Assessing bias across health domains from two online, probability-based panel surveys: Examples from the National Center for Health Statistics Rapid Surveys System

Jim Dahlhamer, PhD

Katherine Irimata, PhD

2024 FCSM Research & Policy Conference

October 22-24, 2024

National Center for Health Statistics (NCHS) Rapid Surveys System (RSS)

- Platform that utilizes two commercial probability-based survey panels to collect data: NORC at the University of Chicago AmeriSpeak Panel and Ipsos' KnowledgePanel
- Produce relevant and timely estimates on emerging health topics
- Utilizes statistical weighting methods to leverage the strength of NCHS' core household surveys to improve the reliability of estimates
- Release of public use and restricted-access data, web tables of estimates, documentation, and evaluations of data quality
- RSS website: https://www.cdc.gov/nchs/rss/rapid-surveys-system.html

RSS Rounds 1-3 (1)

- Round 1 of RSS was fielded between August 1, 2023, and September 8, 2023
  - Air cleaners and purifiers, chemical hair products, genetic tests, long COVID, and sunscreen safety
- Round 2 of RSS was fielded between October 16, 2023, and November 7, 2023
  - ADHD, online connectedness, contraception access, swimming, technology-facilitated sexual violence, and hearing protector fit testing

RSS Rounds 1-3 (2)

- Round 3 of RSS was fielded between January 18, 2024, and February 15, 2024
  - Family health history, genetic testing for cancer and heart disease, and sexual health
- For rounds 1 and 2, data were collected from NORC and Ipsos using two different methodologies
  - Method 1: Standard data collection procedures
  - Method 2: Enhanced data collection procedures to study representativity

RSS Rounds 1-3 Sample Sizes, Completed Interviews, and Completion Rates

	RSS1	RSS2	RSS3
Sample size	18,307	18,928	21,112
# of completed interviews	7,599	7,046	8,375
Completion rate	41.5	37.2	39.7

Completion rate = (# of completed interviews / sample size) * 100

Weighting and Combining Methods

Methods: Weighting

- NORC and Ipsos developed final study weights at the conclusion of each round of data collection
- The two sets of panel weights were separately calibrated to control totals based on 2023 National Health Interview Survey (NHIS) Early Release (ER) data for adults
  - 2023/Q1 for RSS-1; 2023/Q2 for RSS-2; 2023/Q3 for RSS-3
- Calibration was performed using raking on selected sociodemographic, health, social limitation, and civic engagement variables using PROC WTADJUST in SUDAAN

RSS Rounds 1-3 Calibration Variables

Description	Source	Subgroups (Levels)
Age group	Panel Profiles	18-34, 35-49, 50-64, 65+
Sex	Panel Profiles	Male, Female
Race and Hispanic origin	Panel Profiles	Hispanic, non-Hispanic black, non-Hispanic white, non-Hispanic other
Educational attainment	Panel Profiles	Less than high school, High school degree equivalent, Some college or above
Household income	Panel Profiles	$0-$49,999; $50,000-$99,999; $100,000+
Region	Panel Profiles	Northeast, Midwest, South, West
Marital status	Questionnaire	Married, Not married
Housing tenure	Panel Profiles	Own or being bought, Rent/other arrangement
Urbanization level	Panel Profiles	Metropolitan, Nonmetropolitan
Ever diagnosed with high cholesterol	Questionnaire	Yes, No
Difficulty participating in social activities due to physical, mental, or emotional condition	Questionnaire	No difficulty/some difficulty; A lot of difficulty/cannot do this at all
Civic engagement	Questionnaire	Yes (attended a public meeting in past 12 months OR voted in last local elections OR volunteered for any organization in past 12 months), No

Methods: Combination

- To produce a single RSS estimate, estimates from NORC and Ipsos are combined using a convex combination of the effective sample sizes

$$\lambda_1 = \frac{n_{e,1}}{n_{e,1} + n_{e,2}} \qquad n_{e,i} = \frac{(\sum_s w_k)^2}{\sum_s w_k^2}$$

$$p = \lambda_1 \, p_1 + \, (1 - \lambda_1) \, p_2$$

$$se(p) = \sqrt{\left(\lambda_1 \, se(p_1)\right)^2 + \left((1 - \lambda_1) \, se(p_2)\right)^2}$$

	Adjustment Factor		
	RSS1	RSS2	RSS3
Panel Provider 1	.271	.261	.391
Panel Provider 2	.729	.739	.609

Bias Evaluation

Evaluation

- 136 health outcomes from RSS Rounds 1-3 were included as benchmark variables to assess estimates compared to the NHIS (ER)
  - RSS-1 (n=39); RSS-2 (n=45); RSS-3 (n=52)

- Individual and combined estimates were evaluated to assess data quality and the benefits of combining data across probability-based panel surveys

- Differences were evaluated using absolute bias:

$$Absolute\ bias = |RSS\ estimate\ - NHIS\ estimate|$$

and standardized bias:

$$Standardized\ bias = \frac{|RSS\ estimate\ - NHIS\ estimate|}{\sqrt{NHIS\ estimate\ * (100 - NHIS\ estimate)}}$$

