

Adjusting samples for nonresponse bias: use of non-response surveys in ESS Round 3

Hideko Matsuo, Jaak Billiet & Geert Loosveldt
CeSO - K.U. Leuven

Nederlandstalig Platform voor Survey-Onderzoek

Vlaamse overheid, Ellipsgebouw – auditorium
11 maart 2010

Structure of presentation

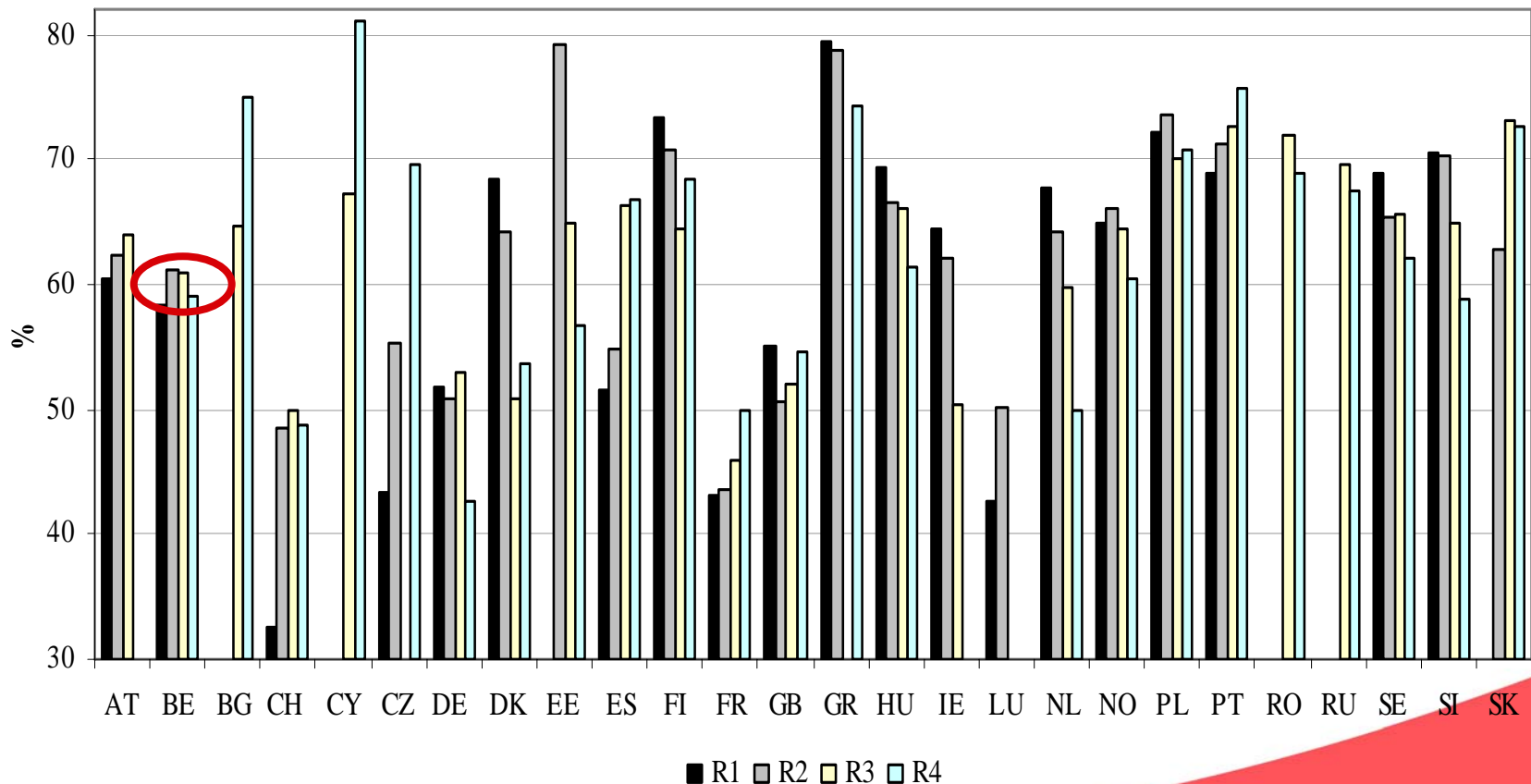
1. Introduction
2. Short Overview
3. Data and method
4. Key results
5. Conclusion & discussion

1. Introduction

Analysis of nr bias still needed:

WHY? Still large differences in RR (NR) rates based on CF R1-R4

Response rates



2. Short overview (1)

- How to deal with nonresponse bias is proposed by Groves (2006) on 5 different methods
- Short overview of approaches to the assessment of bias applied in ESS (*Billiet, Matsuo, Beullens & Vehovar, Research & Methods. ASK. vol 18 (1, 2009), pp. 3-43*).

