

---

# Survey Quality Indicator Measures: Response Rates and Alternatives

2012 Federal Committee on Statistical  
Methodology Research Conference

January 11, 2012

Washington, DC

Donsig Jang • Sixia Chen (Iowa State University)  
Flora Lan (National Science Foundation)

**MATHEMATICA**  
**Policy Research**

# Background

---

- Survey goal is to provide high-quality data
- Response rate was considered an indicator of survey quality
  - Data collection goal: High response rate
- But the threshold for a high response rate has changed due to decreasing response rates (90% to 70%, or even lower) from the early 1990s to mid-2000s

- Non-response bias  $b(\bar{y}) = (1 - \rho)(\bar{y}_r - \bar{y}_{nr})$

# Background (cont'd.)

---

- **Response rates alone are not good indicators of nonresponse bias (Groves and Peytcheva 2008)**
  
- **Alternatives**
  - **Multiple thresholds of response rates by key domains**
  - **R-indicator (Schouten et al. 2009) measuring representativeness**
    - Measure of response propensity rate variation among respondents
    - Leading to focus on less representing subpopulations during a late stage of data collection

# Objectives

---

- **Empirical evaluation of response rates and R-indicators using real survey data**
  - Observe the relationship between the two measures
  - Understand the relationship between each measure (response rate or R-measure) and potential nonresponse bias
- **Decision for data collection closeout**
  - Based on response rate, R-indicator, both or other alternative indicators?

# Methods

---

- **Examine R-indicator and response rate trends over the data collection period**
- **Calculate key survey and frame variable estimates on a weekly basis during data collection**
- **Calculate upper-bound estimates for bias and root mean square errors for weekly estimates during data collection**

# Data: National Survey of Recent College Graduates (NSRCG)

---

- **Sponsored by the National Science Foundation (NSF) and conducted every two or three years since 1974**
- **Targets recent graduates with bachelor's or master's degrees in science, engineering, or health**

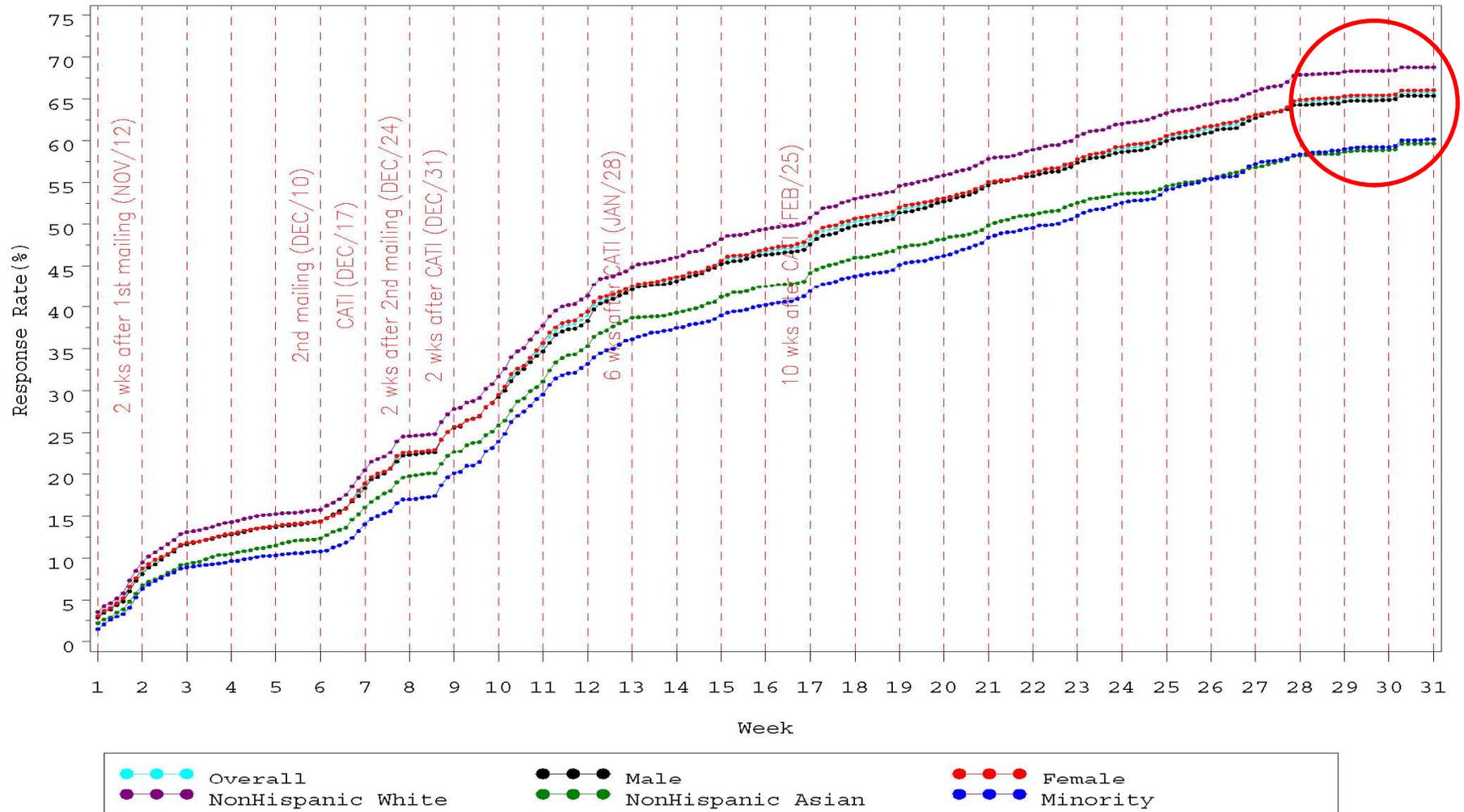
## NSRCG (cont'd.)

