

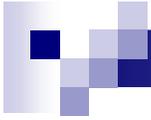
(maintaining quality in)
social research in
straightened times

Patrick Sturgis
University of Southampton,
National Centre for Research Methods



Outline

- Inevitability of severely constrained funding in years ahead
- Pressure to reduce costs of research, particularly data collection
- Risk of ‘bad driving out the good’
- Example: opt-in internet panels
- Conclusion: tempting but very risky



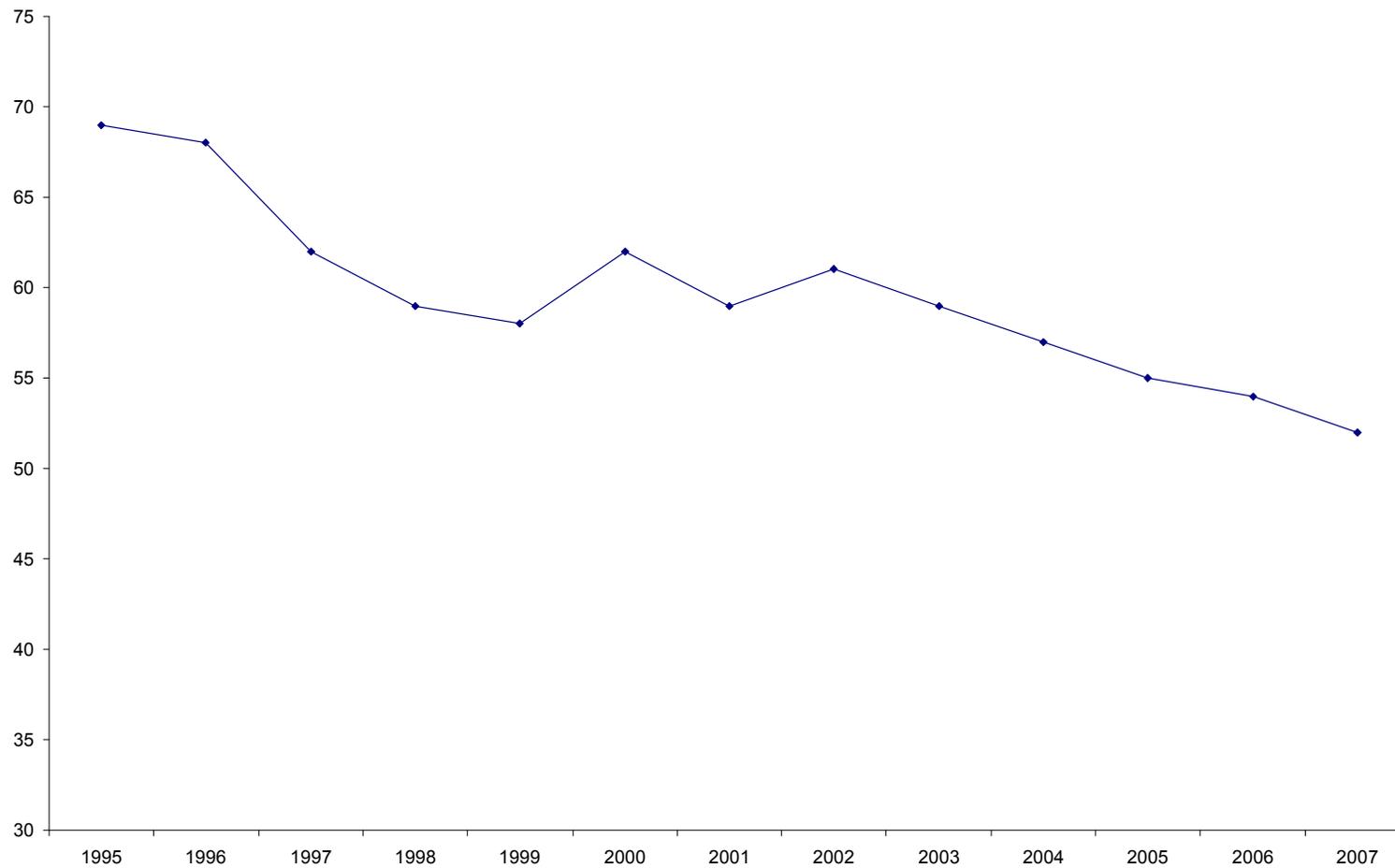
The context



Challenges to random sampling

- Changing lifestyles and demographics
- People are harder to contact and harder to persuade to take part
- Face-to-face costs high and rising
- Slow turnaround
- RDD surveys ~ 5% response rate?
- Increase in 'mobile-only' households

