

Quality report for the European Social Survey, Round 6*

Koen Beullens, Hideko Matsuo, Geert Loosveldt, and Caroline Vandenplas

Centre for Sociological Research, Leuven



*Refer to this report as “Beullens, K. Matsuo, H., Loosveldt, G., & Vandenplas, C., (2014) Quality report for the European Social Survey, Round 6. London: European Social Survey ERIC”

Contents

1	Introduction	3
2	Fieldwork procedures	3
2.1	Synoptic picture of the fieldwork	3
2.2	Assigning interviewers to the field	7
2.3	Nonresponse codes	11
2.4	Compliance with noncontact specifications	12
3	Fieldwork results	15
3.1	Ineligibles	15
3.2	Response rates	17
3.3	Noncontact rates	18
3.4	Refusal rates and refusal conversion	19
3.5	Indications of nonresponse bias	20
4	The actual interview	28
4.1	Interviewer variance	28
4.2	Interviewer effects on interview length	30
4.3	Item nonresponse	31
4.4	Straightlining	35
	References	38

1 Introduction

The purpose of this document is to develop a fieldwork and data quality profile for each participating country in ESS round 6. The total survey error framework (TSE) offers a good framework for this kind of evaluation, identifying sampling, nonresponse, measurement and data processing as possible sources of error (Biemer & Lyberg, 2003) and focusing on the accuracy and the precision of the data as quality criteria.

The current report mainly focusses on the fieldwork process (compliance to contact procedure guidelines and refusal conversion) and output (response, contact, refusal and ineligible rates, as well as indications of nonresponse bias) enriched with some indicators of the quality of the actual interview (interview length, interviewer variance, item non-response and straightlining). This means that not all the sources of the total survey error are discussed. The equivalent and comprehensive report for future ESS rounds should cover all or at least more aspects of the survey life cycle: from translation and sampling to data cleaning and processing. This extension is necessary to assess the overall quality of the produced data.

In this general overview, fieldwork and data quality indicators for all countries participating in the ESS round 6 will be displayed together to permit countries to situate themselves compared to other participating countries.

2 Fieldwork procedures

This section will predominantly focus on the contact phase which aims to establish contact with the sample units and eventually make them participate in the survey. The contact form data are currently the most important data source to explore the ESS fieldwork quality. The first section will deal with the process part of the fieldwork (section 2.1, 2.2, 2.3, 2.4), whereas the subsequent section (section 3) deals with the results of these fieldwork processes. Particularly response rates, noncontact rates and refusal rates will be presented, followed by an assessment of nonresponse bias based on information from interviewer observations.

2.1 Synoptic picture of the fieldwork

We start showing a set of graphs that provide a first overview of the fieldwork of the different ESS6 countries. The graphs indicate when the particular fieldwork took place, how long it lasted and how extensive the fieldwork was. It also informs about when new addresses are allotted to the field and the degree to which follow-up contact attempts have been carried out.

The Figures 1 and 2, provide an overview of the fieldwork courses in the different ESS6 countries. All fieldwork actions take place between August 3, 2012 and December 16, 2013. For each country and for each fieldwork day, a vertical bar is presented, indicating the sum of contact attempts that have been done on that particular day. Red bars indicate Sundays, all other days of the week are in grey. The lighter shades of grey/red represent first contact attempt with sample cases, the darker shades indicate follow-up contact attempts. That way, it can be easily seen when new addresses have been released or first contacted.

The different graphs suggest that ESS fieldwork is not particularly a uniform process among the participating countries. Some countries start early (e.g. Belgium, Switzer-

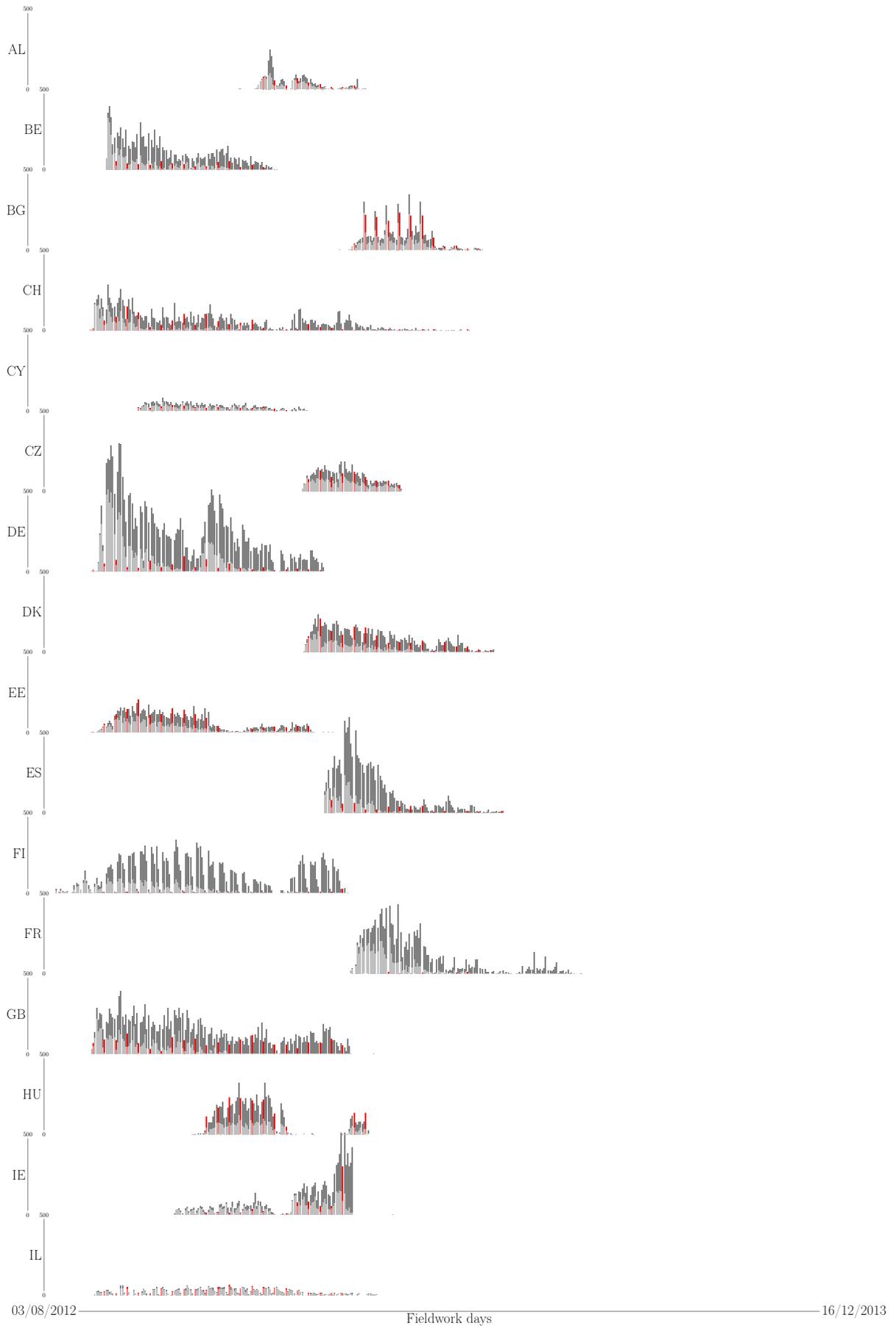


Figure 1: Synoptic picture of the fieldwork (a)

