

Using paradata to identify response styles and respondent profiles for adaptive survey design

Barry Schouten - Stats Netherlands & Utrecht University

RSS meeting – *Paradata: From Survey Research to Practice*



Centraal Bureau
voor de Statistiek

Outline

- Motivation
- Response styles, paradata and profiles
- Case studies LFS 2008 and CVS 2011
- Case study LFS – LISS panel 2013-2014
- Discussion

Motivation – multi-mode surveys

Research and case studies mostly motivated from need to:

- Reduce mode effects;
- Include survey mode as design feature in adaptive survey design;
- Find heuristics to perform costly transition periods and parallel runs;

Options to reduce mode effects:

1. Prevent through questionnaire design;
2. Avoid through data collection design (possibly adaptive);
3. Adjust through the estimation design;
4. Stabilize through estimation design;

Questionnaire profiles, respondent profiles and response styles may help:

- To set a benchmark design;
- To reduce dimensions for multi-purpose, diverse surveys;
- To filter for pre-testing and for transition periods;
- To adapt to respondent;

Adaptive survey design

Adaptive survey designs and responsive survey designs are motivated by two observations:

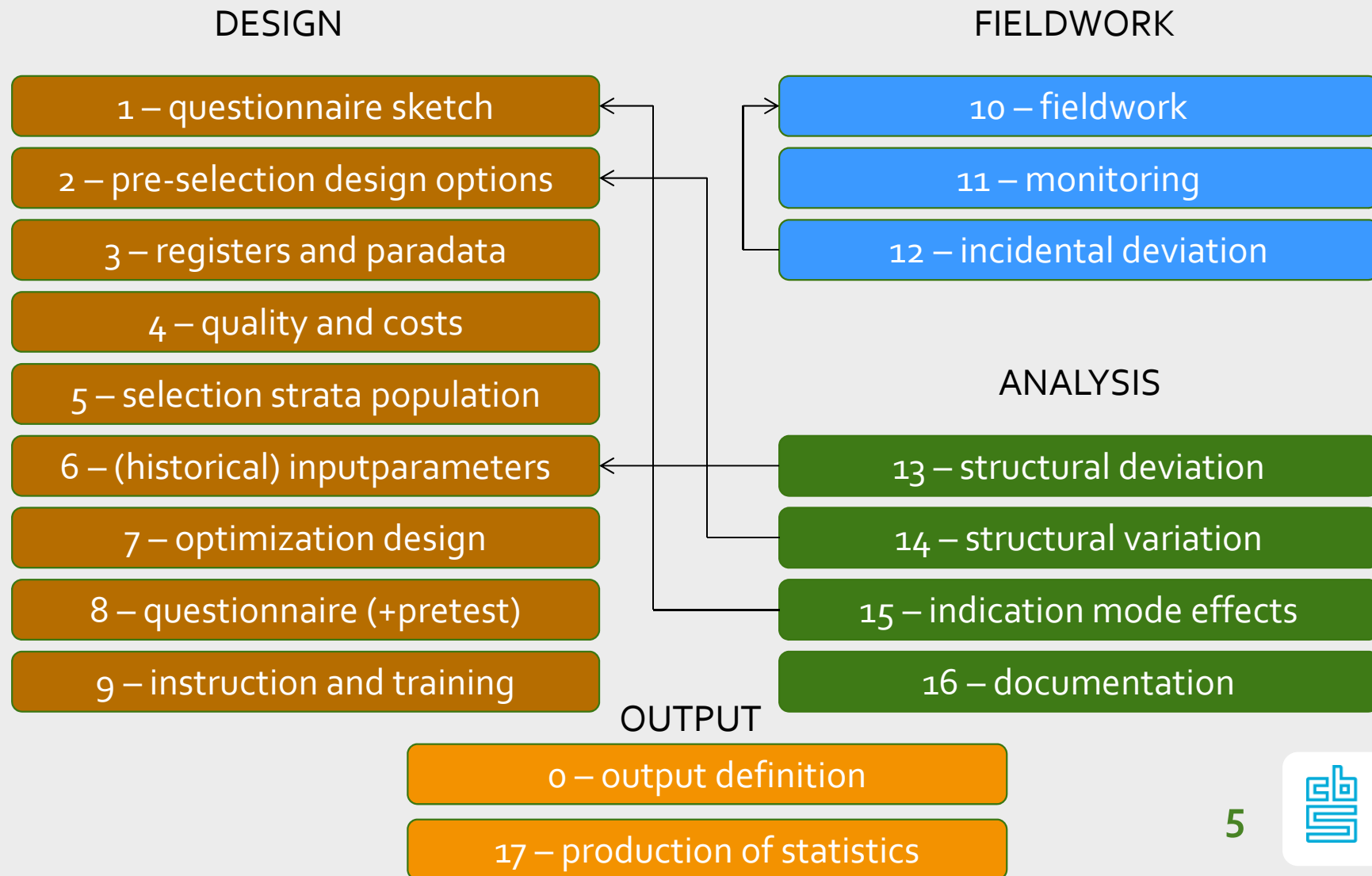
- Persons have different preferences for communication and interview, i.e. respond differently to different data collection strategies;
- Different data collection strategies are associated with different costs per person;

