2013 PAPOR Mini-Conference

Non-Probability Based Samples

and

AAPOR Update

Sarah Cho

Kaiser Family Foundation

June 2013

THE HENRY J.
KAISER
FAMILY
FOUNDATION

2013 PAPOR Mini-Conference

Disclaimer: Slides were obtained with author permission, but any mistakes/misinterpretations are my own!

PewResearchCenter

A Comparison of Results from Dual Frame RDD Telephone Surveys and Google Consumer Surveys

Scott Keeter, Rob Suls, Danielle Gewurz, Michael Dimock, *Pew Research Center*
Leah Christian, *Nielsen*
Jon Sadow, Brett Slatkin, Paul McDonald, Matt Mohebbi, *Google*

skeeter@pewresearch.org

What is a Google Consumer Survey?

- Short web survey (max 2 questions per respondent). Can be filtered to select certain kinds of people
- Various types of questions – single answer, multiple answer, open-end, use of images, etc.
- Web interface to display results and download data
- Nonprobability sample – people are sampled from online publisher websites who have agreed to allow Google to place surveys on their sites (a "survey wall")
- Quota sampling – people sampled based on their gender, age and location
- Sampling and weighting based on Google's inferred demographics

Interests

Banking, and 14 more. Edit
From your previous searches

We use interests from your activity on Google

We use interests from your activity on websites to tailor ads to you. How it works

ADD

Interest

Banking

Bicycles & Accessories

Celebrities & Entertainment News

Consumer Electronics

Cooking & Recipes

East Asian Music

Fitness

Hair Care

Hygiene & Toiletries

Make-Up & Cosmetics

Mobile Phones

Online Video

Rap & Hip-Hop

Search Engine Optimization & Marketing

Social Networks

Interest

Air Travel

Apparel

Arts & Entertainment

Baked Goods

Cards & Greetings

Celebrities & Entertainment News

Computers & Electronics

Cooking & Recipes

Cycling

Fitness

Food & Drink

Golf

Kia

Music & Audio

News

Outdoors

Politics

Programming

Real Estate

Real Estate Listings

Restaurants

Running & Walking

San Francisco Bay Area
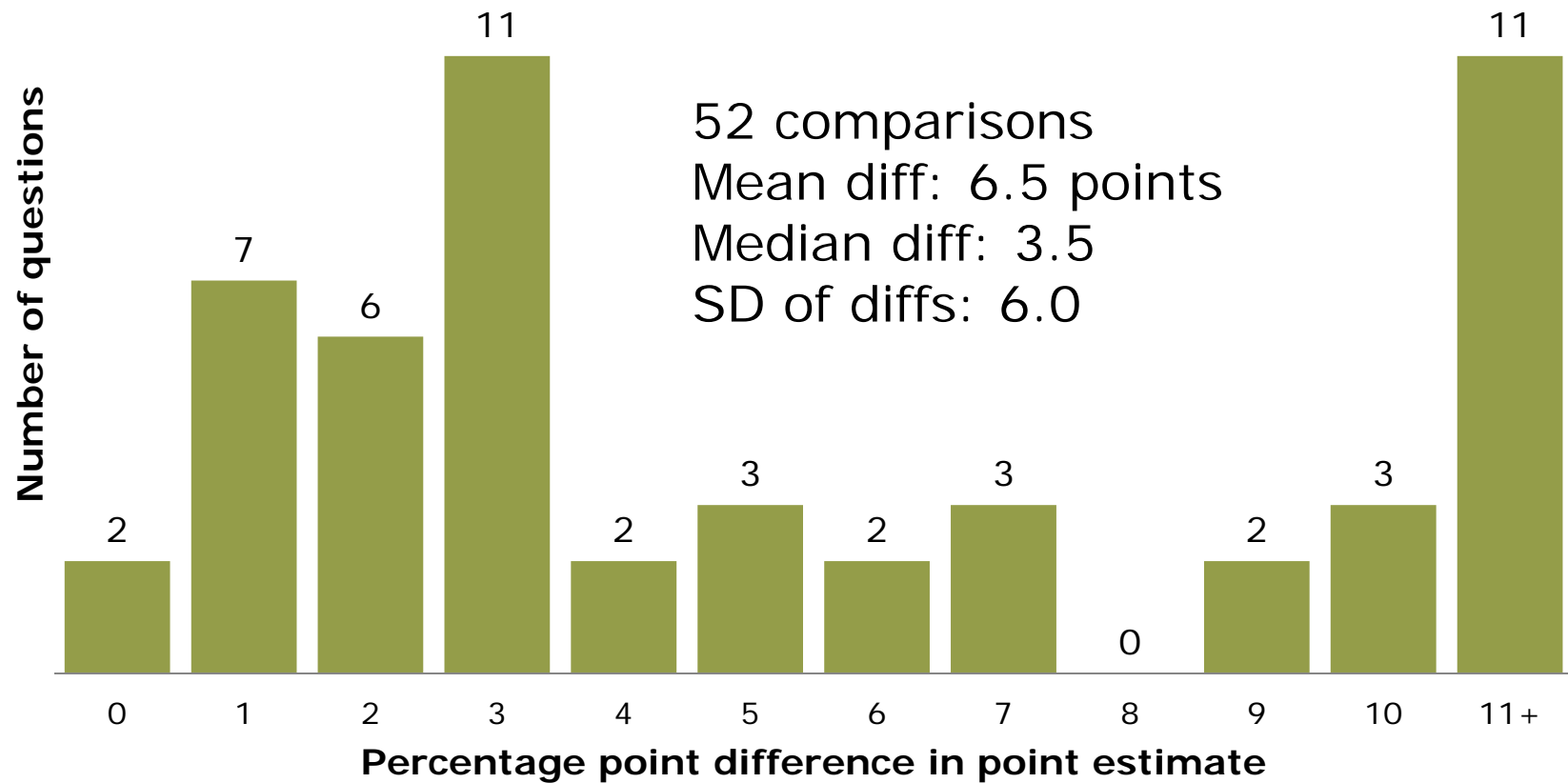
Social Networks

Sports

TV & Video

Travel

Fit for Purpose

Different researchers will have different needs and standards. Any method will fit some purposes better than others.

- National point estimates
- Associations between variables
- Tracking change over time
- Quick reaction measurement
- Pretesting question wording
- Open-end testing
- Diverse question formats

National Point Estimates: Summary of Differences in Point Estimates, Phone vs. Google

52 comparisons
Mean diff: 6.5 points
Median diff: 3.5
SD of diffs: 6.0

Number of questions

Percentage point difference in point estimate

National Point Estimates: Composition of Phone & Google Samples

	Phone (internet users)	Google
White, non-Hispanic	69	68
Black, non-Hispanic	11	10
Hispanic	13	10
Other	7	12
Male	49	53
Female	51	47

National Point Estimates: Composition of Phone & Google Samples

	Phone (internet users)	Google
18-24	16	9
25-34	20	19
35-44	19	17
45-64	34	38
65+	12	16
College graduate	32	45
Some college	30	29
High school or less	37	26

National Point Estimates: Point Estimate Comparisons

	Phone	Google
Republican party ID *(May, April 2013)*	25	27
Democratic party ID	32	31
Independent/other	43	42
"Always" vote *(June '12, April '13)*	50	46
Conservative *(Sept '12, Aug '12)*	37	40
Moderate	35	36
Liberal	23	24

