



NHANES pre-pandemic 2017-2020 sample weights

NCHS BSC Meeting, January 27, 2021

Cynthia L. Ogden, PhD

Stats Team: Te-Ching Chen, PhD, Orlando Davy, MPH, Steven Fink, MA

Outline

- Development of pre-pandemic 2017-2020 file and weights
- Evaluation/comparison of estimates to:
 - ACS
 - NHIS
 - Previous NHANES cycles
- Next steps

Combining 2017-2018 and partial 2019-2020

2019-2020

- Not complete – not representative or usable alone
 - 18 PSUs completed
 - 12 PSUs missed
- 2019: first time single year not nationally representative
- To compensate for missing PSUs
 - Data from completed 2019-2020 PSUs combined with 2017-2018

Pre-pandemic 2017-2020 weights require combining data from 2 sample designs

- Strata different in 2017-2018 and 2019-2020
 - 2017-2018 from 2015-2018 sample design
 - 2019-2020 from 2019-2022 sample design

Because 2 cycles come from different designs (strata), choose either:

- Map PSUs from 2017-2018 to 2019-2022 strata
 - **Not possible:** Incomplete 2019-2020. Not all 2019-2022 strata have PSUs even with PSUs from 2017-2018.

OR

- Map completed 2019-2020 PSUs to 2015-2018 strata
 - Pre-pandemic 2017-2020 data treated as extended 2017-2018 data. All 2015-2018 strata have PSUs

Pre-pandemic 2017-2020 PSUs per strata

2015-2018 Strata

	Cer- tainty	Healthiest				2 nd healthiest			3 rd healthiest				Least healthy		
		A1	A2	A3	A4	B1	B2	B3	C1	C2	C3	C4	D1	D2	D3
2017-2018	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

Pre-pandemic 2017-2020 PSUs per strata

2015-2018 Strata

	Cer- tainty	Healthiest				2 nd healthiest			3 rd healthiest				Least healthy		
		A1	A2	A3	A4	B1	B2	B3	C1	C2	C3	C4	D1	D2	D3
2017-2018	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2019-2020 completed	1	1	1			1		2	2	1	1	4	1	3	

2 PSUs
In healthiest

4 PSUs in least
healthy

Pre-pandemic 2017-2020 PSUs per strata

2015-2018 Strata

	Cer- tainty	Healthiest				2 nd healthiest			3 rd healthiest				Least healthy		
		A1	A2	A3	A4	B1	B2	B3	C1	C2	C3	C4	D1	D2	D3
2017-2018	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2019-2020 completed	1	1	1			1		2	2	1	1	4	1	3	
Pre-pandemic 2017-2020	3	3	3	2	2	3	2	4	4	3	3	6	3	5	2

PSU adjustment factor for pre-pandemic 2017-2020

- Number PSUs per stratum for pre-pandemic 2017-2020 was 2-6 (last slide)
 - Should have been 4 based on sample design so PSU factor created:

$$\text{PSU adjusting factor} = \frac{4}{\text{\# of PSUs fielded for the major stratum}}$$

# of PSUs per stratum	Adjustment factor
6	0.67
5	0.80
4	1.00
3	1.33
2	2.00

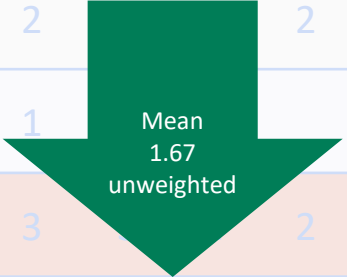
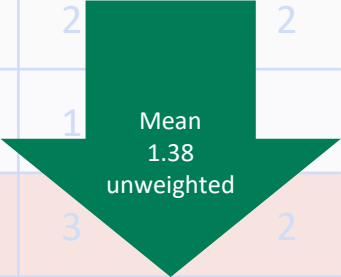
Pre-pandemic 2017-2020 PSUs per strata

2015-2018 Strata

	Cer- tainty	Healthiest				2 nd healthiest			3 rd healthiest				Least healthy		
		A1	A2	A3	A4	B1	B2	B3	C1	C2	C3	C4	D1	D2	D3
2017-2018	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2019-2020 completed	1	1	1			1		2	2	1	1	4	1	3	
Pre-pandemic 2017-2020	3	3	3	2	2	3	2	4	4	3	3	6	3	5	2
Adjustment factor	1.33	1.33	1.33	2	2	1.33	2	1	1	1.33	1.33	0.67	1.33	0.8	2

Pre-pandemic 2017-2020 PSUs per strata

2015-2018 Strata

	Cer- tainty	Healthiest				2 nd healthiest			3 rd healthiest				Least healthy		
		A1	A2	A3	A4	B1	B2	B3	C1	C2	C3	C4	D1	D2	D3
2017-2018	2	2			2	2	2	2	2	2	2	2	2		2
2019-2020 completed	1	1	 Mean 1.67 unweighted				1	2	2	1	1	4	1	 Mean 1.38 unweighted	
Pre-pandemic 2017-2020	3	3					2	3	2	4	4	3	3		
Adjustment factor	1.33	1.33	1.33	2	2	1.33	2	1	1	1.33	1.33	0.67	1.33	0.8	2

