EVIDENCE AND CONTRIBUTIONS SUPPLIED BY...
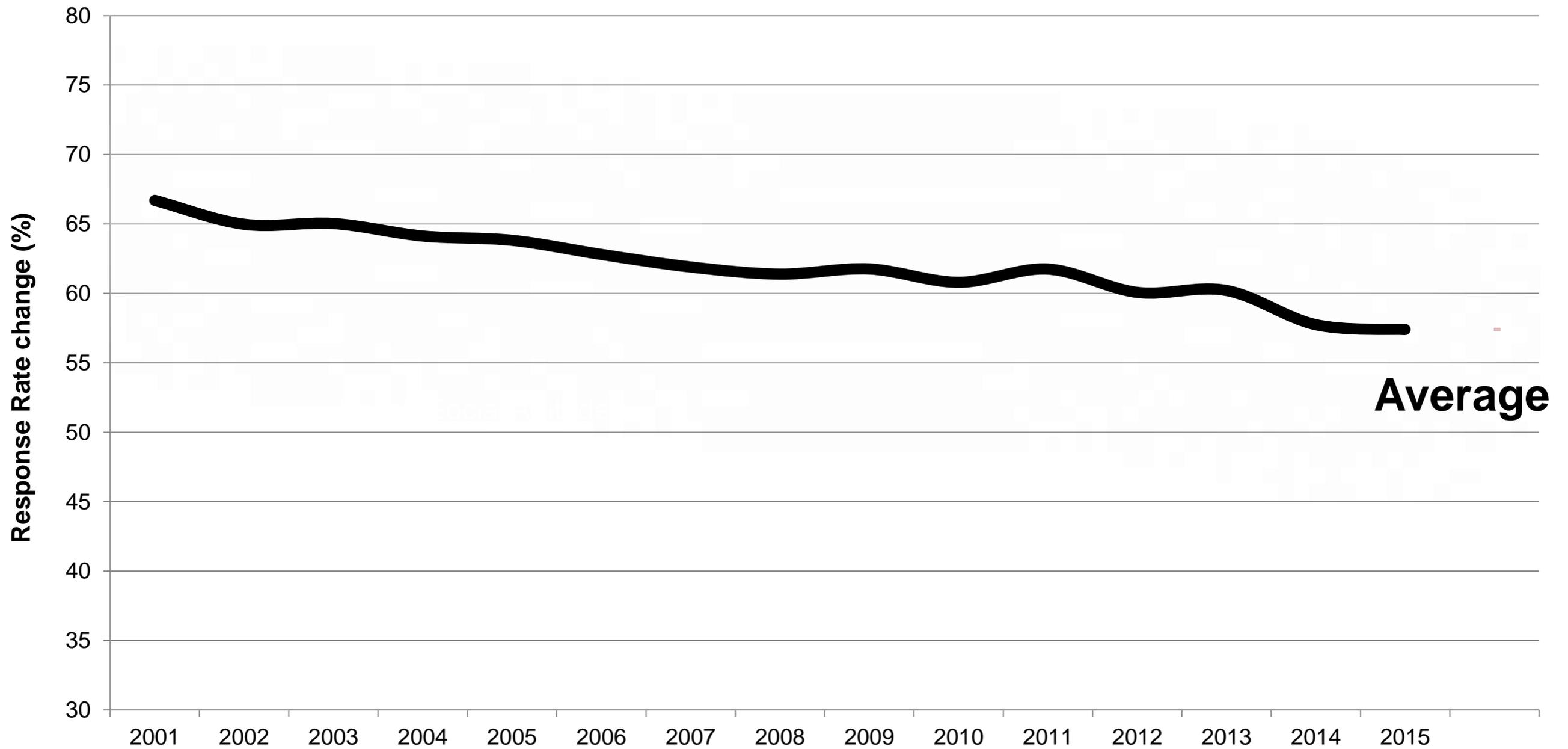
CONTENTS

1| **Background: what has been happening to response rates?**

2| **Looking beneath the bonnet – levels of effort**

3| **Implications and conclusions**

01

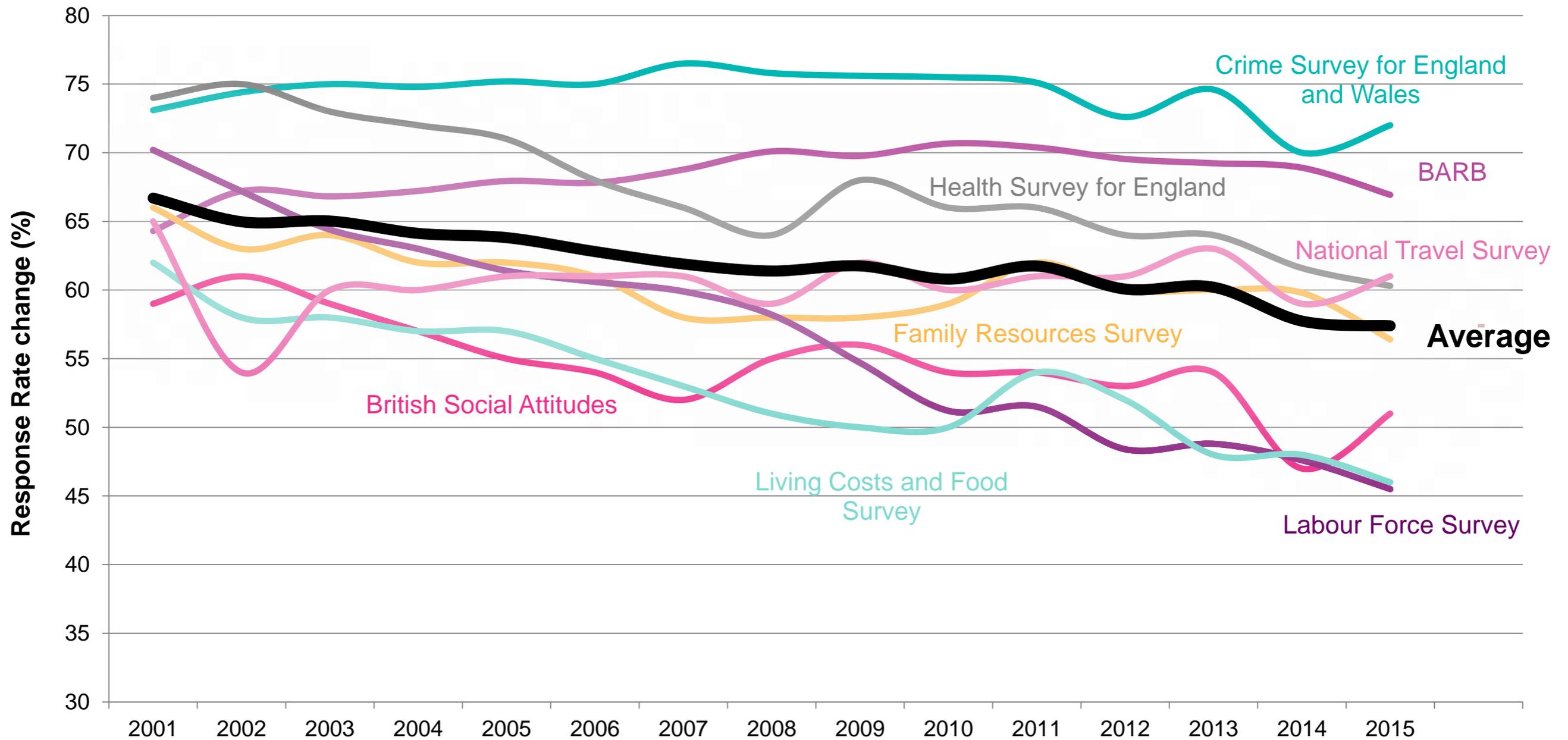## BACKGROUND : WHAT HAS BEEN HAPPENING TO RESPONSE RATES

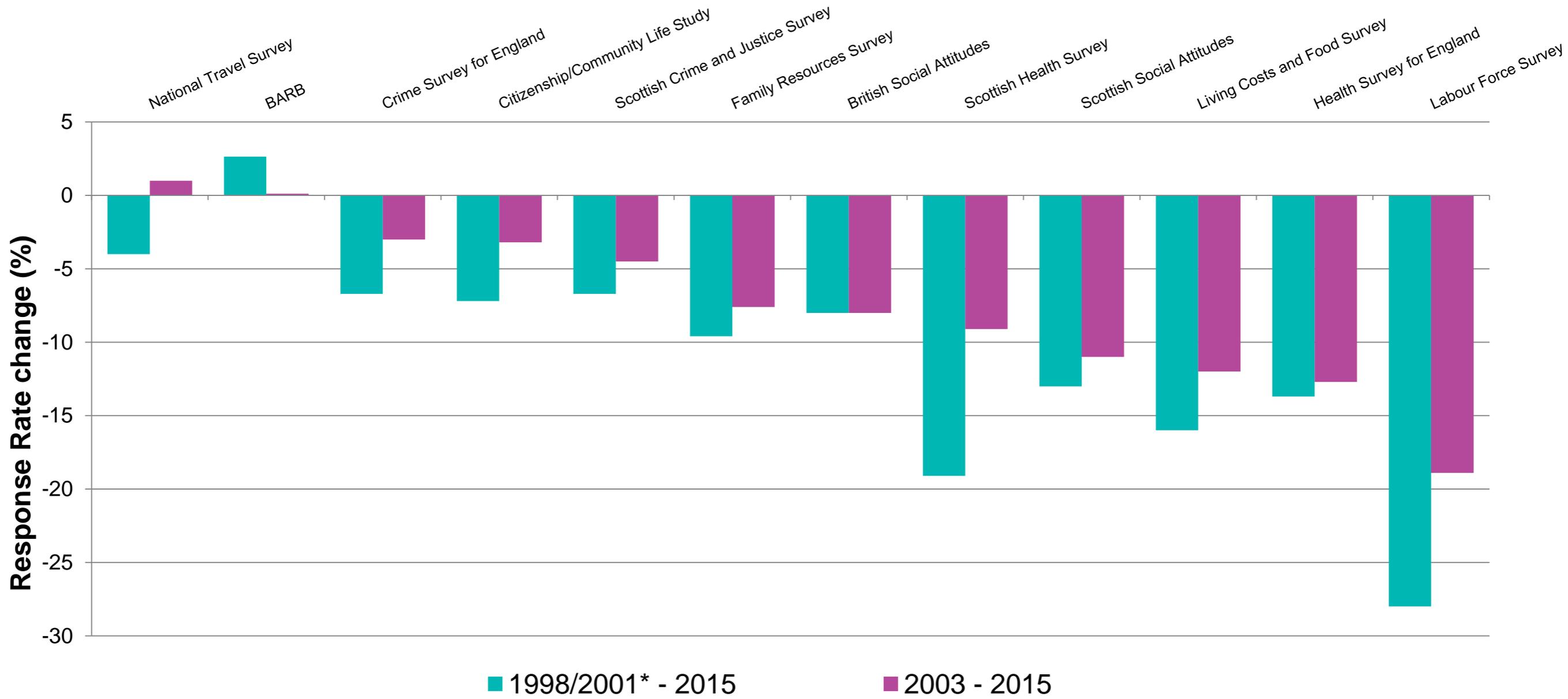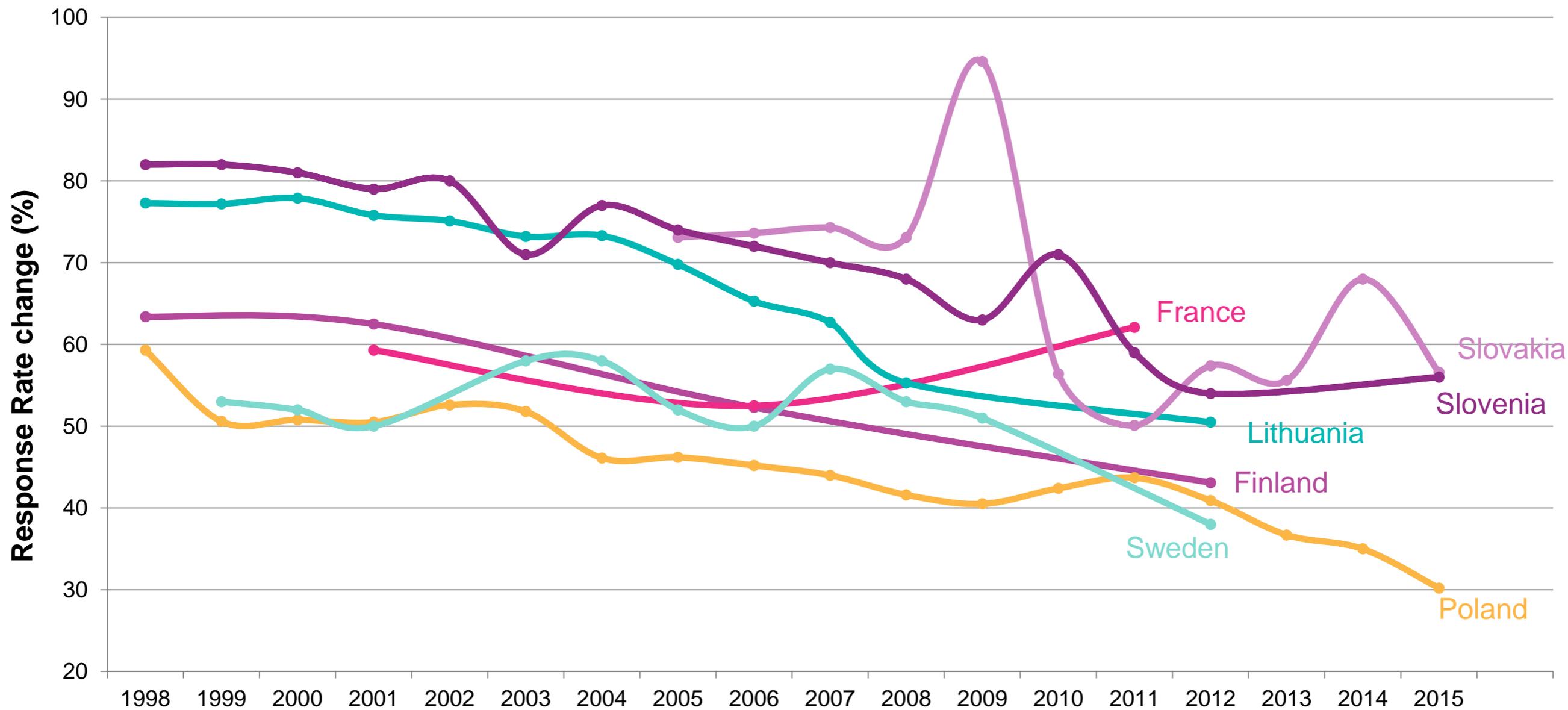# SURVEY RESPONSE RATES BY YEAR: AVERAGE

