

FORS



explore.understand.share.

Unil

UNIL | Université de Lausanne

Translation, wording, question design,
survey design:
an empirical exploration of different
challenges to the comparability of
international survey data

3MC conference, 26th July 2016, Chicago

by Michèle Ernst Stähli, International Surveys, FORS

Research question

- What makes concepts measurements comparable?
- How far the specific wording matters?
- How far the scale matters?
- How far survey design features matter?
(sampling, contact procedures, interview mode, surrounding questions, etc.)
- *What is the impact of the different sources of error?*

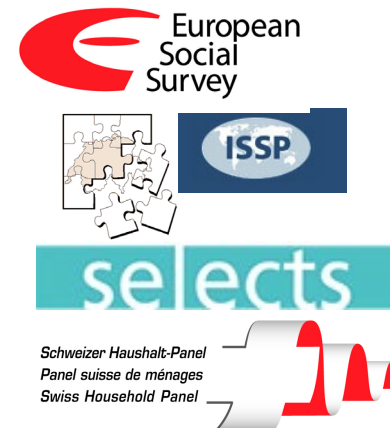
Research strategy

- 1) Comparisons of same or very similar questions on a national level
 - across surveys (different modes, topics, question designs)
 - across question designs (different designs implemented in same survey)
- 2) Comparisons of same question in same survey on an international level
 - across countries (with relative survey designs)
 - across institutional and cultural country characteristics
- 3) Comparisons of very similar question in different surveys on an international level
 - across surveys (with relative question and survey designs)
 - across countries (with relative survey designs)

First step: national comparisons

Same questions compared in different national surveys (in Switzerland):

- Concepts:
Political interest & Satisfaction with democracy in own country
- Surveys considered:
 - **ESS (European Social Survey)**
 - **MOSAiCH/ ISSP (International Social Survey Program)**
 - **SELECTS (Swiss Electoral Study)**
 - **SHP (Swiss Household Panel)**



First step: national comparisons

Same questions compared in different national surveys (in Switzerland):

- **Concepts:**
Political interest & Satisfaction with democracy in own country
- **Surveys considered:**
 - **ESS (European Social Survey)**
 - **MOSAiCH/ ISSP (International Social Survey Program)**
 - **SELECTS (Swiss Electoral Study)**
 - **SHP (Swiss Household Panel)**

All with simple random sampling, same s. frame

All fielded in Switzerland in 2010/2011

All produced by FORS

First step: national comparisons

Same questions compared in different national surveys (in Switzerland):

- **Concepts:**
Political interest & Satisfaction with democracy in own country
- **Surveys considered:**
 - **ESS (European Social Survey)**
 - **MOSAiCH/ ISSP (International Social Survey Program)**
 - **SELECTS (Swiss Electoral Study)**
 - **SHP (Swiss Household Panel)**

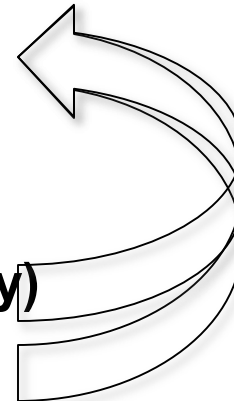
Different populations
Different field procedures
Different modes
Different survey topics

Differences in questions:
- Wording
- Scale

First step: national comparisons

Same questions compared in different national surveys (in Switzerland):

- Concepts:
Political interest & Satisfaction with democracy in own country
- Surveys considered:
 - **ESS (European Social Survey)**
 - **MOSAiCH/ ISSP (International Social Survey Program)**
 - **SELECTS (Swiss Electoral Study)**
 - **SHP (Swiss Household Panel)**



Different question designs implemented in one survey (ESS 2014)

First step: national comparisons

Overview of central features of the surveys and items

Survey	ESS (cross-sectional)	MOSAiCH-ISSP (cross-s)	SELECTS (cross-sect.)	SHP (longitudinal)
Topic	General	General	Politics, postelectoral	General
Mode	FtF	FtF	CATI, WEB	CATI
Target pop	15+	18+	18+ Swiss citizens	Private HH, 14+ for individuals
Periodicity	every 2 years (since 2002)	every 2 years (since 97/05)	every 4 years (since 1995)	every year (since 1999)
Sample	SRS 1500	SRS 1200	SRS 4000 -5000	SRS 4000 -5000 HH
Average response rate	50%	50%	35% (including coverage error)	65% of HH in W1, 85% of individuals within HH, 10% attrition
Edition(s) considered	2010 & 2014	2011	2011	2011
«Political interest» (backtranslated)	<p>“How interested would you say you are in politics?”</p> <p>4 point scale, fully labeled</p> <p>“Very interested”</p> <p>“Quite interested”</p> <p>“Hardly interested”</p> <p>“Not interested at all”</p> <p>With showcards</p>	<p>“How interested would you say you are in politics?”</p> <p>4 point scale, fully labeled</p> <p>“Very interested”</p> <p>“Quite interested”</p> <p>“Hardly interested”</p> <p>“Not interested at all”</p> <p>With showcards</p>	<p>“Generally, how interested are you in politics? Are you...”</p> <p>4 point scale, fully labeled</p> <p>“Very interested”</p> <p>“Rather interested”</p> <p>“Rather not interested”</p> <p>“Not interested at all”</p>	<p>“Generally, how interested are you in politics?”</p> <p>11 point scale, endpoint labels: “Not at all interested”</p> <p>“Very interested”</p>
«Satisfaction w democracy» (backtranslated)	<p>“On the whole, how satisfied are you with the way democracy works in Switzerland?”</p> <p>11 point scale, endpoint labels: “Extremely dissatisfied”</p> <p>“Extremely satisfied”</p> <p>With showcards</p>	<p>“On the whole, are you ... satisfied with the way democracy works in Switzerland?”</p> <p>4 point scale, fully labeled</p> <p>“Very satisfied”</p> <p>“Rather satisfied”</p> <p>“Rather not satisfied”</p> <p>“Not satisfied at all”</p> <p>With showcards</p>	<p>“In general, are you ... satisfied with the way democracy works in our country?”</p> <p>4 point scale, fully labeled</p> <p>“Very satisfied”</p> <p>“Rather satisfied”</p> <p>“Rather not satisfied”</p> <p>“Not satisfied at all”</p>	<p>“Generally, what is your level of satisfaction with the way democracy works in our country?”</p> <p>11 point scale, endpoint labels: “Not at all satisfied”</p> <p>“Completely satisfied”</p>

