SURVEY FUTURES

SURVEY DATA COLLECTION METHODS COLLABORATION

Web-based RDS in High-quality Surveys

Methods and Lessons Learned

Curtis Jessop, Mark Todd, Toby Li (NatCen); Olga Maslovskaya, Luciano Perfetti Villa (University of Southampton); Carina Cornesse (GESIS)

University of Essex | University of Southampton | UKRI Economic and Social Research Council

CITY UNIVERSITY OF LONDON EST 1894 | Ipsos | LSE The London School of Economics and Political Science | MANCHESTER 1824 The University of Manchester | National Centre for Social Research | NCRM National Centre for Research Methods | Office for National Statistics | UCL | Ulster University | University of Leeds | UNIL | Université de Lausanne | verian | WARWICK The University of Warwick

Introduction

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Research Strand 1 of Survey Futures: 'Enhanced Sampling Frames and Procedures'

Sub-project 3: Respondent-driven Sampling Methods.

Aim: to assess effectiveness of RDS for recruiting hard-to-reach groups

- **Evidence review** summarising current knowledge and practice
- **Feasibility study** trialling recruitment of 18-24 year-olds using 'seeds' from a probability-based panel **[Phase 2]**
- **Survey Practice Guide** including recommendations for practitioners

Contents

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

- Context
- ~~Summary of findings from evidence review~~
- ~~Experiences from the feasibility study~~
- Recommendations for practice (and outstanding questions)
- Conclusions

What is RDS?

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

- A ‘chain-referral’ sampling method
  - Researchers recruit a small number of ‘seeds’ who participate, and are then asked to recruit people they know from the target population, who then recruit others (…)
- After several waves sample composition stabilises and becomes independent of seeds
- By tracking recruitment chains, measuring network sizes, and applying appropriate weights, it is possible to make inferences about the wider population, if statistical assumptions hold

Context

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

- RDS as a method for supplementing existing surveys
  - To boost samples sizes for groups that have a small incidence rate
  - To improve sample representativeness and inclusiveness

- RDS with probability-based seeds
  - Giving RDS-recruited sample representativeness a ‘head-start’

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Recommendations

University of Essex
University of Southampton
UKRI Economic and Social Research Council

Seed selection

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

- Diverse seed selection minimises initial bias and enhances population coverage
- Be strategic about your design to vary seed characteristics

- *Should seed samples be representative or purposive?*
- *How many seeds do you need?*
- *Can seeds be from outside the target population?*

Recruitment rules

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

- Target sample sizes depend on analytical requirements
- Chains achieve stable sample composition after 4-6 waves
- 3-5 ‘coupons’ per participant balances feasibility with risks of bias

- *Are ‘super-seeds’ good or bad?*
- *Maximum chain lengths?*

Measuring network sizes

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

- Important to capture information about people's networks (number of contacts in target population)

- *How to word network questions?*
- *How precise does network size data need to be?*

Fieldwork design

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

- Tracing of referral links is essential to the design
- Dual incentives (for both participants & recruiters) consistently boost engagement & should be tailored to circumstances
- Online mode affords scalability & speed, but risks fraud/protocol non-compliance

- *How (de-)centralised should recruitment be?*
- *What fieldwork protocols are effective?*
  - *Interactions with mode of interview? Design of recruitment questions? Reminders? Fieldwork length?*

Adaptive designs

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

- Real-time monitoring of fieldwork progress, network structure, homophily, & sample representativeness
- Pre-plan adaptions to protocols to mitigate implementation risks

- *What interventions can be effective in what circumstances?*

Quality checks

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

- Important to validate that recruitment protocols have been followed
  - Is the participant a member of the target population?
  - Are they a part of their recruiter's network?
  - Are they only participating once?
  - […]

- *What options are there for remote data collection?*
- *What do you do when protocols aren't followed?*

Standardised reporting

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

- Standardised reporting of quality indicators to ensure transparency and wider learning
  - E.g.: The Strengthening the Reporting of Observational Studies in Epidemiology for RDS Studies (STROBE-RDS) guidelines (White et al., 2015).

White, R.G., Hakim, A.J., Salganik, M.J., Spiller, M.W., Johnston, L.G., Kerr, L., Kendall, C., Drake, A., Wilson, D., Orroth, K., Egger, M. and Hladik, W. (2015). Strengthening the Reporting of Observational Studies in Epidemiology for respondent-driven sampling studies: 'STROBE-RDS' statement. *Journal of Clinical Epidemiology*, [online] 68(12), pp.1463–1471. doi:https://doi.org/10.1016/j.jclinepi.2015.04.002.

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Conclusions

University of Essex
University of Southampton
UKRI Economic and Social Research Council

More evidence needed!

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

- RDS is a promising methodology, but evidence on its application in high quality national surveys is thin

- Recent pilots/feasibility studies have demonstrated the feasibility of implementing RDS in practice, but have raised questions:
  - What are the right protocols for RDS to perform well?
  - How can we mitigate violation of underlying statistical assumptions?