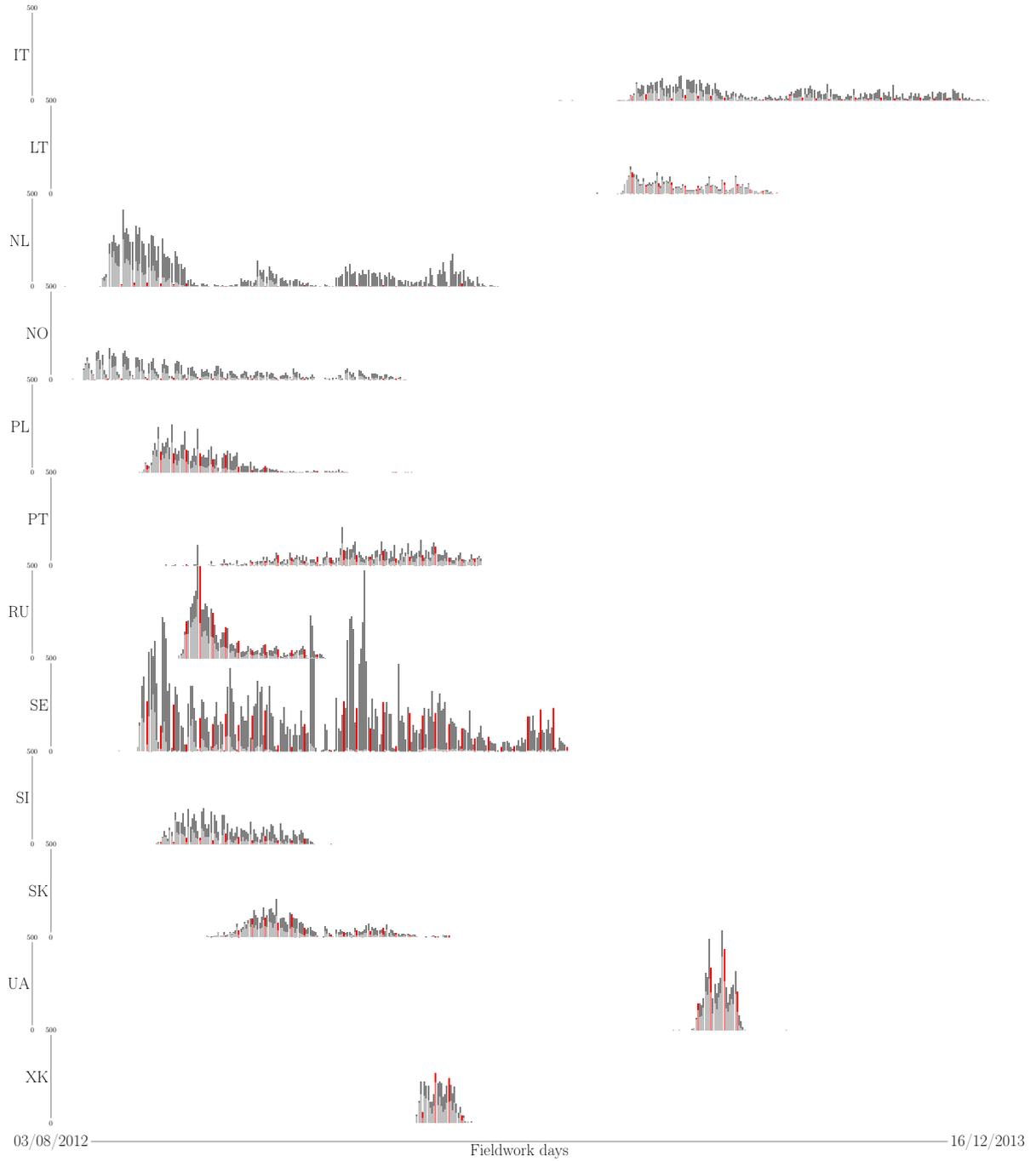


Figure 2: Synoptic picture of the fieldwork (b)

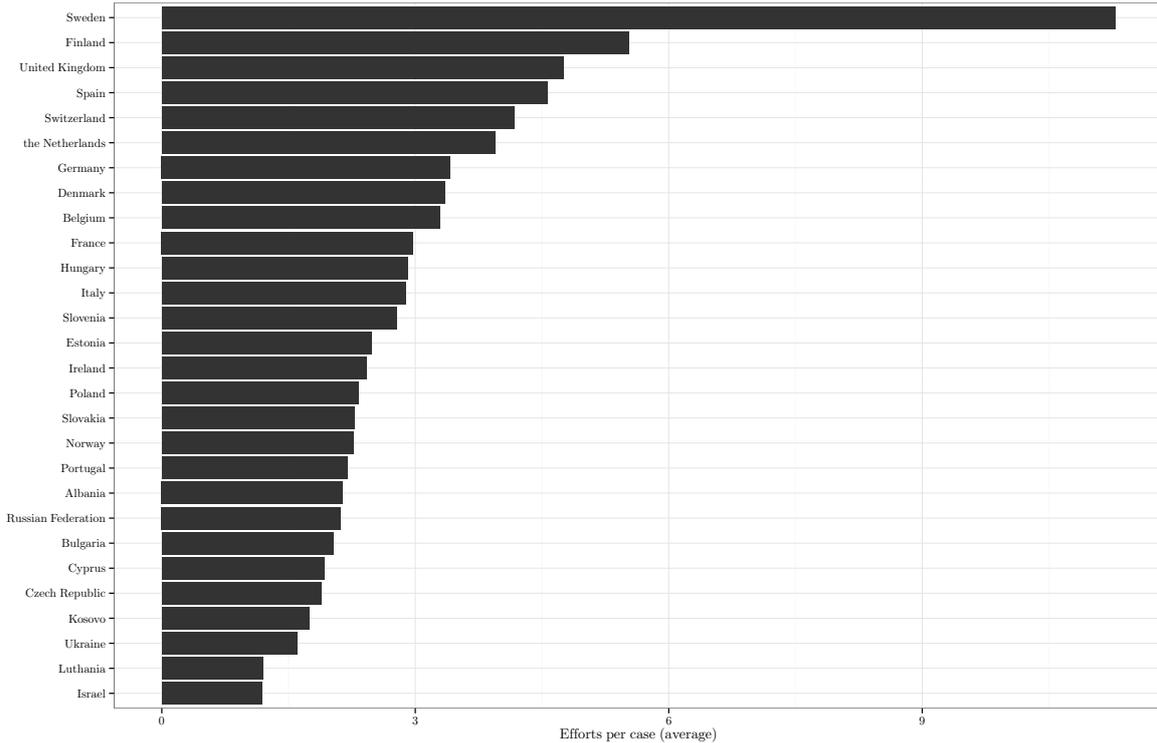


Figure 3: Average efforts per case, 28 ESS6 countries

land, Germany, Estonia, Finland, United Kingdom, Israel, the Netherlands, Norway and Sweden), while other countries only start relatively late. These are countries such as Bulgaria, Czech Republic, Spain or Kosovo. The length of the fieldwork is also a distinctive feature. Switzerland, Germany, Finland, United Kingdom, Israel, the Netherlands, Norway and Sweden seem to have a long fieldwork period as compared to fast countries such as Bulgaria, Czech Republic, Kosovo or Ukraine.