First step: national comparisons

Overview of central features of the surveys and items

Survey	ESS (cross-sectional)	MOSAiCH-ISSP (cross-s)	SELECTS (cross-sect.)	SHP (longitudinal)
Topic	General	General	Politics, postelectoral	General
Mode	FtF	FtF	CATI, WEB	CATI
Target pop	15+	18+	18+ Swiss citizens	Private HH, 14+ for individuals
Periodicity	every 2 years (since 2002)	every 2 years (since 97/05)	every 4 years (since 1995)	every year (since 1999)
Sample	SRS 1500	SRS 1200	SRS 4000 -5000	SRS 4000 -5000 HH
Average response rate	50%	50%	35% (including coverage error)	65% of HH in W1, 85% of individuals within HH, 10% attrition
Edition(s) considered	2010 & 2014	2011	2011	2011
«Political interest» (backtranslated)	<p>“How interested would you say you are in politics?”</p> <p>4 point scale, fully labeled</p> <p>“Very interested”</p> <p>“Quite interested”</p> <p>“Hardly interested”</p> <p>“Not interested at all”</p> <p>With showcards</p>	<p>“How interested would you say you are in politics?”</p> <p>4 point scale, fully labeled</p> <p>“Very interested”</p> <p>“Quite interested”</p> <p>“Hardly interested”</p> <p>“Not interested at all”</p> <p>With showcards</p>	<p>“Generally, how interested are you in politics? Are you...”</p> <p>4 point scale, fully labeled</p> <p>“Very interested”</p> <p>“Rather interested”</p> <p>“Rather not interested”</p> <p>“Not interested at all”</p>	<p>“Generally, how interested are you in politics?”</p> <p>11 point scale, endpoint labels:</p> <p>“Not at all interested”</p> <p>“Very interested”</p>
«Satisfaction w democracy» (backtranslated)	<p>“On the whole, how satisfied are you with the way democracy works in Switzerland?”</p> <p>11 point scale, endpoint labels:</p> <p>“Extremely dissatisfied”</p> <p>“Extremely satisfied”</p> <p>With showcards</p>	<p>“On the whole, are you ... satisfied with the way democracy works in Switzerland?”</p> <p>4 point scale, fully labeled</p> <p>“Very satisfied”</p> <p>“Rather satisfied”</p> <p>“Rather not satisfied”</p> <p>“Not satisfied at all”</p> <p>With showcards</p>	<p>“In general, are you ... satisfied with the way democracy works in our country?”</p> <p>4 point scale, fully labeled</p> <p>“Very satisfied”</p> <p>“Rather satisfied”</p> <p>“Rather not satisfied”</p> <p>“Not satisfied at all”</p>	<p>“Generally, what is your level of satisfaction with the way democracy works in our country?”</p> <p>11 point scale, endpoint labels:</p> <p>“Not at all satisfied”</p> <p>“Completely satisfied”</p>

First step: national comparisons

Aims: Detect and compare

- Coverage error
- Nonresponse bias
- Measurement errors due to:
 - presence of an interviewer
 - mode of data collection
 - question wording
 - number and labels of answer categories

Strategy:

- Compare sample composition (design weighted)
- Weight on socio-demographics
- Compare means
- Retest in same survey

“Political interest” across CH surveys

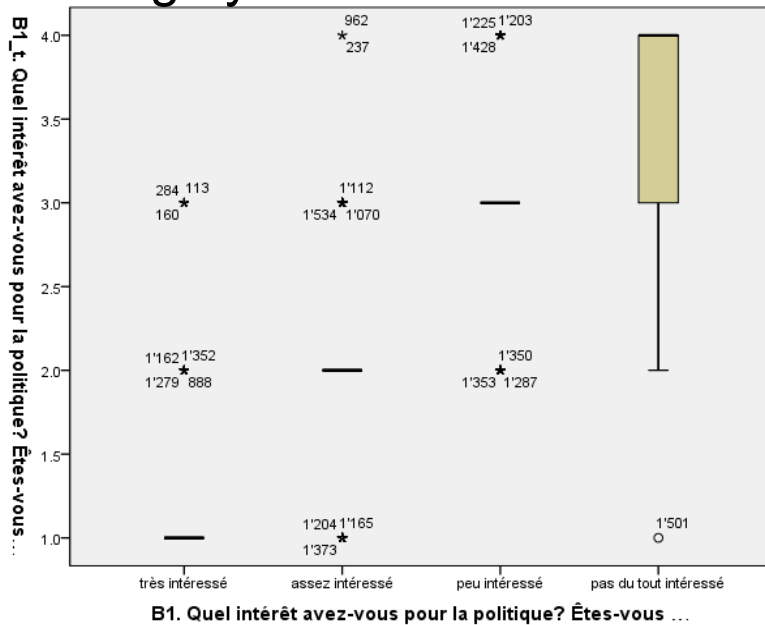
scale	scale wording	quest word	ESS10 FtF (N~=1200)		MOS11 FtF (N~=1000)		SEL11 cati (N~=2300)		SEL11 web (N~=430)		SHP11 cati (N~=6200)		ESS14 FtF (N~=1140/570)	
			mean	(SE)	mean	(SE)	mean	(SE)	mean	(SE)	mean	(SE)	mean	(SE)
4	quite/hardly VS rather/rather not	'-' vs 'generally'	1.711	.024	1.746	.028	1.878	.020	1.871	.037			2.215	0.025
													(2.234 retest)	(.036)
11	not at all/very	generally									5.365	.036	5.717	.112
											(~=1.95)		(~=2.08)	

