Peter Barlow (Uni. Manchester)
peter.barow@manchester.ac.uk

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Prof Alexandru Cernat (Uni. Manchester)
Prof Natalie Shlomo (Uni. Manchester)
Prof Nikos Tzavidis (Uni. Southampton)
Prof Joseph Sakshaug (IAB)
Dr. Thomas O'Toole (Uni. Manchester)

Using integrated non-survey data for monitoring and intervening in survey data collection

RS7: Data Integration

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

NCRM National Centre for Research Methods
Office for National Statistics
UCL
Warwick The University of Warwick
Manchester 1824 The University of Manchester
City University of London Est 1894
National Centre for Social Research
LSE The London School of Economics and Political Science
Ulster University
Unil UNIL | Université de Lausanne
Ipsos
Kantar Public

Different types of survey design

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Survey designs produce data

Standardised survey design

Everyone in the sample treated the same

Adaptive survey designs

Sample treated differently depending on characteristics

Why adaptive survey design?

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Precipitous decline in response rates in recent years

Costs and budgetary pressures have also increased

Adaptive designs optimise data quality against cost

Target sub-groups in the sample

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

Two types of adaptive design

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Targeted survey design	Static approach that involves a planning phase where decisions regarding targeting and interventions occur (Lynn 2017)
Responsive survey design	Approach changes between phases of data collection using the data collected on non-response (Brick and Tourangeau 2017)

Goals of adaptive design

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Aims	Maximise data quality	Keep costs low
How	Targeting and intervention	Budgetary restraints
Requirements	Auxiliary data and process methodology	Some form of optimisation process

Key definitions and jargon

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Targeting, strata and the problem	Intervention, treatment and dosage	Bias, costs and consequences
• Targeting is the identification of subgroups in the sample • Subgroups are called strata • The problem is the difference in response rates across strata	• Intervention is the differential in data collection applied to a stratum • This is also called a treatment • Dosage is the extent to which this treatment is applied to the stratum	• There is a time/human resources cost from planning and preparing the design • Financial cost to the treatment and implementation • Risk of bias due to intervention • Complexities in using the data

Auxiliary essential for all of these

Measures of representativeness

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

2 types of indicator of effectiveness

Survey based indicator	Variable or category based indicator
• R indicator + others • Coefficient of variation	• Partial R-indicator
Evaluates representativeness of the survey as a whole	Evaluates representativeness with respect to 1 variable

Require good auxiliary data

Non-survey data and its value

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Type of data	Examples	Use for targeting
Sample frame data	Age/gender linked to sample frame	Can target based on this combined with other data
Geo-spatial data	Data on region and LSOA. IMD data and subdomains Demographic breakdown	Make bespoke language for nations/regions Profiles of LSOA – targeting the LSOA
Administrative data	Demographic and health data and interactions with the NHS	Detailed data on health and behaviour Excellent source internationally Not so great in the UK
Commercial data	CACI data, Experion data on commercial behaviour	Commercial data provide data on behaviour giving insight to non-response

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

Criteria for selecting variables from non-survey data

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Two key dimensions on which variables are judged	
Variables should be correlated with key outcomes and participation	Variables should not be collinear with each other
Maximises how much of the variation in the key outcomes being explained	Maximises the breadth of the variance in the key outcome being explained
Intuition and consideration of correlation coefficients are key here	Use previous experience, research and data to understand the best model

2 ways of targeting

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Subgroup targeting	Data based targeting
Use past experience/intuition	Classify strata based on model of response propensity
Data used purely to identify strata within the sample	Data used for stratification and selection for intervention
Most often used in UK Many examples – Generations and Gender Survey	Used in TLFS, DESNZ public attitudes tracker

Spectrum of interventions

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Fieldwork role out	Invitation language	Follow up reminders	Knock to nudge	Differential incentives	Different modes
Role out fieldwork to lowest propensity parts of sample first	Change the language of invitation based on propensity or region	Issue more follow up reminders to higher non-response propensity people	In-person contact to help complete a CAWI survey	Allocate a greater or lesser incentive to different strata	Allocate a CAWI mode to some strata and a CAPI/CATI mode to others
Millenium Cohort Study	Understanding Society	Community Life Survey	TLFS	Children of the 2020s study	Dutch Labour Force Survey

Sequencing, dosage and costs

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Questions

When should the intervention be applied?