Efforts during fieldwork strongly varies between the countries. In Bulgaria, Cyprus, Czech Republic, Estonia, Hungary, Israel, Portugal, Slovenia, Slovakia and Kosovo, the total amount of fieldwork efforts is relatively modest, either attributable to (a combination of) short fieldwork courses or relatively few contact attempts per day. In other countries such as Belgium, Germany, Switzerland, Spain, Finland, Sweden or the United Kingdom, interviewer efforts appear to be higher. These differences might partially be explained by the fact that some countries like Germany or the United Kingdom used a larger sample than small countries (e.g. Cyprus). Other particular reasons for these difference may relate to the way contact forms are filled out. For example, many telephone calls are documented in Sweden, whereas many other countries did not particularly (or not extensively) make use of telephone as a means to contact the sample units. Another interesting explanation for these differences might pertain to the fact that some countries needed to do much more efforts to attain a satisfactory (or reasonably high) response rate, due to the unfavourable survey climate in their country. Possibly, there will be a relationship between extensive fieldwork efforts (containing a lot of dark grey follow-up contacts), and the level of response rates, where, somewhat counter-intuitive, low response rates relate to much fieldwork efforts. An overview of the contact efforts per over 28 ESS6 participating countries can be found in Figure 3

In some countries, a weekly or short-wave structure can be observed in the graphs. For example, Bulgaria seems to have done a lot of efforts during the weekend (high red bars and high grey bars preceding the red bars), other countries on the other hand have avoided weekend days. This can clearly be observed in countries such as Spain, Finland or Norway. Also, long-wave patterns can be observed in the field pictures. A good example are the Netherlands, where fieldwork episodes of rather extensive efforts are alternated with periods of inactivity. Such rest periods may be explained by holidays such as Christmas. Another possibility is that the fieldwork sometimes needs a period to reflect on the current state of affairs, in order to adapt the fieldwork strategy. Such fieldwork adaptations may imply a reshuffle of the interviewer force (e.g. sending more experienced interviewer to reluctant cases), releasing new addresses, call-back checks, the introduction of new incentives, It would be interesting to have some feedback from the different NC's or local fieldwork managers as to why such hold period took place, possibly learning from their experience and insights or to assess the results of the (adapted) fieldwork tactics.

A last point that may draw some attention is the distribution of first contact attempts over the fieldwork. Usually, one may expect that all addresses are released at the beginning of the fieldwork. As a result, the first contact attempts (lighter grey bar) should also be in the beginning of the field process. In some countries, however, first contact attempts also appear even when the fieldwork has been going on for a while. In some instances, these initial attempts even appear after a period of relative inactivity. Clear examples can be found in Germany, Ireland and the Netherlands (particularly during the second wave of contact attempts). This phenomenon may be explained by the fact that in some countries in some regions the fieldwork started somewhat later than elsewhere. Alternatively, the field may also have been provided with more addresses. Another related fieldwork strategy consists of keeping some addresses for later, when interviewers are asked to do (refusal) conversion. New (unapproached) addressed may motivate interviewers to still stay in the field.

2.2 Assigning interviewers to the field

Figure 4 shows how the interviewer force in each ESS country has been deployed during the fieldwork. In order to explain how the different graphs should be interpreted, consider the case of Belgium (BE). Each horizontal line in the graph (probably consisting of many segments of different shades of grey) represents one interviewer. Each long rectangular segment represents a period of two weeks in the fieldwork. The darker the segment, the more contact attempts the interviewer have been doing for that period of two weeks. Per country, the interviewers have been sorted by the 'average date of all their contact attempts'. The interviewers at the top of the graph have usually concentrated their activities at start of the fieldwork, whereas the interviewers at the bottom of the graph have usually concentrated their activities at the end of the fieldwork. This can be seen fairly well in the case of Israel (IL). In the Netherlands, on the contrary, most interviewer seem to be working continually, without clearly visible concentrations of efforts during specific episodes of the fieldwork. The numbers shown next to each country code refer to the number of interviewers deployed during the fieldwork.

A first, rather trivial, observation about this figure is the relatively strong relationship between the number of deployed interviewers on the one hand and the efforts done per interviewer. Countries deploying a relatively low number of interviewer in the ESS field-

work such as Cyprus (CY), Spain (ES), Slovenia (SI) or Kosovo (XK) clearly show more dark segments on the graphs, suggesting extensive efforts per interviewer as compared to countries deploying many interviewers such as Russia (RU), Bulgaria (BG). Sweden (SE) is in this respect a somewhat outlying country as it counts extremely many efforts per interviewer, as suggested by Table 1. This is probably due to the extensive use of contact attempts by telephone. Nevertheless, some countries may choose to assign many interviewers who only offer a restricted amount of efforts or assign relatively few interviewers combined with extensive efforts per interviewer. In some countries such as Germany (DE) or the United Kingdom (GB), a combination of a large interviewer force and many efforts per interviewer might have been opted for because of a large gross sample. Similarly, some countries may also have chosen a similar combination because of a low expected response rate.

Looking at Figure 4, it seems that countries or their respective fieldwork management may have implemented interviewers relatively differently into the field. In general, and as working hypothesis, four different interviewer assignment strategies may be observed. The first one is strongly reflected in the Israeli graph. Interviewers seem to choose their own planning during the course of the fieldwork. Some start and end their activities relatively early, whereas others start and end later, but most interviewers only spend about one month in the field. Bulgaria, Cyprus, Russia and to a lesser extent Slovakia show similar patterns when assigning interviewers in the survey fieldwork. Belgium can be used as a good example of the second type of interviewer assignment strategy. All interviewers seem to start their fieldwork activities at the same time. After 8 weeks, about half of the interviewer force stops, whereas the other half continues. This may be the moment where the fieldwork management decides to change its fieldwork strategy by possibly opting for a more extensive (refusal) conversion program where more experienced interviewers are deployed. Apart from Belgium, Germany also shows traces of such a fieldwork silhouette, as well as Spain or Norway to some extent. The third type of interviewer assignment can be seen in the Netherlands. Finland, the United Kingdom and Sweden are in this respect also very similar. Under this regime, all interviewers seem to take part for as long as the fieldwork lasts. Somewhat striking is the fact that for some periods during the fieldwork, there hardly is any interviewer activity. A quite trivial explanation for these hold periods may be holidays such as Christmas, but they may also reflect a period of reflection on how to further develop the fieldwork strategy. The fourth type is shown by the graphs of Ireland and Portugal: Only a few interviewers are activated at the start of the fieldwork, gradually more interviewers get involved as the fieldwork continues. This might be due to an overlapping or concurrent fieldwork project that needed to be finished first. On the other hand, the pressure on the fieldwork in order to achieve a 70% response rate might have led to this particular fieldwork picture.