National Point Estimates: Point Estimate Comparisons

	Phone	Google
Get tougher with China *(Oct 2012)*	49	54
Build stronger relationship w/China	42	46
Attend religious services more than once a week *(May '12, Aug '12)*	12	12
Once or twice a month	26	23
A few times a year	18	14
Seldom/Never	30	43

Associations Between Variables: Correlations by Age

% Prefer smaller gov't

Google

43 51 60 61 61

Phone

34 48 54 55 55

18-24 25-34 35-44 45-54 55+

% Who always vote

Phone

61 56

48

Google

41 42

23 31 33 38

21

18-24 25-34 35-44 45-54 55+
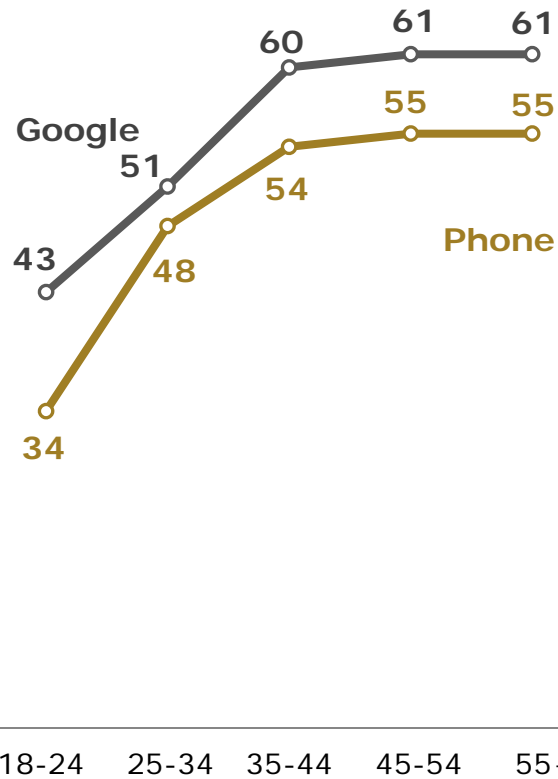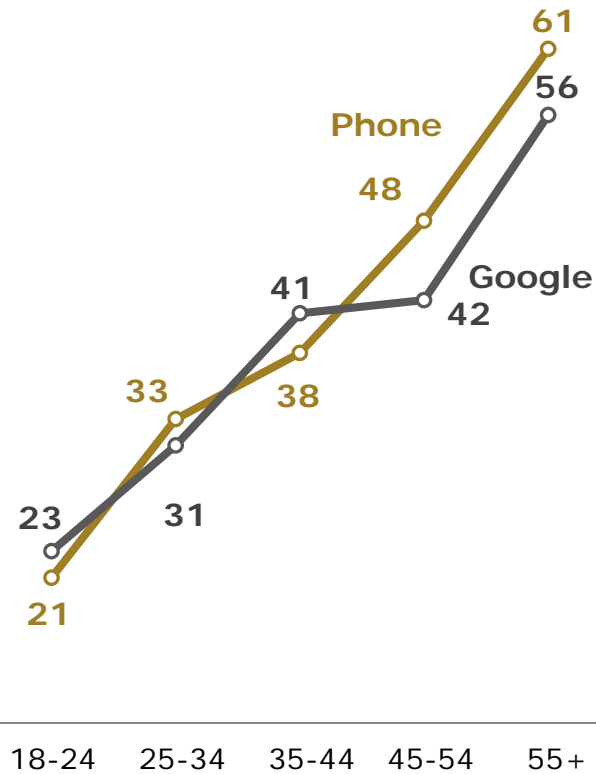
Google demographics based on inferred information.

Reliability of Selected Estimates

% Always/nearly always vote

67 68 72

% Prefer smaller gov't

60 59 69

% Earth is warming

50 52 54

% Oppose same sex marriage

38 40 40

% Republican

29 27 30

Apr 15　　　　　Apr 30　　　　　May 15
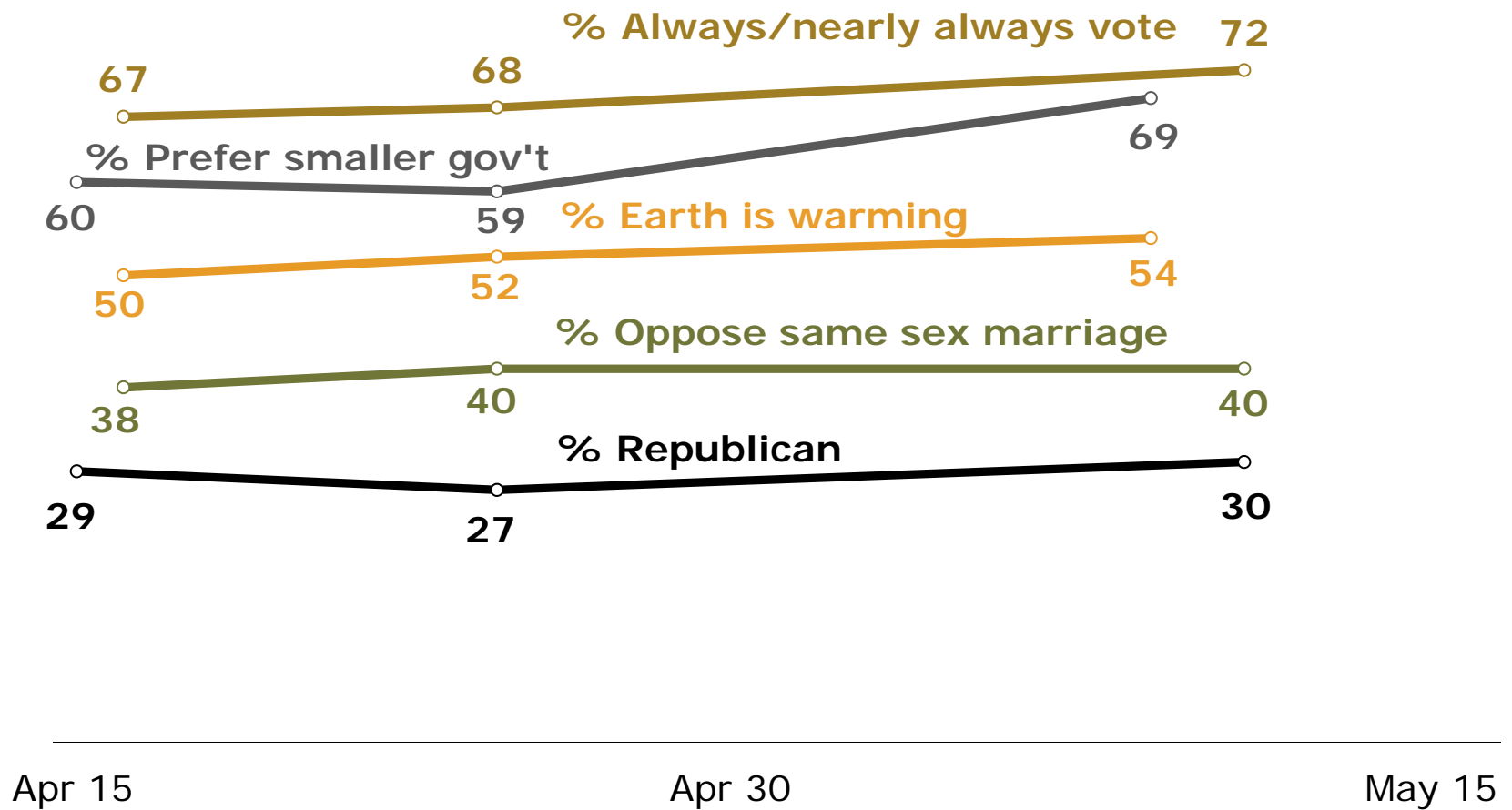
Quick Reaction Surveys

