

DATA QUALITY OF TAIWAN'S ELECTION AND DEMOCRATIZATION STUDY

RETEST INTERVIEW AND INTERVIEWER ASSESSMENT

TSUNG-WEI LIU

Associate Professor
Chung-cheng University

twliu1972@gmail.com

CHI-LIN TSAI

Post-Doctoral Fellow
Cheng-chi University

chilin.tsai.clt@gmail.com

KUANG-HUI CHEN

Associate Professor
Chung-cheng University

ckhglory@gmail.com

RESEARCH QUESTION

- Is **INTERVIEWER ASSESSMENT (I.A.)** an informative indicator of survey data quality?



INTERVIEWER ASSESSMENT (I.A.)

- a.k.a. **INTERVIEWER PERCEPTION**
- interviewers' assessments of the interviews that they have just completed
- Whenever an interviewer completes an interview, the survey may ask that **INTERVIEWER** to answer following questions:
 - *How **cooperative** was the respondent?*
 - *How well do you think the respondent **understood the questions**?*
 - *How **trustworthy** do you think the respondent's answers are?*

RESEARCH QUESTION

- Is **INTERVIEWER ASSESSMENT (I.A.)** an informative indicator of survey data quality?
- We argue that the I.A. may be a good **supplement** to (not a **replacement** for) conventional measures of data quality

CONVENTIONAL MEASURES OF DATA QUALITY

- **RELIABILITY**, E.G.:
 - Test-retest
 - Split-half
 - Parallel-forms
- **VALIDITY**, E.G.:
 - Content
 - Face validity
 - Criterion-related
 - Construct



LIMITATIONS OF THE CONVENTIONAL MEASURES

- public opinion is not static by nature
- mainly for individual survey measurement
- relatively costly and inconvenient to use

I.A. AS A SUPPLEMENTARY MEASURE OF DATA QUALITY

- overall evaluation
- contextualisation
- relatively affordable and convenient to use

I.A. IN TEDS

		INTERVIEWER ASSESSMENT
2001		No
2002C	Taipei	Yes
2002C	Kaohsiung	Yes
2003		Yes
2004P		Yes
2004L	Questionnaire A	Yes
2004L	Questionnaire B	Yes
2005M	Taipei	Yes
2005M	Taichung	Yes
2005M	Yunlin	Yes
2005M	Kaohsiung	Yes
2006C	Taipei	Yes
2006C	Kaohsiung	Yes
2008L		Yes

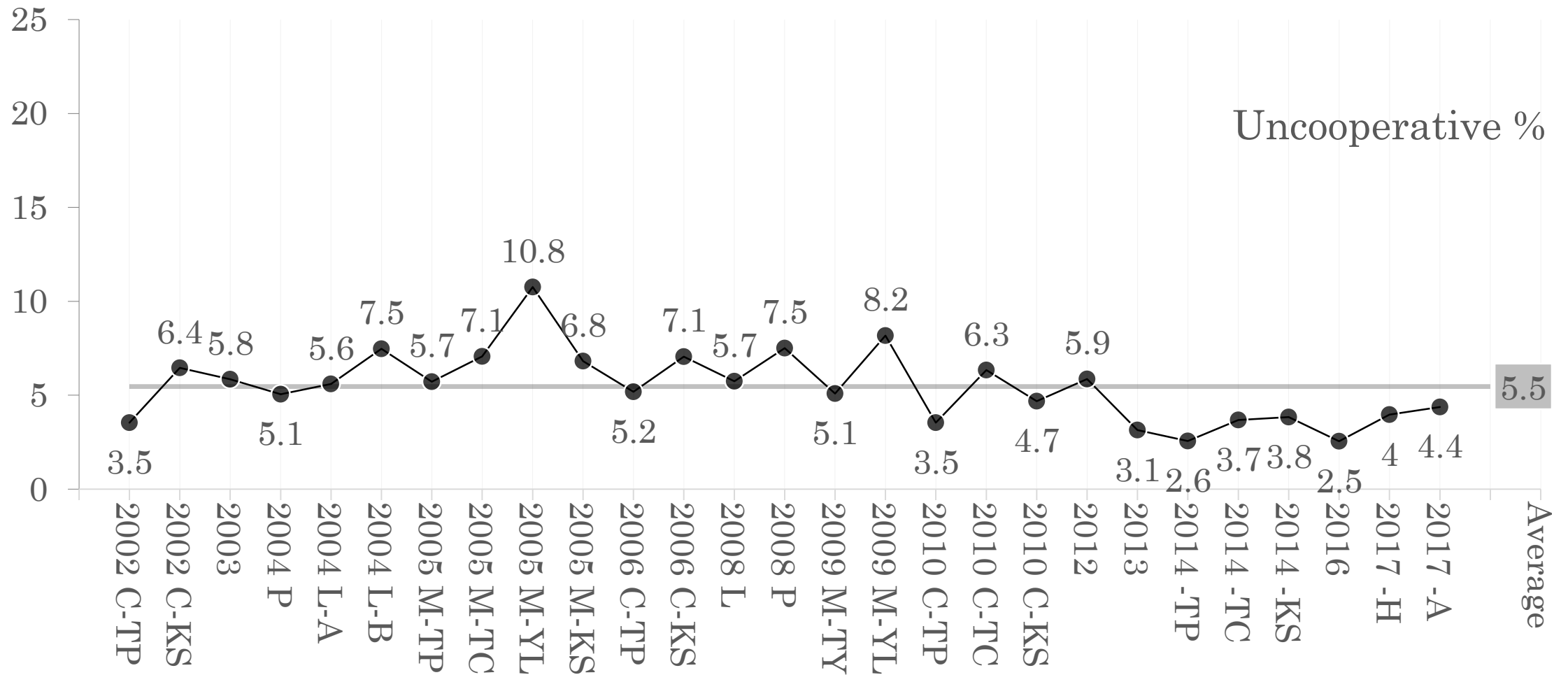
		INTERVIEWER ASSESSMENT
2008P		Yes
2009M	Taoyuan	Yes
2009M	Yunlin	Yes
2010C	Taipei	Yes
2010C	Taichung	Yes
2010C	Kaohsiung	Yes
2012		Yes
2013		Yes
2014	Taipei	Yes
2014	Taichung	Yes
2014	Kaohsiung	Yes
2016		Yes
2017	Household	Yes
2017	Address	Yes

