Evaluating Diary Collection Mode Changes in the Context of the COVID-19 Pandemic

Brett McBride and Nikki Graf

Federal Committee on Statistical Methodology

2021 Research and Policy Conference

Session E-5

November 3, 2021

BLS

Consumer Expenditure Surveys

- Surveys produced by the Bureau of Labor Statistics (BLS) collecting detailed expenditure information for a national sample of households
- Diary Survey
  - 2 one-week self-administered (paper) diaries
  - Collect daily, small and easily forgotten expenses
  - Approximately 17,800 sampled households per year, 53 percent response rate (2019)

Time Periods and Protocol Changes

■ **Baseline:** January – March 18, 2020

  ▶ Involved usual procedures

■ **Period 1:** March 19 – May 2020

  ▶ Census stopped personal visits and the Census National Processing Center closed

  ▶ Diary completion: FRs called respondents, transcribed entries into a paper diary

■ **Period 2:** June 2020

  ▶ Online diaries with an alternative of phone transcription

■ **Period 3 (this talk):** July – December 2020

  ▶ Diary completion: online diary, telephone transcription, or 'traditional' personal visits

Data Analyzed

- Consumer Units (CUs) with complete diaries from July to December, 2020

- Excluded: diaries with no entries (no reported expenditures), diaries from CUs temporarily away from home during the reporting period