Within the context of ESS Joint Research Activities (JRA) on nonresponse bias, 4 methods for bias detection and estimation are studied:

In all rounds (R1, R2, R3, R4.....)

1. Bias as deviation between obtained sample and population (or 'Golden standard' survey) = post-stratification and evaluations of samples before and after weighting
2. Bias as difference between cooperative and converted refusals collected via refusal conversion = comparison of cooperative with reluctant respondents (converted refusals)
3. Bias as difference in 'observable' data among all sampling units (collected in contact forms)= sample based comparison between all respondents and all nonrespondents

2. Short overview (2)

In context of R3

4. Bias as difference between respondents and **non-respondents** collected via post hoc nonresponse survey = surveys among nonrespondents after R3
in PL, NO and CH (real NRS)
in BE (at moment of refusal only among refusals = **Doorstep Questions Survey**)

National Co-ordinators* and their respective team members in Non Response Survey

Belgium	Marc Jacquemain*, Geert Loosveldt*, Sebastiane Fontaine, Frederic Herremans, Hideko Matsuo, Katrien Symons, Lee Vandecasteele,
Norway	Kristen Ringdal*, Frode Berglund, Einar Bjørshol, Øyvin Kleven
Poland	Pawel Sztabinski*, Franciszek Sztabinski, Anna Dyjas-Pokorska Teresa Zmijewska-Jedrzejczyk
Switzerland	Dominique Joye*, Alexandre Pollien, Nicole Schöbe

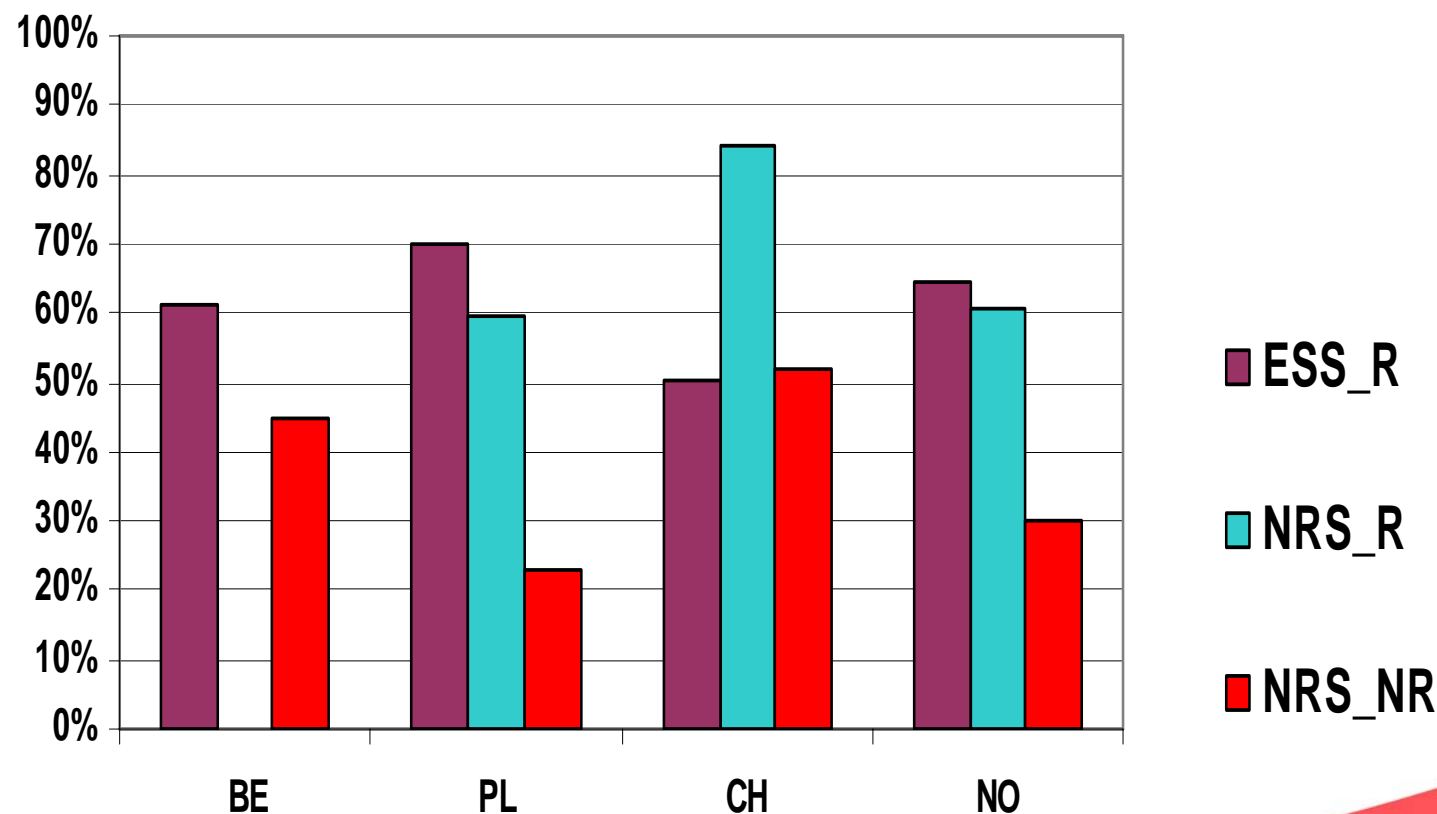
3. Data and method (1)

