2022 FCSM Research & Policy Conference

Matrix Sampling in General Population Web Surveys

Z. Tuba Suzer-Gurtekin

Zhen Sun

Yuxuan Chen

Xinran Wang

James Wagner

Joanne Hsu

Outline

- Motivation
- Research design
- Preliminary results
- Future research

Motivation- Questionnaire length

- Nonresponse
  - Perceived burden
    - Mention of questionnaire length
    - Visual cues
  - Alternative survey modes and devices
    - Once the interview starts, breakoffs are not likely in interviewer-administered modes
    - Assess respondents' situation and respondent motivation, the same reasons for data quality
  - Self-administered modes (Eisele et al., 2022; Lugtig & Luiten, 2021)
    - No interviewer recruitment
    - Shorter questionnaires
- Measurement error
  - Higher respondent fatigue and burden as the questionnaire is longer (Herzog and Bachman, 1981; Adams and Gale, 1982; Peytchev and Peytcheva, 2017; Andreadis and Kartsounidou, 2020)
- Fitness of use:
  - Robustness to disruptions in communication channels that would allow to mode switching

Matrix Sampling (Split Questionnaire Design)

- Split Questionnaire Designs (SQDs) (Gonzalez and Eltinge, 2007; Adiguzel and Wedel, 2008))
  - First method:
    - Examine correlations among data based on complete questionnaire and identify those that are most related
    - Allocate questions with high correlations to different sub-questionnaires
    - Use **multiple imputation** to analyze the data collected from the multiple matrix sampling forms.
  - Second method:
    - To develop an algorithm that would automatically distribute items among a set number of forms based on some constraint.
    - Index of predictive value= (difference between variance of the no-imputation estimator and variance of imputation estimator)/(difference between variance of no imputation estimator and variance that would have been obtained with complete data) (Thomas et al. (2006))
  - Third method:
    - Determine optimal number of forms given quantified information loss (Stuart and Yu, 2022)

Matrix Sampling (Split Questionnaire Design)

- **Variation (1): Constrained by “high priority” questions: there are also questions that may not be randomly or statistically allocated to different forms**
- **Variation (2): Subjective judgment determines an optimal number of forms**

Constraints/Gaps

- Ongoing surveys
  - Multi-item indicators
    - Scientific
      - Example: Need for Cognition Scale (long vs. short)
    - Comparability
  - Covering different domains
    - Monthly Surveys of Consumers
    - Future income prospects: Optimism+Certainty
    - Consumer sentiment, income expectations, unemployment expectations, inflation expectations
  - Context effects (Schuman & Presser ,1981)
    - Validity and reliability of responses
    - Bivariate associations
    - Form resistant correlations
  - Estimation
    - Imputation vs. no imputation
  - Survey weights

Need for Cognition Scale

Table 1
18-Item Need for Cognition Scale

Item Number	Item Wording
1	I would prefer complex to simple problems.
2	I like to have the responsibility of handling a situation that requires a lot of thinking.
3	Thinking is not my idea of fun.*
4	I would rather do something that requires little thought than something that is sure to challenge my thinking abilities.*
5	I try to anticipate and avoid situations where there is likely chance I will have to think in depth about something.*
6	I find satisfaction in deliberating hard and for long hours.
7	I only think as hard as I have to. *
8	I prefer to think about small, daily projects to long-term ones.*
9	I like tasks that require little thought once I've learned them.*
10	The idea of relying on thought to make my way to the top appeals to me.
11	I really enjoy a task that involves coming up with new solutions to problems.
12	Learning new ways to think doesn't excite me very much.*
13	I prefer my life to be filled with puzzles that I must solve.
14	The notion of thinking abstractly is appealing to me.
15	I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought.
16	I feel relief rather than satisfaction after completing a task that required a lot of mental effort.*
17	It's enough for me that something gets the job done; I don't care how or why it works.*
18	I usually end up deliberating about issues even when they do not affect me personally.

* Reverse scoring is used on this item.