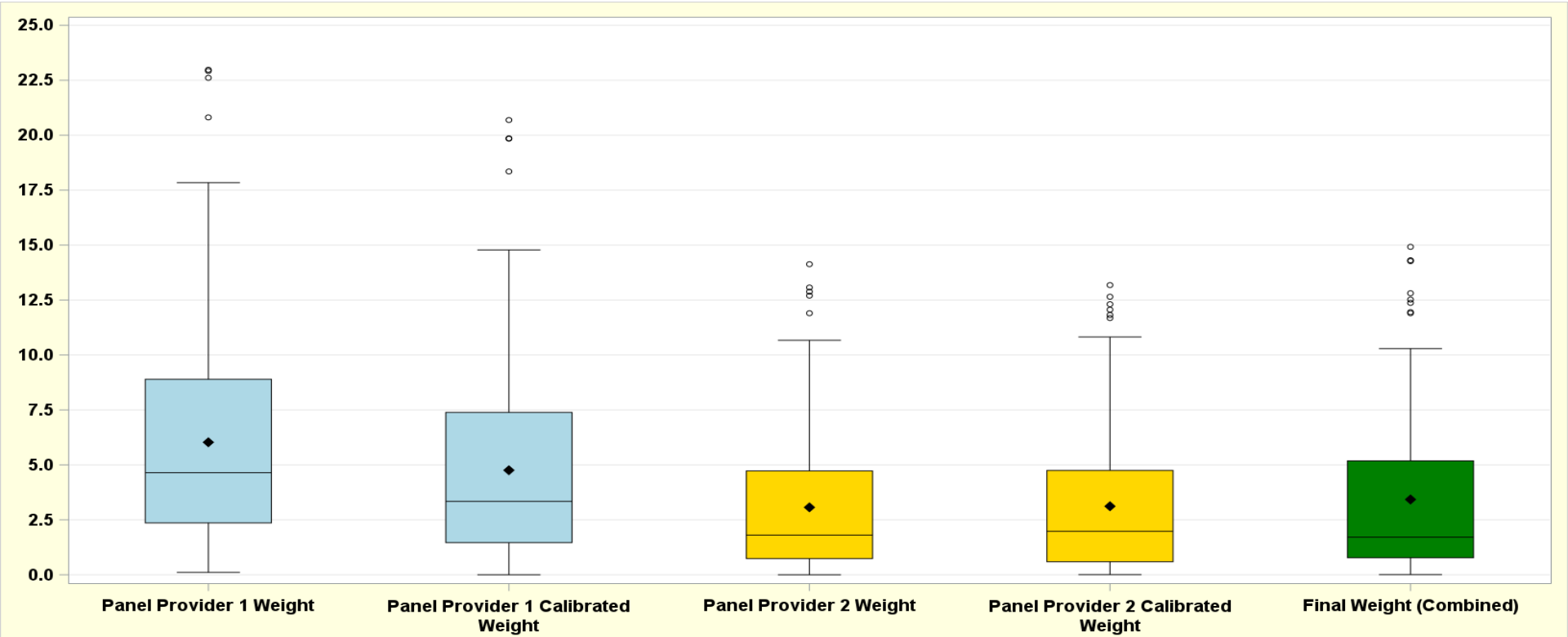
Benchmarks

- 136 benchmark estimates were organized into 7 health domains:

Health status: chronic conditions (n=19) Ex: Ever diagnosed with cancer	Healthcare access (n=29) Ex: Usual source of care
Health status: disability and work days missed due to illness or injury (n=11) Ex: Vision difficulty	Healthcare utilization (n=28) Ex: Doctor visit in past 12 months
Health status: mental and self-rated health (n=21) Ex: Feel depressed daily/weekly	Food security, trouble paying housing costs, lack of reliable transportation (n=16) Ex: Very low food security
Health behaviors (n=12) Ex: Current cigarette smoking	

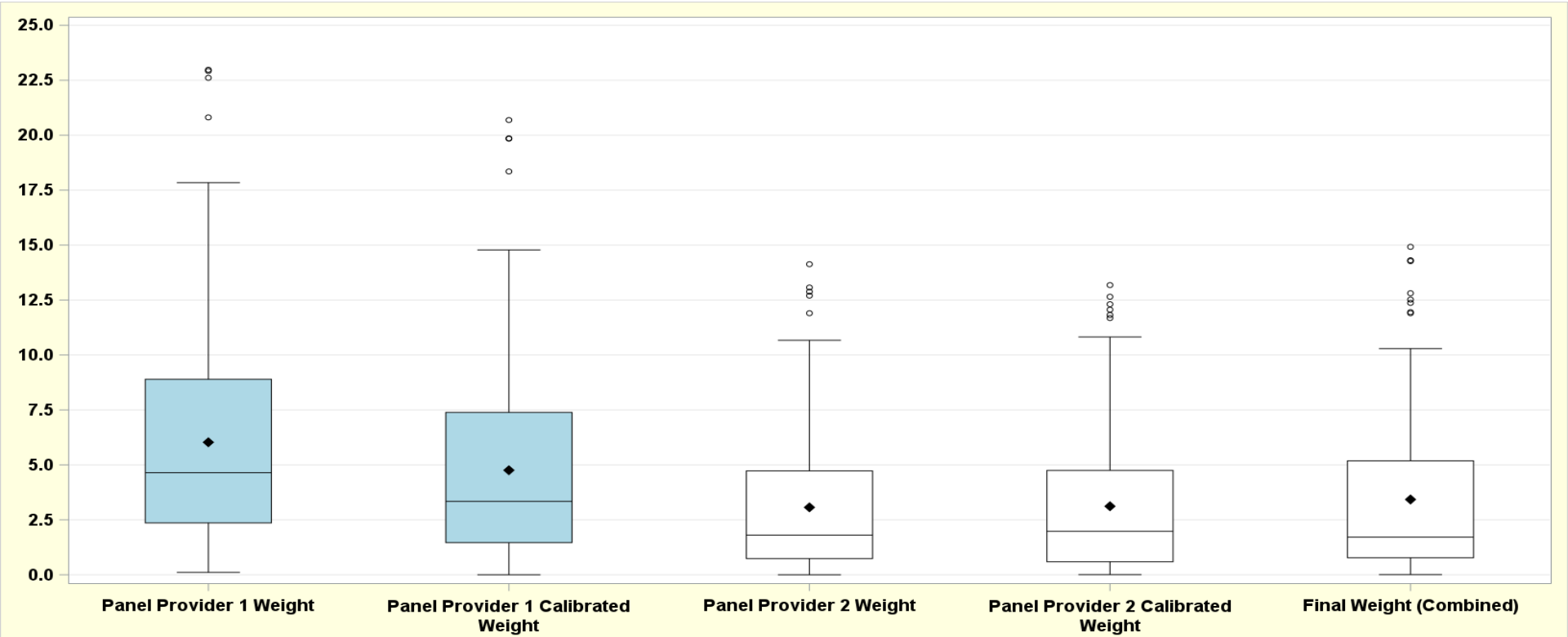
  - 13 benchmark estimates were repeated across all 3 rounds