Survey among nonrespondents

	Target group	Timing	Mode	Use of Incentives	Type of questionnaire	Sample size	Response Rates (R/NR)
BE	ESS3 refusers	Same as ESS	PAPI at door	NO	1 short	303	DQS_R: 44.7%
CH	ESS3_R & ESS3_NR	After ESS	Mail/ Web/ CATI	10 Swiss FR.	2 short & long	1023	NRS3_R: 84% NRS3_NR: 51.8%
NO	ESS3_R & ESS3_NR	After ESS	Mail/ Web/ CATI	NO	1 long	487	NRS3_R: 60.79% NRS3_NR: 30.25%
PL	ESS3_R & ESS3_NR	After ESS	Mail	Notepad	2 short & long+	1208	NRS3_R: 59.04% NRS3_NR: 23.24%

3. Data and method (2)

Response rates between ESS respondents and ESS nonrespondents who participated in NRS



3. Data and method (3)

The questions

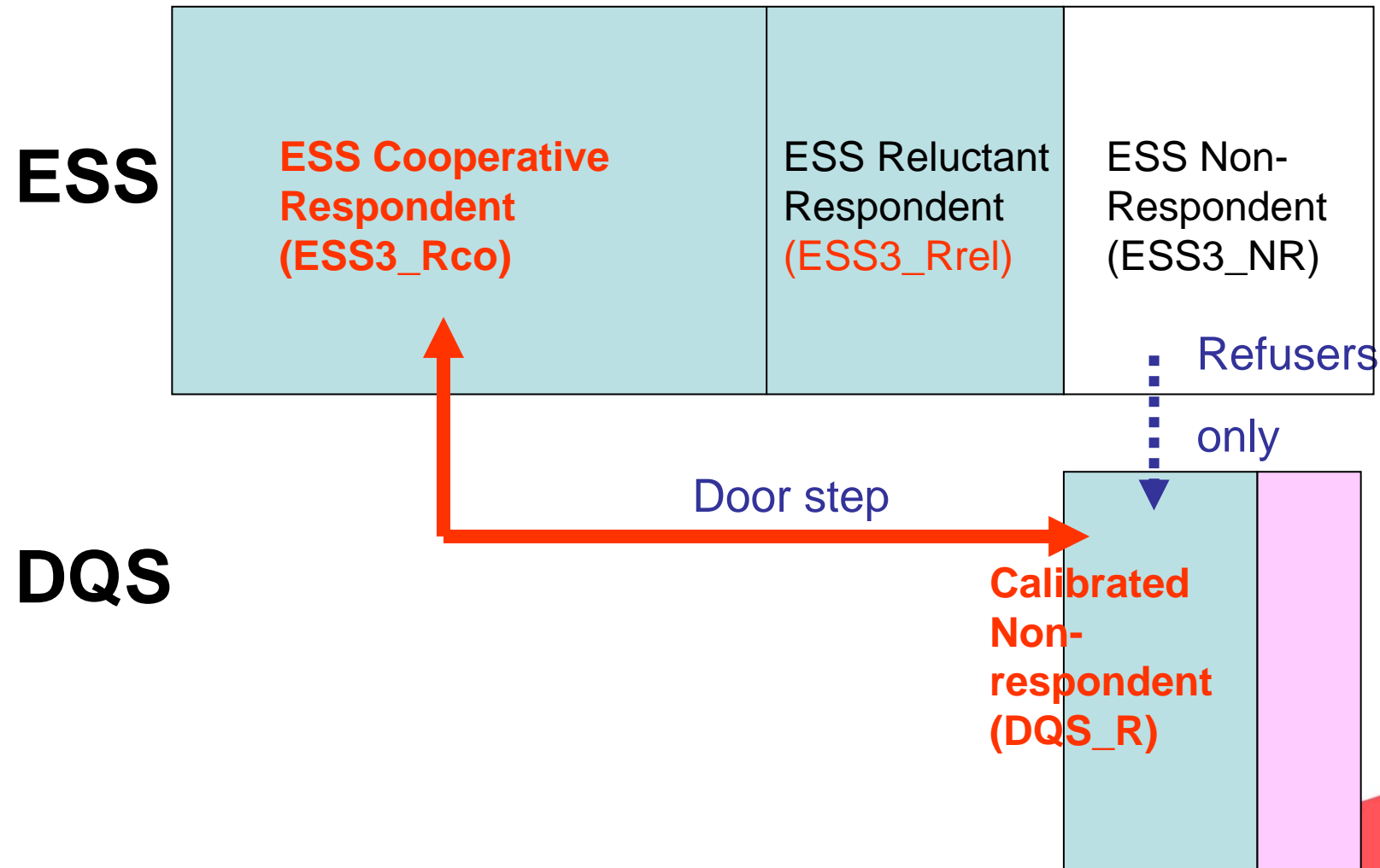
Key questions procedure (*Pedaksi approach*)

Short 7 question module (+ at door): work situation, highest level of education, # of members in household, frequency of social activities, feeling (un)safe, interest in politics, attitude towards surveys

Normal 16 questions module: same as short + gender, year of birth, TV watching, voluntary work, trust in people, satisfied with democracy, trust in politicians, immigration good/worse for country, (+ reasons for refusal (closed) in one subgroup)

