

Evaluating Respondents' Burden via Indirect Indicators of Data Quality: Item vs. Index Scores

Daniel K. Yang
Office of Survey Methods Research
U.S. Bureau of Labor Statistics
CE Survey Methods Symposium
July 18, 2017

The views expressed in this paper are those of the author(s) and do not necessarily reflect the policies of the Bureau of Labor Statistics



Overview

- ❑ Consumer Expenditure Surveys (CE) Redesign and burden measurement.
- ❑ Burden questions (items) and burden index scores.
- ❑ Correlations between burden and indirect indicators of data quality.
- ❑ Effects of burden on expenditure estimates.
- ❑ Conclusion and follow-ups.

Respondents' Burden Perception

- ❑ Gemini: redesign the CE Surveys to improve data quality, through a verifiable reduction in measurement error.
- ❑ Important: able to measure respondent burden (could contribute to data quality).
- ❑ How to best evaluate respondents' perceived level of burden is still an open question.

Burden Items

- ❑ Between October 2012 and September 2013, a series of questions were asked in the interview survey at the end of the final wave, including ten questions assessing respondents' perceived burden:

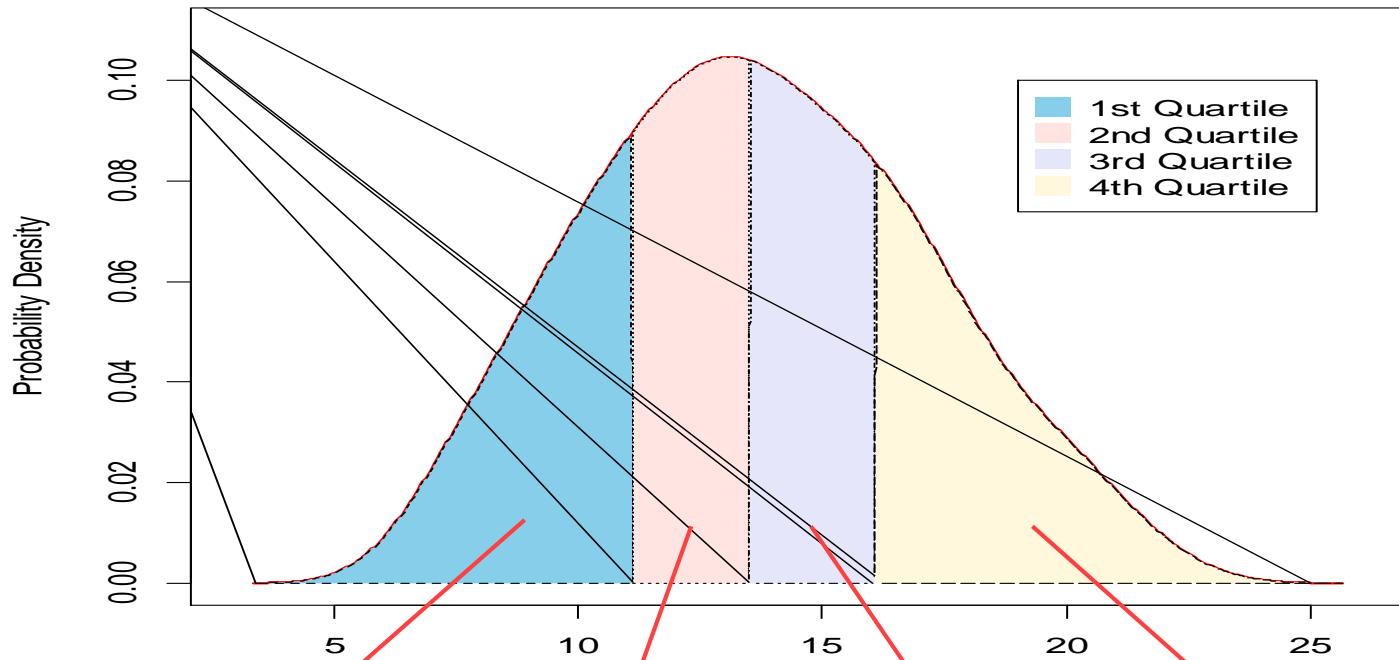
e.g. How burdensome was this survey to you? (bbur)

- Not at all burdensome
- A little burdensome
- Somewhat burdensome
- Very burdensome

Burden Index Scores

- ❑ Composite burden index scores (weighted, involving a correlation matrix of level of measurements, Yang 2015).
- ❑ Likert scales summation scores: a simplified alternative by compute a simple summation (of burden items Likert scales).