“Satisfaction with democracy” across CH surveys

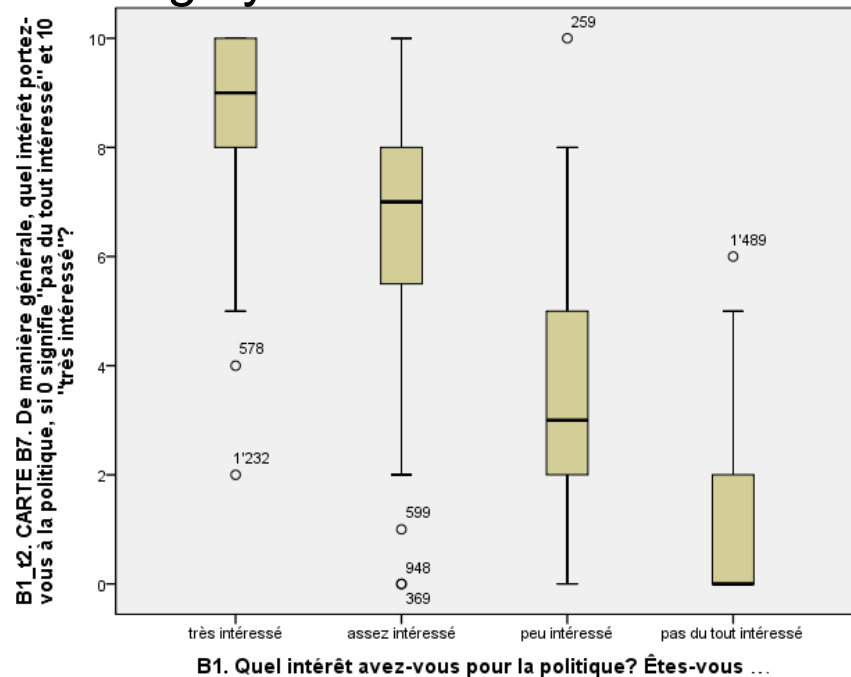
11	not at all / completely VS extr diss/ extr sat	generally , level of sat VS 'on the whole'	7.022	.056							6.171	.024	7.235	.06
			(~=2.55)								(~=2.24)		(~=2.63)	
													(7.058)	(.08)
													(~=2.57))	
4	very/ rather/ rather not/ not at all	globally			2.019	.020	2.004	.015	1.820	.031			1.944	0.022

“Political interest” 4pt scale (ESS) vs 11pt scale (SHP)

Orig by Retest



Orig by SHP version



81% keep same answer
(gamma = .93***)

9% say to be more interested
10% say to be less interested

81% keep «same» answer
(gamma = -.85 ***)

8% say to be more interested
11% say to be less interested

First step: national comparisons

Results:

- Compared and corrected for coverage error
- Compared and corrected for nonresponse bias
- ⇒ Coverage and NR errors are mostly along the same socio-demographic characteristics: we miss foreigners, single households, non-married
- ⇒ But poststrat. weights do not correct or alter the means of the considered two items
- Measurement errors due to:
presence of an interviewer, mode of data collection, question wording, number and labels of answer categ.
- ⇒ **The differences in means in a single country are bigger across surveys than across question versions**

Second step: international comparisons in same survey

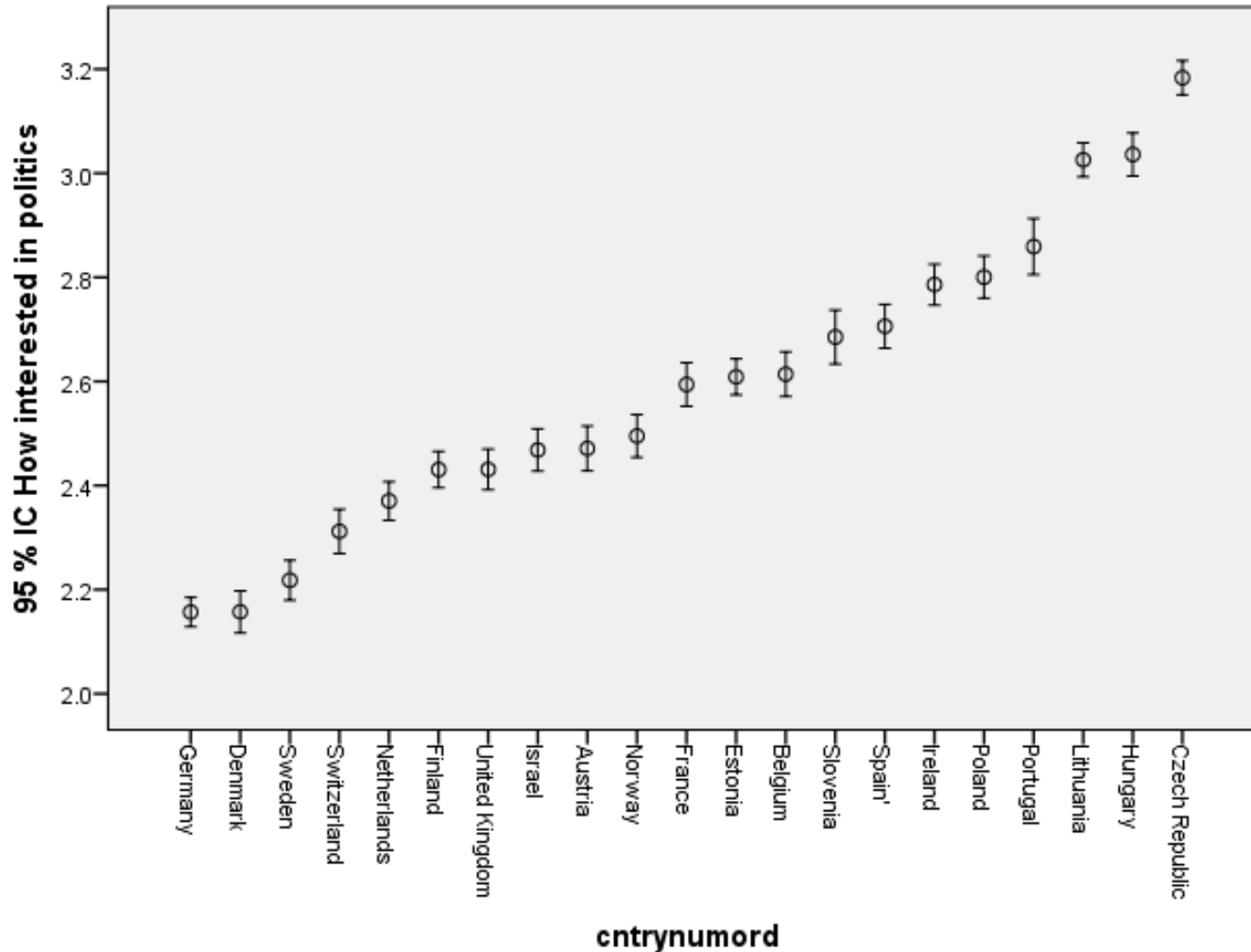
Same question compared in different countries of the same survey (ESS):

