



Methodology Seminar Series

Probability based on-line panels in Great Britain

Curtis Jessop, NatCen Social Research

Rory Fitzgerald & Gianmaria Bottoni, European Social Survey

Thursday 21 June 2018, 6pm

Research Context

- **‘Gold standard’ is not always appropriate**
 - Research budgets may not cover face-to-face fieldwork
 - May need to respond quickly to unfolding events
- **Question marks over the quality of alternatives**
 - Coverage & self-selection biases on web panels;
 - ‘Early responders’ due to quotas or short fieldwork periods

The NatCen Panel

- The first probability-based panel in Great Britain
- Open to be used by anyone (with funding!) for social research
- Aims to provide a quality alternative to ‘traditional’ probability-based approaches, producing reliable population estimates in a shorter-time frame and at lower cost

The CRONOS Panel – proof of concept

- CRONOS: first attempt to set up a cross-national probability-based web panel under an input-harmonisation framework.
- Ultimate aim to provide open access data resource and additional questionnaire space
- Panel recruitment, setup and maintenance were guided by the same methodological principles in all participating countries.
- CRONOS aims to evaluate the effectiveness of panel recruitment on the back of an existing cross-national survey in terms of costs, sample representativeness, participation and attrition rates, and data quality.
- Proof of concept for the possible future development of a larger European level, methodologically harmonised, probability-based, web panel.
- Pilot funded by European Commission.

Panel designs

CRONOS (GB)

NatCen Panel

| | CRONOS (GB) | NatCen Panel |
|-----------------------------|--|--|
| <i>Target Population</i> | General Population, 18+ in GB | General Population, 18+ in GB |
| <i>Sample design</i> | Probability-based | Probability-based |
| <i>Recruitment strategy</i> | Piggy-backing approach (ESS) | Piggy-backing approach (BSA) |
| <i>Recruited panel size</i> | 1218 recruited | 6,785 recruited, 5948 'active' |
| <i>Waves</i> | 6 main surveys plus a welcome survey | Continuous (every 1-2 months) |
| <i>Interview length</i> | 20 min | 15 min |
| <i>Data collection</i> | Single mode (Online) | Sequential mixed-mode (Online & telephone) |
| <i>Survey incentives</i> | Unconditional £5 per wave <u>OR</u> one-off unconditional £30 | Conditional £5 gift card |
| <i>Offline coverage</i> | 142 tablets provided to those without web access (5.2% of invited respondents) | Follow-up in telephone mode 10% of panellists report no internet access |
| <i>Mobile completion</i> | 33% (average) | c.25% (average) |

Response rates

CRONOS (GB) wave 1

NatCen Panel – May 2017

| | Number of cases | Cumulative RR (%) | Number of cases | Cumulative RR (%) |
|---|-----------------|-------------------|-----------------|-------------------|
| <i>Issued sample (eligible)</i> | 4,447 | - | 14,897 | - |
| <i>Participated in recruitment interview (ESS or BSA)</i> | 1,844 | 41.5 | 7,270 | 49% |
| <i>Agreed to join the panel</i> | 1,225 | 27.5 | 4,205 | 28% |
| <i>Issued at this wave (minus attrition)</i> | 1,218 | 27.3 | 3,729 | 25% |
| <i>Participated at this wave</i> | 685 | 15.4 | 2,223 | 15% |

Method for assessing representativeness

- We are going to assess the representativeness of CRONOS wave 1 achieved sample compared both to ESS Round 8 (2016) and to population data.
- The data about the CRONOS achieved sample comes directly from answers to ESS face-to-face questionnaire (ESS worked as recruitment interview for CRONOS).
- **This allows us to compare characteristics of the population that participated in the web panel with those who did not without any mode effects.**
- In order to assess discrepancies between CRONOS, ESS and population data, we used:
 - Duncan dissimilarity index.
 - Logistic regression model for demographics characteristics
 - Logistic regression for attitudinal and behavioural variables

Benchmark data

Variables to assess the representativeness of CRONOS wave 1 achieved sample compared to ESS Round 8 (2016) and to population data.