Figure 1. Distribution of Composite Burden Index Scores

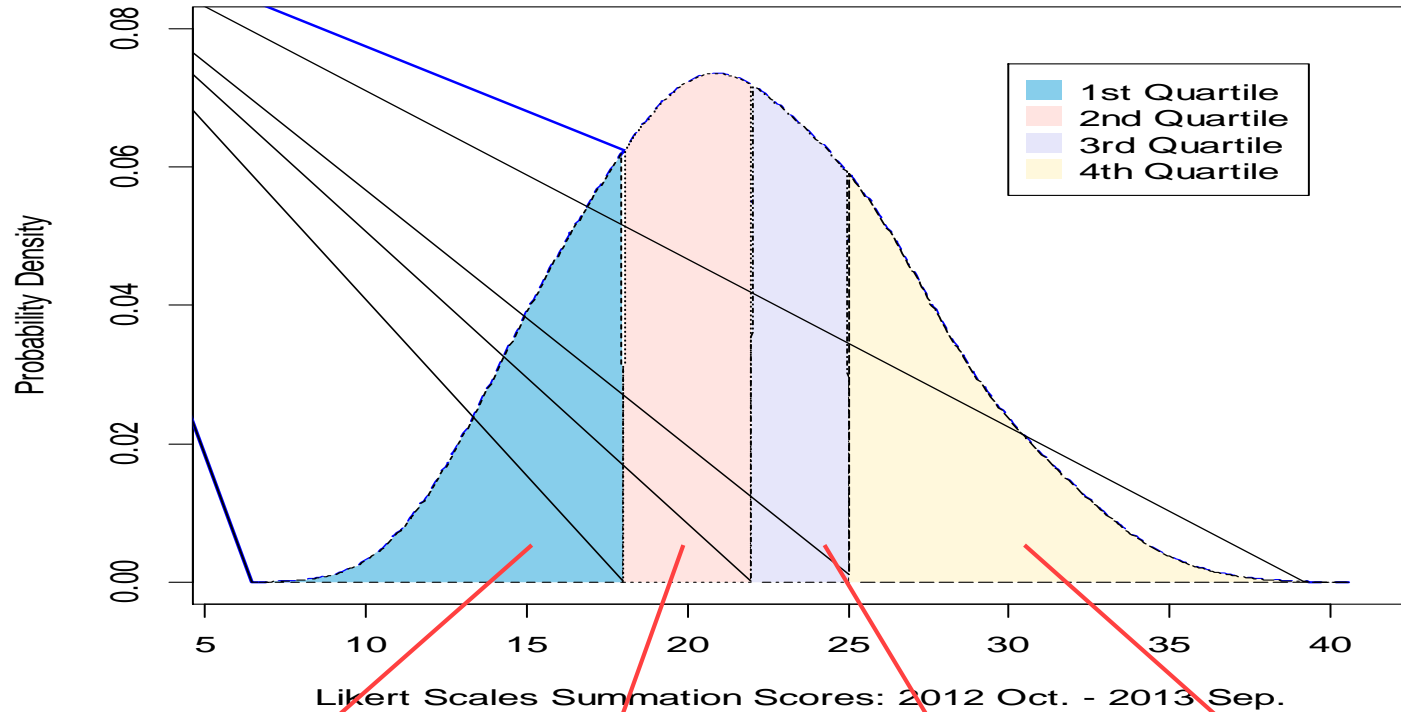


Composite Burden Index Scores: 2012 Oct. - 2013 Sep.

