Weights

What's delivered with the original data and what to do about weighting linked data

CPS Weights

- There are several
- Choice of which weight to use depends on which data you are using
- With one exception, all weights provided by CPS are cross-sectional weights

Weights

IPUMS name	Original name	Purpose
WTFINL	PWSSWGT	Use for most analyses of basic monthly data
COMPWT	PWCMPWGT	Use to replicate published BLS labor force estimates
EARNWT	PWORWGT	Use for outgoing rotation group analyses
PANLWT	PWLGWWT	Use for gross flows analysis

Supplement Weights

- Many supplements have a weight that you should use if you are using data from that supplement
  - Noted on "sample notes" pages
- Not every supplement universe is "All Persons" and there is supplement non-response
- IPUMS supplement weight naming convention
  - [SUPP STEM]SUPPWT (e.g., VOSUPPWT for voting)
  - Exception: ASECWT

Weights for Linked Data

- PANLWT is the closest…
  - Used for gross flows analysis of adjacent months of data
  - Weight on month X applies to flows between month X-1 and month X
  - Reference: “Estimating gross flows consistent with stocks in the CPS” by Harley J. Frazis, Edwin L. Robison, Thomas D. Evans and Martha A. Duff (Monthly Labor Review, September 2005, pp. 3-9.)

- IPUMS CPS has constructed a set of test weights for use with linked data
  - NOTE: these are BETA

Why Care about Weights?

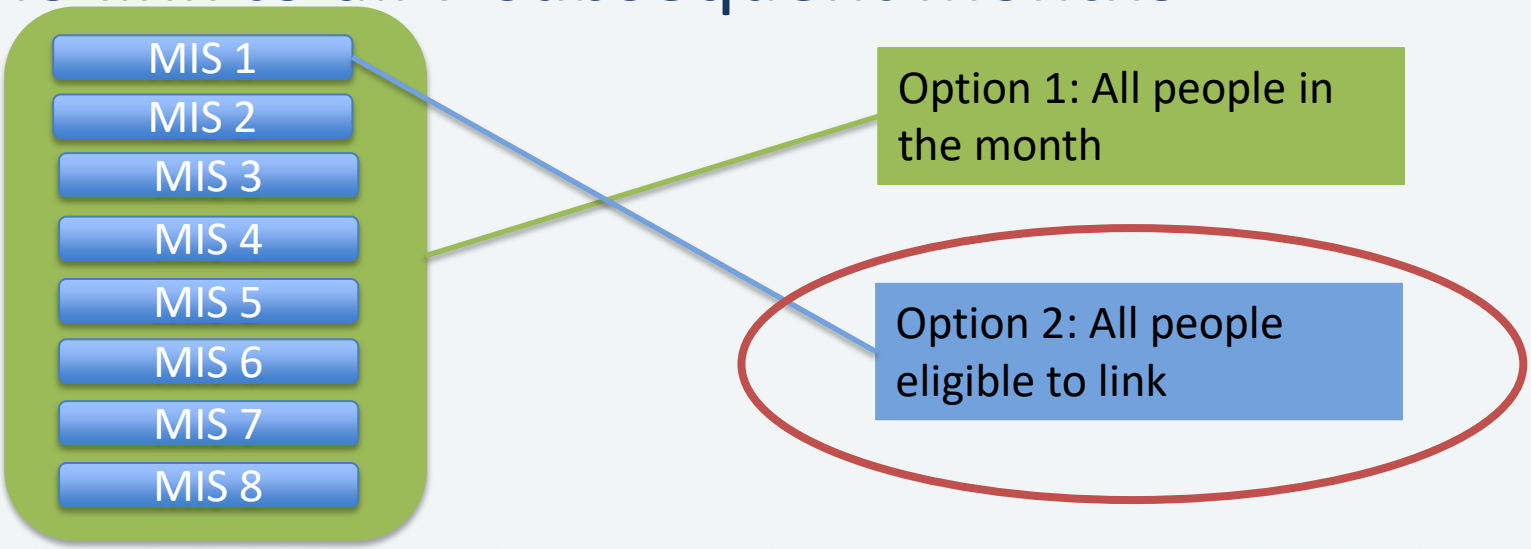
- Panel attrition and sample representativeness
  - Panel attrition is more severe as more time passes between observations
  - January - February 2009 (month to month)
    - 96% retention; 95% match on AGE, SEX, RACE
  - March 2009 - March 2010 (one year apart)
    - 79% retention; 75% match on AGE, SEX, RACE
  - All 8 months for cohort beginning in January 2009
    - 68% retention; 65% match on AGE, SEX, RACE

Options We Considered

- Doing nothing
- Normalizing the weights
- Generating regression-based adjustment factors – predicted probability of being observed in following month(s) and multiple T1 weight by 1/predicted probability
- Iterative Proportional Fitting

What population should we weight to?

- General problem: only a subset of any month can link to another month or set of months
- Specific example: only the subset of people who link to all 7 subsequent months



Construction of IPUMS Weights

- Iterative Proportional Fitting (IPF) or Raking in Python (validated in Stata)
- Builds on what CPS does
- Anchored to the month for which they are available and (currently) for forward linkages
- Idea is for subset of individuals who *actually* link to be representative of the subset of who *should* link

Eligible Subset

MIS (January)	MIS (February)
1	2
2	3
3	4
4	OUT
5	6
6	7
7	8
8	OUT

Eligible Subset

Construction of IPUMS Weights

- Sum of WTFINL for "eligible subset"
- Create cross-categories of:
  - AGE, SEX, RACE
  - AGE, SEX, HISPAN
  - AGE, SEX, STATE
  - Essentially sum of WTFINL at each category of every intersection
- Rake WTFINL for the eligible subset who actually link, for example, to the next month

IPUMS Weights for Linked Data

- Not comprehensive of all linking possibilities
- Based on types of linkages that we think people might use
- Based on ***forward*** linking
- Currently available from 1976 to present where linkages are possible

IPUMS Weights for Linked Data

IPUMS name	Purpose
LNKFW1MWT	Use for one month forward linkages
LNKFW1YWT	Use for one year forward linkages
LNKFW8WT	Use for a full 8 observation panel linked forward
LNKFWMIS14WT	Use for full MIS 1-4 observations linked forward
LNKFWMIS45WT	Use for MIS 4 linked forward to MIS 5
LNKFWMIS58WT	Use for full MIS 5-8 observations linked forward

Some Caveats

- Idea is for subset of individuals who *actually* link to be representative of the subset of who *should* link
- *Actual* linkages = "mechanical" linkages
- Current IPUMS panel weights are
  - For forward in time linkages
  - Created based on adjustments to WTFINL, not supplement weights

So What to do about Weights?

- Do something…
  - IPUMS recommends using weights for analysis of CPS data
  - Lots of sensitivity analyses
- Help us test our new weights for linked data and give us feedback!
- We are planning to conduct analyses with original and new weights to better understand the extent to which weight adjustment methods (or not) matter for estimates

Creating Weights for Linked Data

- Stata **ipfraking** program written by Stanislav Kolenikov
  - Documented in https://www.stata-journal.com/sjpdf.html?articlenum=st0323

- IPUMS CPS code is available