17.

(2) If the number of youth in the baseline population who participated in the outcome data collection at age 17 is greater than 5,000, the State agency must calculate the sample size using the formula in appendix C of this part, without the FPC. The State agency must increase the resulting number by 30 percent to allow for attrition, but the sample size must not be larger than the number of youth who participated in data collection at age 17.

Appendix C can be found at
http://www.law.cornell.edu/cfr/text/45/1356/appendix-C.

No state had more than 5000 youth in their cohort, so the Finite Population Correction (FPC) applies to all states.

SURVEY ADMINISTRATION

Under NYTD rules, states have the discretion to choose the methods used to administer the outcomes survey to youth (e.g., in person, via the Internet or over the phone) provided that the survey is administered to the person directly. No one can answer for the youth, nor can data from other sources be used to answer questions. Participation in the survey is completely voluntary on the part of the youth.

WAVE 1

The overall response rate (the number of youth who completed the survey divided by the number in the baseline population) for the initial (FY2011) baseline survey was 54%. Response rates varied dramatically by state, perhaps reflecting variations in data collection procedures.

For Wave 1, the response rate is the number of youth in the Cohort divided by the number in the Baseline Population. In terms of the variables in the dataset, this would be:

$$Wave\ 1\ Response\ Rate = \frac{Baseline = 1\ and\ FY11Cohort = 1}{Baseline = 1}$$

Response Rates for NYTD Wave 1 (Age 17 in Foster Care)

FIPS	State	Baseline Population	Wave 1 Responses	Wave 1 Rate
01	Alabama	262	102	39%
02	Alaska	66	49	74%
04	Arizona	673	83	12%
05	Arkansas	262	150	57%
06	California	5,116	1,819	36%
08	Colorado	552	467	85%
09	Connecticut	469	362	77%
10	Delaware	102	71	70%
11	District of Columbia	138	92	67%
12	Florida	1,170	530	45%
13	Georgia	529	375	71%
15	Hawaii	72	31	43%
16	Idaho	67	43	64%
17	Illinois	818	551	67%
18	Indiana	664	517	78%
19	Iowa	543	472	87%
20	Kansas	563	443	79%
21	Kentucky	670	516	77%
22	Louisiana	372	342	92%
23	Maine	74	55	74%
24	Maryland	267	266	100%
25	Massachusetts	924	632	68%
26	Michigan	635	233	37%
27	Minnesota	353	252	71%
28	Mississippi	310	95	31%
29	Missouri	698	373	53%
30	Montana	85	58	68%
31	Nebraska	395	167	42%

Response Rates for NYTD Wave 1 (Age 17 in Foster Care)

FIPS	State	Baseline Population	Wave 1 Responses	Wave 1 Rate
32	Nevada	176	107	61%
33	New Hampshire	65	51	78%
34	New Jersey	419	171	41%
35	New Mexico	65	46	71%
36	New York	1,878	282	15%
37	North Carolina	584	335	57%
38	North Dakota	95	87	92%
39	Ohio	1,075	361	34%
40	Oklahoma	287	249	87%
41	Oregon	477	116	24%
42	Pennsylvania	1,249	1,022	82%
44	Rhode Island	170	170	100%
45	South Carolina	359	287	80%
46	South Dakota	71	68	96%
47	Tennessee	1,004	196	20%
48	Texas	1,563	1,227	79%
49	Utah	323	256	79%
50	Vermont	48	48	100%
51	Virginia	552	352	64%
53	Washington	456	378	83%
54	West Virginia	398	252	63%
55	Wisconsin	714	272	38%
56	Wyoming	101	42	42%
72	Puerto Rico	126	75	60%
		29,104	**15,596**	**54%**

WAVE 2

Response rates for Wave 2 – the age 19 follow-up – averaged 27%. This national rate is affected by the fact that two states – New York and Puerto Rico – did not participate in the Wave 2 survey. For Wave 2, states that elected to survey just a sample of those who responded to the Wave 1 survey attempted to contact only those youth who were in the sample . States that elected to sample are marked with an asterisk in the table below.