---

- **2008: AY06, AY07**
- **Two-stage sample design: school sample (first stage) and graduate sample (second stage)**
  - Sample sizes: 300 schools and 18,000 graduates
  - For more information, visit [www.nsf.gov/statistics/srvyrecentgrads](http://www.nsf.gov/statistics/srvyrecentgrads)
- **Information collected on demographics, education, employment, etc.**
- **Mixed mode: Mail/Web with CATI follow-up**
- **Final response rates**
  - 71.4 (unweighted), 69.7 (weighted)

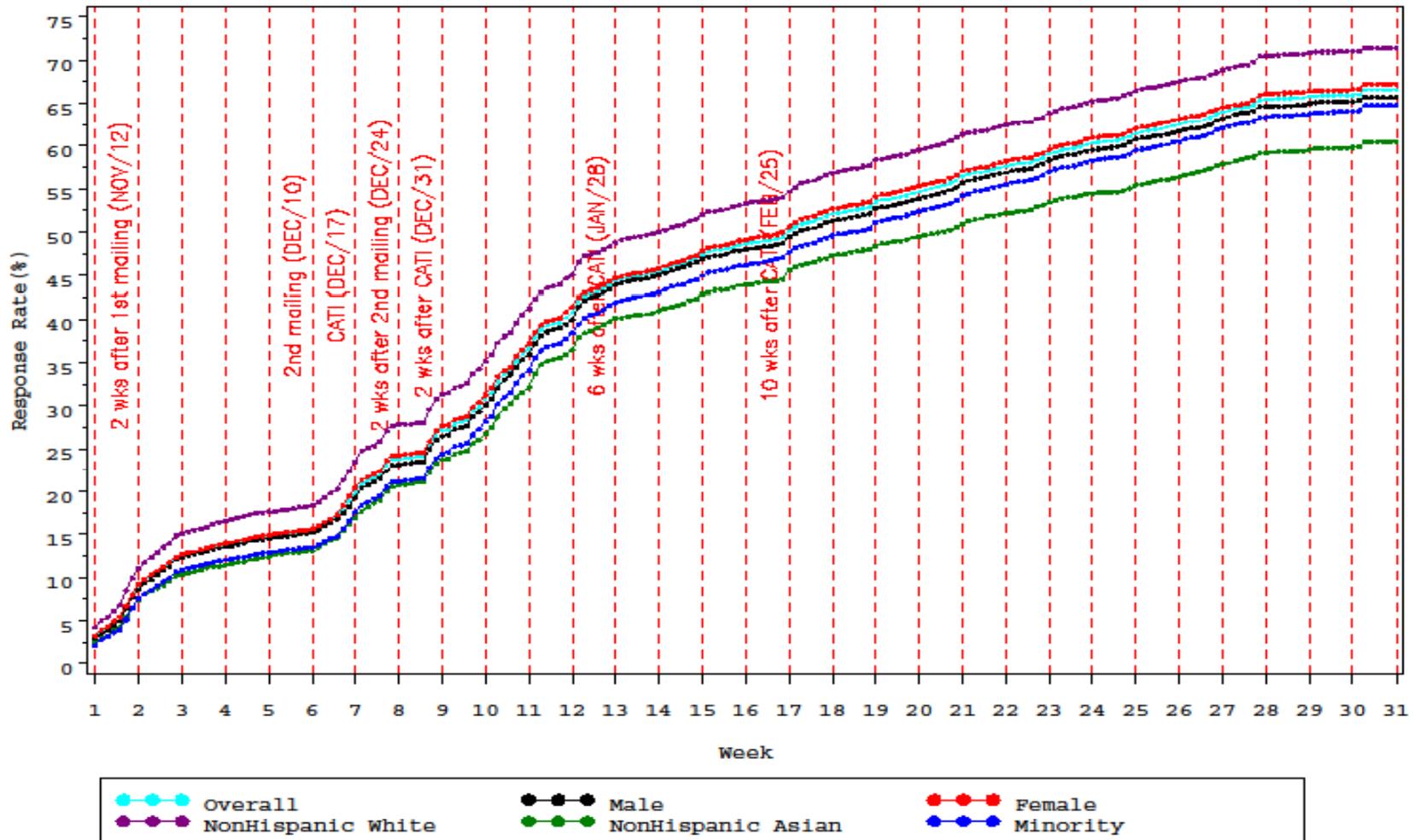
# Response Rates (Weighted)