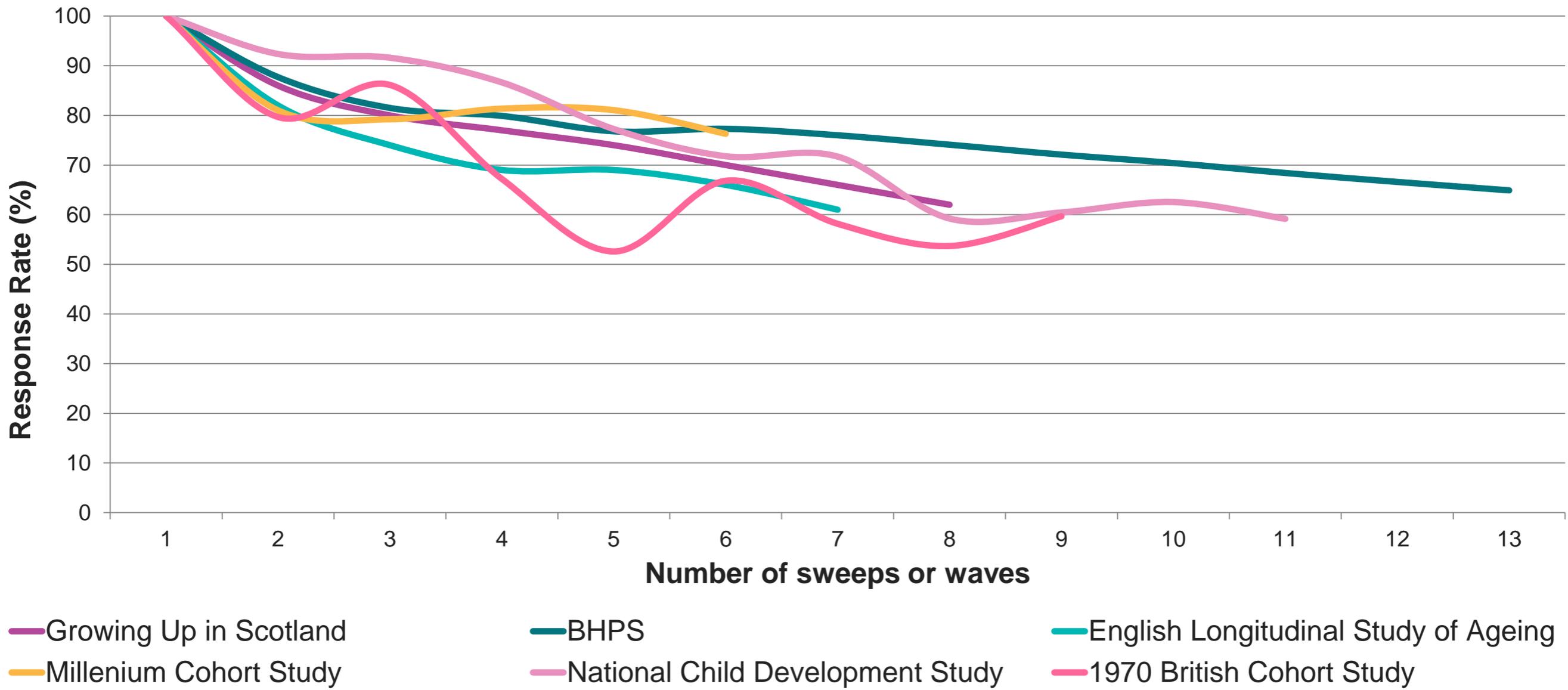# SURVEY RESPONSE RATES BY YEAR: DETAIL

SURVEY RESPONSE RATES: CHANGE OVER TIME

RESPONSE RATES FOR THE HOUSEHOLD BUDGET SURVEY

LONGITUDINAL STUDIES ATTRITION RATES

POSSIBLE EXPLANATION

- **Less trust in government, brands and professions**
- **Survey fatigue**
- **Availability**
- **Accessibility**

LESS TRUST IN GOVERNMENT AND POLITICIANS

Almost never

MAJOR FACTORS AFFECTING RESPONSE

FIELDWORK

- Volumes & allocation
- Management
- Reissues

PARTICIPANT COMMUNICATIONS

- Messaging
- Channels
- Volume
- Timing

Societal change

SURVEY DESIGN

- Sample type
- Geography
- Clustering
- Sample data
- Selection method
- Interview length
- Topic
- Mode

Response rate

INCENTIVE

- Type
- Value

INTERVIEWER/NURSE

- Training
- Motivation
- Management
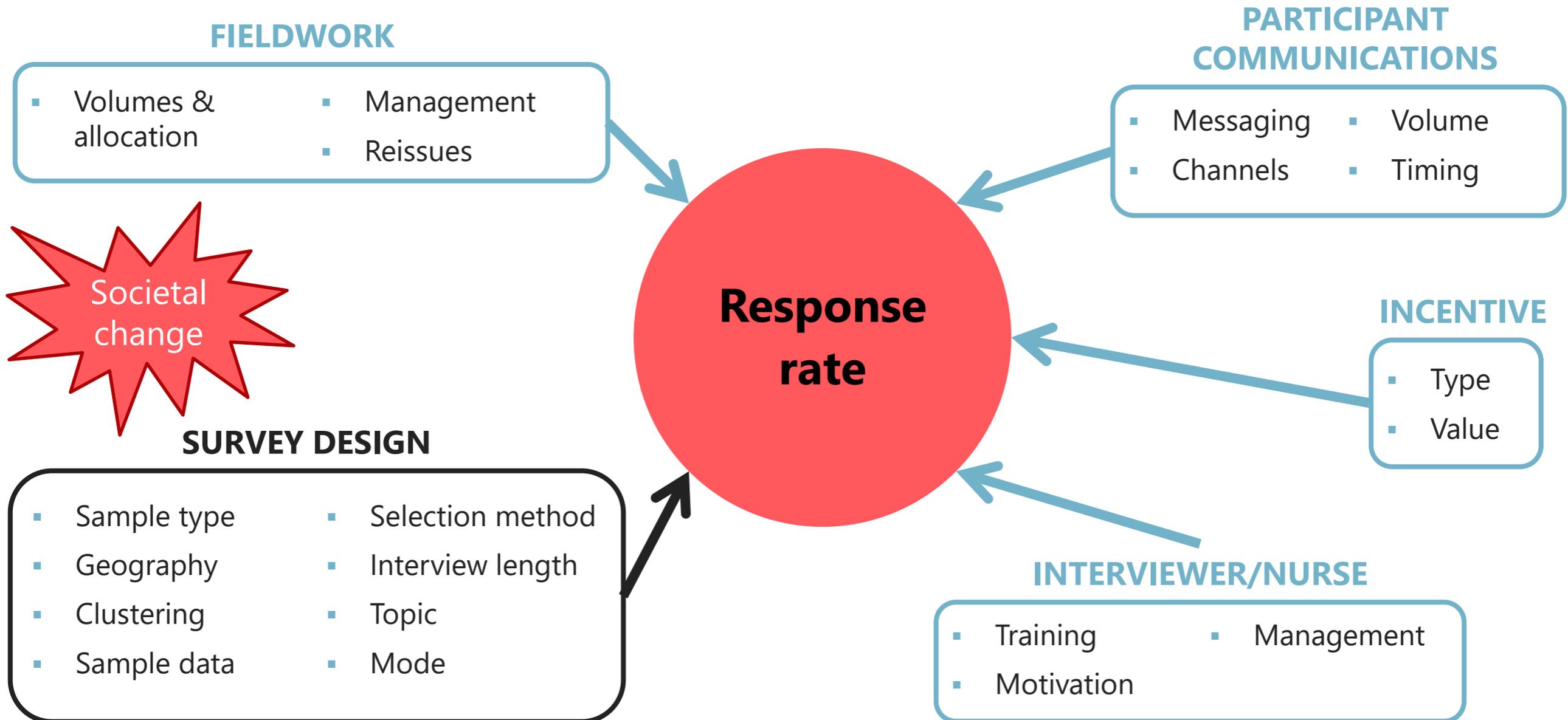
02

LOOKING BENEATH THE BONNET : LEVELS OF EFFORT

DETAILS OF THREE CASE STUDIES

|	BARB	BSA	CSEW
First conducted	1991	1983	1981
Sample design	Tightly clustered	Clustered	Unclustered
PSUs	Based on OAs	Based on postcode sectors	Bespoke but based on LSOAs
Selection	Household level interview	One adult (18+) randomly selected	One adult (16+) randomly selected
Assignment size (issued sample)	c. 17-18 addresses	c. 26 addresses	c. 32 addresses
Incentives	None	Conditional voucher (£10 from 2014)	Stamps sent with advance letter
Interview length	c. 15 mins	c. 65 mins	c. 50 mins
Fieldwork length	Up to 6 weeks	Around 4 months	Max. 6 months

CASE STUDY RESPONSE RATES
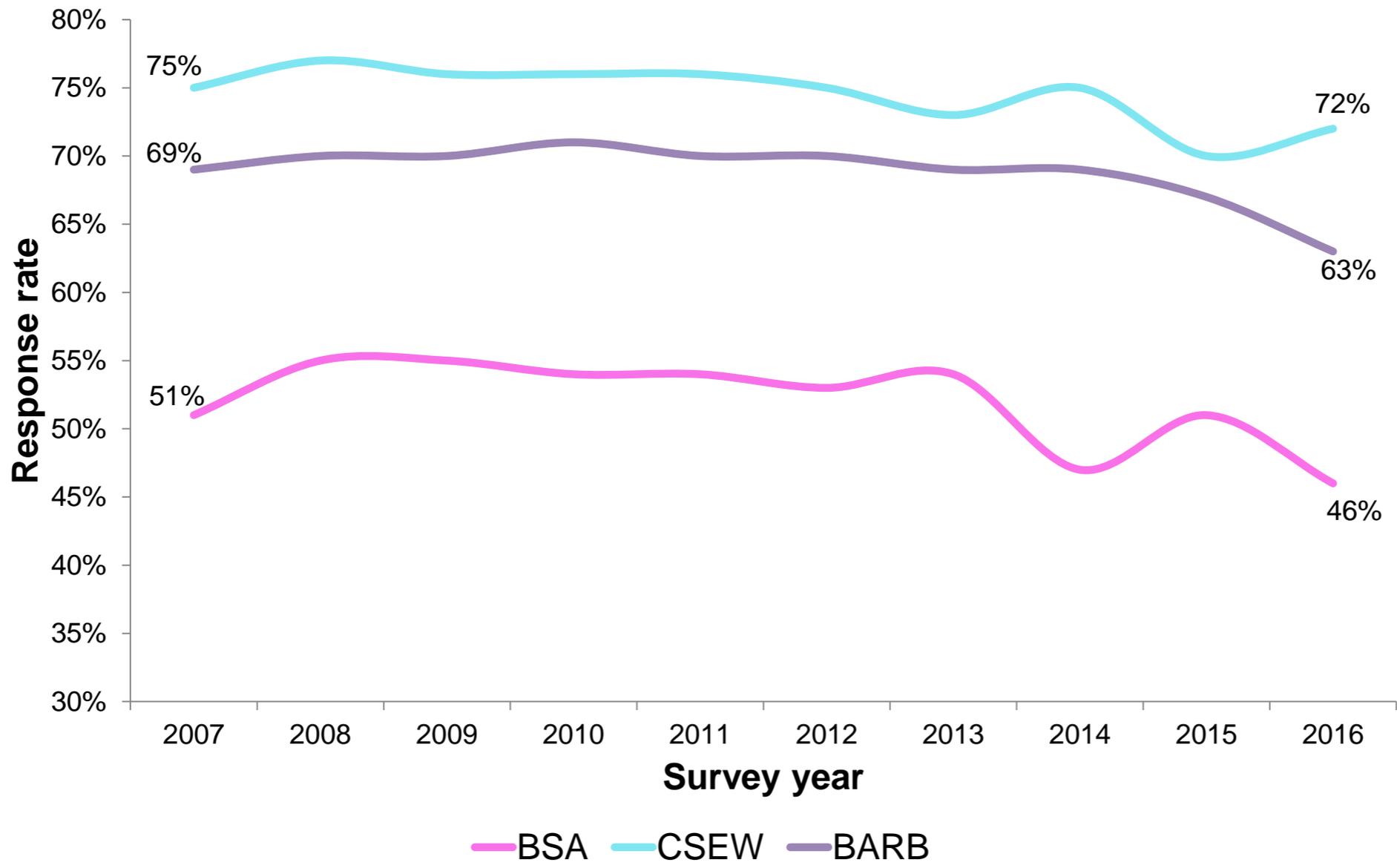
NON-CONTACT RATES HAVE BEEN KEPT AT LOW LEVELS
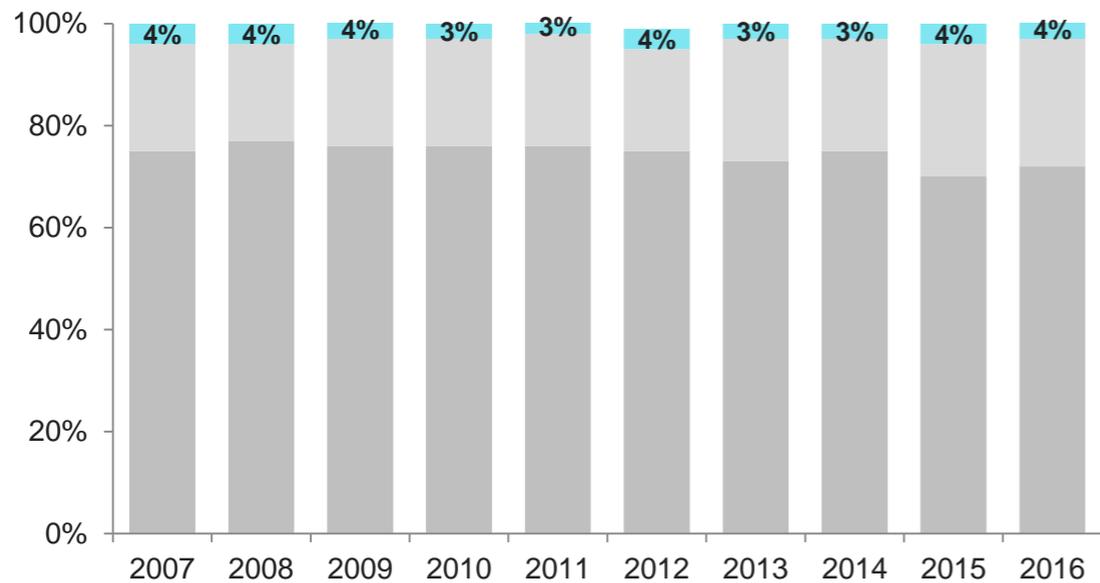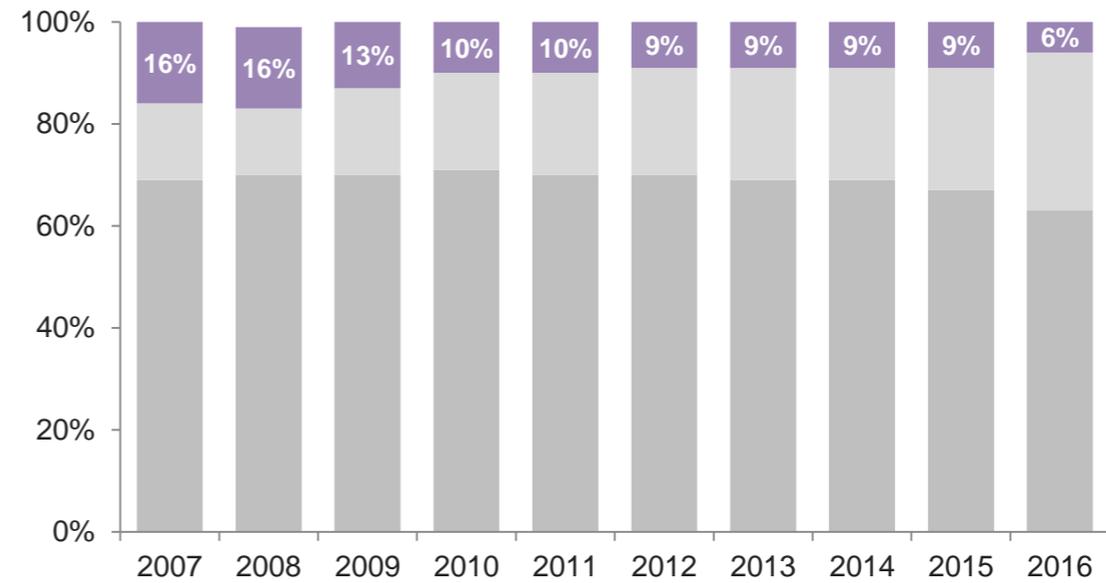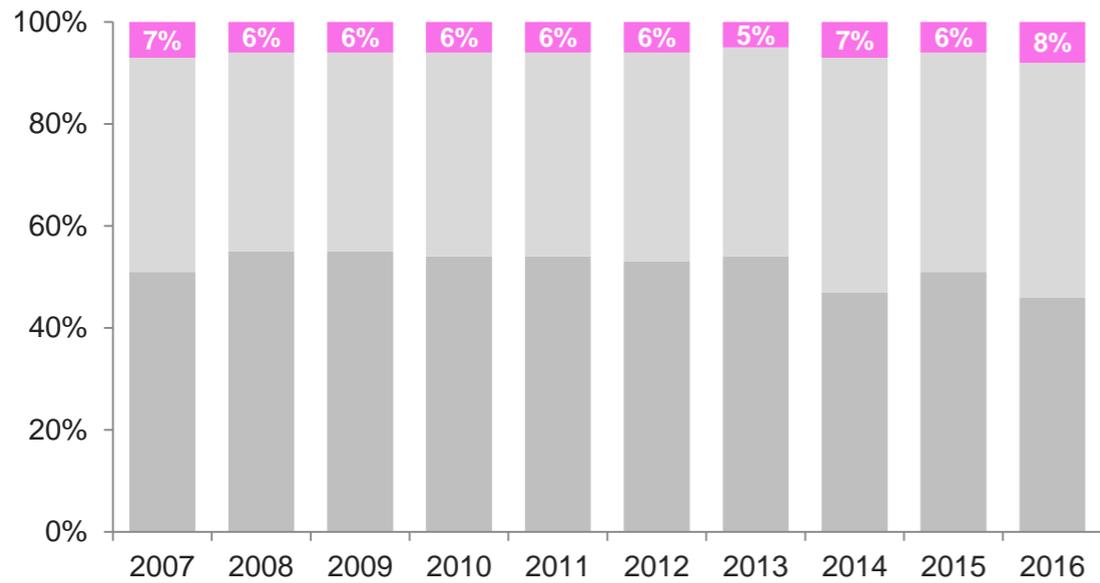
BSA
BARB
CSEW