Not at all burdensome A little burdensome Somewhat burdensome Very burdensome



Figure 2. Distribution of Likert Scales Summation Scores



Not at all
burdensome

A little
burdensome

Somewhat
burdensome

Very
burdensome



Burden Measures

- Now we have three burden measures:
 - *(single) burden question or (single) burden item,*
 - *(composite) burden index (scores), and*
 - *Likert scales summation (scores) or Likert scales sum.*

Single Burden Question vs. Burden Index Scores

- ❑ The single burden question (item) had been used to examine the impact of the respondent's perceived burden on data quality (Yan 2015).
- ❑ Objective: whether we can use the single burden question (item) to monitor burden via comparing correlations between indirect indicators of data quality and the single burden question (item) vs. burden index scores.

Data Sample

- ❑ Burden data were collected between October, 2012 (Q4) and September, 2013 (Q3).
- ❑ Excluded households with missing values in any of the burden questions (items).
- ❑ Final sample total has 6,378 households.

Evaluation of Single Burden Question (Item) vs. Index Scores

- I. Visualize the descriptive statistics.
- II. Numerically examine how burden index scores are correlated to indirect indicators of data quality.
- III. Compare methods to identify most-burdened respondents (via burden measures) and the effects of burden measures on statistical estimates of expenditures.



Evaluation of Single Burden Question (Item) vs. Index Scores

- I. Visualize the descriptive statistics.



Figure 3. Example: Distribution of Single Burden Question and Burden Index Score Quartiles

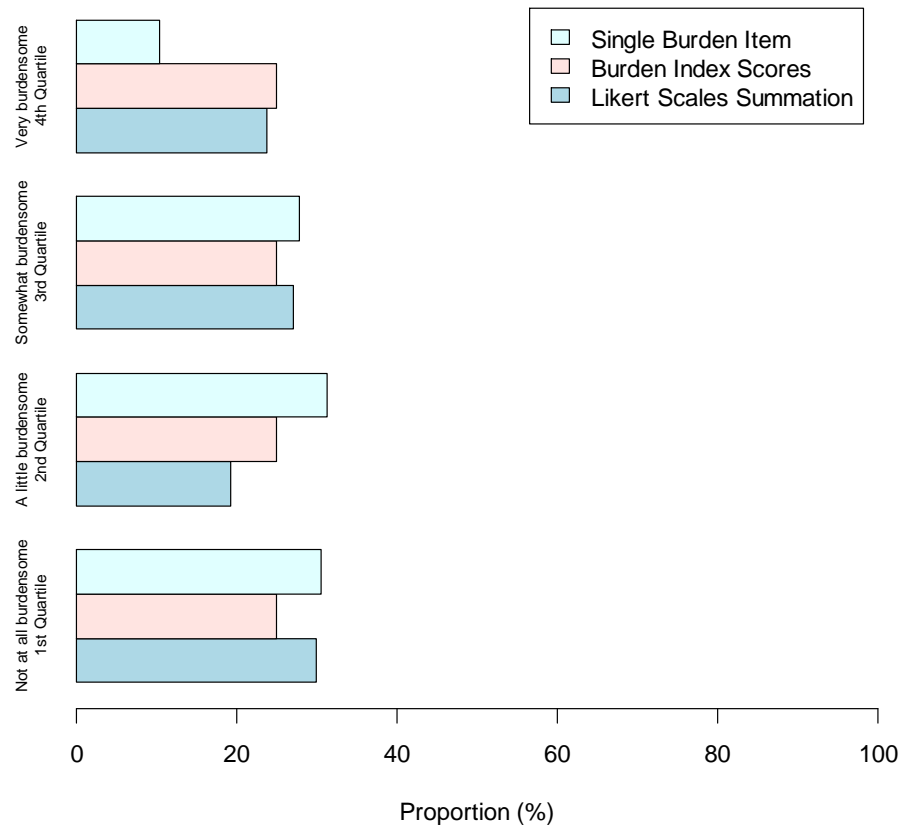


Table 1. Mean and Standard Deviation (SD) of Burden Index Scores at (Single) Burden Question (Item) Level

Comparing by (Level) of Single Burden Question (Item, bbur)		Composite Burden Index	Likerts Scale Sum
	n	Mean (SD)	Mean (SD)
1 = Not at all burdensome	1944	10.39 (2.11)	17.17 (3.08)
2 = A little burdensome	1994	13.45 (2.11)	21.44 (3.08)
3 = Somewhat burdensome	1773	15.57 (2.15)	24.48 (3.18)
4 = Very burdensome	667	18.89 (2.24)	29.42 (3.43)