Constraints/ Gaps

Case Study: Design (I)

Question	Pair	Type	Present	Future	Past	12 months	5-10 years	Absolute vs. Comparative
A2-A2a	1	Personal			x	x		C
A2b	1	Personal			x		x	C
A3	2	Personal		x		x		C
A3b	2	Personal		x			x	C
A4	3	Business		x		x		A
A5	4	Business			x	x		C
A6-A6a	5	News		x				A
A7	6	Business		x		x		C
A8	6	Business		x			x	A
A9	7	Govt						A
A10	7	Unemp		x		x		C
A11	7	Int		x		x		C
A12-A12a-A12b-A12b10-A12c	8	Price		x		x		C
A13-A13a-A13b-A13b10-A13c	8	Price						
A14	9	Real Income Exp.		x		x		C
A15-A15a	9	Income exp.		x		x		C
A16-A16a	10	Home buying	x					A
A17-A17a	10	Home selling	x					A
A18-A18a	11	Durables	x					A
A19-A19a	11	Vehicle	x					A
A20-A20a	11	Gas price		x			x	C
A20b-A20c	11	Gas price		x		x		C

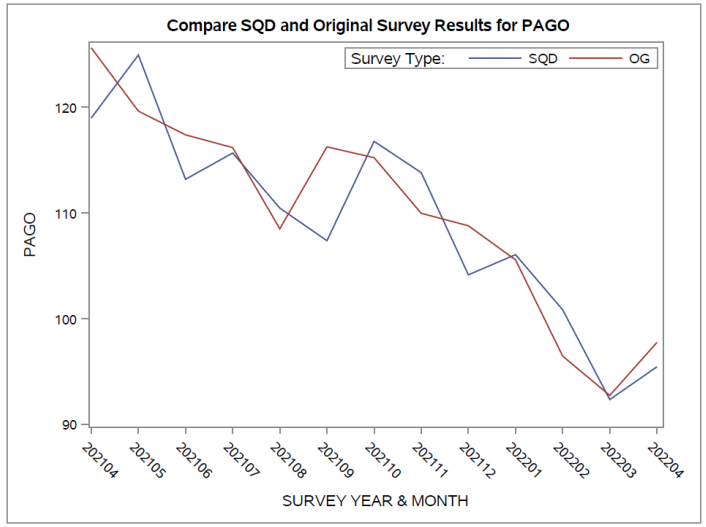
Design (II): SQDs

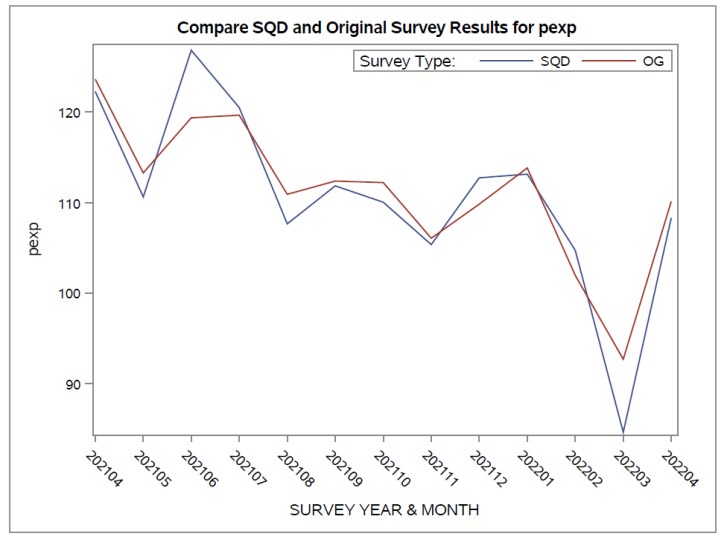
SQD	Var1	Var2	Var3	Var4	Var5	Var6	Var7	Var8	Var9	Core1	Core2	Core3
1	A2-A2a	A3	A7	A11	A13-A13a-A13b-A13b10-A13c	A15-A15a	A17-A17a	A19-A19a	A20b-A20c	A4	A5	A6-A6a
2	A2b	A3b	A7	A11	A13-A13a-A13b-A13b10-A13c	A15-A15a	A17-A17a	A19-A19a	A20b-A20c	A4	A5	A6-A6a
3	A2b	A3	A7	A10	A13-A13a-A13b-A13b10-A13c	A14	A16-A16a	A19-A19a	A20-A20a	A4	A5	A6-A6a
4	A2b	A3b	A7	A11	A12-A12a-A12b-A12b10-A12c	A14	A16-A16a	A18-A18a	A20-A20a	A4	A5	A6-A6a
5	A2b	A3b	A7	A9	A12-A12a-A12b-A12b10-A12c	A14	A16-A16a	A18-A18a	A20b-A20c	A4	A5	A6-A6a
