Experimental methods to increase online response rates in Australian Social Surveys

Stephen Cohen stephen.cohen@abs.gov.au

Australian Bureau of Statistics Informing Australia's important decisions



Context

The Australian Labour Force & History

Our Household Survey Innovation Panel

- Panel design
- Recent results
- Surprising findings
- Some reflections

Australian Labour Force Survey (LFS)

- 8 months, 8 rotation groups
- Address based sample
- 2 week reference period
- Proxy reporting for all individuals in household
- Mixed mode
 - First month = >50% CAPI, 20% eForm (at 2017)
 - Months 2-7 mostly CATI or eForm
- 92% response rate
 - Slowly falling
 - No incentives

		Australia	Canada	NZ	UK	US
	Frequency	Monthly	Monthly	Quarterly	Quarterly	Monthly
	Responding sample size	26,000 households	56,000 households	15,000 households	40,000 households	74,000 households
		50,000 people	100,000 people	30,000 people	100,000 people	105,000 people
	Response rates	~ 92%	~ 90%	~ 86%	~ 49%	~ 86%
	Rotation groups	8	6	8	5	8
	Population compared to Aus		1.5 times the size	19% the size	2.7 times the size	13.3 times the size

Australian Labour Force Survey (LFS)

High costs

- LFS accounts for very large share of ABS collection costs.
 - ABS facing increasing budget pressures
- Eform mode introduced in 2012
 - Seen as the primary lever to reduce costs
 - Only 16% eform takeup (2017)

Low risk appetite

- Rich, Highly scrutinized & politically sensitive series
- Efficient sample design, sensitive to shocks
- Brittle systems & limited staff to maintain them

Change is necessary ... yet constrained

MPS: Current state



LFS eform response rates





How do we maximise self-initiated response?

Responsive design

IPND to call respondents

Monthly attrition

HCDF timing

Letters

HCDF extension

Optimised approach timing

Clearer instructions

Priority post

Personalise letters with names

Envelope design

Improve the form

Interviewers as case managers

Email & SMS reminder content & timing

Rolling reference periods

How do we maximise self-initiated response?

Single use signon

Revised survey sales pitch

XIAM secret question removal

Market segme	ntation
--------------	---------

Magic

Approach strategy

Gradual engagement

Better survey name

The proposal: An quasi-embedded experimental program

- Monthly Randomised Controlled Trials
 - Small number (n=700) additional households selected, monthly, commencing October 2017
 - Field procedures and materials mimic LFS except for the *experimental treatment* that is hypothesised to increase the likelihood of self-initiated response.
 - Response rates compared to those of mainstream LFS, enabling ABS to identify the causal factors driving e-collection take-up.
- Minimised risk to Labour Force by segregating the experimental sample and systems from mainstream MPS
- Qualitative followup to contact nonrespondents & understand why they did not respond (i.e. barriers identification)
- Implementation risk minimised & managed through demonstrated, experimental/scientific approach





Cumulative HCDF completion rate (cumulative %)





New

Old





Control Treatment

Results:

- 6.2pp improvement from new materials
- Treatment (mostly) implemented in live LFS survey
 - Comparable results being achieved

Cumulative HCDF completion rate (cumulative %)



Cumulative HCDF completion rate (cumulative %)



First reminder...

...Second "overdue" reminder





Results:

• 13pp improvement from second reminder

Cumulative HCDF completion rate (cumulative %)



Cumulative HCDF completion rate (cumulative %)





Some highlights

What works...

- Visual improvements to letter (+6pp)
- Improved messaging and sales pitch (+5pp)
- Addition of 'bureaucratic' second reminder (+13pp) examination of tone soon
- Soft close (+2pp)
- Personalisation/Postit notes
 - Handwritten +8pp (but...)
 - Pre-printed +4pp
 - Image of postit n.s.

What hasn't...

- "Radically Simplified" approach letters (-10pp)
- Envelope teasers (n.s.)
- Compulsion messaging (n.s.)
- Extended deadlines with hard close (n.s)
- Plain "official" envelopes (-4.5pp vs coloured envelopes)



Please contact us It'd really help us out! teve

MERCE

p-26

riease really he.