Which candidate did the better job?	-------Google-------			Phone
	Night of	Next day	Wkend	Phone
First presidential debate				
Obama	32	16	16	20
Romney	44	59	57	72
Same (Both, Neither, DK vol.)	24	25	27	(7)
Second presidential debate				
Obama	48	50	50	48
Romney	33	32	32	37
Same (Both, Neither, DK vol.)	20	19	18	(15)

Asked of registered voters who watched each debate. Pew Research survey and Google surveys conducted Oct 2012

Question Wording Experiment

The best way to ensure peace is through military strength	Phone	Google
Agree	53	55
Disagree	42	45
Don't know	4	-
The best way to ensure peace is through military strength	31	33
Good diplomacy is the best way to ensure peace	58	67
Don't know	11	-

Pew Research survey conducted among general public March 2011 / Google survey conducted June 2012

Open Ended Question Testing

What ONE subject should schools emphasize more than they do now?	Phone	Google
Math, mathematics, arithmetic	30	27
English, grammar, writing, reading	19	18
Science	11	8
History, social studies, civics, govt.	10	10
Art, arts, music	6	7
Computers, computer science	4	3
Physical education, health, sex ed.	2	3

Pew Research survey conducted among general public March 2013 / Google survey conducted March 2013

Conclusions

- Google Consumer Surveys produces results quickly, cheaply and timely (for specific times/days/events, etc.)
- Allows for the use of multiple question types
- But because of the reliance on nonprobability sampling it is difficult to predict when it works well and when it doesn't
- Google Consumer Surveys continues to evolve – evaluating asking more questions, adding more publishers, and testing new question types

Conclusions

- **Pew Research plans to continue to use Google Consumer Surveys for quick reaction polls, for testing of survey questions – including question wording, order and format as well as testing open-ended questions to help inform development of closed-ended questions**

- **We are interested in exploring how well it can measure media use at various times of day**

- **We hope to other explore types of non-probability methods to see how they might supplement our traditional probability based surveys**

Probability vs non-probability samples.
Is Accuracy only for Probability Samples?

Johan Martinsson, Stefan Dahlberg and Sebastian Lundmark

Department of Political Science

University of Gothenburg

Surveys from online panels

Survey company	Mode	Sampling method	Participation rate
Novus	Web panel	Prob. based recruitment	59
TNS Sifo	Web panel	Prob. based recruitment	38
YouGov	Web panel	Self-recruitment	40
Cint	Web panel	Self-recruitment (85%)	24

Cross-sectional surveys with different modes

Survey company	Mode	Sampling method	Response rate
SOM institute	Mail	Random population sample	53
Detector	Telephone	Random population sample	51
LORe	Web	Random population sample	8

Comparability of surveys

- a set of identical questions were included

- approximatley same period of field work, except for the SOM-institute, which was conducted a few months later

- however, field work length and nr of reminders differ

- we focus on basic demographics and political attitudes

- for demographics, we use census data from Statistics Sweden as benchmark

- target population: 18-70 yrs old, in the Gothenburg region (west sweden, approx. 1 million inhabitants)

Demographics: average absolute deviation from Statistics Sweden (unweighted estimates)

	Mail (SOM)	Phone (Detector)	Web (LORE)
Average 5 indicators	4.3	3.8	6.9

average abs deviations, demographics
(weighted estimates)
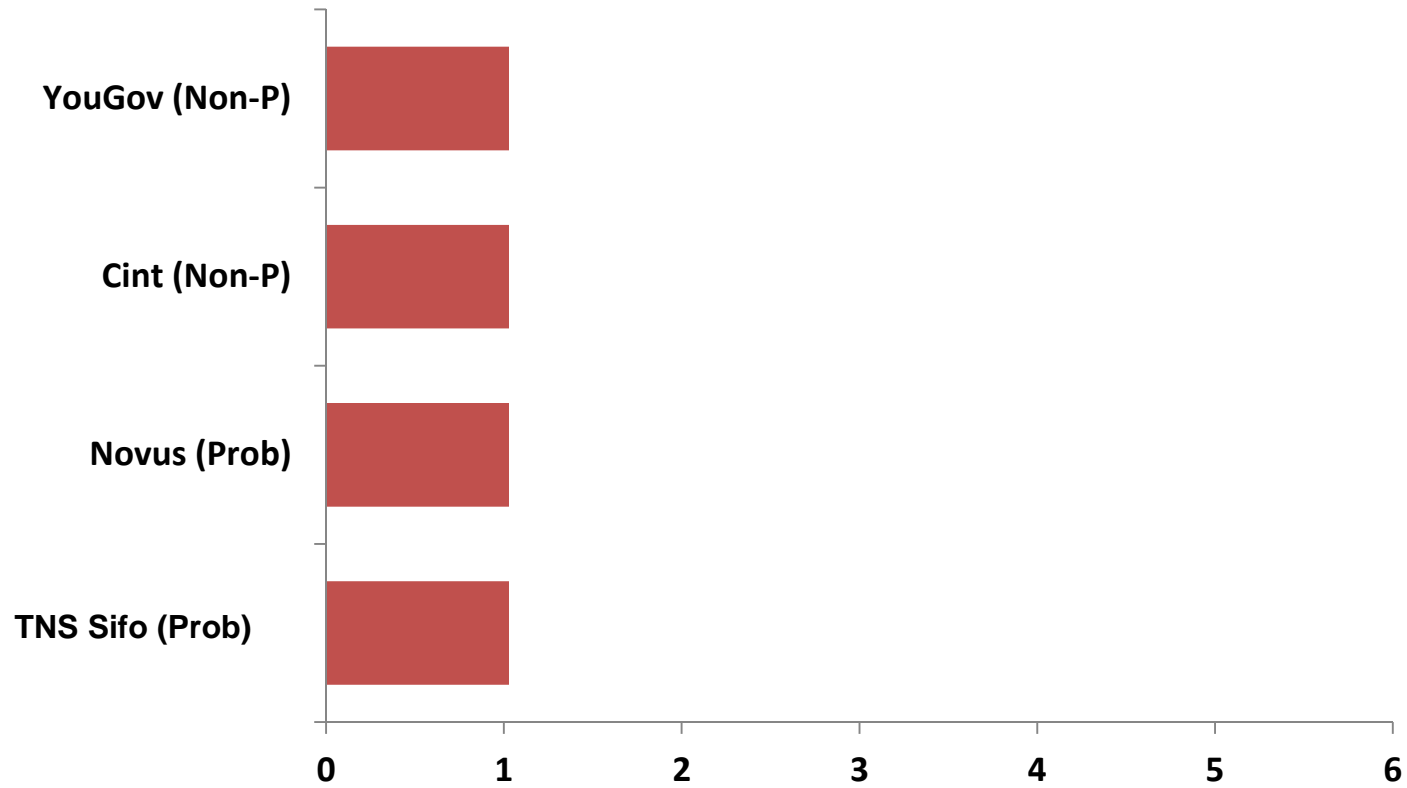
average abs deviations, demographics (unweighted estimates)