Note: SD – standard deviation



Figure 4. Boxplot of Composite Burden Index Scores at (Single) Burden Question (Item) Level

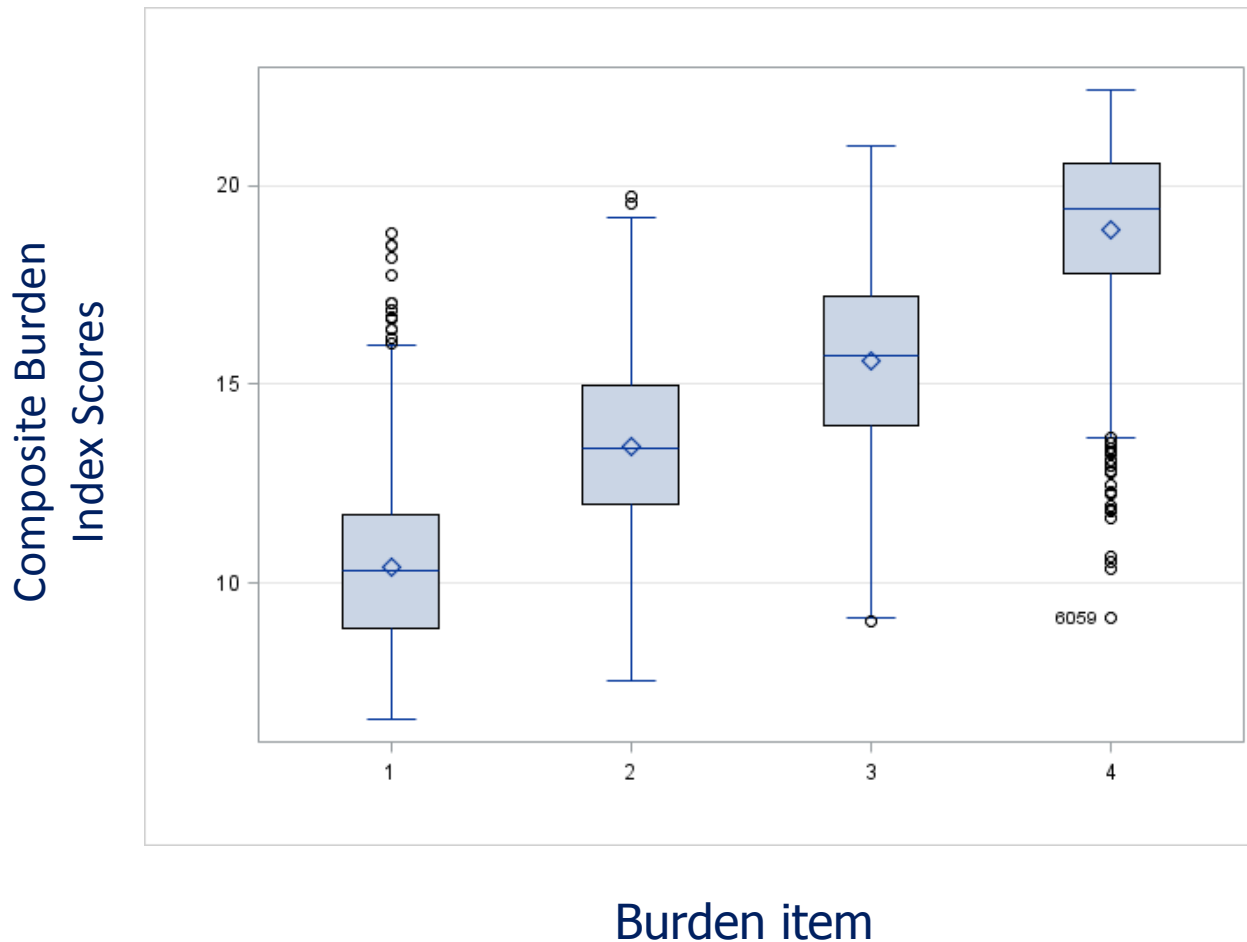
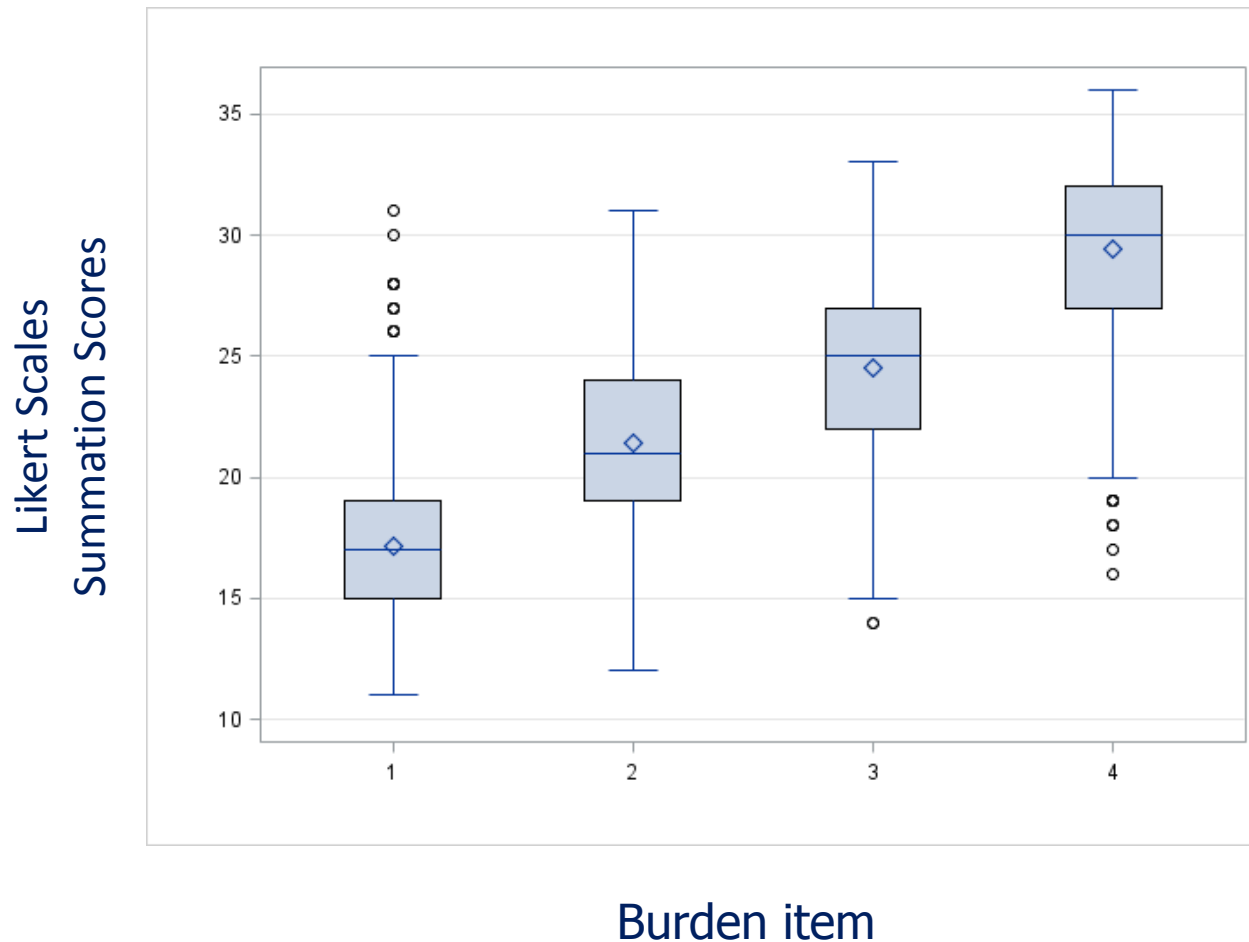


Figure 5. Boxplot of Likerts Scale Summation Scores at (Single) Burden Question (Item) Level



Evaluation of Single Burden Question (Item) vs. Index Scores

- II. Numerically examine how burden index scores are correlated to indirect indicators of data quality as compared to the correlation between the single burden question (item) and indirect indicators of data quality.