6	A2-A2a	A3	A8	A9	A13-A13a-A13b-A13b10-A13c	A14	A17-A17a	A19-A19a	A20-A20a	A4	A5	A6-A6a
7	A2-A2a	A3b	A7	A10	A13-A13a-A13b-A13b10-A13c	A15-A15a	A16-A16a	A19-A19a	A20b-A20c	A4	A5	A6-A6a
8	A2b	A3	A8	A11	A12-A12a-A12b-A12b10-A12c	A14	A16-A16a	A18-A18a	A20b-A20c	A4	A5	A6-A6a
9	A2b	A3b	A8	A9	A13-A13a-A13b-A13b10-A13c	A15-A15a	A16-A16a	A19-A19a	A20b-A20c	A4	A5	A6-A6a
10	A2-A2a	A3b	A7	A9	A12-A12a-A12b-A12b10-A12c	A14	A16-A16a	A19-A19a	A20b-A20c	A4	A5	A6-A6a
11	A2b	A3b	A8	A10	A13-A13a-A13b-A13b10-A13c	A15-A15a	A16-A16a	A18-A18a	A20-A20a	A4	A5	A6-A6a
12	A2b	A3b	A7	A10	A12-A12a-A12b-A12b10-A12c	A14	A16-A16a	A19-A19a	A20-A20a	A4	A5	A6-A6a
13	A2-A2a	A3	A8	A10	A13-A13a-A13b-A13b10-A13c	A15-A15a	A17-A17a	A19-A19a	A20b-A20c	A4	A5	A6-A6a
14	A2-A2a	A3	A8	A9	A12-A12a-A12b-A12b10-A12c	A14	A16-A16a	A19-A19a	A20-A20a	A4	A5	A6-A6a
15	A2b	A3	A7	A11	A13-A13a-A13b-A13b10-A13c	A15-A15a	A17-A17a	A18-A18a	A20b-A20c	A4	A5	A6-A6a
16	A2-A2a	A3b	A8	A11	A13-A13a-A13b-A13b10-A13c	A14	A17-A17a	A18-A18a	A20b-A20c	A4	A5	A6-A6a
17	A2-A2a	A3b	A8	A11	A12-A12a-A12b-A12b10-A12c	A14	A17-A17a	A19-A19a	A20-A20a	A4	A5	A6-A6a
18	A2b	A3b	A8	A9	A13-A13a-A13b-A13b10-A13c	A15-A15a	A16-A16a	A18-A18a	A20-A20a	A4	A5	A6-A6a
19	A2b	A3	A8	A11	A13-A13a-A13b-A13b10-A13c	A15-A15a	A17-A17a	A19-A19a	A20b-A20c	A4	A5	A6-A6a

