Do low-response rate online surveys provide equal or better data quality than high response rate face-to-face designs? Separating sample selection from measurement effects

Eliud Kibuchi[1], Patrick Sturgis[2], Gabriele Durrant[3], Olga Maslovskaya[3], Joel Williams[4]

[1]University of Glasgow, [2]London School of Economics, [3]University of Southampton, [4]Kantar Public

European Survey Research Association (ESRA) 8th Conference

16th July 2019

University of Zagreb, Croatia

Background and motivation

- Mixed-mode (MM) designs are used in many surveys (de Leeuw, 2005)
- MM designs may introduce differential measurement effects
- The key challenge is to disentangle selection and measurement effects which are completely confounded
- How to disentangle selection and measurement in observation studies?
  - Use of propensity score (PS) matching (Lugtig et al., 2011)

Research Objectives & empirical strategy

- RO1: Evaluate the performance of different formulations of propensity score models for mixed-mode data from complex sample surveys
- RO2: What can matching tell us about whether Address Based Online Sampling (ABOS) compares to 'gold standard' of face-to-face interviewing
- ES1: Mixed-mode observational studies based on face-to-face and online surveys
- ES2: Three different methods of estimating propensity scores
- ES3: Three different analytical methods of estimating measurement effects in matched sample

Data sets

Face-to-Face Survey

- Multi-stage random sample, 1 randomly selected adult aged 16+
- Response rate 60% (n=666)

Online (follow up) Survey

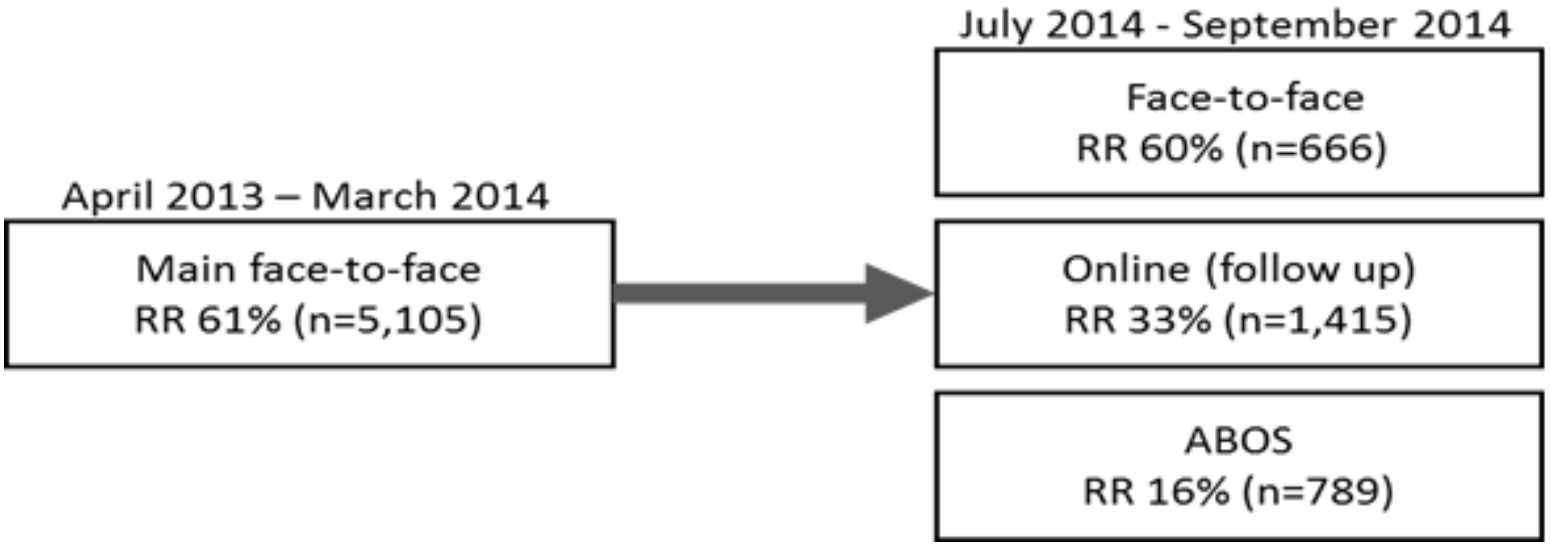
- Respondents who participated in 2013 face-to-face main CLS (2013)
- Respondents invited to participate using web and postal (i.e. those without internet access)
- Response rates 37% (n=1,576 online=1, 415 and postal=161)
- Only online respondents used in this analysis

Address based online surveying (ABOS)

- Up to to maximum of four individuals in a household for participation
- Algorithm used to clean fraudulent completions
- Had postal option for the population not covered by internet
- Response rate 17% (n=834, online=789 and postal=48)
- Only online respondents considered for this analysis