Distribution of Absolute Bias for 136 RSS Round 1-3 Benchmark Estimates, by Panel Provider and Weight (1)



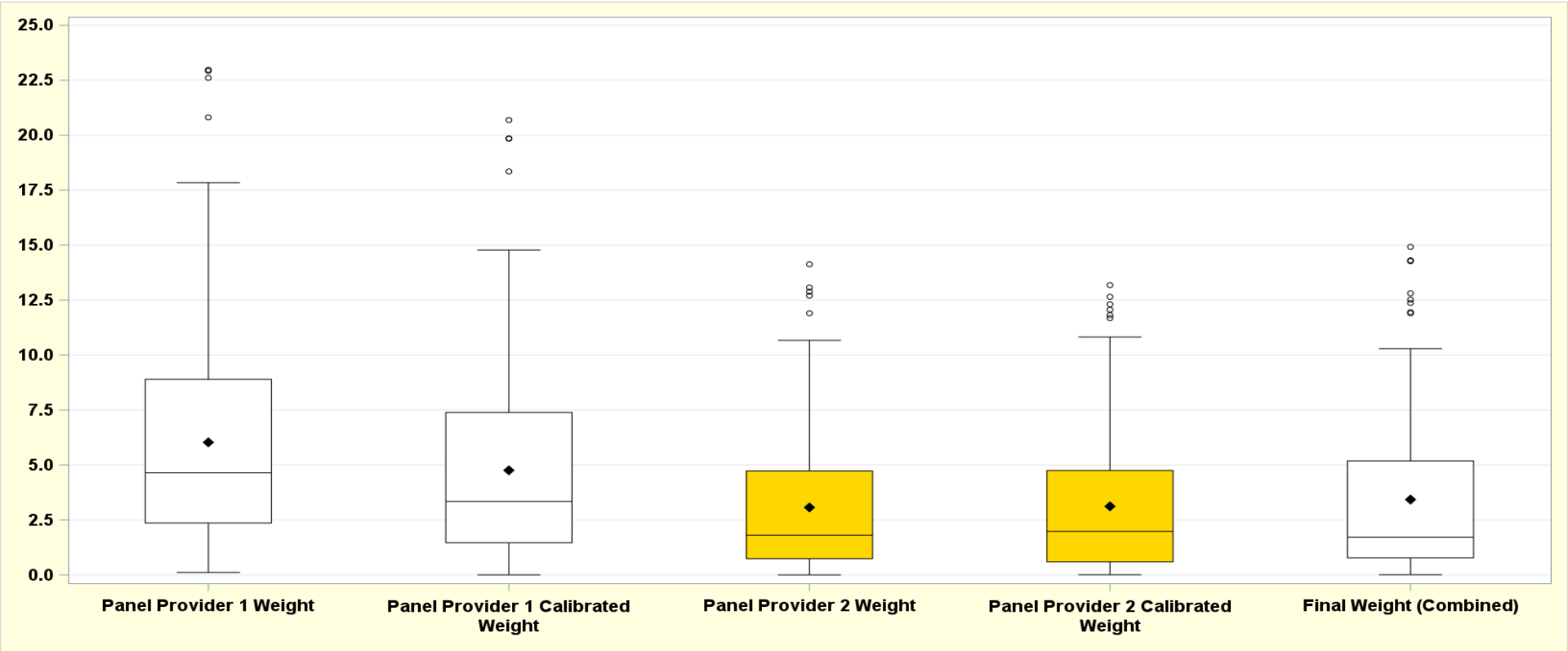
Mean:	6.03	4.76	3.07	3.12	3.43
Median:	4.65	3.34	1.81	1.98	1.71

Distribution of Absolute Bias for 136 RSS Round 1-3 Benchmark Estimates, by Panel Provider and Weight (2)



