

Internet and Electronic Data Collection in Demographic Surveys

Benjamin Reist

Census Scientific Advisory Committee Meeting

April 10, 2014

Current State

- Predominantly CAPI data collection
- Mix of in-person and telephone interviewing
- Limited use of self response
- Web data collection only in a few surveys
- Experiencing declining response rates
- One size fits all data collection
- Collect until time or money are exhausted

Future State

- Wider use of mixed modes
- Increased use of self response
- Leverage web data collection in many surveys
- Moving away from response rate as key indicator of quality
- Tailoring data collection methodology at the case level

Questions to Answer

- What is the quality impacts of mixing modes?
 - What modes should be used?
 - **In what sequence should modes be offered?**
- How do we target cases for tailored data collection methodology?
- What **alternative quality measures** should be used?

Mode Sequence Experiment

- The 2010 cycle marked the introduction of the web data collection mode to the National Survey of College Graduates (NSCG)
- 60,875 were eligible and randomly assigned to four treatment groups
 - Web first group (n = 5,000)
 - Mail first group (n = 5,000)
 - Telephone first group (n = 3,500)
 - Choice group (n = 47,375)

Mode Sequence Experiment

- Experiment was evaluated using
 - Survey response rates
 - Proportion of response by each mode
 - Sample representation of responding sample
 - Data quality Measures
 - Survey cost