For Wave 2, the response rate is the number of youth that took the survey divided by the number who were eligible to take the survey. In terms of the variables in the dataset, this would be:

$$Wave\ 2\ Response\ Rate = \frac{Elig19 = 1\ and\ OutcmRpt = 1}{Baseline = 1}$$

Response Rates for Wave 2 (Age 19 Follow-up)

FIPS	State	Baseline Population	Wave 2 Responses	Wave 2 Rate
01	Alabama	262	83	32%
02	Alaska	66	46	70%
04	Arizona	673	31	5%
05	Arkansas	262	92	35%
06	California	5,116	1,239	24%
08	Colorado	552	256	46%
09	Connecticut	469	268	57%
10	Delaware	102	50	49%
11	District of Columbia	138	79	57%
12	Florida	1,170	327	28%
13	Georgia*	529	153	29%
15	Hawaii	72	26	36%
16	Idaho	67	27	40%
17	Illinois*	818	152	19%
18	Indiana*	664	169	25%
19	Iowa*	543	129	24%
20	Kansas	563	313	56%
21	Kentucky*	670	167	25%
22	Louisiana*	372	98	26%
23	Maine	74	28	38%
24	Maryland	267	200	75%
25	Massachusetts*	924	189	20%
26	Michigan	635	222	35%
27	Minnesota	353	195	55%
28	Mississippi	310	74	24%
29	Missouri	698	295	42%
30	Montana	85	44	52%
31	Nebraska	395	61	15%
32	Nevada	176	93	53%
33	New Hampshire	65	40	62%
34	New Jersey	419	110	26%
35	New Mexico	65	28	43%
36	New York	1,878	0	0%
37	North Carolina	584	221	38%
38	North Dakota	95	55	58%
39	Ohio*	1,075	142	13%
40	Oklahoma	287	163	57%
41	Oregon	477	86	18%
42	Pennsylvania*	1,249	128	10%
44	Rhode Island	170	138	81%
45	South Carolina	359	220	61%
46	South Dakota	71	61	86%
47	Tennessee*	1,004	91	9%
48	Texas*	1,563	265	17%
49	Utah	323	178	55%

Response Rates for Wave 2 (Age 19 Follow-up)

FIPS	State	Baseline Population	Wave 2 Responses	Wave 2 Rate
50	Vermont	48	32	67%
51	Virginia	552	207	38%
53	Washington*	456	187	41%
54	West Virginia	398	125	31%
55	Wisconsin	714	116	16%
56	Wyoming	101	11	11%
72	Puerto Rico	126	0	0%
		29,104	**7,710**	**26%**

ANALYTIC CONSIDERATIONS

RecNumbr is the encrypted child identifier. This ID is only guaranteed to be unique within a state, so *RecNumbr* must always be used in combination with the state ID when counting or otherwise analyzing particular children.

To facilitate working with *St-RecNumbr*s, NDACAN added the derived variable *StChID* to the dataset. *StChID* is the concatenation of *St* and *RecNumbr*. For example, when *St* = "CA" and *RecNumbr* = "123456789012", then *StChID* would be "CA123456789012".

A *StChID* is unique within a wave. A *StChID-Wave* combination is unique in the entire longitudinal file. The current file has two waves. The final file will have three.

LINKING TO OTHER FILES.

The variable *RecNumbr* is an encrypted version of the child's unique identifier used by the state agency. The ID may go by different names in the various linkable files. These are:

- NYTD Outcomes File: RecNumbr
- AFCARS Foster Care File: RecNumbr
- AFCARS Adoption File: RecNum
- NCANDS Child File: AFCARSID

The AFCARS ID is encrypted, but is done so in the same way for all these datasets, so it serves as an indicator of the same child across datasets and across years. Be careful, though. These commonalities are generally reliable, but are not applicable to all states in all years. Contact NDACAN Support for further information regarding which states can be linked across which years.

