Paradata and Survey Management in the National Survey of Family Growth (NSFG), 2006-2012

William D. Mosher, NCHS; James Wagner, Mick Couper, & Nicole Kirgis, U of Michigan ISR

National Conference on Health Statistics

Washington, DC, August 7, 2012

wmosher@cdc.gov

Outline

- Goals & Features of the NSFG
- Using Paradata for Responsive Design in the NSFG
- The NSFG has used Paradata to manage:
  - survey costs & response rates.
  - data quality & interviewer performance
- Summary
- Future work on paradata in the NSFG

Main Features of the National Survey of Family Growth , 2006-2010

- National multi-stage area probability sample
- Topics: Pregnancy, contraception, infertility, marriage, parenting
- Interview men & women age 15-44
- All interviews in person. No phone, no internet.
- Continuous design, using four 12-week “quarters” per year.
- 14,000 addresses selected each year, to yield at least 5,000 interviews per year.
- Release data files with every 2-4 years of data.

NSFG (continued)

- Two *phases* of data collection in each 12-week quarter:
  - Phase 1: 10 week data collection
  - Phase 2: Subsample remaining cases
    - Select a sample of about 1/3 of the remaining cases;
    - That reduces each interviewer's case load by 2/3
    - Oversample cases with large weights (adults, whites)
    - Change data collection model:
      - added interview token of appreciation,
      - interviewer behavior change

        --(triple the number of hours per case)

  - Combine data and response rates from two phases using weights

Goals of Responsive Design in the NSFG

- The major expense in an in-person area probability sample is interviewer time, including labor and local travel costs.
- So the NSFG's initial goals in using responsive design were to use paradata to ration and allocate the interviewers' time to manage costs and response rates.
- The NSFG is trying to balance the following goals:
  - Control costs & prevent cost increases;
  - keep response rates at acceptable levels

    (70- 80% overall & for sub-groups)
  - get at least 5,000 interviews per year.

Key Elements of Responsive Design

- Monitoring key indicators using paradata
  - The use of statistical process control methods to track interviewer effort
- Statistically-based decisions to alter the design
  - Targeted at one or more measurable outcomes like response rates, sample yield, hours of interviewer labor per interview, or data quality.
- Interventions targeted at particular subsets of the sample
  - As opposed to changing a strategy for the entire sample
- Documentation of the decision process
- Evaluation of the success of the intervention *(some work, some don't; you have to drop things that don't work.)*

NSFG Responsive Design Interventions

- Weeks 1-10: Phase 1 interventions
  - Targeting interviewer effort at screener versus main calls at selected points in the quarter
  - Identifying and targeting “high priority” cases, often with embedded randomization
- Weeks 11-12: Phase 2 intervention
  - Subselection of high priority cases targeted at improving response rates and minimizing bias
- All interventions informed by
  - Response propensity models run daily (overnight)
  - Dashboard:
  - Tracking of key indicators on daily basis,
  - comparisons to survey goals and prior quarters

The NSFG Dashboard

Effort

- I'rs working
- hours
- % production
- calls/day
- calls/hour
- % peak calls
- scrn'r/main calls

X

Active Sample

- occupied
- eligible
- nonworked
- noncontacts
- mean calls
- 8+ calls
- locked bldgs
- resistant
- hard appt.
- propensity

