

Learn Time Change in Response Propensity from a Mixed-mode Historic Survey

Shiya Wu ^{1,2}

1. Utrecht University, Department of Methodology and Statistics
2. Statistics Netherlands, Department of Statistical Methods

Outline

1. Overview
2. Research Question
3. Time-series Model
4. A Case Study on the Dutch Health Survey (GEZO)
5. Discussion

Overview

Response Propensity

- the theoretical probability of a sampled unit/group to respond to a survey protocol
- necessary for monitoring (Kreuter 2013) and adapting (Schouten et al. 2017) survey designs
 - relates nonresponse to bias in a survey estimate
- estimated by a multivariate statistical model
 - provides the point estimate
 - without acknowledging the uncertainty

Overview

Recent Research Accounts for Uncertainty in Response Propensity(RP) Predictions

- In time stable context
 - Existing internal historic data sets
 - Bayesian (Schouten et al., 2018) updates predictions
 - Access to external historic data sets
 - Expert elicitation of data collection manager (Coffey et al., 2020) or staff (Wu et al., 2021) constructs prior knowledge
- In time-series context
 - Observe time change in response rates (e.g., downward trend, seasonality, etc.)
 - In a mixed-mode survey, modeling RP at each mode can benefit the adaption of survey designs

Research Question (RQ)

1. What time-series factors contribute to the variation and mode-related correlation in RPs?
2. How prediction accuracy depends on the length of historic survey time series?

Time-series Model

A Bayesian Multilevel Time-series Model (Boonstra & van den Brakel, 2019)

- Identify potential time-series factors linked to variations and correlations
- Use logit transformation for $[0,1]$ response propensities to infinite signal
- Signal takes the linear regression of identified factors

The Dutch Health Survey (GEZO)

- A stratification of Age and Ethnicity forms 13 subgroups, each of them approached via Web and then F2F
- For RQ1
 - 2014 – 2019 Web historic data to select the “best” model by criteria
 - The widely applicable information criterion (WAIC) for a reasonable balance between model fit, model complexity and efficient computation
 - posterior predictive p-values of mean and variance to check model adequacy
 - 2014 – 2017 Web-F2F data to analyze if each of the “best” model components is reasonably correlated with mode
- For RQ2
 - 2014 – 2019 Web historic data to evaluate the length-dependent performance by criteria
 - Root mean squared error (RMSE), bias (B), and standard deviation (SD)
 - Compare the length-dependent performance of F2F evaluated by F2F data to Web-F2F data of 2014-2017, i.e., individual model vs overall model