| Variable | Categories | Comparison | Target population |
|---------------------|--|--|-------------------|
| Gender | <ul style="list-style-type: none"> Female Male | ESS, National statistics offices | 18+ |
| Age | <ul style="list-style-type: none"> 18-24 25-34 35-54 55-64 65-74 75+ | ESS, National statistics offices | 18+ |
| Education | <ul style="list-style-type: none"> Primary Secondary Tertiary | ESS, EU-LFS | 18-74 |
| Marital Status | <ul style="list-style-type: none"> Married Not Married | ESS, Eurostat, Office for National Statistics (just for GB*) | 20+ |
| Employment relation | <ul style="list-style-type: none"> Self-employed Not self-employed | ESS, EU-LFS | 25-54 |
| Work status | <ul style="list-style-type: none"> In paid work Not in paid work | ESS, EU-LFS | 25-54 |
| Citizenship | <ul style="list-style-type: none"> Citizen Not citizen | ESS, Eurostat | 20+ |
| Household size | <ul style="list-style-type: none"> Single person Household Household with at least 2 individuals | ESS, EU-LFS | 25-54 |

Assessing discrepancies

Duncan dissimilarity index between CRONOS, ESS and Population

| | Estonia | | GB | | Slovenia | | | |
|------------------------|----------------|----------------|----------------|----------------|----------------|----------------|---------------------------|---------------------------|
| | CRONOS/ ESS | CRONOS/ POP | CRONOS/ ESS | CRONOS/ POP | CRONOS/ ESS | CRONOS/ POP | Average CRONOS/ ESS | Average CRONOS/ POP |
| Gender | 3.7 | 3.9 | 1.0 | 4.1 | 1.8 | 5.2 | 2.2 | 4.4 |
| Age | 11.1 | 8.3 | 6.5 | 10.1 | 12.3 | 9.4 | 10.0 | 9.3 |
| Education | 9.1 | 6.2 | 5.7 | 2.3 | 7.5 | 7.6 | 7.4 | 5.4 |
| Employment relation | 1.0 | 3.5 | 2.1 | 2.3 | 0.9 | 1.3 | 1.3 | 2.4 |
| Work status | 2.7 | 2.2 | 1.2 | 1.8 | 0.1 | 2.9 | 1.3 | 2.3 |
| Citizenship | 4.0 | 8.9 | 1.3 | 0.8 | 0.6 | 4.0 | 2.0 | 4.6 |
| Household size | 3.2 | 5.6 | 1.0 | 1.9 | 0.7 | 6.6 | 1.6 | 4.7 |
| Marital status | 0.4 | 3.0 | 5.4 | 7.5 | 1.6 | 3.4 | 2.5 | 4.6 |
| Average | 4.4 | 5.2 | 3.0 | 3.9 | 3.2 | 5.1 | 3.5 | 4.7 |

Education and Age

| Education | | <i>Estonia</i> | <i>GB</i> | <i>Slovenia</i> |
|------------------|--------|----------------|-----------|-----------------|
| Primary | CRONOS | 8.5 | 20.3 | 8.8 |
| | ESS | 13.6 | 25.6 | 16.4 |
| | EU-LFS | 14.8 | 21.1 | 15.2 |
| Secondary | CRONOS | 52.9 | 42.5 | 65.4 |
| | ESS | 56.9 | 42.9 | 62.9 |
| | EU-LFS | 50.2 | 40.2 | 57.8 |
| Tertiary | CRONOS | 38.6 | 37.2 | 25.7 |
| | ESS | 29.5 | 31.5 | 20.7 |
| | EU-LFS | 35.1 | 38.8 | 26.9 |

| Age distribution | | <i>Estonia</i> | <i>GB</i> | <i>Slovenia</i> |
|-------------------------|--------|----------------|-----------|-----------------|
| 18-24 | CRONOS | 10.1 | 10.2 | 8.1 |
| | ESS | 7.8 | 9.3 | 8.6 |
| | POP | 8.8 | 11.2 | 8.4 |
| 25-34 | CRONOS | 20.4 | 13.6 | 19.7 |
| | ESS | 16.5 | 15.6 | 14.5 |
| | POP | 18 | 17.2 | 16.1 |
| 35-54 | CRONOS | 37.3 | 36.5 | 41 |
| | ESS | 32.3 | 34.4 | 33.8 |
| | POP | 33.1 | 34 | 35.9 |