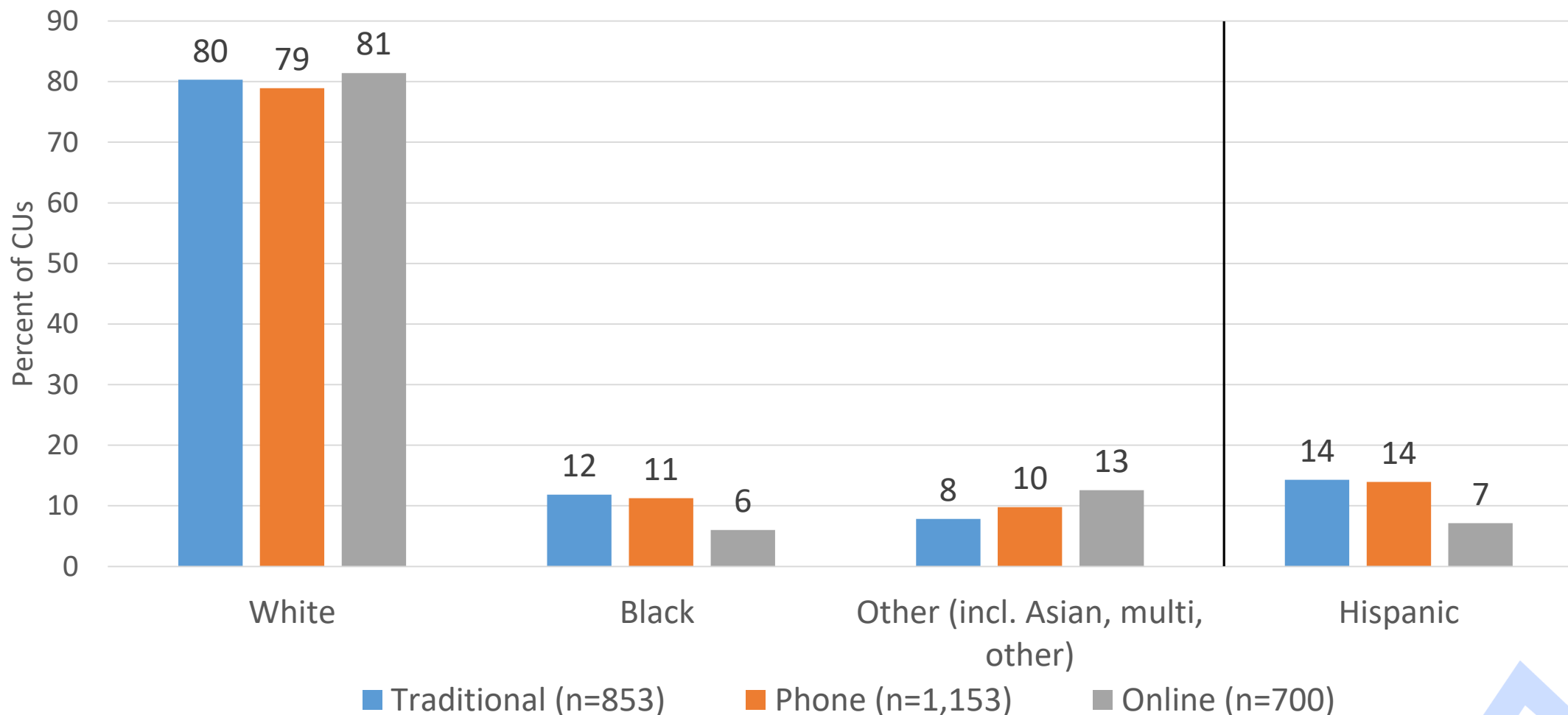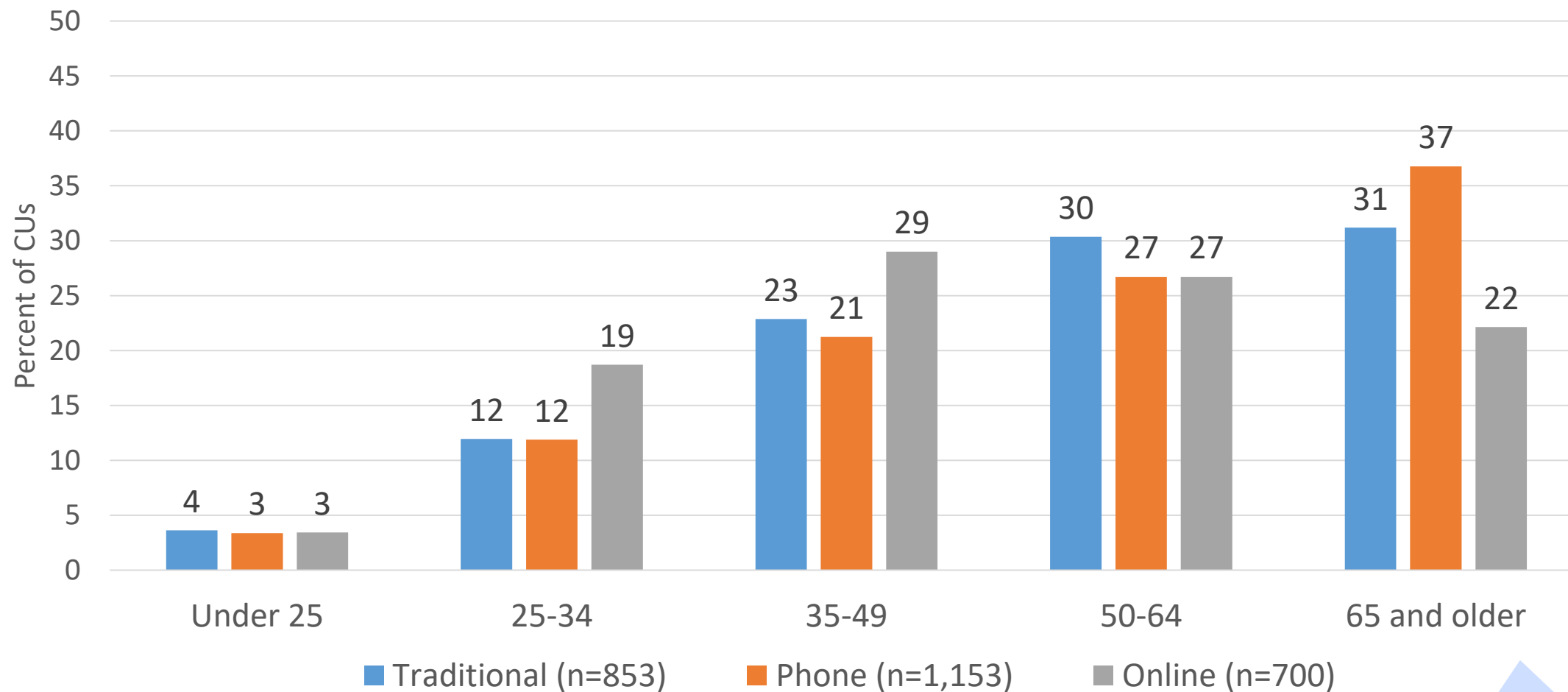- Final analysis dataset: 5,205 diaries (2,706 CUs)