USE OF WEIGHTS.

The weights provide estimates for all AFCARS 17.

DATA FILE INFORMATION

NDACAN Dataset #182 contains data for the first two waves of outcome data for the FY2011 Cohort. The variable "Wave" distinguishes between the waves. Wave 1 (Wave=1) includes all youth in the baseline population, regardless of whether they responded to the survey. Wave 2 (Wave=2) includes only youth who were in the FY2011 cohort and were eligible for the age 19 follow-up.

SEE NYTD GUIDE TO THE DATA ELEMENTS FOR VARIABLE INFORMATION

Technical support for this dataset is provided by NDACAN.

Please send your inquiries to NDACANsupport@cornell.edu

APPENDIX:
NYTD WEIGHTING METHODOLOGY FOR THE FY 2011 COHORT AT BASELINE AND TWO-YEAR FOLLOW UP

OVERVIEW

The National Youth in Transition Database (NYTD) is a federally mandated data collection system established in response to the John H. Chafee Foster Care Independence Program (Public Law 106-169, Sec. 477 of the Social Security Act). In addition to reporting information on independent living services, states collect and report information on outcomes of youth in the child welfare system who are transitioning to adulthood. States are required to report on a cohort of youth who turn age 17 and continue data collection efforts every two years until the youth turns age 21. A new cohort of 17 year olds is identified every three years. The first cohort of 17 year olds was surveyed at baseline in federal fiscal year (FY) 2011 and surveyed again in FY 2013 at age 19.

As specified in 45 CFR 1356.81 and explained further in Technical Bulletin #5, the NYTD survey and sampling methodologies create a number of scenarios in which the population of youth surveyed can decrease over time. Youth who are eligible to take the survey but who choose not to participate or who cannot be located also reduces the size of the survey population. In both instances, response rates (the number of youth surveyed as a proportion of the number of youth eligible to take the survey) are lowered (See Figure 1). However, demographic information for the non-responders is still included in the dataset.

In FY 2011, 29,104 youth were identified as eligible to take the survey at age 17. Of those, 15,596 completed the survey, resulting in a national response rate of 54%. At the state level, response rates varied widely from 12% to 100%. In FY 2013, 11,353 youth were eligible for follow up at age 19, and 7,717 (68%) completed the survey. State-level response rates also varied widely at follow up, with two states not reporting and the remaining 50 states ranging from 26% to 95%.

The combination of response rate variation and survey design constraints has the potential to produce biased results if the respondents are significantly different from the non-respondents. In that case, results would not adequately represent

the outcomes of the population of 17 or 19 year olds for whom the survey is intended to assess. The Children's Bureau, therefore, has employed a weighting methodology with the NYTD survey responses to identify and correct potential non-response bias.

WEIGHTING

Weighting to adjust for non-response entails distributing the weight of non-respondents across respondents, so that the sum of the non-response-adjusted weights matches the total number of cases selected for the survey. The reduction of non-response bias can be improved, however, by performing this adjustment within more granular adjustment classes. In this way, respondents are weighted to specifically represent non-respondents who are similar to them in response-relevant attributes used to form weight adjustment classes.

For example, if sex is associated with response (e.g., females are disproportionately more likely to respond than males), then male respondents can be weighted to represent all selected males, and likewise for female respondents. This more granular weighting class adjustment ensures that groups that differ in response behavior are represented by members of those groups in the weighted dataset. To the extent that these differences are also related to survey outcomes, such weighting class adjustments will reduce the bias of weighted estimates.

NYTD WEIGHTING METHODOLOGY

This report outlines the methodology used to weight the first two NYTD survey data collection efforts. Wave 1, conducted in 2011, serves as the baseline year in which states conducted a census of all 17-year-old youths in the eligible population. Wave 2, conducted in 2013, was the first follow-up of Wave 1 respondents, then at age 19. Two states failed to report on any youth in Wave 2 and therefore are not included in the weighting methodology in Wave 2.