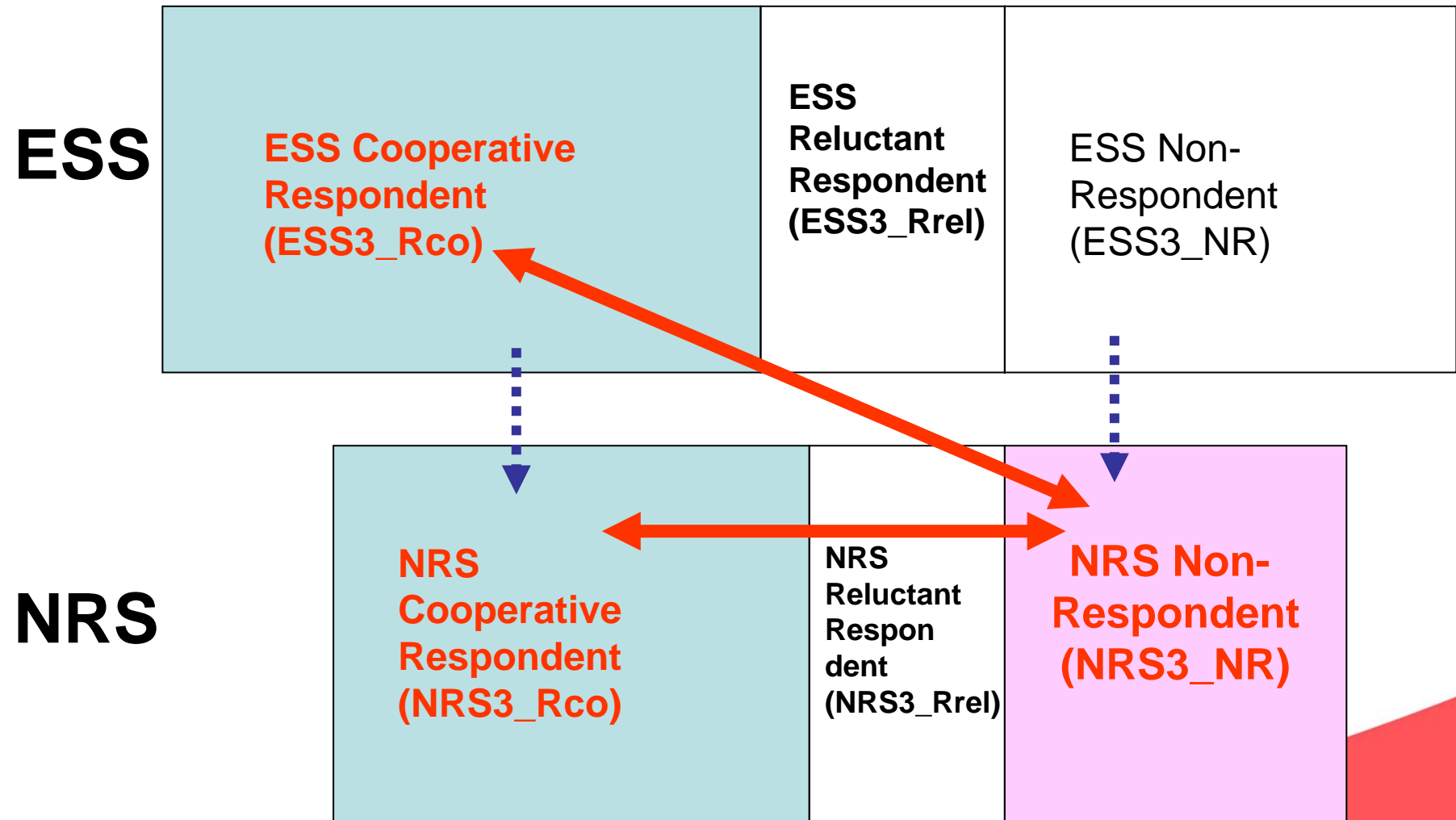
3. Data and method (4)

Kinds of respondents in data analyses [BE]



3. Data and method (5)

Kinds of respondents in data analyses [NO, CH & PL]



3. Data and method (6)

Kinds of respondents in NRS/DQS

decisions to take in view of computing propensity scores for weighting the sample

NRS/(cooperative vs. nrs)

NRS/(cooperative vs. main)

(NRS+reluctant) vs (cooperative (nrs or main?))

NRS/cooperative vs reluctant/cooperative

3. Data and method (7)

Method used for adjusting the sample for nonresponse bias

1. **Identify** survey **response differences** on key explanatory variables between types of respondent (*'nonrespondent vs. cooperative respondent'*).
2. Study **net effects** of key explanatory variables on **response probabilities** via logistic regression model (*dependent variable: prob ratio's 'nonrespondent/cooperative'*).
$$\log \left[e(\mathbf{x}) / (1 - e(\mathbf{x})) \right] = \alpha + \beta' f(\mathbf{x})$$
3. Obtain **propensity scores** on all cases on non-response probabilities via logistic regression model (*dependent variable: prob ratio's 'cooperative/nonrespondent'*).

4. Transform propensity scores into **weights via stratification method** (*Rosenbaum & Rubin 1984; Little 1986; Lee & Vaillant 2008*):
 - ✓ Form 10 strata with equal number of cases after sorting on ps;
 - ✓ Assign each sample unit into correct corresponding sub-strata
 - ✓ $\text{Weight} = \frac{\text{expected probability}}{\text{observed probability}}$ of the coop. respondent (or nonrespondent) in the corresponding sub-strata.