- **Concepts:**
Political interest
- **Surveys considered:**
 - **ESS (European Social Survey)**
Round 7 (2014)
21 countries
 - (- ISSP 2014)

Identical question
Identical context (place in questionnaire)
Very similar field procedures

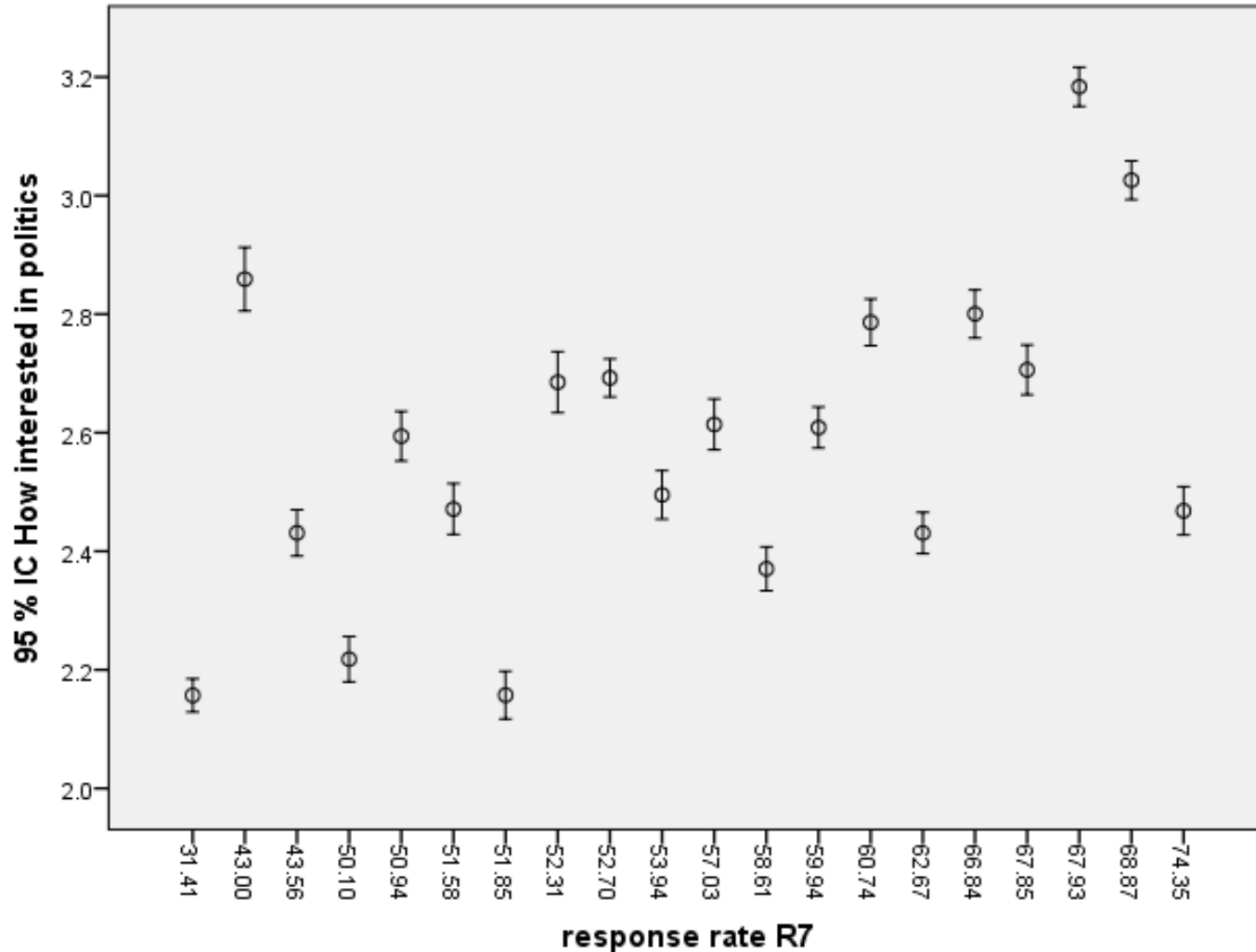
Different sampling frames,
Different translations,
Same mode but CAPI/PAPI
Produced by different institutions