These different country graphs and associated conjecture or hypotheses may give rise to some interesting questions as to how the fieldwork is managed in the different countries and what particular objectives they served. These objectives involve the response rate to be achieved, fieldwork costs, the availability of the interviewer staff, the extent to which interviewers are specialized in some particular tasks (e.g. refusal conversion), Therefore, some questions can be considered, based on Figures 1, 2, 3 and 4:

- Were there any concurrent (face-to-face) studies fielding at the same time as ESS6?
- Does the fieldwork management build in reflection periods? Are there periods of low activity due to holidays?

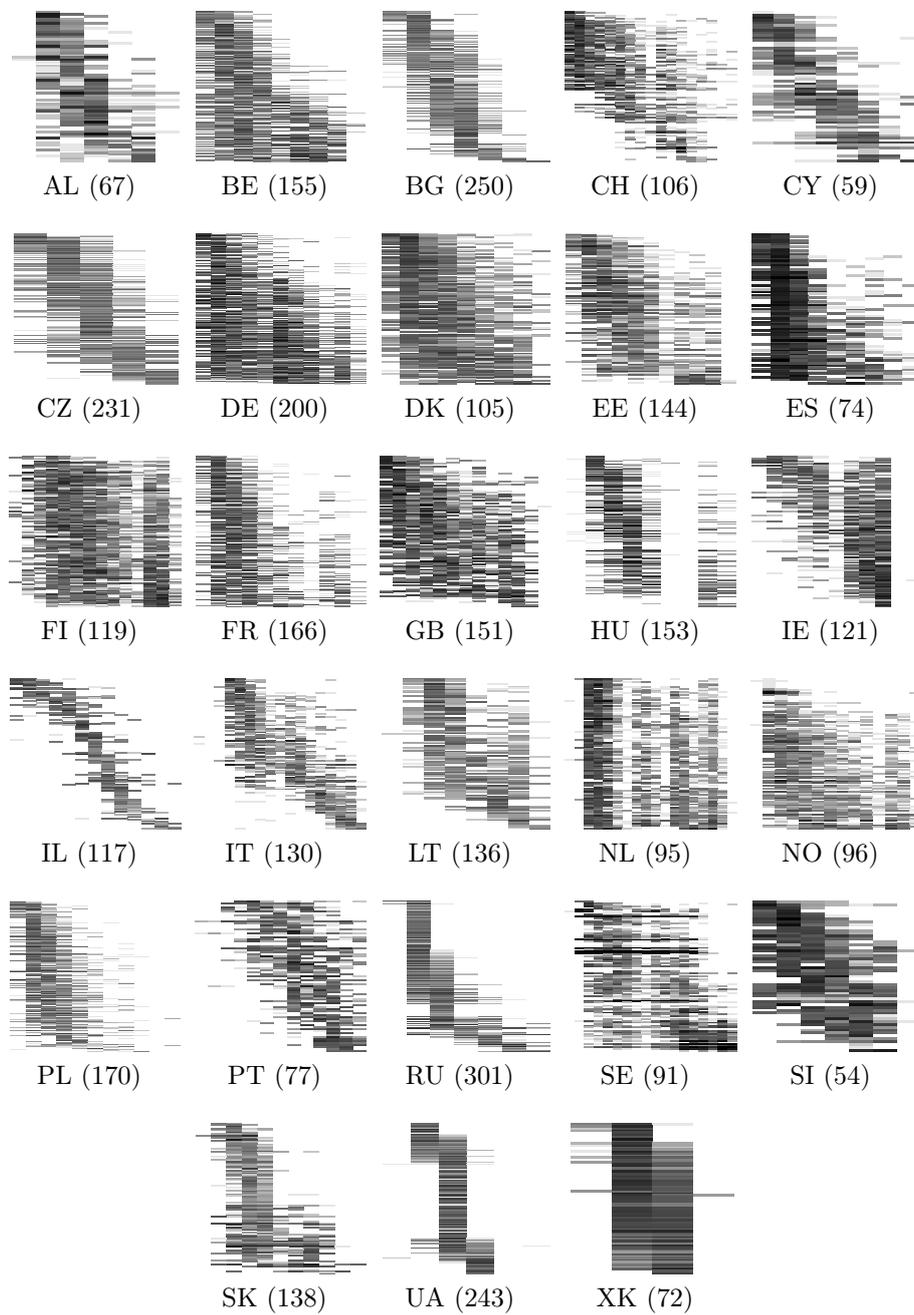


Figure 4: Interviewer intensity during the fieldwork

Table 1: Interviewers per country

Country	n	n_r	$\frac{n_r}{n}$	n_{int}	$\bar{attempts}$	\bar{cases}
AL	1602	1215	0.76	67	51.04	23.91
BE	3267	1869	0.57	155	69.29	21.08
BG	3200	2263	0.71	250	26.01	12.80
CH	2907	1529	0.53	106	114.36	27.42
CY	1588	1122	0.71	59	51.63	26.92
CZ	3010	2009	0.67	231	24.61	13.03
DE	8904	2964	0.33	200	151.99	44.52
DK	3372	1652	0.49	105	93.07	32.11
EE	3707	2382	0.64	144	63.76	25.74
ES	2868	1889	0.66	74	176.89	38.76
FI	3296	2197	0.67	119	152.97	27.70
FR	4200	1969	0.47	166	75.28	25.30
GB	4520	2287	0.51	151	142.39	29.93
HU	3194	2020	0.63	153	60.22	20.88
IE	4358	2633	0.60	121	87.31	36.02
IL	3230	2511	0.78	117	32.69	27.61
IT	2778	973	0.35	130	60.35	21.37
LT	4470	2109	0.47	136	26.82	32.87
NL	3537	1845	0.52	95	141.31	37.23
NO	3041	1628	0.54	96	70.92	31.68
PL	2706	1903	0.70	170	37.04	15.92
PT	3040	2152	0.71	77	86.51	39.48
RU	3772	2490	0.66	301	26.54	12.53
SE	3750	1863	0.50	91	465.12	41.21
SI	2250	1261	0.56	54	115.57	41.67
SK	2506	1853	0.74	138	41.41	18.16
UA	3692	2184	0.59	243	24.33	15.19
XK	2282	1295	0.57	72	55.40	31.69

 n : Total sample size n_{int} : Number of interviewers n_r : Completed interviews $\bar{attempts}$: Attempts per interviewer $\frac{n_r}{n}$: \sim response rate \bar{cases} : Cases per interviewers

- Does the payment of interviewers or their employment status explain why some interviewers are more episodic or short term deployed whereas others serve the whole fieldwork period?
- Have some interviewers been deployed during the fieldwork, more than initially planned?
- Did the response rate 70%-rule put pressure, particularly the end of the fieldwork, to increase the interviewer force.
- Why are some addresses only released in the field much later than the initially released addresses.
- Is there any plausible explanation why some days of the week show more activity than others.
- Could there be any over- or under-reporting of contact events, possibly biasing the fieldwork synopses?