Scope/implementation/approach dependent on setting:

- Length of data collection period and number of instances for intervention
- Refusal conversion
- Strength of prior knowledge (frame/registry data, paradata previous waves)
- Focus on learning during data collection or from wave to wave
- Focus on structural and incidental deviations from optimal quality

So far, frameworks and implementation focussed on nonresponse error.

The data collection process



Response style propensities

Response style: A deficiency in one or more of the cognitive steps of the answering process that persists throughout a significant part of the interview;

Response style propensity: Propensity that a respondent with certain characteristics will show a response style for a certain questionnaire.

Strong analogy to a response propensity

Note:

- Multiple styles may occur in an interview;
- Styles may interact (strengthen, attenuate or mask each other);
- Styles may only be evoked by certain items or blocks of items;

Explaining answering behaviour

Three steps:

- Detect answering behaviour using paradata;
- Model answering behaviour with person level, survey item level and survey level characteristics;
- Evaluate impact of behaviour using validation data;

Model: Explain variation in answering behaviour using multi-level models:

- Between persons;
- Between items;
- Between surveys;

Low between item variation hints at response style

Low between survey variation hints at consistency of response style

Types of answering behaviour

Deviant (not necessarily erroneous) forms of answering behaviour:

- Respondent gives a social desirable answer
- Respondent gives an answer that avoids follow-up questions
- Respondent gives the same answer as to previous question
- Respondent gives one of the first answer options
- Respondent gives one of the last answer options
- Respondent answers more quickly than average
- Respondent answers slower than average
- Respondent gives a don't know answer
- Respondent provides no answer
- Respondent rounds the answer
- Respondent changes answer to a question

Paradata on answering behaviour

Web:

- Time measurements;
- Browser and device;
- Break-offs (Blaise \$work files);
- Audit trails;
- Mouse movements;

Telephone-F2F:

- Time measurements;
- Audit trails;
- CARI recordings;

Interviewer observations:

- Presence of other household members;
- Concentration of respondent (in general and changes during interview);
- Difficulty (assistance by interviewer);

Respondent profiles

Analogy with response propensities:

- Some person characteristics generally relate to answering behaviour;
- Some person characteristics only evoke answering behaviour for certain survey topics, certain survey designs and certain survey items;

What characteristics come into play in the cognitive steps of the answering process
(Interpretation & comprehension, Information retrieval, Judgment, Reporting)

General characteristics: educational level, age, ethnicity, personality traits;

Specific characteristics: motivation, relevance of survey topic;

Questionnaire profiles

Survey item = Introduction/help + question + answer options + context

Survey item question characteristics

Content of question (objective - behaviour, other objective, subjective - opinion, subjective - satisfaction, other subjective)

Time reference (past, present, future)

Complexity – length of question in words

Complexity – complicated words or language

Complexity – use of conditions and exceptions

Complexity – recall needed

Complexity – hypothetical setting or situation

Complexity – need for calculations

Complexity – double-barreled question

Emotional content (question may arouse emotions)

Question posed as statement

Item is filter question

Item may be perceived as being a filter question

Survey item sensitive to socially desirable answering

Centrality of survey item

Questionnaire profiles - continued

Survey item answer characteristics

- Mismatch between question and answer
- Number of answer categories
- Measurement level
- Complexity – difficult words or language in answer categories
- Ordinal scale – direction
- Ordinal scale – balance
- Ordinal scale – neutral category
- DK answer category available

Survey item context characteristics

- Item is part of a battery of items
- Position in questionnaire
- Explanation provided to survey item

Complications

There are number of complications in modelling answering behaviour:

- Survey items may not be randomized (or only to some extent);
- Direction of ordinal rating items may not be randomize;
- Questionnaire is not an arbitrary selection of items but a coherent, logical series of items (ideally);
- No true value may be available → validation data;
- Answering behaviour is unobservable process → paradata;
- Available person characteristics may be limited;
- Survey item characteristics may not fully capture context;
- Survey item characteristics may have low reliability among coders;

Case study 1 – LFS 2008

Motivated underreporting on number of employments (given administrative data): maximize response rate with or without propensity to show underreporting subject to constraints on budget and representativeness.