Second step: international comparisons in same survey



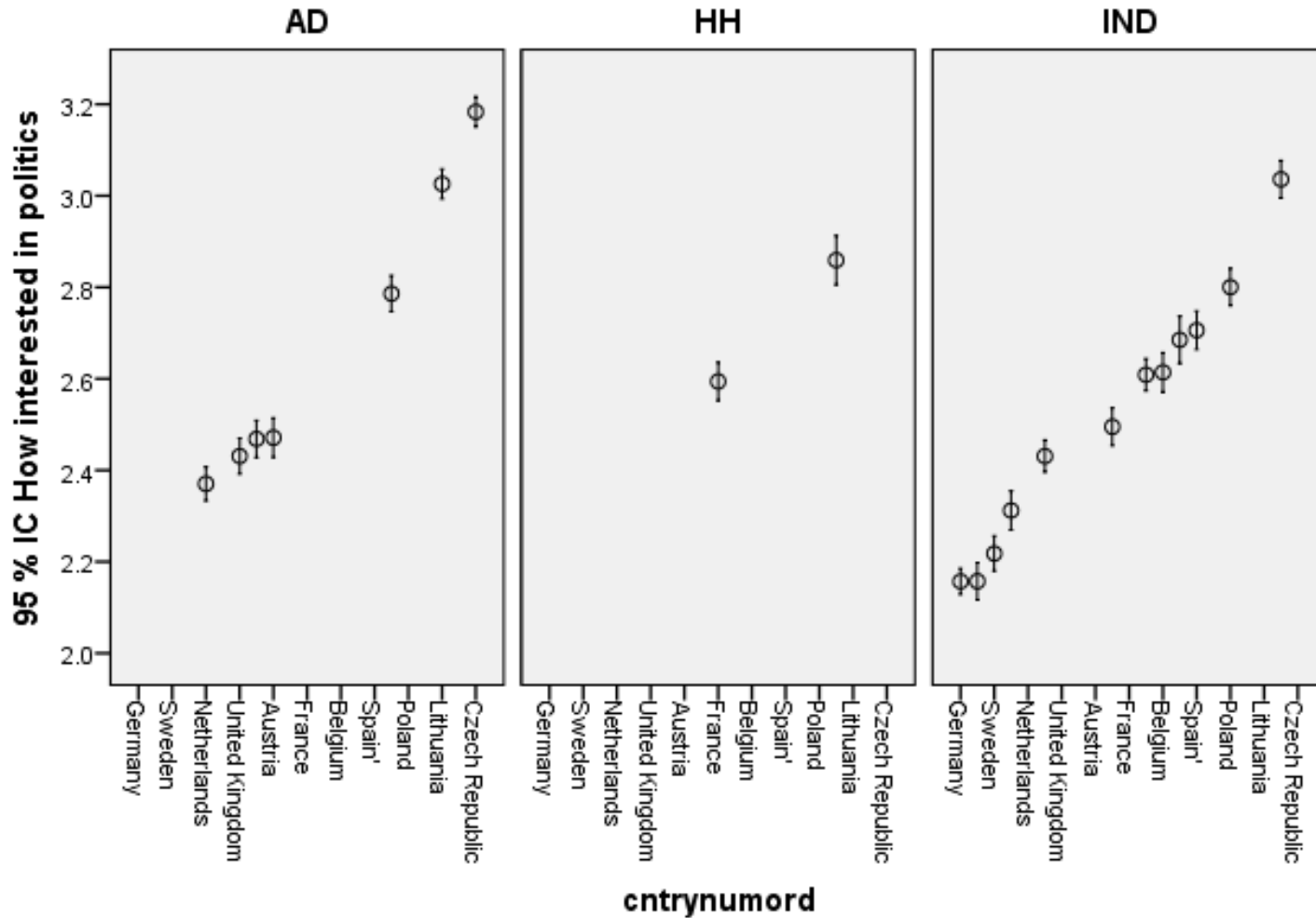
Observations pondérées par Design weight

Second step: international comparisons in same survey



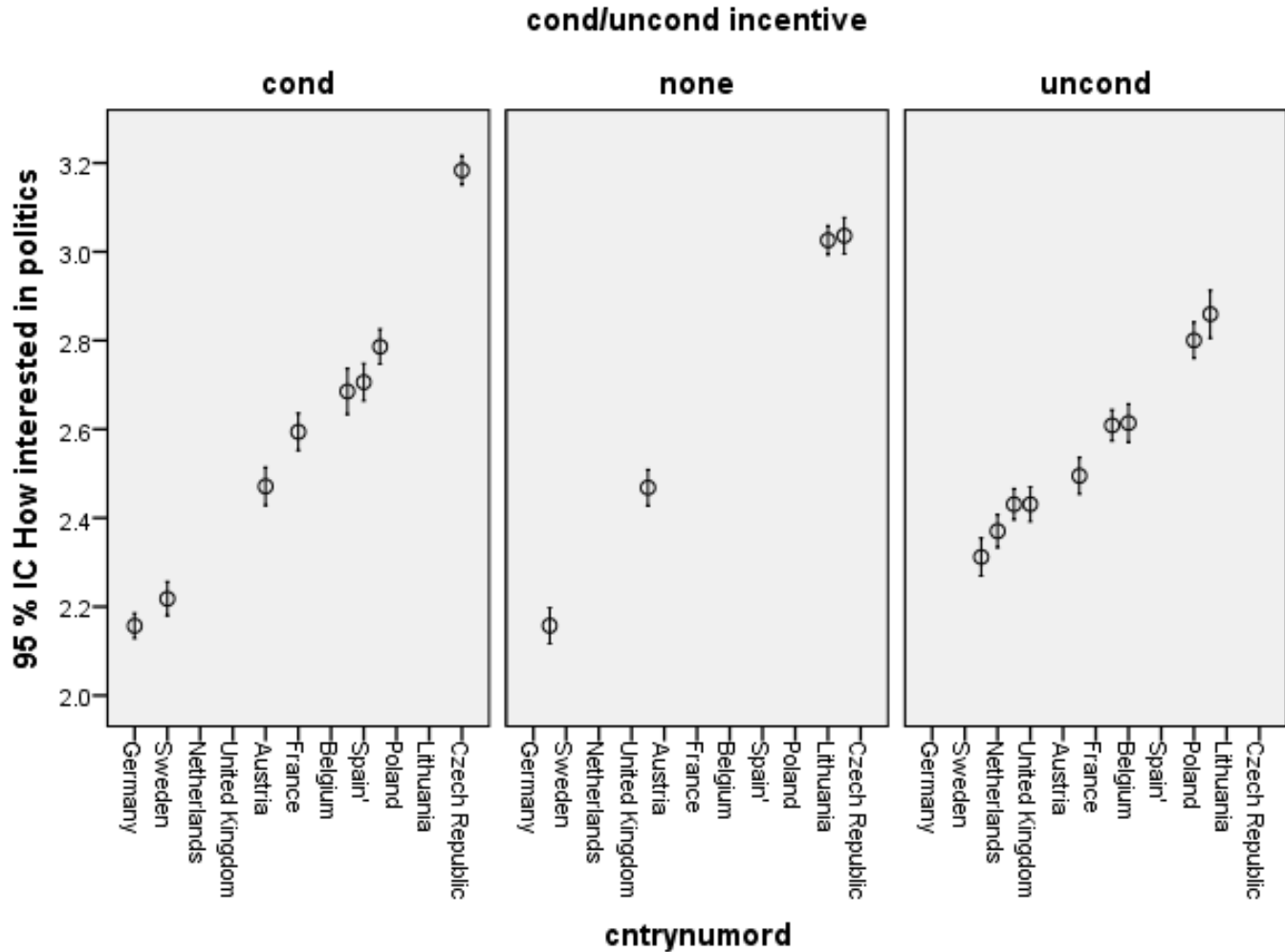
Observations pondérées par Design weight