Steve

re contend us-

Stell

really help us out!

it well,

Current work

- Improved sales pitch for the survey & Social norms
- Tone of letter content & escalation over multiple contacts
- Additional email and SMS reminders
- Targeted strategies: Secure apartment buildings, remote populations
- Reducing the barriers reported by nonrespondents
- Improvements to website landing page & response process
- "Retaining" respondents over 8 months of LFS



"Why don't you use a factorial design?

...it's much more efficient...

...and would tell you about the interactions"

- 10



0



Which design?

Factorial design

- Efficient use of sample
- Estimate interaction terms
- Quickly become large and complex
- High effort & management costs
 = single, one off experiment
- Vulnerable to implementation errors

RCT program is...

- highly responsive,
- maintains ecological and internal validity,
- is relatively robust to operationalisation errors,
- conducted at very low cost
- dramatically increasing adoption



...highly responsive, maintains ecological and internal validity, is relatively robust to operationalisation errors, and yet is conducted at very low cost while dramatically increasing adoption

The monthly, low-effort design:

- Increases willingness to test more risky/unconventional approaches
- Is capable of informing immediate issues
 - Easily integrate new ideas from outside the organisation
 - Adapts to emerging and evolving interests, theory and problems
- Frequent results maintain interest in the experimental program and respondent behaviour generally

Internal & External Validity

...highly responsive, maintains ecological and internal validity, is relatively robust to operationalisation errors, and yet is conducted at very low cost while dramatically increasing adoption

- Internal validity of a well-designed RCT
- Externally valid and *Demonstrative*: uses existing systems and processes.
 - Evaluates the theory *and* it's implementation, together.
- Always up-to-date control condition:
 - ...minimises the theoretical leaps needed in implementation
- Qualitative followup

Robust to errors

...highly responsive, maintains ecological and internal validity, is relatively robust to operationalisation errors, and yet is conducted at very low cost while dramatically increasing adoption

When live systems go wrong:

- We run the trial again next month
- And opportunistically benefit from "natural experiments" when errors do occur
- In the worst case, we have "spent" n=700
- Errors in a factorial design would be more concerning.
 - Complex designs can *cause* errors.

...highly responsive, maintains ecological and internal validity, is relatively robust to operationalisation errors, and yet is conducted at very low cost while dramatically increasing adoption

Low cost

- Quasi-embedded:
 - 700 topup sample per month
 - powered to detect ~3.5-4pp differences
 - Control group "for free"
 - total n=4200 each month, for the cost of 700*3 letters. (< \$5k / month)

Adoption

...highly responsive, maintains ecological and internal validity, is relatively robust to operationalisation errors, and yet is conducted at very low cost while dramatically increasing adoption

- From a baseline of 16%...
- High ROI
 - 16pp+ improvement to date (without incentives)
 - A further 5-10pp further identified (would achieve ~40% sample without incentives)
 - Despite our account system has not been resolved do we have additional 'latent' electronic responders?
- Now scaling across ABS survey program
 - Adoption throughout our social survey program
 - Addressing unique needs
 - Looking to expand into a similar business survey program

Future directions

- 1. Future experiments planned
 - Addressing key barriers reported by nonrespondents
 - Where is Australia's web takeup 'ceiling'?
- 2. Combining/consolidate across 20+ RCTs?
 - "roll up" into a fractional factorial design; other network approaches?
- 3. Can this vehicle be used for content experiments, not just approach/eform takeup?
 - Consent questions
- 4. Elaborate the model across the ABS survey program
 - Noting some key differences for business surveys

Survey Methods in Multinational, Multiregional, and Multicultural Contexts

Questions? stephen.cohen@abs.gov.au

METHODS FOR TESTING AND Uating Survey Questionnaires

Context The Australian Labour Force Survey

Our Household Survey Innovation Panel

- Panel design

- Recent results

- Surprising findings

Some reflections

TOURANGEAU / RIPS / RASINSKI

IMPROVING INTERVIEW METHOL AND QUESTIONNAIRE DESIGN

COUPER

DESIGNI

Juestions &

Designing and Cons for Social Research

Protocol Ana

Cognition and Survey Research

Handbook of Usal