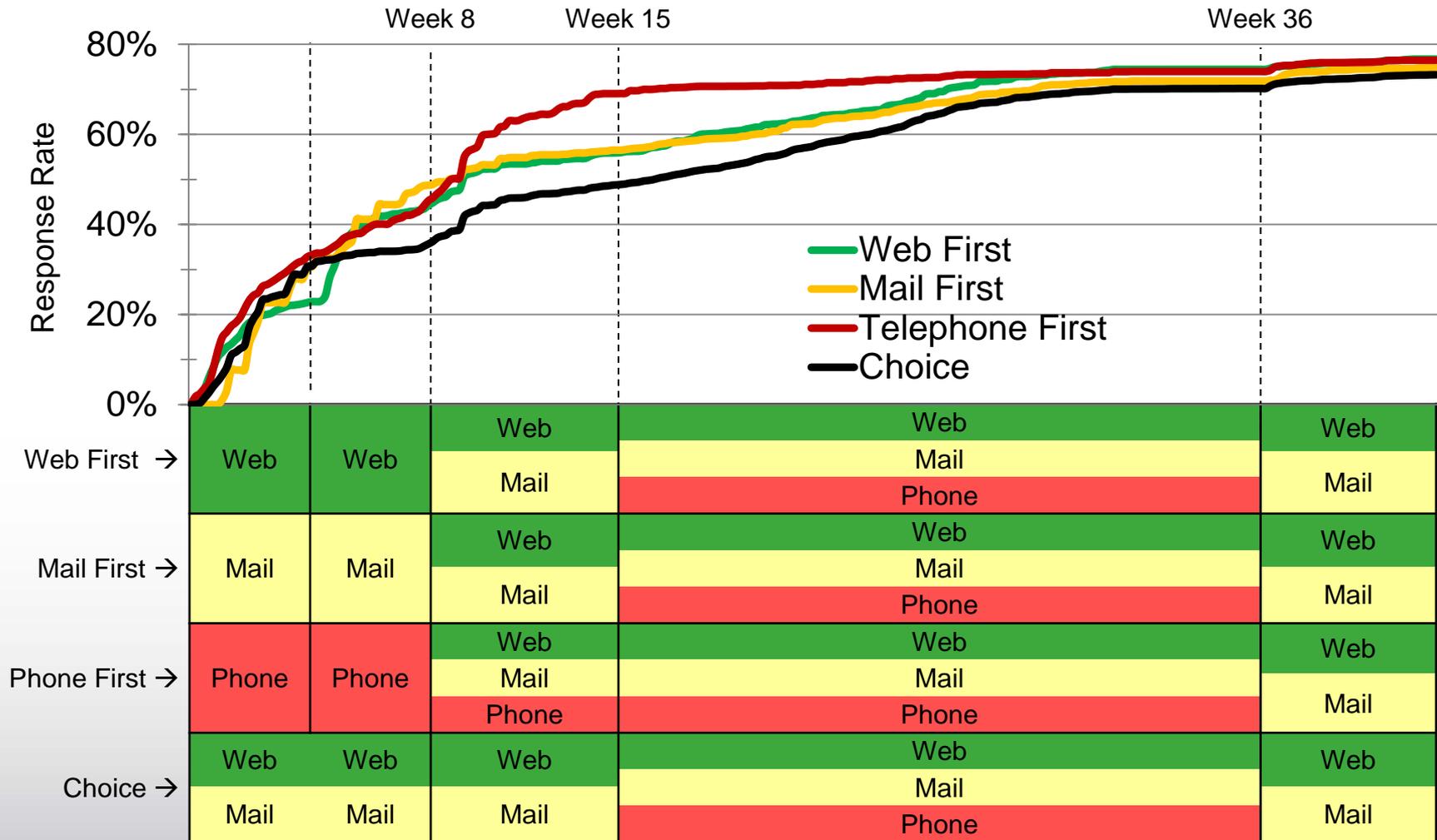
Data Collection

- The contact strategy was similar across treatment groups, but available modes differed as shown in the next slide
- The NSCG data collection effort included five distinct phases:
 - Invitation (weeks 1-4)
 - Initial Reminder (weeks 5-7)
 - Additional Mode (weeks 8-14)
 - Production CATI (weeks 15-35)
 - Incentive (weeks 36-42)

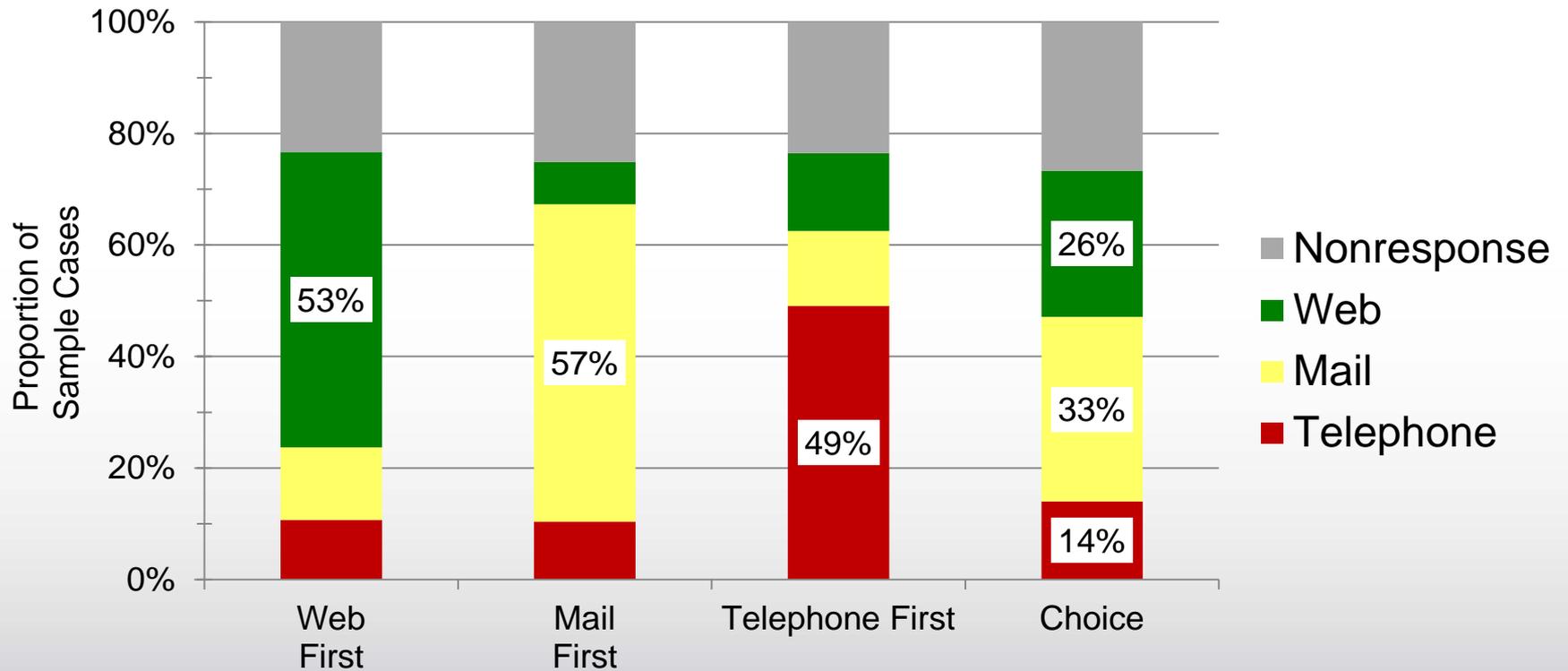
Response Options in Each Phase of Data Collection