| Age distribution | | <i>Estonia</i> | <i>GB</i> | <i>Slovenia</i> |
|-------------------------|--------|----------------|-----------|-----------------|
| 55-64 | CRONOS | 16.6 | 17 | 18 |
| | ESS | 17.6 | 16.1 | 19.7 |
| | POP | 16.2 | 14.7 | 17.2 |
| 65-74 | CRONOS | 11.1 | 18 | 9.8 |
| | ESS | 13.9 | 15.5 | 12.7 |
| | POP | 12.1 | 12.6 | 11.8 |
| 75+ | CRONOS | 4.5 | 4.7 | 3.4 |
| | ESS | 11.8 | 9.3 | 10.6 |
| | POP | 11.8 | 10.3 | 10.5 |

Assessing CRONOS respondents' characteristics

In order to assess whether specific demographic variables predict the propensity to join the CRONOS panel, we performed a logistic regression

| | B | S.E. | Sig. | Exp(B) | 95% C.I. for EXP(B) | |
|------------------------------|--------|-------|------|--------|---------------------|-------|
| Female | 0.134 | 0.065 | .039 | 1.143 | 1.007 | 1.298 |
| 65+ | | | | | | |
| 18-34 | 0.022 | 0.116 | .850 | 1.022 | 0.814 | 1.283 |
| 35-64 | 0.117 | 0.103 | .258 | 1.124 | 0.918 | 1.376 |
| Primary | | | | | | |
| Secondary | 0.326 | 0.094 | .001 | 1.385 | 1.152 | 1.665 |
| Tertiary | 0.489 | 0.106 | .000 | 1.631 | 1.326 | 2.006 |
| Voted | 0.400 | 0.080 | .000 | 1.492 | 1.275 | 1.746 |
| Paid Work | -0.250 | 0.080 | .002 | 0.779 | 0.665 | 0.912 |
| Living comfortably or coping | 0.199 | 0.095 | .037 | 1.220 | 1.012 | 1.470 |
| Internet Use – Never | | | | | | |
| Occasionally/Most days | 1.167 | 0.128 | .000 | 3.212 | 2.498 | 4.130 |
| Every day | 1.789 | 0.125 | .000 | 5.986 | 4.687 | 7.645 |

