Respondent Centric Survey Design and Data Collection – Transformed Labour Force Survey

Maria Tortoriello
Principal Social Researcher

Colin Beavan-Seymour
Principal Social Researcher

UNECE Expert Meeting 2023
Talk outline

Part 1 – Survey Design
- What is the purpose of the Transformed Labour Force Survey?
- Survey Design – sample, collection modes
- Return rates

Part 2 – Implementation of an Adaptive Survey Design
- Why use an Adaptive Survey Design?
- How was it developed?
- How was it implemented?
- Initial findings
Part 1 – Survey Design

Colin Beavan-Seymour
What is the Transformed Labour Force Survey?

• A new survey which will collect data on key labour market measures
• Developed with a respondent centric approach
• Qualitative and quantitative research
• Online first

• A rationalisation and redevelopment / rethink of how to measure core labour market concepts
• Extensive qualitative research with members of the public, interviewers, data users
The journey so far...

2017
Tests 1 & 2
Online response rates
Engagement strategies

2018
Test 3
Mixed mode (online & F2F)
Statistical outcomes

2019
Test 4
Online attrition test – response rates across 3 waves

2020
TLFS Beta
Online only in response to pandemic

2022
Addition of Telephone
Online & telephone collection

2022/23
Knock-to-nudge
Using an Adaptive Survey Design
Sample Design

Transformed Labour Force Survey Wave 1

140,000 households

- TLFS Wave 2
- TLFS Wave 3
- TLFS Wave 4
- TLFS Wave 5

40,000 households

- Opinions Survey
- Other Social Surveys
- Other Social Surveys
What data did this give us?

- A return rate (complete returns & partials) of around **37.5%** - a great start!
- However, we were still seeing similar biases in the responding sample that other voluntary surveys in the UK were experiencing, despite the online mode and user-centric design:
  - A large proportion of respondents were **over 55**, many over 65 – fewer respondents of working age, more economically inactive
  - A majority of respondents **owned their homes**, many without a mortgage or loan
  - Respondents with a **white ethnic background** comprised the vast majority of the data, under-representation from other ethnic backgrounds
  - The vast majority of data was from the **online mode** – only a small percentage was from telephone collections
  - The 2018 test indicated that interviewers visiting households can increase response from under-represented areas
  - But… with a large scale survey of over 500,000 a year… *how can we increase the quality of the data collection but keep the cost of the operation down?*
Part 2 – Adaptive Survey Design

Maria Tortoriello
What is an Adaptive Survey Design (ASD)?

In November 2022 we implemented an ASD for the TLFS.

- What is an ASD?
  - Dividing a sample into smaller groups that have similar characteristics (segmentation)
  - Applying alternative survey design features for different groups:
    - modes, materials, incentives
  - Objective is to improve targeted survey outcomes
    - reduce bias, reduce costs
Why use an Adaptive Survey Design?

• TLFS data collection strategy same for all sampled addresses = no adaptive survey design
• Experiencing differential non-response bias which affects estimates
• Statistical processing enables weighting of sample to account for some bias, but confidence in estimates would only improve with higher quality input data.

• Next step for TLFS was to introduce additional modes - Face to Face follow up
• One size does not fit all!
• ASD allows you to target the right respondents in the right way, rather than targeting all respondents in the same way = more efficient use of field resources
How was the Adaptive Survey Design developed?

- Closely followed work of Statistics Netherlands (Schouten, B et al.)
- A key objective of ASD is to divide the sample into strata in order to define targeted protocols for each of the strata
- A logistic regression model was applied to historical TLFS data to identify auxiliary variables strongly associated with response to formulate the ASD strata.
- Variables considered were **Index of Multiple Deprivation (IMD)**, **Urban/Rural Classification**, **Country of Birth, Age & Ethnicity** (limited by available data).
- Derived and examined **CV, R-Indicators and Partial R-Indicators** to identify the variables and categories of variables driving variation in response propensities
- Strongest predictors of response:
  - Age (<45)
  - Urban/Rural Classification (Urban)
  - Index of Multiple Deprivation (IMD deciles 1-4)

Constructed 8 strata based on these variables
ASD: Iteration 1

- Potential to include numerous interventions in the ASD (e.g., mode, incentive, materials..)
- Keeping it simple with 1 intervention = ‘Knock to Nudge’ (KtN) follow up
- ASD will target KtN data collection at under-represented strata based on response propensities in order to reduce the variation in response propensities for a selected set of auxiliary variables.

- STRATA 1 = Urban, less deprived areas, 45+
- STRATA 2 = urban. more deprived areas, 16-44
- STRATA 3 = urban, less deprived areas, 16-44
- STRATA 4 = urban, more deprived areas, 45+
- STRATA 5 = non-urban, more deprived areas, 16-44
- STRATA 6 = non-urban, more deprived areas, 45+
- STRATA 7 = non-urban, less deprived areas, 16-44
- STRATA 8 = non-urban, less deprived areas, 45+

- This will ensure that data collection resources are used in the most efficient way whilst increasing response from historically underrepresented population groups.
ASD Optimisation approach

• We are following a structured ‘trial and error’ approach to optimising our ASD.
• The optimum solution is unknown and experimental testing is needed
• Start with a simple design that can be accommodated using existing systems
• Document, evaluate, learn, extend…
• Grow – add features to the ASD as technical and admin systems improve over time
Early results

• ASD Evaluation project - ongoing
  • **Operational evaluation** – evaluating optimal set up of KtN
    o Optimal number of visits = 2/3
    o Best days to make contact: Monday, Tuesday, Sunday
    o Best time of day to make contact between 3pm-8pm
    o KtN not working as well in London and North West regions
  • **Data quality evaluation**
    • Improving variability in response across strata
    • Small improvements in representativity of data
      o Statistically significant increase in response from 'hard to reach' groups

First ‘full’ knock-to-nudge month
Thank you for listening!

Any questions?

Contact details:
maria.tortoriello@ons.gov.uk
colin.beavan-seymour@ons.gov.uk