Evaluating the sample weights

Weighting adjustments beyond those done for 2017-2018*

	2017-2020 W1
Added 18 PSUs from 2019-20 to the 2017-18 data collection	X
Created factors to adjust base weights for 48 completed PSUs in a 60PSU sample design	X

*non-response adjustments and raking for demographics, education and income

Compare demographic estimates from ACS

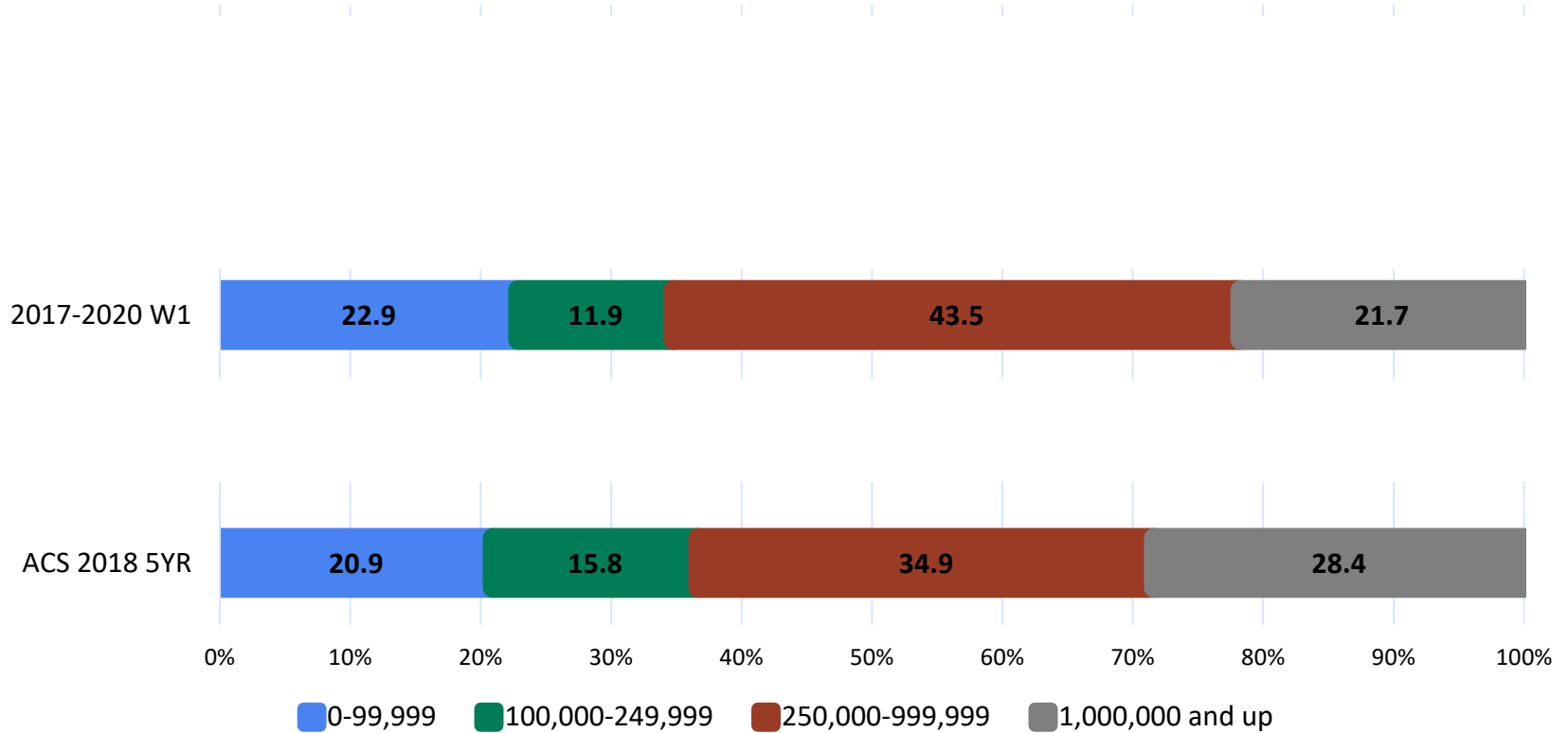
- No population size or urban/rural adjustments for 2017-2018
- Because PSU adjustment factors for 2017-2020, compare weighted sample to ACS on:
 - Population size
 - Urban/rural distribution