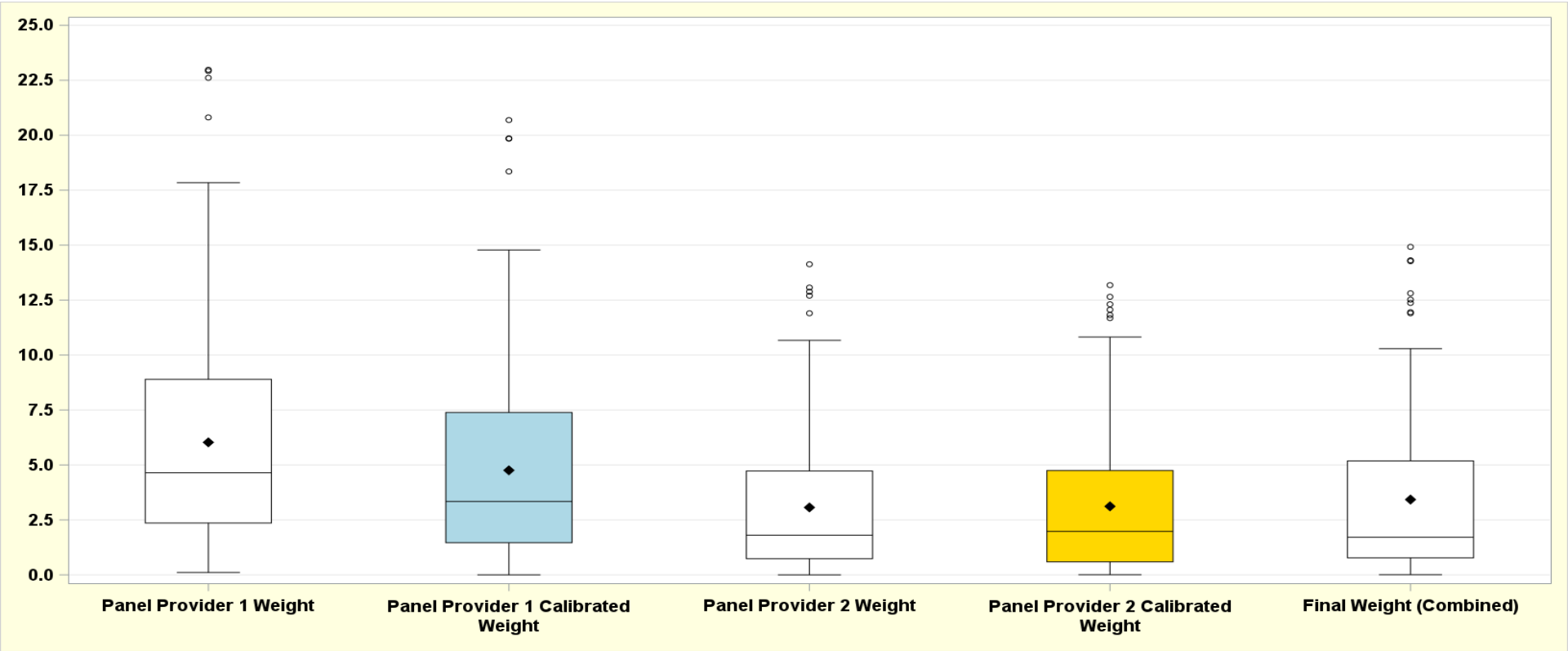
Mean:	6.03	4.76	3.07	3.12	3.43
Median:	4.65	3.34	1.81	1.98	1.71

Distribution of Absolute Bias for 136 RSS Round 1-3 Benchmark Estimates, by Panel Provider and Weight (3)



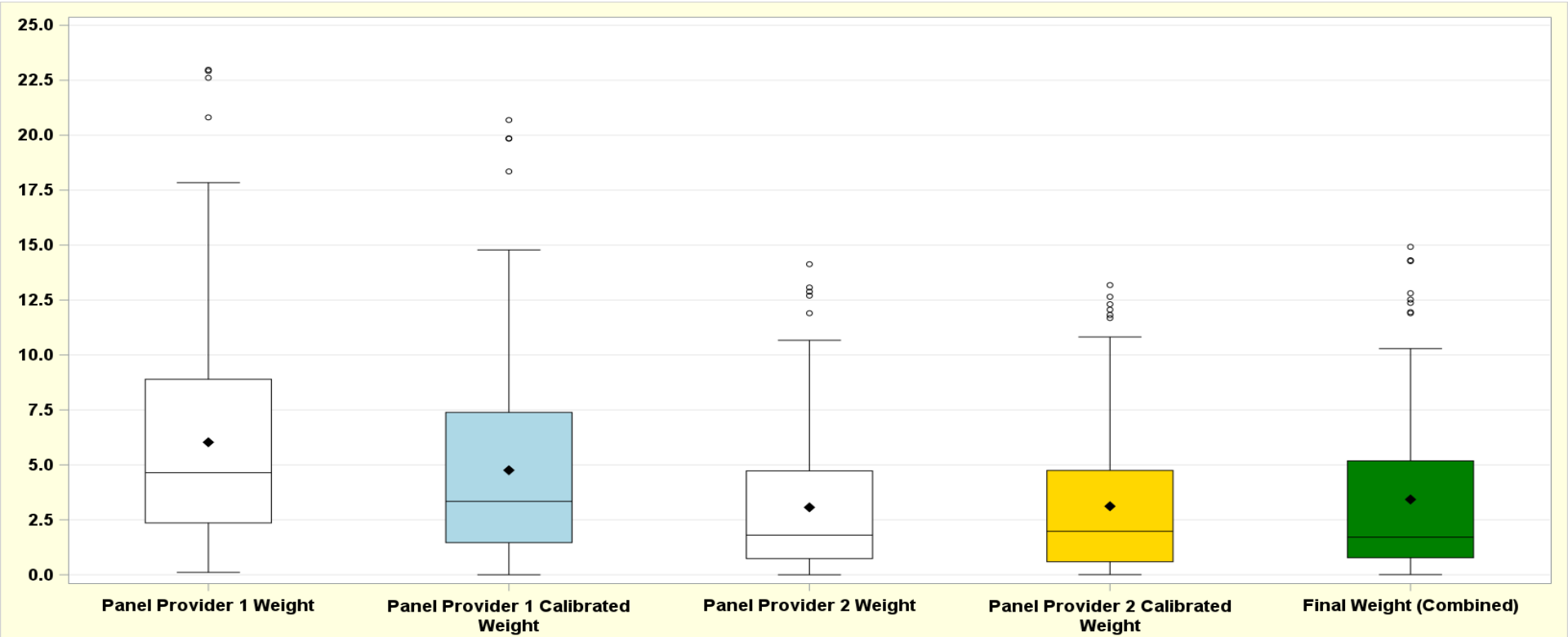
	Panel Provider 1 Weight	Panel Provider 1 Calibrated Weight	Panel Provider 2 Weight	Panel Provider 2 Calibrated Weight	Final Weight (Combined)
Mean:	6.03	4.76	3.07	3.12	3.43
Median:	4.65	3.34	1.81	1.98	1.71

Distribution of Absolute Bias for 136 RSS Round 1-3 Benchmark Estimates, by Panel Provider and Weight (4)