REFUSAL RATES HAVE INCREASED IN THE LAST FEW YEARS
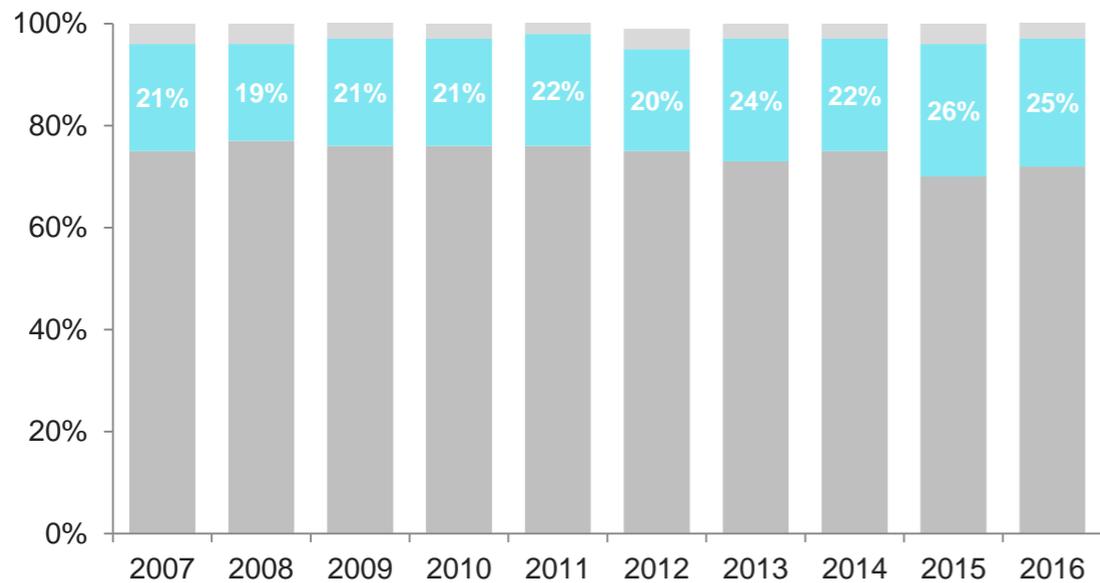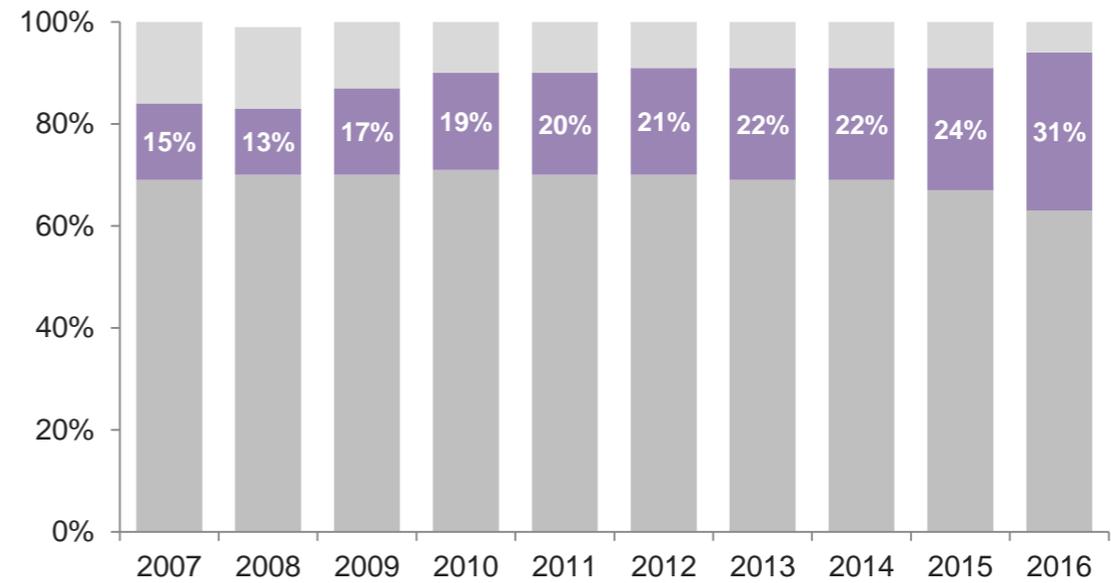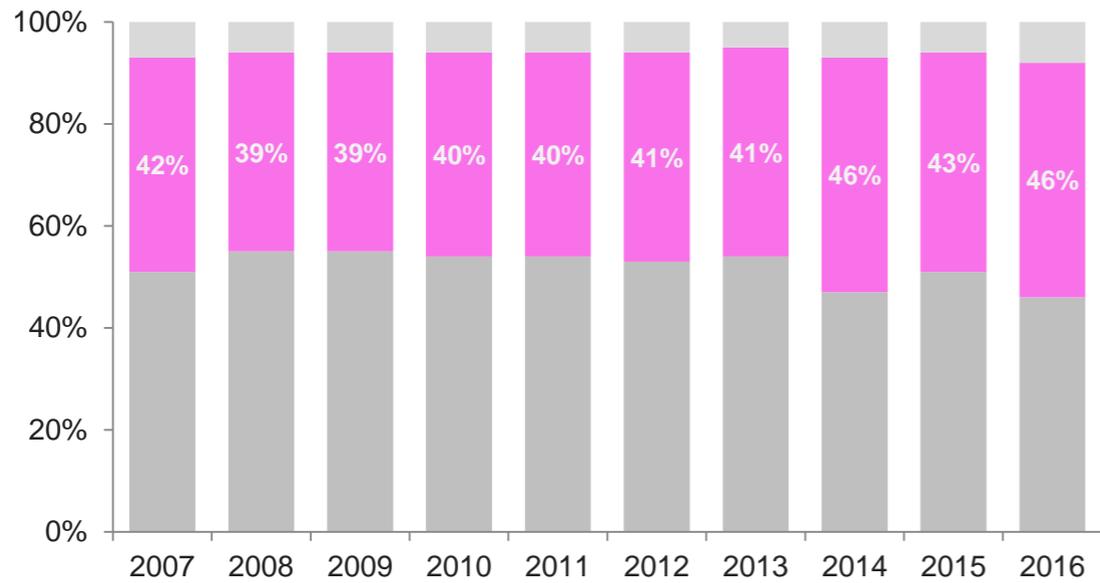
MORE EFFORT BEING PUT IN TO KEEP RESPONSE RATES UP
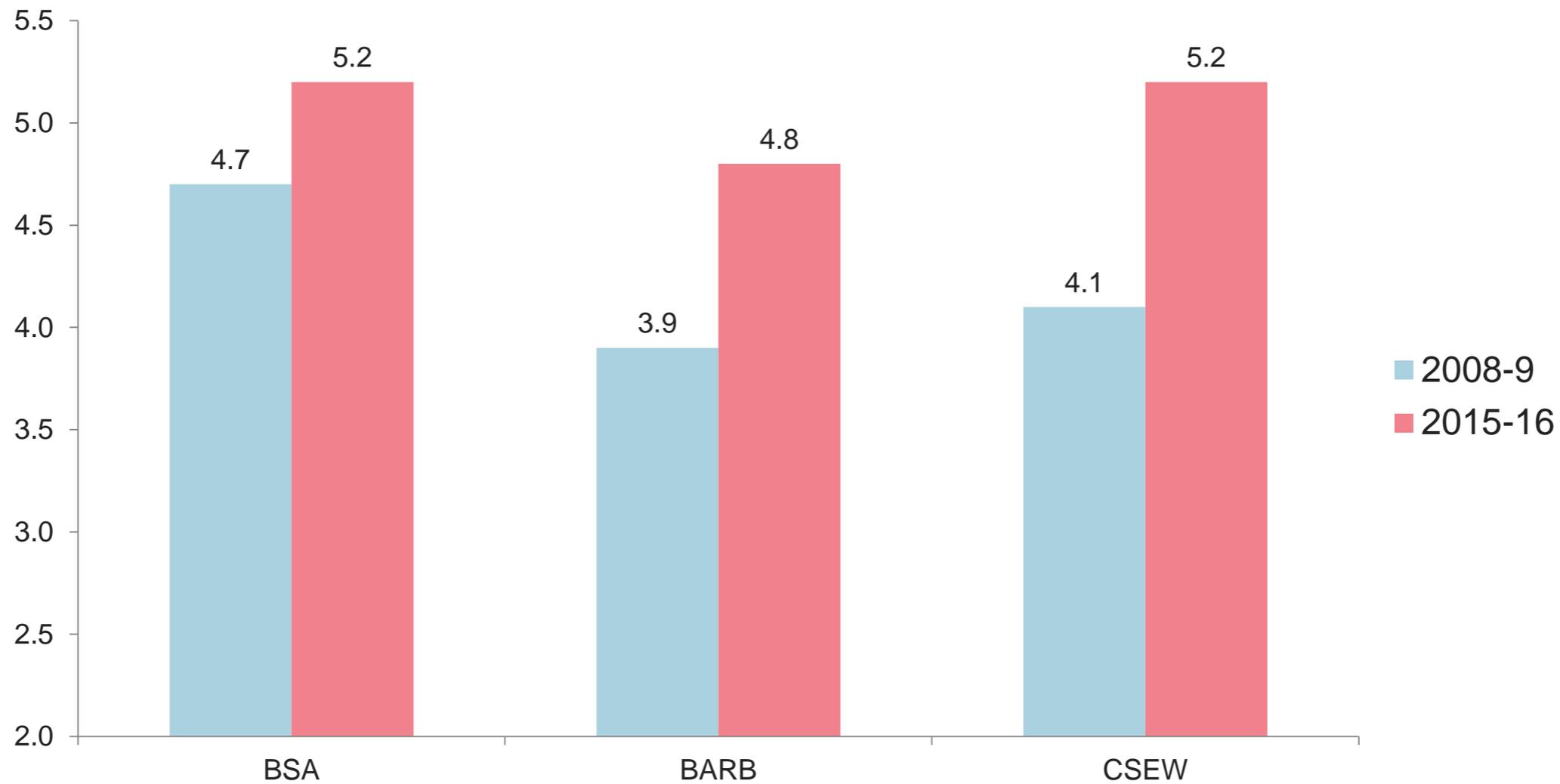
MORE EFFORT BEING PUT IN TO KEEP RESPONSE RATES UP

	2008/9	2015/16	% increase	Extra calls per year
BSA (2008-2016)	4.7	5.2	11%	c. 4,000
BARB (2009-2015)	3.9	4.8	23%	c. 72,000
CSEW (2008-2016)	4.1	5.2	27%	c. 55,000