I.A. IN TEDS

- Three items are of particular interest:
 - *How cooperative was the respondent?*
 - *How well do you think the respondent understood the questions?*
 - *How trustworthy do you think the respondent's answers are?*

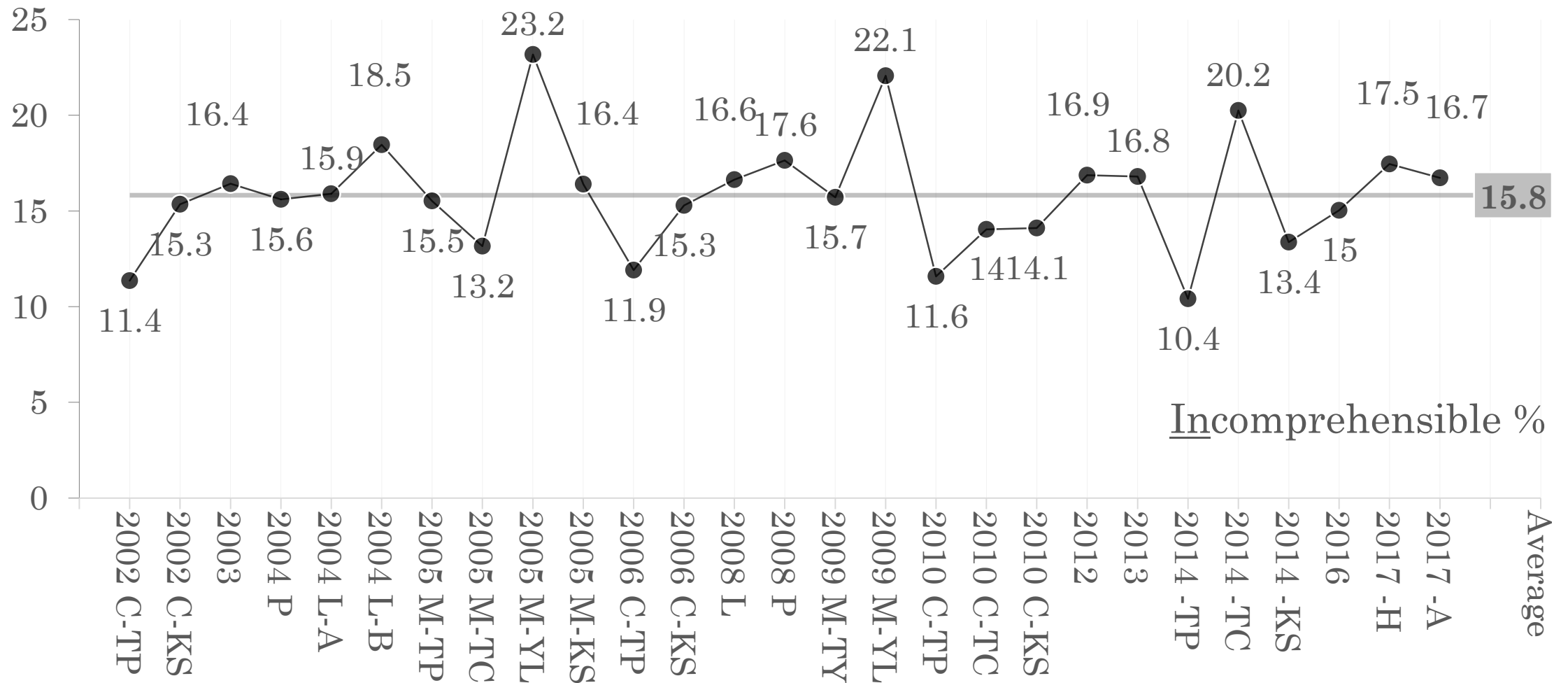
I.A. IN TEDS: COOPERATIVENESS

How cooperative was the respondent?



I.A. IN TEDS: COOPERATIVENESS

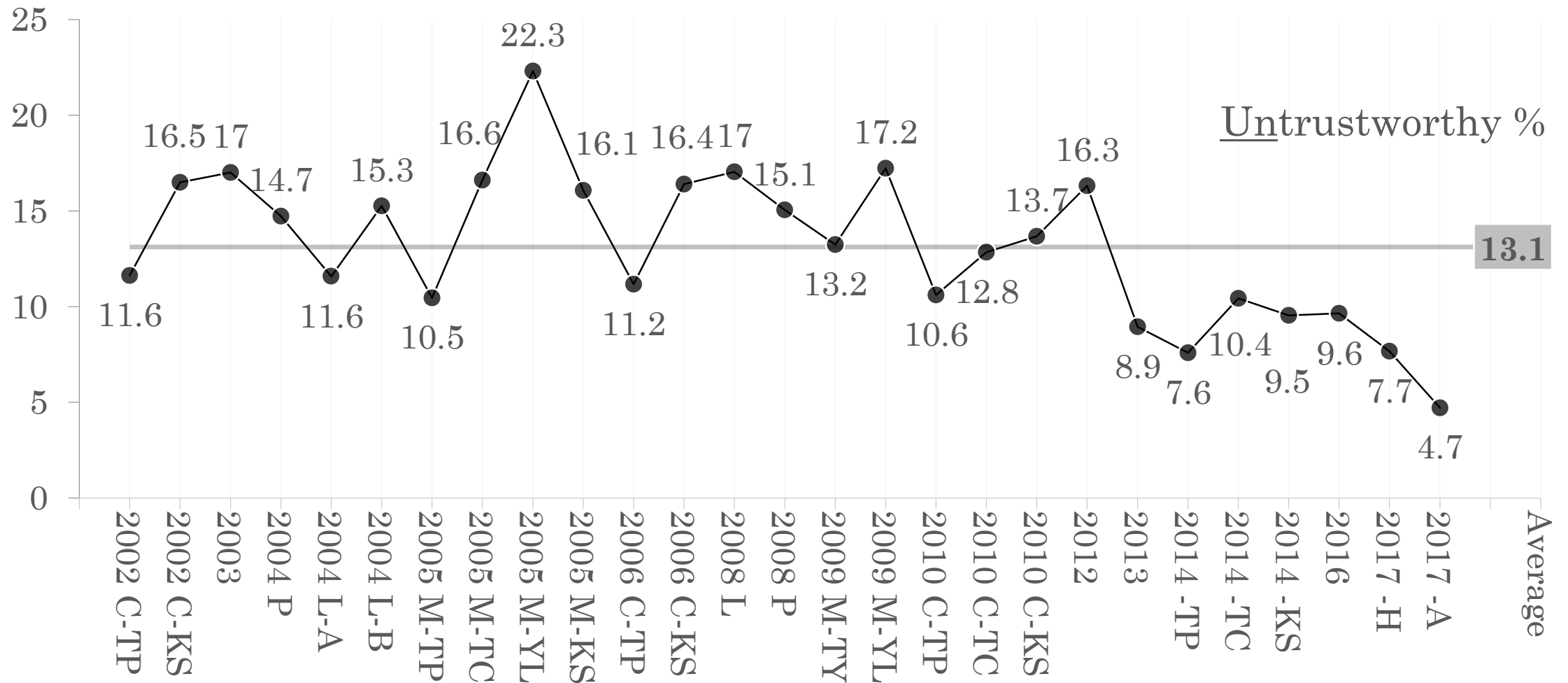
How well do you think the respondent understood the questions?