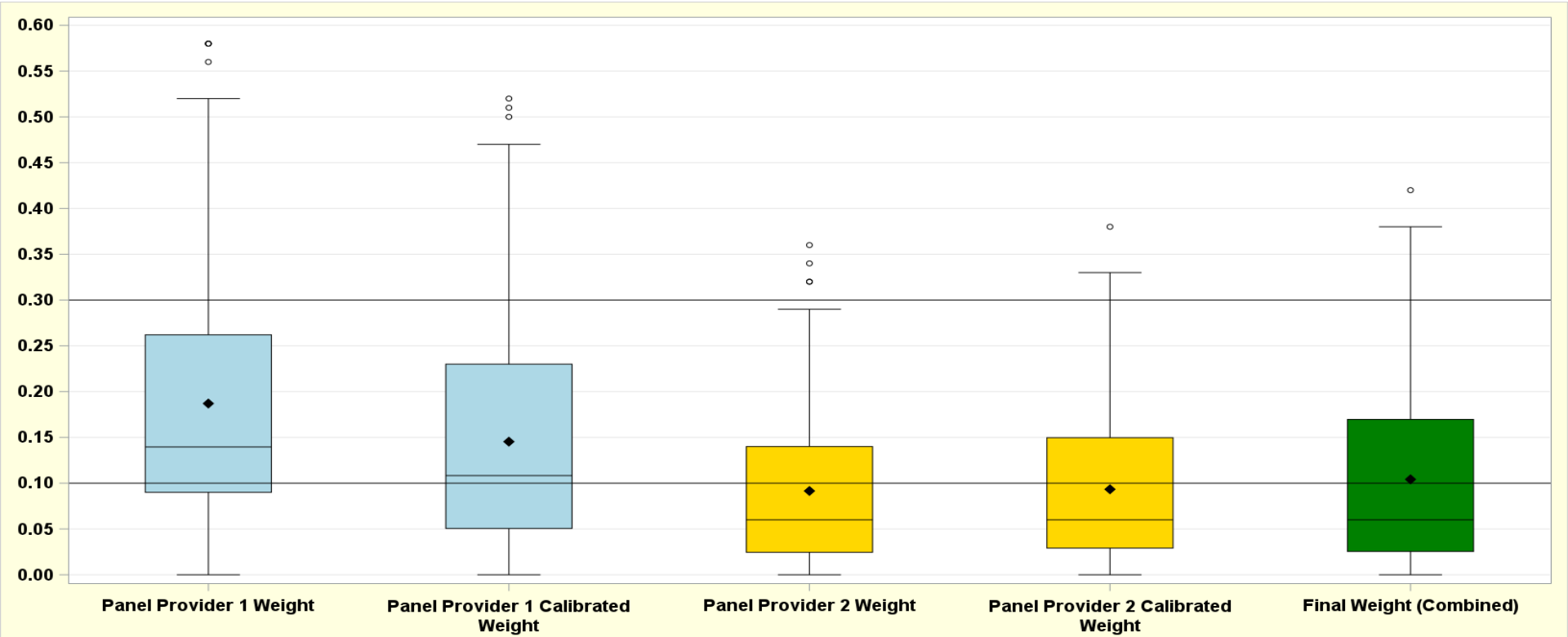
	Panel Provider 1 Weight	Panel Provider 1 Calibrated Weight	Panel Provider 2 Weight	Panel Provider 2 Calibrated Weight	Final Weight (Combined)
Mean:	6.03	4.76	3.07	3.12	3.43
Median:	4.65	3.34	1.81	1.98	1.71

Distribution of Absolute Bias for 136 RSS Round 1-3 Benchmark Estimates, by Panel Provider and Weight (5)



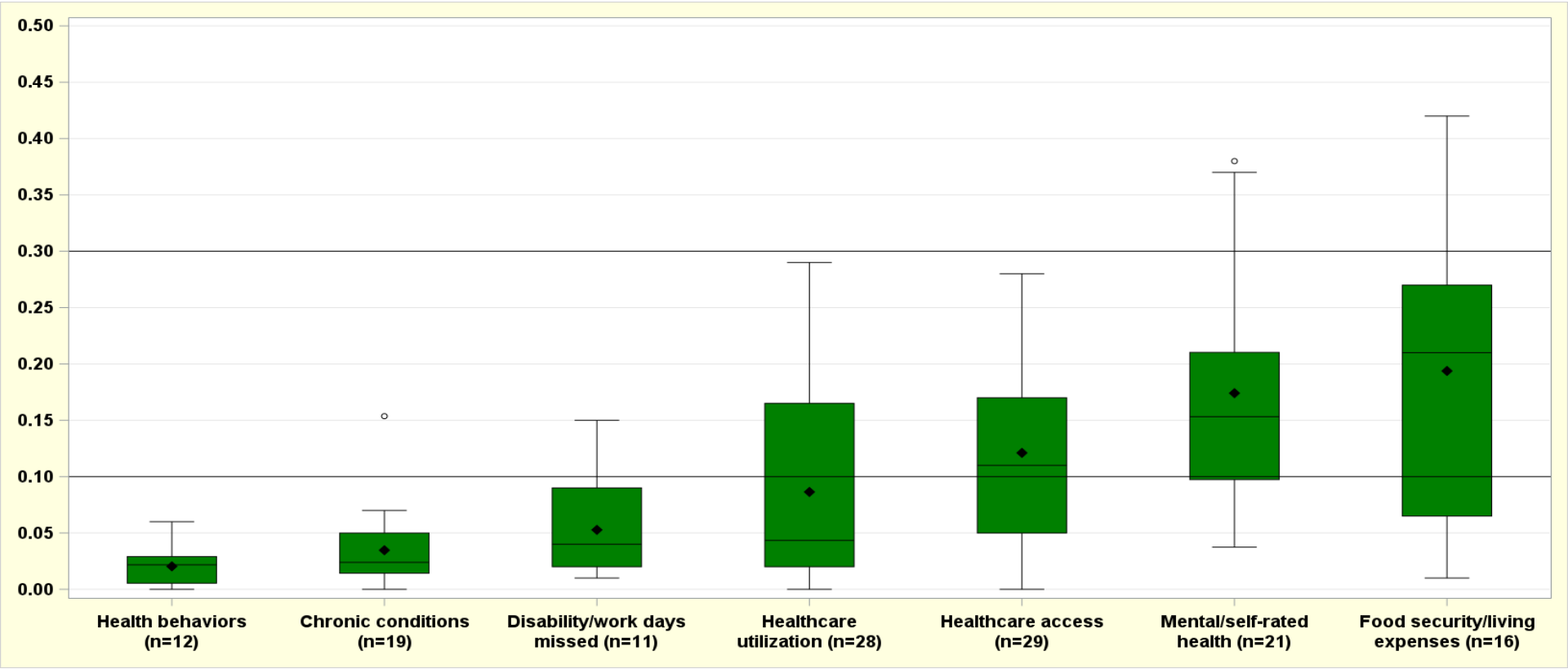
	Panel Provider 1 Weight	Panel Provider 1 Calibrated Weight	Panel Provider 2 Weight	Panel Provider 2 Calibrated Weight	Final Weight (Combined)
Mean:	6.03	4.76	3.07	3.12	3.43
Median:	4.65	3.34	1.81	1.98	1.71