5. **Evaluate effects of propensity score weighting** via two main criteria:

- ✓ Tests between unweighted & weighted sample on cooperative respondents (NRS3_Rco & ESS3_Rco).
1b. In case of significant differences: test differences between parameters of relevant substantive explanatory models
- ✓ Study differences in distributions on key questions between types of respondents (*NRS3_Rco vs NRS3_NR or ESS3_Rco vs. NRS3_NR*).

4. Key Results(1): Differences (ESS coop & nr)

	Belgium		Norway	
	ESS coop (N=1658)	DQS res (N=303)	ESS coop(N=1646)	NRS nr (N=242)
(Age)				
14-29	23.40	8.25	22.24	17.77
30-39	16.04	21.45	17.25	14.05
40-49	20.69	16.83	21.26	21.07
50-59	16.22	18.81	17.07	19.83
60+	23.64	34.65	22.17	27.27
<i>Chi² ; df = 4; prob.**</i>		47.544; <i>p < 0.0001</i>		7.534, <i>p=0.110</i>
(Gender)				
Male	46.68	47.52	51.64	46.28
Female	53.32	52.48	48.36	53.72
<i>Chi² ; df = 1; prob.**</i>		0.073; <i>p=0.787</i>		2.783; <i>p=0.095</i>
(Educational level)^[1]				
Lower (basic)	25.41	33.92	(1+2)	(1+2)
Lower sec (hum.)	10.56	9.19	17.77	34.30
Higher secondary	35.67	33.22	35.61	34.71
Higher tertiary	28.36	23.67	46.62	30.99
<i>Chi² ; df = 3; prob.**</i>		9.349; <i>p=0.025</i>		49.920; <i>p<0.0001</i>
(Work status)				
Employed	49.76	49.65	67.60	56.43
Unemployed	50.24	50.35	32.40	43.57
<i>Chi² ; df=1; prob.**</i>		0.034; <i>p=0.853</i>		13.728; <i>p=0.0002</i>
(HH composition)				
One person household	11.59	15.57	19.10	19.50
More person household	88.41	84.43	80.90	80.50
<i>Chi² ; df=1; prob.**</i>		3.651; <i>p=0.056</i>		0.025; <i>p=0.874</i>

4. Key Results (2): Differences (ESS coop & nr)

	Belgium		Norway	
	ESS coop(N=1658)	DQS res (N=303)	ESScoop(N=1646)	NRS nr (N=242)
<i>(Neighbourhood security)</i>				
Very safe				
Safe	19.84	18.18	50.91	45.00
Unsafe	60.07	61.89	40.83	44.17
Very safe	17.06	16.78	6.01	8.33
	3.02	3.15	2.25	2.50
<i>Chi² ; df =3; prob.**</i>		0.499; p=0.919		4.517; p=0.211
<i>Means/SD</i>	<u>2.033/0.700</u>	<u>2.049/0.689</u>	<u>1.596/0.704</u>	<u>1.683/0.732;</u>
<i>T-value; df; prob.**</i>		0.36;df=1937;p=0.716		-1.79;df=1884;p=0.074
<i>(Social participation)</i>				
Much less than most	15.30	27.27	3.28	12.45
Less than most	29.20	24.36	17.69	17.43
About the same	35.25	37.09	61.16	60.17
More than most	15.90	6.55	16.05	8.30
Much more than most	4.35	4.73	1.82	1.66
<i>Chi² ; df = 4; prob.**</i>		36.294; p<0.0001		70.776; p<0.0001
<i>Means/SD</i>	<u>2.648/1.055</u>	<u>2.371/1.094</u>	<u>2.954/0.735</u>	<u>2.693; 0.855;</u>
<i>T-value; df; prob.**</i>		-4.01; df=1927; p<0.0001		5.05; df=1884; p<0.0001
<i>(Political interest)</i>				
Very interested	8.75	4.55	9.72	3.32
Quite interested	37.15	19.58	39.00	31.12
Hardly interested	33.17	31.12	44.29	50.21
Not at all interested	20.93	44.76	6.99	15.35
<i>Chi² ; df = 3; prob.**</i>		84.070; p<0.0001		40.047; p<0.0001
<i>Means/SD</i>	<u>2.663/0.904</u>	<u>3.161/0.896</u>	<u>2.485/0.764</u>	<u>2.776/0.741</u>
<i>T-value; df; prob.**</i>		8.61;df=1942;p<0.0001		-5.53; df=1885; p<0.0001