Second step: international comparisons in same survey



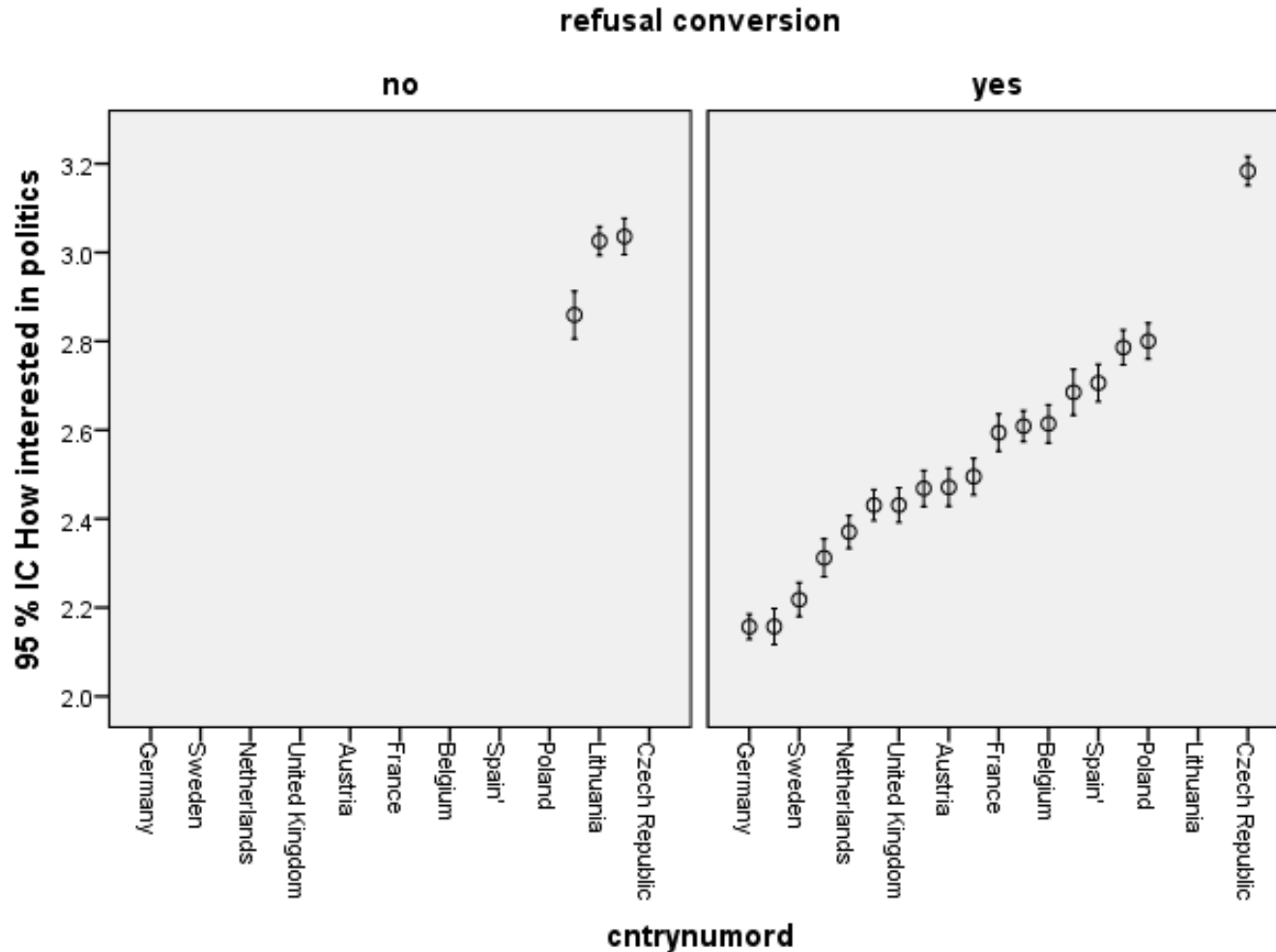
Observations pondérées par Design weight

Second step: international comparisons in same survey



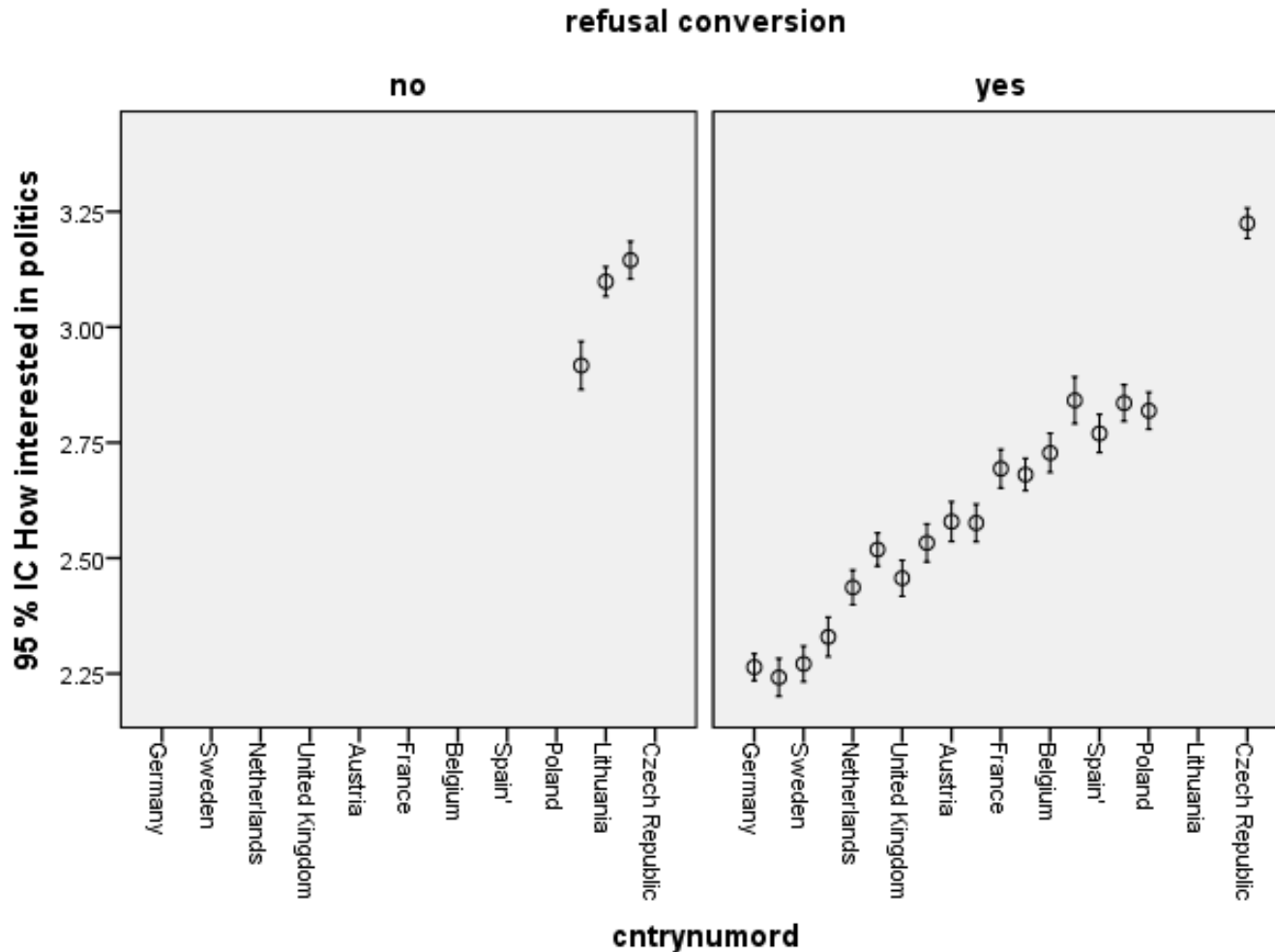
Observations pondérées par Design weight