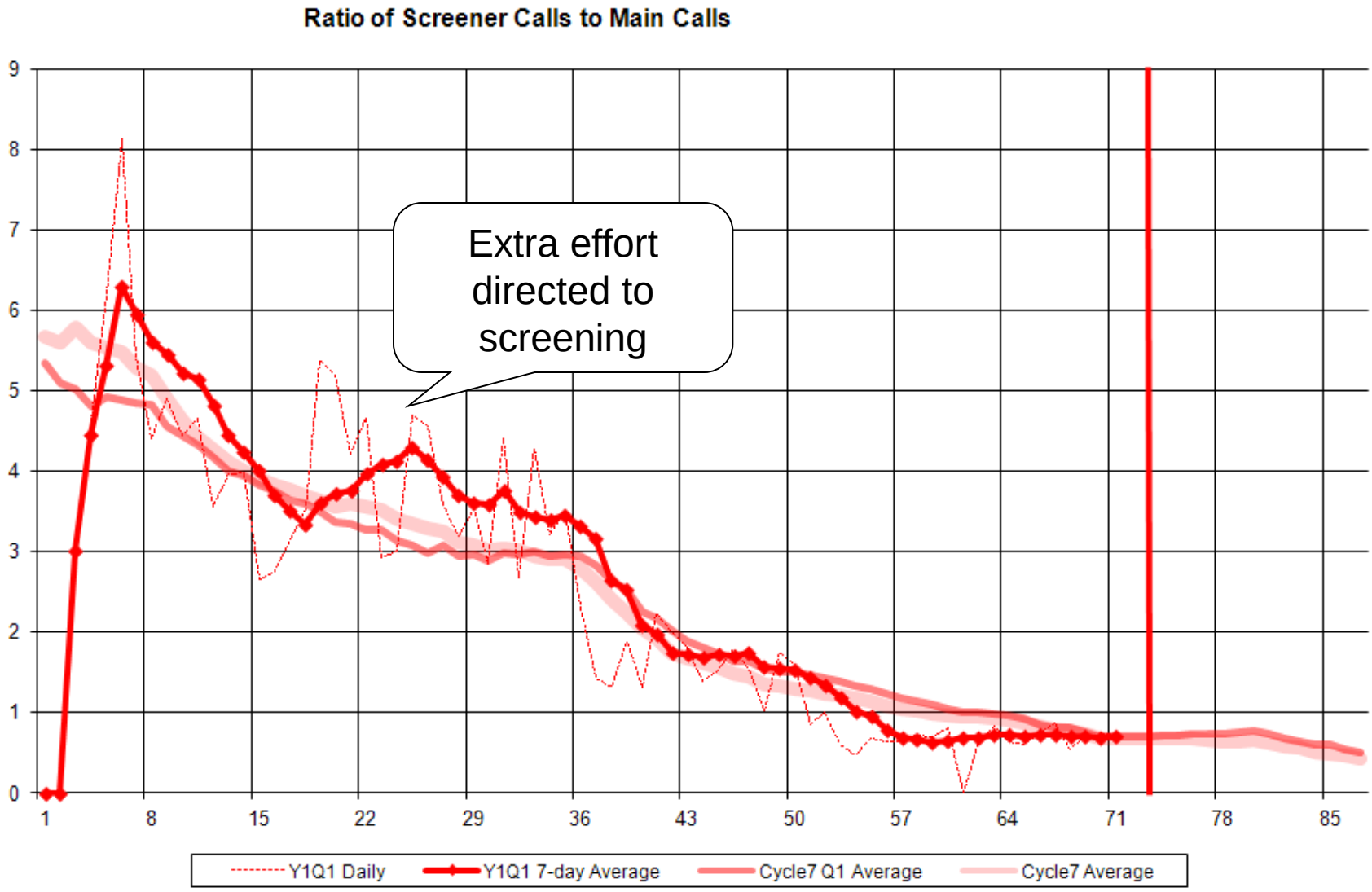
Productivity

- interviews
- cum. interviews
- hours/interview
- calls/interview

Data Set Balance

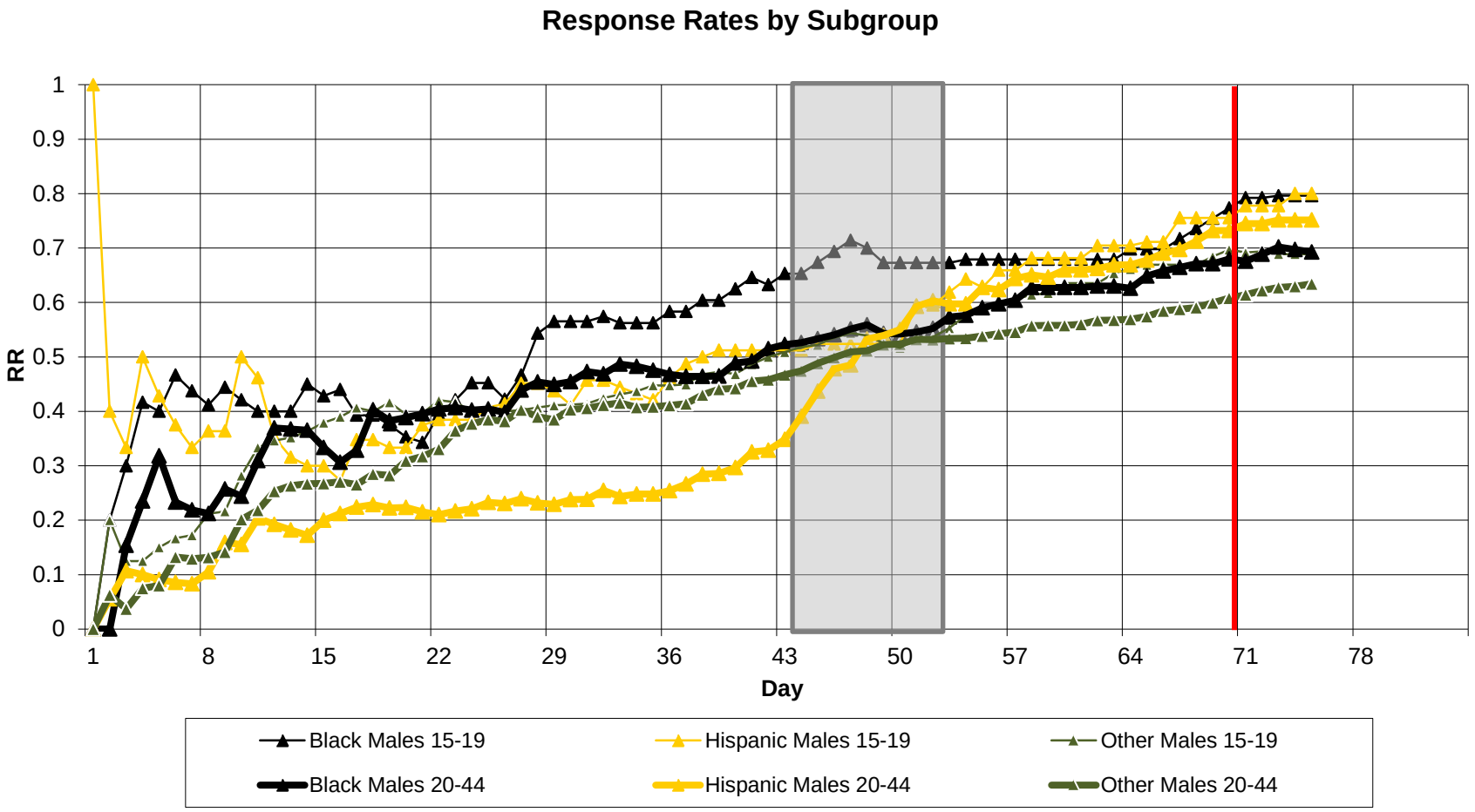
- response rate