Distribution of Standardized Bias for 136 RSS Round 1-3 Benchmark Estimates, by Panel Provider and Weight



	Panel Provider 1 Weight	Panel Provider 1 Calibrated Weight	Panel Provider 2 Weight	Panel Provider 2 Calibrated Weight	Final Weight (Combined)
Mean:	0.19	0.15	0.09	0.09	0.10
Median:	0.14	0.11	0.06	0.06	0.06

Distribution of Standardized Bias for 136 Benchmark Estimates by Health Domain: RSS Rounds 1-3 (combined, final weight)



	Health behaviors (n=12)	Chronic conditions (n=19)	Disability/work days missed (n=11)	Healthcare utilization (n=28)	Healthcare access (n=29)	Mental/self-rated health (n=21)	Food security/living expenses (n=16)
Mean:	0.02	0.03	0.05	0.09	0.12	0.17	0.19
Median:	0.02	0.02	0.04	0.04	0.11	0.15	0.21

Distribution of 136 Benchmark Estimates by Health Domain and Standardized Bias: RSS Rounds 1-3 (combined, final weight)

Health Domain	Bias					
	Low (< 0.10)		Medium (0.10 - < 0.30)		High (>= 0.30)	
Health behaviors	12	100.0	0	0.0	0	0.0
Chronic conditions	18	94.7	1	5.3	0	0.0
Disability/work days missed	9	81.8	2	18.2	0	0.0
Healthcare utilization	18	64.3	10	35.7	0	0.0
Healthcare access	13	44.8	16	55.2	0	0.0
Mental and self-rated health	6	28.6	12	57.1	3	14.3
Food security/housing costs/trans.	5	31.3	8	50.0	3	18.7
TOTAL	81	59.6	49	36.0	6	4.4

Distribution of 121 Benchmark Estimates by Health Domain and Standardized Bias: RSS Rounds 1-3 (combined, final weight)

Health Domain	Bias					
	Low (< 0.10)		Medium (0.10 - < 0.30)		High (>= 0.30)	
Health behaviors	12	100.0	0	0.0	0	0.0
Chronic conditions	18	94.7	1	5.3	0	0.0
Disability/work days missed	9	81.8	2	18.2	0	0.0
Healthcare utilization	18	64.3	10	35.7	0	0.0
Healthcare access	13	44.8	16	55.2	0	0.0
Mental and self -rated health	6	28.6	12	57.1	3	14.3
Food security/housing costs/trans.	5	31.3	8	50.0	3	18.7
TOTAL	81	59.6	49	36.0	6	4.4

Distribution of 121 Benchmark Estimates by Health Domain and Standardized Bias: RSS Rounds 1-3 (combined, final weight)

Health Domain	Bias					
	Low (< 0.10)		Medium (0.10 - < 0.30)		High (>= 0.30)	
Health behaviors	12	100.0	0	0.0	0	0.0
Chronic conditions	18	94.7	1	5.3	0	0.0
Disability/work days missed	9	81.8	2	18.2	0	0.0
Healthcare utilization	18	64.3	10	35.7	0	0.0
Healthcare access	13	44.8	16	55.2	0	0.0
Mental and self-rated health	6	28.6	12	57.1	3	14.3
Food security/housing costs/trans.	5	31.3	8	50.0	3	18.7
TOTAL	81	59.6	49	36.0	6	4.4

Summary/Future Directions (1)

- Among 136 benchmark estimates evaluated, 81 had a low standardized bias, 49 had a medium standardized bias, and 6 had a high standardized bias compared to the NHIS

- The average standardized bias of estimates varied by health domain
  - Chronic health conditions, health behaviors, and disability tended to have the lowest average standardized bias
  - Mental and self-rated health and food security had the highest
    - Among domains with higher standardized bias, on average, considerable variability
    - Question sensitivity--mode effects?

Summary/Future Directions (2)

- Evaluations of data from rounds 1-3 led to a change in the mix of NHIS calibration variables for round 4
  - Continue to evaluate calibration variables to improve estimates/reduce bias for health domains with higher bias
  - Consider a different mix of calibration variables for each provider?
  - Consider different weighting and combining methods?
- Release of round 4 data is tomorrow, with release of round 5 data anticipated by end of the year

Thank you!