YouGov (Non-P): 2.9
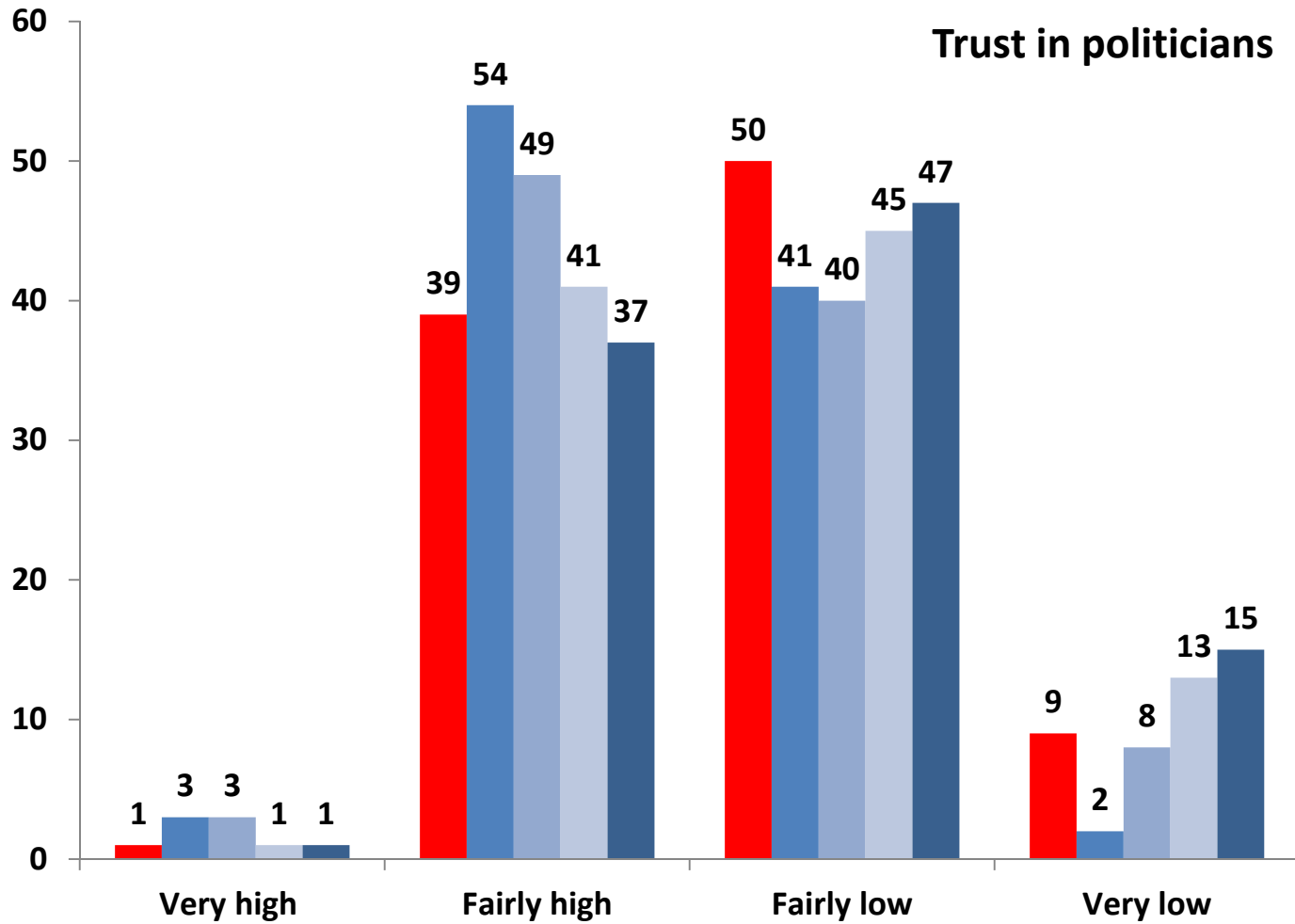Cint (Non-P): 3.1
Novus (Prob): 3.7
Sifo (Prob): 7
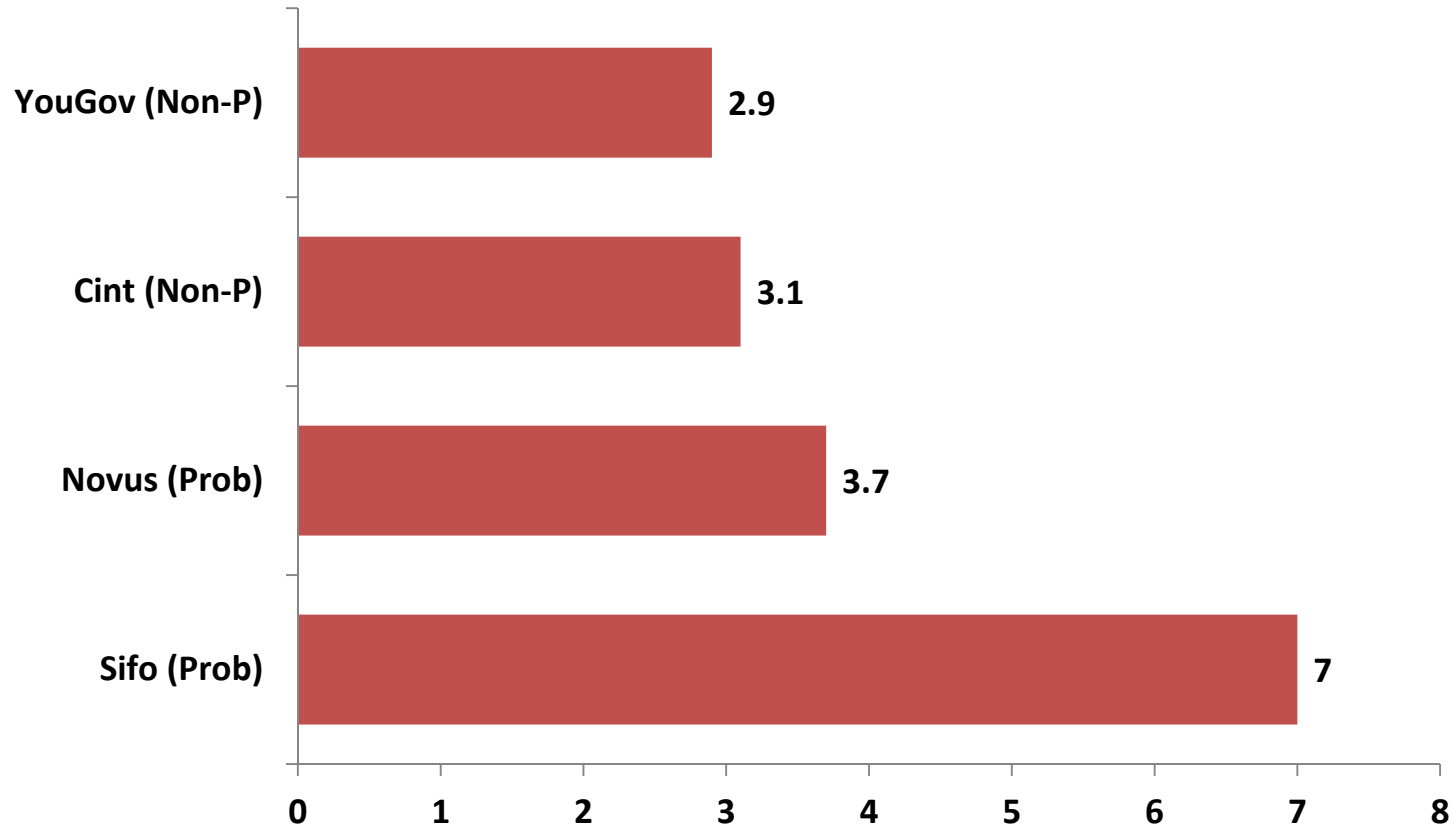
Political attitudes