Weighted RR by CompletionDates



# Response Rates (Unweighted)

Unweighted RR by CompletionDates



# R-Indicator

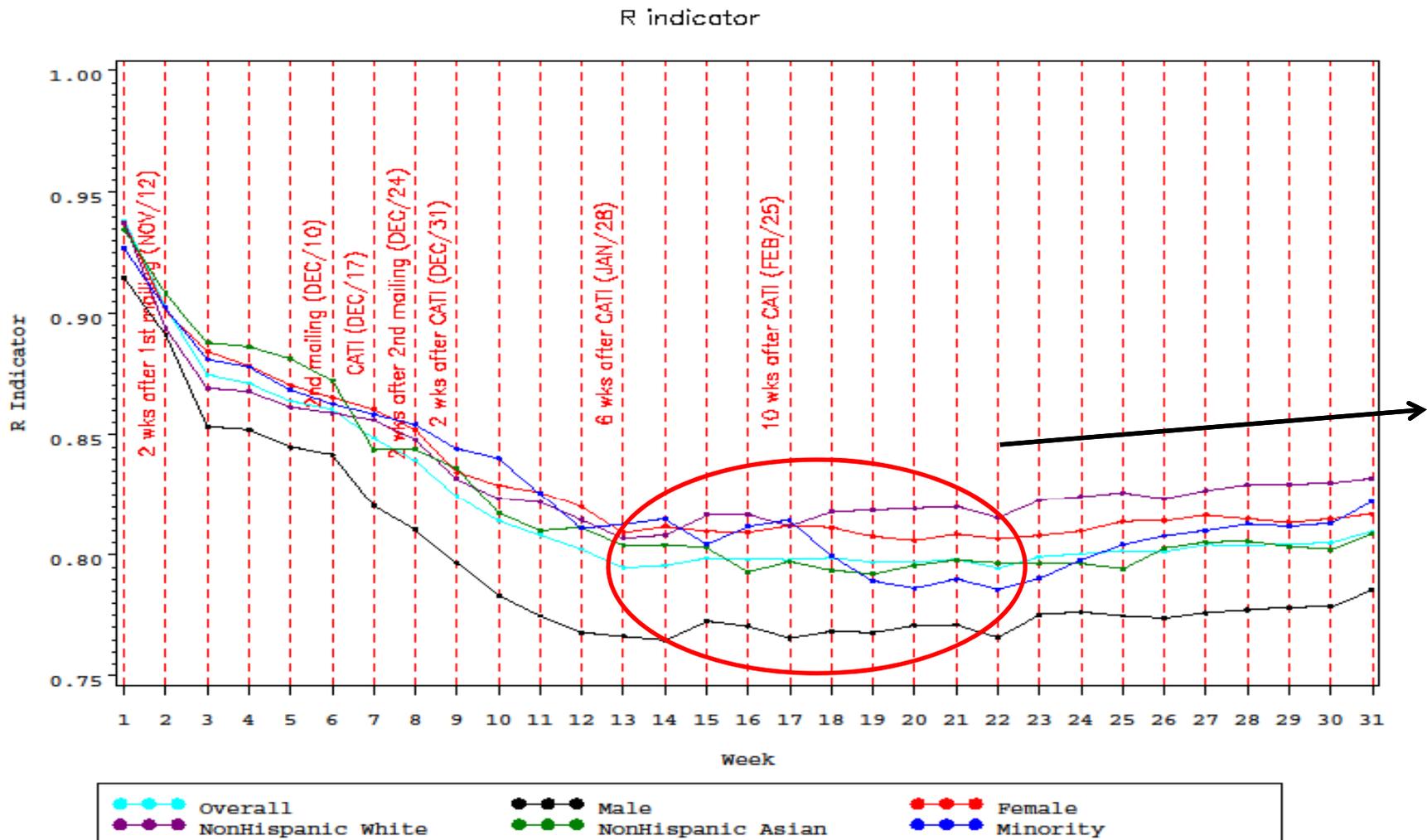
---

- Measures similarity or dissimilarity of response propensities
- A measure independent of specific outcome variables (similar to response rate)

