Expert prior elicitation in Bayesian adaptive survey design: two case studies

Shiya Wu (Utrecht University)
Barry Schouten (Utrecht University and Statistics Netherlands)
Ralph Meijers (Statistics Netherlands)
Mirjam Moerbeek (Utrecht University)

6th Workshop on Advances in Adaptive and responsive survey design
US Census Bureau, November 4-5, 2019

Motivation

Context

- New survey or major redesign of existing survey
- Knowledgeable data collection staff
- Rich set of historic survey data
- Mix of administrative data and paradata

Data collection staff make predictions of response and costs on a continuous basis.

Goal: Mix data collection expertise and adaptive survey design procedures in a structured way to inform decisions

Data collection practice

Data collection staff:

- Select historic survey data with similar “features”
- Predict response rates based on a mix of the selected surveys
- Response rates predictions are detailed in case of stratified sampling (i.e. in most surveys)
- Predict costs given the anticipated design of the new survey and given the predicted response rates

- Do not account for uncertainty in response rate predictions
- Do not use a uniform procedure across staff members
- Do not consider relevant strata for nonresponse

Here, focus on prediction of stratum response propensities

Methodology

Perform a Bayesian analysis that constructs a prior with the help of data collection staff and updates the prior for observed data of the new survey.

Employ a power prior where the power α_k represents the similarity of a historic survey k to the new survey.

Response to a stratum g in data set k follows a Binomial distribution. Beta distribution is a conjugate prior to binomial distribution. Assuming a Beta(1,1), i.e. uniform prior, gives

$$\pi\left(\rho_g \middle| D_{k,g}, \alpha_k\right) \propto \rho_g^{\alpha_k r_{k,g}^0} \left(1-\rho_g\right)^{\alpha_k\left(n_{k,g}^0 - r_{k,g}^0\right)}$$

Expert prior elicitation – staff assistance

Data collection staff assists in four steps:

1. The construction of the list of design features on which surveys are compared
2. The choice of weights for the features to construct an overall score
3. The selection of the set of historic surveys included in the analysis
4. The actual scoring of the features for the selected historic surveys

Steps 1 and 2 are performed once or at low frequency. Steps 3 and 4 are repeated for each new survey/redesign

Expert prior elicitation – survey design features

- Topics/themes of the survey
- Target population
- Time elapsed since last fieldwork
- Unit of observation
- Mode strategy (including contact and reminder)
- Incentive strategy
- Questionnaire length (respondent effort)
- Bureau effect relative to Statistics Netherlands

Expert prior elicitation – feature importance

	Expert 1	Expert 2	Expert 3	Average	Weights
Topic	2	5	1	2.7	0.08
Population	3	4	4	3.7	0.15
Time	4	4	2	3.3	0.13
Observation	1	4	3	2.7	0.14
Mode	5	5	5	5.0	0.20
Incentive	4	4	4	4.0	0.16
Response Effort	2	1	1	1.3	0.05
Bureau Effect	2	1	3	2.0	0.08

Expert prior elicitation – construction of prior

Procedure:

1. Set a non-informative Beta prior to the historic survey data sets
2. Construct similarity parameters α_k for each historic survey k
   a) Ask three experts to independently derive similarity scores for each of the eight features.;
   b) Ask the three experts to reach a consensus on each of the eight features. Let $\alpha_{k,l}$ be the consensus score for survey k for feature l.
   c) Construct the overall score by weighting using weights w_l. The survey score becomes $\alpha_k = \sum_{l=1}^{8} w_l \alpha_{k,l}$.
3. Derive sample sizes $n_{k,g}$ and respondent numbers $r_{k,g}$ per stratum and data set
4. Construct the new survey Beta prior
5. Update the Beta prior of step 4 with data from each new wave of the new survey

Evaluation study

Consider two main aggregate quality indicators for ASD decisions

1. Weighted response rate
2. Coefficient of variation of response propensities over a set of relevant sample strata

Compare power prior approach to a non-informative prior using Mean Square Error (MSE) of prediction.

