



A Delicate Transaction: The Data Quality Program for the 2016 and 2019 Survey of Consumer Finances

Kate Bachtell, NORC @ University of Chicago

2019 FedCASIC
April 16, 2019
Washington, D.C



CO-AUTHORS

Catherine Haggerty, NORC @ University of Chicago

Shannon Nelson, NORC @ University of Chicago

Kevin Moore, Board of Governors of the Federal Reserve System

Richard Windle, Board of Governors of the Federal Reserve System

The analysis and conclusions set forth are those of the authors and do not indicate concurrence by colleagues, the Board of Governors of the Federal Reserve System, or NORC.



DELICATE IN MANY STAGES

Collecting complex financial data from U.S. households is extremely difficult

- **Every interview is an accomplishment**

Tough to give constructive feedback about these accomplishments!

SURVEY OF CONSUMER FINANCES (SCF)



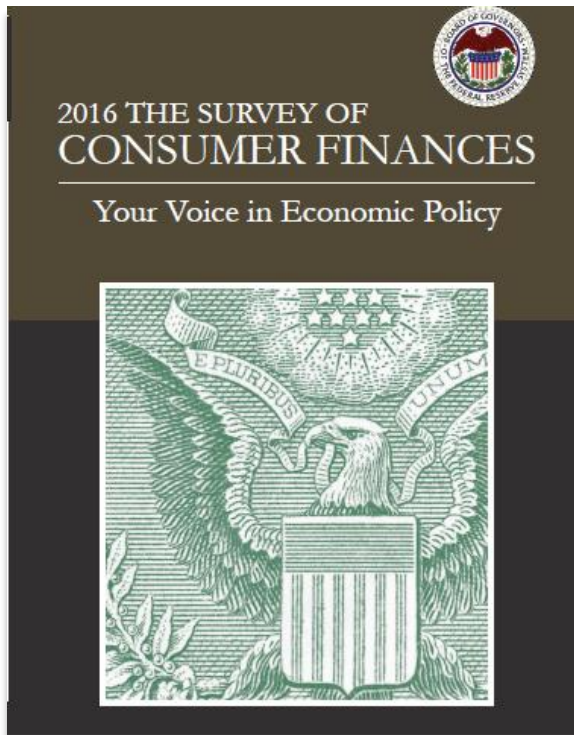
Triennial survey

Sponsored by the Board of Governors of the Federal Reserve System (FRB)

Premier source of data on U.S. household finances

Dual frame sample

Public use datasets available through FRB website



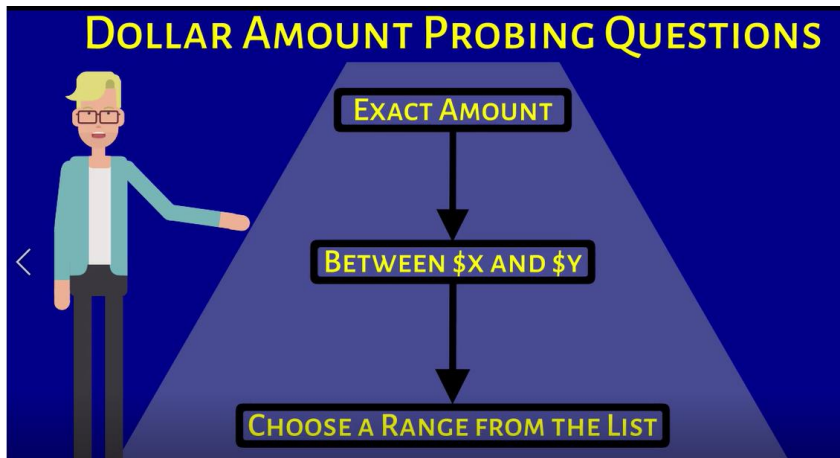
“HIGH QUALITY”