- for attitudes and opinions, there is no true benchmark
- however, as second best option, we use the mail survey as quasi-benchmark
- why? :
  - well known and high quality survey
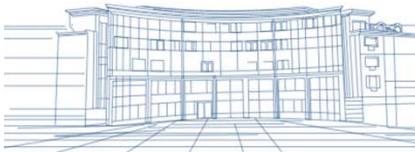 - excellent sampling frame and high response rates
  - mode (mail) most similar to web surveys
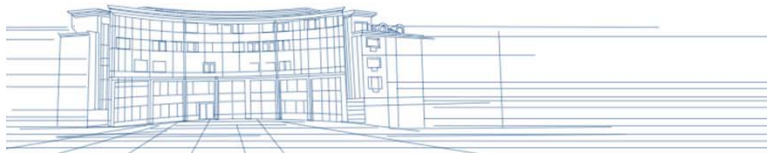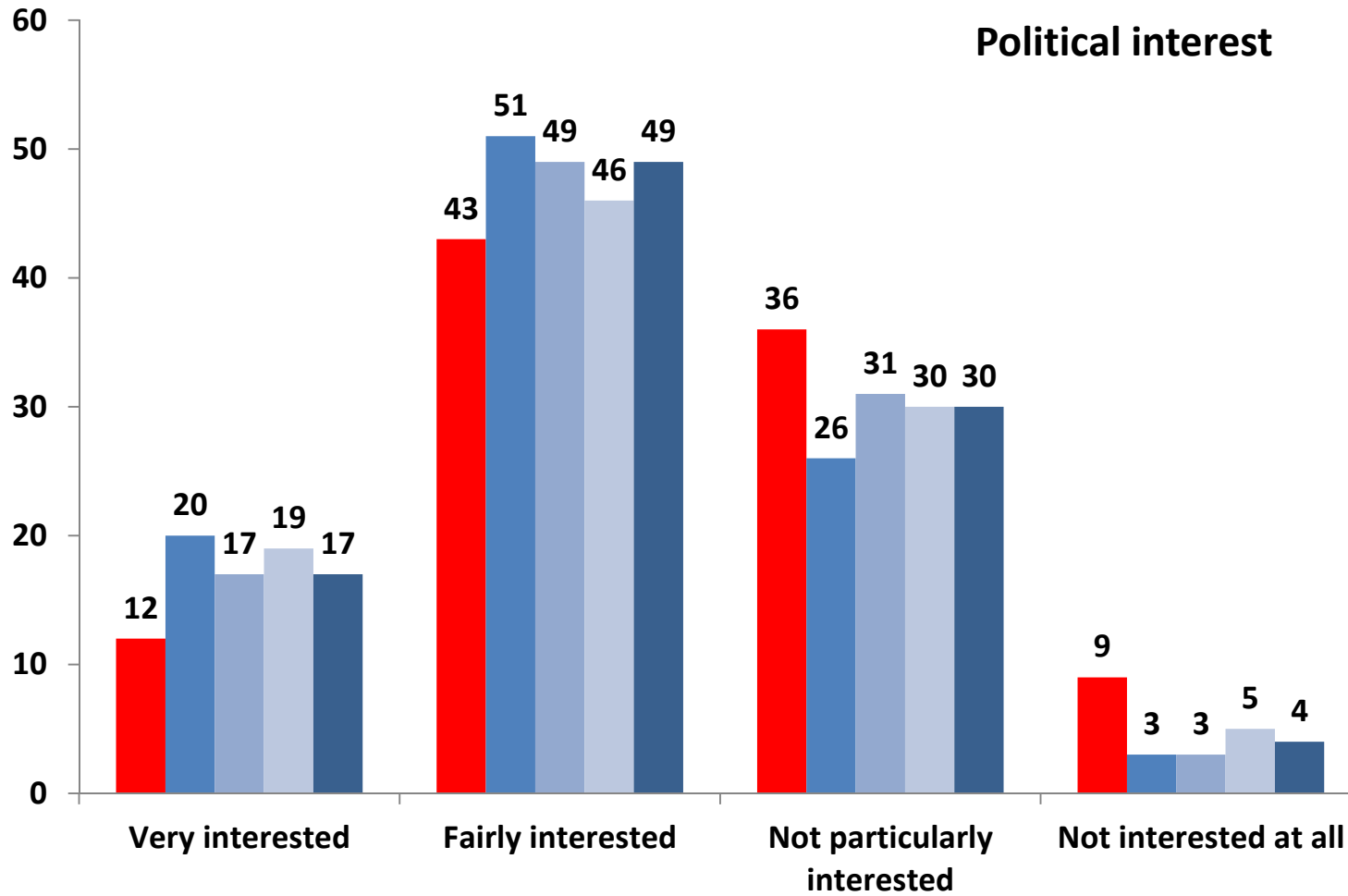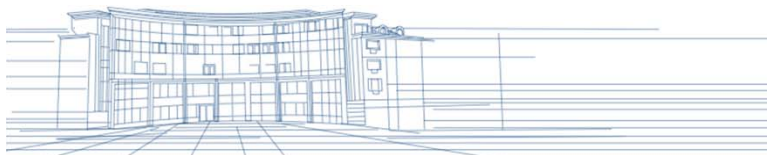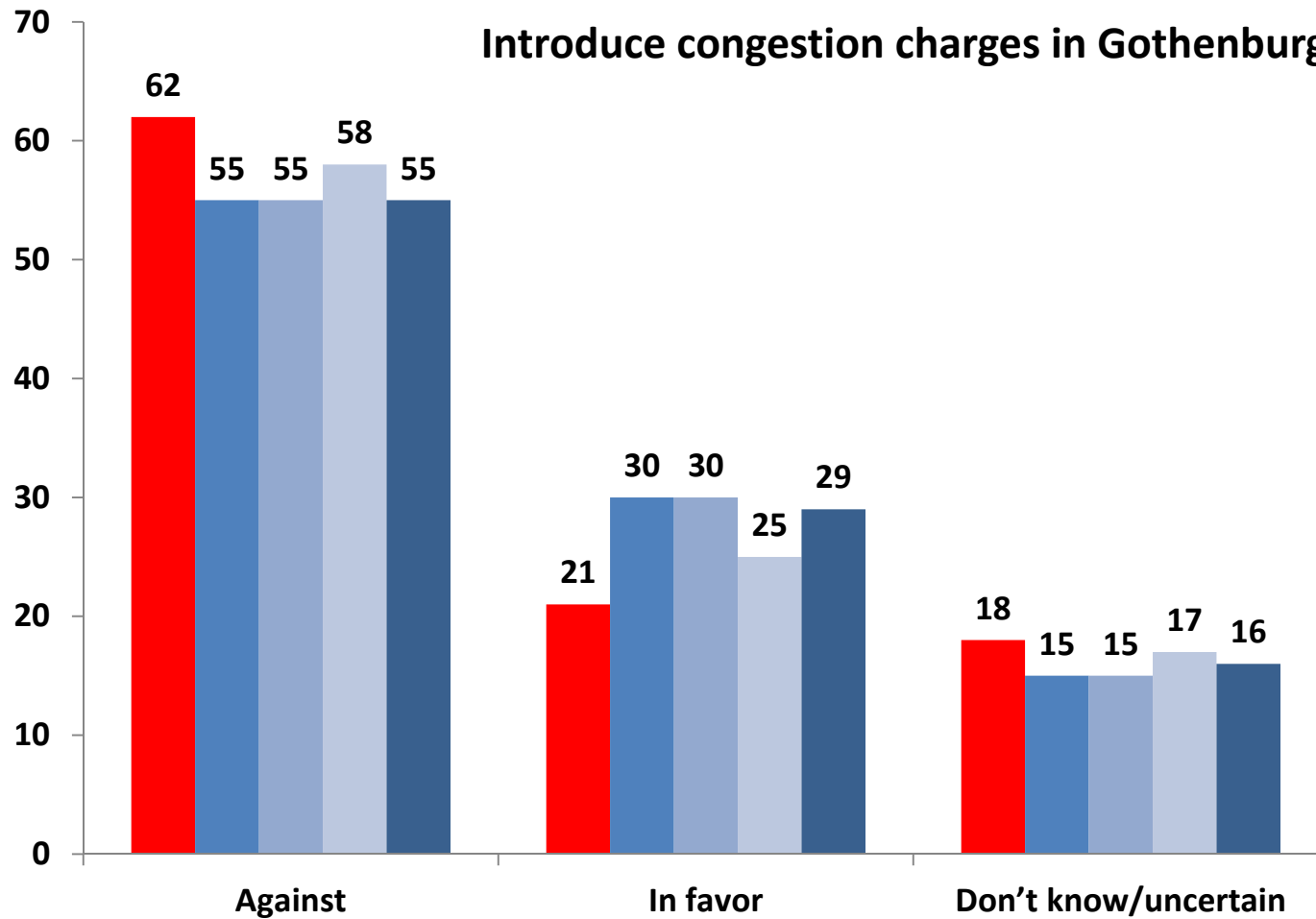
Conclusions