Data sets

Representation of the CLS study



Methodology

Propensity Score (PS) Matching

- Three different formulations of propensity scores models based on how survey weights are incorporated considered:
  - No sampling weights included in estimation (unweighted model)
  - Survey weights incorporated as a covariate (weight as covariate model)
  - Survey weights incorporated in estimation (weighted model)

- Also considered are three different analytical methods of estimating measurement effects from matched sample:
  - No sampling weights included on the outcome analysis
  - Matched control units retain their natural sampling weights
  - Matched control units inherit sampling weights of the treated units to which they were matched to

Methodology

- Logistic regression used to estimate propensity scores
- The outcome is the survey modes assigned to survey respondents
- Social-demographic and area variables for the respondents are used as baseline covariates for PS model
- The PS model as a balancing score is evaluated by checking area of common support using histograms

Methodology

- Matching sample obtained using one-to-one greedy nearest neighbour matching without replacement.

- Quality of the matched samples is assessed using covariate balance defined in terms of absolute standardised mean differences (SMD)

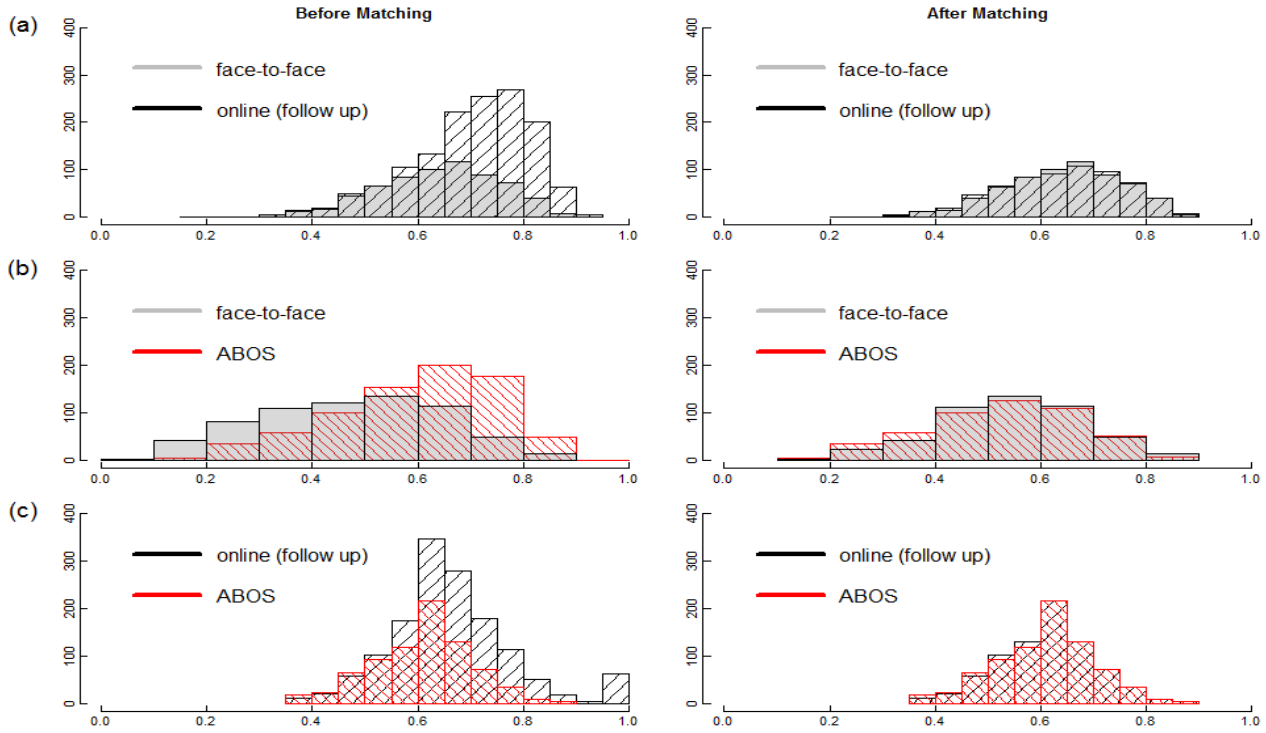
- SMD <0.10 indicate that the matched sample has adequate covariate balance

Methodology

- The selection and measurement effects are evaluated using Absolute Percentage Differences (APD) estimates
- APD estimated from 70 variables that consists of attitudinal and behavioural questions
- APD is the difference between the proportion obtained for a response in a certain question for a given mode minus the proportion in another mode as follows:
  - Given a categorical variable with K response levels, $K-1$ APD estimates are derived, where the omitted response level is the one with the lowest frequency.
  - Let $\hat{\pi}_{ijA}$ and $\hat{\pi}_{ijB}$ denote the estimated percentage proportions for question j and categorical level i for survey modes A and B respectively
  - Then APD estimate denoted as y_{ij} is defined $y_{ij}= \left|\hat{\pi}_{ijA} - \hat{\pi}_{ijB}\right|$
- The effectiveness of PS matching in disentangling selection and measurement effects is measured by comparing the median of the APD estimates before and after matching