RQ1: Primary Time-series Model Components

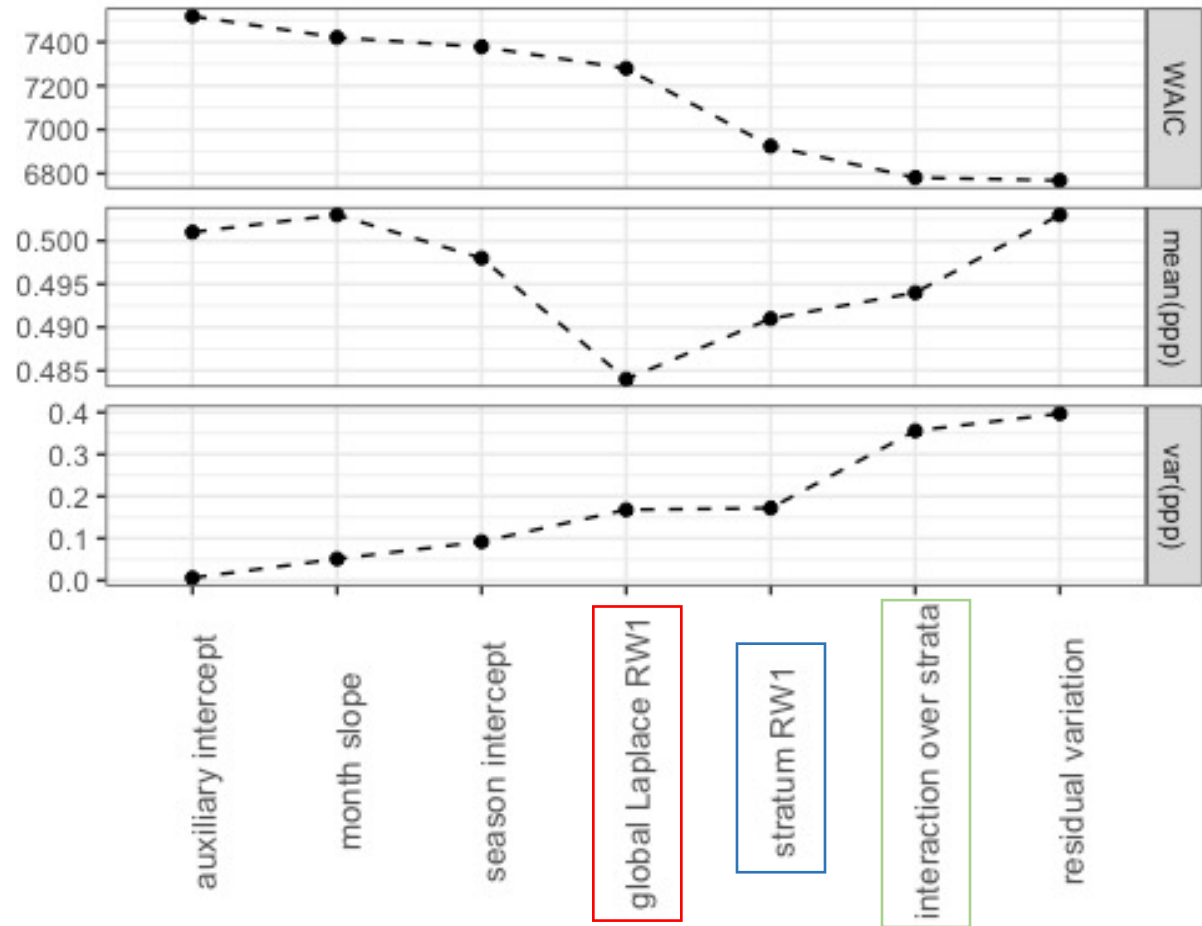


Figure 1.1. Optimize the model components by cumulatively adding the current factor to its left side. The “best” formulation with minimum WAIC (Widely Applicable Information Criteria) and close-to-0.5 mean of posterior predictive p-value (ppp).

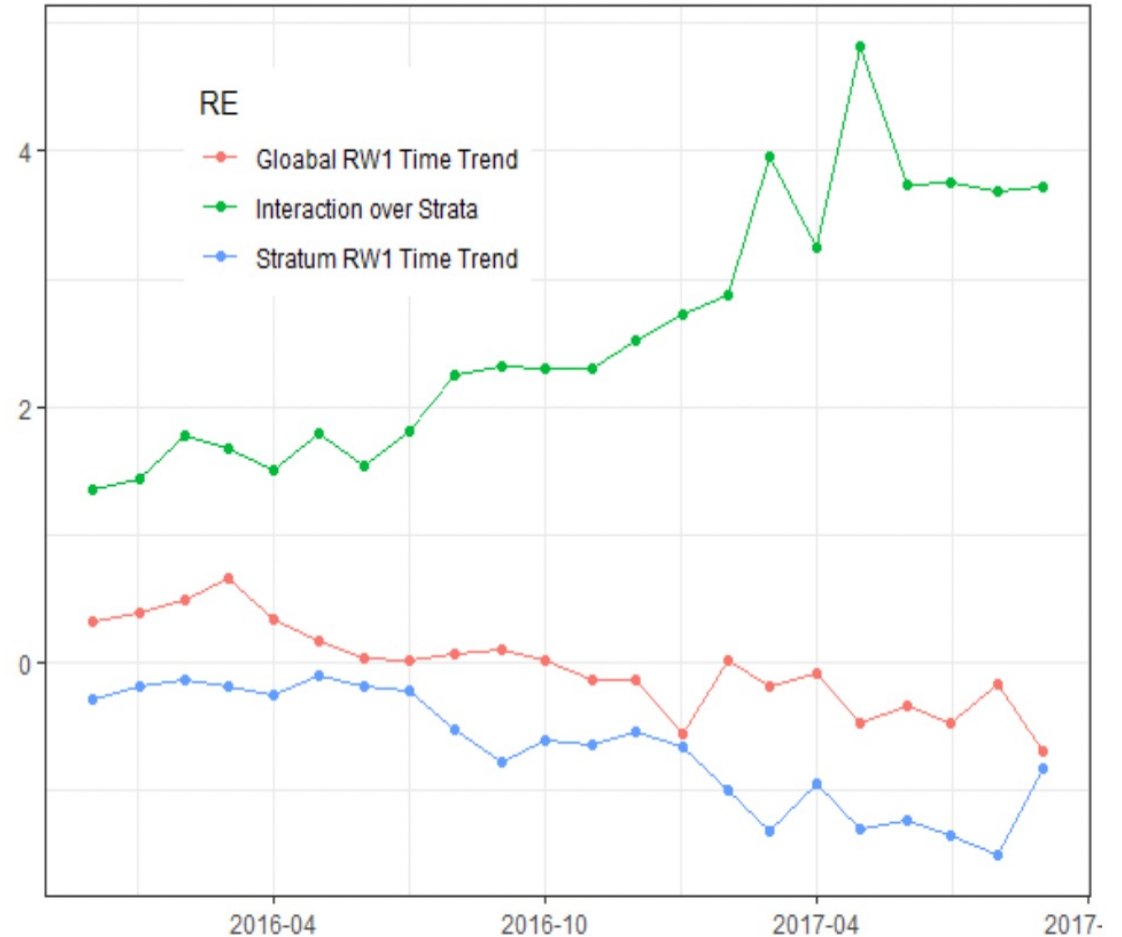


Figure 1.2. T-values of correlations in RPs between Web and F2F are learned from cumulative historic data.