W1

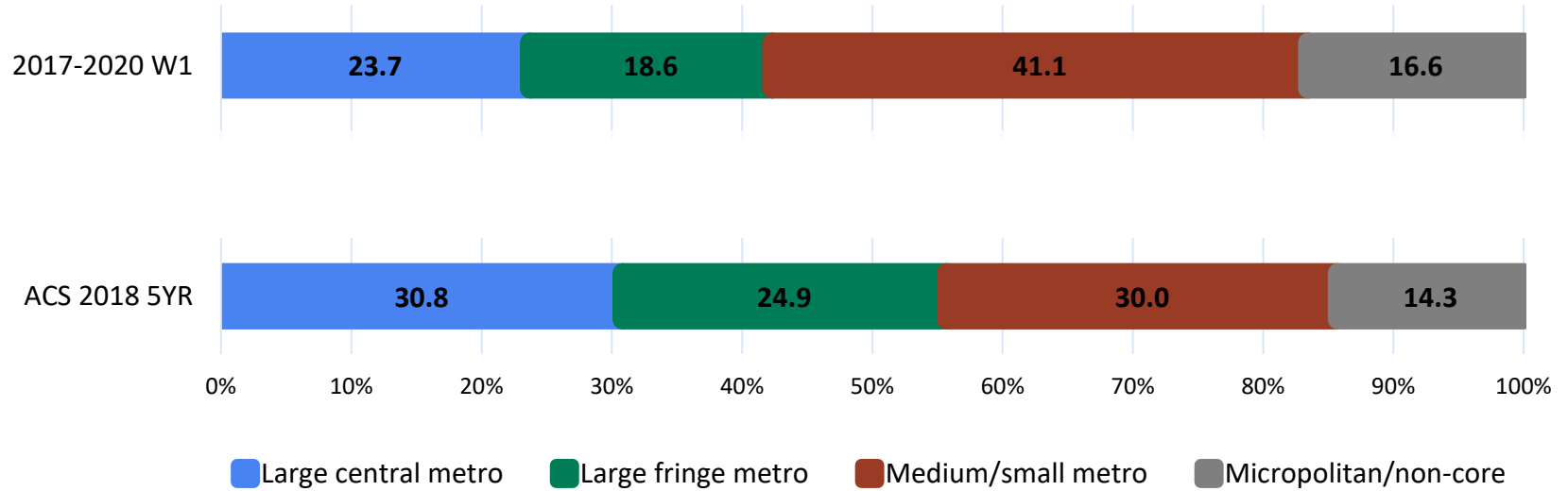
- NHANES pre-pandemic 2017-2020 differed from ACS
 - Population size
 - Urban/rural distribution

NHANES vs. ACS

Population Size Distribution



NHANES vs. ACS Urban/Rural Distribution



Weighting adjustments beyond those done for 2017-2018*

	2017-2020 W1	2017-2020 W2
Added 18 PSUs from 2019-20 to the 2017 -18 data collection	X	X
Created factors to adjust base weights for 48 completed PSUs in a 60PSU sample design	X	X
Raking to 4 PSU pop size		X

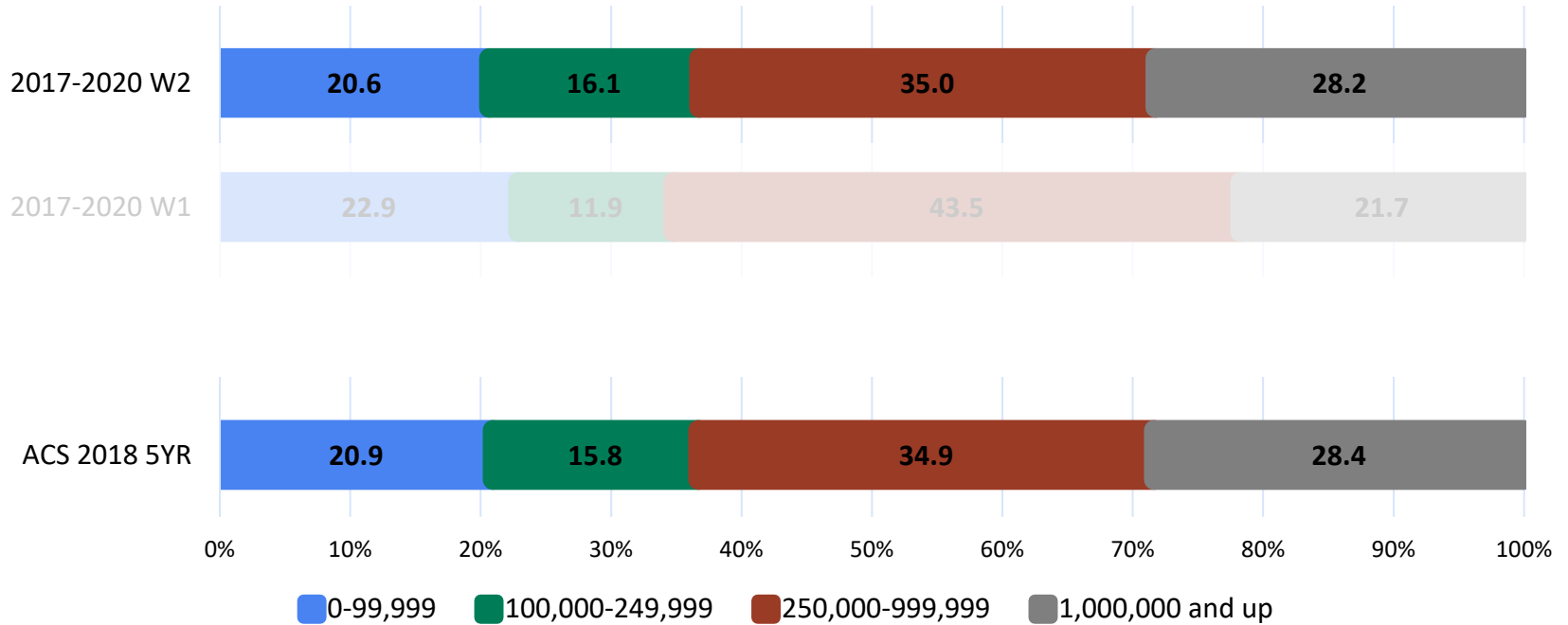
*non-response adjustments and raking for demographics, education and income

