

**Fragile Families  
Census Tract Measures  
Restricted Use Data Appendage**

5-year and 9-year Follow-Up Waves

February 2015

***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/>

Prepared by Louis Donnelly at the Bendheim-Thoman Center for Research on Child Wellbeing (CRCW), Princeton University. For more information about Fragile Families, please visit our web site at <http://fragilefamilies.princeton.edu/> or email [ffdata@princeton.edu](mailto:ffdata@princeton.edu)

## DATA APPENDAGE OVERVIEW

The Fragile Families Census Tract Measures Restricted Use Data Appendage (ffgeo2\_all\_pub1.dta) contains Census tract-level contextual data that correspond to the residence of respondents during the 5-year and 9-year follow-up waves of the Fragile Families and Child Wellbeing Study (FFCWS). The current data appendage is an update to the previously released data appendage (ffgeo\_all\_pub1.dta), which contained the same contextual data from prior survey waves. The contextual variables included in the current data appendage are identical to the contextual variables that appeared in the previous data appendage. Variables consist of aggregate characteristics on individuals, families, households, and housing units within Census tracts. All contextual data correspond to the current residence of the biological mother and biological father at the wave(s) in which they were interviewed.

## FILE LAYOUT

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

## VARIABLE NAMING CONVENTION

Contextual variable names are 7-8 characters long. The first 3 characters contain the variable prefix. The latter 4-5 characters contain the variable suffix, which includes an abbreviated description of the contextual data. The variable names are constructed as follows:

<u>Position</u>	<u>Character</u>	<u>Indicates</u>
1	t	Census tract-level measure (all variables in appendage)
2	m	Mother residence
	f	Father residence
3	4	Fourth follow-up interview (5-year)
	5	Fifth follow-up interview (9-year)
4-8	**** (e.g., phisp)	Contextual data (e.g., Percent of population Hispanic)

For example, variable tm4phisp: the prefix (position 1-3) tm4 refers to a Census tract-level measure (t) corresponding to the mother's residence (m) at the fourth follow-up interview (5-year) (4); the suffix phisp (position 4-8) refers to percent of population Hispanic.

## 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 wave. 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 the current data appendage use 2000 Census tract boundary definitions. 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 at 5-year and 9-year follow-up waves were used to geo-code respondents into U.S. Census tracts according to 2000 Census tract boundary definitions. The geo-coding process was completed by Daniel M. Sheehan as part of the Built Environment and Health Project at The Institute for Social and Economic Research and Policy, Columbia University.

### ***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, like the previous data appendage (ffgeo\_all\_pub1.dta), the current data appendage includes pseudo Census tract identifier variables (t\*tract). These variables do not represent actual U.S. Census tracts, but are provided 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. The same pseudo Census tract identifiers are used in the previous data appendage (ffgeo\_all\_pub1.dta). 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.

## **DATA SOURCES**

### ***U.S. Census Bureau***

All contextual data included in the current data appendage was obtained from the U.S. 2000 Decennial Census. Data was downloaded on December, 2014 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 <i>Not in wave</i> | The respondent did not participate in interview/assessment   |
| -7 <i>N/A</i>         | The respondent participated in interview/assessment; however, the respondent’s residence could not be determined.              |
| -3 <i>Missing</i>     | The respondent participated in interview/assessment and the residence was determined; however, contextual data is unavailable. |

## **ACKNOWLEDGEMENTS**

This work was supported by Mary Clare Lennon and Anthony Buttaro at City University of New York, and the Built Environment and Health Project at Columbia University.

## 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.

---

<b>Variable</b>	<b>Source</b>	<b>Description</b>
phisp	2000 Census	Percent of population Hispanic
pwhite	2000 Census	Percent of population non-Hispanic White
pblck	2000 Census	Percent of population non-Hispanic Black
pamin	2000 Census	Percent of population non-Hispanic American Indian/Native American
pasia	2000 Census	Percent of population non-Hispanic Asian
ppaci	2000 Census	Percent of population non-Hispanic Native Hawaiian/Pacific Islander
pothr	2000 Census	Percent of population non-Hispanic other race
pfrgn	2000 Census	Percent foreign born
pchbr	2000 Census	Percent of female population of childbearing age (15-49)
pfhhr	2000 Census	Percent of family households with kids <18 headed by females
mpphh	2000 Census	Mean number of persons per household
psch	2000 Census	Percent of pop 16-19 enrolled in school
p25hs	2000 Census	Percent of 25+ population with HS+ educ
p25b	2000 Census	Percent of 25+ population with bachelor's or higher
pemp	2000 Census	Percent of civilian labor force (16+) employed

---

<b>Variable</b>	<b>Source</b>	<b>Description</b>
puemp	2000 Census	Percent of civilian labor force (16+) unemployed
pvach	2000 Census	Percent of housing units vacant
pbf40	2000 Census	Percent of housing units built before 1940
prent	2000 Census	Percent of occupied housing units renter-occupied
pnoph	2000 Census	Percent of occupied housing units without phone service
pplum	2000 Census	Percent of housing units lacking complete plumbing
rphi	2000 Census	Median gross rent as percent of HH income, 1999
mrent	2000 Census	Median monthly gross rent, 1999
mhval	2000 Census	Median housing value, 1999
myear	2000 Census	Median year housing structure built
ppuba	2000 Census	Percent of households on public assistance
pfbpl	2000 Census	Percent of families below poverty level in 1999
pb10k	2000 Census	Percent of families with 1999 income <\$10K
p1014	2000 Census	Percent of families with 1999 income \$10K to \$14,999
p1524	2000 Census	Percent of families with 1999 income \$15K to \$24,999
p2534	2000 Census	Percent of families with 1999 income \$25K to \$34,999
p3549	2000 Census	Percent of families with 1999 income \$35K to \$49,999

<b>Variable</b>	<b>Source</b>	<b>Description</b>
p5074	2000 Census	Percent of families with 1999 income \$50K to \$74,999
p7599	2000 Census	Percent of families with 1999 income \$75K to \$99,999
p100k	2000 Census	Percent of families with 1999 income \$100K to \$149,999
p150k	2000 Census	Percent of families with 1999 income \$150K+
mhinc	2000 Census	Median HH income in 1999
p50pl	2000 Census	Percent of persons with income less than .50 of poverty line
p74pl	2000 Census	Percent of persons with income .50 to .74 of poverty line
p99pl	2000 Census	Percent of persons with income .75 to .99 of poverty line
p124p	2000 Census	Percent of persons with income 1.00 to 1.24 of poverty line
p149p	2000 Census	Percent of persons with income 1.25 to 1.49 of poverty line
p174p	2000 Census	Percent of persons with income 1.50 to 1.74 of poverty line
p184p	2000 Census	Percent of persons with income 1.75 to 1.84 of poverty line
p199p	2000 Census	Percent of persons with income 1.85 to 1.99 of poverty line
p200p	2000 Census	Percent of persons with income 2.00 & over poverty line