Incomprehensible %

I.A. IN TEDS: COOPERATIVENESS

How trustworthy do you think the respondent's answers are?



IS THE I.A. A GOOD INDICATOR OF DATA RELIABILITY?

- We use TEDS **TEST-RETEST RELIABILITY** as the benchmark
- The **INTERVIEW ASSESSMENT** is considered a good indicator of data reliability, if its results match the **TEST-RETEST RELIABILITY**

TEDS RETEST INTERVIEW

		RETEST INTERVIEW
2001		Yes
2002C	Taipei	Yes
2002C	Kaohsiung	Yes
2003		Yes
2004P		Yes
2004L	Questionnaire A	Yes
2004L	Questionnaire B	Yes
2005M	Taipei	Yes
2005M	Taichung	Yes
2005M	Yunlin	Yes
2005M	Kaohsiung	Yes
2006C	Taipei	Yes
2006C	Kaohsiung	Yes
2008L		Yes

		RETEST INTERVIEW
2008P		Yes
2009M	Taoyuan	Yes
2009M	Yunlin	Yes
2010C	Taipei	Yes
2010C	Taichung	Yes
2010C	Kaohsiung	Yes
2012		Yes
2013		No
2014	Taipei	Yes
2014	Taichung	Yes
2014	Kaohsiung	Yes
2016		Yes
2017	Household	No
2017	Address	No

TEDS RETEST INTERVIEW

- The procedure of the TEDS retest interview is to:
 - sample roughly **20%** of the respondents from the sample of the main interview
 - invite them to answer some questions that are chosen from the questionnaire of the main interview

RELIABILITY OF A QUESTION

AN EXAMPLE

- Question B2 in TEDS 2016

“Do you usually talk about politics or elections with other people? Is it often, sometimes, seldom, or never?”

<i>n</i> = 320		RETEST INTERVIEW				
		often	sometimes	seldom	never	refuse
MAIN INTERVIEW	often	4	9	2	1	0
	sometimes	6	74	42	5	0
	seldom	0	21	72	18	0
	never	0	4	23	39	0
	refuse	0	0	0	0	0

- Reliability of Question B2 = n in diagonal cells / $n = (4+74+72+39+0) / 320 = 59\%$

