Fragile Families
Census Tract Measures
Restricted Use Data Appendage

Baseline, 1, 3, 5, 9, and 15 year Waves

January 2022

Bendheim-Thoman
Center for Research on Child Wellbeing
Wallace Hall
Princeton University
Princeton, NJ 08544
http://crcw.princeton.edu

Columbia Population Research Center
1255 Amsterdam Avenue, Room 715
Columbia University
New York, NY 10027
http://cupop.columbia.edu/

DATA APPENDAGE OVERVIEW

The Fragile Families Census Tract Measures Restricted Use Data Appendage (ffgeo5_all_pub.dta) contains Census tract-level contextual data that correspond to the residence of respondents at Baseline and during the 1, 3, 5, 9, and 15-year follow-up waves of the Fragile Families and Child Wellbeing Study (FFCWS). Variables consist of aggregate characteristics on individuals, families, households, and housing units within Census tracts and identifiers for the pseudo-tract of the residence. Contextual data at Baseline through Year 15 correspond to the residence of the biological mother and biological father at the year(s) in which they were interviewed. Contextual data for Year 15 correspond to the residence of the focal child's primary caregiver (PCG; biological mother, biological father, other family, or other non-family caregivers) at the time of the Year 15 interview.

In January 2022, variable *tp6tract_cen00* was added to the most recent data appendage to allow researchers to convert Year 15 pseudo tracts to 2000 census tract boundaries, harmonizing these data with prior waves.

FILE LAYOUT

The file contains 4,898 observations (one per family), sorted by *idnum*.

VARIABLE NAMING CONVENTION

Contextual variable names are 7-14 characters long. The first 3 characters contain the variable prefix. The next 4-5 characters contain the variable suffix, which includes an abbreviated description of the contextual data. The last 4 characters for the Year 15 data, indicate if the variable was constructed from 2015 ACS 5-year estimates, the 2010 Census or the 2000 Census. The variable names are constructed as follows:

Position	Character	Indicates
1	t	Census tract-level measure (all variables in appendage)
2	m	Mother residence
	f	Father residence
	p	Primary Caregiver (PCG) residence (15-year)
3	1	Baseline Interview
	2	First follow-up interview (1-year)
	3	Second follow-up interview (3-year)
	4	Third follow-up interview (5-year)
	5	Fourth follow-up interview (9-year)
	6	Fifth follow-up interview (15-year)
4-8	**** (e.g., phisp)	Contextual data (e.g., Percent of population Hispanic)
9-14	_cen00	Data source – 2000 Decennial Census
	_cen10	Data source – 2010 Decennial Census
	_acs15	Data source – American Community Survey, 2011-2015 5-year estimates

For example, variable tp6pfrgn_acs15: the prefix (position 1-3) tp6 refers to a Census tract-level measure (t) corresponding to the PCG's residence (p) at the sixth follow-up interview (15- year) (6); the suffix (pfrgn_acs15) refers to percent of population foreign born, as found in the ACS 2015 5-year estimates data.

GEOGRAPHY

U.S. Census tracts

All contextual data are reported at the Census tract level and correspond to the current residence of respondents at each year. Census tracts are statistical subdivisions of a county, averaging about 4,000 inhabitants, and designed by the U.S. Census Bureau to be relatively permanent over time. All Census tracts that appear in Baseline through Year 9 use 2000 Census tract boundary definitions; tracts in Year 15 use the 2010 Census tract boundary definitions. The variable *tp6tract_cen00* allows researchers harmonize Year 15 pseudo tracts with 2000 census tract boundaries. For more information, see: https://www.census.gov/geo/.

Geo-coding process

The primary residential addresses reported by focal children's biological mothers and fathers or PCG at each wave were used to geo-code respondents into U.S. Census tracts according to 2000 Census tract boundary definitions for Baseline to Year 9, and U.S. Census tracts according to the 2010 Census tract definitions for Year 15. To create the variable *tp6tract_cen00,* Fragile Families staff used the Longitudinal Tract Database (https://s4.ad.brown.edu/Projects/Diversity/Researcher/Bridging.htm) to code Census tracts from Year 15 with 2000 Census tracts boundaries.