Attitudes and behaviours

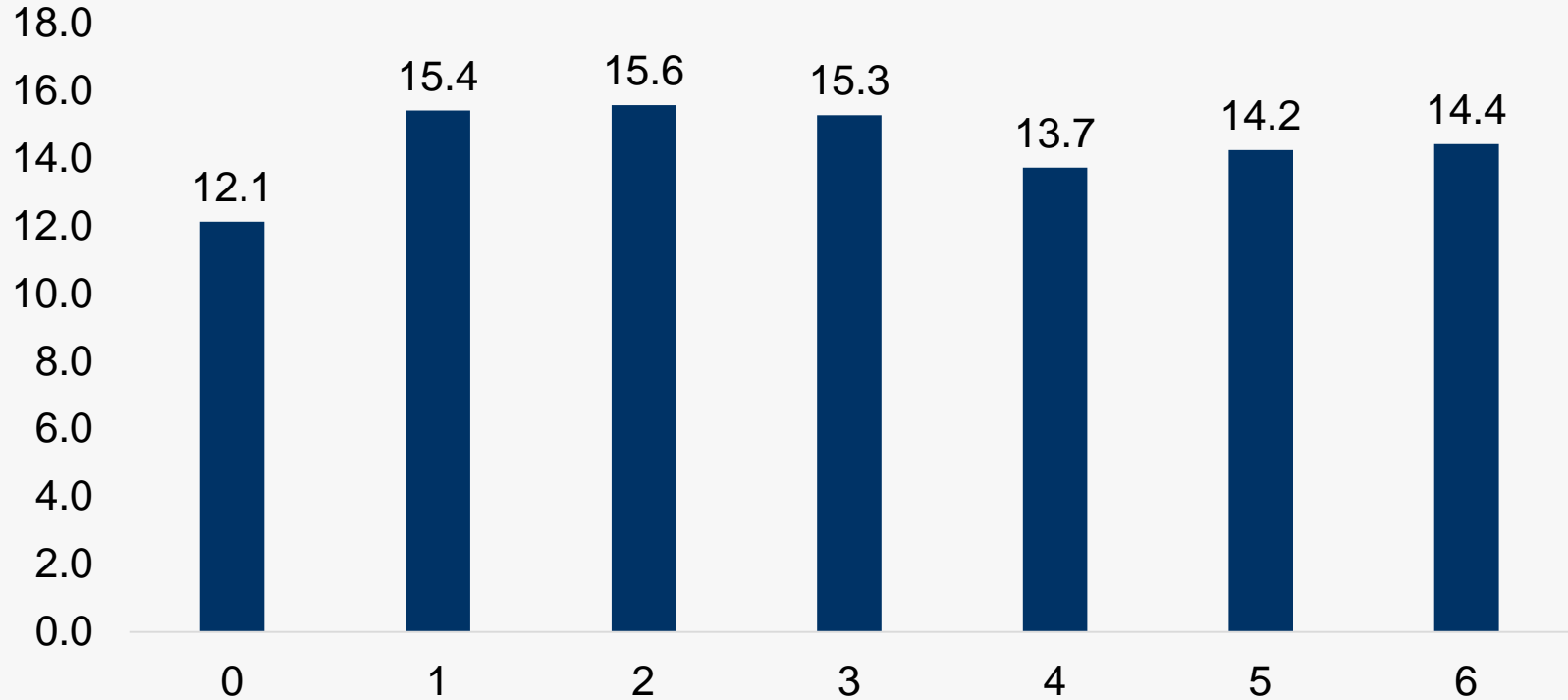
As Couper noted (2000), even though the demographics characteristics of web survey respondents match those of the population, the fundamental question is whether they are also similar on the substantive variables of interest concerning attitudes and behaviours

| | B | S.E. | Sig. | Exp(B) | 95% C.I. for EXP(B) | |
|---|--------|-------|-------|--------|---------------------|-------|
| Political participation | 0.524 | 0.069 | 0.000 | 1.688 | 1.475 | 1.931 |
| Health (self reported) | 0.270 | 0.077 | 0.000 | 1.311 | 1.128 | 1.523 |
| Household's total net income | 0.038 | 0.014 | 0.006 | 1.039 | 1.011 | 1.067 |
| Most people can be trusted | 0.051 | 0.016 | 0.002 | 1.052 | 1.019 | 1.086 |
| Trust in country's parliament | 0.036 | 0.015 | 0.017 | 1.037 | 1.007 | 1.068 |
| How worried about climate change | 0.207 | 0.074 | 0.005 | 1.231 | 1.065 | 1.422 |
| How satisfied with life as a whole | 0.035 | 0.019 | 0.059 | 1.036 | 0.999 | 1.075 |
| Gays and lesbians free to live life as they wish | 0.095 | 0.079 | 0.229 | 1.099 | 0.942 | 1.283 |
| Immigrants make country better place to live | 0.013 | 0.016 | 0.418 | 1.013 | 0.982 | 1.046 |
| How often socially meet with friends, relatives or colleagues | 0.003 | 0.023 | 0.884 | 1.003 | 0.959 | 1.049 |
| Domicile - Country | | | | | | |
| Town/Small City | 0.035 | 0.081 | 0.670 | 1.035 | 0.883 | 1.213 |
| City | -0.025 | 0.082 | 0.764 | 0.976 | 0.831 | 1.146 |
| Gender | 0.181 | 0.067 | 0.007 | 1.198 | 1.051 | 1.366 |
| Age | -0.007 | 0.002 | 0.001 | 0.993 | 0.989 | 0.997 |
| Years of education completed | 0.050 | 0.011 | 0.000 | 1.051 | 1.029 | 1.074 |