British Social Attitudes Survey response rate 1995-2007





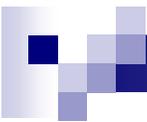
Internet surveys

- The ‘big story’ in survey research over the past decade
- US market share = \$2 billion in 2009
- Fast data turnaround
- Can target ‘niche’ groups
- Powerful integrative IT capability
- And cheap!



Types of internet survey

- Probability-based panels
 - Selected by RDD or face-to-face survey with subsequent questionnaires administered online
 - Online access provided by survey organisation as incentive to participate
- 'Opt-in' non-probability based panels
 - Panel members recruited via online banners, pop-ups, referrals etc.
 - Rewarded with points/prizes/cash for each survey completed



10 standard questions, general population omnibus survey

- NatCen = £22,500
 - Achieved sample size = ~1600
 - Results delivered ~ 11 weeks after fieldwork starts
 - Response rate = 55%
- YouGov = £3,200
 - Achieved sample size = ~2000
 - Results delivered within 2 working days of the survey going live
 - Response rate = ?



But...

- 30% of UK households do not have internet access (ONS, 2009)
- Those who are online are quite different from those who are not
- Those who agree to do online surveys are even more different from everyone else
- Because most become panelists for remuneration, there are major problems of measurement accuracy (satisficing, flat-lining)



Where do panelists come from?

- People can join panels at the company website
- And are 'harvested' from a wide range of different online environments (pop-ups, registration lists, etc.)
- Incentivised by points awarded for survey completion (redeemable as cash/vouchers)
- Need to do ~ 50+ surveys to get £50



media & commercial uses

- The majority (85%) of English football fans believe that England will reach the quarter finals or beyond in the 2010 football World Cup
- 68% agreed that Cheryl Cole was right to divorce Ashley Cole
- 58% of the British public would watch free-to-air TV on their mobiles



Appropriate topics/applications?

- For the majority of Brits (61%) international aid is the number one area that should be targeted for cuts
- An overwhelming 72% of British adult males would support the government if it were to send Navy ships to the Falklands to protect its mining interests
- Opt-in panels are used by a number of academic research projects funded by the ESRC to study public opinion and electoral behaviour



Does it matter?

- Evidence of validity mostly based on high-profile 'successes'
- E.g. predicting outcome of elections or X-factor winners
- But this is an empirical not a theoretical argument
- And this cuts both ways - increasing evidence that online polls often get things 'wrong' too



What is the theory?

- Draw random sample from within quotas set on known population marginals
- Samples then match population on e.g. age, sex, employment status, newspaper readership, attitudes etc.
- But this approach rests on a big assumption:
 - *There is no difference within quota classes on the variable of interest*
 - *e.g. women on an internet panel are no different in terms of their political outlook than women who are not on an internet panel*



American Association of Public Opinion Research (AAPOR) Report on On-line panels (2010)

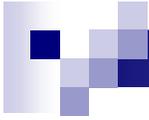
*“researchers should avoid nonprobability
online panels when one of the research
objectives is to accurately estimate
population values”*

Available free at: <http://www.aapor.org>



Concluding remarks

- As research funding becomes more constrained, pressure will grow to reduce data collection costs
- In responding to these challenges, it is essential to maintain basic quality standards
- Poor quality data is generally worse than no data at all
- Opt-in internet panels have their uses
- But these do not currently extend to social research where population inference is the goal



Don't waste clean thoughts on dirty data!