- % with kids
- % sexually active
- group rates
- CV group rates

NSFG Intervention: Screener vs Main



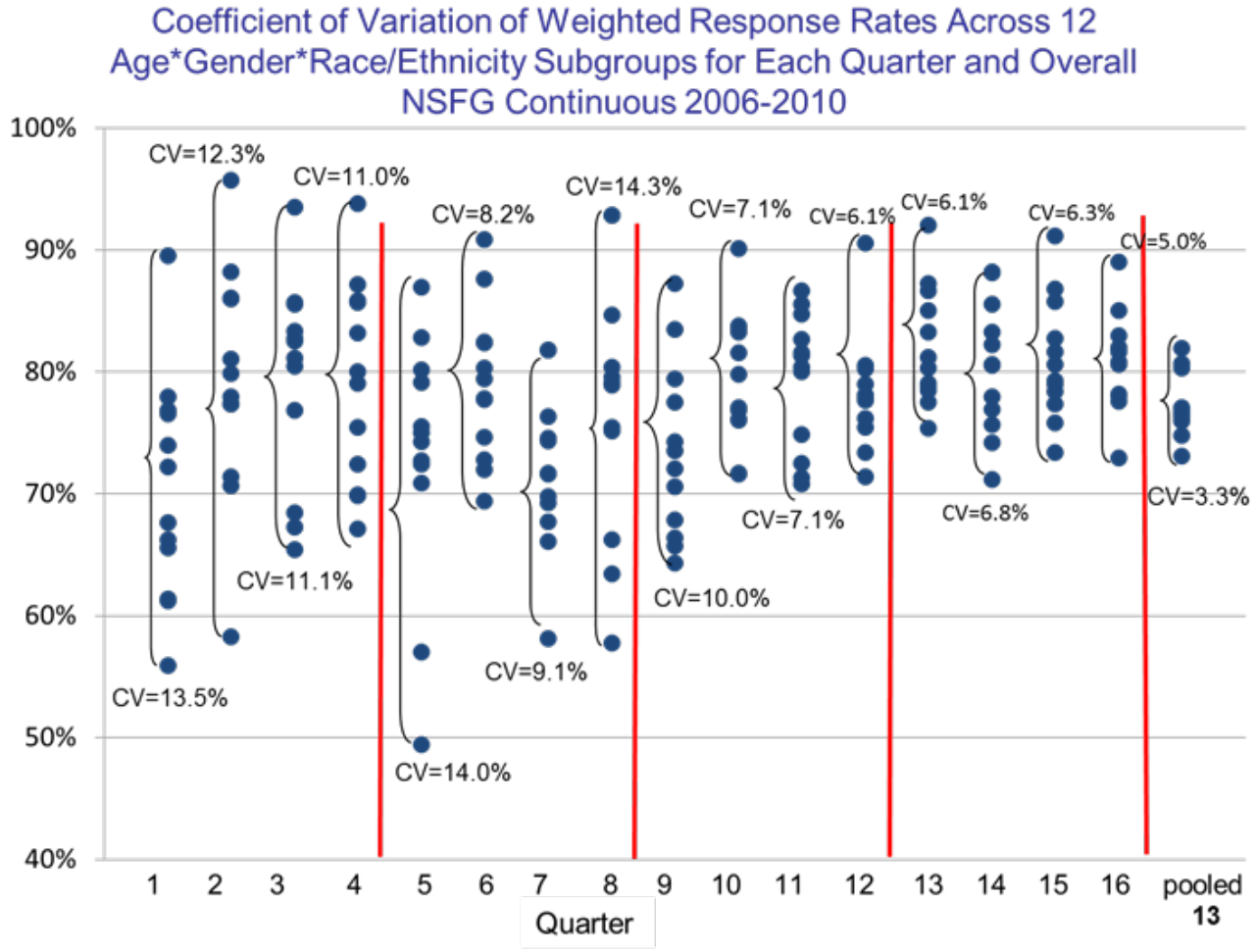
NSFG Intervention: Subgroup Response Rates