INCREASE IN THE PROPORTION OF ADDRESSES WHICH HAVE MORE VISITS THAN MINIMUM CALL REQUIREMENTS
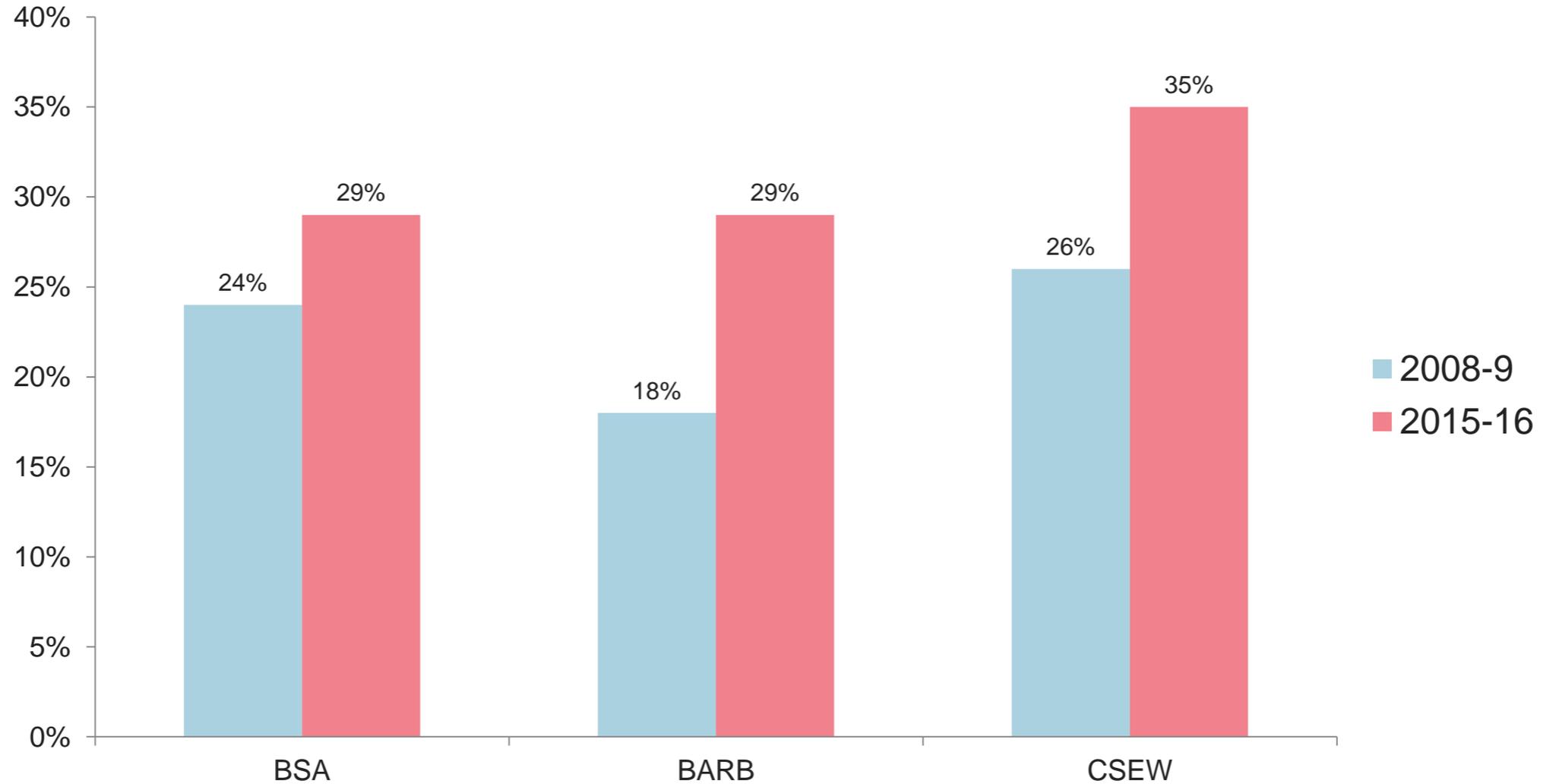
UNPRODUCTIVE ADDRESSES ARE REISSUED TO BOOST RESPONSE
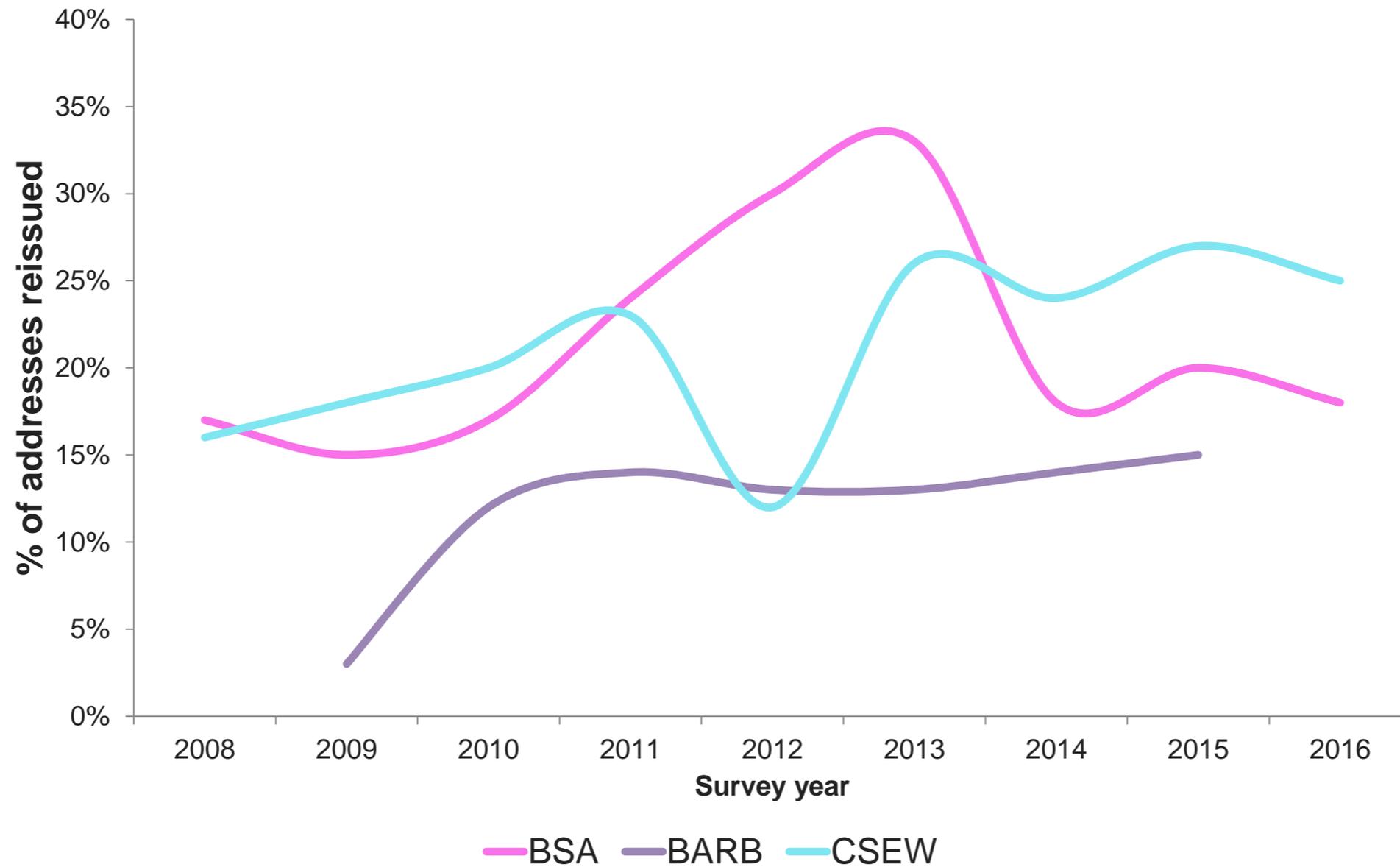
RESPONSE ON REISSUES IS LOWER ALTHOUGH HAS REMAINED FAIRLY STEADY
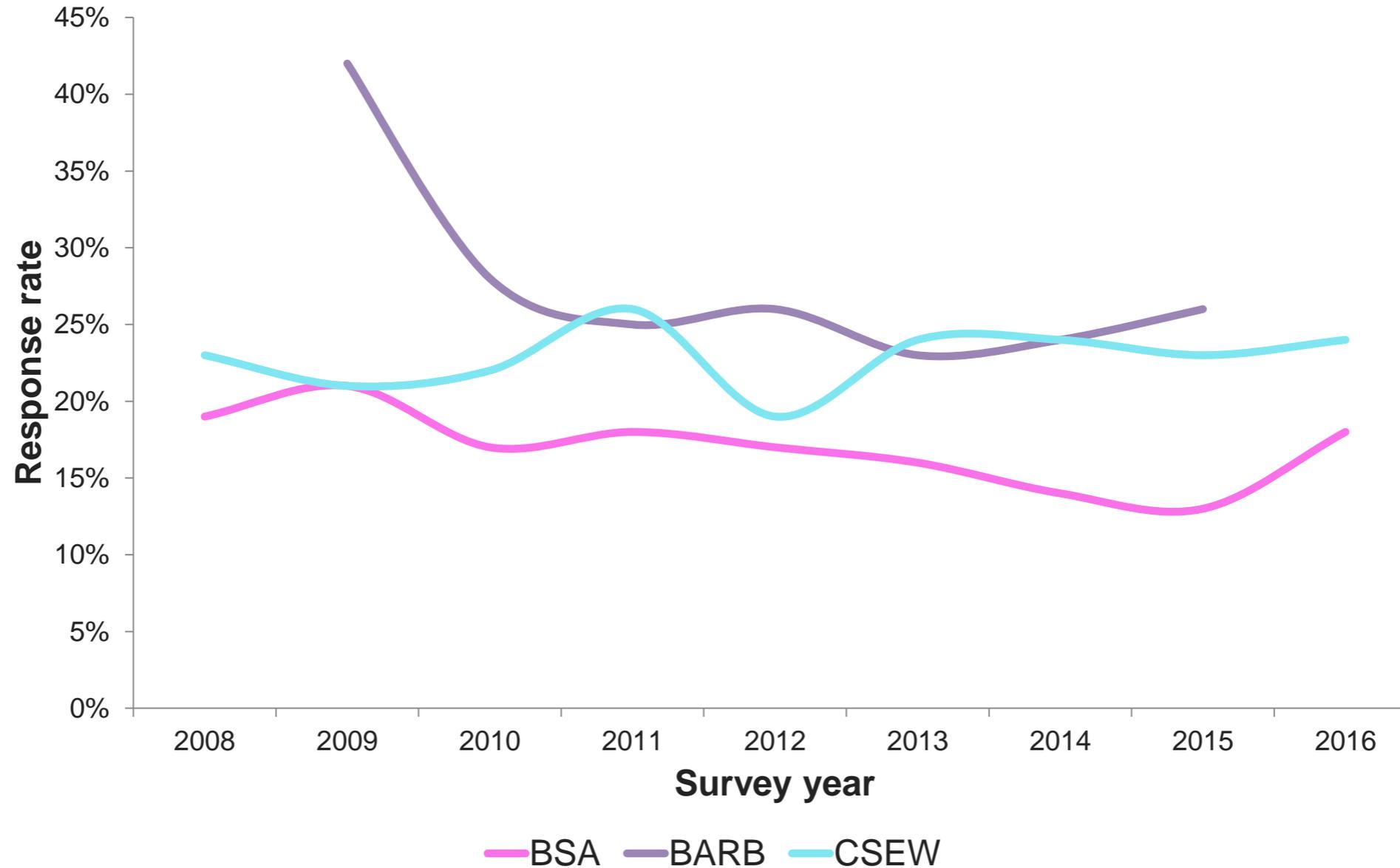
CUMULATIVE RESPONSE RATE AFTER X VISITS, 2015-16
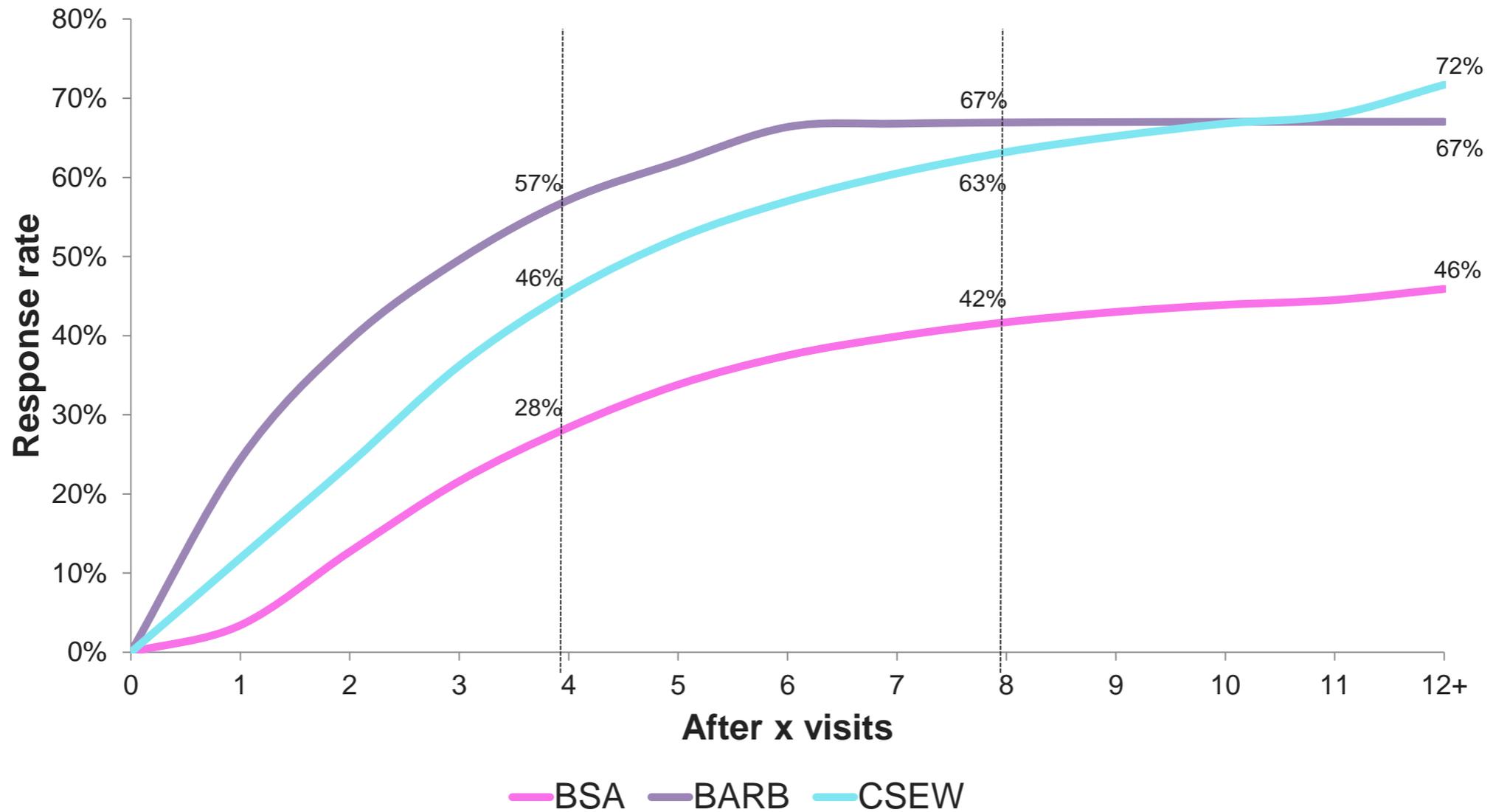
RELATIONSHIP BETWEEN RESPONSE RATE AND CALLS
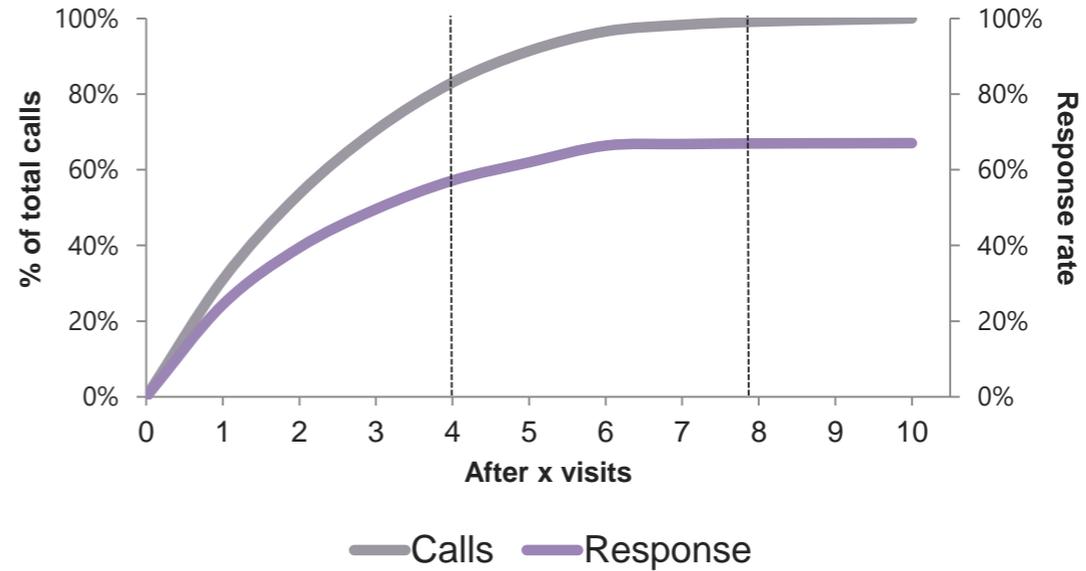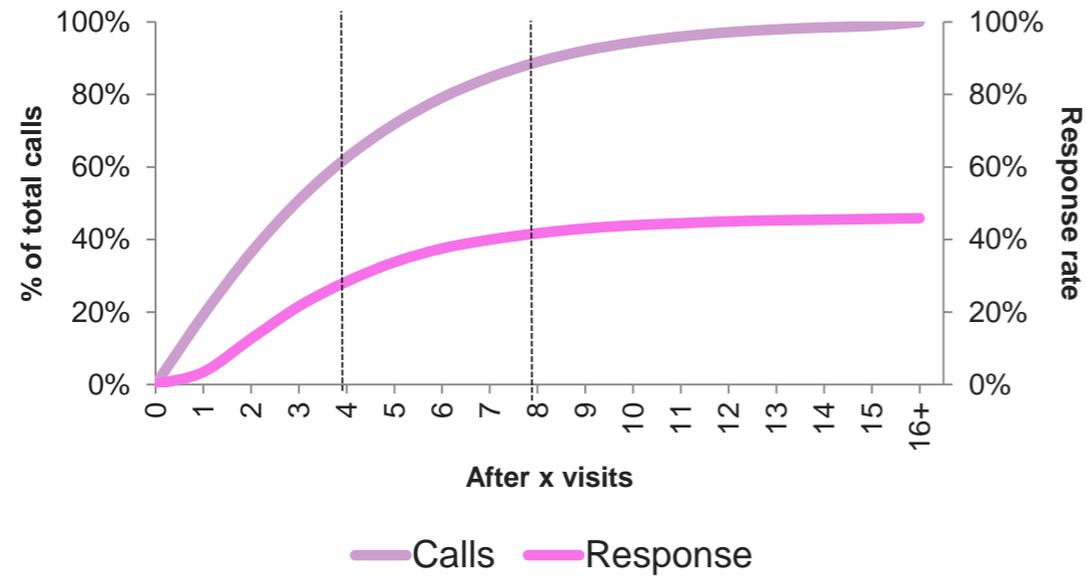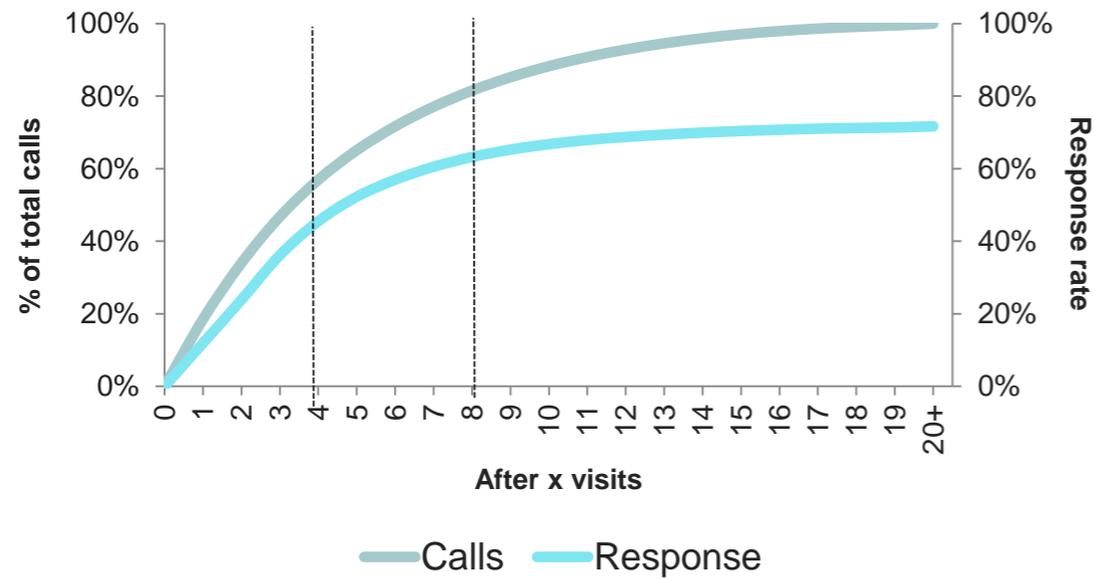
03