$$R(\rho) = 1 - 2\sqrt{\frac{1}{N-1} \sum_{i=1}^N (\rho_i - \bar{\rho})^2} \longrightarrow \hat{R}(\rho) = 1 - 2\sqrt{\frac{1}{\hat{N}-1} \sum_{i \in R} w_i (\hat{\rho}_i - \hat{\bar{\rho}})^2}$$

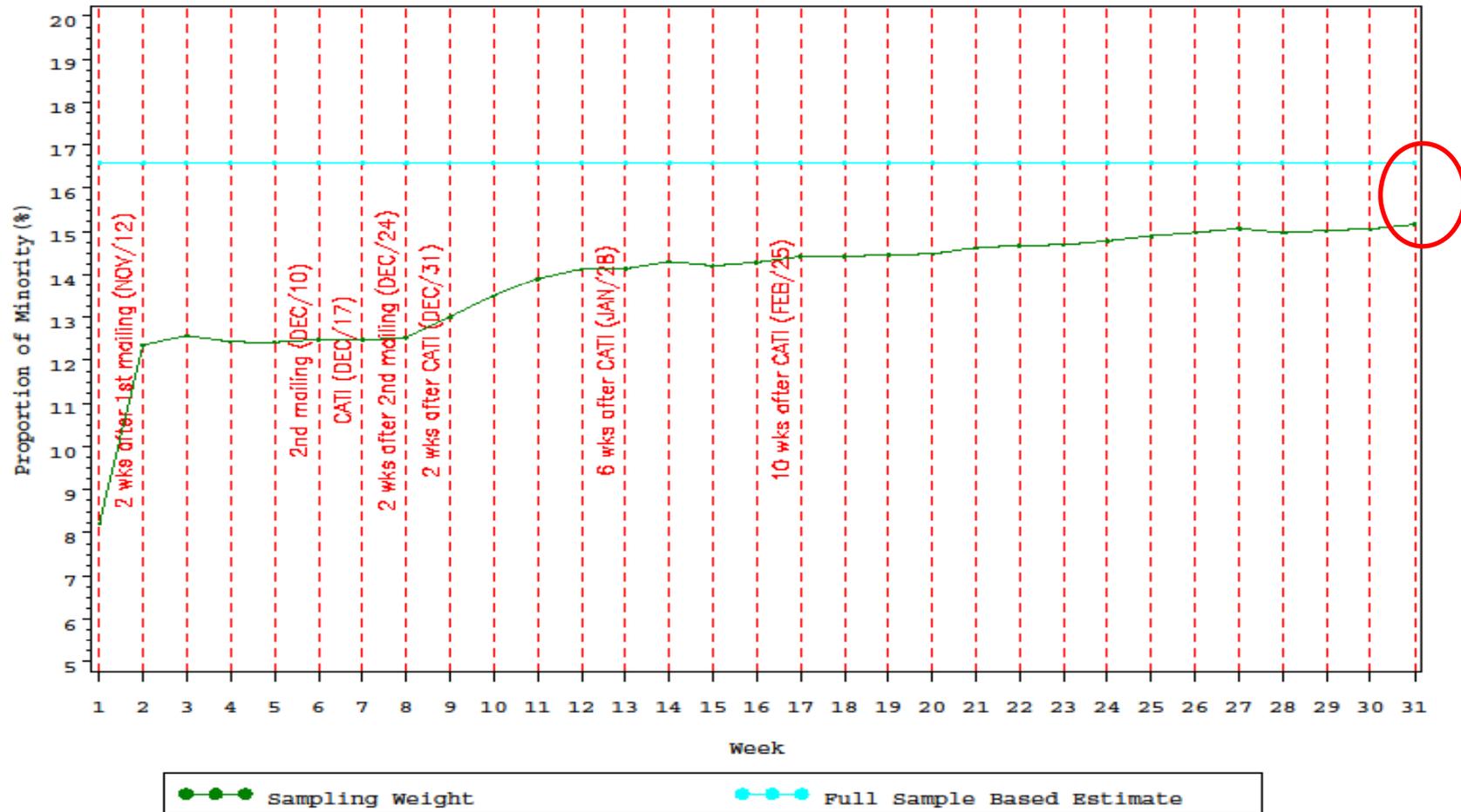
where  $\rho_i$  is an individual response propensity