Response rates for Hispanic adult males lag before intervention; higher after it.



NSFG Intervention: Response Rates for sub-groups varied more in Quarter 1 than in Quarter 16

(Male & Female, 15-19 & 20-44; Hispanic & Black & other race)



"Interventions" used in the 16 quarters of NSFG fieldwork in 2006-2010

- At least 16 experimental (randomized) "interventions" were used, to increase response rates or sample yield
- Most interventions increased rates for the group they were aimed at, but not always significantly.
- Some will be continued in the new NSFG, others won't be.
- Paper on the interventions:

J Wagner et al, "Use of Paradata in a Responsive Design Framework to Manage a Field Data Collection," *Journal of Official Statistics*, forthcoming.

Using Paradata to predict the survey's outcome variables

- We can compare means of key statistics for subgroups defined by the paradata. We tracked 19 key statistics by several paradata indicators.
- For example, we can look at means of key statistics by an interviewer observation as to whether the interviewer thinks that the adults living here are in a marriage or cohabiting relationship. The percentage of these women who had ever been pregnant was:
- Example: Proportion of women ever pregnant
    - Married or cohabiting: 69% ever been pregnant
    - Not married or cohab: 27% ever pregnant

How are we using paradata now, in the 2011-2015 NSFG?

- Continuing development and use of dashboard to guide responsive design decisions
  - Identifying and improving key indicators
  - Identifying and documenting interventions

    (some successful, some not)
- Challenges of dealing with the federal security requirements (we can only access paradata in secure locations).
- Development and evaluation of data quality indicators (see next slides)

Paradata for Data Quality Evaluation

- Many surveys now use Computer audio-recorded interviewing (CARI) to evaluate data quality and interviewer performance.
- We don't, for two reasons:
  - sensitive nature of NSFG content raises privacy & data security issues.
  - Budget reasons.
- Exploring alternatives using item-level paradata (keystroke files)

Indirect Data Quality Indicators

(R = Respondent)

- Example indicators derived from keystroke data
  - How long does it take R's to answer the question?
  - how many times did an interviewer back up and change an answer?
  - Error messages and actions taken by interviewer.
  - Item missing data (DK, RF) rates.
  - How often does the interviewer access the Help function?
  - How often does the interviewer enter a comment on R's answer?

- **Used factor analysis to combine indicators**
- **Goal= identify outlier interviewers on multiple factors**