- surprisingly, the demographic accuracy of the non-probability based panels are better

- compared to a benchmark mail survey, the non-probability panels also came closer to political attitudes

- in this comparison, we find no evidence that self-recruited on-line panels have less accuracy than probability based on-line panels

Discussion

- Too much uncertainty about this result

- we would need more demographic indicators than 4-5!

- Sweden also has an extremely high internet coverage

- The self-recruited panels seem to attract more people with low SES, to their advantage

- The probability-based panels are not revealing enough about how succesful their recruitment is, are they really high quality probability based panels?

Statistical Adjustments for Internet Opt-in Panel Surveys

Sunghee Lee, University of Michigan

Catherine Okoro, CDC

Satvinder Dhingra, CDC

Data

	Probability sample	Web opt-in panel sample

Study 1 - National	2010 National Health Interview Survey (NHIS), n=27157
	
	• Area probability
	• Face to face
	• 8 Subjective Quality of Life Qs
	   • SRH apart from the rest
	   • End of cancer module
		2012 Well-Being Study (WBS), n=3948
		
		• Sample matching & weighting
		• Conducted only in English
		• 8 Subjective Quality of Life Qs
		   • Within various subjective quality of life measures

Study 2 - Selected states	2011 Behavioral Risk Factor Surveillance System (B-RDD), n=38,143 • Dual-frame RDD • GA, IL, NY, TX • Subjective and objective Qs	2013 BRFSS State Pilot Study (B-Web), n=4,000 • Sample matching & weighting • GA, IL, NY, TX • Conducted only in English • Subjective and objective Qs

Study 2 – Demographic variables

	B-RDD (n=38143)	B-Web (n=4000)
Male	48.7	47.4
18-34 yrs	32.1	31.8
35-49 yrs	**26.9**	**24.0**
50-64 yrs	**24.8**	**28.2**
65+ yrs	16.2	16.0
Hispanic	**20.6**	**18.0**
NH White	56.8	59.0
NH Other	22.6	23.1
Married	**48.7**	**45.8**
Children in hhld	**40.1**	**35.7**

	B-RDD (n=38143)	B-Web (n=4000)
<High school	**17.1**	**7.6**
High school	**27.9**	**32.8**
Some College	**28.7**	**35.2**
College or more	26.3	24.4
<$20K income	21.6	20.7
$20-50K income	**31.2**	**28.7**
>$50K income	38.8	41.1
Employed	**55.4**	**47.9**
Own home	**62.9**	**53.5**

Bold font indicates statistical significance (*p*<0.05).

Study 2 – Demographic variables

	Non-Hispanic Whites		Hispanics		Non-Hispanic Other	
	B-RDD (n=25936)	B-Web (n=2550)	B-RDD (n=5151)	B-Web (n=573)	B-RDD (n=7056)	B-Web (n=877)
18-34 yrs	26.3	24.2	43.7	49.2	36.2	37.7
35-49 yrs	25.0	23.9	**30.4**	**22.7**	28.6	25.4
50-64 yrs	28.1	30.4	17.5	22.4	23.0	27.3
65+ yrs	20.6	21.5	8.4	5.8	12.2	9.6
Children in hhld	**33.0**	**29.8**	**58.0**	**44.9**	41.6	43.6
<$20K income	13.5	15.8	36.4	30.4	28.0	25.7
$20-50K income	30.1	27.3	33.2	27.4	32.0	33.1
>$50K income	49.9	47.3	**17.6**	**33.0**	30.9	31.7
<High school	**9.4**	**6.0**	**40.8**	**13.5**	**14.8**	**7.1**
High school	28.3	29.6	**26.8**	**42.7**	28.0	33.2
Some College	**30.7**	**35.8**	**21.3**	**30.0**	**30.4**	**37.7**
College or more	**31.5**	**28.6**	11.1	13.8	**26.8**	**22.0**
Own home	**74.5**	**63.1**	47.1	41.9	**48.3**	**38.0**

Bold font indicates statistical significance (*p*<0.05).