Indirect Indicators of Data Quality

- ❑ Two indirect indicator variables of data quality were adopted:
 - NUMDK: the number of “Don’t Know” answers and
 - NUMRF: the number of “Refused” answers provided by respondents.

Figure 6. Average Number of “Don’t Know” Answers by Levels of Single Burden Question (Item) and Burden Index Scores Quartile Groups

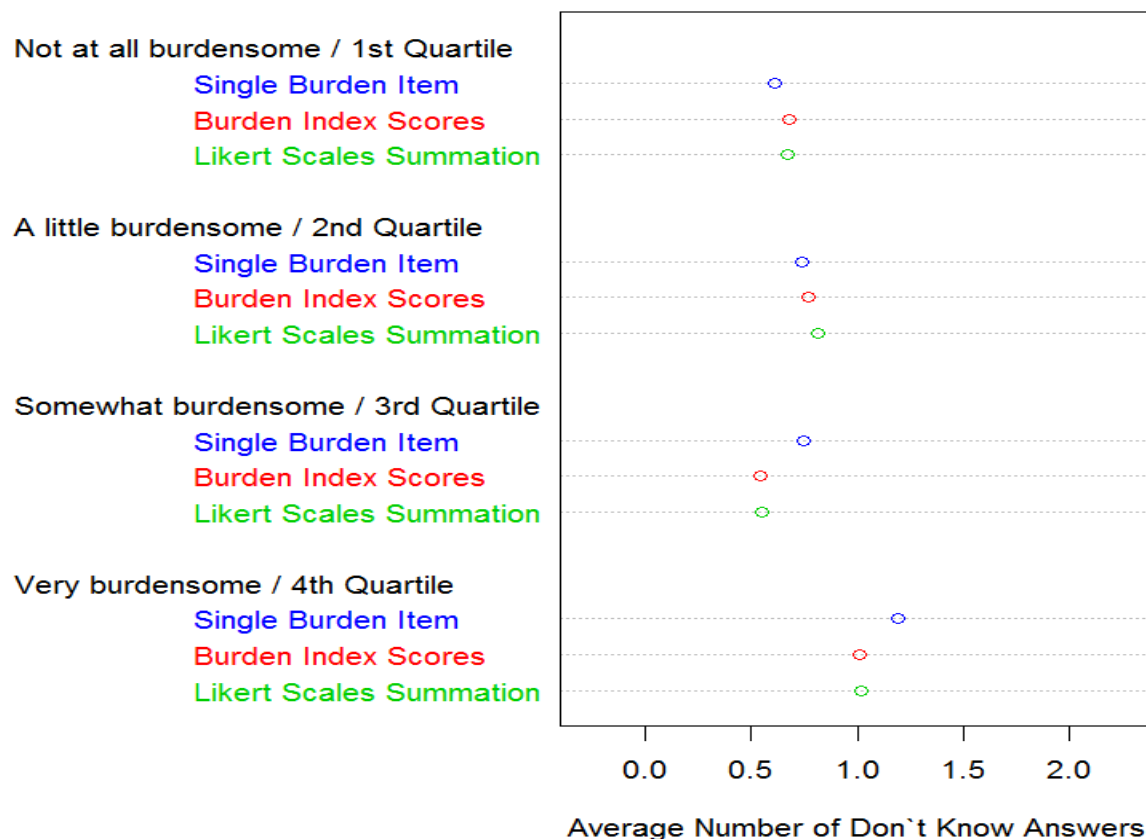
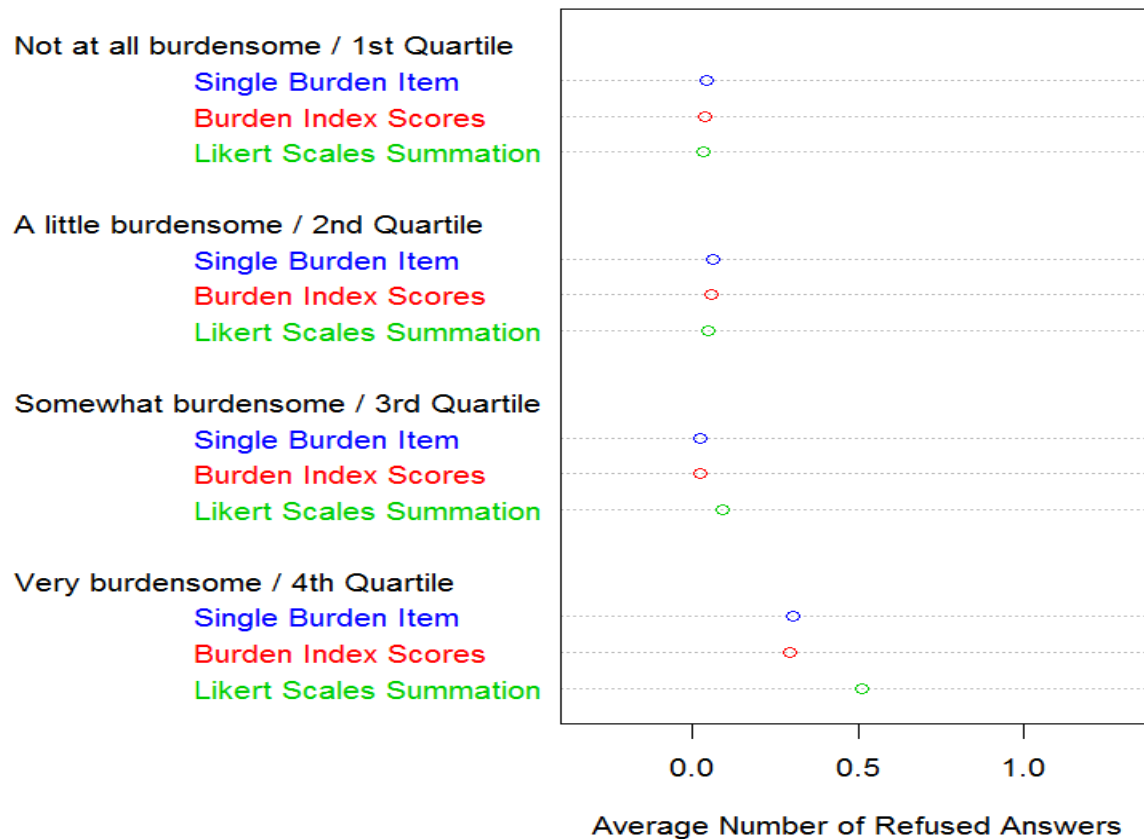