IMPLICATIONS AND CONCLUSIONS

CONCLUSIONS

Response rates are falling but the level of effort is increasing. Poses a number of questions:

- **What can we realistically control in influencing response?**
- **What is the cost-benefit analysis?**
- **What are the best measures of performance?**

F2F INTERVIEWING IN 2017
AN OPERATIONS' PERSPECTIVE

Sophie Ainsby (NatCen) & Mark McLaughlin (GfK)

With Caroline Baxter (Kantar Public) & Jeremy May (Ipsos MORI)

18th May, 2017

EVIDENCE AND CONTRIBUTIONS SUPPLIED BY...

INTERVIEWING DAYS AND INTERVIEWER PANEL YEAR ON YEAR 2010-2016
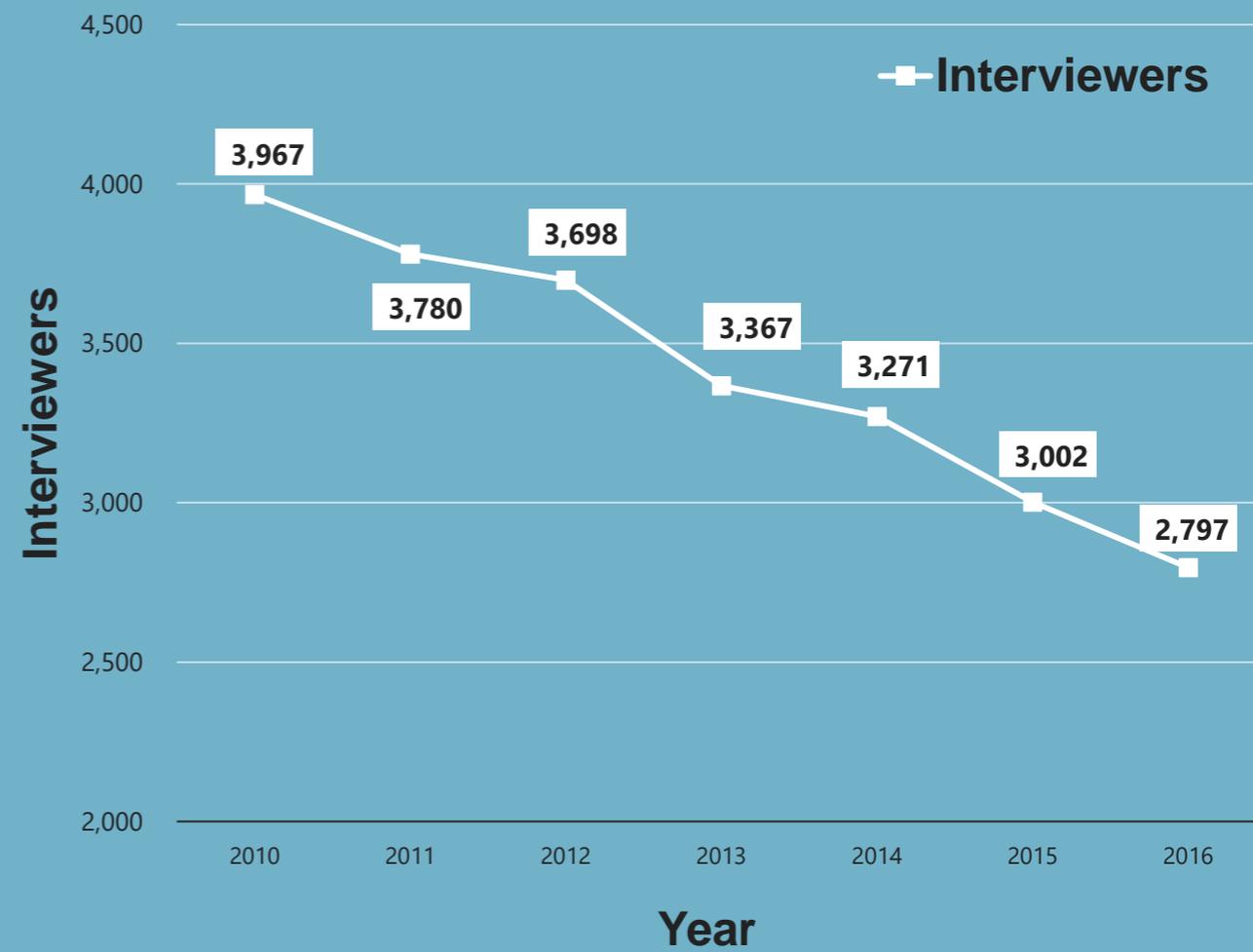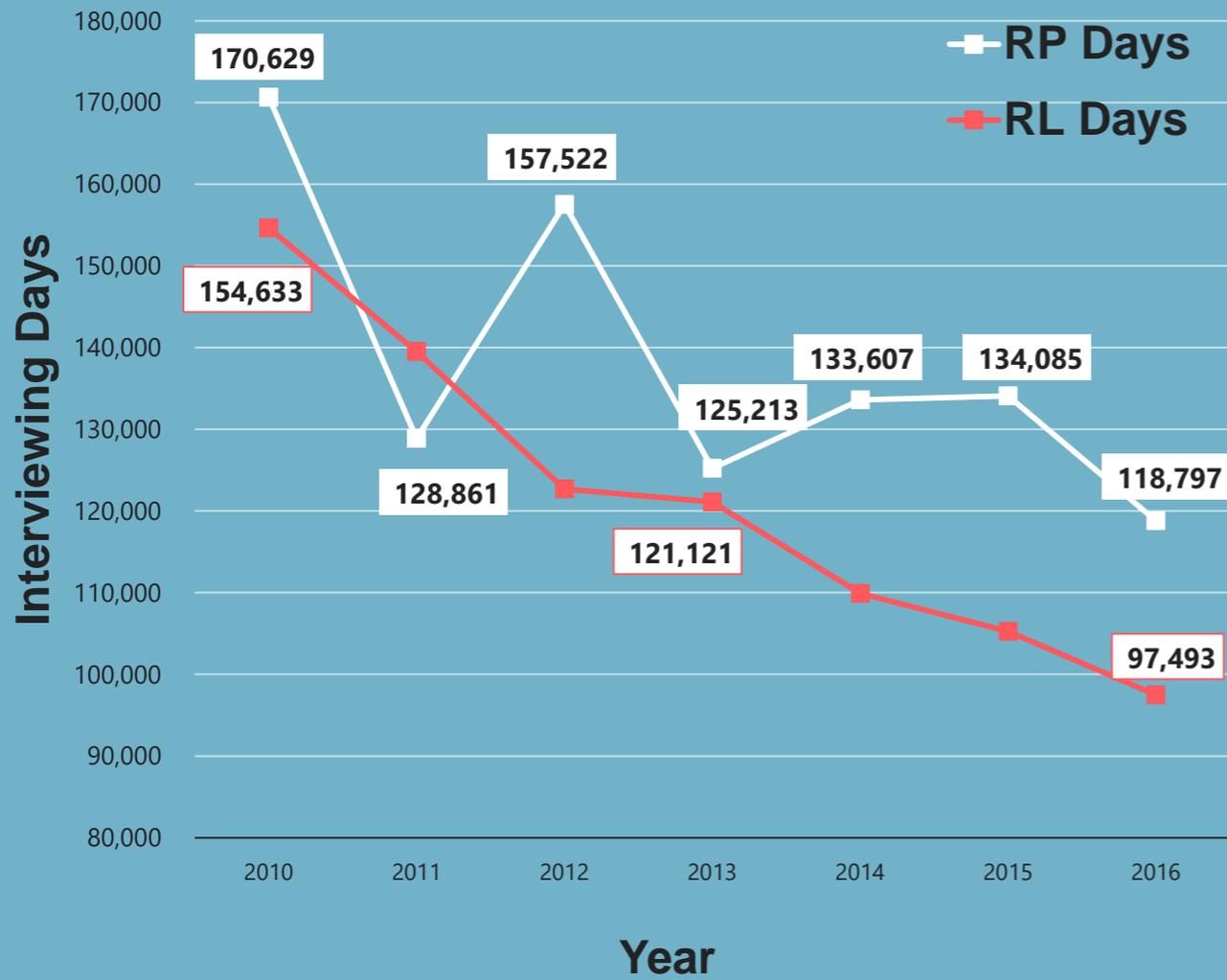
What are our operational challenges today?
F2F 2017

Recruitment
Response to recruitment advertising fluctuates with few responses translating to new panel members. We face competition for candidates from a wide range of alternative employers offering less challenging work with guaranteed income. We have tried alternative payment methods but this has largely been unsuccessful. We can still build panels but it's increasingly challenging to do so quickly.