Pseudo Census tract identifiers

To protect the privacy of respondents, the actual state, county, and Census tract of residence are not included in the current data appendage. However, it includes pseudo Census tract identifier variables (t*tract). These variables do not represent actual U.S. Census tracts, but are used to indicate which respondents live in the same Census tract. The pseudo Census tract identifiers are consistent between mothers and fathers, and across waves and contextual data appendages for Baseline through Year 15. Therefore, they can be used as an indicator of whether the focal children's mothers and fathers resided in the same Census tract from one survey wave to the next.

The 2010 pseudo tract identifiers (tp6tract_cen10) are not consistent with those in previous waves. Boundaries of tracts are subject to change between Census years, and the Census reports tracts where more than 2.5% of the area changed as "substantially changed" tracts. A list of substantially changed tracts may be obtained at https://www.census.gov/geo/maps-data/data/tract_rel_download.html. Around 10-15% of the Fragile Families sample resides in these "substantially changed" tracts. In January 2022, variable *tp6tract_cen00* was added to the most recent data appendage to allow researchers to convert Year 15 pseudo tracts to 2000 census tract boundaries, harmonizing these data with prior waves.

DATA SOURCES

U.S. Census Bureau

Contextual data in Baseline through Year 9 included in the current data appendage was obtained from the U.S. 2000 Decennial Census. Data was downloaded from http://factfinder2.census.gov.

Contextual data for Year 15included in the current data appendage was obtained from the U.S. 2010 Decennial Census as well as the American Community Survey 2015 5-year estimates.

American Community Survey

In addition to using data from the 2010 Decennial Census, the contextual data for Year 15 includes data from the American Community Survey (ACS) 2015 5-year estimates. The ACS is an annual survey conducted by the Census Bureau, sampling about 3.5 million housing unit addresses in every county in the nation. It is designed to provide more up-to-date information about the population at a local level between each decennial census. The 5-year estimates used in Year 15 are based on data gathered between January 1, 2011 and December 31, 2015. The 5-year estimates are preferable to the one-year ACS estimates because they give a more precise measure for small geographies, like Census tracts.

Data from the ACS is used in place of some Census data due to the change in Census data reporting. From 1970-2000, the Census had both a long- and short-form questionnaire. Every household received the short-form, and a sample received the long-form questionnaire. The long-form questionnaire included questions about educational attainment, income, employment, wealth, and country of origin. The 2010 Census did away with the long-form questionnaire, opting instead for a shorter, 10-question survey sent to all households. Since 2010, the ACS has gathered the data previously reported in the Census long-form questionnaire.

Variables found in both the 2010 Census and the 2015 ACS are both included in the current appendage. Those variables found in the ACS include the suffix "_acs15", and the identifier "ACS 2015 5-year estimates" in the variable label.

For the 2015 ACS poverty level variables (p50pl-p200p), each level was reported by the ACS in age groups (i.e. under 6, 6 to 11, 12 to 17, etc.), while that same data was simply reported at each ratio level in the 2000 Census. To get a percentage of the total population at each poverty level, each age group was summed at each poverty level and divided by the total population of each tract.

The 2015 ACS variable for median rent as a percent of income (mrphi) had values of 10.0- and 50.0+, which were recoded to 9.9 and 50.1 respectively.

Additional information about the ACS and multi-year estimates is available at https://www.census.gov/programs-surveys/acs.html. Data was downloaded from http://factfinder2.census.gov.

Random "noise"

Some random noise has been introduced into the data to ensure that respondents' Census tracts cannot be identified on the basis of the characteristics provided in this file. This noise should have no impact on analyses.