Study 2 – Health variables
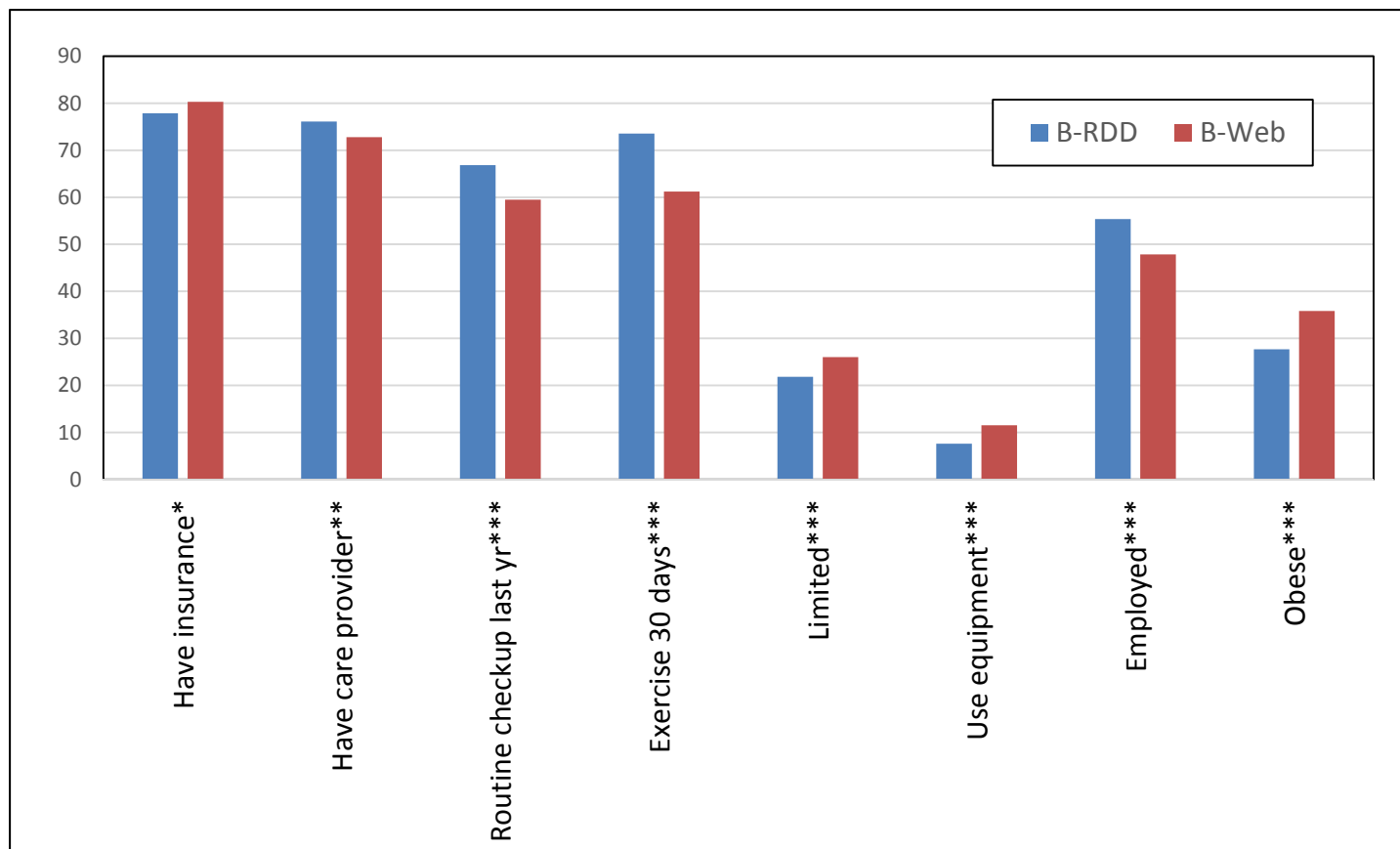
- Social desirability?
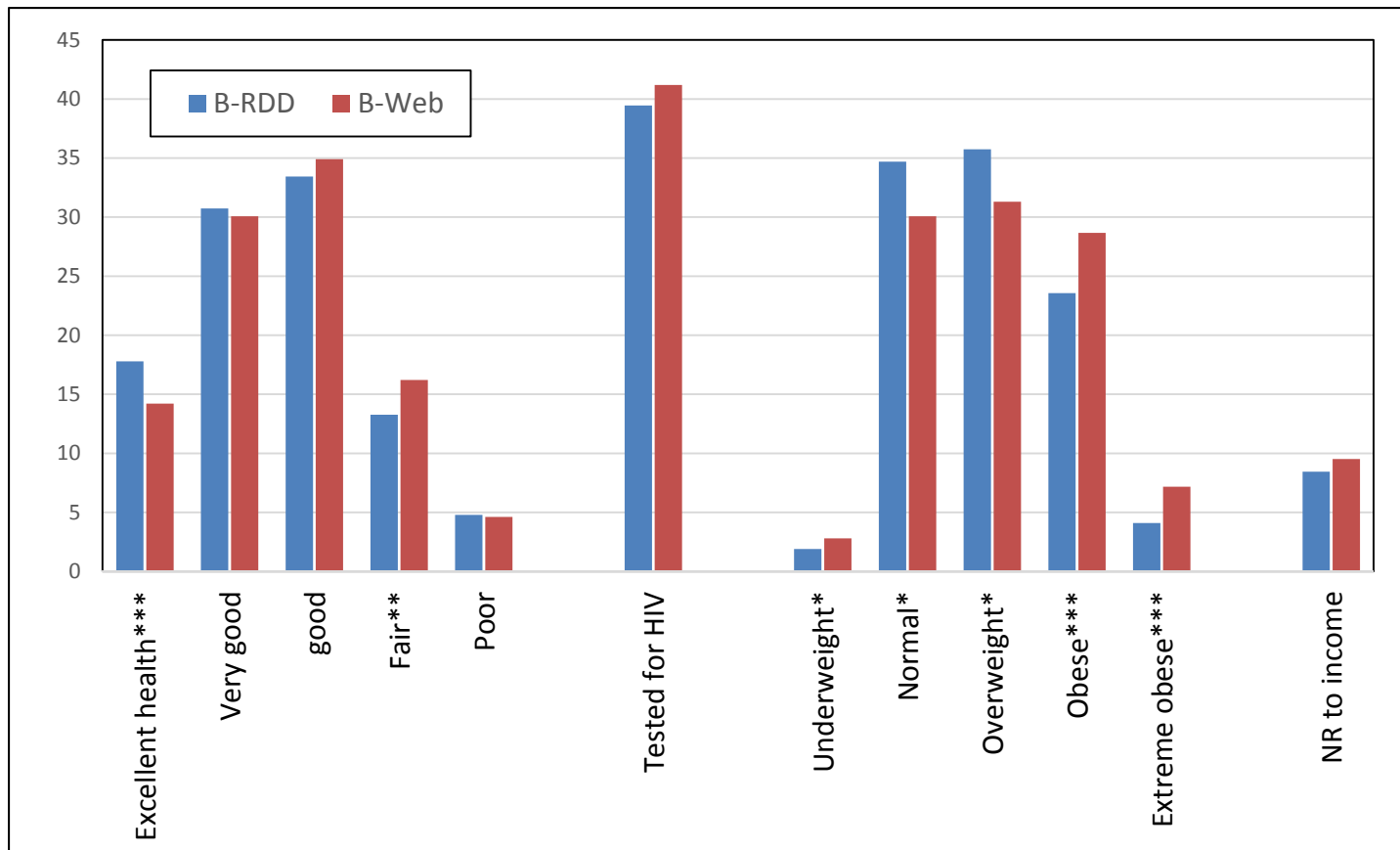
* *p*<0.05, ** *p*<0.01, *** *p*<0.001

Study 2 – Health variables

- Web opt-in panel members not as healthy?