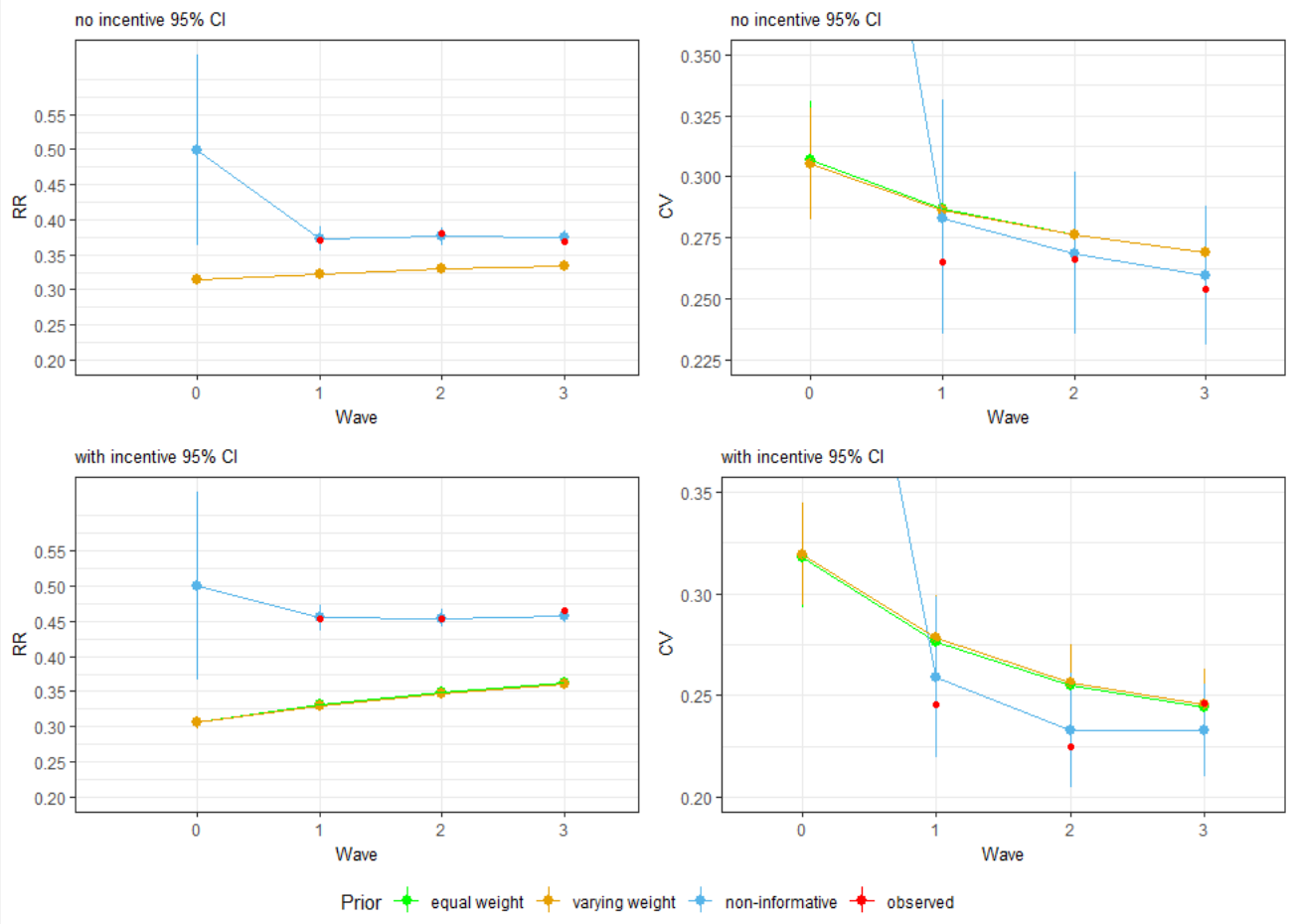
Questions:

- Does data collection expertise help?
- In particular, can we move to ASD decisions sooner when expert elicitation is performed?