2.3 Nonresponse codes

Apart from looking at the fieldwork from a temporal dimension or from an interviewer force perspective, the ESS fieldwork for round 6 counts more than (recorded) 250.000 contact attempts. Most of these attempts are unsuccessful. Figure 5 shows the distribution of these nonresponse codes per country. The number in each box represents the number of contacts during the entire fieldwork that have been recorded. For example, in Belgium 1219 refusals by the target person have been recorded (multiple occasions per unit are possible). The shade of the box expresses the likelihood of a renewed contact attempts right after such a nonresponse code occurred. The darker the box, the higher the probability of a renewed contact attempt. For example, the Belgian box for '*refusal.proxy*' is darker than for '*refusal.target*', indicating that refusals by proxy are more likely to be re-issued than refusals by the target person.

The nonresponse categories are:

- *Refusal.target*: Refusal of respondent
- *Refusal.proxy*: Refusal by proxy or interviewer didn't know if it was the target person
- *Noncontact*: No contact at all
- *Away*: Respondent is unavailable/not at home until ...
- *Disable.short*: Mentally/physically unable/ill/sick (short term and therefore could revisit during the fieldwork period)
- *Disable.long*: Mentally/physically unable/ill/sick (long term and would be unable to complete interview during the fieldwork period)
- *Language*: Language Barrier
- *Other*

- *Moved.known*: Respondent has moved, still in country
- *Moved.unknown*: Respondent moved to unknown destination
- *Ineligible*: Respondent has moved out of country, Respondent is deceased, address is not valid

Information on successful interviews or (preceding) appointments is excluded from Figure 5.

With regard to the frequencies of nonresponse codes it can be observed that some boxes in the graph are outliers. For example, the Czech Republic has an extraordinary number of partial interviews, in some countries the ‘other’-nonresponse category has been recorded relatively frequently (Finland, the Netherlands). The disability code (short term) has been used extremely frequent in Sweden. . . . Some of these differences may be attributed to real countries differences or for obvious reason be attributed to the survey design. For example, the frequency with which the ‘moved’-category is observed, can easily be related to the type of sampling course. Countries using individual based sample frames seem to have to deal more with this nonresponse categories. Also, due the survey climate, some countries may be faced with more reluctance than other countries. On the other hand, countries (or even interviewers within countries) may differ with regard to the decision they make as to how to categorize the result of a doorstep interaction. In this respect, it is not inconceivable that nonresponse categories ‘away’ and ‘refusal by proxy’ are somewhat interchangeable, which may have led to prioritize one category in one country, and the other category in the other country.

Looking at the rows of Figure 5, it seems that the overall re-issue tendency between the countries can be very different. Israel, for example, hardly applies any follow-up contacts, which is consistent with Figure 2. Other countries such as Finland, Spain or Hungary show a relatively high re-issue inclination. Some countries, such as Czech Republic are inclined only to re-issue specific classes of nonrespondents. In the case of Czech Republic, these are noncontacts, ‘away’, long-term disability, and partial interviews.

Looking at the columns of Figure 5, some remarkable differences between codes can also be noticed with regard to conversion practices. Most noncontacts are re-issued, as is also stipulated by the ESS fieldwork protocol: at least four contact attempts should be done, of which one during the weekend, one in the evening and there should be at least two weeks between the first and the last contact attempt. Looking at the darkness of the boxes, only Israel seems to be an exception. Also, most visits coded as ‘away’ and ‘disable.short’ have been re-approached later.

Special attention should be given to refusal conversion. These two categories are probably the most important classes of nonresponse. For countries that have not been able to achieve the 70% response rate target, refusal conversion might be the only solution to substantively enhance the response rates. This may explain why countries such as Cyprus, Czech Republic, Slovakia, Portugal and Kosovo hardly do any refusal conversion. This topic will be elaborated further on.

2.4 Compliance with noncontact specifications

The ESS requires at least four contact attempts, from which at least one in the evening and one at the weekend, best spread over at least two weeks to increase the chance of contact .

	refusal.target	refusal.proxy	noncontact	away	disable.short	disable.long	language	other	moved.known	moved.unknown	ineligible	partial
AL	234	294	6	31	8	7	4	213	49	20	76	31
BE	1219	74	4223	1058	42	138	75	237	58	90	86	26
BG	361	291	2308	167	64	24	7	2	9	74	184	11
CH	1708	451	5106	71	30	49	97	1	61	201	32	13
CY	104	85	1034	187	21	12	64	3	5	7	129	3
CZ	337	546	1715	238	19	28	10	1	3	4	70	178
DE	4584	600	12491	5035	122	441	171	20	53	486	92	5
DK	946	100	3659	524	74	123	42	511	89	73	16	21
EE	717	151	2649	900	65	106	5	274	314	252	31	16
ES	741	249	5753	1081	33	114	21	372	115	734	212	16
FI	1206	52	5246	19	20	101	44	7247	15	19	44	3
FR	292	1116	6672	69	48	0	347	0	60	7	435	0
GB	1409	1209	12317	1455	109	102	27	556	0	0	351	24
HU	2410	460	2795	277	23	117	0	240	0	389	115	8
IE	703	393	5521	74	11	89	77	29	1	3	333	0
IL	185	262	531	160	46	11	18	0	0	4	17	16
IT	874	385	3456	775	28	50	12	298	30	80	180	7
LT	107	460	460	20	1	4	6	0	6	17	295	0
NL	1956	606	6117	245	29	64	97	751	0	0	214	5
NO	956	106	1083	729	64	91	68	120	56	61	77	3
PL	427	138	1465	535	49	82	0	291	276	210	35	18
PT	83	332	3033	107	6	7	1	69	0	0	251	1
RU	190	875	3718	142	14	8	0	0	0	2	76	15
SE	2613	280	26111	707	397	12	0	0	167	101	249	1
SI	591	97	2352	435	35	53	3	367	42	99	42	6
SK	273	280	2067	90	11	7	0	0	0	0	8	32
UA	805	263	1774	306	25	29	3	222	30	32	1	27
XK	145	420	843	247	5	13	11	43	0	3	389	17

Figure 5: Occurrence and treatment of different types of nonresponse codes

Table 2: Percentages of noncontacts for which there was at least four contact attempts, one attempt in the evening, one attempt at the weekend and the contact attempts were spread over two weeks.