# R-Indicator (cont'd.)



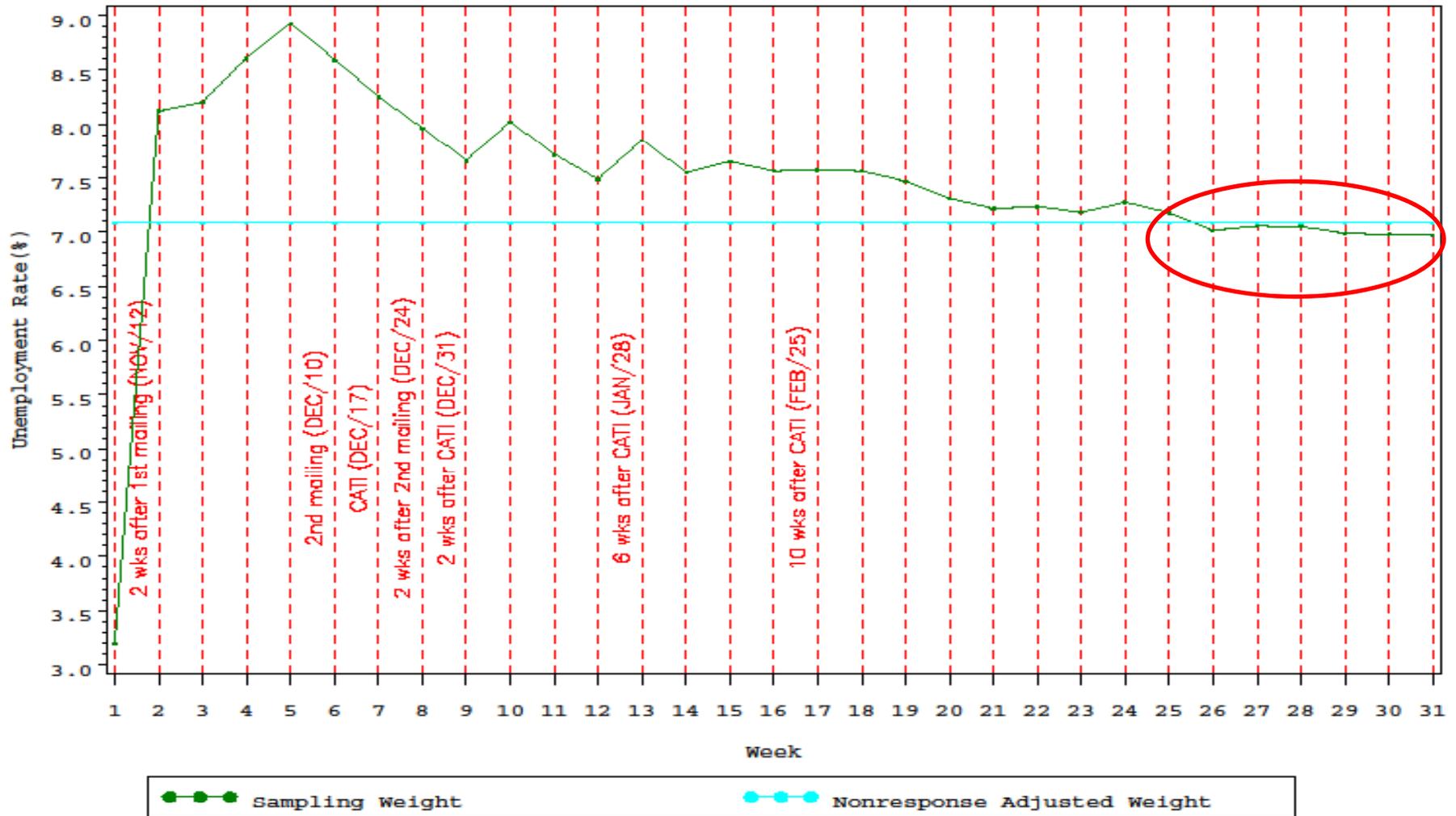
# Tracking Estimates for Frame Variables

