What's New in the HRS

Kathleen McGarry
UCLA and NBER

Health and Retirement Study: Sample

- Nationally representative panel survey
  - ~ Representative of the pop 50+ and their spouses
  - Begun in 1992 with sample born 1931-1941
  - In 1998 merged with
    - AHEAD born 1923 or earlier, interviewed in 1993, 1995
    - “War Babies” born 1942-1947
    - “Children of the Depression Era” born 1923-1930
  - New cohorts added every six years
    - In 2004 added “Early Baby Boomers” born 1948-1953
    - In 2010 added “Middle Baby Boomers” born 1954-1959
    - Plan to continue refresher cohorts in the future
  - Total respondents to date over 30,000

Health and Retirement Study: Content

- Very Detailed Data on Income and Wealth
- Family (children, parents, siblings)
- Health
  - Disease conditions, health status, ADL limitations
  - Biomarkers, genetics

- Innovative questions
  - Bracketed responses for financial variables
  - Subjective expectations
  - Experimental modules

Health and Retirement Study: Content

- Off-years surveys
  - CAMS (Consumption and Activities Mail Surveys)
  - HUMS (Human Capital and Educational Expenses Mail Survey)
  - Prescription Drug Surveys
  - Disability Vignettes
- ADAMS (Aging, Demographics and Memory Study)

Health and Retirement Study: Sample

- Innovative Design
  - Individuals are followed into nursing homes
  - Exit interviews
  - Mode
    - Phone and face to face
  - Links to administrative records
    - Social Security
    - Medicare
    - Pension data
    - Geographic data

What's New Today?

- Family
  - New Data Set on the Horizon
- Consumption
  - Newly release imputations and weights
- Health
  - Biomarkers, genetics

HRS / RAND Family Data

- HRS family data is an extremely rich resource that is currently under used
- Information on respondent, children, parents, siblings for 18 years +
- Develop user friendly data sets to facilitate empirical research on the family

For Each Child:

- Characteristics of Each Child
  - Sex, birth year, marital status, own home, coresident, lives near respondent, household income (earnings), education, employment, contact frequency
- Financial and other help *given to* child
  - Transfers of money, grandchild care, inclusion in wills, trust, deed to home, who coresidency helps
- Financial and other help received from child
  - Transfers of money, help with ADL limitations and chores, expect future care

Parent / In-law Variables

- Characteristics of each Parent / Parent-in-law
  - Age, marital status, own home, lives near, region of country, relative financial situation, needs help with ADL limitations, can be left alone, has memory related disease
- Transfers to Parents
  - Financial Transfers, time helping with ADL limitatations, time with chores, who coresidency helps

Parent Variables

- If the parent has died :
  - Date of death
  - If they had an illness last more than 3 mos
  - Ever in a nursing home
  - Left inheritance

Sibling Variables

- Number of siblings of respondent & spouse
- For each sibling:
  - Sex, if sib provided money and/or care to parents, amount and type of help
- More detail on up to 4 siblings:
  - Distance from parents, financial status, co-reside with parent

Data Have Not Been Exploited

- Large Amount of Information
  - Large number of variables per child
  - Large number of children (20+)
  - Number of children changes over time (step children)
- Questions change across waves
  - Some variables are not asked repeatedly and need to be carried forward
- Answers not always straight forward
  - “All my children equally”
  - Helper data: who, how often, how much?
- Typical measurement error

Challenges of Linking Children

- Respondent & spouse split
  - Multiple reports on a child
  - Remarry, new spouse's kids added
  - Remarry each other
- IDs not always consistent across time
- Relationship difficulties
  - Relationship varies
  - Treatment of children's spouses varies
  - Treatment of grandchildren varies

Challenges of Linking Children

- Strategy for Linking Children
  - Use OPN / SUBHH
  - Algorithm checks consistency of gender, birth year, relationship codes
  - Restricted data: child names

Key Variables in new files

- Unique child, parent, and sibling identifiers facilitate linking over waves
- Cleaned and consistent measures of demographic and economic measures across waves
  - Carried forwards, imputed values,
  - Derived variables
    - e.g. time and money transfers
- Flags identifying relationship to respondent
- Flags identifying imputed variables