RELIABILITY OF A SURVEY

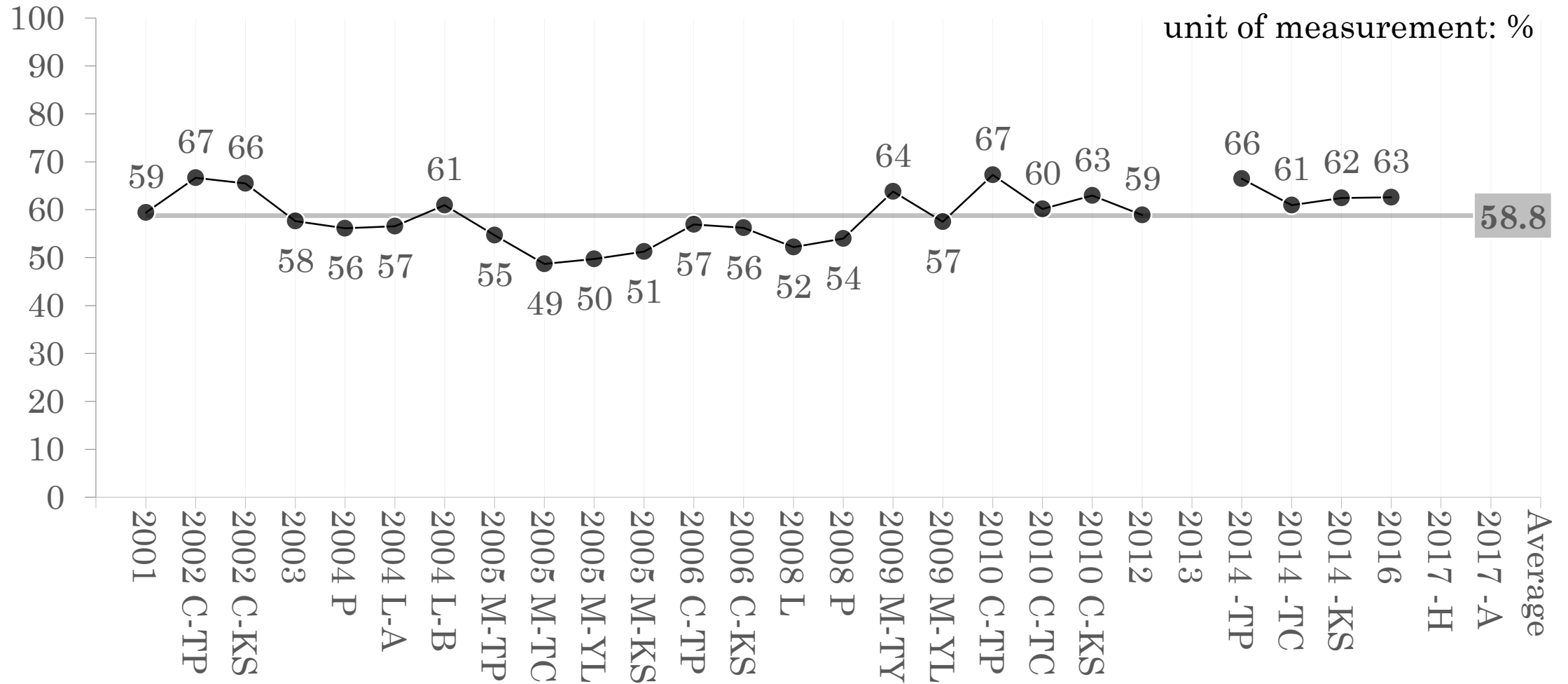
AN EXAMPLE

- There are 25 questions in the questionnaire of the TEDS 2016 retest interview
- We exclude follow-up questions and questions that allow multiple answers
- Reliabilities of the remaining 17 questions are:

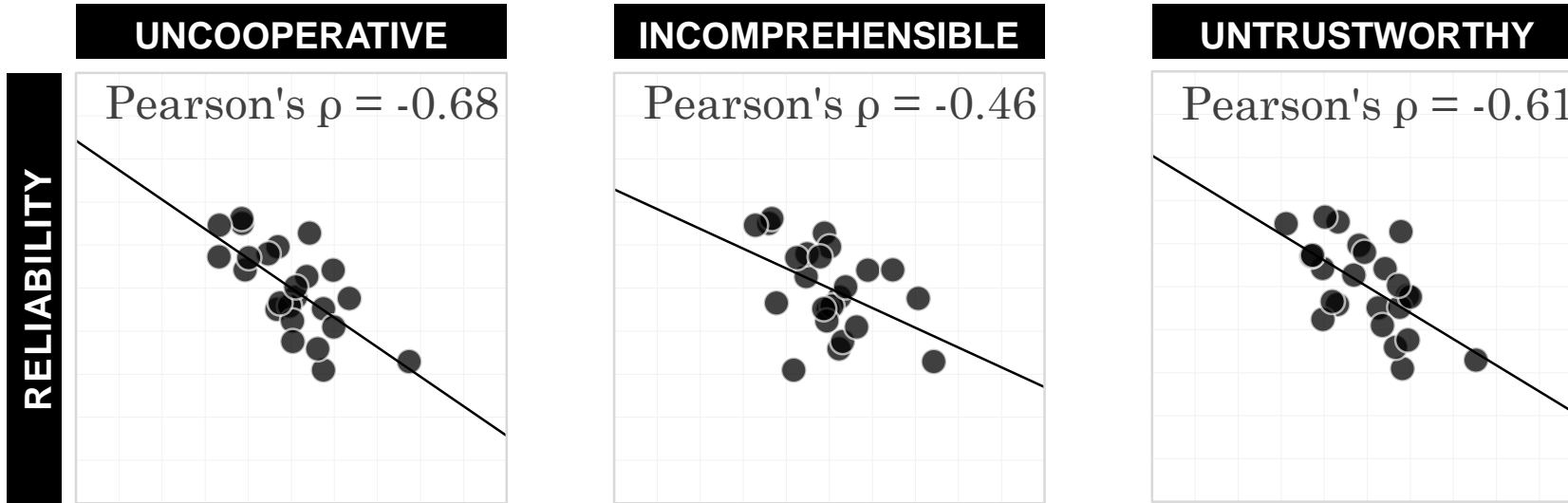
QUESTION	RELIABILITY	QUESTION	RELIABILITY	QUESTION	RELIABILITY
A2	80%	K6b	28%	P4	59%
B2	59%	K6c	31%	Q1	76%
H6	51%	N1	52%	Q4	49%
J1	98%	N3	56%	S4	87%
J4	80%	N3a	76%	S11	73%
K6a	30%	P1	79%		

- Reliability for TEDS 2016 = Average over all these reliabilities = 63%

TEDS TEST-RETEST RELIABILITY (AS THE BENCHMARK)