Country	Contact rate at t_1	Number of non-contacts	Four attempts	In the evening	At the weekend	Over two weeks
Albania	99,6	6	0,0	0,0	0,0	0,0
Belgium	60,6	63	100,0	100,0	100,0	100,0
Bulgaria	61,4	0				
Cyprus	67,7	55	100,0	85,5	70,9	52,7
Czech Republic	65,1	89	100,0	77,5	100,0	30,3
Denmark	63,2	70	91,4	92,9	88,6	91,4
Estonia*	74,6	119	100,0	84,9	93,3	83,2
Finland*	73,4	2	100,0	0,0	0,0	100,0
France	44,2	350	100,0	100,0	100,0	92,9
Germany*	56,8	408	27,9	60,5	31,6	53,2
Hungary	76,0	161	94,4	44,1	54,0	10,6
Ireland	58,7	613	100,0	82,9	88,9	46,0
Israel	86,6	282	0,0	41,1	23,0	0,4
Italy	58,7	327	82,6	80,7	67,3	41,9
Kosovo	77,3	0				
Lithuania	90,7	18	100,0	100,0	94,4	16,7
Netherlands	55,9	68	100,0	92,6	80,9	100,0
Norway*	87,0	104	42,3	75,0	39,4	39,4
Poland	77,9	24	83,3	95,8	79,2	83,3
Portugal	53,2	107	82,2	86,9	84,1	20,6
Russian Federation	56,0	270	51,5	91,5	81,9	15,9
Sweden*	99,4	22	0,0	22,7	50,0	0,0
Slovenia	64,0	130	55,4	67,7	63,1	64,6
Slovakia	64,8	97	83,5	60,8	74,2	43,3
Spain	59,9	15	100,0	100,0	100,0	93,3
Switzerland	43,6	173	43,4	70,5	50,3	56,1
Ukraine	76,0	289	70,9	77,9	72,0	0,0
United Kingdom	44,4	323	89,8	92,3	92,3	92,0

*This country contacted a large proportion of their sample by telephone first, the first contact attempt refers here to the first face-to-face contact attempt(Estonia 12.5%, Finland 58.9%, Germany 20.1%, Norway 56.3% and Sweden 95.6%)

Table 2 displays the percentage of noncontacts that have been contacted at least four times, in the evening, at the weekend and for which the contact attempts were spread over two weeks complying with the ESS guidelines.

In Belgium, all noncontacts (63) were visited four times, for all units there was an attempt in the evening, at the weekend and the attempts were spread over two weeks. This is in fact the only country that perfectly meets the requirements. Almost all countries that achieved less than 3% noncontact rate requirement followed the guidelines for the contact procedures to a high degree; the exceptions being Finland and Sweden with respectively 2 and 15 noncontacts. From the countries that obtained noncontact rates between 3 and 6%, Estonia followed the guidelines for a large percentage of the noncontact as well as Cyprus, Portugal and Slovakia; the last three failed to spread their contact attempts over more than two weeks. Norway, Germany, Hungary, Slovenia and Switzerland have lower degrees of compliance with the guidelines. Despite a high degree of compliance with the guidelines the United Kingdom and Ireland (not spread over two weeks) had higher noncontact rates. Finally, about half of the noncontacts in the Russian Federation did not get four contact attempts and the visits were rarely spread over more than two weeks. Israel has a low compliance with the contact procedure guidelines.

3 Fieldwork results

Many of the fieldwork processes that are discussed in the previous section may be related to the fieldwork results in terms of response rates, contact rates or refusal rates. The causal order of the relationship may not always be so obvious. On the one hand, extended fieldwork efforts or strategies may result in better response rates. On the other hand, and this turns the causal chain, the expectation of a certain response rate may make fieldwork managers anticipate and alter their strategy to save on fieldwork costs. Therefore, one might also expect that low levels of fieldwork efforts are related to high response rates.

Additional to this presentation of fieldwork results, we will also provide some indications of nonresponse bias, based on the information from interviewer observable data.

3.1 Ineligibles

The percentage of ineligibles found in the sample is mainly related to the quality of the sampling frame. Figure 6 shows the ineligible rates on the sample by country and for each ineligible category. The countries are split by the type of sampling frame that they used.

Two groups of countries are distinguishable: countries with less than 4% ineligibles (Slovakia, Israel, Switzerland, Finland, Germany, Russian Federation, Czech Republic, Hungary, Belgium, Norway, and Slovenia) and countries with more than 4% (United Kingdom, Netherlands, Estonia, Spain, Poland, Ireland, Cyprus, Portugal and Kosovo). Noteworthy is that in many countries that used an address of household frame the ‘not occupied’ code for ineligibles is the most frequent (Israel, Russian Federation, Czech Republic, United Kingdom, Netherlands, Ireland, Cyprus, Portugal and Kosovo), this can be due to the specific situation in the country. The second most frequently used category is ‘respondent move out of country’ in countries that used an individual frame (Switzerland, Germany, Hungary, Sweden, Norway, Slovenia, Estonia, Spain and Poland).