The two waves of data collection were weighted separately, so that estimates can be generated using both datasets -- the 2011 survey (youth at age 17) and the 2013 survey (age 19 follow-up). In both cases, weights are constructed so that

weighted estimates represent the full baseline population[1] (excepting the loss of two states in Wave 2).

A similar weighting plan was implemented for both waves and executed in two stages. First, a non-response adjustment was applied so that respondents to the survey in the given wave represent all youth who were eligible to take that survey (i.e., respondents and non-respondents). This weight also adjusts for selection of a sample, rather than a census, when applicable. This weight (W_1) was used to calculate the final weight included in the publically available datasets. Second, a post-stratification adjustment was applied so that key demographic distributions in the response data match those in the population, reducing coverage bias and improving face validity. The adjusted (post-stratified) weight (W_2) is the final weight included in the publically available datasets. For Wave 2, the post-stratification adjustment simultaneously adjusts the Wave 2 follow-up respondents to represent the full baseline population (rather than Wave 1 respondents only).

ADJUSTING FOR NON-RESPONSE (W_1)

When creating non-response adjustment classes for the NYTD data, each state served as the base class. Within each state, additional adjustment class dimensions were drawn from a list of potential response covariates.

For Wave 1, these were 32 variables from the FY 2011 Adoption and Foster Care Analysis Reporting System (AFCARS) that were available for both respondents and non-respondents, as well as sex, race (five levels), and Hispanic origin. See Table 4 for a list of AFCARS variables that associated with a response in Wave 1, and were significant covariates in at least one state.

For Wave 2, these were 42 Wave 1 NYTD survey outcomes, as well as sex, race, and Hispanic origin. See Table 5 for the list of Wave 1 variables that were associated with a response in Wave 2 for at least one state.

---

1 Because the baseline population included all youth that states had identified as having turned 17 years of age while in foster care in FY 2011, select FY 2011 AFCARS variables were used as adjustment variables in the weighting procedures. In some cases, however, youth record numbers did not match with the AFCARS record numbers which resulted in approximately 2% of the records not receiving weights. Weighted results represent 29,104 (98%) records in the baseline population at age 17 and 27,101 (92%) records at age 19 due to the two states that did not report follow-up information.

All potential response covariates were dichotomized: For categorical covariates, this occasionally required collapsing levels; for continuous covariates, a median split was applied. Next, missing values for the covariates were imputed using a recursive hot deck algorithm seeded with a sort list of state by sex.

Following imputation, the potential covariates were tested for association with response using (2 × 2) Pearson Chi-Square tests. Given the somewhat small sample sizes within each state, a liberal alpha level of .10 was used when testing associations for significance. In each state, up to four significant response covariates were selected, in descending order of significance, to define the most granular non-response adjustment matrix for that state. Thus, the most granular non-response adjustment matrix definition in any state would have four dichotomous dimensions, yielding a maximum of $2^4 = 16$ independent adjustment cells.

Slicing the response data at such a granular level often results in empty cells; however, each cell must contain at least one respondent to carry the weight of the non-respondents in that cell. Moreover, allowing only one respondent to represent a potentially large number of non-respondents leads to large weights that increase the weighting variance and lower the precision of weighted estimates. For this reason, a minimum of three respondents were required in every cell of an adjustment matrix for it to be used in a given state.

Specifically, in each state, the most granular non-response adjustment matrix was tested to make sure it met the minimum three-respondents-per-cell criterion. If it did not, the least-significant response covariate was dropped from the adjustment matrix definition (reducing the number of cells by a factor of 2) and the collapsed matrix was retested. This process was repeated until a suitable adjustment matrix was found, or until all response covariates were dropped, leaving only the state to define the (one-dimensional) adjustment matrix.

Table 1 summarizes the number of additional covariates (i.e., other than state itself) used to define non-response adjustment matrices across states for Wave 1 and Wave 2. Although higher numbers of covariates can improve bias reduction, sample sizes in many states were insufficient to outweigh the increase in variance that would result from weighting the data at such a granular level.