Survey data

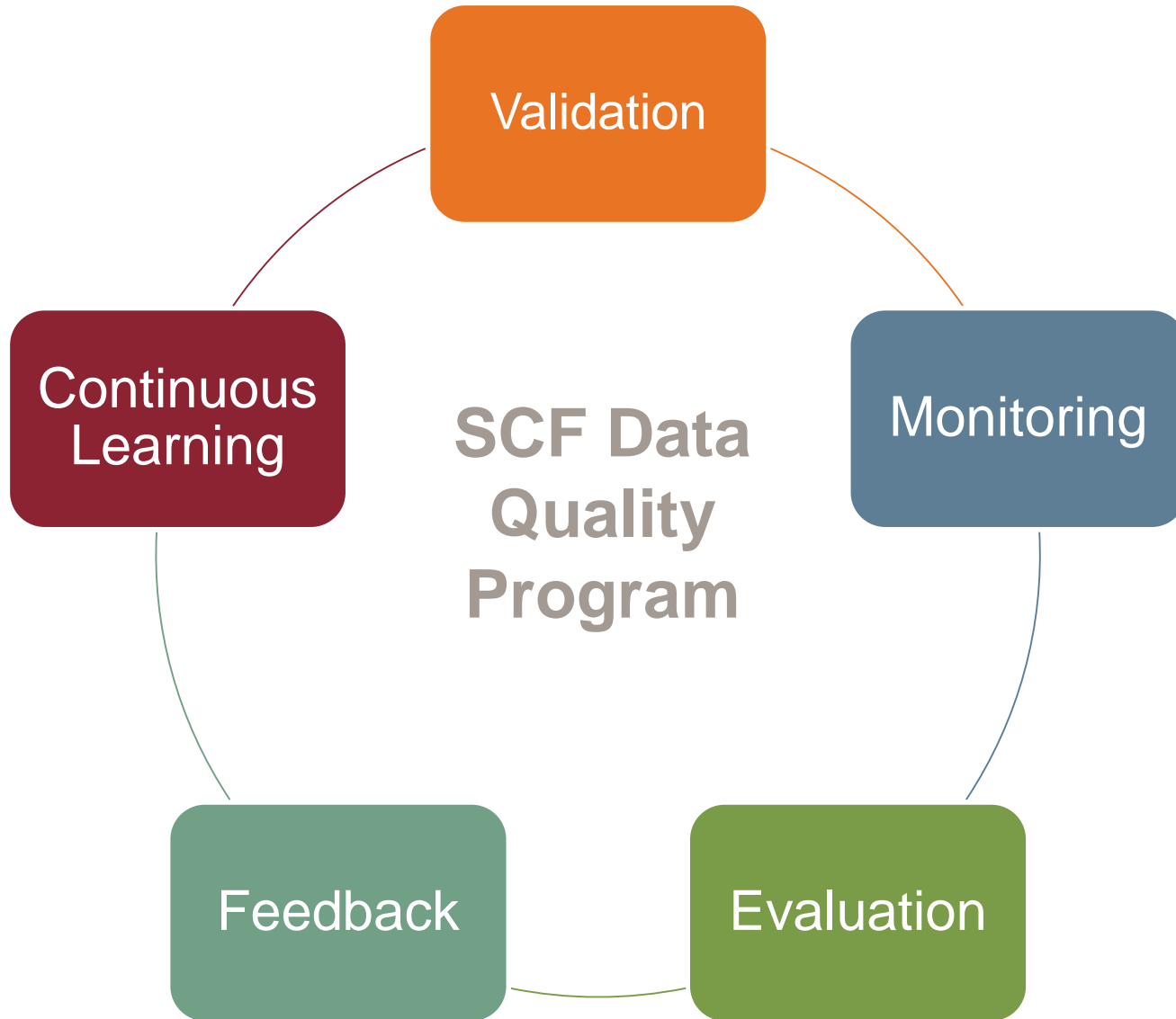
- ✓ Precise dollar amounts
- ✓ Ranges

Data Quality Feedback

- ✓ Relevant
- ✓ Informed
- ✓ Coherent
- ✓ Actionable
- ✓ Impactful



CORE FEATURES OF SCF DATA QUALITY PROGRAM



RESULTS FROM 2016 SCF

What drives the probability of having missing dollar values?



METHODS

- Total 2016 SCF Completed Interviews (n=6,261)
- Area Probability Sample (n=4,756)
- Unweighted
- Descriptive statistics, bivariate and multivariate logistic regression

VARIABLES

- Dependent Variables:
 - Percent missing dollar values (DK/REF)
 - Any missing dollar values

VARIABLES

- Independent Variables:
 - Region of country
 - Number of refusals at week 11
 - FI assessment of Respondent:
 - Unlikely to participate (at week 11)
 - Interest (during interview)
 - Suspicion (before & after interview)
 - Period of interview completion (early/middle/late)
 - Self-rated financial knowledge
 - Financial literacy

DESCRIPTIVE STATISTICS

70% have ZERO missing dollar values

27% have less than a quarter missing

Only 3% have a quarter or more missing

...So it's a small slice of the pie. But can we predict this behavior?

BIVARIATE LOGISTIC REGRESSION RESULTS

Variable	Mean	Effect
Region of country	n/a	
Number of refusals	.42	
Unlikely to participate	n/a	+ (nearly significant)
Period of completion	n/a	+ (Late)
Interest	3.64	-
Suspicion	2.62	+
Financial knowledge	7.22	
Financial literacy	2.17	

MULTIVARIATE LOGISTIC REGRESSION RESULTS

Variable	Estimate		Std. Error	Odds Ratio
<i>Intercept</i>	-1.57	***	0.22	0.21
Northeast	0.05		0.09	1.05
South	-0.01		0.09	0.99
West	0.02		0.09	1.02
# refusals	0.00		0.04	1.00
Unlikely to participate	0.15		0.12	1.16
Middle	-0.09		0.09	0.91
Late	0.48	**	0.19	1.62
Interest	-0.02		0.04	0.98
Suspicion	0.19	***	0.04	1.21
Financial knowledge	0.02		0.02	1.02
Financial literacy	0.06		0.04	1.07

n = 4,718

-2 Log Likelihood = 5688.23

What drives the probability of having any missing dollar values?

Late in field period

Respondent suspicion

IMPROVEMENTS FOR 2019 SCF DATA QUALITY PROGRAM

IMPROVEMENTS FOR 2019 SCF DATA QUALITY PROGRAM

✓ Coding of FRB editors' open-ended comments

Monitoring

Feedback

▪ Systematic identification & communication of exemplary interviewing

Continuous Learning

Evaluation

- Pre-determined continuous learning exercises mapped to *specific* data quality issues
- Monitoring completion of prescribed continuous learning activities
- Enhanced monitoring of data quality metrics over time

Get the public-use SCF data at:

<https://www.federalreserve.gov/econres/scfindex.htm>

<https://www.federalreserve.gov/econres/scfindex.htm>

Learn more about NORC's involvement at:

<http://www.norc.org/Research/Projects/Pages/survey-of-consumer-finances-SCF.aspx>

Kate Bachtell
Bachtell-kate@norc.org

Thank You!