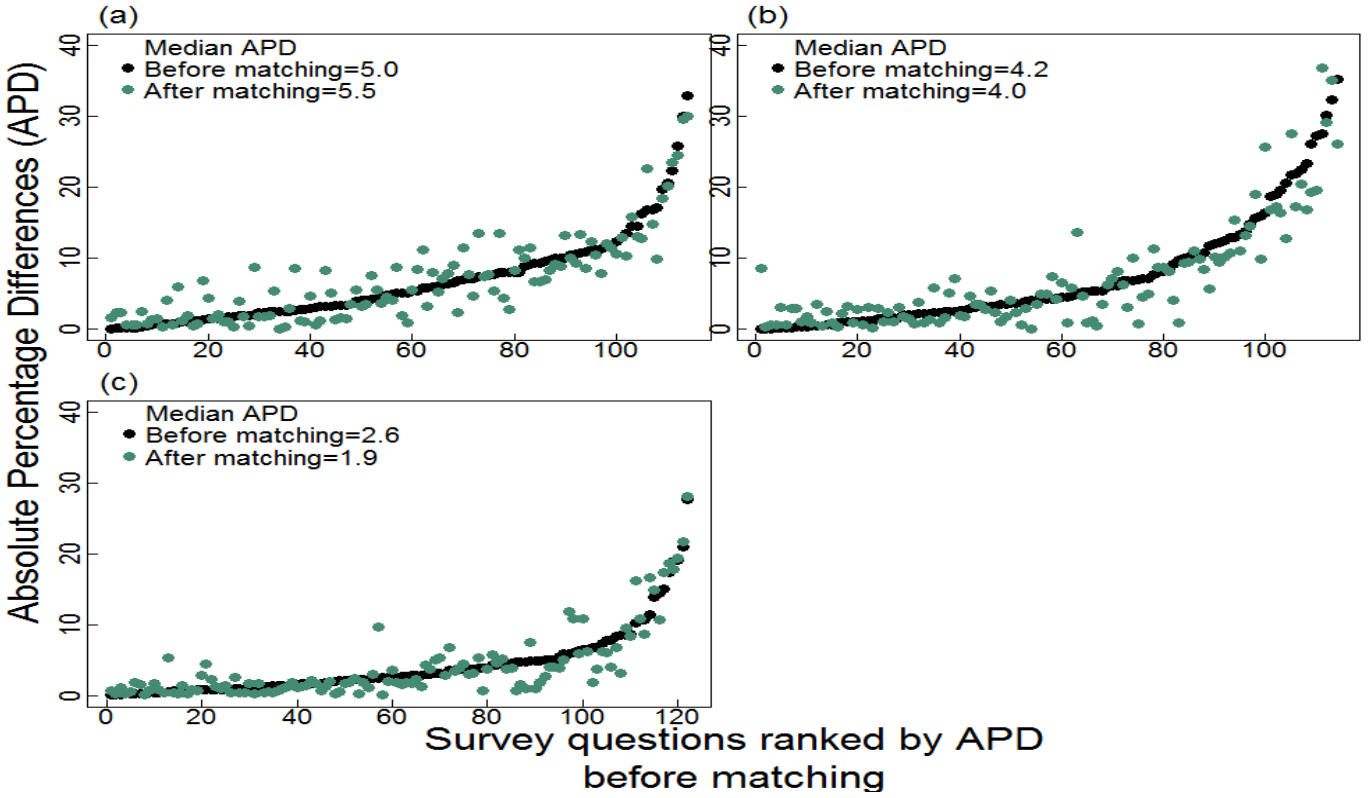
Results

Histograms of propensity scores distributions before and after matching for face-to-face and online (follow up) (a), face-to-face and ABOS (b) and ABOS and online (c)



Results

Estimated mode effects by Question before and after matching for face-to-face and online (follow up) (a), face-to-face and ABOS (b), and ABOS and online (follow up) (c)



Conclusion

- Majority of total mode effects between the online and face-to-face surveys is due to measurement rather than selection effects
- Neither of the two online surveys was similar to the face-to-face interview after matching. Therefore, online surveys do not provide equal or better data quality than higher response rate face-to-face interviews.
- Propensity score matching cannot be assumed to be a completely effective method for removing selection effects in surveys with different modes of data collection.
- Specification of different formulations of survey weights in propensity score models and outcome analysis were found to have no impact on the estimates of mode effects.

Thank You

Email: Eliud.Kibuchi@glasgow.ac.uk

Twitter: @e_Kibuchi