Use of Data Quality Indicators
by interviewer, and by question

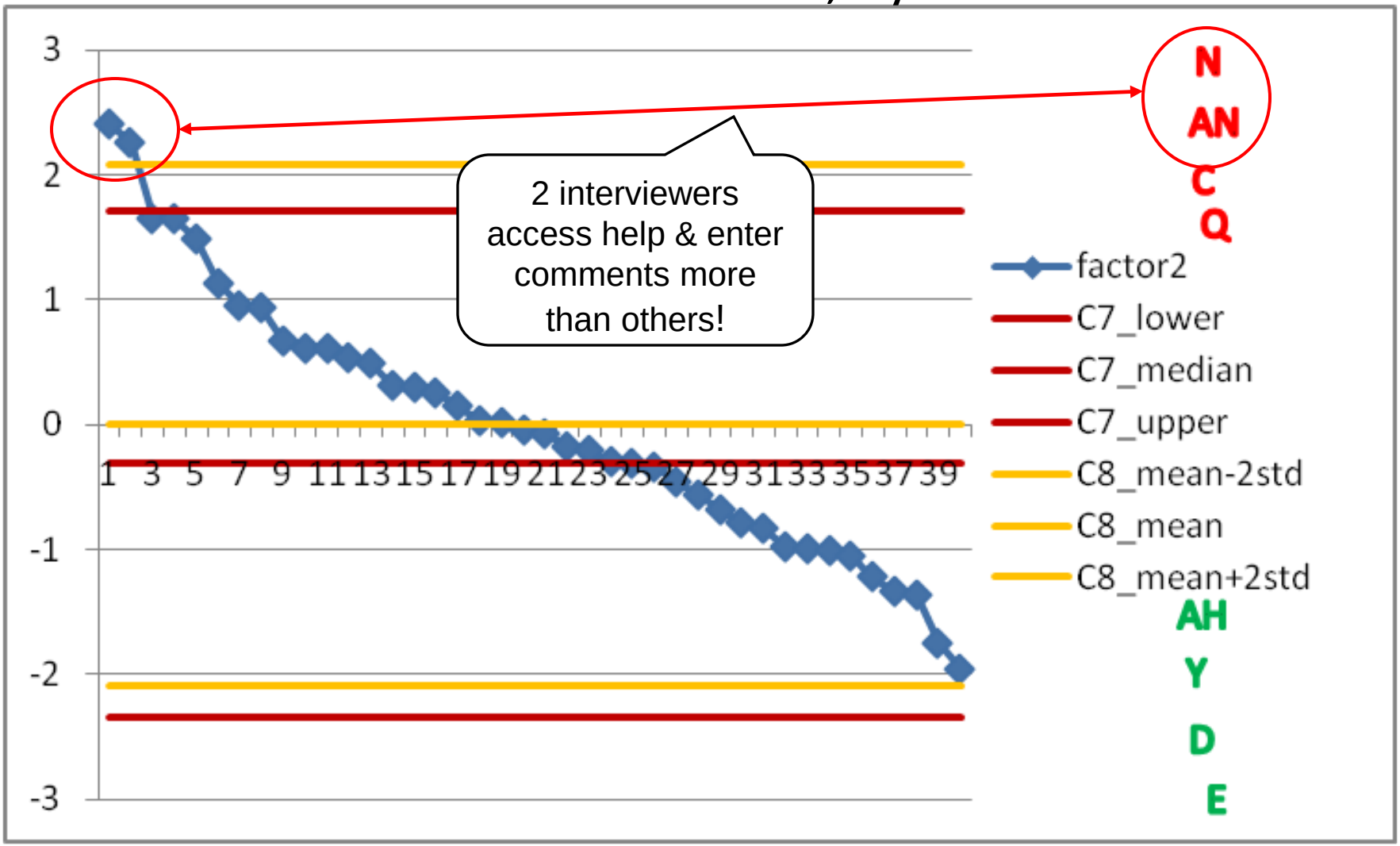
By Interviewer:

- We are currently using these indicators to identify interviewers for special attention, including re-training in the weekly calls with supervisor.

By Question:

- But showing these indicators for specific QUESTIONS (instead of INTERVIEWERS) can help identify potentially problematic items or questions.

How much do interviewers access Help and enter comments? Standard scores, by interviewer.



Summary: Use of Paradata in the NSFG

- In the NSFG, paradata were used in 2006-2010 to (a) limit costs, (b) get adequate sample sizes, and (c) equalize response rates across subgroups by age-race-gender that are strongly correlated with NSFG outcomes
- Uses included targeting groups with lagging response rates, encouraging adequate screening effort, and designing the non-response follow-up.
- In 2011-12, we have also been using paradata to monitor data quality and interviewer performance. Intent is to improve questionnaire design and interviewer training.
- Thinking about a future paradata data file: Which paradata elements would be useful for a limited release paradata file?

References and further information

1. NSFG web site: http://www.cdc.gov/nchs/nsfg.htm
2. ISR/U of Michigan methodology research on NSFG:
http://www.psc.isr.umich.edu/pubs/series-list.html?sc=ng

3. Groves et al, "Planning & Development of the Continuous National Survey of Family Growth." NCHS, Vital & Health Statistics, Series 1, No. 48, Sept 2009, pages 23-27.

4. Lepkowski et al, The 2006-2010 National Survey of Family Growth: Sample Design & Analysis of a Continuous Survey. Vital & Health Statistics, Series 2, No. 150, June 2010, pages 14-17.