Once a suitable adjustment matrix was defined for each state, the non-response adjustment was computed as the ratio of cases selected to be surveyed in that cell to the number of responding cases in that cell, $w_1 = n_{selected}/n_{responded}$.

Table 1: Number of Response Covariates Used to Define Non-Response Adjustment Matrices in States, by Wave

	Wave 1		Wave 2	
N Covariates	Frequency	Percent	Frequency	Percent
0	7	13%	11	22%
1	14	27%	15	30%
2	21	40%	18	36%
3	5	10%	4	8%
4	5	10%	2	4%
	52	**100%**	**50**	**100%**

POST-STRATIFYING TO BASELINE POPULATION TOTALS (W_2)

Post-stratification adjustments are used to correct for undercoverage and improve the face validity of the weighted dataset by matching distributions in the data to known population distributions (i.e., control totals). For both Wave 1 and Wave 2 of the NYTD, the non-response adjustment weight w_1 was post-stratified using a raking (sample-balancing) algorithm, which iteratively adjusts w_1 to match control totals along specified dimensions. By iteratively adjusting to marginal control totals (e.g., sex, race, etc.), raking makes it possible to match multiple post-stratification dimensions at once, as opposed to simpler post-stratification ratio adjustments.

For Waves 1 and 2, the sum of w_1 in each state was post-stratified to sex and race control totals. Race was coded to have four levels (non-Hispanic white, non-Hispanic Black, non-Hispanic Other, and Hispanic); however, in some states the response data did not include cases in all four levels, which necessitated collapsing race to two levels (non-Hispanic white vs. Other).

Sex and race control totals for the state-level post-stratification were taken from the FY 2011 AFCARS dataset, which served as the population frame for the Wave 1 survey. The result is that the sum of the post-stratified weight w_2 in each state matches AFCARS totals by sex and by race (with either four or two levels).

For Wave 1, the non-response adjustment weighted back to the AFCARS frame, so that the sum of w_1 for this wave matched the population size in each state. For Wave 2, the non-response adjustment weighted back only to the number of Wave 1 respondents in a state, so that the sum of w_1 for this wave did not yet match the AFCARS population totals. By post-stratifying to AFCARS control totals, the sum of w_2 for Wave 2 is also made to match these population totals. In short, weighted estimates from Wave 1 or Wave 2 using w_2 will represent the full AFCARS population at the time of the survey.

Table 2 and Table 3 provide summary statistics for the post-stratification weight overall and by state, for Waves 1 and 2 respectively. To gauge the impact of unequal weighting effects on survey precision (variances), we computed the design effect due to weighting (DEFF) as $1 + CV^2$, where CV is the coefficient of variation of the weights. The effective sample size (n Eff) is the actual number of responses (N Obs) divided by the DEFF, which is the size of a random sample that would achieve the same level of precision.

WEIGHTED RESULTS

For Wave 1, the post-stratification adjustment introduced little additional weighting variance. For Wave 2, however, the post-stratification adjustment introduced a slightly larger increase in weighting variance at the national level. This is likely due to the fact that Wave 2 allowed states to sample rather than census the survey frame (i.e., Wave 1 respondents), introducing more variability across states when adjusting to population control totals during the post-stratification adjustment. At the state level, however, the design effect due to weighting generally remains quite low. As a result, at the national level, weighted results are not markedly different from unweighted results in terms of the distribution across categories, though it should be noted that the number of youth is larger overall. This congruence may be because non-response occurred at random (indicating no systematic bias) and/or unknown or unavailable data points were not included in the models (indicating unmeasured variance).