Implications

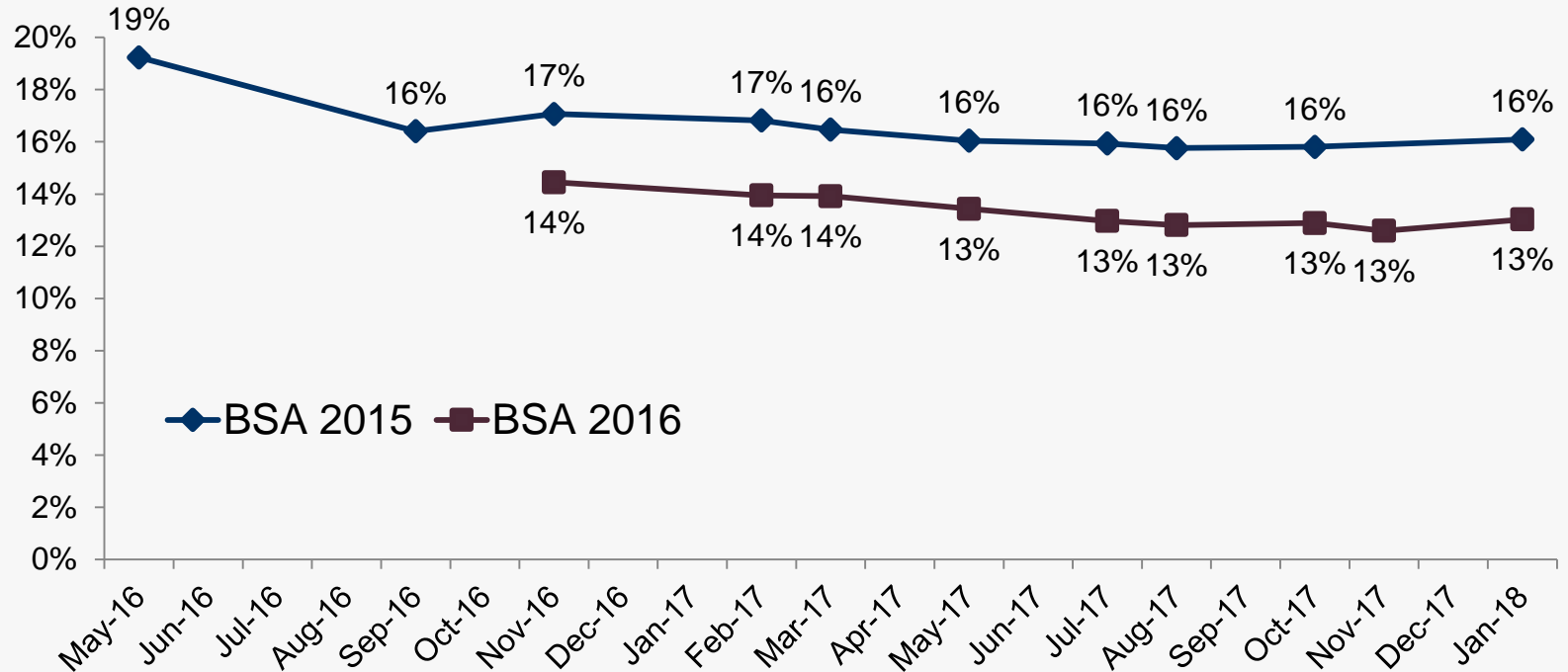
- Age, internet use and especially the interaction between these variables suggest that web surveys remain an unsuitable way to survey those aged over 75 years, particularly in those countries with a low internet penetration rate.
- Analysis shows that the key driver of panel participation is internet use (NatCen panel found similar but less strong effect).
- Taking into account that internet penetration rates are very likely to increase in the future - and ITC literacy along with it - web surveys and web panels should be progressively less affected by sample composition bias and non-response error in future.
- At the same time, we should be aware that as internet penetration rates increase, the discrepancy between the online and offline population might increase as well. The off-line group may represent a more and more highly isolated and extremely differentiated niche.