Analysis Outline

■ Distribution of CU Demographics by Mode

■ Diary Entries by Mode of Diary Completion

■ Diarist Adherence to Itemization Instructions

■ Diary Entries by whether Data Editing (Allocation) Occurred

■ Regression of Diary Entries on Mode and Controls

BLS

Race/Ethnicity of Reference Person

Age of Reference Person

Education of Reference Person

CU Size

Percent of CUs

	1 person	2-3 persons	4+ persons
Traditional (n=853)	29	50	21
Phone (n=1,153)	36	49	16
Online (n=700)	26	52	22
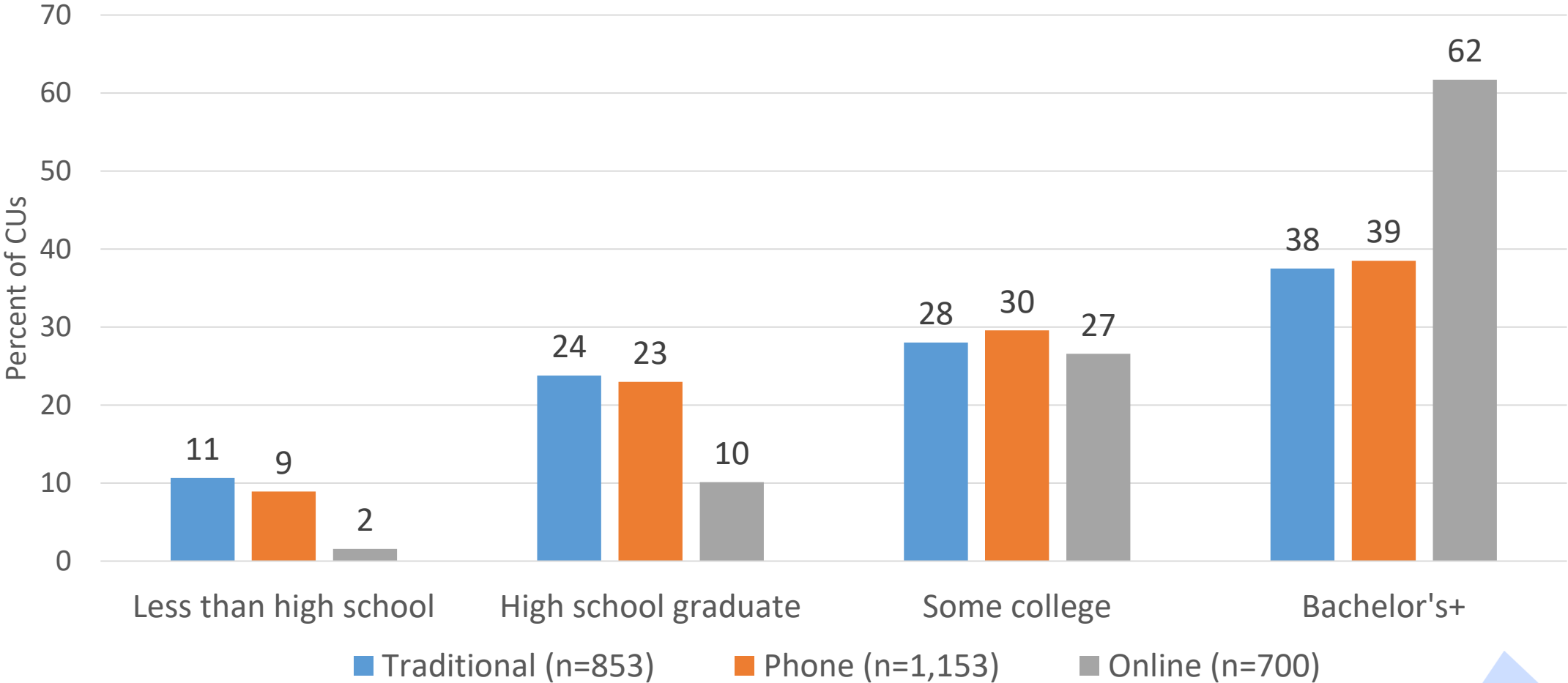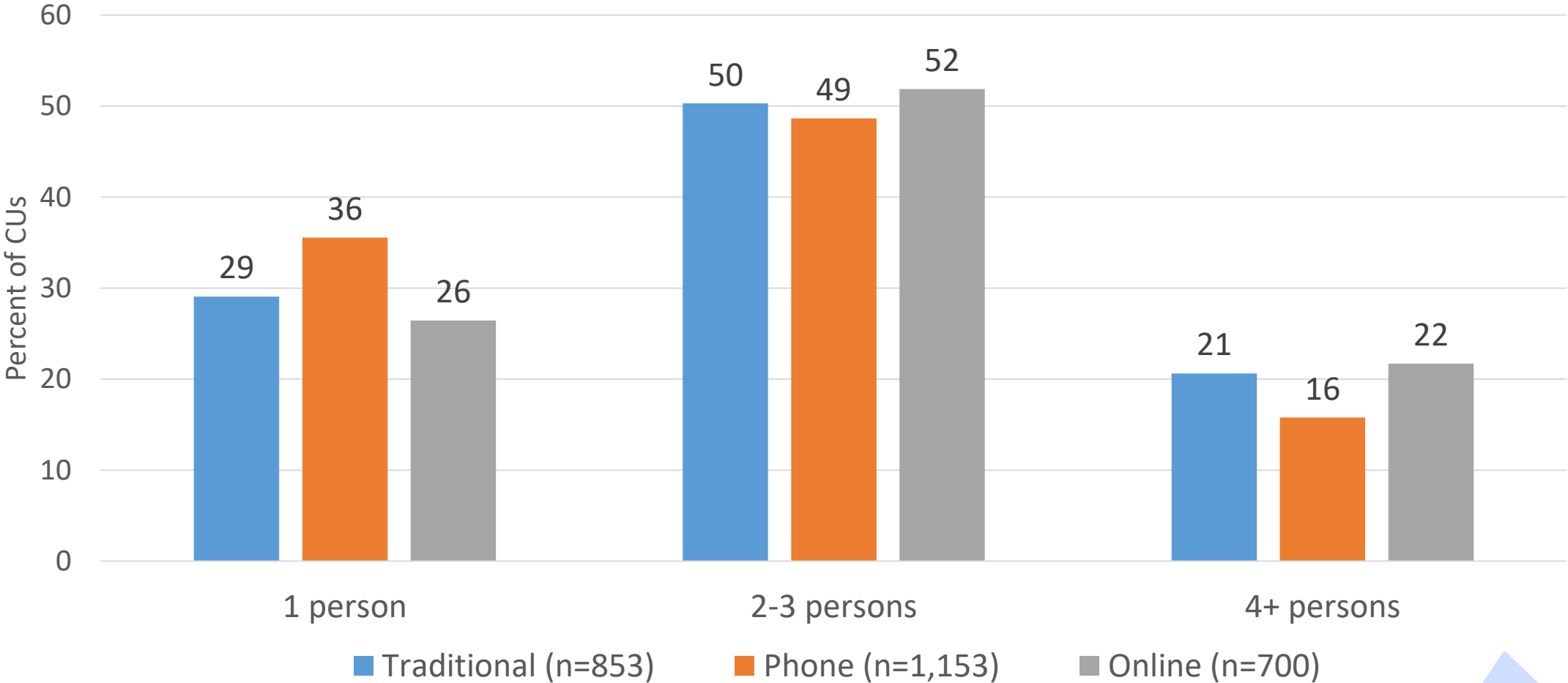
CU Income by Mode

Mode	N (CUs)	Mean income before tax	Median income before tax
Traditional	853	$77,905	$54,695
Phone transcription	1,153	$77,477	$53,164
Online	700	$114,136	$90,800