Table 2. *Summary Statistics for Wave 1 Weights*

Post-Stratification Weights (w_2) for Wave 1								
State	N Obs	Min	Mean	Max	CV	DEFF	n Eff	Sum
Overall	**15,597**	**0.70**	**1.87**	**18.65**	**0.69**	**1.47**	**10,601**	**29,106**
AK	49	0.76	1.35	3.31	0.35	1.13	44	66
AL	102	2.44	2.57	4.11	0.08	1.01	101	262
AR	150	1.15	1.75	3.88	0.24	1.06	142	262
AZ	83	1.88	8.11	18.65	0.50	1.25	66	673
CA	1,819	1.90	2.81	6.12	0.27	1.07	1,698	5,116
CO	467	1.01	1.18	1.35	0.07	1.01	465	552
CT	362	1.17	1.30	2.68	0.18	1.03	350	469
DC	92	1.00	1.50	2.99	0.24	1.06	87	138
DE	71	0.94	1.44	2.49	0.29	1.09	65	102
FL	530	1.59	2.21	2.85	0.15	1.02	519	1,170
GA	375	1.18	1.41	2.55	0.13	1.02	369	529
HI	31	1.00	2.32	3.99	0.35	1.12	28	72
IA	472	0.92	1.15	1.51	0.10	1.01	468	543
ID	43	1.16	1.56	2.95	0.25	1.06	40	67
IL	552	1.37	1.48	2.00	0.05	1.00	550	819
IN	517	0.80	1.28	2.36	0.22	1.05	493	664
KS	443	0.95	1.27	2.28	0.16	1.03	432	563
KY	516	1.11	1.30	1.87	0.11	1.01	510	670
LA	342	0.99	1.09	1.12	0.02	1.00	342	372
MA	632	1.08	1.46	2.34	0.18	1.03	612	924
MD	266	1.00	1.00	1.01	0.00	1.00	266	267
ME	55	0.79	1.35	1.87	0.27	1.08	51	74
MI	233	1.88	2.73	4.29	0.20	1.04	224	635
MN	252	1.02	1.40	3.39	0.30	1.09	231	353
MO	373	1.14	1.87	4.35	0.20	1.04	359	698
MS	95	1.69	3.26	7.33	0.47	1.22	78	310
MT	58	0.95	1.47	1.98	0.18	1.03	56	85
NC	335	1.43	1.74	2.83	0.18	1.03	324	584
ND	87	0.98	1.09	1.36	0.12	1.01	86	95
NE	167	1.57	2.37	5.85	0.29	1.09	154	395
NH	51	0.95	1.27	2.43	0.23	1.05	48	65
NJ	171	1.36	2.45	4.92	0.33	1.11	154	419
NM	46	0.70	1.41	3.56	0.33	1.11	41	65
NV	107	0.80	1.64	3.36	0.28	1.08	99	176
NY	282	3.57	6.66	11.43	0.31	1.10	257	1,878

Post-Stratification Weights (w_2) for Wave 1								
State	N Obs	Min	Mean	Max	CV	DEFF	n Eff	Sum
OH	361	1.69	2.98	14.57	0.57	1.33	272	1,075
OK	249	1.10	1.15	2.85	0.15	1.02	243	287
OR	116	2.17	4.11	13.17	0.40	1.16	100	477
PA	1,022	1.07	1.22	1.49	0.07	1.01	1,017	1,249
PR	75	1.00	1.68	2.05	0.20	1.04	72	126
RI	170	1.00	1.00	1.00	0.00	1.00	170	170
SC	287	1.18	1.25	1.35	0.05	1.00	286	359
SD	68	0.87	1.04	1.47	0.10	1.01	67	71
TN	196	1.69	5.12	8.77	0.38	1.15	171	1,004
TX	1,227	1.08	1.27	2.12	0.12	1.01	1,210	1,563
UT	256	1.09	1.26	2.64	0.20	1.04	246	323
VA	352	1.36	1.57	2.94	0.17	1.03	342	552
VT	48	1.00	1.00	1.00	0.00	1.00	48	48
WA	378	1.02	1.21	2.00	0.17	1.03	367	456
WI	272	2.11	2.63	4.16	0.14	1.02	267	714
WV	252	0.81	1.58	2.24	0.14	1.02	247	398
WY	42	1.79	2.40	6.43	0.36	1.13	37	101

