Study Timeline

- Year 1: 1998
- Year 3: 2000
- Year 5: 2002
- Year 9: 2006
- Year 15: 2010
Accessing Public Data

♦ Includes nearly all data collected during course of study

♦ To apply/download:
  ♦ Register at OPR Data archive: opr.princeton.edu/archive

♦ Submit short application and wait for approval.
  ♦ Receive link to download data by e-mail
# Public Data

## File Structure and Components

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Year 1</th>
<th>Year 3</th>
<th>Year 5</th>
<th>Year 9</th>
<th>Year 15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother core survey</td>
<td>Baseline File</td>
<td>Y1 File</td>
<td>Y3 File</td>
<td>Y5 File</td>
<td>Y9 File</td>
<td></td>
</tr>
<tr>
<td>Father core survey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Childcare provider / Teacher survey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Home assessments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interviewer observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary caregiver survey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child/teen survey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Accessing Contract Data

✧ Two step application

- Preliminary application - send a draft of your extended abstract, description of data requested, and CV to ffdata@princeton.edu.

- Full application – includes a data protection plan, contract data agreement, IRB approval, etc.

For more info on Contract Data contents and the application process, see https://fragilefamilies.princeton.edu/restricted).
Getting started with data

- Review survey documentation
  - Read the User Guides for the waves you plan to use
  - Browse through questionnaires

- Browse or search on the Metadata Explorer

- Read Fragile Families Fact Sheet

- Review FAQ

- Sign up for data alerts and ListServ

- E-mail ffdata@princeton.edu with questions
# Variable prefixes for survey questions

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Year 1</th>
<th>Year 3</th>
<th>Year 5</th>
<th>Year 9</th>
<th>Year 15</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mother core</strong></td>
<td>m1</td>
<td>m2</td>
<td>m3</td>
<td>m4</td>
<td>m5</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Father core</strong></td>
<td>f1</td>
<td>f2</td>
<td>f3</td>
<td>f4</td>
<td>f5</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Home assessments</strong></td>
<td>NA</td>
<td>NA</td>
<td>h3*</td>
<td>h4*</td>
<td>h5*</td>
<td>h6</td>
</tr>
<tr>
<td><strong>Interviewer observations</strong></td>
<td>NA</td>
<td>NA</td>
<td>o3*</td>
<td>o4*</td>
<td>o5</td>
<td>o6</td>
</tr>
<tr>
<td><strong>Primary caregiver survey</strong></td>
<td>NA</td>
<td>NA</td>
<td>p3*</td>
<td>p4*</td>
<td>p5</td>
<td>p6</td>
</tr>
<tr>
<td><strong>Child/teen survey</strong></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>k5</td>
<td>k6</td>
</tr>
<tr>
<td><strong>Childcare center survey/obs.</strong></td>
<td>NA</td>
<td>NA</td>
<td>d3/e3*</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Family care survey/obs.</strong></td>
<td>NA</td>
<td>NA</td>
<td>r3/s3*</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Teacher survey</strong></td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>t4*</td>
<td>t5</td>
<td>NA</td>
</tr>
</tbody>
</table>

* Changed in Fall 2018 re-release of data files
Constructed Variables

- Constructed variables start with letter c, then the correct prefix from the previous slide (e.g., cm1, cf5)
- May be constructed using survey variables (e.g., meets CIDI depression criteria) or based on data not available from survey variables (e.g., interviewed conducted in Spanish)

Commonly used baseline constructed variables:
- cm1bsex – focal child sex
- cm1lbw – focal child low birth weight
- nm1numb – focal child twin/not twin
- c*1ethrace – mother/father race/ethnicity categories
- c*1edu – mother/father highest education categories
Commonly used constructed variables at follow-up waves

- `cm*relf` – mother relationship with father at * wave
- `c**age` – mother/father age at * wave core interview
- `c**edu` – mother/father highest education categories
- `c**hhinc` – mother/father household income
- `c**in pov` – mother/father household income / poverty threshold
- `c****jail` – father current/ever in jail at * wave
  - Based on various reports and measurement techniques
- `cf1span` – mother father interview conducted in Spanish
- `c**kid / c**adult` – number of children/adults in household
- `c***int` – mother/father participated in * wave core interview
- `**intmon / **intyr` – month/year of interview
Missing Data Codes

- **-9 Not in wave** – *Did not participate in survey/data collection component*
- **-8 Out of range** – *Response not possible; rarely used*
- **-7 Not applicable** (also -10/-14) – *Rarely used for survey questions*
- **-6 Skipped** “Valid skip” – *Intentionally not asked question; question does not apply to respondent or response known based on prior information.*
- **-5 Not asked** “Invalid skip” – *Respondent not asked question in the version of the survey they received* (present in two “pilot” cities).
- **-3 Missing** – *Data is missing due to some other reason*
- **-2 Don’t know** – *Respondent asked question; Responded “Don’t Know”.*
- **-1 Refuse** – *Respondent asked question; Refused to answer question.*
Weights for Core Interviews

◊ “National” weights make data representative of:
  ◊ Births occurring in large U.S. cities (the 77 U.S. cities with populations over 200,000 in 1994) between 1998 and 2000.
  ◊ Applicable to: 16 randomly selected cities (73% of data at baseline)
  ◊ Variables: *natwt (e.g., m1natwt, f2natwt)
    ◊ *natwtx excludes one pilot city – some questions different

◊ “City” weights make data representative of:
  ◊ Applicable to: All 20 cities (100% of data at baseline)
  ◊ Variables: *citywt (e.g., m1citywt, f2citywt)
Note about Weights and FFCWS

◊ How and whether to use weights is a methodological decision of individual data users
 ◊ Decision influenced by many factors, including academic discipline, individual preferences, research questions/objectives

◊ Common approach for many FF data users and PIs:
 ◊ Present weighted descriptive (univariate) statistics
 ◊ For efficiency, conduct unweighted multivariate analyses
    ◊ Include controls for important baseline characteristics related to sampling design (e.g., mothers’ marital status and age)
Learn More about the FFCWS

Publications Database
https://ffpubs.princeton.edu/
- Text search of title and abstract and/or search by author
- Filter by publication type and/or keyword
- Contains journal articles, books, book chapters, and working papers

Documentation Page
http://fragilefamilies.princeton.edu/documentation
- Questionnaires
- User Guides
- Weights documentation and design paper
- Codebooks (tabulations of all variables)
Important FFCWS Links

Study website:
www.fragilefamilies.princeton.edu

Download FFCWS data:
www.opr.princeton.edu/archive/ff/

Search FFCWS Metadata Explorer:
http://metadata.fragilefamilies.princeton.edu/

Email questions to:
ffdata@princeton.edu

Follow us on twitter @FFCWS