Response rates over time (CRONOS GB)



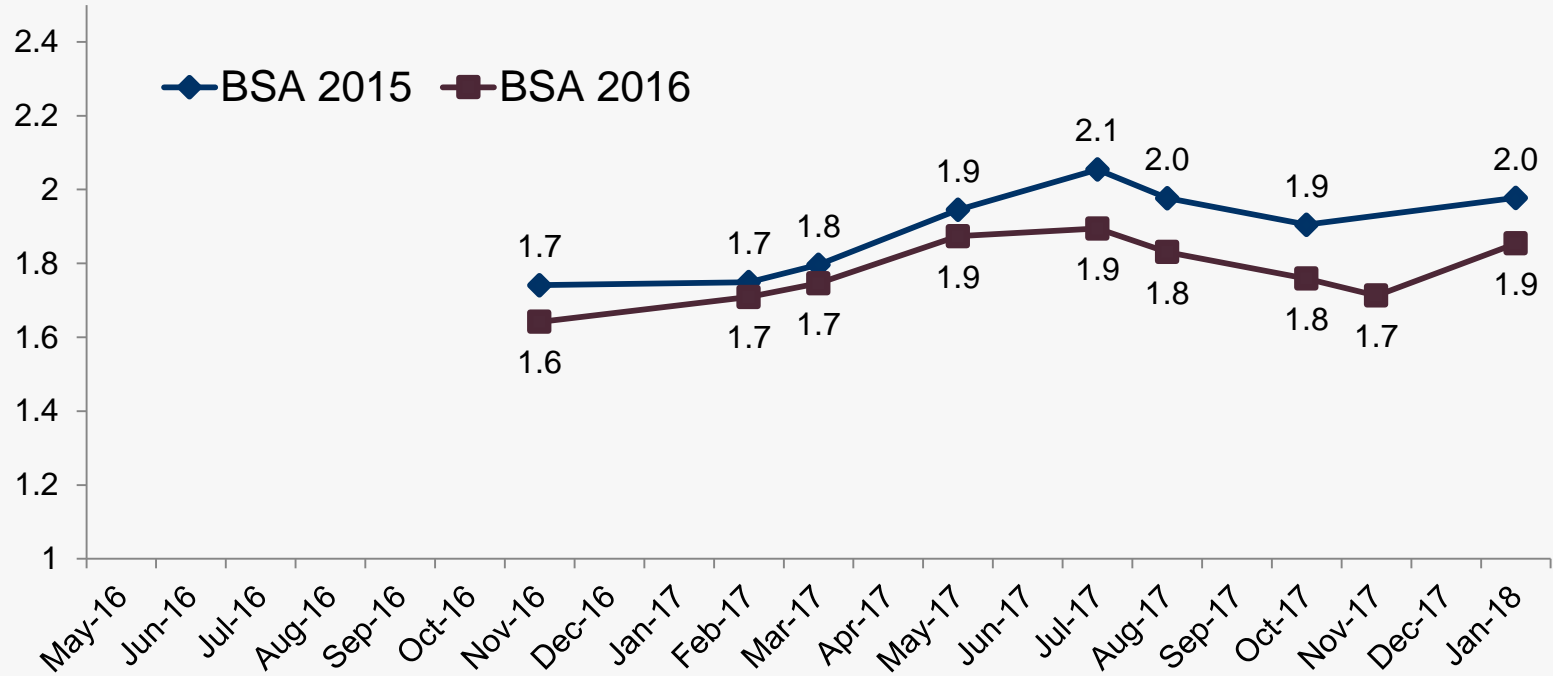
Base: all eligible cases issued to ESS (4,447)

Response rates over time (NatCen Panel)