RQ2: Prediction accuracy as a function of time length

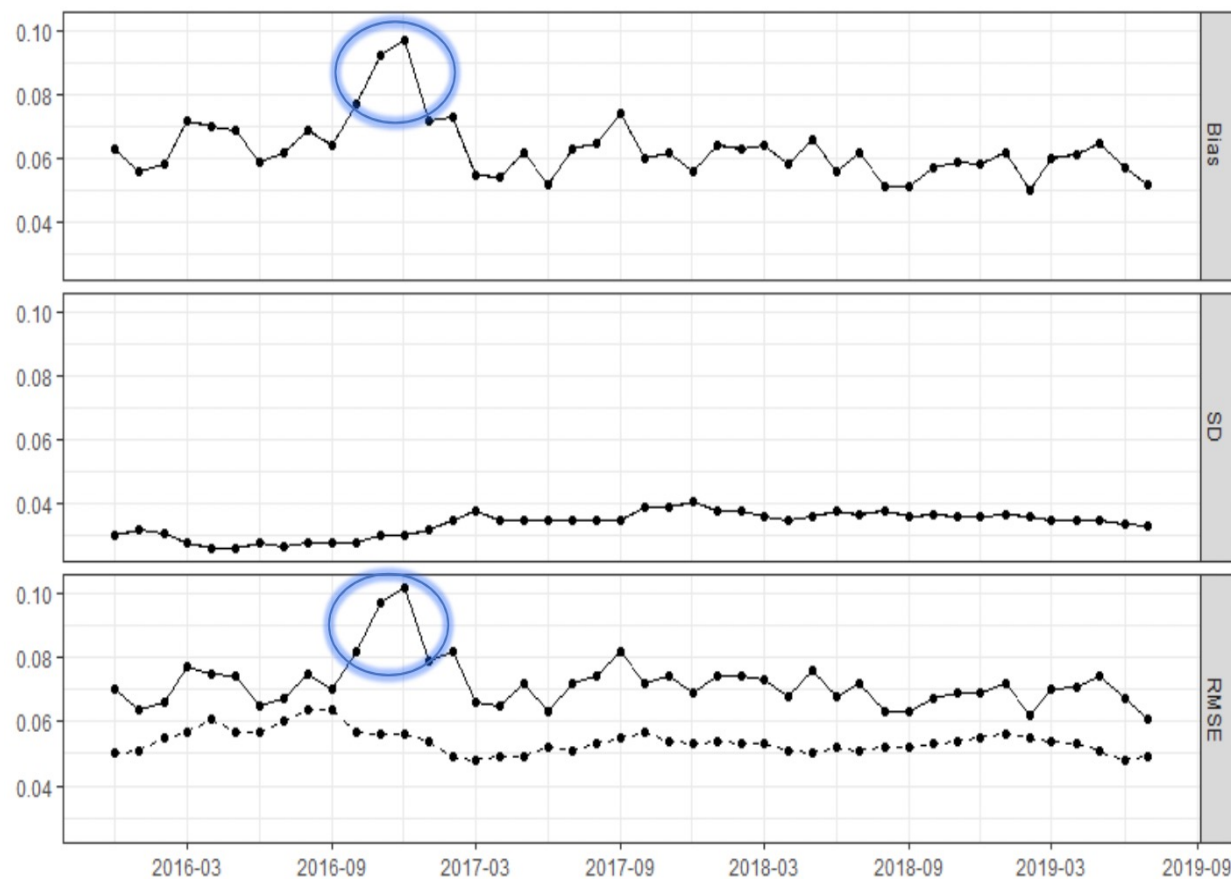


Figure 2.1. Quarter-averages of Web prediction accuracy in different historic time window. The accuracy is estimated by without-correlation model. The “dotted” RMSE defined by the posterior mean of the binomial variance is a lower bound.