4. Key Results (3): Differences in NO

	ESS res.(N=1646)	NRS nr.(N=242)
How satis. Democracy works (0: dissatisfy- 10: satisfy) Chi^2 ; $df =10$; $prob.$ ** <u>Means/SD</u> <i>t-value; df; prob.**</i>	<u>6.632/1.941</u>	70.060; $p < 0.0001$ <u>5.907/2.119</u> 5.31; $df=1864$; $p < 0.0001$
Trust in politicians (0: No trust-10: complete trust) Chi^2 ; $df =10$; $prob.$ ** <u>Means/SD</u> <i>t-value; df; prob.**</i>	<u>4.457/1.997</u>	26.044; $p=0.004$ <u>4.261/2.226</u> 1.40; $df=1876$; $p=0.162$
Imm. Make country worse/better place to live (0: worse – 10: better) Chi^2 ; $df =10$; $prob.$ ** <u>Means/SD</u> <i>t-value; df; prob.**</i>	<u>5.117/2.040</u>	119.682; $p < 0.0001$ <u>4.356/2.558</u> 5.17; $df=1871$; $p < 0.0001$
TV watching time per day (0: No time – 7: more than 3 hours) Chi^2 ; $df =7$; $prob.$ ** <u>Means/SD</u> <i>t-value; df; prob.**</i>	<u>3.709/1.764</u>	27.349; $p=0.0003$ <u>4.248/1.750</u> -4.38; $df=1878$; $p < 0.0001$
Involved in work for vol. and charity org. (0: at least once a week – 6: never) Chi^2 ; $df =5$; $prob.$ ** <u>Means/SD</u> <i>t-value; df; prob.**</i>	<u>4.211/1.735</u>	14.634; $p=0.012$ <u>4.531/1.625</u> -2.69; $df=1885$; $p=0.007$
Most people trusted /can't be too careful (0: can't be too careful – 10: most trusted) Chi^2 ; $df =10$; $prob.$ ** <u>Means/SD</u> <i>t-value; df; prob.**</i>	<u>6.844/1.805</u>	59.173; $p < 0.0001$ <u>6.600/2.250</u> 1.89; $df=1884$; $p=0.058$

4. Key Results(4): Summary in 4 countries

	BE	CH	NO	PL
Age	***	N.A.	N.S.	N.A.
Gender	N.S.	N.A.	N.S.	N.S.
Education level	*	N.S.	***	**
Work status	N.S.	***	**	*
Household composition	N.S.	**	N.S.	N.S.
Neighborhood security	N.S./N.S.	**/**	N.S./N.S.	**/*
Social participation	***/**	***/**	***/**	***/**
Political interest	***/**	**/*	***/**	***/**

(X²/T-tests): NA: not applicable; NS: not significant; *** p<.0001;
 **p<.01; *p<.05

4. Key Results (5): BE

Logistic regression model: nr/cooperative

R ² =0.055 / H&L= 10.979		Odds ratio's
Age	30-39	1.451*
	40-49	0.836
	50-59	1.199
	60+	1.793**
	<i>Ref: 14-29</i>	
Education	Lower secondary	0.801
	Higher secondary	1.002
	Higher tertiary	0.917
	<i>Ref: Lower basic educ.</i>	
Work	employed <i>Ref: unemployed</i>	1.205
Household composition	One <i>Ref: more than one.</i>	1.113
Social participation	Almost the same	1.223
	More than most/much more than most <i>Ref: much less than most/less than most</i>	0.722*
Political interest	Hardly/not at all interested <i>Ref: very /quite interested</i>	1.574***