W2

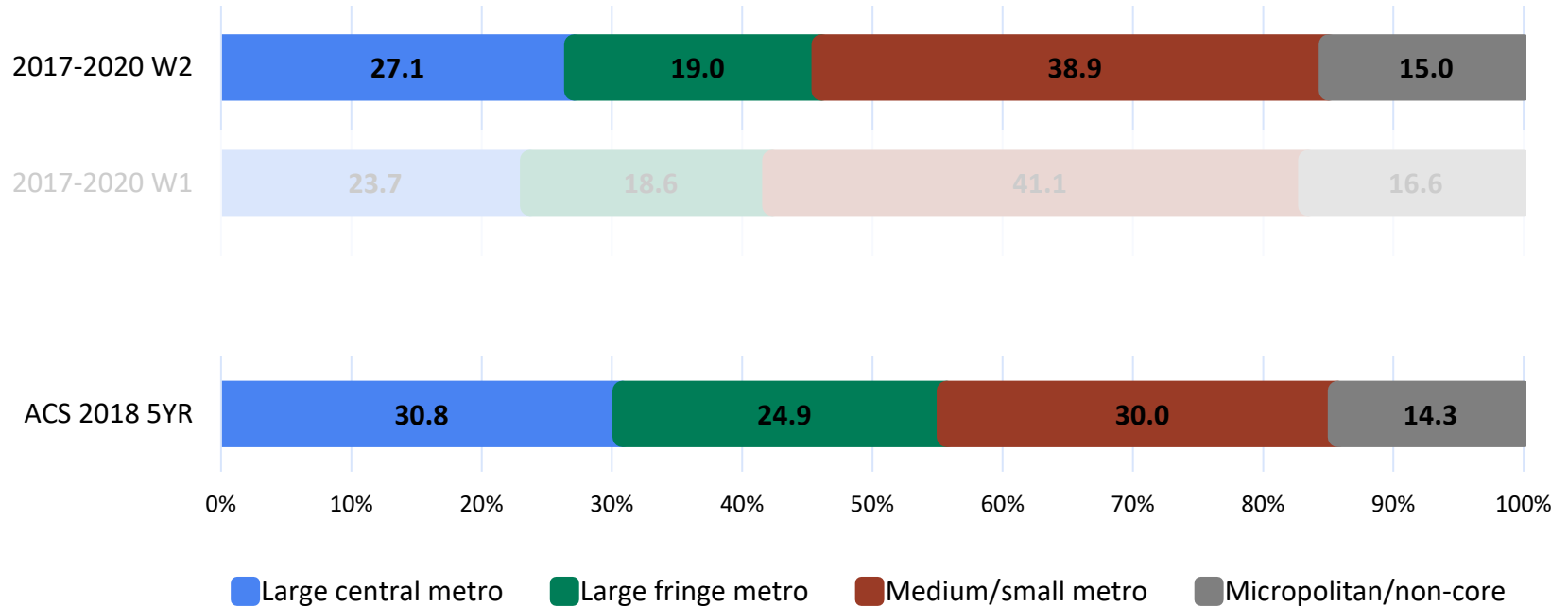
- NHANES pre-pandemic 2017-2020 compared to ACS
 - Better match to ACS on population size than W1
 - Urban/rural distribution continued to differ

NHANES vs. ACS

Population Size Distribution



NHANES vs. ACS Urban/Rural Distribution



Weighting adjustments beyond those done for 2017-2018*

	2017-2020 W1	2017-2020 W2	2017-2020 W3
Added 18 PSUs from 2019-20 to the 2017-18 data collection	X	X	X
Created factors to adjust base weights for 48 completed PSUs in a 60PSU sample design	X	X	X
Raking to 4 PSU pop size		X	
Raking to NCHS rural-urban codes (4 levels)			X