Retention
New people need an introduction to F2F to help maintain motivation and build doorstep skills and experience and this can be challenging due to work types available. Payment by results also has an impact- on first projects trainees may see their earning potential limited.

Response Rate Targets
On paper response rates remain as they were 10 years ago. However, saturation of customer satisfaction and cold calling makes target achievement more challenging than it was. Research frequently makes headlines, but there isn't a great deal of public awareness regarding the vital role they play.

Operations/Field Challenges Today

2016 Recruitment: from applicant to interviewer

c17K
Respond to adverts

C1.7K
Invited to Train

1180
Join interviewer panel
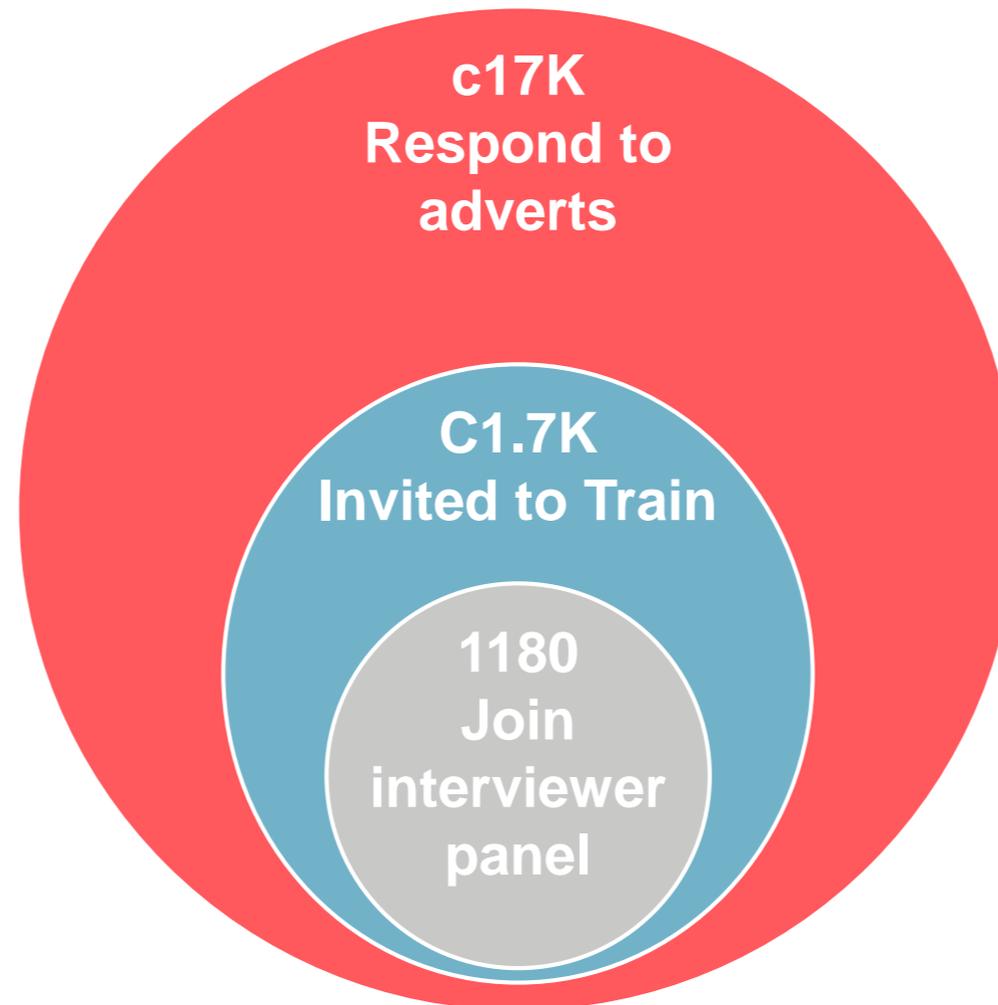
An average of 7% of applicants actually became interviewers in 2016

JOINERS & LEAVERS 2012-2016

Year	32%	47%	35%	46%	53%

Joiners / Leavers / Panel Size data:
- 2012: Joiners 1444, Leavers 1186, Panel Size 3698
- 2013: Joiners 1021, Leavers 1593, Panel Size 3367
- 2014: Joiners 1117, Leavers 1161, Panel Size 3271
- 2015: Joiners 1322, Leavers 1371, Panel Size 3002
- 2016: Joiners 1180, Leavers 1495, Panel Size 2797

Legend:
- Joiners
- Leavers
- Panel Size
- Leavers as % Total Panel
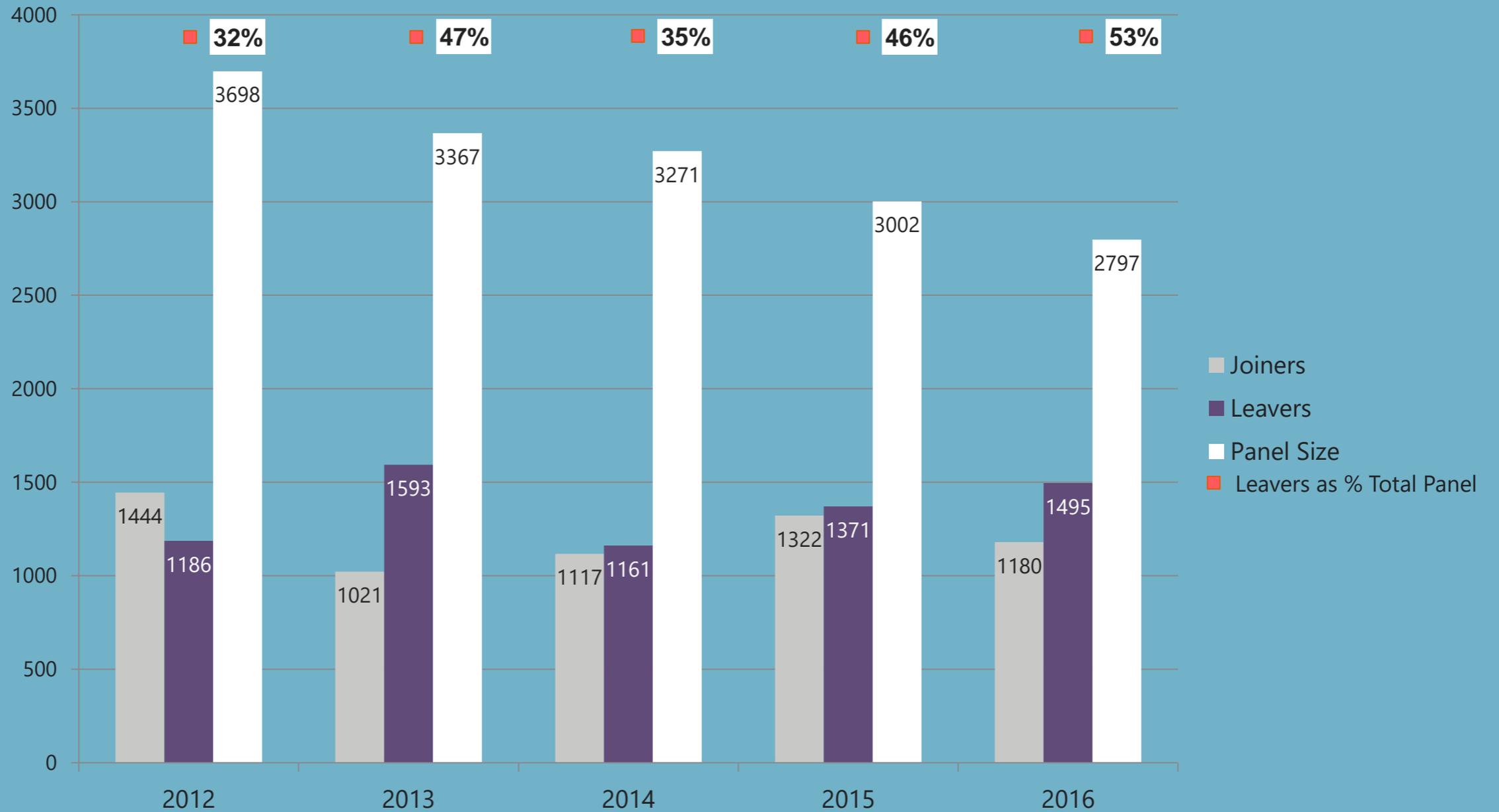
INTERVIEWER LENGTH OF SERVICE

2006 vs 2016
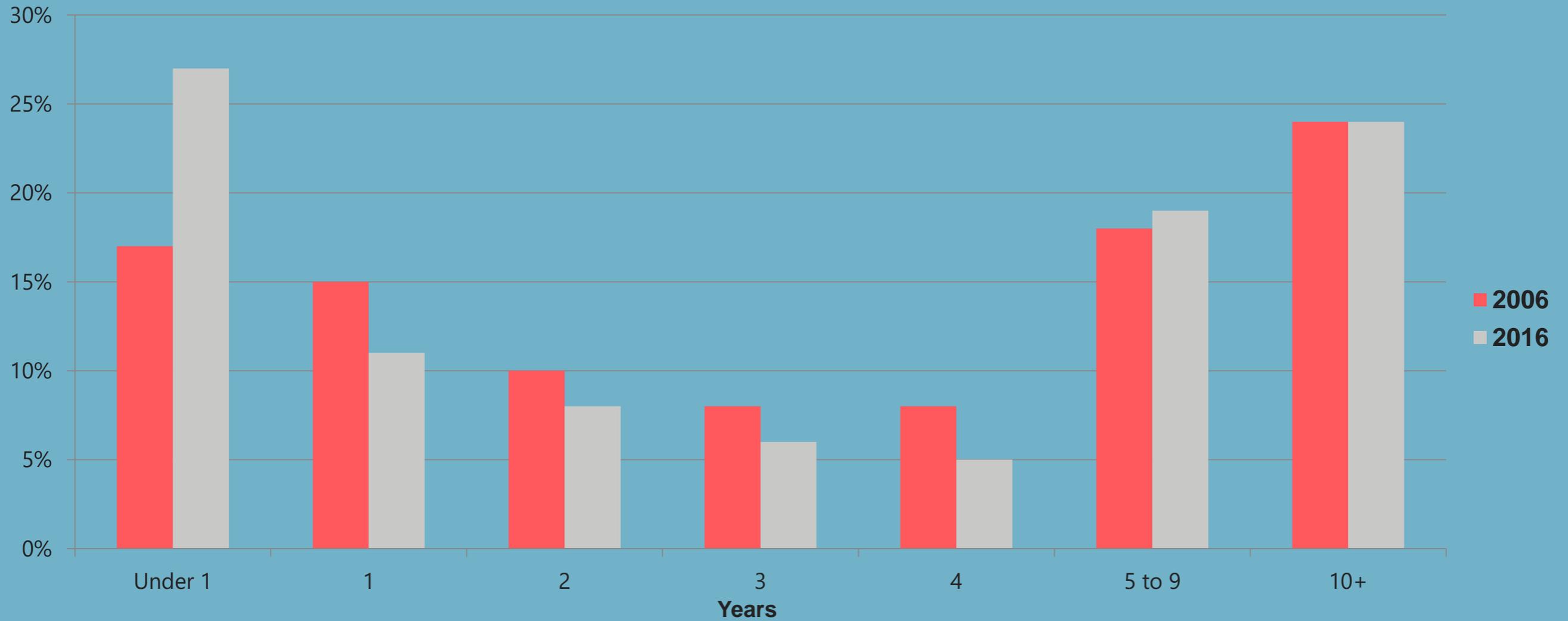
MITIGATING THE ISSUES

Operations Actions

Cost

Over the last three years c£3M spent on mitigating recruitment & retention challenges. This cost is fully absorbed and not passed on.

Support

Increased support for new people; mentoring programmes and engagement measures

Challenges

Trying to understand and address our interviewers' challenges via a cross company interviewer survey

INTERVIEWER SURVEY

As an industry we are keen to better understand our interviewers' challenges

- Produced in conjunction with the MRS, the survey began in 2016. 2017 wave has just been completed

- 93% of interviewers feel pay levels influence why people choose to stop interviewing

- 86% feel the challenges of getting participants to agree to be interviewed is also a significant factor

- Experienced interviewers really enjoy random probability work- 81% rate interest of subject matter highly

INTERVIEWER COMMENTS ON WHAT THEY FIND CHALLENGING

"It's now very difficult to get people to take part in a survey, much more so than in the past. Consequently if being paid per interview, the rate per hour is less."

"People are becoming untrusting- newer interviewers get disheartened."

"Working unsocial hours- especially in winter."