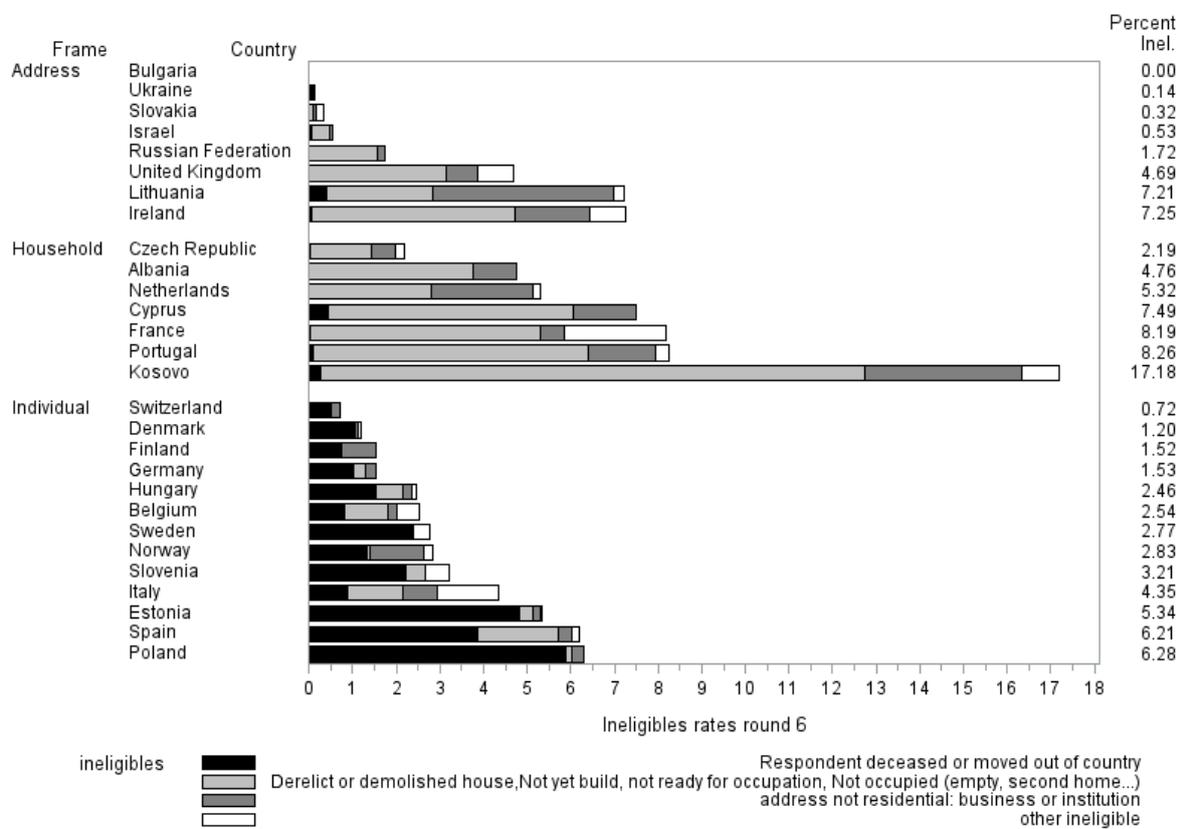


Figure 6: Ineligibles by country, ESS6

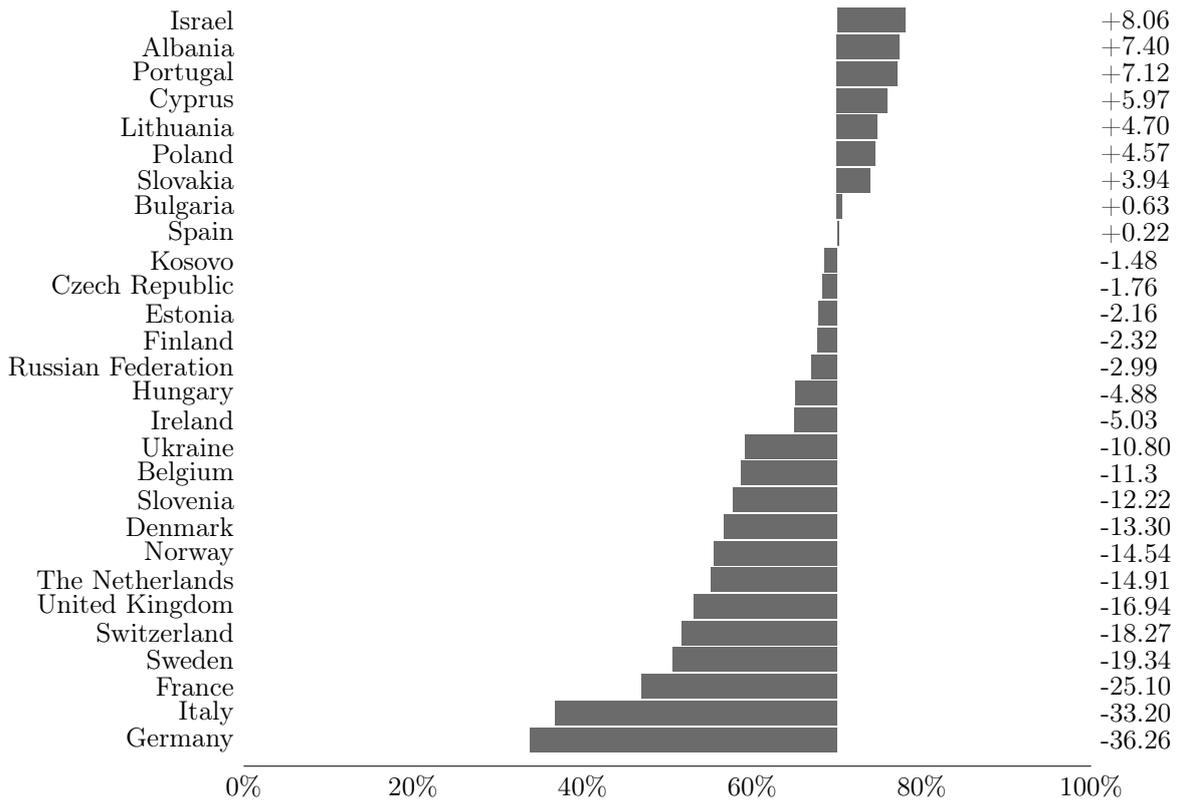


Figure 7: Response rates, ineligibles excluded, ESS6

3.2 Response rates

The European Social Survey requirements state that each participating country should aim for a response rate of 70% or more. The response rates are calculated as the total number of completed interviews divided by the sample size from which the identified ineligible cases are subtracted (AAPOR RR1). Ineligibles are defined as ‘Respondent deceased’, ‘Respondent moved out of the country’, ‘Derelict or demolished house’, ‘Not yet built, not ready for occupation’, ‘Not occupied’, ‘Address not residential: business’, ‘Address not residential: institution’, ‘other ineligible’.

Figure 7 gives an overview of the obtained response rate in each participating country compared to this goal. The benchmark of 70% is clearly marked in the figure and also shows the deviation from this margin (see right column of Figure 7).

Nine countries obtained response rates higher than the target response rate (Israel, Albania, Portugal, Cyprus, Lithuania, Poland, Slovakia, Bulgaria and Spain). Six countries came really close to the target response rate, response rates above 65% (Kosovo, Czech Republic, Estonia, Finland, Russian Federation and Hungary). All remaining countries with the exception of Germany (33.7%) and Italy (36.80%) obtained response rates above 50% (Ireland, Belgium, Slovenia, Norway, Netherlands, United Kingdom, Switzerland, and Sweden).

Figure 8 shows the evolution of the response rates for each country since round 1. Only ESS6 participating countries have been taken into account.

The ESS mean response rate is calculated as the mean over all countries that participated in round 6. These average response rates per ESS round are indicated in grey.

Some countries have been increasing their response rate rather consistently (Switzerland, Cyprus, Czech Republic, Estonia, Slovakia and Portugal). A few countries have