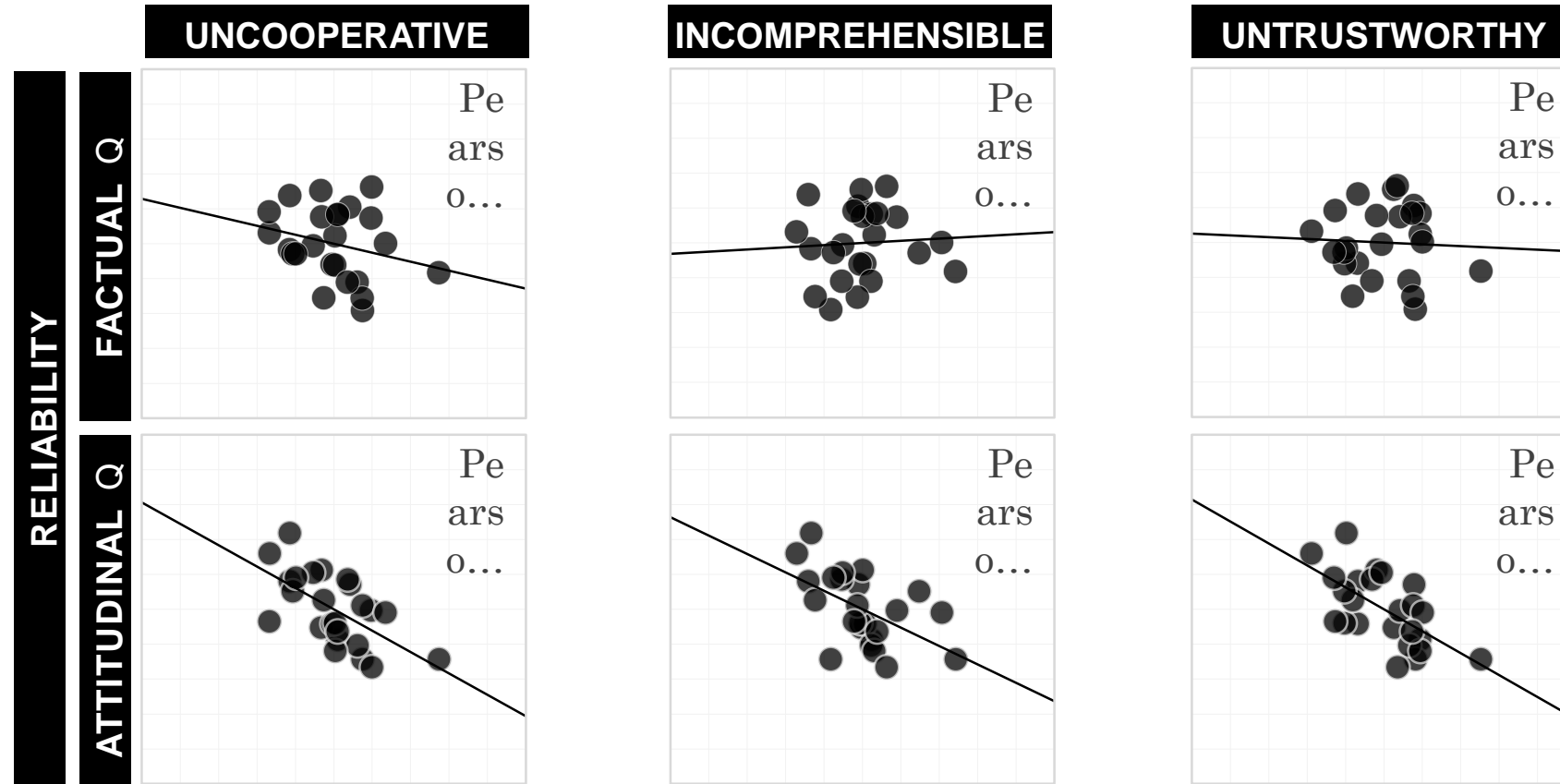


IS THE I.A. A GOOD INDICATOR OF DATA RELIABILITY?



- Match quite well:
 - The higher the proportion of respondents flagged as uncooperative, incomprehensible and/or untrustworthy in a survey, the poorer the overall test-retest reliability of that survey
- The interviewer assessment appears to be an useful indicator of data reliability

IS THE I.A. A GOOD INDICATOR OF DATA RELIABILITY?



- The interviewer assessment is more informative in terms of the reliability of attitudinal questions than that of the factual questions.

IS THE I.A. A GOOD INDICATOR OF DATA VALIDITY?

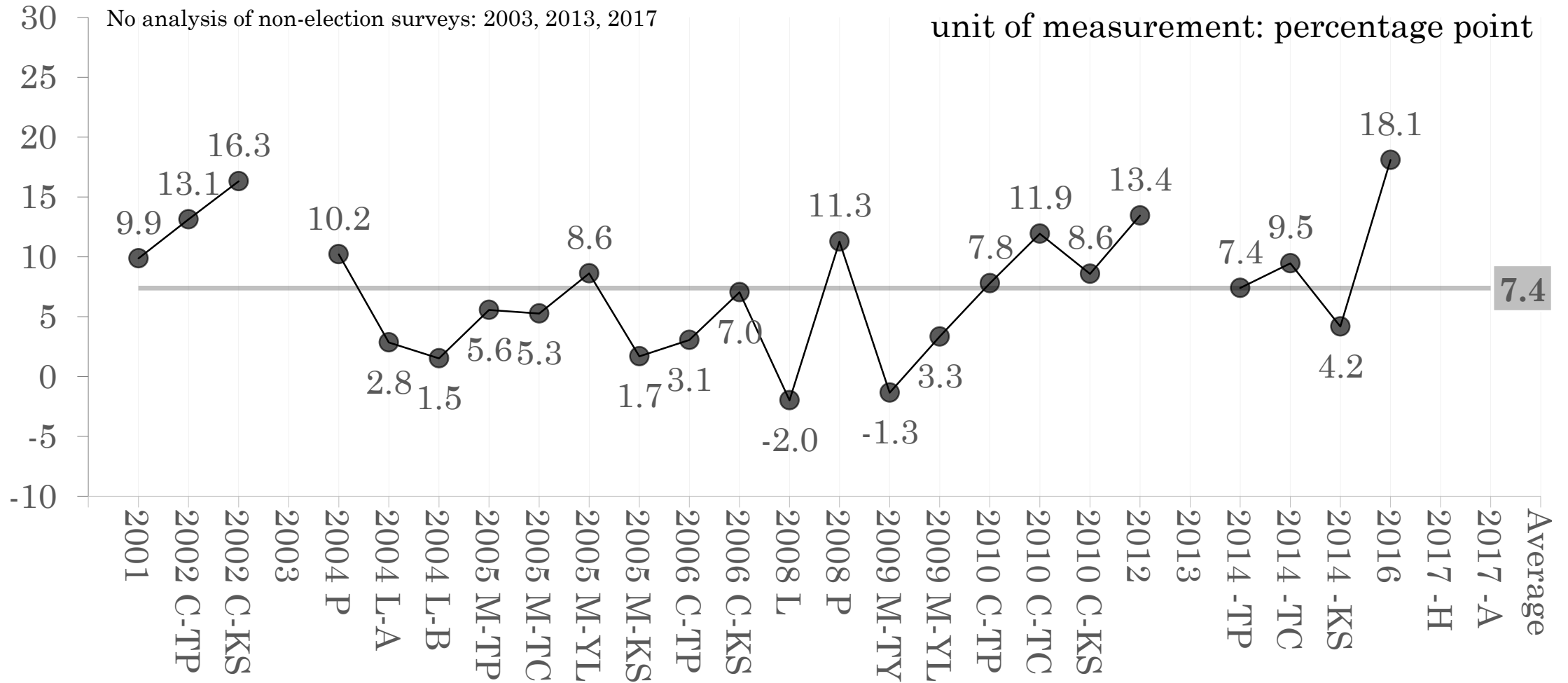
- TEDS official validity evaluation:
 - Based on the results of pre-test interviews
 - Not standardised as much as the test-retest reliability evaluation
 - Not available for end users
- We are unable to use that as a benchmark for examining the interviewer assessment
- We construct **three alternative benchmarks**