"Having to continually work evenings and weekends"

OVERCOMING OUR CHALLENGES

Response Rate Targets

▼

- Achievable targets impact on interviewer earnings and on retention
- With the correct resource levels we can deliver

Costs

▼

- Resourcing and maintaining a field force grows increasingly expensive
- Supporting this infrastructure requires investment

Awareness & Engagement

▼

- Raise awareness regarding the part the public plays in collection of this valuable data
- Focus on continued efforts, make more of the value of research
- Encourage commissioners to be champions of research

IN CONCLUSION

- As an industry we face challenges in attracting people to become interviewers

- Interviewer retention also requires continued focus, with increasing polarisation in experience levels making succession planning a challenge

- We <u>can</u> still build panels but it has become more difficult and costly to do so

- Raising awareness and encouraging promotion of research can help

- Continuing to invest in incentive strategies for target samples

- Response rate targets need to reflect underlying changes in society

- We plan to continue as individual companies and collectively to work to address this and to engage with interviewers

RESPONSE RATES AND POPULATION INFERENCE

Patten Smith (Ipsos MORI) & Joel Williams (Kantar Public)

18th May, 2017

THIS PRESENTATION

Aims

1. Examine relationship between survey *response rates* and *quality of population inferences*

2. Describe relevant empirical work

3. Draw conclusions for future survey practice

Outline

- Conceptual framework

- Empirical findings

- Practical implications for survey commissioners / practitioners

1

CONCEPTUAL
FRAMEWORK

DEFINING TERMS

Response rate

- AAPOR standard definitions: survey response rate is proportion of *eligible* population taking part

Quality of population inference

- Use Total Survey Error (TSE) framework - generally accepted framework for assessment of random probability survey quality

- Inventory of survey processes and errors identified with each

- Places errors on common metric: mean square error (MSE - error variance from all sources + square of bias (from all sources))

- Hard to measure TSE, but framework clarifies quality issues

SURVEY PROCESSES AND TSE
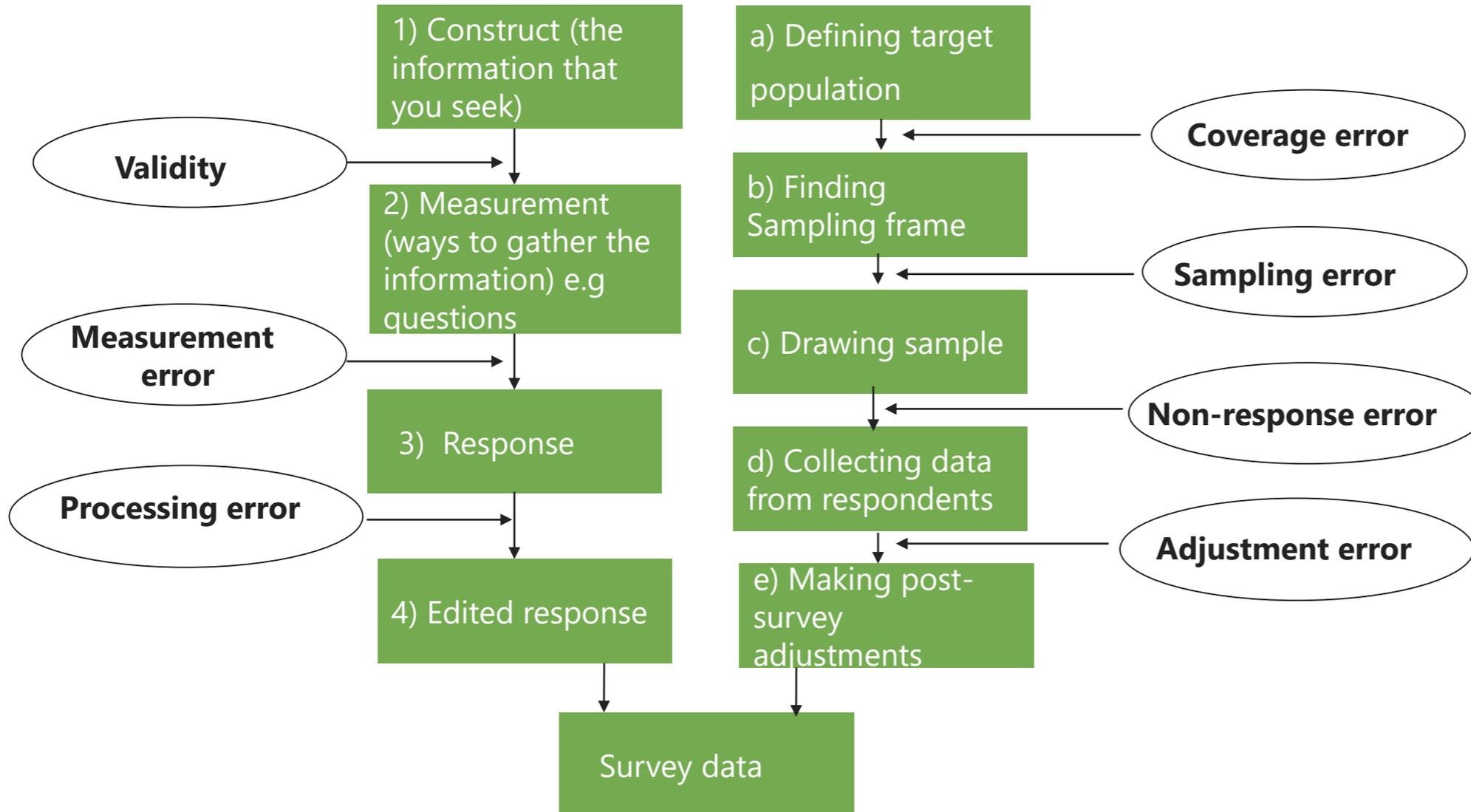
RESPONSE RATE AND NON-RESPONSE BIAS

Non-response bias: difference between estimates for survey respondents and overall population

Non-response bias (NR bias) for variable Y increases with:

1. correlation between Y and likelihood of responding
2. non-response rate

NR bias:

- is variable-specific; wide range of NR bias values in a single survey
- increases as response rate decreases *only if non-zero correlation between response rates and variables*
- but not by much unless correlation is substantial

RESPONSE RATE AND NON-RESPONSE BIAS

Expressed algebraically, (P=sample member's propensity to respond (0-1)):

$$\text{NR Bias}(\bar{Y}_R) \approx \frac{cov(PY)}{\bar{P}} = \frac{Corr(PY) \times SD(Y) \times SD(P)}{\bar{P}}$$

2 EMPIRICAL EXAMINATIONS OF RESPONSE RATE-TO-BIAS RELATIONSHIP

RESPONSE RATE AND NON-RESPONSE BIAS

Relationship between RR and NR bias is open; requires empirical assessment

Two types of empirical study depending on availability of validation data

1. Where good estimate of the "true" value of a variable, compare this with estimate for survey respondents; difference is estimate of **total NR bias**
   – Aggregate population estimates
   – Sample based estimates from "gold-standard" surveys
   – Sample frame information about issued cases (direct or linked)

2. Otherwise, assess how survey estimates change with increasing fieldwork effort (e.g. number of contact attempts, extent of reissuing); assumption - not a logical necessity! - that this will measure **relative NR bias**

TOTAL AND RELATIVE NON-RESPONSE BIAS

Ideally, measure total NR bias using benchmark validation data

Often measure *relative* NR bias because:

- – Surveys usually commissioned to study the un-benchmarked!
- – Benchmarks usually demographic / structural – nothing on opinions & very little on behaviour (voting excepted)
- – Estimating relative NR bias vs that obtained with a 'gold standard' RR (not 100%!) is a more useful way of judging VFM of each component of data collection effort

EMPIRICAL EVIDENCE

US: Groves et al (2006/08, US): meta-analysis of studies of total NR bias

UK:

- Census link study: total bias estimates (demographic only) for six surveys
- Hall et al (2013, UK) - relative bias and FW effort in three surveys
- Sturgis, Williams et al (2016, UK) - relative bias and FW effort in 541 non-demographic variables in six surveys
- Williams (2016) - relative bias and FW effort in CSEW subpopulations
- Unpublished Ipsos analysis of relative bias in 2009-10 Citizenship Survey

GROVES (2006 & 2008 W/PEYTCHEVA, US)

Meta-analyses (30 studies / 319 estimates; 59 studies / 959 estimates)

Measured absolute NR bias

Very low correlation between RR and NR bias

Greater variation *within* studies than between them

Preliminary efforts to identify survey design features that target NR bias
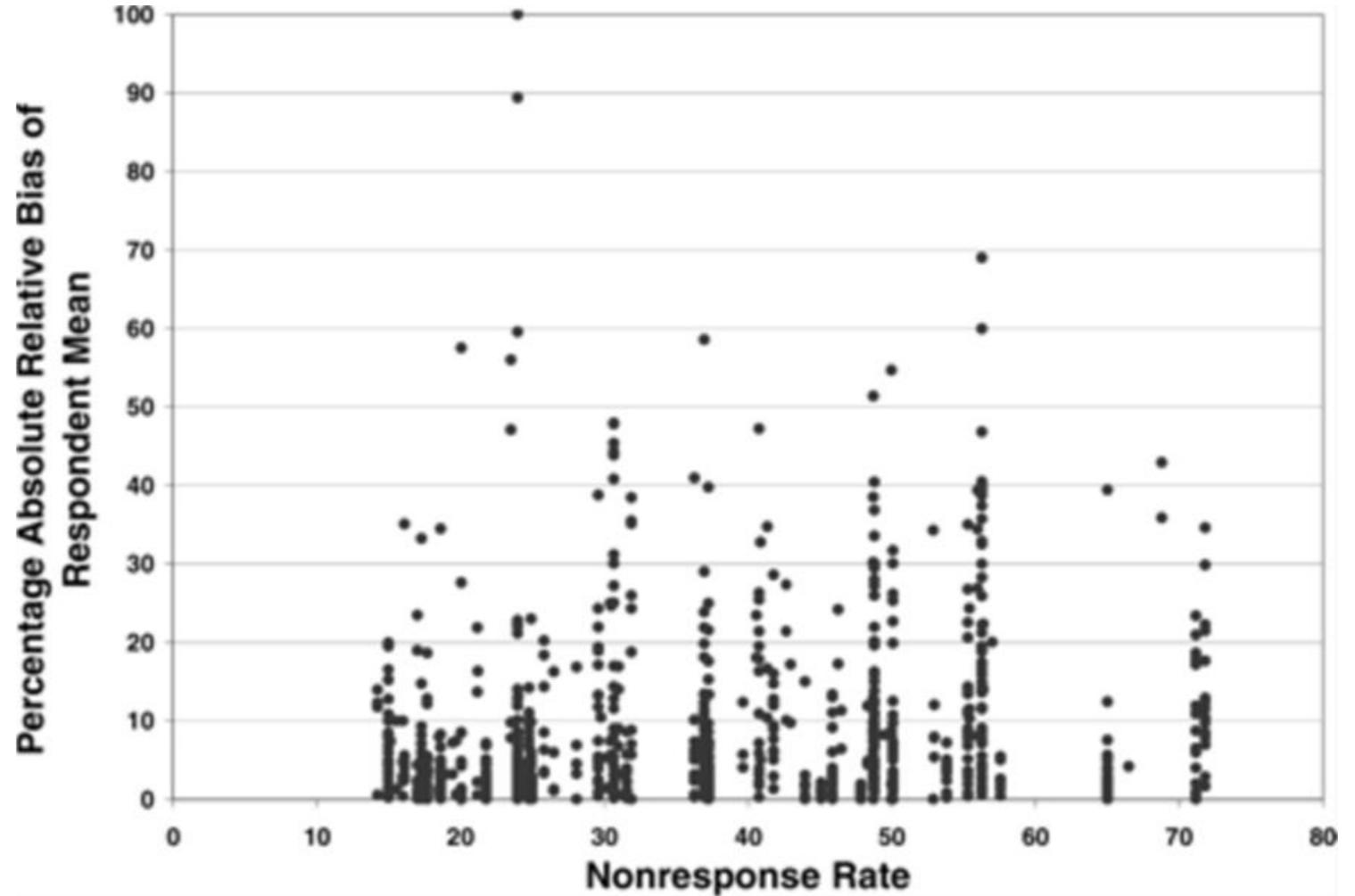
HALL, BROWN, NICOLAAS AND LYNN (2013)

Three NatCen Surveys: BSA, FRS and HSE (final RRs 51% - 61%)

Compared demographic and survey variables before and after extended FW efforts (6+ calls to get contact *or* reissued after initial refusal)

Almost no significant differences between calibration-weighted estimates before and after these efforts

(As noted earlier, substantial share of total effort allocated to this reissue phase)

STURGIS, WILLIAMS & BRUNTON-SMITH (2016, UK)

Six surveys from 2011-14 (all Kantar Public); all (541) non-demographic items

Relative bias (*distance from final estimate*) measured after each FW effort level:

- 1 call (RRs=7-22%), 2 calls, 3 calls, 5 calls (cf. final RRs = 55-76%)

- Average error after call 1 = 1.6%pts; 1.1%pts after calibration

- Average error after call 2 = 1.0% pts; calibration barely improves on this

Study estimated that a 4-call rule leads to lowest # of total calls (=lowest cost)

On average, questions about beliefs (and attitudes) tended to respond to FW effort slightly more than did questions about behaviour

Small number of variables with higher relative bias (internet use, freq. of being in during day, freq. of going to pub) – but convergence after 3 calls

IMPACT OF 1ST, 2ND, 3RD, 4TH/5TH CALLS
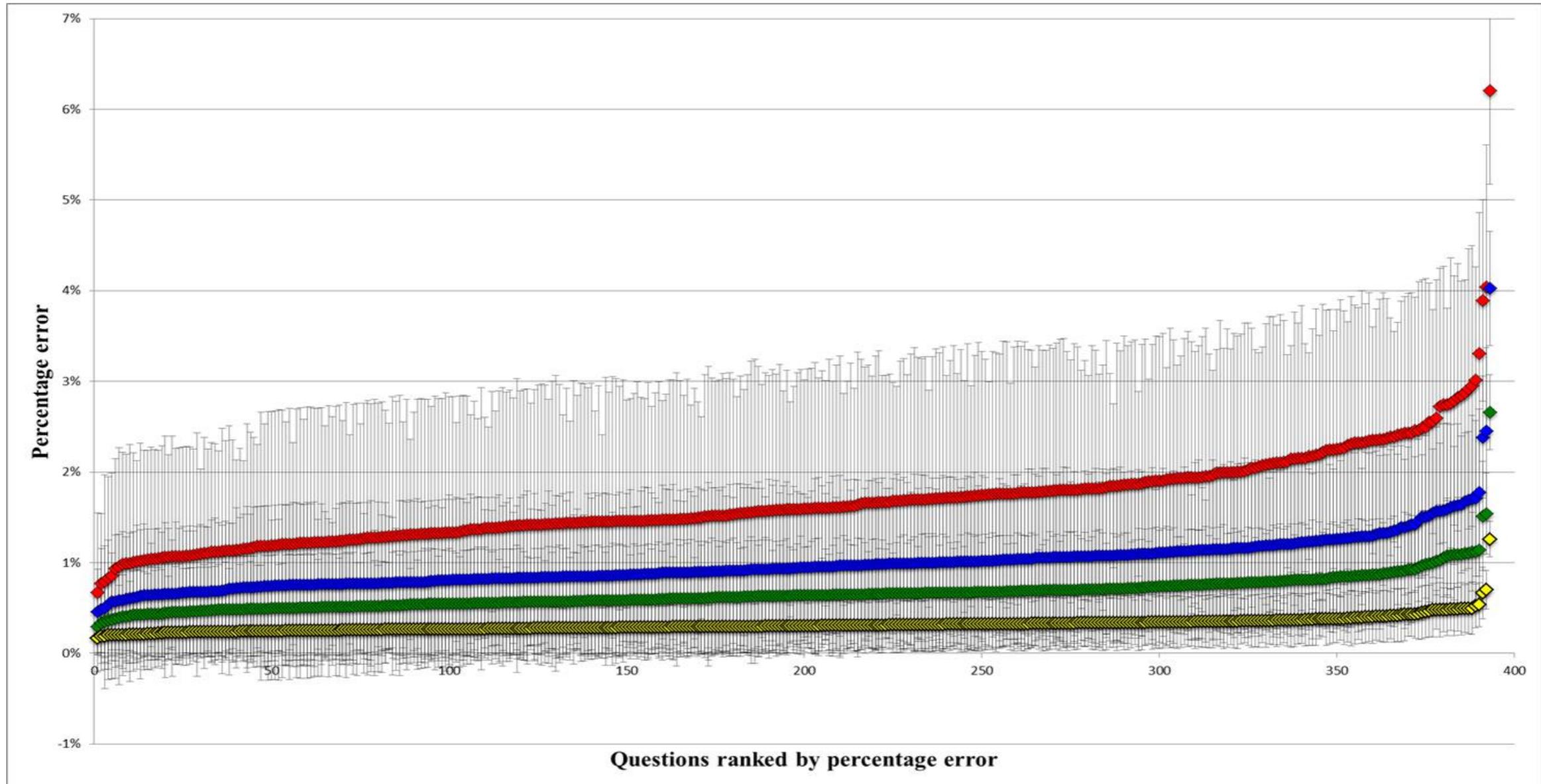
WILLIAMS (2016, UK)