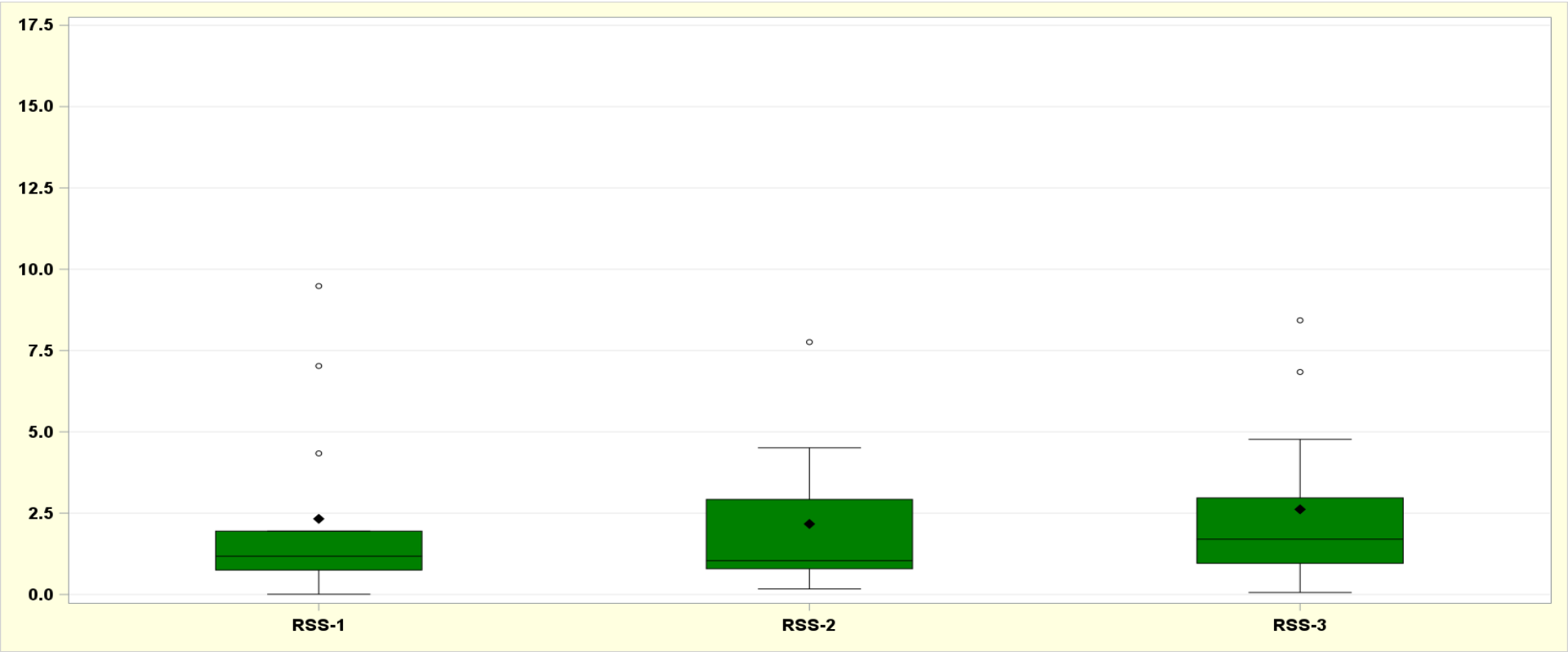
jmdahlhamer@cdc.gov

RSS website: https://www.cdc.gov/nchs/rss/rapid-surveys-system.html

For more information, contact CDC
1-800-CDC-INFO (232-4636)
TTY: 1-888-232-6348 www.cdc.gov