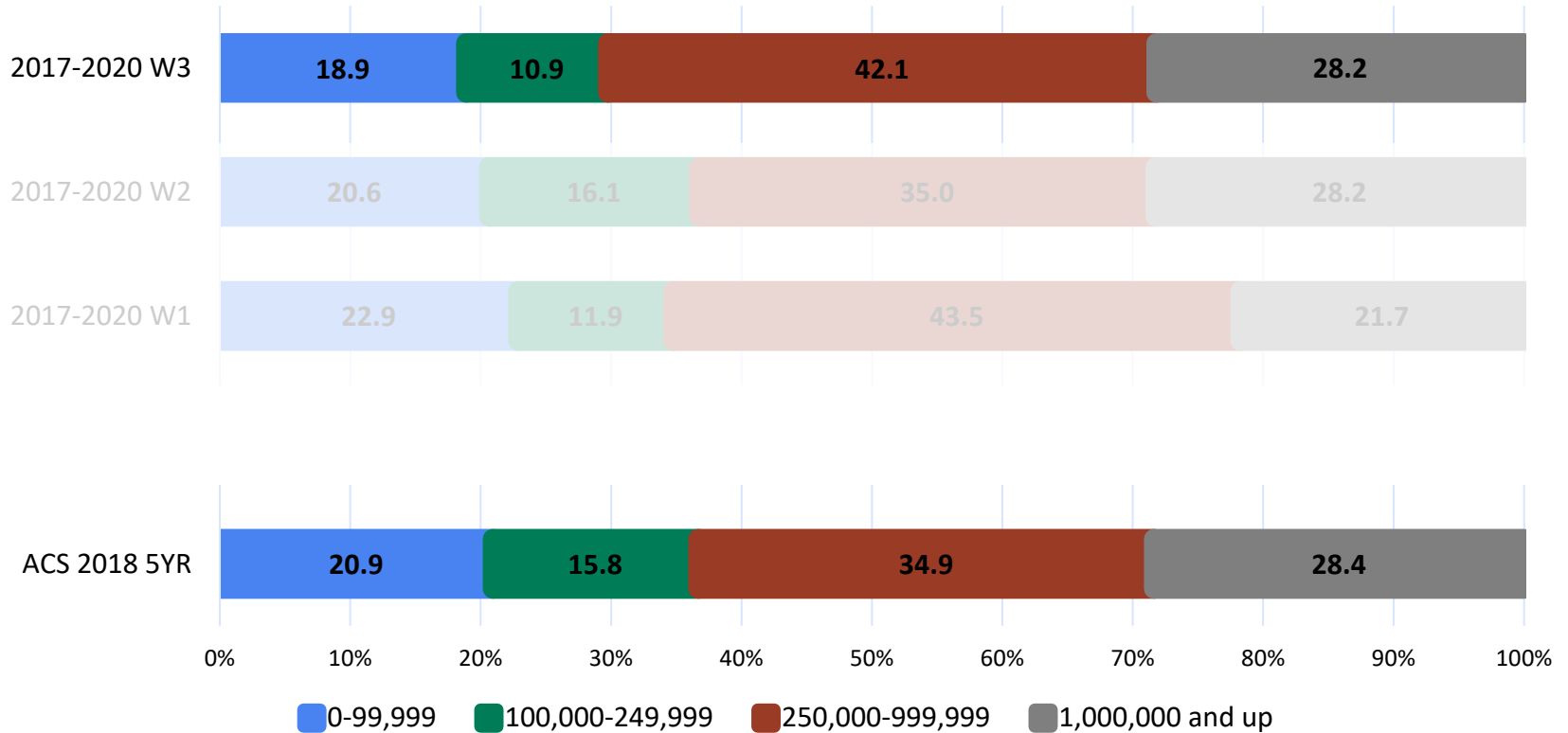
*non-response adjustments and raking for demographics, education and income

W3

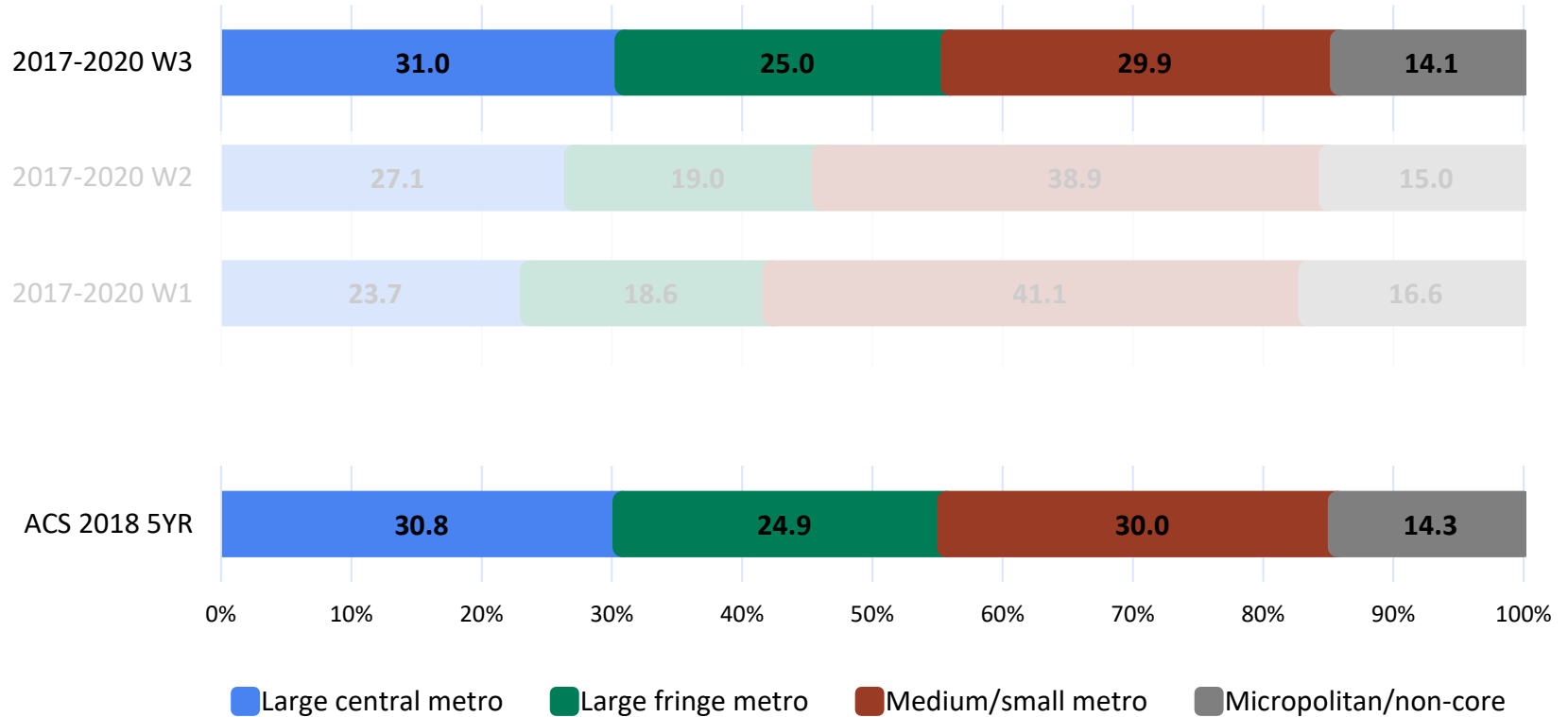
- NHANES pre-pandemic 2017-2020 compared to ACS
 - Population size distribution deviated further from ACS
 - Better match on urban/rural distribution compared to W1 and W2