1ST BENCHMARK

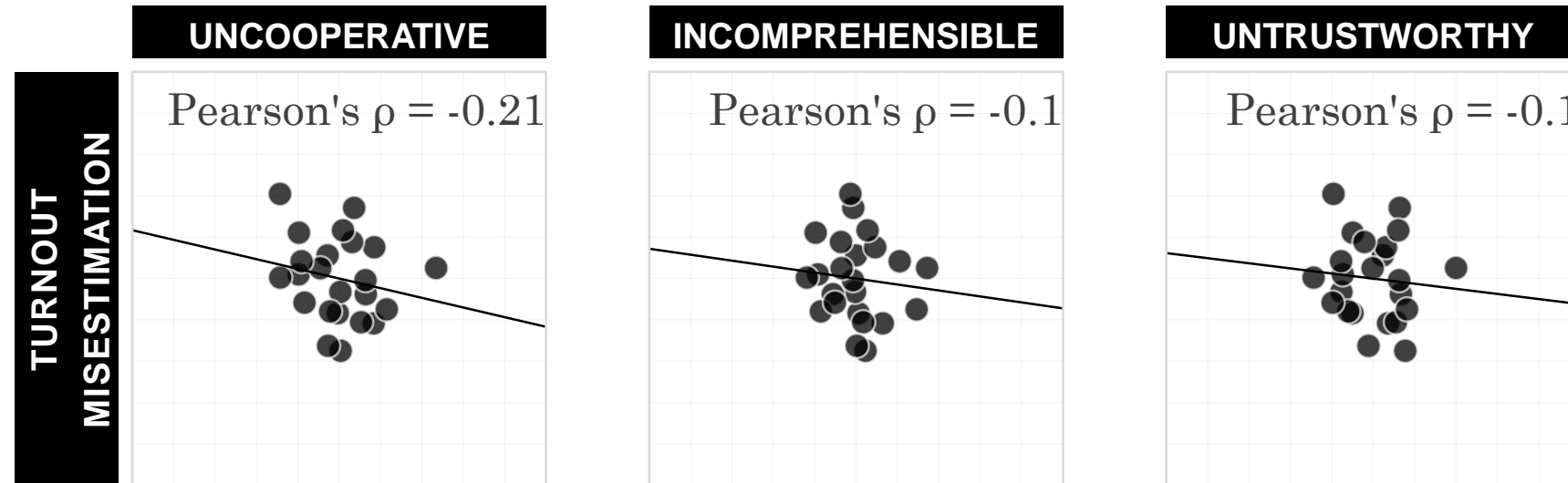
▪ **TURNOUT MISESTIMATION**

- Turnout as one of the most important measurements for TEDS surveys
 - Misestimation = TEDS estimate of turnout rate minus the official rate
 - Assume that serious misestimation of a survey = low validity
 - an application of the conventional criterion-related validity
-
- The **INTERVIEW ASSESSMENT** is considered a good indicator of data validity, if it is **positively** correlated with turnout mis-estimation:
 - e.g. TEDS surveys with **serious** turnout mis-estimation receive **poor** I.A.

TEDS TRUNOUT MISESTIMATION (AS BENCHMARK 1)



IS THE I.A. A GOOD INDICATOR OF DATA VALIDITY?