4. Key Results (7): CH/PL

Logistic regression model: nr/cooperative

	CH Odds ratio's R ² =0.036 / H&L= 3.907	PL Odds ratio's R ² =0.042 / H&L= 3.8355
Education Lower secondary Higher secondary Higher tertiary <i>Ref: Lower basic educ.</i>		0.894 1.365* 1.772**
Work employed <i>Ref: unemployed</i>	0.860**	0.742**
Household composition One <i>Ref: more than one.</i>	0.777**	
Feeling unsafe after dark Safe Unsafe&very unsafe <i>Ref: very safe</i>	1.151* 0.961	1.335* 1.344*
Social participation Almost the same More than most Much more than most <i>Ref: much less t'n most/less t,n most</i>	0.793** 0.760** 0.834	(2 categories) 0.548*** 1.287
Political interest Hardly/not at all interested <i>Ref: very /quite interested</i>	1.099*	0.895

4. Key Results (8)

Evaluation of the propensity score weighting

Approach 1: Is the weighted sample of ESS coop res significantly different from the unweighted sample?

Yes means that the adjustment had effect on sample estimates

Results: No significant differences ($p < 0.05$) at all.

4. Key Results (9)

Evaluation of the propensity score weighting

Approach 2: does the initial significant differences by type of respondents in the unweighted sample disappear in the weighted sample?

(NMAR to MAR?)

Yes means that the adjustment had effect on sample estimates

Results: Largely successful: all differences disappeared except for some variables (**look further**)

4. Key Results (10)

Evaluation of the propensity score weighting Belgium example: ESS coop vs DQS res

	Unweighted		Weighted	
	Chi ²	prob	Chi ²	prob
Age categories (df=4)	47.544	<.0001	2.008	0.734
Educational level (df=3)	9.349	0.025	3.887	0.274
Social participation (df=4)	36.295	<.0001	17.912	<.0001
Political interest (df=3)	84.070	<.0001	25.938	<.0001

all differences disappeared

except social participation & political interest

4. Key Results (11)

Evaluation of the propensity score weighting Norwegian example 1: ESS coop vs NRS nr

	Unweighted		Weighted	
	Chi ²	prob	Chi ²	prob
Education category (df=2)	40.552	<.0001	1.813	0.404
Work status (df=1)	11.594	0.0007	0.470	0.493
Political interest (df=3)	33.014	<.0001	13.247	0.010
Social participation (df=4)	48.105	<.0001	2.301	0.681

all differences disappeared
except political interest

4. Key Results (12)

Evaluation of the propensity weights

Norwegian example 2: ESS coop vs NRS nr

	Unweighted		Weighted	
	T-value	prob	T-value	prob
How satisfied democracy (df=1864)	5.31	<.0001	1.01	0.313
Imm. make worse/better (df=1871)	5.17	<.0001	0.56	0.577
TV watching time per day (df=1878)	-4.38	<.0001	-0.81	0.420
Involved in voluntary work (df=1885)	-2.69	0.007	-0.69	0.492

all differences disappeared

4. Key Results (13)

Evaluation of the propensity score weighting on approach 2: summary in all countries

	BE	CH	NO	PL
Age categories	N.S.	N.A.	N.A.	N.A.
Education level	N.S.	N.S.	N.S.	N.S.
Work status	N.A.	N.S.	N.S.	N.S.
Neighb. security	N.A.	N.S.	N.A.	N.S.
Household comp.	N.A.	N.S.	N.A.	N.A.
Social participation	**	N.S.	N.S.	***
Political interest	***	N.S.	*	***

NA: not applicable; NS: not significant; *** $p < .0001$; ** $p < .01$; * $p < .05$

4. Conclusion & discussion (1)

Conclusion from the study

1. Number of variables are detected for non-response bias (education/work/social participation/political interest);
2. Characteristics of nonrespondents (vs. cooperative respondent) are country specific (PL case);
3. Adjustments of propensity score weighting were mostly effective (approach 2); BUT some not;

4. Conclusion & discussion (2)

Discussion

1. Response propensities may take different form;
2. Evaluating weights were performed through 2 approaches,
 - a. Rather small adjustments: small nonresponse bias or small NRS sample size;
 - b. Different approaches to evaluate effectiveness of weights can be explored: use of variables from contact file – it is crucial to study which information to use
3. Future research
 - a. Follow-up NRS/DQS study may be possible through standardized sample design (cost-effective?);
 - b. Use of observable data – data quality enhancement proposed for Round 5.