NHANES vs. ACS

Population Size Distribution



NHANES vs. ACS Urban/Rural Distribution



Additional comparisons to ACS

- Education
- Poverty income ratio (PIR)
- Household Composition
- Marital Status
- Health Insurance Coverage

ACS versus NHANES

- Differences for education, PIR, household composition, marital status and health insurance
 - Similar for W2 and W3
 - Similar to differences between ACS and previous NHANES cycles

Comparison to NHIS

- Self-Reported Health Status
- Health Insurance
- Current Smoking
- Diagnosed diabetes

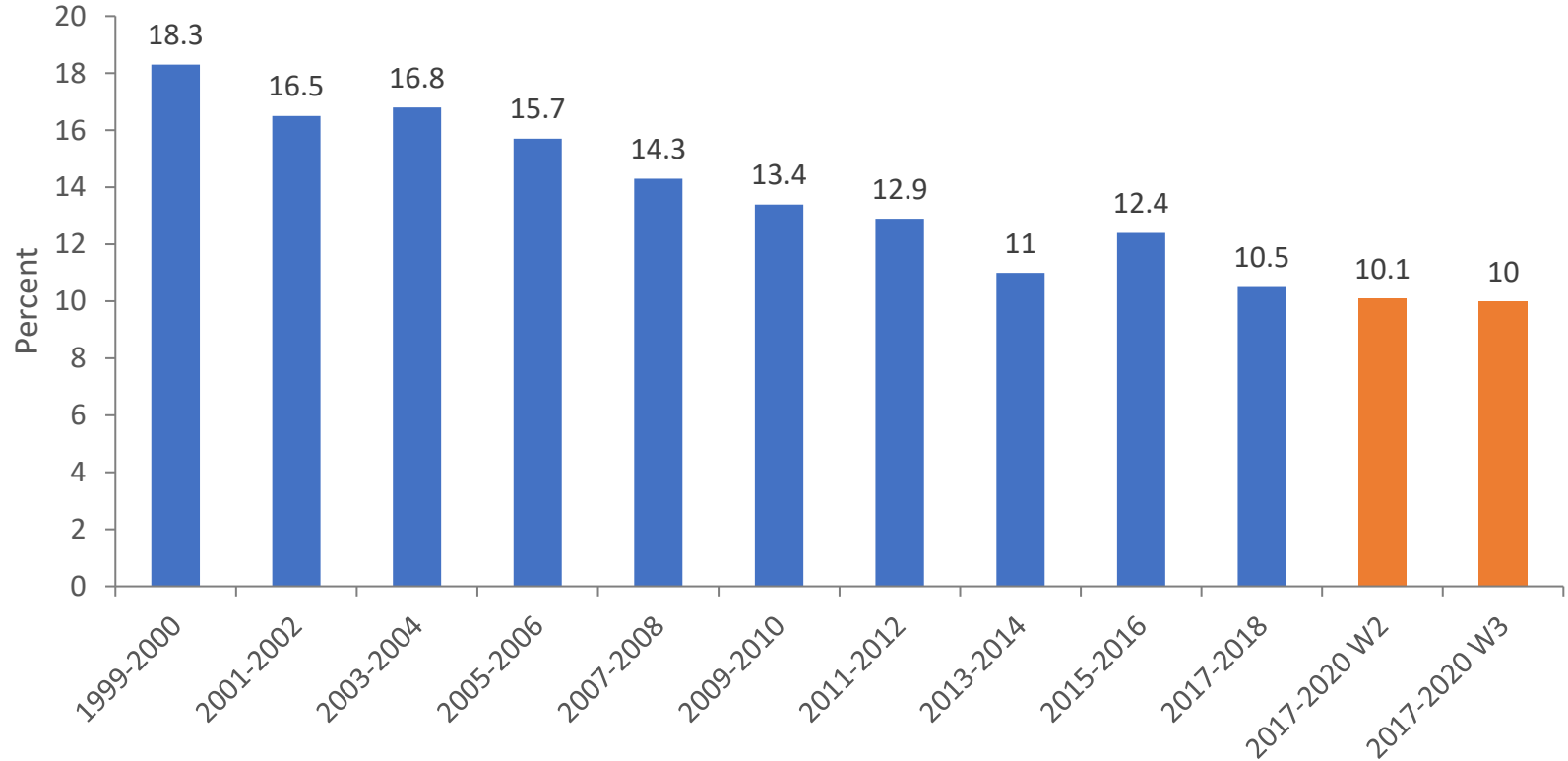
NHIS versus NHANES

- Consistent differences over time for 4 outcomes examined
 - Regardless of W2 or W3

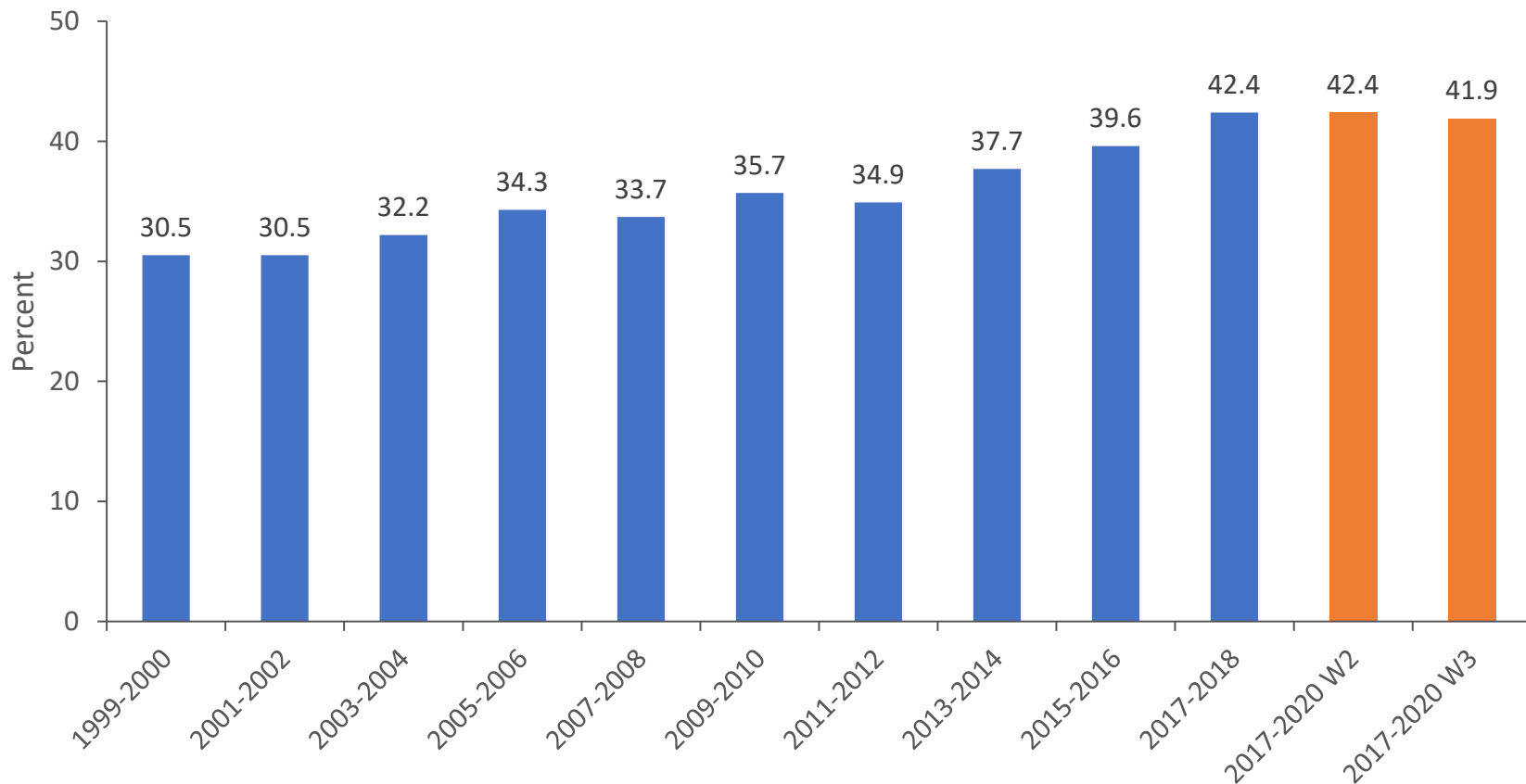
Comparison to previous NHANES

- Diagnosed diabetes
- Low HDL-cholesterol
- Hypertension
- Dental caries (youth)
- Tooth loss (65+)
- High total cholesterol
- Obesity (adults and youth)