Percentages of Minority group  
(Hispanic, Black, American Indians)



# Tracking Estimates for Survey Variables

Unemployment rate estimates for Minority group  
(Hispanic, Black, American Indians)



# Upper Bounds for Bias and Root Mean Square Error

---

- Upper bounds of bias and root mean square errors can be estimated using survey response data:

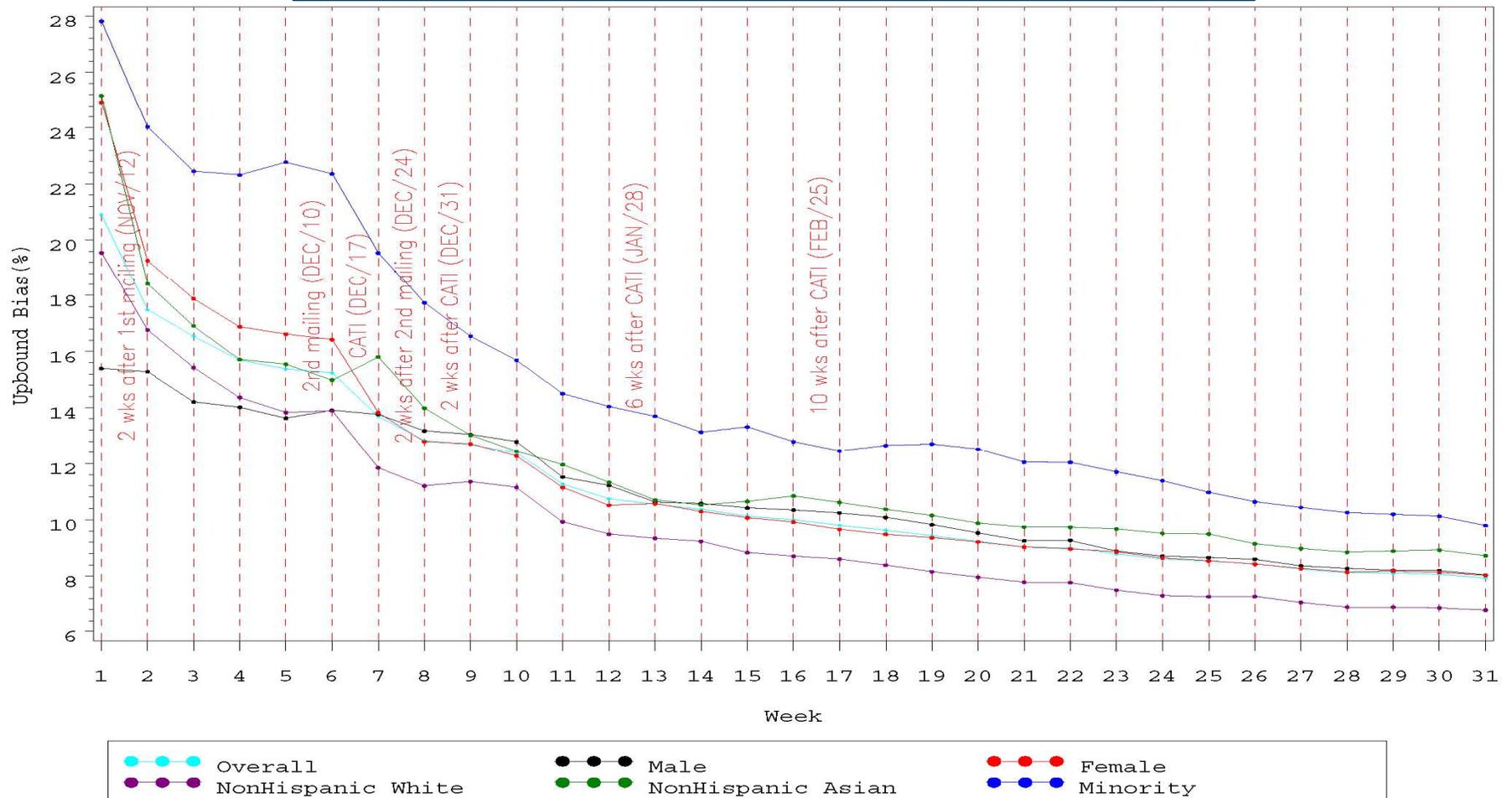
$$\hat{B}_U(\hat{y}) = \frac{(1 - \hat{R}(\rho))\hat{S}(y)}{2\hat{\rho}} \quad \widehat{RMSE}_U(\hat{y}) = \sqrt{\hat{B}_U^2(\hat{y}) + v(\hat{y})}$$