Figure 7. Average Number of “Refused” Answers by Levels of Single Burden Question (Item) and Burden Index Scores Quartile Groups



Williams' t-test: Compare Two Non-independent Correlations

- Let n be the sample size, let r denote correlation and let j , k and h index three separate samples, respectively, then

- $$t_{n-3} = (r_{jk} - r_{jh}) \sqrt{\frac{(n-1)(1+r_{kh})}{2\left(\frac{n-1}{n-3}\right)|R| + \frac{(r_{jk}+r_{jh})^2}{4}(1-r_{kh})^3}}$$

- where $|R| = 1 - r_{jk}^2 - r_{jh}^2 - r_{kh}^2 + 2r_{jk}r_{jh}r_{kh}$.

Table 2. Compare Non-Independent Correlations (ρ) with a Variable in Common (NUMDK/NUMRF among Burden Measures)

Common Variable: Indirect Indicators of Data Quality	Burden Measures	ρ (jk, jh)	Williams' t- test p-value of $H_0: \rho_{jk} = \rho_{jh}$
Number of Do Not Know (NUMDK, j)	Burden Question (k)	0.07	
	(h): Composite Burden Index	0.10	0.03
	Likert Scale Sum	0.10	0.03
Number of Refused (NUMRF, j)	Burden Question (k)	0.12	
	(h): Composite Burden Index	0.15	0.26
	Likert Scale Sum	0.14	0.27

Evaluation of Single Burden Question (Item) vs. Index Scores

- III. Compare methods to identify most-burdened respondents (via burden measures) and the effects of burden measures on statistical estimates of expenditures.