MISSING VALUES
All missing values for contextual variables are coded using similar conventions found in other FFCWS data files. More specifically, the following three missing codes are used:

-9 Not in wave	The respondent did not participate in interview/assessment
-7 *N/A*	The respondent participated in interview/assessment; however, the respondent's residence could not be determined.
-3 *Missing*	The respondent participated in interview/assessment and the residence was determined; however, contextual data is unavailable.

ACKNOWLEDGEMENTS

DATA DICTIONARY

Note: Variable names listed in the data dictionary exclude the 4-character length prefixes. Please refer to the *Variable Naming Convention* section for more information about variable prefixes.

Variable	Description	U.S. Census 2000[1]	U.S. Census 2010	ACS 2015 5-year Estimates[2]
phisp	Percent of population Hispanic	X	X	X
pwhte	Percent of population non-Hispanic White	X	X	X
pblck	Percent of population non-Hispanic Black	X	X	X
pamin	Percent of population non-Hispanic American Indian/Native American	X	X	X
pasia	Percent of population non-Hispanic Asian	X	X	X
ppaci	Percent of population non-Hispanic Native Hawaiian/Pacific Islander	X	X	X
pothr	Percent of population non-Hispanic other race	X	X	X
pfrgn	Percent foreign born	X		X
pchbr	Percent of female population of childbearing age (15-49)	X	X	X
pfhhr	Percent of family households with kids <18 headed by females	X	X	X
mpphh	Mean number of persons per household	X	X	X
psch	Percent of population 16-19 enrolled in school	X		X
p25hs	Percent of 25+ population with HS+ education	X		X
p25b	Percent of 25+ population with bachelor's or higher	X		X
pemp	Percent of civilian labor force (16+) employed	X	X	X

[1] Data from the U.S. Decennial Census 2000 used in Baseline to Year 9.
[2] Data from the U.S. Decennial Census 2010 and ACS 2015 5-year Estimates used in Year 15.

Variable	Description	U.S. Census 2000	U.S. Census 2010	ACS 2015 5-year Estimates
puemp	Percent of civilian labor force (16+) unemployed	X	X	X
pvach	Percent of housing units vacant	X	X	X
pbf40	Percent of housing units built before 1940	X		X
prent	Percent of occupied housing units renter-occupied	X	X	X
pnoph	Percent of occupied housing units without phone service	X		X
pplum	Percent of housing units lacking complete plumbing	X		X
rphi	Median gross rent as percent of HH income	X		X
mrent	Median monthly gross rent	X		X
mhval	Median housing value	X		X
myear	Median year built	X		X
ppuba	Percent of households on public assistance	X		X
pfbpl	Percent of families below poverty level	X		X
pb10k	Percent of families with income <$10K	X		X
p1014	Percent of families with income $10K to $14,999	X		X
p1524	Percent of families with income $15K to $24,999	X		X
p2534	Percent of families with income $25K to $34,999	X		X
p3549	Percent of families with income $35K to $49,999	X		X
p5074	Percent of families with income $50K to $74,999	X		X
p7599	Percent of families with income $75K to $99,999	X		X

Variable	Description	U.S. Census 2000	U.S. Census 2010	ACS 2015 5-year Estimates
p100k	Percent of families with income $100K to $149,999	X		X
p150k	Percent of families with income $150K+	X		X
mhinc	Median HH income	X		X
p50pl	Percent of persons with income less than .50 of poverty line	X		X
p74pl	Percent of persons with income .50 to .74 of poverty line	X		X
p99pl	Percent of persons with income .75 to .99 of poverty line	X		X
p124pl	Percent of persons with income 1.00 to 1.24 of poverty line	X		X
p149p	Percent of persons with income 1.25 to 1.49 of poverty line	X		X
p174p	Percent of persons with income 1.50 to 1.74 of poverty line	X		X
p184p	Percent of persons with income 1.75 to 1.84 of poverty line	X		X
p199p	Percent of persons with income 1.85 to 1.99 of poverty line	X		X
p200p	Percent of persons with income 2.00 & over poverty line	X		X