Adaptive survey designs that minimize nonresponse and measurement risk

Melania Calinescu (VU University, Amsterdam)
Barry Schouten (Statistics Netherlands and University of Utrecht)

6th International Total Survey Error Workshop, Sept 2 – 4, Santpoort

Outline

- Why adaptive survey designs for NR and ME errors?
- The mathematical framework
- LFS case study
- Discussion

Schouten, Calinescu, Bhulai (2011), general framework for nonresponse
Luiten, Schouten (2012), example for Consumer Sentiments Survey

Adaptive survey designs for NR and ME

General motivation ASD: Different population subgroups react differently to different combinations of survey design features. Gain in efficiency is possible by explicit trade-off between cost and quality.

Setting at Statistics Netherlands:

- wide variety of registry data and to some extent also of paradata
- many recurring surveys with similar designs and topics
- pressure on survey budgets
- strong quality-cost differential for survey modes; focus on NR too simple due to mode effects

Adaptive survey designs for NR and ME

ASD's may be viewed as extensions to sampling designs as they consider multiple strategies and extend to non-sampling errors.
A TSE view is needed in order to include design features like the survey mode.

How to account for measurement when a survey has many key items?

Option: replace measurement error on individual items by measurement profile(s) over multiple items

Mathematical framework

Steps:

1. Choose quality and cost functions, Q and C
2. Identify candidate strategies, $S=\{ø,s1,...,sk\}$
3. Identify subgroups (registry/frame data, paradata), $X=\{x1,...,Xm\}$ with population distribution $q(x)$
4. Estimate parameters: response propensities $\Psi(x,s)$, measurement profile propensities $\Pi(x,s)$ and cost functions $c(x,s)$
5. Determine strategy allocation probabilities $p(s|x)$:
   - either, maximize quality Q given constraints on cost C
   - or, minimize cost C given constraints on quality Q
6. Monitor data collection closely, and possibly repeat steps

Mathematical framework

Static designs: employ registry data only

Dynamic designs: include paradata

Two options:

- Paradata collected before interview
- Paradata collected during interview; only useful for panels, multiple waves

Mathematical framework

Without ME error:

- Quality functions = response rate $\rho = \sum_{s,x} q(x)\, p(s \mid x) \rho(s,x)$

- Cost constraint = total budget $\sum_{s,x} q(x)\, p(s \mid x) c(s,x) \leq b$

- Optional constraint = representativeness of response (R-indicator)

$$R(X) = 1 - 2\sqrt{\sum_x q(x)\left(\sum_s p(s \mid x)\rho(s,x) - \rho\right)^2} \geq a$$

With ME error:

- Option 1: Quality function = rate of response without ME profile

$$\sum_{s,x} q(x)\, p(s \mid x) \rho(s,x)(1 - \theta(x,s))$$

- Option 2: Additional constraint = ME profile risk

$$\frac{1}{\rho} \sum_{s,x} q(x)\, p(s \mid x) \rho(s,x) \theta(x,s) \leq \Theta$$

Labour Force Survey – case study

- Restriction to CAPI first wave

- Strategies: reporting type (self or proxy) × number of visits (max 10)
- Three subgroups based on age: 15- 25, 26 – 55, 56 - 65
- Measurement profile: satisficing, derived from inconsistencies with employment registry data
- Cost constraint: expected number of visits
- Additional constraint: type of reporting not allowed to change
- Parameters: estimated from LFS 2008

Example of solution to ASD optimization

setting	Group	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	Group response (%)	Response (%)
$\alpha = 0.80$	15-25	1	2	0	0	0	0	0	0	0	0	32.7	
b=20000	26-55	1	2	2	0	0	0	0	0	0	0	46.9	44.3
$\Theta = 4.0\%$	56-65	2	1	0	0	0	0	0	0	0	0	48.0	

Self reporting =1, proxy reporting allowed = 2

Without ME

Min R	Max visits	Response rate	Number of visits	R-indicator	Profile risk
-	*20000*	50.5	19960	0.718	4.2
	25000	62.4	24954	0.924	4.5
	30000	64.2	26011	0.939	4.5
0.80	*20000*	50.0	19738	0.806	4.4
	25000	62.4	24954	0.924	4.5
	30000	64.2	25689	0.939	4.5
0.85	*20000*	48.1	19135	0.872	4.4
	25000	62.4	24954	0.924	4.5
	30000	64.2	25689	0.939	4.5
0.90	*20000*	46.5	18612	0.911	4.4
	25000	62.4	24954	0.924	4.5
	30000	64.2	25689	0.939	4.5