Data Files

- One Data Set at the Respondent Level
  - Data on all family members e.g.
    - Each child, parent, and sibling
  - Consistently named variables across waves
    - Child's age in each wave, employment status, etc
  - Aggregate measures
    - e.g. Average age of children
    - Number of grandchildren
    - Oldest parent/parent-in-law
- Three Data Sets at Family-Member Level
  - One each for children, parents, siblings

Documentation

- Codebook
  - Follows RAND HRS codebook format
  - Basic descriptive statistics
  - How variables constructed
  - All raw HRS variables listed
  - Cross wave differences noted
- SAS Programs used to construct data sets

First Release

- Child Data for 1998, 2000, 2002
- Variables include:
  - Relationship, alive, sex, birth year, marital status, number of own children, proximity, contact frequency, in school, years of schooling, working
  - HRS respondent provides child care, financial transfers to child,
  - HRS respondent receives financial help, help with ALDs / IADLs, chores, expected future help

Timeline

- Preliminary Release in April 2011
- Full preliminary release in May 2011
- All wave of child data in September 2011
- Parents in December 2011
- Siblings in December 2011
- Continue to update

Number of Obs per child

Num Obs	‘92 cohort	’98 cohorts	’04 cohorts	All
1	2,502	2,450	648	7,688
2	1,450	2,464	650	6,149
3	1,363	2,224	***5,016***	9,205
4	1,016	2,134	--	4,416
5	1,147	2,541	--	4,712
6	993	***11,512***	--	12,706
7	1,236	--	--	2,038
8	1,994	--	--	2,409
9	***13,985***	--	--	***13,985***

Multiple Observations per child

- When parents split have two reports per child.
  - If parents reunite we are back to one report in that wave
  - Several cases where they then split again.
  - Viewed as hold up in past

Stories per child

“Stories”	Number of Obs
1	59,269
2	1,881
3	87
4	4

Part II of Child Data: Proximity

Whether a child…	Which Years
…is resident	All
…lives within 10 miles	All
…lives nearest among those living 10 miles away	1995-2002 if no child lives within 10 miles; 2004-2008 can be inferred from child proximity file
…lives nearest among those living within 10 miles	2004-2008
…lives within 2 blocks	2004-2008

Geographic Proximity: Zip Code Data

- In 2004, 2006, 2008 HRS collected zip code information on children (or city and state).
  - 2004 for all children
  - Subsequent years if child moved / new info
  - 86% of families have at least one report
  - 78% of children have reported information
- Data are available as restricted data

Geographic Proximity: Zip Code Data

- Wanted more readily accessible data
- Summary measures that would be useful to researchers and could be made publicly available
- Chose to begin with pairwise distances between parents and children.
- Conference last September for new work using these data and similar data in PSID.

What should be made available?

- Currently have:
  - Distance between parents and each child in each wave
- Adding :
  - Distance between children themselves (up to four)
  - Distance across waves for respondents
- Other options:
  - Distance between ex-spouses
  - Driving times between family members
  - Distance to hospitals / schools / work

CAMS

- Biennial mail survey
  - 2001, 2003, 2005, 2007, 2009…
  - Self-administered
- Random sub sample 5000 HRS households in 2001
  - Response rate was 77.3%
  - Interviewed biennially
  - 850 respondents from new cohort added in 2005
  - Anticipated collection Oct 2011 with new cohort

CAMS

- Part A. 36 activities (time-use) categories:
  - How many hours did you _____ last week ?
  - Activities: Reading, tv, working, with friends, …
- Part B. 32 consumption categories:
  - 6 big ticket items (e.g. cars, appliances)
  - 26 non-durable items (e.g. clothes, electricity)
- Categories designed to follow CEX
  - Matches extremely well

CAMS- CEX Comparison

CAMS – CEX Comparison Average of 2001, 2003, 2005 (2003$)		
Age	CAMS	CEX
55-64	38,970	39,677
65-74	34,276	32,436
75+	28,761	24,066
All	34,472	33,096
CAMS and CEX close at 55-54 but at 75+ CAMS is 20% higher		

Going Forward

- Added MBB cohort in 2010 (1954-1959)
- Added minority sample
  - Aiming for ~ 3,000 individuals
  - Expected composition
    - 1,250 blacks
    - 1,000 Hispanics
    - 750 non-minority

Minority sample sizes – core respondents

	1998	2004	2010 (est.)
Black	3,001	2,874	4,400
Hispanic	1,652	1,927	3,100
Other	16,731	15,328	16,300
Total	21,384	20,129	23,800