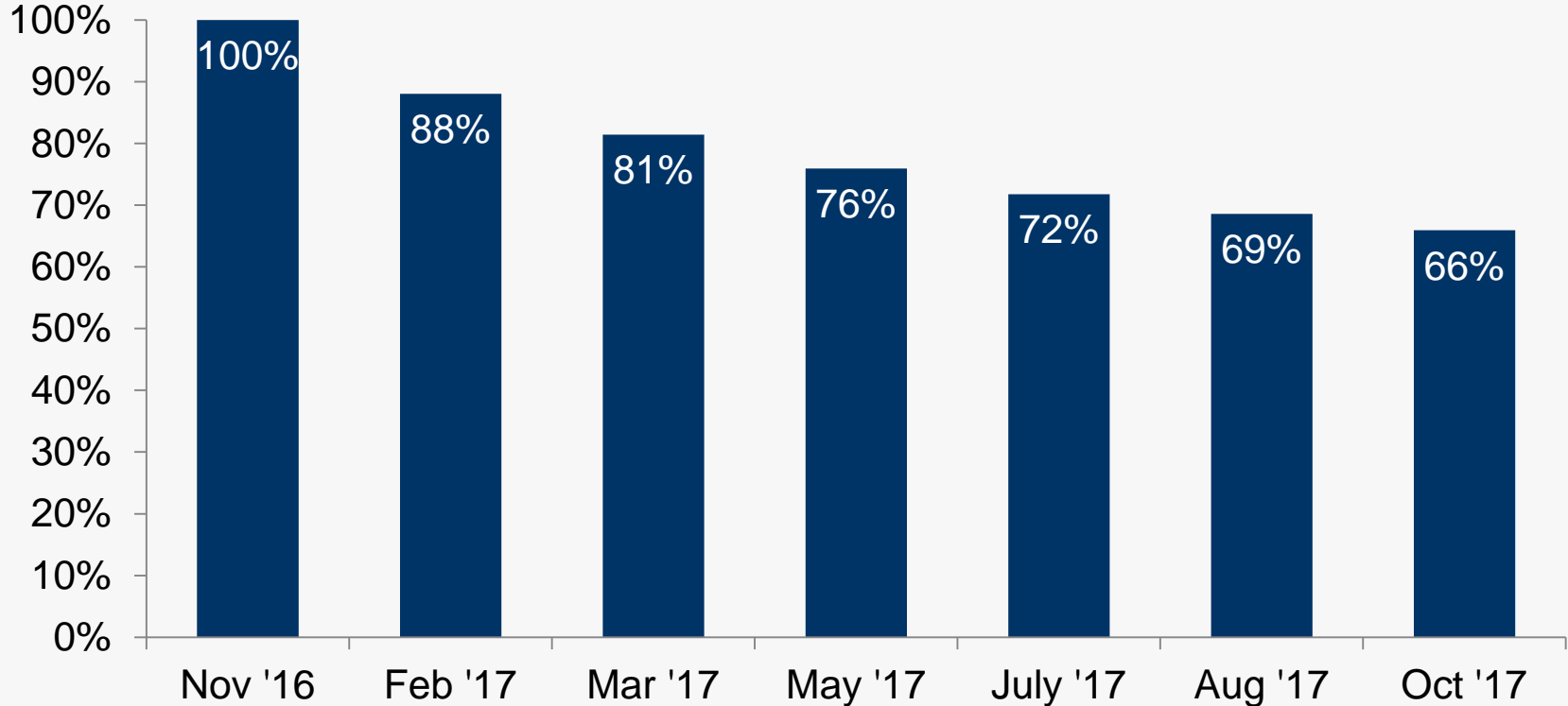


Base: all eligible cases issued to BSA 2015 (8,489) & BSA 2016 (6,408)

DEFF over time (NatCen Panel)



Re-interview rates (NatCen Panel)



Base: all productive cases in Nov '16 (2,375)

Using online panels

- **Range of data collected on a range of substantive topics, for different organisations:**
 - Attitudinal, behavioural, experiential, knowledge, experimental
 - Cross-sectional, longitudinal, data linkage, qualitative follow-ups
 - Health, wellbeing, personality traits, income, internet use, politics...
 - Government departments, charities, academics, commercial organisations...
- **But also used for ongoing experimentation:**
 - Comms strategies: mode, timing, messaging, incentives
 - Questionnaire testing
 - Fieldwork design: targeted designs, fieldwork length, role of telephone interviewers...

Summary

- Probability-based panels provide an effective alternative to ‘traditional’ probability-based methods
 - Fieldwork can be conducted quicker, and at lower cost
 - Initial analysis suggest the produce robust estimates, and this quality is maintained over time
- Lots of opportunities to make the most of this infrastructure
- But, as with all methods, we should be aware of their potential limitations

Future work

- Further analysis of the methodology
 - Comparisons to benchmarks; impact of attrition/panel conditioning/ageing cohorts; understanding sampling/measurement effects
 - Comparisons between the two panels
- Continued experimentation and development
 - R-indicators; targeted design; comms & incentive strategies
- Explore potential utility
 - Longitudinal analysis; experimental designs; qualitative follow-ups; bio-medical