where  $v(\hat{y})$  is a variance estimator of  $\hat{y}$  and

$$\hat{S}(\hat{y}) = nv(\hat{y}) / deff(\hat{y})$$

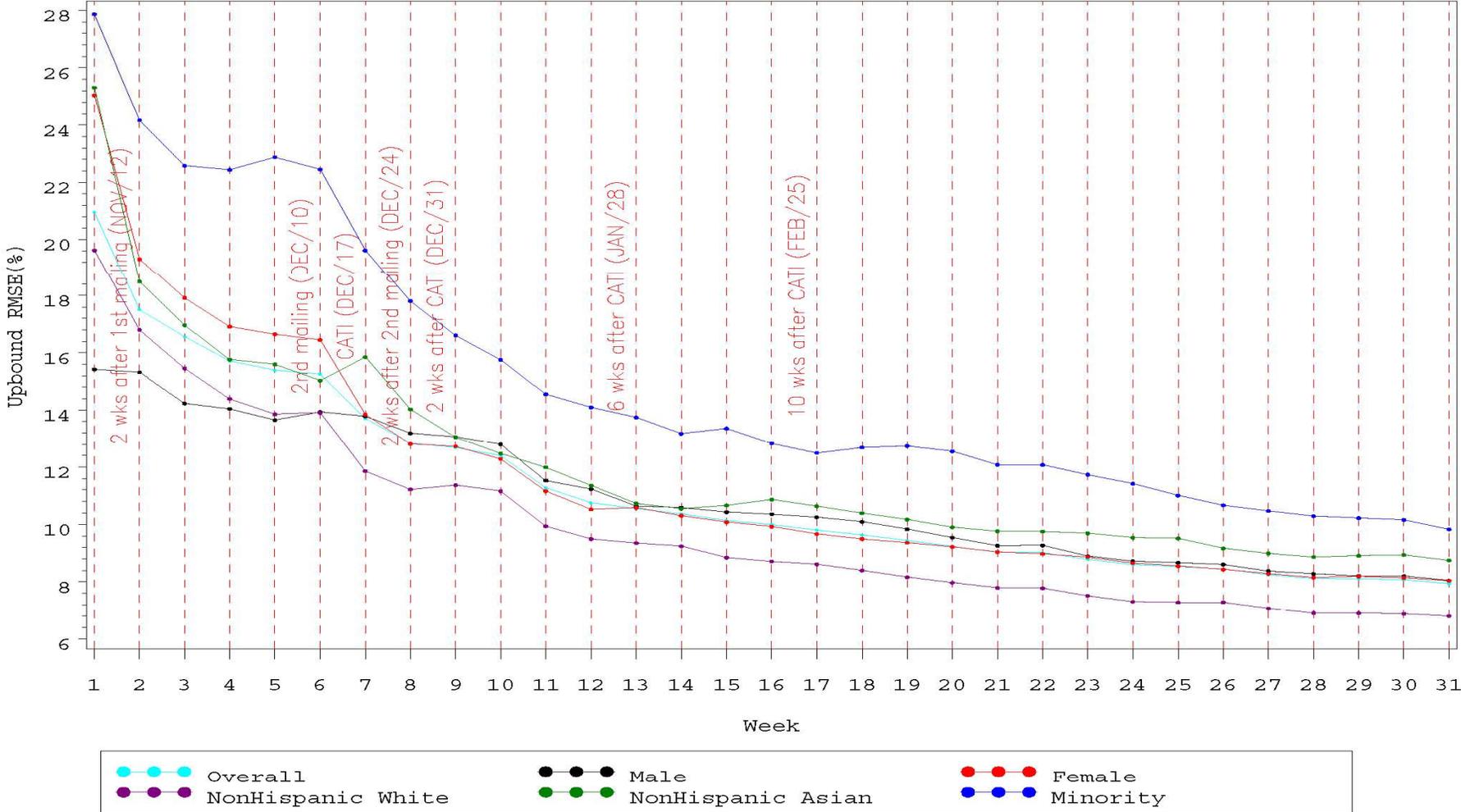
# Upper Bound Bias

## Upper Bound Bias of Unemployment Rate Estimates



# Upper Bound Root Mean Square Error

Upper Bound Root Mean Square Error for Unemployment Rate Estimates



# Results

---

- **Response rate curves are monotonically increasing, but in increments of only 1% across all key domains (gender and race/ethnicity) during the last three weeks (weeks 29–31)**
- **R-indicator curves are U-shaped, with the lowest values between week 13 (RR = 44%) and week 22 (58%) overall and for key domains**
  - **After week 22, R-indicator values for most domains steadily increase but not much**

## Results (cont'd.)

---

- **The percentage of minority graduates among respondents is less than that of the full sample at the end of data collection, indicating the importance of weighting adjustment**
- **Survey estimates for “unemployment rate” of Minority group seem steady after week 25**
- **The upper bounds of potential bias indicators (bias and RMSE) for “unemployment rate” estimates are steadily decreasing, although the rates of decrease over the last few weeks are minimal**

# Discussion

---

- **The last three weeks of 2008 NSRCG data collection added, at most, 1% point to the response rate—supporting the data collection closeout decision at week 31**
- **Other measures may have supported the decision made**