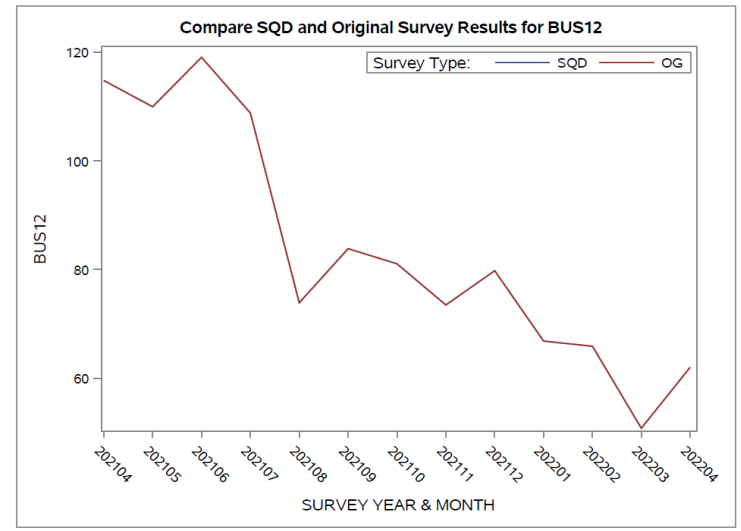
Results

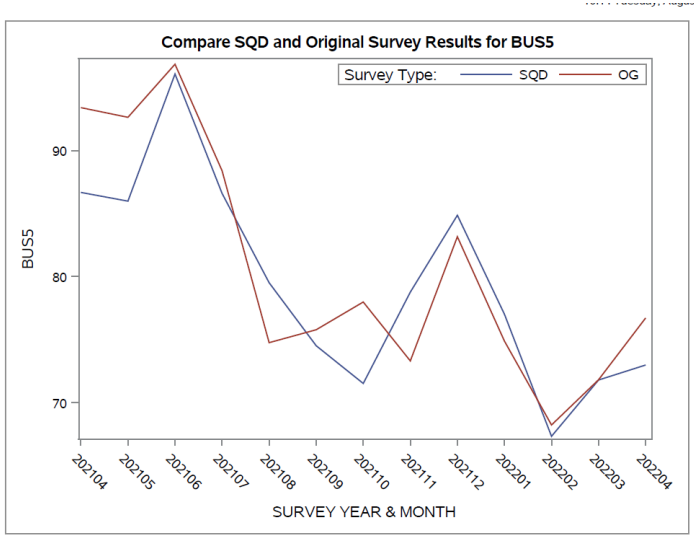
- Time Series Correlations

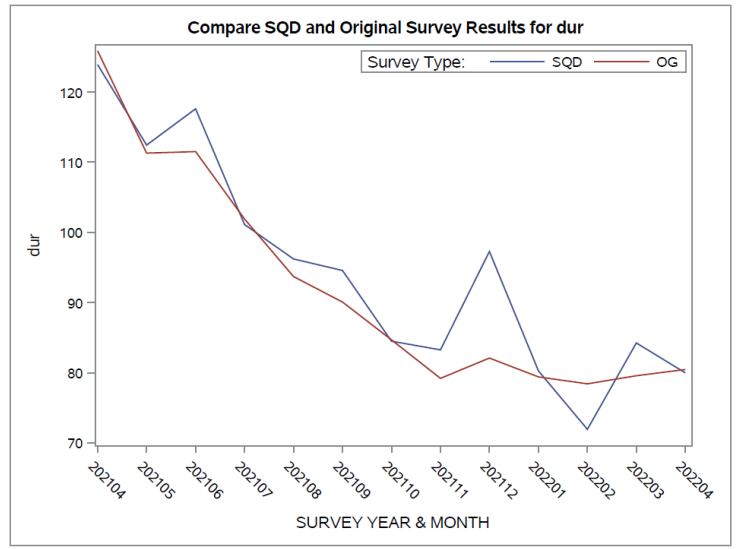
Variable	Question Number	Correlation
Personal finances a year ago	A2 (PAGO)	0.90
Personal finances in the next year	A2b (PEXP)	0.95
Business condition in the next 12 months	A4 (BUS12)	1.00
Business condition in the 5 years	A5 (BUS5)	0.90
Household durable purchase conditions	A18 (DUR)	0.95
Short term inflation expectations	A12 (PX1)	0.95
Long term inflation expectations	A13 (PX5)	0.93
Unemployment expectations	A10 (UMEX_M)	0.82