	<u>Data Collection Phase</u>				
	Invitation Phase (Weeks 1-4)	Initial Reminder Phase (Weeks 5-7)	Additional Mode Phase (Weeks 8-14)	Production CATI Phase (Weeks 15-35)	Incentive Phase (Weeks 36-42)
Web First	Web	Web	Web	Web	Web
			Mail	Mail	Mail
Mail First	Mail	Mail	Web	Web	Web
			Mail	Mail	Mail
Telephone First	Telephone	Telephone	Web	Web	Web
			Mail	Mail	Mail
Choice	Web	Web	Web	Web	Web
	Mail	Mail	Mail	Mail	Mail

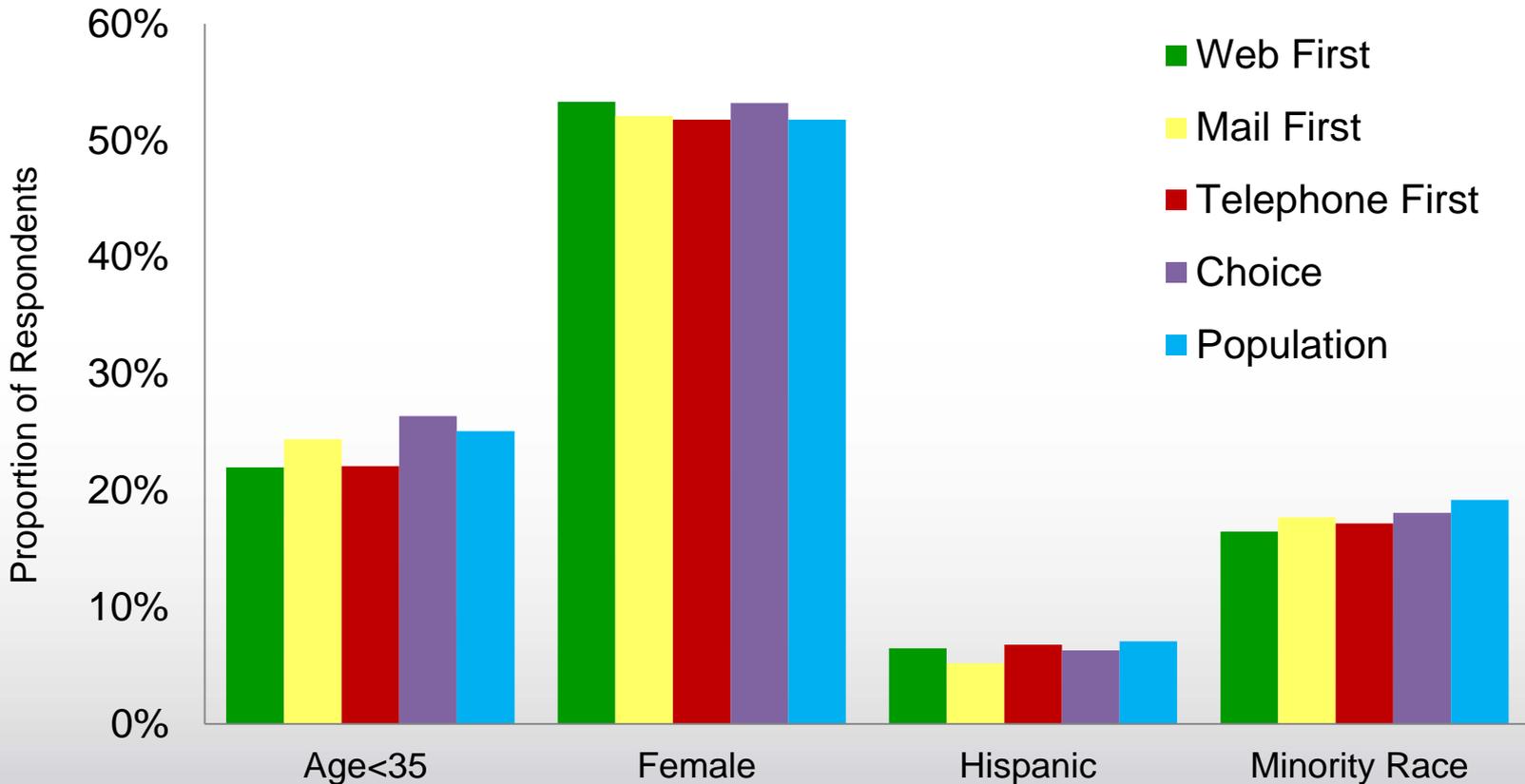
Does mode sequence impact response rates?