Diary Data Collection Modes, Entries During Pandemic

Mode Share and Mean Entries

Y-axis: Mean Entries (Non-Blank Diaries) — 0, 5, 10, 15, 20, 25, 30, 35, 40

- Traditional (n=1,649): **31.7%**
- Phone (N=2,212): **42.5%**
- Online (n=1,344): **25.8%**
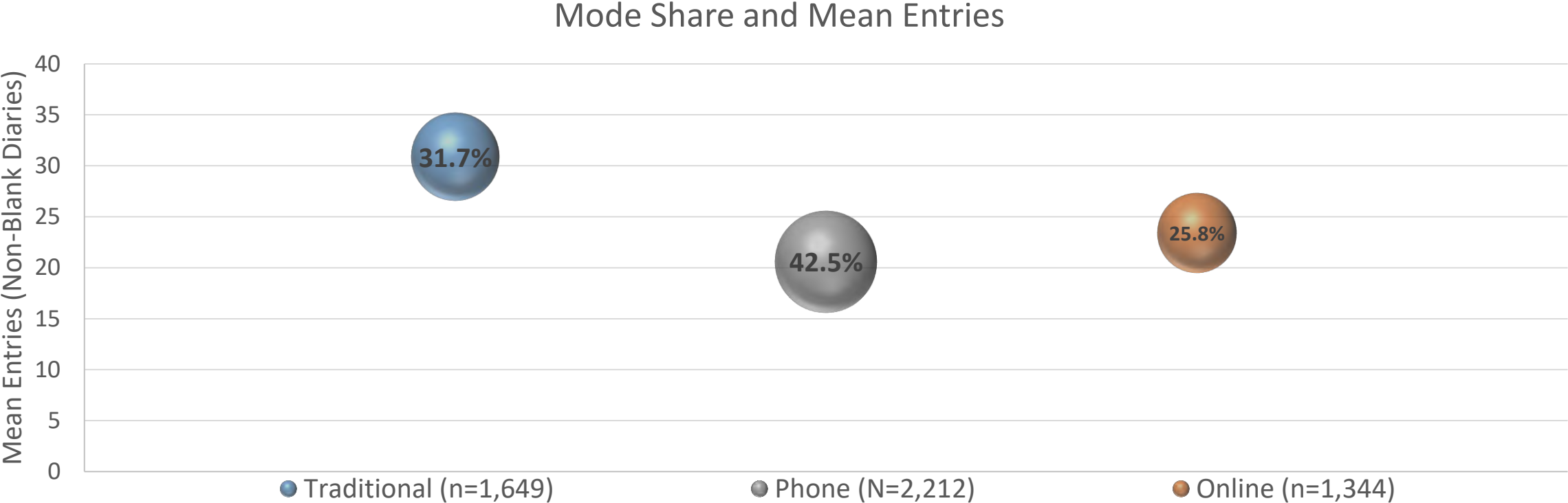
Note: size of the ball indicates mode share of total diaries, height represents mean entries in that mode

Diary Grocery Reporting Instructions

■ Respondents are instructed to itemize: report each individual food item purchased

Food and Drinks for Home Consumption

	What did you buy or pay for?	Is this item: Mark (X) one				Cost without tax		Mark (X) If purchased for someone not on your list	Name of Store or Website where purchased
		fresh	frozen	bottled/ canned	other				
101	bread	X (1)	(2)	(3)	(4)	1	49		Foodway Grocery Store
102	eggs	X (1)	(2)	(3)	(4)	1	50		
103	chicken wings	(1)	X (2)	(3)	(4)	6	78		
104	apples	X (1)	(2)	(3)	(4)	2	80		
105	beer	(1)	(2)	X (3)	(4)	4	29		
106	milk	X (1)	(2)	(3)	(4)	2	99		
107	orange juice	(1)	(2)	X (3)	(4)	3	99		
108	candy	(1)	(2)	(3)	X (4)	2	50		

Level of detail needed

BEEF – Specify the cut and describe, such as round roast, ground beef, etc.

PORK – Specify the cut and describe, such as whole ham, bacon, spareribs, etc.