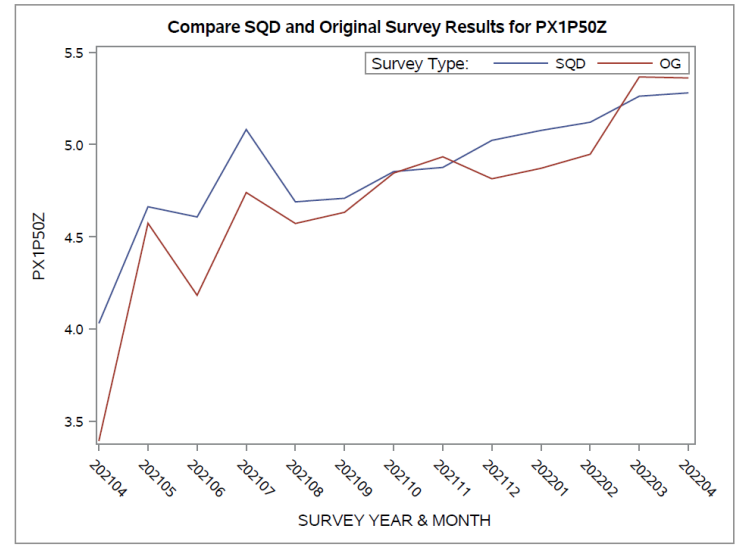

Compare SQD and Original Survey Results for PAGO
Survey Type: SQD OG
PAGO
120
110
100
90
202104 202105 202106 202107 202108 202109 202110 202111 202112 202201 202202 202203 202204
SURVEY YEAR & MONTH


Compare SQD and Original Survey Results for pexp
Survey Type: SQD OG
pexp
120
110
100
90
202104 202105 202106 202107 202108 202109 202110 202111 202112 202201 202202 202203 202204
SURVEY YEAR & MONTH


Compare SQD and Original Survey Results for BUS12
Survey Type: SQD OG
BUS12
120
100
80
60
202104 202105 202106 202107 202108 202109 202110 202111 202112 202201 202202 202203 202204
SURVEY YEAR & MONTH


Compare SQD and Original Survey Results for BUS5
Survey Type: SQD OG
BUS5
90
80
70
202104 202105 202106 202107 202108 202109 202110 202111 202112 202201 202202 202203 202204
SURVEY YEAR & MONTH


Compare SQD and Original Survey Results for dur
Survey Type: SQD OG
dur
120
110
100
90
80
70
202104 202105 202106 202107 202108 202109 202110 202111 202112 202201 202202 202203 202204
SURVEY YEAR & MONTH


Compare SQD and Original Survey Results for PX1P50Z
Survey Type: SQD OG
PX1P50Z
5.5
5.0
4.5
4.0
3.5
202104 202105 202106 202107 202108 202109 202110 202111 202112 202201 202202 202203 202204
SURVEY YEAR & MONTH

Future Research

- Constraint the sample size to 300 per block element
- Estimate variance
- Field Study to test context effects
  - Bivariate associations
  - Form resistant correlations

Contact Information

tsuzer@umich.edu

References

- Adigüzel, F., & Wedel, M. (2008). Split questionnaire design for massive surveys. Journal of Marketing Research, 45(5), 608-617
- Andreadis, I.& Kartsounidou, E. (2020) The Impact of Splitting a Long Online Questionnaire on Data Quality
- Eisele, G., Vachon, H., Lafit, G., Kuppens, P., Houben, M., Myin-Germeys, I., & Viechtbauer, W. (2022). The Effects of Sampling Frequency and Questionnaire Length on Perceived Burden, Compliance, and Careless Responding in Experience Sampling Data in a Student Population. Assessment, 29(2), 136–151. https://doi.org/10.1177/1073191120957102
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. Advances in Experimental Social Psychology, 19(C), 123–205. https://doi.org/10.1016/S0065-2601(08)60214-2
- Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The Efficient Assessment of Need for Cognition. Journal of Personality Assessment, 48(3), 306–307. https://doi.org/10.1207/s15327752jpa4803_13
- Lugtig, P., & Luiten, A. (2021). Do shorter stated survey length and inclusion of a QR code in an invitation letter lead to better response rates ? Survey Methods: Insights from the Field.
- Schuman, H., & Presser, S. (1981). Questions and Answers in Attitude Surveys Experiments on Question Form, Wording, and Context. Academic Press, Inc.
- Stuart, M. & Yu C. (2022) A Computationally Efficient Method for Selecting a Split Questionnaire Design
- Thomas, N., Raghunathan, T.E., Schenker, N., Katzoff, M.J., & Johnson, C.L. (2006). An Evaluation of Matrix Sampling Methods Using Data from the National Health and Nutrition Examination Survey.