Effects of Burden Measures on Expenditure Estimates

- ❑ We examined whether or not the impact of burden index scores on expenditure estimates would be different from the single burden question (item).
- ❑ We computed the mean estimates with all respondents and the mean estimates without “most-burdened” respondents for each expenditure variable in two ways:

Effects of Burden Measures on Expenditure Estimates

- ❑ Excluding the “Very burdensome” level for the single burden question (item),
- ❑ Excluding the 4th quartile level for the composite burden index scores and the simple Likerts Scale summation scores, respectively.

Note: Standard error of mean (SE) was computed after accounting for complex design and weights.

Table 3A. Complex Weighted Estimates of Mean Expenditures with and without Most-Burdened Respondents

	Mostly in Person						
	With	Single Burden Question (Item)		Composite Burden Index Scores		Likerts Scale Summation Scores	
Expenditures (\$)		Without*	Difference	Without†	Difference	Without‡	Difference
Total Expenditure (SE)	12268 (204)	12269 (216)	1	12399 (236)	131	12328 (225)	60
Food (SE)	1812 (27)	1789 (25)	-23	1780 (24)	-32	1772 (22)	-40
Housing (SE)	3871 (54)	3859 (56)	-12	3854 (63)	-17	3839 (61)	-32
Transportation (SE)	2276 (70)	2300 (84)	24	2366 (97)	90	2357 (95)	81

Note: for without Burdened-out Respondents:

* Burden Item excludes level 4 (Very burdensome),

† Burden Index scores < 16.08,

‡ Likerts Scale Sum < 25



Table 3B. Complex Weighted Estimates of Mean Expenditures with and without Most-Burdened Respondents

	With	Mostly Telephone				Likerts Scale Summation Scores	
		Single Burden Question (Item)		Composite Burden Index Scores			
Expenditures (\$)		Without*	Difference	Without†	Difference	Without‡	Difference
Total Expenditure (SE)	12296 (228)	12213 (226)	-83	12365 (246)	69	12340 (253)	44
Food (SE)	1868 (26)	1859 (29)	-9	1847 (30)	-21	1852 (27)	-16
Housing (SE)	4219 (91)	4174 (99)	-45	4208 (109)	-11	4206 (109)	-13
Transportation (SE)	2103 (107)	2079 (111)	-24	2081 (1467)	-22	2074 (145)	-29

Conclusion

- ❑ The single burden question (item, 'bbur') can be used as a sole indicator of respondent's perceived burden for monitoring respondent's burden perception over time based on available data.
- ❑ If possible, include the composite burden index scores as a “check point” in case when the simple Likert Scale summation scores behave differently from the single burden question (item).

Proposed Follow-up Steps

- ❑ Exploring Burden Item's Proxy-Indicators: e.g. household income before tax, number of expenditures (unedited), mortgage indicator, whether it is a converted refusal, information booklet usage, records usage, interview mode and CHI variables: interview length and door step concerns?
- ❑ How does indirect indicators of data quality correlate with the single burden question (item) comparing to those proxy-indicator candidates?
- ❑ How to quantify the associations? (e.g. extrapolate into a new data set?)

Acknowledgement

- ❑ Branch of Research and Program Development (BRPD), DCES.
 - Adam N. Safir, Chief (DCES)
 - Laura P. Erhard, Supervisory Economist (BRPD)
- ❑ Office of Survey Methods Research (OSMR)
 - Wendy L. Martinez, PhD, Director, Mathematical Statistics Research Center (MSRC)
 - Brandon M. Kopp, PhD, Research Psychologist
 - Scott S. Fricker, PhD, Research Psychologist

THANK YOU!

CONSUMER EXPENDITURE SURVEY (CE)
Survey Methods
SYMPOSIUM
AND
Microdata Users'
WORKSHOP



Contact Information

Daniel K. Yang

Research Mathematical Statistician

Office of Survey Methods Research (OSMR)

www.bls.gov/osmr/home.htm

yang.daniel@bls.gov

Disclaimer: Any opinions expressed in this paper are those of the author(s) and do not constitute policy of the Bureau of Labor Statistics.