High total cholesterol (age-adjusted) adults 20 +: NHANES

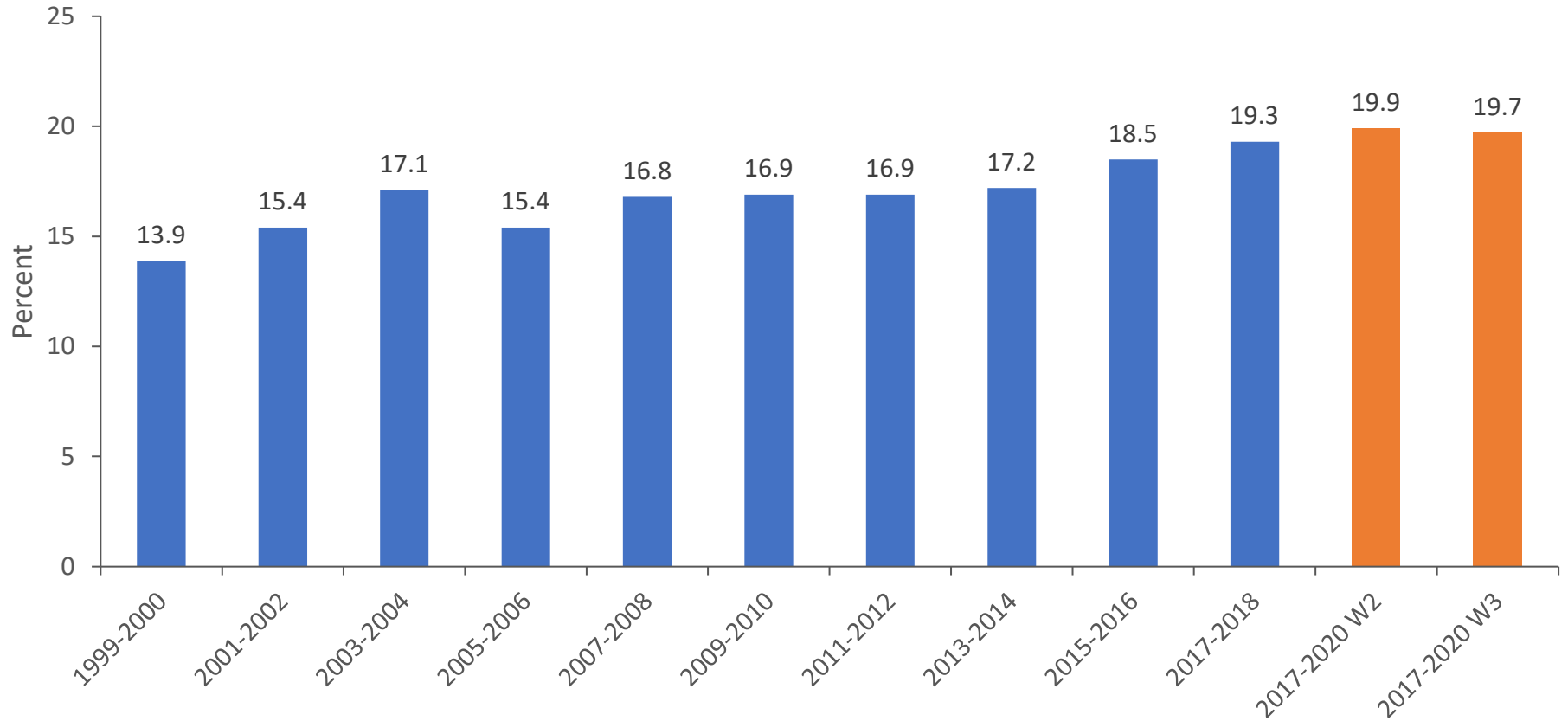


Trends in age-adjusted obesity, adults 20+: NHANES



NOTE: Estimates were age-adjusted by the direct method to the 2000 U.S. census population using the age groups 20–39, 40–59, and 60 and over.
SOURCE: NCHS, National Health and Nutrition Examination Survey.

Trends in childhood obesity, 2-19 years: NHANES



NHANES health outcomes over time

- 2017-2020 pre-pandemic compares well to earlier years
 - Regardless of W2 or W3

Next steps

- Evaluate pre-pandemic 2017-2020 fasting weights
- Simulations of 2017-2020 pre-pandemic PSU weight adjustment using 2013-2016 incomplete datasets examining effects on 6 health outcomes
- Leaning to weight 3 since urban/rural differences more closely tied to health than population size alone

Pre-pandemic 2017-2020 file

- **DO NOT** separate 2019-2020 from 2017-2018 in file
 - 2019-2020 not nationally representative due to interrupted data collection
 - Using 2017-2020 pre-pandemic weights does not result in nationally representative estimates for 2019-2020
 - File will include renumbering of sequence ID numbers and other changes to discourage identifying 2017-2018 records

Questions for BSC

- Thoughts related to messaging about the file
 - Stand-alone analysis of 2019-2020 data not possible (implications for trend analyses)
- Interpretation of estimates
 - What language should be used about the time period represented (i.e., most of 2020 missing)