Table 3. *Summary Statistics for Wave 2 Weights*

Table 3: Post-Stratification Weights (w_2) for Wave 2								
State	**N Obs**	**Min**	**Mean**	**Max**	**CV**	**DEFF**	**n Eff**	**Sum**
Overall	**7,591**	**0.85**	**3.57**	**80.71**	**0.77**	**1.59**	**4,762**	**27,101**
AK	46	1.16	1.43	2.00	0.15	1.02	45	66
AL	83	2.50	3.16	8.00	0.23	1.05	79	262
AR	85	1.07	3.08	5.37	0.33	1.11	77	262
AZ	31	5.74	21.71	80.71	0.69	1.47	21	673
CA	1,175	2.69	4.35	9.51	0.29	1.08	1,083	5,116
CO	256	1.21	2.16	4.09	0.31	1.10	233	552
CT	268	1.15	1.75	3.43	0.23	1.05	255	469
DC	78	1.00	1.77	4.44	0.31	1.09	71	138
DE	50	1.67	2.04	2.49	0.18	1.03	48	102
FL	327	2.09	3.58	4.51	0.15	1.02	320	1,170
GA	153	3.00	3.46	4.71	0.09	1.01	152	529
HI	26	1.84	2.77	4.39	0.25	1.06	24	72
IA	129	2.94	4.21	12.23	0.41	1.17	111	543
ID	27	1.03	2.48	4.29	0.37	1.14	24	67
IL	133	4.00	6.16	11.16	0.27	1.07	124	819
IN	169	2.20	3.93	5.25	0.24	1.06	160	664
KS	313	1.37	1.80	3.38	0.16	1.03	305	563
KY	167	2.44	4.01	9.36	0.27	1.07	156	670
LA	84	2.16	4.43	24.62	0.70	1.49	56	372
MA	189	2.16	4.89	12.67	0.25	1.06	178	924
MD	200	0.95	1.34	3.36	0.22	1.05	191	267
ME	28	1.50	2.64	5.00	0.22	1.05	27	74
MI	222	2.24	2.86	4.98	0.17	1.03	216	635
MN	193	1.12	1.83	3.57	0.24	1.06	183	353
MO	295	1.04	2.37	3.23	0.15	1.02	288	698
MS	74	3.48	4.19	6.76	0.20	1.04	71	310
MT	44	0.85	1.93	3.02	0.26	1.07	41	85
NC	221	2.19	2.64	4.15	0.16	1.03	215	584
ND	49	1.76	1.94	2.39	0.12	1.01	48	95
NE	52	3.57	7.60	25.67	0.55	1.30	40	395
NH	40	1.11	1.63	7.00	0.58	1.34	30	65
NJ	110	2.62	3.81	13.00	0.37	1.14	97	419
NM	28	1.33	2.32	5.00	0.39	1.15	24	65
NV	93	0.92	1.89	4.15	0.26	1.07	87	176
NY								
OH	190	2.89	5.66	8.06	0.27	1.07	177	1,075

Table 3: Post-Stratification Weights (w_2) for Wave 2								
State	**N Obs**	**Min**	**Mean**	**Max**	**CV**	**DEFF**	**n Eff**	**Sum**
OK	163	1.05	1.76	2.90	0.20	1.04	157	287
OR	86	3.22	5.55	10.47	0.32	1.11	78	477
PA	135	3.89	9.25	14.05	0.39	1.15	117	1,249
PR								
RI	138	1.15	1.23	4.00	0.20	1.04	133	170
SC	220	1.09	1.63	2.23	0.15	1.02	215	359
SD	61	0.88	1.16	1.53	0.13	1.02	60	71
TN	71	3.90	14.14	47.00	0.43	1.18	60	1,004
TX	265	5.19	5.90	9.11	0.14	1.02	260	1,563
UT	178	1.19	1.81	3.67	0.25	1.06	167	323
VA	175	2.23	3.15	8.44	0.29	1.09	161	552
VT	32	0.90	1.50	2.20	0.30	1.09	29	48
WA	187	2.18	2.44	2.87	0.08	1.01	186	456
WI	116	4.34	6.16	10.28	0.21	1.04	111	714
WV	125	2.10	3.18	4.24	0.24	1.06	118	398
WY	11	6.60	9.18	29.00	0.72	1.52	7	101