Second step: international comparisons in same survey



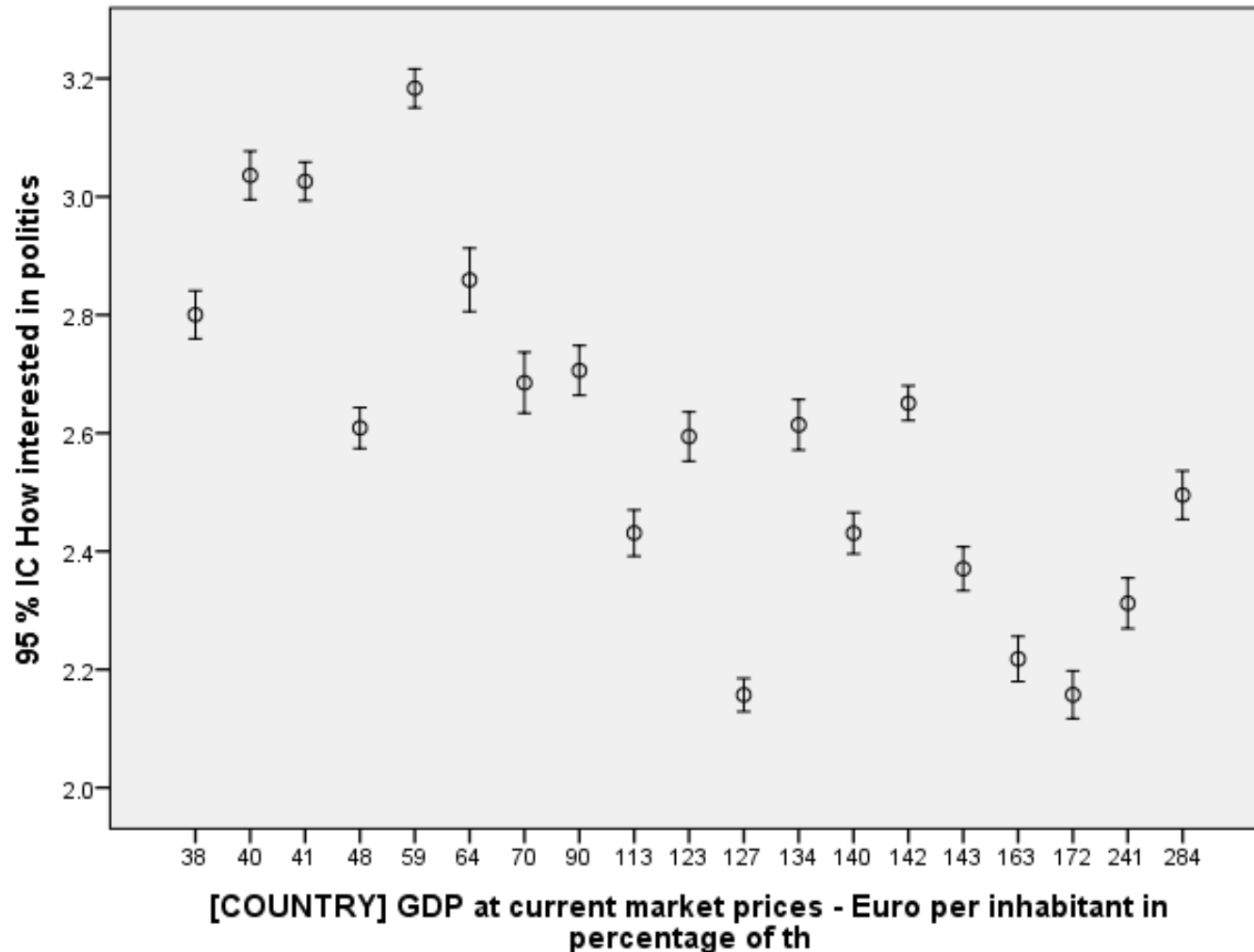
Observations pondérées par Design weight