Constraint R-indicator	Constraint number of visits	Realized R-indicator		Realized response rate	
		Without UR	With UR	Without UR	With UR
-	20000	0.718	0.499	50.5	48.5
	25000	0.924	0.885	62.4	59.6
	30000	0.939	0.937	64.2	61.5
0.80	20000	0.806	0.804	50.0	47.9
	25000	0.924	0.885	62.4	59.6
	30000	0.939	0.937	64.2	61.5
0.85	20000	0.872	0.919	48.1	47.5
	25000	0.924	0.885	62.4	59.6
	30000	0.939	0.937	64.2	61.5
0.90	20000	0.911	0.919	46.5	47.5
	25000	0.924	0.917	62.4	59.5
	30000	0.939	0.937	64.2	61.5



Case study 1 – LFS 2008 - continued

Motivated underreporting: maximize response rate given constraints on underreporting propensity, budget and representativeness.

Constraint underreporting	Constraint R-indicator	Constraint number of visits	Realized R-indicator	Realized response rate
3.5%	0.80	20000	0.862	40.6
		25000	0.852	49.3
		30000	0.820	59.7
	0.85	20000	0.862	40.6
		25000	0.852	49.3
		30000	0.852	49.3
	0.90	20000	-	-
		25000	-	-
		30000	-	-
4.0%	0.80	20000	0.885	44.3
		25000	0.816	55.8
		30000	0.928	63.5
	0.85	20000	0.885	44.3
		25000	0.893	54.2
		30000	0.928	63.5
	0.90	20000	0.935	42.8
		25000	0.905	52.6
		30000	0.928	63.5

Case study 2 – CVS 2011

- In 2011 multi-mode experiment was linked to Crime Victimization Survey;
- Experiment consisted of randomized allocation of sample persons to four modes (Web, mail, phone, F2F) and full sample received a F2F follow-up;
- Follow-up questionnaire with various survey attitudes and repeated key CVS variables;
- Experiment allowed a decomposition of mode-specific selection and measurement bias for key variables;

Research questions:

1. Can we model mode-specific answering behaviour with item characteristics?
2. Can we model mode-specific measurement bias with item characteristics?

BUT:

- Limited paradata;
- One questionnaire;
- No validation data;

Case study 2 – CVS 2011

- Mode-specific measurement bias estimated at the category level for 125 CVS survey items (excluding LFS survey items);
- A number of response styles that are assumed to be mode-related was investigated: DK-answers, primacy, recency, extreme response style and straightlining. DK-answers very strongly mode-dependent and clusters strongly within individuals. Other styles less pronounced but some traces were found.
- CVS study too limited: Replication needed over respondents and wider range of surveys and survey items.

Case study 3 – LFS & LISS panel 2013-2014

Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
Health	LFS1	Religion & ethnicity	Social intrst & leisure	LFS2	Work & schooling	Personality	LFS-3
LISS02	Politics & values	LISS22	LISS32	Family & household	LISS52	LISS62	Income/ living/asset
LISS03	LISS13	LISS23	LISS33	LISS43	LISS53	LISS63	LISS73
	LFS-HH1 Web CATI CAPI			LFS-HH2 CATI			LFS-HH3 CATI
	LFS-Person Web Cati Web→Cati						
Paradata (logfiles, time measurements, interviewers observations)							
Validation data = Administrative data Social Statistical Database							



Case study 2 – CVS 2011

Case study design:

- Multiple measurements on same respondents, and some of the items are repeated or very similar (LFS);
 - Rich set of person characteristics through panel surveys and recruitment, and through linkage to administrative data;
 - Validation data available from administrative data for some items;
 - Paradata measurements in all modes;
 - Randomization over reporting (self versus proxy) condition;
 - Parallel to LFS at Statistics Netherlands → multiple survey modes with some randomization;
-
- No randomization within questionnaires;
 - Paradata specific to survey modes;

Case study 3 - Methodological issues

So far some methodological issues have been identified:

- Relatively low coder reliability on key survey item characteristics: content of survey item, sensitivity, complexity of language in question and centrality;
- Filter questions;
- Aggregation of survey item characteristics and answering behaviour to questionnaire block level and to questionnaire overall level;
- Simultaneous occurrence of multiple response styles/forms of answering behaviour;

Discussion

What is most powerful paradata in detecting answering behaviour?

- Time measurements;
- Audit trails/logfiles;
- Interviewer observations;
- Mouse movements;
- Other;

How to deal with mode-dependency of paradata?

Other points for discussion:

- Is there consistency in answering behaviour over multiple surveys?
- Can we deduce questionnaire and respondent profiles?
- How to use such profiles in adaptive survey design?
- Methodological issues?
- Similar experiences and research projects?