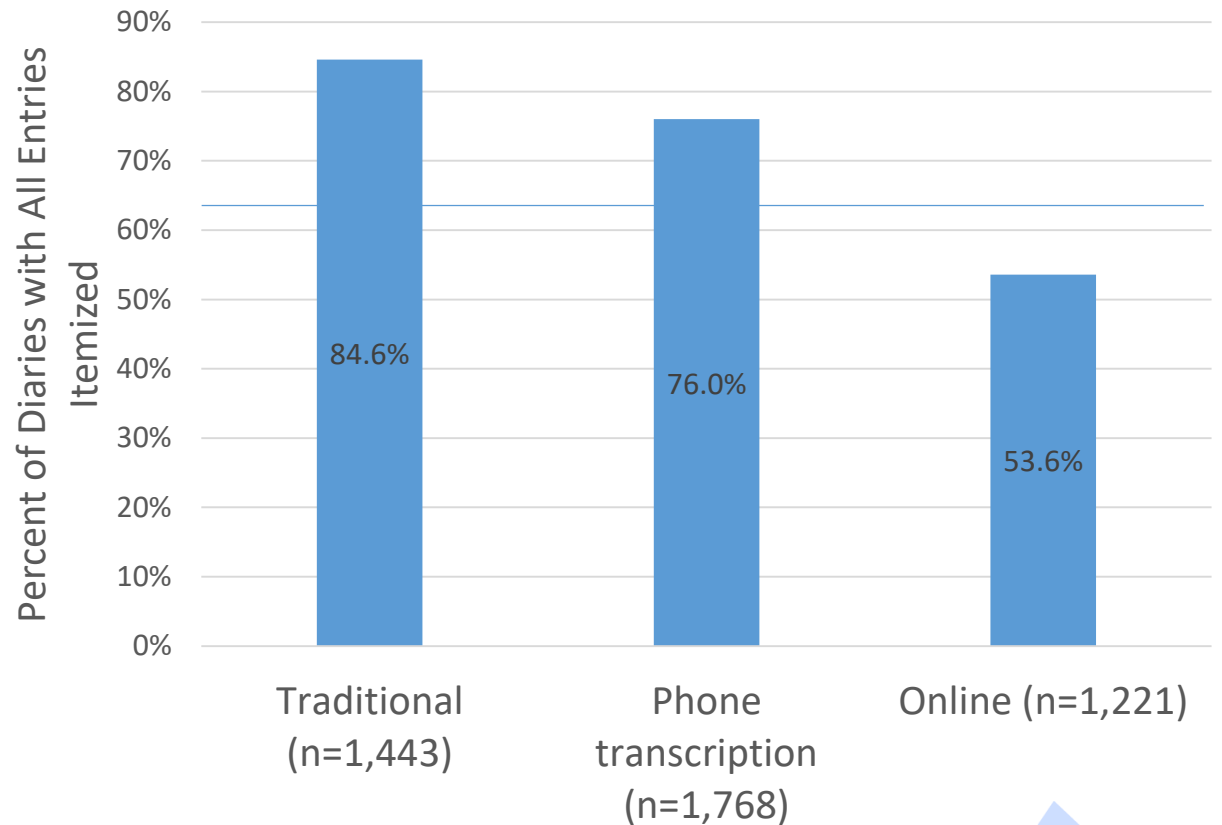OTHER FOOD – Give a complete description, such as scalloped potatoes.

Adherence to Instructions Varied by Mode

- ## Overall compliance was good

	Number (diaries)	Percent
Itemized entries	3,218	72.6%
1+ non-itemized	1,214	27.4%
Total	4,432	100.0%

- ## Less compliance with itemization in online diaries

Bar chart — Percent of Diaries with All Entries Itemized:
- Traditional (n=1,443): 84.6%
- Phone transcription (n=1,768): 76.0%
- Online (n=1,221): 53.6%

Diaries without Allocation

- Lack of itemization results in need for allocation (i.e., specific items determined using established patterns of itemized reporting)

- Yet allocation happened in most diaries in this period (79%)

- Unexpectedly, those with allocation had a larger number of entries (before allocation)

Presence of Allocation	Number (diaries)	Average number of entries (before allocation)	Median number of entries (before allocation)
No allocation carried out	1,107	11.4	7
Allocation carried out	4,098	28.2	21