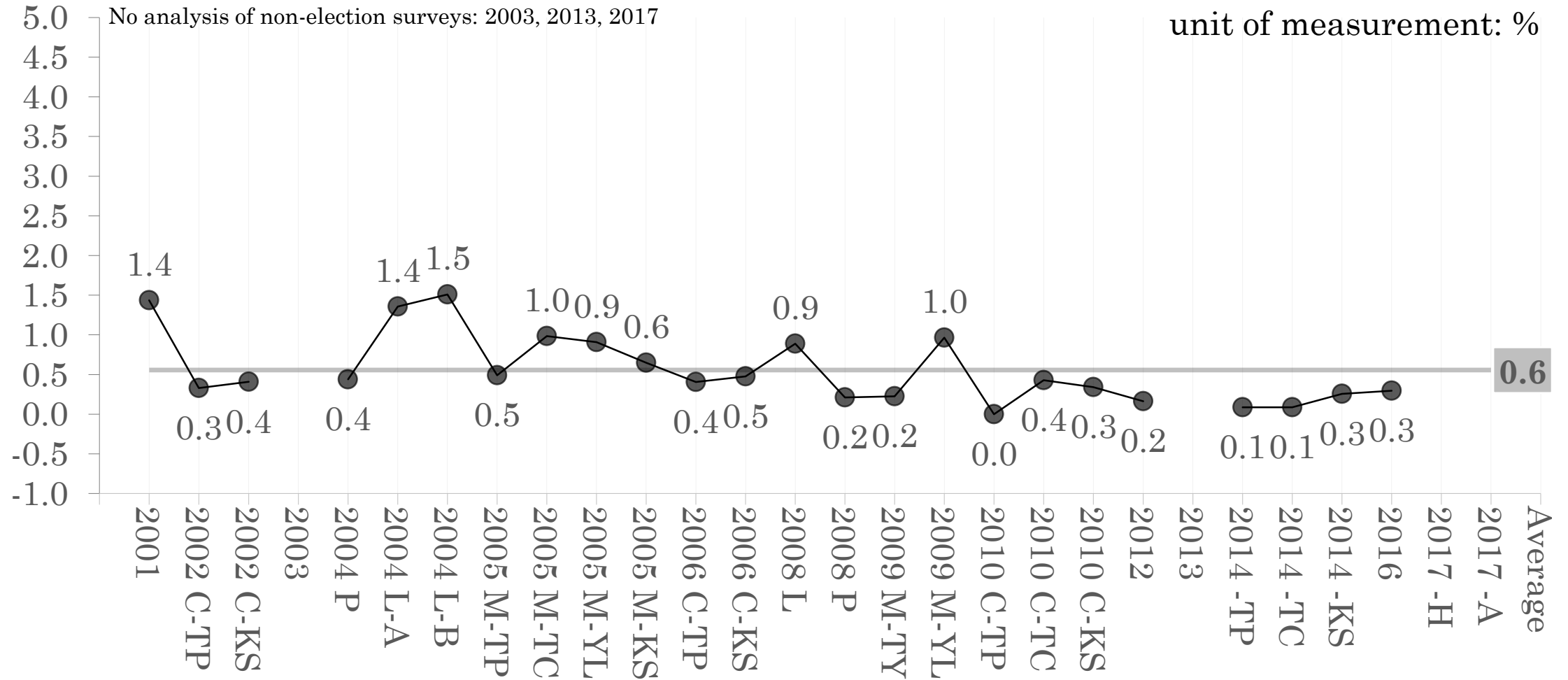


- I.A. does not seem to be a good indicator in terms of the 1st benchmark
 - The interviewer assessment gives **incorrect** indication as to how serious a turnout rate is misestimated by a TEDS survey

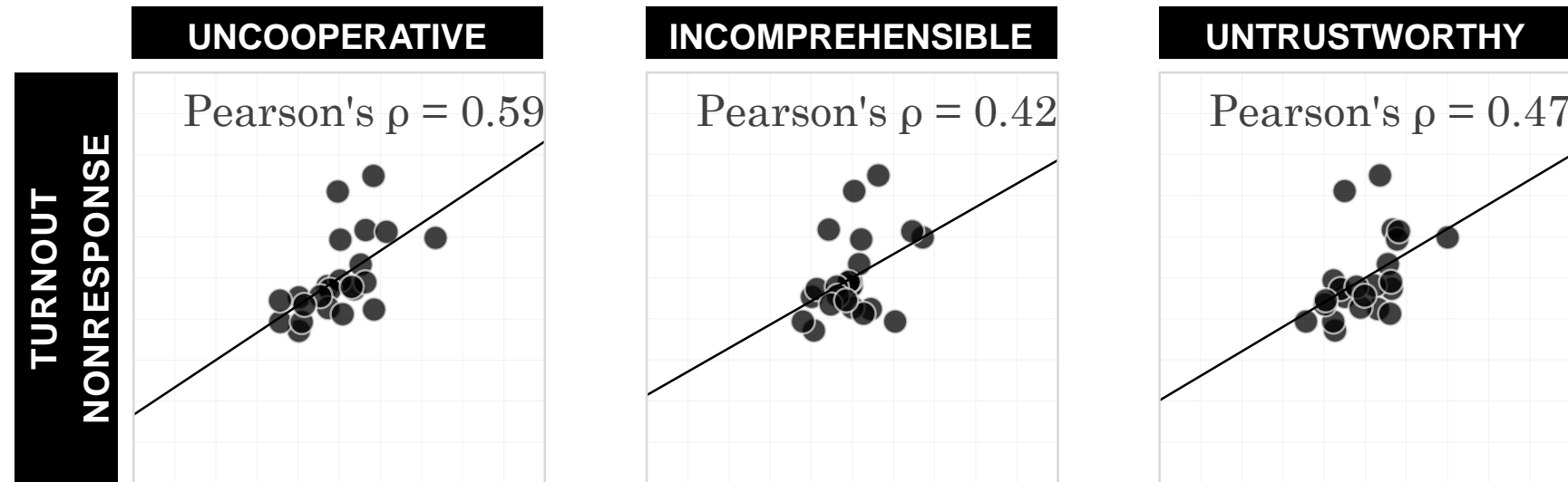
2ND BENCHMARK

- Item nonresponse to the turnout question
 - Assume that more nonresponses = low validity
 - An application of the conventional face validity
- The **INTERVIEW ASSESSMENT** is considered a good indicator of data validity, if it is **positively** correlated with Item nonresponse to the turnout question:
 - e.g. TEDS surveys with **more** nonresponses receive **low** I.A.

TEDS TRUNOUT NONRESPONSE (AS BENCHMARK 2)



IS THE I.A. A GOOD INDICATOR OF DATA VALIDITY?