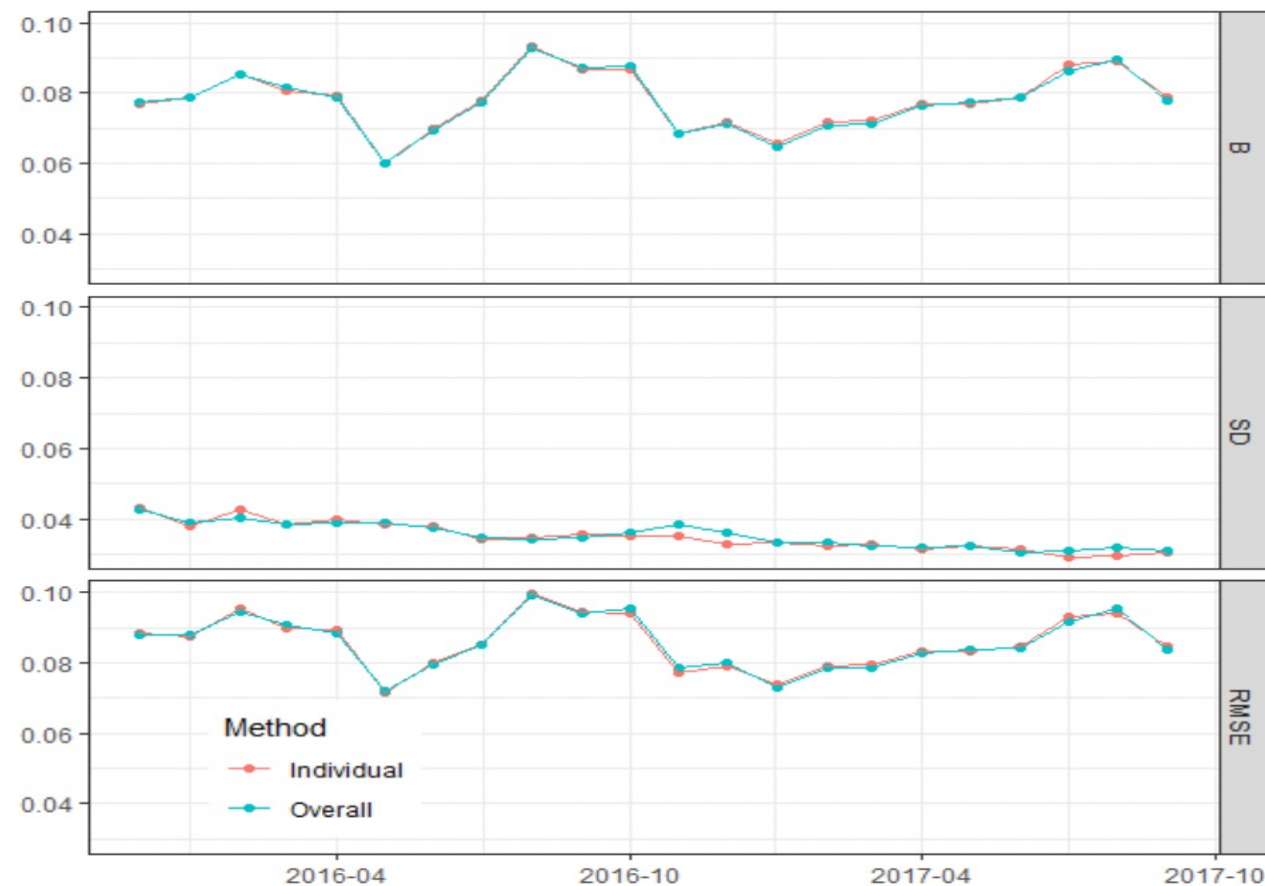


Figure 2.2. Quarter-averages of F2F prediction accuracy in different historic time window. “Individual” is a separate F2F RP model from Web. “Overall” is a RP model accounting for the correlation between Web and F2F.

Discussion

- A time series approach is sensitive to sudden external events and redesigns of a survey. How can the models be made more robust?
- We found relatively low correlations between web response propensities and face-to-face response propensities. How to improve the model that may better grasp any relations between the two modes?
- The stratification is not determined by variable selection. How does the model performance to such stratification?
- Next step is to construct an optimization problem in a Bayesian framework. The objective is to maximizing the performance of an adaptive survey against a non-adaptive survey.

Reference

Kreuter, F. (2013). Facing the Nonresponse Challenge. *Annals of the American Academy of Political and Social Science*.

Schouten, B., Peytchev, A., & Wagner, J. (2017). Adaptive survey design. In *Adaptive Survey Design*.

West, B. T., Wagner, J., Coeffey, S., & Elliott, M. R. (2020). The elicitation of prior distributions for Bayesian survey design: Historical data analysis vs. literature review.

Boonstra, H. J., & van den Brakel, J. A. (2019, Dec). Estimation of level and change for unemployment using structural time series models. *Survey Methodology*, 45(3), 395-425.