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Study 2 – Implications

- Can we blend the two data sets?
- Data comparability
  - Sampling differences: bringing in different people?
    - Potentially, yes
    - Web somewhat higher SES status
      - Bias not consistent across race/ethnicity (e.g., education, income)
    - Web not as healthy; higher risk behaviors
  - Mode effects
    - Unclear evidence
  - Question context effects on SRH
    - Potentially, yes

Report of the AAPOR Task Force on Non-Probability Sampling

Conclusions and Recommendations

- Unlike probability sampling, there is no single framework that adequately encompasses all of non-probability sampling

- Researchers and other data users may find it useful to think of the different non-probability sample approaches as falling on a continuum of expected accuracy of the estimates

- Transparency is essential

- Making inferences for any probability or non-probability survey requires some reliance on modeling assumptions

Conclusions and Recommendations

- The most promising non-probability methods for surveys are those that are based on models that attempt to deal with challenges to inference in both the sampling and estimation stages

- One of the reasons model-based methods are not used more frequently in surveys may be that developing the appropriate models and testing their assumptions is difficult and time-consuming, requiring significant statistical expertise

- Fit for purpose is an important concept for judging survey quality, but its application to survey design requires further elaboration

- Sampling methods used with opt-in panels have evolved significantly over time and, as a result, research aimed at evaluating the validity of survey estimates from these sample sources should focus on sampling methods rather than the panels themselves

Conclusions and Recommendations

- If non-probability samples are to gain wider acceptance among survey researchers there must be a more coherent framework and accompanying set of measures for evaluating their quality

- Although non-probability samples often have performed well in electoral polling, the evidence of their accuracy is less clear in other domains and in more complex surveys that measure many different phenomena

- Non-probability samples may be appropriate for making statistical inferences but the validity of the inferences rests on the appropriateness of the assumptions underlying the model and how deviations from those assumptions affect the specific estimates

TRANSPARENCY INITIATIVE COORDINATING COMMITTEE REPORT

Timothy Johnson

Survey Research Laboratory

University of Illinois at Chicago

AAPOR Code, Section III Summary

Report Immediately (a)
- Who sponsored, conducted & funded the research
- Exact question wording
- Definition of the population
- Geographic location
- Sample frame description
- Sample design
- Sample size & error
- Weighting & cluster adjustments
- Results based on parts of sample only
- Method(s) and dates of data collection

Within 30 Days (b-d)
- Interviewer/respondent instructions
- Relevant stimuli (show cards)
- Sampling frame's coverage
- Methods of panel recruitment (for pre-recruited panels)
- Sample design details (eligibility, screening, oversamples, incentives)
- Sample dispositions
- Weighting details
- Data verification details
- Response rates
- All of the above for each if multiple samples or modes

Joining theTransparency Initiative
(as currently envisioned by TICC)

1. Organization completes TI Certification Agreement
   - Including promise that all relevant employees have completed AAPOR's online educational modules
2. Organization appoints representative to coordinate compliance with AAPOR
3. Organization provides TI compliant documentation from two recent surveys for review
4. Organization pays application fee to AAPOR
5. TICC reviews and approves applications
6. Once approved, organization becomes TI certified
   - Receives letter from AAPOR President
   - Receives TI logo to display on website
   - Organization's name added to AAPOR web site list of TI members

Transparency Initiative Monitoring & Enforcement Methods
(as currently envisioned by TICC)

- Focus will be on continuous education
- TI members asked to reconfirm commitment on annual basis by re-signing the Certification Agreement.
- Agree to cooperate in an evaluation of the transparency of a sample of studies once every two years
  - Reports will be shared with organization and otherwise kept confidential
- Complaints from public about disclosure will be reviewed by TICC within 30 days
  - Where complaint found to have merit, organizations will be given the opportunity to address the problem

AAPOR Task Force Update

- Non-Probability Task Force – report released in May

- Active Task Forces – reports forthcoming:
  - Public Opinion and Leadership Task Force
  - Survey Refusal Task Force
  - Emerging Technologies Task Force

68th Annual Conference
May 16 – 19, 2013
Seaport Boston Hotel &
Seaport World Trade Center
Boston, Massachusetts

Asking Critical Questions:
Toward a Sustainable Future for
Public Opinion and Social Research

Membership Highlights

➢ Expanded outreach to students
  ➢ Added a second student event at this year's conference
  ➢ Increased email communication with student members throughout the year

➢ Increased outreach to members about Honorary Lifetime Membership status

➢ Revisited plans for ongoing membership surveys
  ➢ Member/post-conference survey recently sent out via email – Don't be a non-respondent!

68th Annual Conference
May 16 – 19, 2013
Seaport Boston Hotel &
Seaport World Trade Center
Boston, Massachusetts

Asking Critical Questions:
Toward a Sustainable Future for
Public Opinion and Social Research

AAPOR
AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

Seymour Sudman Student Paper Competition

- Seeks papers in any field related to the study of public opinion
- Open to current students and those who received degree during prior calendar year
- Submissions due in January of the conference year
- Winner gets $750, plus airfare, hotel, and conference registration

AAPOR Student Travel Award

- Open to those enrolled in a masters or doctoral graduate program related to public opinion research or survey methodology
- Submissions due in February of the conference year
- Up to 8 awardees get $500 to defray cost of travel to conference

Burns "Bud" Roper Fellow Award

- Open to those who recently started career (currently work for pay & have primary work responsibilities related to survey research or public opinion)
- Submissions due in February of the conference year
- Up to 10 awardees get up to $700 for conference-related expenses and up to $300 for short course

AAPOR Membership Benefits

NEW! Journal of Survey Statistics and Methodology

▶ New quarterly, interdisciplinary journal being launched by AAPOR and the American Statistical Association (ASA)

▶ Will publish cutting-edge articles on statistical and methodological issues for surveys and censuses, empirical and theoretical papers, applied papers and review papers.

▶ Aims to be the flagship journal for research on survey statistics and methodology.

AAPOR Membership Benefits

Public Opinion Quarterly

- ▸ Free subscription for members (including hard copy and online access)
- ▸ Among the most frequently cited journals of its kind
- ▸ Important theoretical contributions to opinion and communication research
- ▸ Analyses of current public opinion
- ▸ Investigations of methodological issues involved in survey validity-- including questionnaire construction, interviewing and interviewers, sampling strategy, and mode of administration

AAPOR Membership Benefits

Survey Practice

▸ Online AAPOR publication

▸ Provides current information on issues in survey research and public opinion that is useful to survey and public opinion practitioners, new survey researchers, and those interested in survey and polling methods.

AAPOR Membership Benefits

▸ Webinars

  ▸ Members get discounted rates

  ▸ Upcoming webinars:

    ▸ **Survey Coding**, Jon Krosnick and Skip Lupia, July 2013

    ▸ **Designing Effective Online Questions**, Scott Crawford, August 2013

    ▸ **Smartphone Surveys**, Trent Buskirk, September 2013

    ▸ **Questionnaire Design**, Allyson Holbrook, October 2013

    ▸ **The Questionnaire Design Pitfalls of Multiple Modes**, Gerry Nicolaas and Pamela Campanelli, November 2013

    ▸ **Item Response Theory**, Bryce Reeve, December 2013