Note: New York and Puerto Rico dropped out of NYTD data collection after Wave 1

Table 4. AFCARS Elements used in the Wave 1Weighting Models

AFCARS Element	Variable Label	N States
StFIPS	State FIPS Code	52
AgeAtEntry	Age at Entry of most recent episode	6
Sex	Child Gender	5
HisOrgin	Child Hispanic Origin	3
AmIAKN	Child Race: American Indian or Alaska Native	2
BlkAfrAm	Child Race: Black / African American	2
White	Child Race: White	7
ClinDis	Diagnosed Disability	2
DSMiii	Disability: Emotionally Disturbed	4
MR	Disability: Mental Retardation	5
PhyDis	Disability: Physically Disabled	1
VisHear	Disability: Visually Hearing Impaired	1
NoCope	Removal Reason - Caretaker Inability Cope	2
ChBehPrb	Removal Reason - Child Behavior Problem	6
ChilDis	Removal Reason - Child Disability	1
DAChild	Removal Reason - Drug Abuse Child	1
Housing	Removal Reason - Inadequate Housing	2
Neglect	Removal Reason - Neglect	4
PhyAbuse	Removal Reason - Physical Abuse	2
CaseGoal	Most Recent Case Plan Goal	8
CurPlSet	Current Placement Setting	11
NumPlep	Number of Placement Settings (Current FC Episode)	3
TotalRem	Total Number of Removals	2
LOSAtEnd	Length of stay (in days) for Children in care at end of FFY	3
LOSAtExit	Length of stay (in days) for Exits	2
LOSEndMonths	Length of stay (in months) for Children in care at end of FFY	2
SettingLOS	Length of Time in Current Setting (in months)	4

***Table 5**. NYTD Elements used in the Wave 2 Weighting Models*

NYTD Element	Variable Label	N States
StFIPS	State FIPS Code	50
SubAbuse	Outcome: Substance Abuse Referral	11
Incarc	Outcome: Ever Incarcerated	10
Sex	Gender	5
White	Child Race: White	4
DelinqntSv	Services: Youth has ever been adjudicated as a delinquent	4
SocSecrty	Outcome: Youth receiving Social Security	3
HlthEdSv	Services: Health Education And Risk Prevention	3
HisOrgin	Child Hispanic Origin	2
BlkAfrAm	Child Race: Black / African American	2
OthrHlthIn	Outcome: Youth has health insurance other than Medicaid	2
EmplySklls	Outcome: Youth has obtained employment-related skills	2
Medicaid	Outcome: Youth is receiving Medicaid	2
AcSuppSv	Services: Academic Support	2
CareerSv	Services: Career Preparation	2
HousEdSv	Services: Housing Education And Home Management Training	2
OthrFinAs	Services: Other Financial Assistance	2
PSEdSuppSv	Services: Post-Secondary Educational Support	2
CurrEnroll	Outcome: Current enrollment and attendance	1
CurrPTE	Outcome: Current Part-Time Employment	1
Homeless	Outcome: Ever been Homeless	1
PubFoodAs	Outcome: Public Food Assistance	1
BudgetSv	Services: Budget And Financial Management	1
EmplyTrSv	Services: Employment Programs Or Vocational Training	1
FamSuppSv	Services: Family Support And Healthy Marriage Education	1
ILNASv	Services: Independent Living Needs Assessment	1
MentorSv	Services: Mentoring	1
OthrFinaSv	Services: Other Financial Assistance	1
SpecEdSv	Services: Special Education	1