Predictors of Number of Diary Entries (1 of 2)

Parameter	Coefficient	SE
Intercept	34.80	(1.750)*
Phone (v. Traditional)	-7.97	(1.364)*
Online (v. Traditional)	-11.11	(1.371)*
Involved Recall of Expenses	-5.08	(1.269)*
No Allocation	-13.75	(0.830)*
Ref. person HS or less (v. Col grad)	-7.12	(1.185)*
Ref. person some college (v. Col grad)	-4.64	(1.017)*
2 or 3 CU members (v. 1 member)	6.68	(1.019)*
4 or more CU members (v. 1 member)	14.52	(1.585)*
$R^2 = 0.20$, n=5,205 diary weeks		

$* \ P < 0.05$

Predictors of Number of Diary Entries (1 of 2)

Parameter	Coefficient	SE
Intercept	34.80	(1.750)*
Phone (v. Traditional)	-7.97	(1.364)*
Online (v. Traditional)	-11.11	(1.371)*
Involved Recall of Expenses	-5.08	(1.269)*
No Allocation	-13.75	(0.830)*
Ref. person HS or less (v. Col grad)	-7.12	(1.185)*
Ref. person some college (v. Col grad)	-4.64	(1.017)*
2 or 3 CU members (v. 1 member)	6.68	(1.019)*
4 or more CU members (v. 1 member)	14.52	(1.585)*
$R^2 = 0.20$, n=5,205 diary weeks		

* $P < 0.05$

Predictors of Number of Diary Entries (1 of 2)

Parameter	Coefficient	SE
Intercept	34.80	(1.750)*
Phone (v. Traditional)	-7.97	(1.364)*
Online (v. Traditional)	-11.11	(1.371)*
Involved Recall of Expenses	-5.08	(1.269)*
No Allocation	-13.75	(0.830)*
Ref. person HS or less (v. Col grad)	-7.12	(1.185)*
Ref. person some college (v. Col grad)	-4.64	(1.017)*
2 or 3 CU members (v. 1 member)	6.68	(1.019)*
4 or more CU members (v. 1 member)	14.52	(1.585)*
$R^2 = 0.20$, n=5,205 diary weeks		

* $P < 0.05$

Predictors of Number of Diary Entries (1 of 2)

Parameter	Coefficient	SE
Intercept	34.80	(1.750)*
Phone (v. Traditional)	-7.97	(1.364)*
Online (v. Traditional)	-11.11	(1.371)*
Involved Recall of Expenses	-5.08	(1.269)*
No Allocation	-13.75	(0.830)*
Ref. person HS or less (v. Col grad)	-7.12	(1.185)*
Ref. person some college (v. Col grad)	-4.64	(1.017)*
2 or 3 CU members (v. 1 member)	6.68	(1.019)*
4 or more CU members (v. 1 member)	14.52	(1.585)*
R^2 = 0.20, n=5,205 diary weeks		

* $P < 0.05$

Predictors of Number of Diary Entries (2 of 2)

Parameter	Coefficient	SE
Ref. person aged 18-34 (v. 65+)	-3.60	(1.113)*
Ref. person aged 35-49 (v. 65+)	-0.58	(1.118)
Ref. person aged 50-64 (v. 65+)	1.25	(0.915)
Ref. person Black (v. White)	-4.92	(1.821)*
Ref. person Other (v. White)	-5.25	(1.241)*
CU in Northeast (v. West)	-2.11	(1.418)
CU in Midwest (v. West)	-0.72	(2.049)
CU in South (v. West)	-3.27	(1.504)*

* $P < 0.05$

$R^2 = 0.20$, n=5,205 diary weeks

Summary

- Fewer entries when interviews conducted in non-traditional manner (phone, online)

- Weaker adherence with diary-keeping instructions in new modes

- Allocation (needed if the entry is vague) associated with more entries

- Education, CU size, race, geography all factors in diary entries

- Return to baseline/average entries?

Contact Information

Brett McBride & Nikki Graf

mcbride.brett@bls.gov,

graf.nikki@bls.gov

BLS