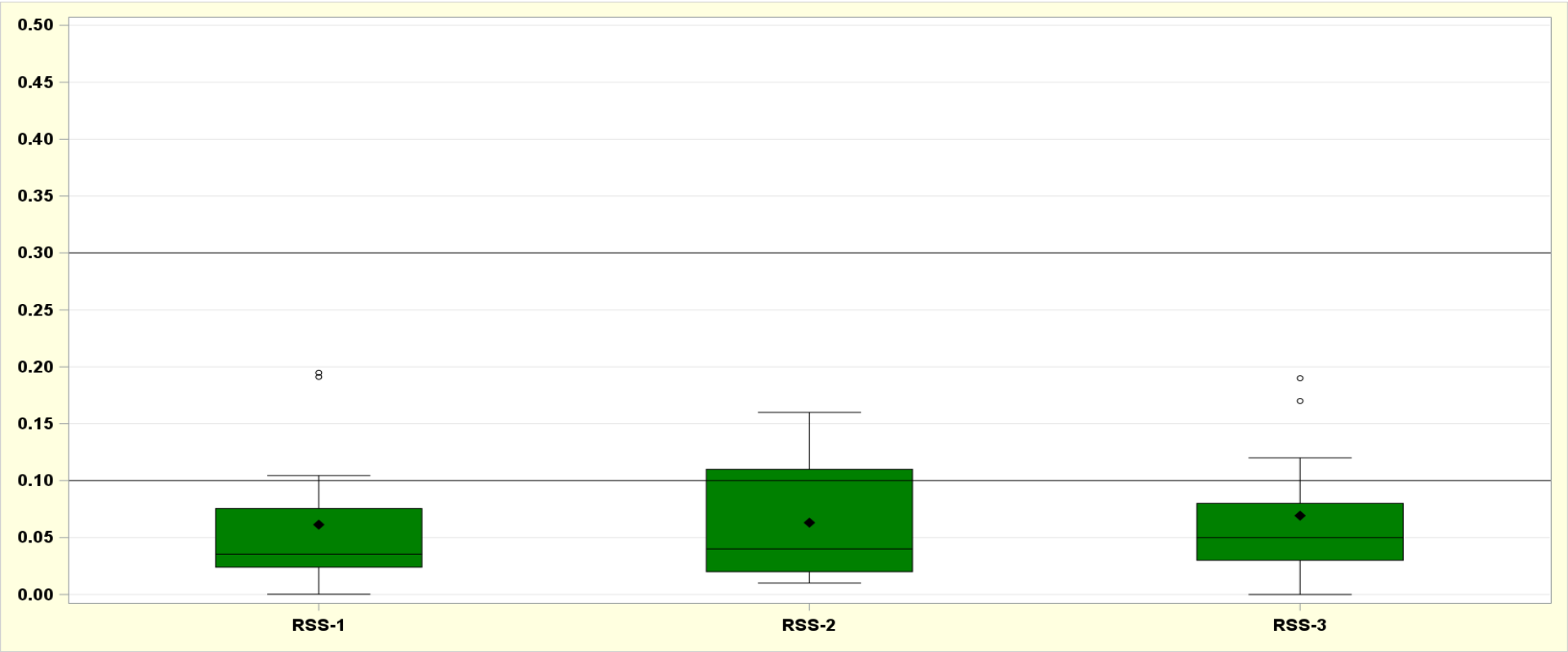
Reserve Slides

Distribution of Absolute Bias for 13 Benchmark Estimates Included on RSS Rounds 1-3, by Round (combined, final weight)



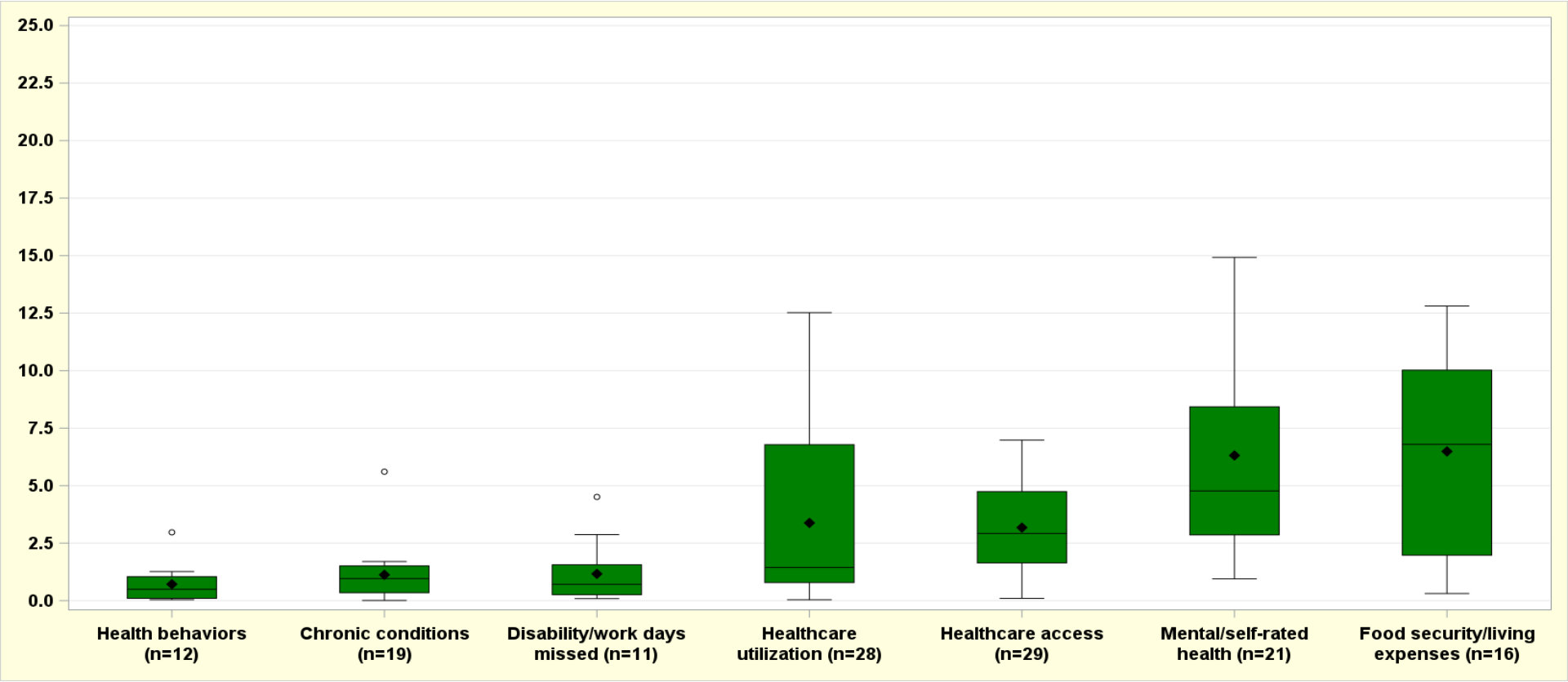
	RSS-1	RSS-2	RSS-3
Mean:	2.33	2.17	2.62
Median:	1.18	1.04	1.70

Distribution of Standardized Bias for 13 Benchmark Estimates Included on RSS Rounds 1-3, by Round (combined, final weight)



	RSS-1	RSS-2	RSS-3
Mean:	0.06	0.06	0.07
Median:	0.04	0.04	0.05

Distribution of Absolute Bias for 136 Benchmark Estimates by Health Domain: RSS Rounds 1-3 (combined, final weight)



	Health behaviors (n=12)	Chronic conditions (n=19)	Disability/work days missed (n=11)	Healthcare utilization (n=28)	Healthcare access (n=29)	Mental/self-rated health (n=21)	Food security/living expenses (n=16)
Mean:	0.72	1.12	1.16	3.38	3.18	6.31	6.49
Median:	0.50	0.96	0.71	1.44	2.92	4.77	6.80

Background: National Health Interview Survey (NHIS)

- **Purpose**: To monitor the health of the US population through the collection and analysis of data on a broad range of health topics

- **Sample**: Address-based, multi-stage, clustered national sample of housing units from every state, to be representative of the civilian noninstitutionalized US population

- **Mode**: In-person interviews by Census interviewers, with telephone interviewing as needed

- **Data collection**: Continuous, with quarterly and annual data files

- **Target sample size**: Complete interviews for 35,000+ households