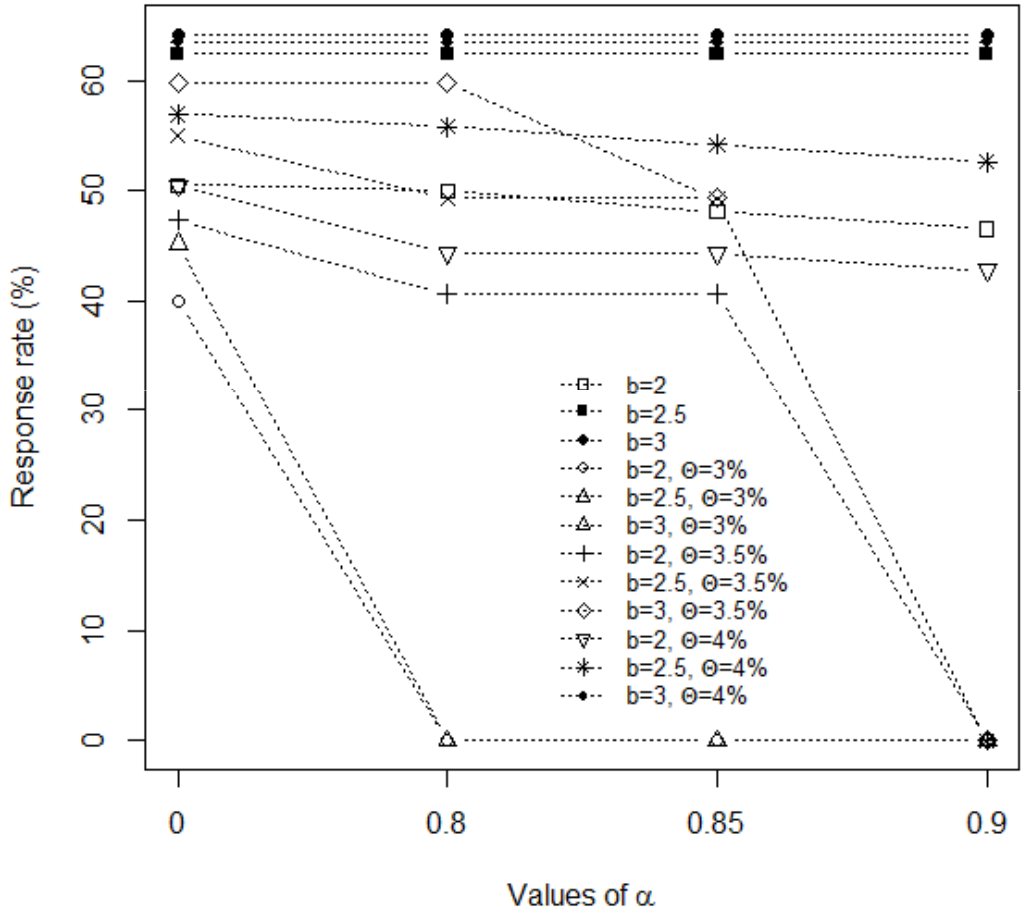
ME in quality objective function

Min R	Max visits	Response rate without profile	Number of visits	R-indicator
-	*20000*	48.5	19989	0.499
	25000	59.6	24995	0.885
	30000	61.5	29199	0.937
0.80	*20000*	47.9	19970	0.804
	25000	59.6	24995	0.885
	30000	61.5	29199	0.937
0.85	*20000*	47.5	19859	0.919
	25000	59.6	24995	0.885
	30000	61.5	29199	0.937
0.90	*20000*	47.5	19859	0.919
	25000	59.5	24957	0.917
	30000	61.5	29199	0.937

ME in cost constraint

Max risk	Min R	Max visits	Response rate	Number of visits	R-indicator	Profile risk
3.5%	0.80	20000	40.6	19847	0.862	3.5
		25000	49.3	24503	0.852	3.5
		30000	59.7	29317	0.820	3.5
	0.85	20000	40.6	19847	0.862	3.5
		25000	49.3	24503	0.852	3.5
		30000	49.3	24503	0.852	3.5
	0.90	20000	IN	FEA	SI	BLE
		25000	IN	FEA	SI	BLE
		30000	IN	FEA	SI	BLE
4.0%	0.80	20000	44.3	19521	0.885	4.0
		25000	55.8	24644	0.816	4.0
		30000	63.5	28809	0.928	4.0
	0.85	20000	44.3	19521	0.885	4.0
		25000	54.2	24879	0.893	4.0
		30000	63.5	28809	0.928	4.0
	0.90	20000	42.8	19892	0.935	4.0
		25000	52.6	23853	0.905	4.0
		30000	63.5	28809	0.928	4.0

Overview evaluation



Conclusions

- Adaptive survey designs including NR and ME errors can be formulated and optimized.
- Computation times modest but may be large when large number of subgroups is considered.
- Constraints on representativeness of response and risk of measurement profile may have strong impact on optimal solutions
- Inclusion of ME in quality objective function is more straightforward as it does not ask for constraints on measurement profile risks

Discussion

Side remarks

- ASDs depend strongly on precision of input probabilities and validity of cost evaluation: modesty is required, robustness and close monitoring are crucial
- ASDs optimize indirect measures of non-sampling error. Sampling error need not be ignored; still central quality component
- Choice of subgroups is very important: needs to be based on main publication domains and range of survey key variables

Questions:

- Are measurement profiles useful concepts?
- If so, should we aim for maximal response without such profiles or should we constrain the occurence of profiles?