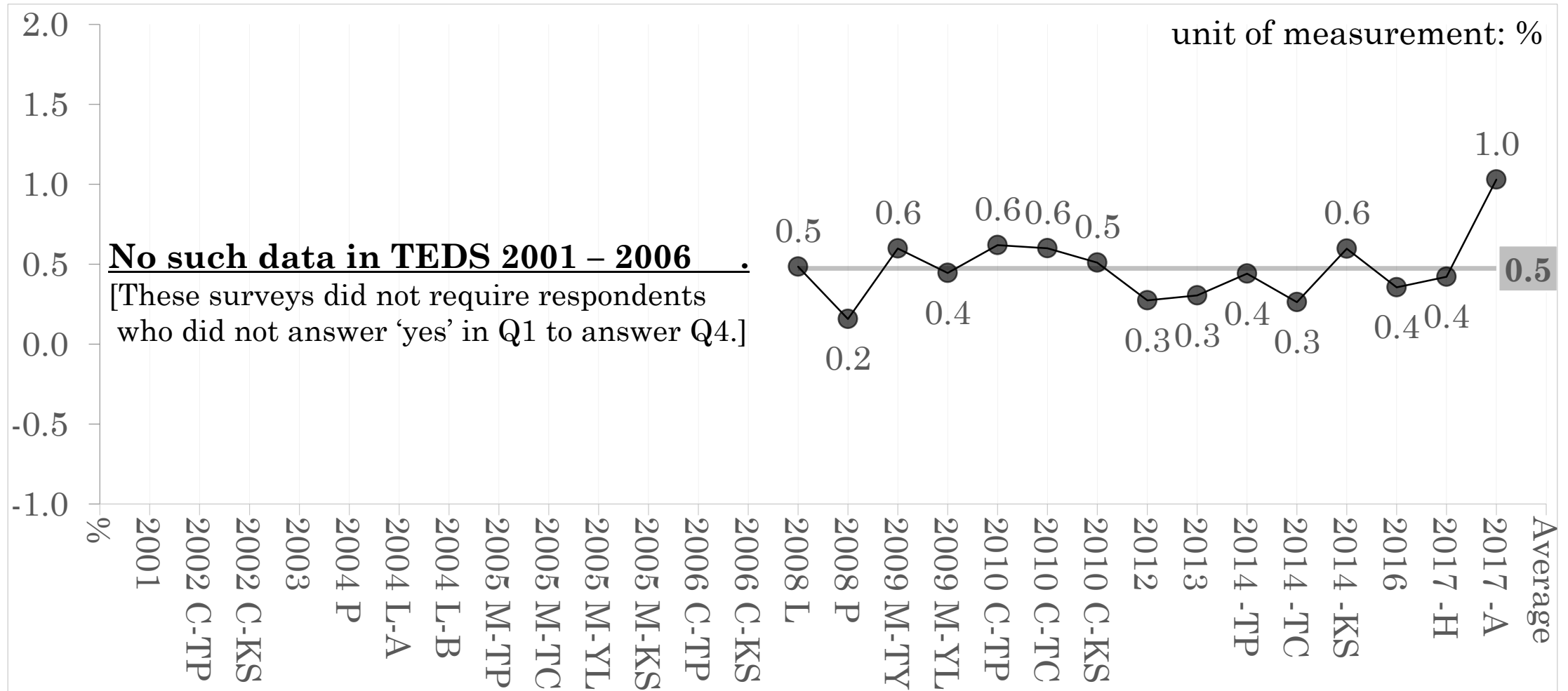


- I.A. matches the 2nd benchmark fairly well:
 - The interviewer assessment gives **correct** indication as to the relative proportion of respondents who did not answer the turnout question

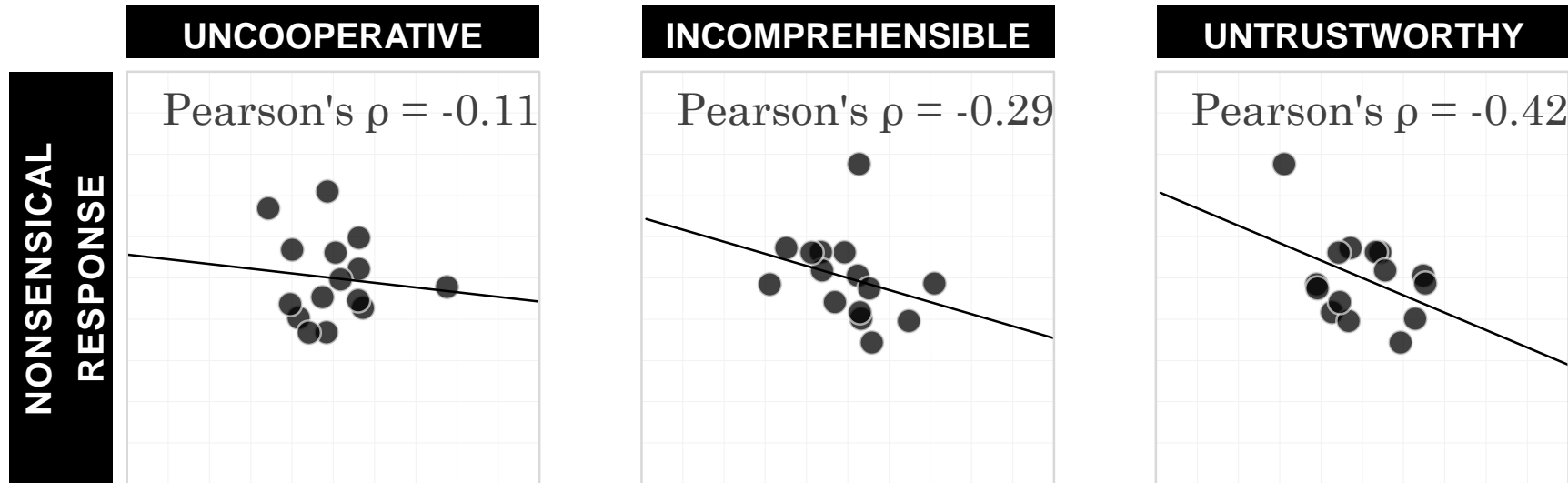
3RD BENCHMARK

- TEDS questions for measuring party identification:
 - *Q1. Do you usually think of yourself as close to any particular party?*
 - *Q2. Do you feel yourself a little closer to one of the political parties than the others?*
 - *Q3. Which party do you feel closest to?*
 - *Q4. Do you feel very close to this party, somewhat close, or not very close?*
- Nonsensical responses:
 - Answer 'yes' in Q1 but answer 'very close' in Q4
 - Assume more nonsensical responses = low validity
 - Another application of the conventional face validity
- The **INTERVIEW ASSESSMENT** is considered a good indicator of data validity, if it is **positively** correlated with Item nonresponse to the turnout question:
 - e.g. TEDS surveys with **more** nonsensical responses receive **low** I.A.

TEDS PID NONSENSICAL RESPONSE (AS BENCHMARK 3)



IS THE I.A. A GOOD INDICATOR OF DATA VALIDITY?



- I.A. does not seem to be a good indicator in terms of the 3rd benchmark
 - The interviewer assessment gives **incorrect** indication as to the proportion of respondents who gave nonsensical answer to the party identification questions

CONCLUSIONS

- As expected, the interviewer assessment can be informative to the evaluation of data quality, but:
 - More informative as to the data reliability than the validity
 - More informative as to the reliability of attitudinal questions than that of factual questions
- The test-retest reliability evaluation is becoming increasingly difficult to do
 - Many respondents refuse consent for the retest interview
 - The interview assessment may be a way out

ISSUES NEEDED FURTHER INVESTIGATION

- **Why** does the interviewer assessment fail to serve as a good indicator of the reliability of factual questions and data validity?
- Can **other kinds of interviewers' assessments** provide information for data evaluation (e.g. interviewers' assessments of respondents' knowledge and interest with respect to survey questions)?
- How can we evaluate and improve **the quality of the interviewer assessment itself**?
- More systematic and theory-driven analyses are in progress...