Subpopulation analysis in CSEW (e.g. age group) – pre-reissue v final (both calibrated)

Compares error distribution v null hypothesis error distribution

Subpopulation error levels slightly smaller than total population error levels (greater within-group homogeneity?)
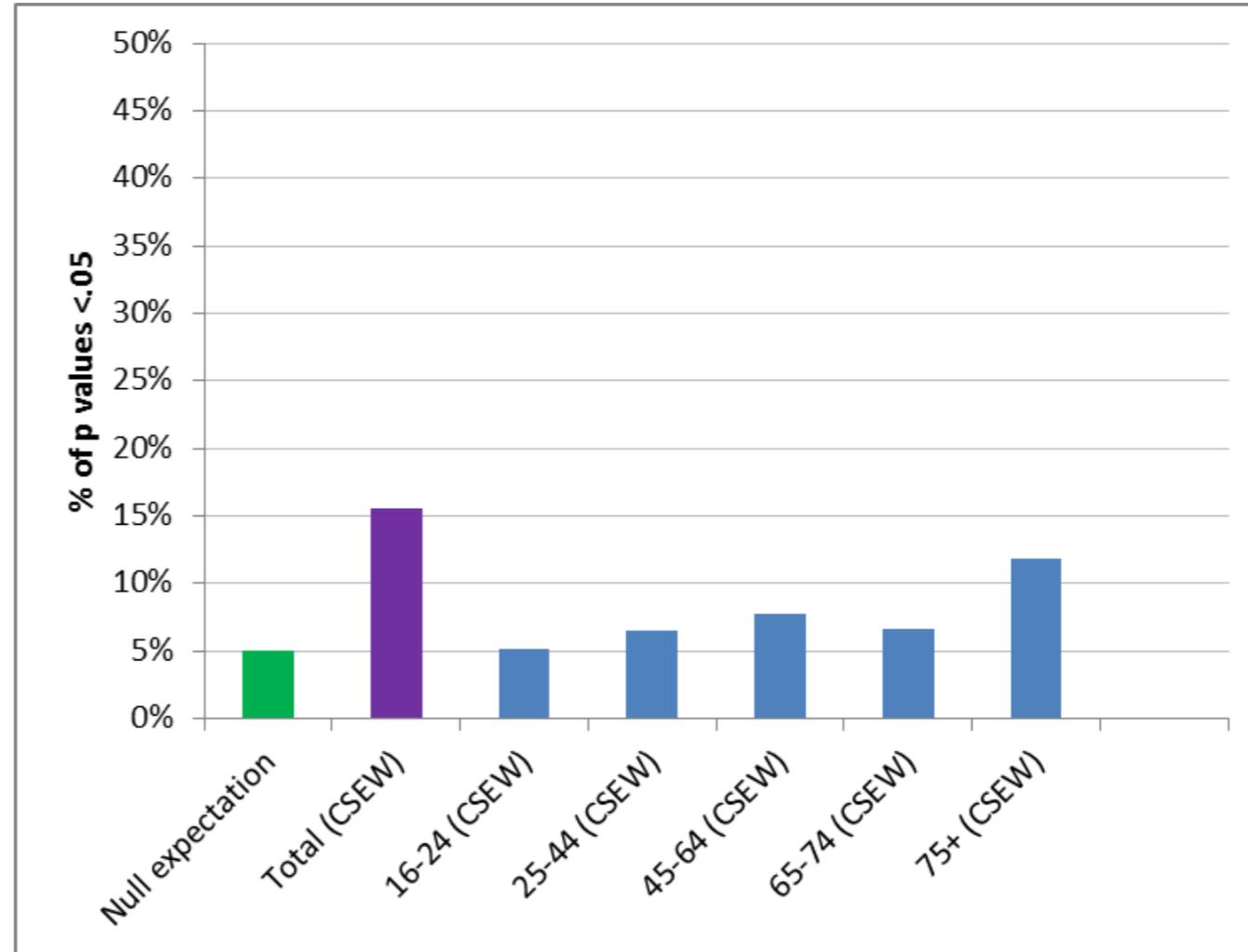
THE POTENTIAL FOR EXCEPTIONS

Occasionally extended FW efforts can reduce NR bias more substantially

2009-10 Citizenship survey reissues (50>>>56% RR)

Formal and informal volunteering prevalence decreased by 1% point after reissuing (CI: 0.5-1.5%)

Fits with US experience also (e.g. see Pew, 2012)
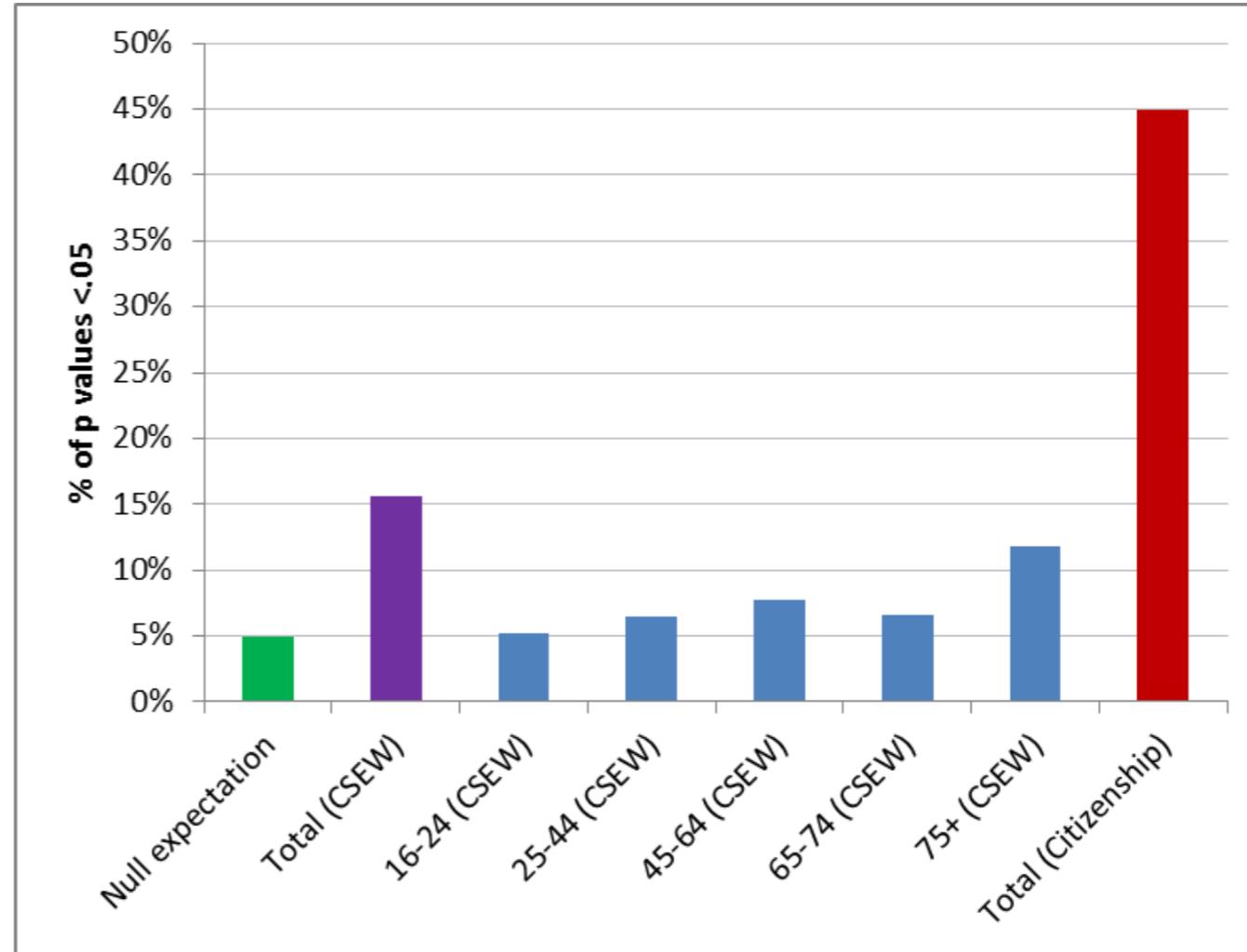
BROAD CONCLUSIONS FROM EMPIRICAL EVIDENCE

Absolute bias studies: relationship between RR and NR bias generally very weak

Relative bias studies: extended FW efforts increase RRs, but small impact on survey estimates

Lack of relative bias should not lead to assumption of *no* bias but should lead to questions over VFM of additional DC efforts

For some variables extended FW effort makes a difference – but hard to predict in advance

Most UK work on relative bias based on f2f interview data – unclear whether findings generalisable to data collection modes with lower 'maximum' RRs

3 IMPLICATIONS FOR SURVEY PRACTICE

QUESTIONS WE SHOULD ASK OURSELVES BEFORE SETTING RESPONSE RATE TARGETS

Question 1: is this a repeat wave of longitudinal survey?

Response rate maximisation often more important because:

- Cumulative attrition produces shrinkage in sample numbers in key groups
- Can lead to large losses in precision

QUESTIONS WE SHOULD ASK OURSELVES BEFORE SETTING RESPONSE RATE TARGETS

Question 2: how vulnerable to NR bias are your key variables?

Weighted survey estimates for *most* variables largely unaffected by response rate variations within observed ranges

Rational starting assumption that response rate levels will make little difference

For some variables declining RR has greater impact on NR bias: e.g. volunteering and web usage

Do you have any such variables? Check literature / earlier survey data

QUESTIONS WE SHOULD ASK OURSELVES BEFORE SETTING RESPONSE RATE TARGETS

Question 3: will marginal reductions in NR bias compromise your conclusions?

Even with highest achievable response rates, levels of *absolute* NR bias may still be substantial

Often trends of greater interest than point estimates. NR bias probably relatively constant over time - little impact on trend lines

QUESTIONS WE SHOULD ASK OURSELVES BEFORE SETTING RESPONSE RATE TARGETS

Question 4: are there better ways of reducing NR bias?

Blanket increase in RR beyond those obtained by reasonable good practice poor method for reducing NR bias

Address NR bias in other ways? –reduce correlation between key variables and response propensity?

Increase FW effort for poor responders differing on key variables using supplementary variables / paradata to identify them

OVERALL CONCLUSION

Observed impact of response rates on survey estimates usually modest

Pursuit of highest possible response rates may not be best use of limited survey resources – may be better ways of addressing survey errors

At times making considerable efforts to maximise response rates *can* be justified:

- Waves 2 – *n* in longitudinal surveys;
- where minimising NR bias critical to conclusions

IMPLICATIONS AND RECOMMENDATIONS

1 | **Issues from the previous sessions**

2 | **Current trends and observations**

3 | **Initial thoughts for discussion**

1 ISSUES FROM THE PREVIOUS SESSIONS

ISSUES FROM THE PREVIOUS SESSIONS

Face to face survey responses are declining

- **Secular trend arising from social changes e.g. declining privacy, trust, social capital, participant time**

- **Strains in the labour market e.g. interviewers harder to recruit and retain**

- **Importance of understanding methodology -- response rates vs. bias in affecting descriptive estimates and generalising inference**

  - **Requires understanding and EMPIRICAL data**

  - **Choices in use of resources**

  - **Better informed design**

2 CURRENT TRENDS AND OBSERVATIONS

CURRENT TRENDS AND OBSERVATIONS: I

1. Increasing use of other data sources e.g. admin data
2. Primary data collection still needed for DEPTH of information
3. Different methodologies face different challenges for descriptive accuracy and generalisability: focus here is on FACE TO FACE
4. Accept competition in the research supplier market
5. Social research projects do not yield large profits
6. SHARED interests in robust quality, design, and maintenance of capacity

CURRENT TRENDS AND OBSERVATIONS: II

7. **Current tendering often results in specific research design BEFORE data collection expertise/ experience brought in**

    ▪ **Leads to less than optimal designs, lack of empirical pilots, dilemma of non-compliant bids, lack of specification of key AIMS of research**

8. **Importance of linking AIMS with value for money in minimising total survey error**

9. **Requires move away from viewing response rates as sole (contractual) measure of data quality**

    ▪ **Alternative approaches in tenders, procurement**

    ▪ **Alternative approaches to research design**

3 SOME INITIAL THOUGHTS FOR DISCUSSION

DISCUSSION POINTS: SPECIFIC STUDIES

1. **Promote earlier expert engagement with research design**

   ➤ **Identify key variables and aims**

   ➤ **Commission / tender for work on research design – BEFORE main stage design fixed**

   ➤ **Use existing empirical data to consider total survey error**

   – **What additional data – admin. data, paradata, etc. – available to test impact of non-response?**

   ➤ **Pilot different designs if helpful**

   ➤ **Will affect timetable and budget allocation, possibly budget size, but with better outcome**

DISCUSSION POINTS: GENERAL & STRATEGIC

2. Strategic engagement with data regulators (ICO) about data linkage, use of department data sets, consents, opt-ins/outs (stressing data protection but also consent protections)

3. Strategic engagement of commissioners, data collection experts and procurement to move away from specification of response rates and focus on agreed PROCESSES to address non-response bias (and other sources of error)

- Better methodologically

- Better allocation of resources

- Better ethically / 'politically'

4 OVER TO WIDER DISCUSSION