What is the appropriate dosage of the intervention?

How should we minimise the cost of the intervention?

Optimisation approaches to adaptive design

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Trial and error approach	Use practitioner intuition with only a rough guide on trying to keep costs within a budget – Most UK examples – only relies on non-survey data to administer intervention (Schouten and Shlomo 2017)
Stopping rules	Select a point at which the targeting will stop (Wagner 2010, Rao et al. 2008) – e.g. a particular response rate from a stratum or a particular coefficient of variation – Stopping rule depends on a particular variable derivable from non-survey data
Propensity based prioritisation	Develop a model of non-response and apply intervention to lower propensity groups (Peytchev et al 2018, Wagner 2013)
Mathematical optimisation	Use a mathematic procedure to ensure that the application of incentives is optimised to a budget constraint so as to maximise a particular quality (e.g. coefficient of variation) (van Berkel 2022)

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

Recommendations

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Policymaker	• Access to better register data in the UK would be useful • Facilitating access to linked data • Training in strengths and weaknesses of adaptive design
Survey agencies	• Significant opportunity for adaptive design with what is currently extant in the UK • Opportunities to use more advanced approaches like mathematical optimisation and Bayesian Adaptive Design • More training opportunities on how to implement adaptive designs

References

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

- Brick, J.M. and Tourangeau, R. (2017), “Responsive Survey Designs for Reducing Nonresponse Bias”, *Journal of Official Statistics*, Vol. 33 No. 3, pp. 735–752, doi: 10.1515/jos-2017-0034.
- Lynn, P. (2017), “From standardised to targeted survey procedures for tackling non-response and attrition”, *Survey Research Methods*, European Survey Research Association, Vol. Vol 11, pp. 93-103 Pages, doi: 10.18148/SRM/2017.V11I1.6734.
- Peytchev, A., Riley, S., Rosen, J., Murphy, J. and Lindblad, M. (2010), “Reduction of Nonresponse Bias through Case Prioritization”, *Survey Research Methods*, European Survey Research Association, Vol. Vol 4, pp. 21-29 Pages, doi: 10.18148/SRM/2010.V4I1.3037.
- Rao, R.S., Glickman, M.E. and Glynn, R.J. (2008), “Stopping rules for surveys with multiple waves of nonrespondent follow-up”, *Statistics in Medicine*, Vol. 27 No. 12, pp. 2196–2213, doi: 10.1002/sim.3063.
- Schouten, B., Peytchev, A. and Wagner, J. (2017), *Adaptive Survey Design*, Chapman and Hall/CRC.
- Schouten, B. and Shlomo, N. (2017), “Selecting Adaptive Survey Design Strata with Partial R-indicators”, *International Statistical Review*, Vol. 85 No. 1, pp. 143–163, doi: 10.1111/insr.12159.
- Van Berkel, K. (2022), “Adaptive Survey Design for the Dutch Labour Force Survey”, *American Journal of Theoretical and Applied Statistics*, Vol. 11 No. 4.
- Wagner, J. (2012), “Adaptive Contact Strategies in Telephone and Face-to-Face Surveys”, *Survey Research Methods*, European Survey Research Association, Vol. Vol 7, pp. 45-55 Pages, doi: 10.18148/SRM/2013.V7I1.5037.

Peter Barlow (Uni. Manchester)
peter.barow@manchester.ac.uk

SURVEY FUTURES
SURVEY DATA COLLECTION METHODS COLLABORATION

Prof Alexandru Cernat (Uni. Manchester)
Prof Natalie Shlomo (Uni. Manchester)
Prof Nikos Tzavidis (Uni. Southampton)
Prof Joseph Sakshaug (IAB)
Dr. Thomas O'Toole (Uni. Manchester)

RS7: Data Integration

University of Essex
University of Southampton
UKRI Economic and Social Research Council
NCRM National Centre for Research Methods
Office for National Statistics
UCL
Warwick The University of Warwick
Manchester 1824 The University of Manchester
City University of London Est 1894
National Centre for Social Research
LSE The London School of Economics and Political Science
Ulster University
Unil UNIL | Université de Lausanne
Ipsos
Kantar Public