Second step: international comparisons in same survey



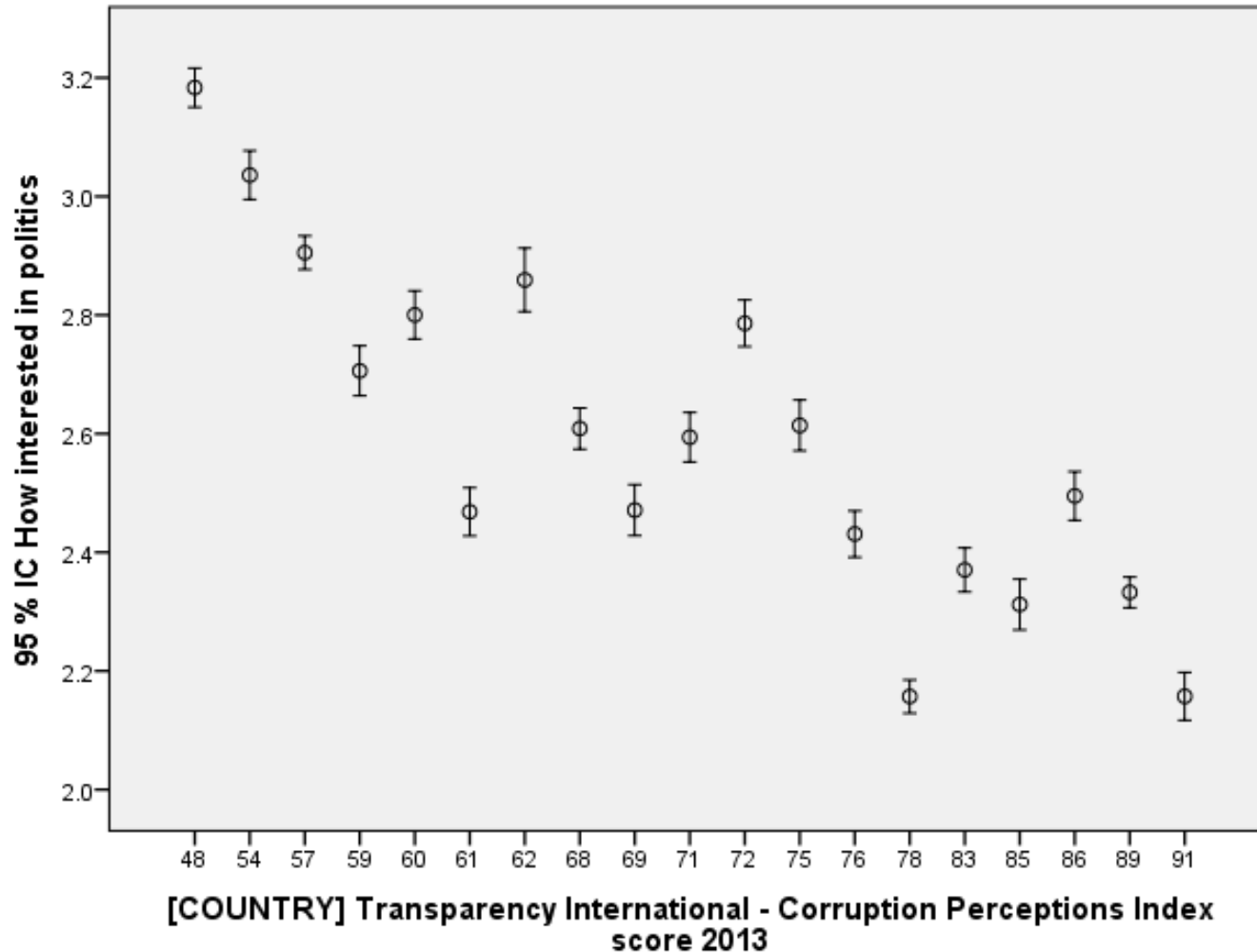
Observations pondérées par Post-stratification weight including design weight

Second step: international comparisons in same survey



Observations pondérées par Design weight

Second step: international comparisons in same survey



Observations pondérées par Design weight

Conclusions

- There are some differences in means across surveys in a single country, even if controlled for SD coverage and NR bias
- These differences are bigger across surveys than across question versions
- (Still not totally disentangled the effect of the different sources of error)
- Additional effects of the way to present the survey?
- In an international context, need to disentangle survey design effects from real country differences, but even apparently slight differences of design could play a role...

More information

- **Links and references:**

Vandenplas, C. & Lipps, O. (2014). Robustness of items within and across surveys. FORS Working Paper Series, paper 2014-3.

Lausanne: FORS.

http://forscenter.ch/wp-content/uploads/2014/06/FORS_WPS_2014-03_Vandenplas.pdf

- **Contact:**

Dr. Michèle Ernst Stähli

Head of unit “International surveys”

FORS, c/o University of Lausanne

Tél. +41 21 692 37 36

Michele.ErnstStaehli@fors.unil.ch

THANK YOU!

And now your questions ...

Overview of central features of the surveys and items

Study	Topic	Periodicity	Target pop	Sample size	Mode	Average response rate	«Political interest» item (backtranslated)	«Satisfaction with democracy» item (backtranslated)
ESS (cross-sect)	General	every 2 years (since 2002)	15+	SRS 1500	FtF	50%	<p>“How interested would you say you are in politics?” 4 point scale, fully labeled “Very interested” “Quite interested” “Hardly interested” “Not interested at all” With showcards</p>	<p>“On the whole, how satisfied are you with the way democracy works in Switzerland?” 11 point scale Endpoint labels: “Extremely dissatisfied” “Extremely satisfied” With showcards</p>
MOSAICH-ISSP (cross-sect)	General	every 2 years (since '97/2005)	18+	SRS 1200	FtF	50%	<p>“How interested would you say you are in politics?” 4 point scale, fully labeled “Very interested” “Quite interested” “Hardly interested” “Not interested at all” With showcards</p>	<p>“On the whole, are you ... satisfied with the way democracy works in Switzerland?” 4 point scale, fully labeled “Very satisfied” “Rather satisfied” “Rather not satisfied” “Not satisfied at all” With showcards</p>
SELECTS (cross-sect)	Electoral politics	every 4 years (since 1995)	18+ Swiss citizens	SRS 4000 - 5000	CATI, (web)	35%	<p>“Generally, how interested are you in politics? Are you...” 4 point scale, fully labeled “Very interested” “Rather interested” “Rather not interested” “Not interested at all”</p>	<p>“In general, are you ... satisfied with the way democracy works in our country?” 4 point scale, fully labeled “Very satisfied” “Rather satisfied” “Rather not satisfied” “Not satisfied at all”</p>
SHP (longitudinal)	General	every year (since 1999)	Private HH, 14+ for individuals	SRS 4000 - 5000 HH	CATI	65% of HH in W1, 85% of individuals within HH, 10% attrition	<p>“Generally, how interested are you in politics?” 11 point scale Endpoint labels: “Not at all interested” “Very interested”</p>	<p>“Generally, what is your level of satisfaction with the way democracy works in our country?” 11 point scale Endpoint labels: “Not at all satisfied” “Completely satisfied”</p>