  - **Though R-indicator showing a steady upward trend, the slope was very small**
  - **Key survey estimates stabilized after 20+ weeks**

## Discussion (cont'd.)

---

- **Other measures may have supported the closeout of the data collection (cont'd.)**
  - Bias and RMSE upper bound measures showed consistent results with response rates, R-indicators, and survey-estimate tracking
- **Importance of tracking various measures during data collection—response rates, R-indicators, frame variables, key survey estimates, bias indicator measures**
- **The decision to close out data collection can be based on quality measures and other practical considerations: budget and data-dissemination schedule**

# For More Information

---

- **Please contact:**
  - **Donsig Jang**
    - [djang@mathematica-mpr.com](mailto:djang@mathematica-mpr.com)
  
  - **Sixia Chen**
    - [sixia@iastate.edu](mailto:sixia@iastate.edu)
  
  - **Flora Lan**
    - [flan@nsf.gov](mailto:flan@nsf.gov)

# References

---

- Groves, R.M., and E. Peytcheva. “The Impact of Nonresponse Rates on Nonresponse Bias.” *Public Opinion Quarterly*, vol. 72, 2008, pp. 1–23.
- Schouten, B., F. Cobben, and J. Bethlehem. “Indicators for the Representativeness of Survey Response.” *Survey Methodology*, vol. 35, 2009, pp. 101–113.
- Sarndal, C.E. “Three Factors to Signal Non-Response Bias with Applications to Categorical Auxiliary Variables.” *International Statistical Review*, vol. 79, 2011, pp. 233–254.