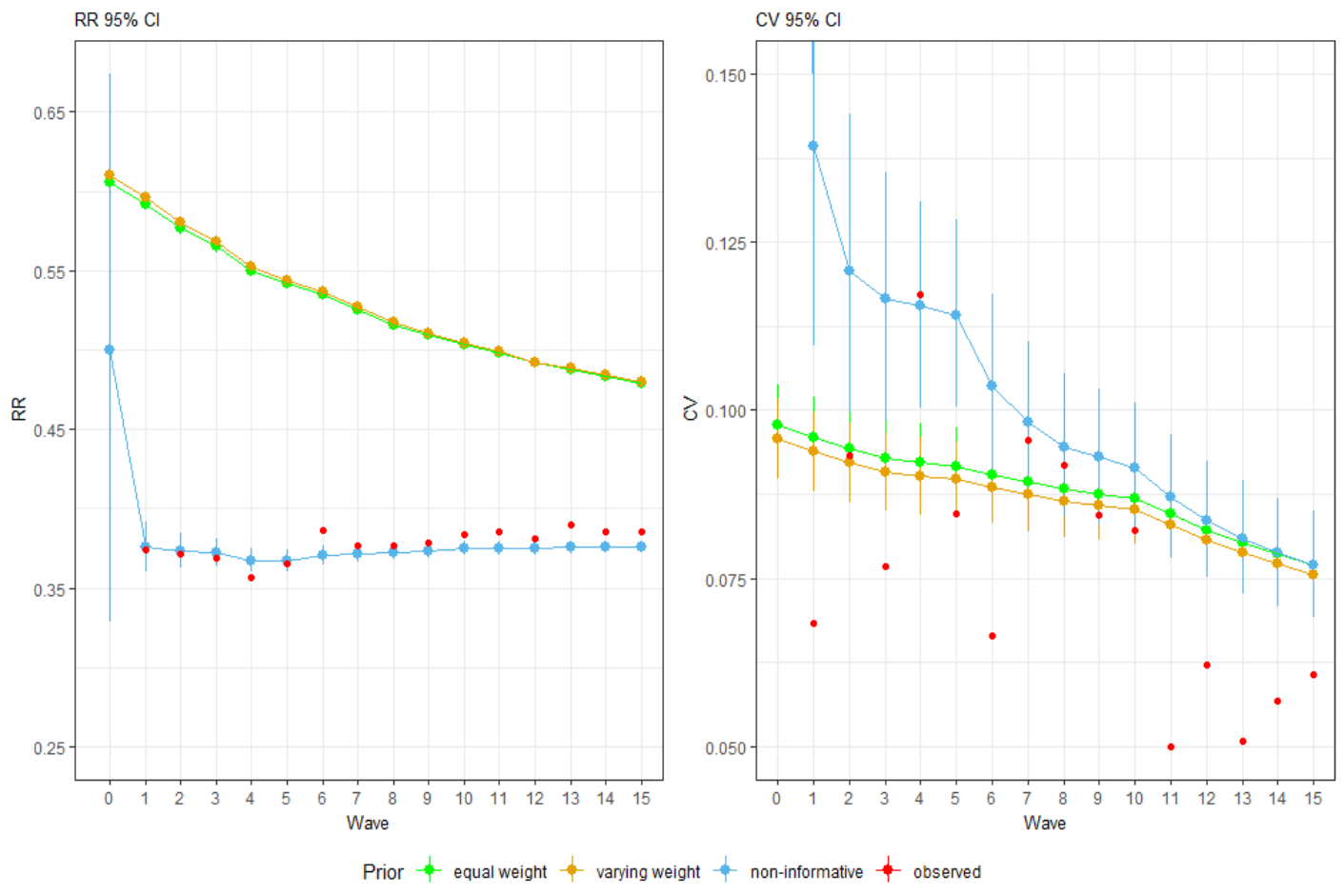
Two case studies

1. EU-SILC redesign 2016: Survey conducted since 2005 in a similar design. Redesign: 1) introduced web and a sequential mixed-mode design with telephone, 2) positioned SILC as a separate survey instead of sixth LFS wave;
2. Energy survey 2018: Infrequent survey performed by Stat Netherlands for the first time. Complex design with web, telephone and F2F;

EU-SILC


no incentive 95% CI
RR
Wave
no incentive 95% CI
CV
Wave
with incentive 95% CI
RR
Wave
with incentive 95% CI
CV
Wave
Prior
equal weight
varying weight
non-informative
observed

Energy survey 2018



Summary case studies

1. EU-SILC redesign 2016: Only at first wave does elicited prior help. At later waves, observed data is large/powerful enough to assist ASD decisions;
2. Energy survey 2018: Elicited response rate far off observed response rate, but CV closer than for non-informative prior;

All in all, modest support for added value of proposed prodecure of data collection expert prior elicitation

Potential improvement: Data collection staff build response rate predictions by adding relative shifts, e.g. 10% increase when introducing an incentive.

Discussion

1. Is the proposed prior elicitation approach practical/feasible?
2. Did we select the right survey design features?
3. How can we combine such an elicitation procedure with adaptation/intervention?