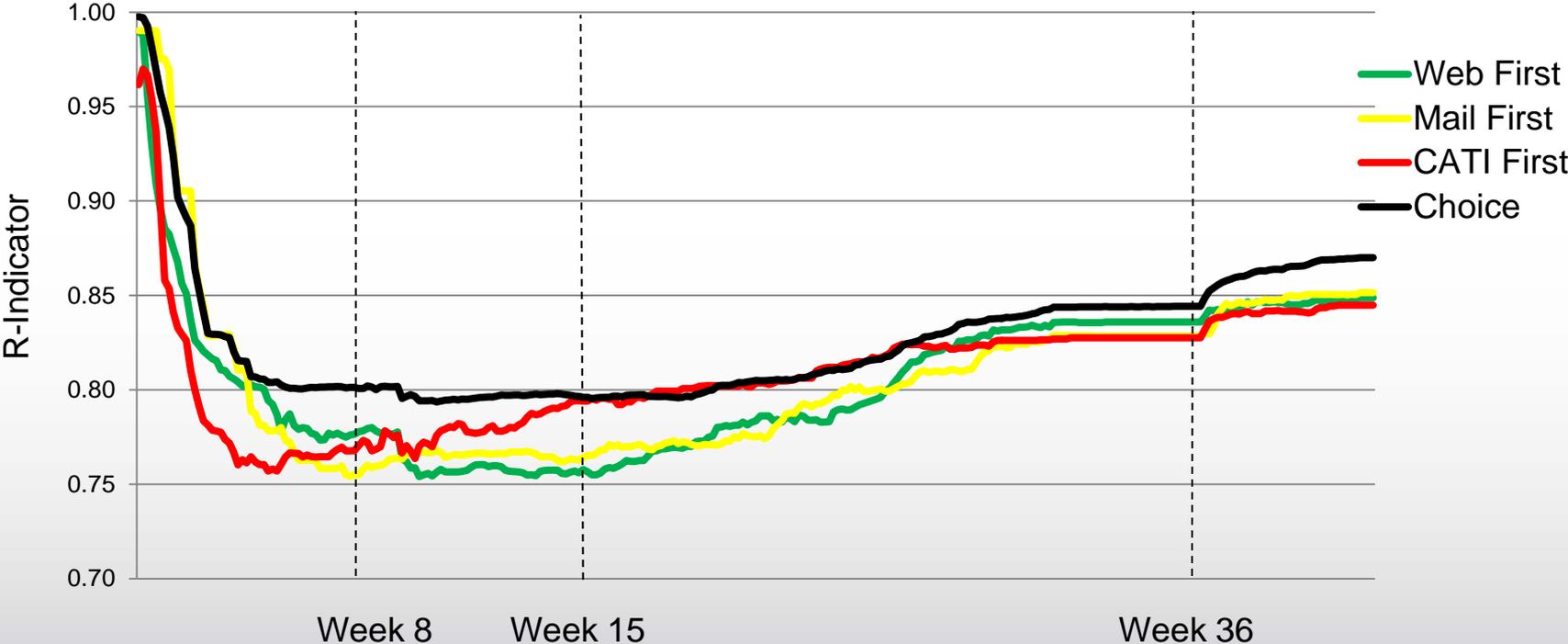
Mode of response gravitates toward initial mode offered



Mode sequence has little impact on the demographic composition of respondents



Sequentially offering modes improves representativeness of responding sample



How mode sequence impacts data quality for three indicators

	Data Editing Score	Item Nonresponse Score	Verbatim Length (Job Description Question)
Web First	6.64	0.66	77.53
Mail First	7.17	1.38	68.59
Telephone First	5.38	0.69	76.21
Choice	6.71	0.98	74.50

How did mode sequence affect survey costs?

	Sample Size	<u>Week 8</u>		<u>End of Data Collection</u>	
		Response Rate	Cost per Response	Response Rate	Cost Per Response
Web First	5,000	42.9%	\$13.80	77.1%	\$48.24
Mail First	5,000	47.1%	\$37.44	75.0%	\$66.04
Telephone First	3,500	42.4%	\$70.42	76.7%	\$74.77
Choice	47,375	34.4%	\$26.02	73.4%	\$58.38

Questions

Are there other modes of data collection that we should investigate?

What measures should we evaluate when considering adding new modes of data collection?

Thank You

John Finamore

National Center for Science and Engineering Statistics

National Science Foundation

&

Dr. Don Dillman

Washington State University