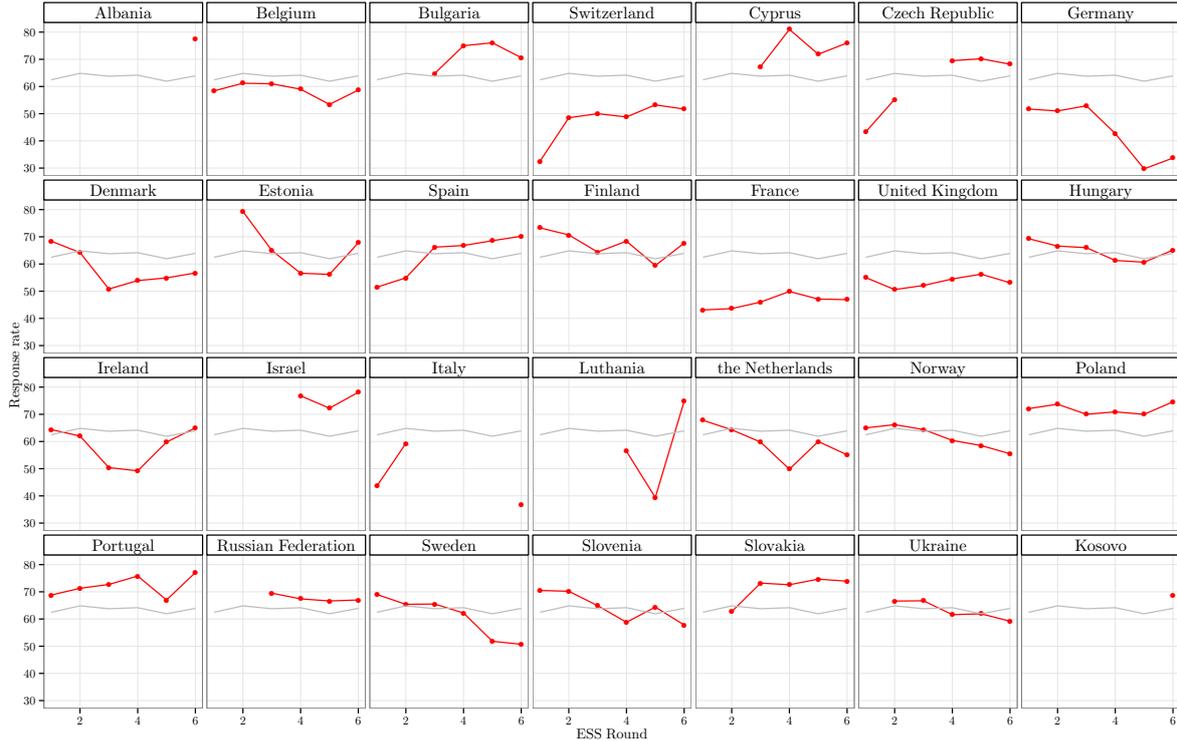


Figure 8: Response rates, ineligible excluded, over different ESS rounds

had rather stable response rates over the rounds (Belgium, United Kingdom, Israel and Poland). A bunch of countries displays a slight decrease in response rates (Finland, Hungary, Netherlands and Russian Federation) whilst others have a more pronounced drop (Germany, Estonia, Norway, Sweden, Slovenia and Lithuania).

Moreover, some countries obtained systematically response rates over the ESS mean response rate (Bulgaria, Cyprus, Finland, Israel, Poland, Portugal, Russian Federation, Slovakia) whilst others systematically lie under the ESS mean response rate curve (Belgium, Switzerland, Germany, United Kingdom and Ireland). A last group of countries obtained for some rounds a response rate above the ESS mean and under for others (Czech Republic, Estonia, Spain, Hungary, Netherlands, Norway, Sweden and Slovenia).

3.3 Noncontact rates

On top of the 70% requirement, the European Social Survey also set a goal to obtain a noncontact rate of maximum 3%. Figure 9 shows which countries managed to achieve that goal (or not) and in how far they deviate from this goal.

Bulgaria and Kosovo obtained a noncontact rate of 0%. An explanation for this result is necessary. Nine other countries also kept their noncontact rate below 3% (Finland, Albania, Spain, Sweden, Lithuania, Poland, Belgium, the Netherlands and Denmark); moreover the Czech Republic had a noncontact rate of 3.02%. Nine countries managed to keep their noncontact rate below 6 % (Estonia, Norway, Cyprus, Portugal, Republic of Slovakia, Germany, Hungary, Slovenia and Switzerland). The Russian Federation, the United Kingdom, Ukraine and Israel obtained noncontact rate below 9%. Ireland had a high noncontact rate of 15%, Italy of 12.6% and France of 9.08%.

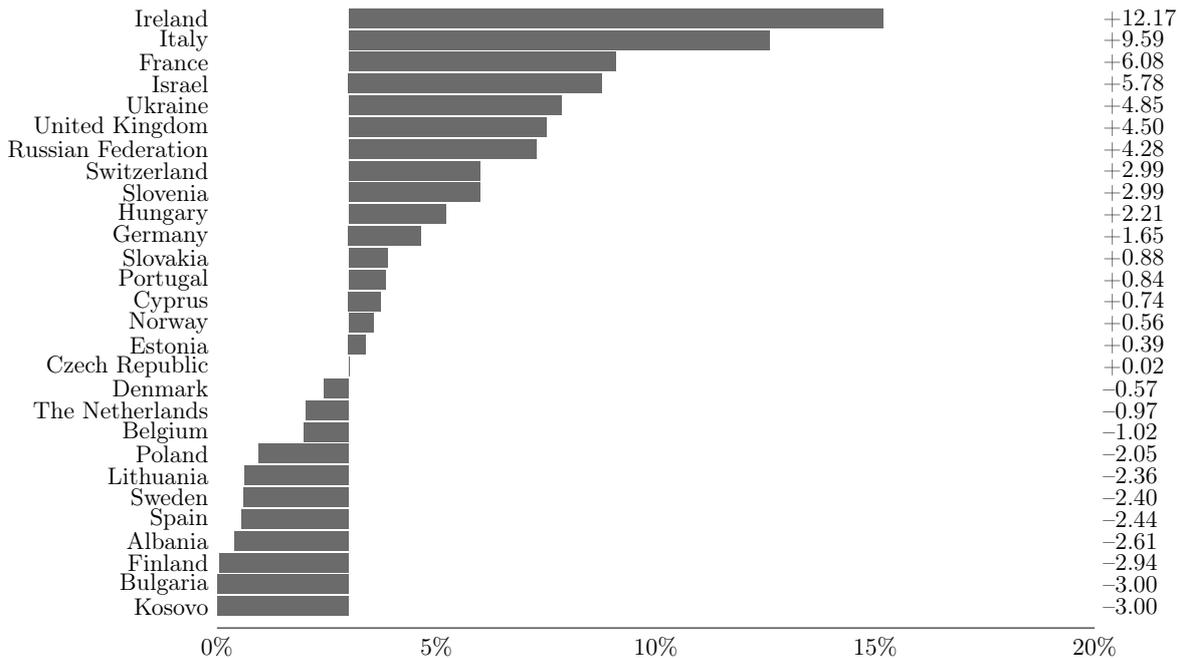


Figure 9: Noncontact rates, ineligible excluded, ESS6

In Figure 10, we display the countries in a scatter plot with the contact rate as y-axis (output) and the contact procedure score as x-axis (amount of effort). The contact procedure scores are constructed as follows. For each noncontact, we count the number of guidelines that have been followed, an individual score. For instance, if a noncontact has been visited four times, once in the evening and the attempts were spread over two weeks but no attempt was made at the weekend, that unit would get a score of 3. We then calculate the percentage of noncontact that has obtained score 1, score 2, score 3 or score 4. The country score is then calculate has the sum of the product of the individual score (1, 2, 3, or 4) times the percentage of non-contact with that score and can thus vary between 0 (no compliance at all) and 400 (complete compliance). A country with score 0 would not have applied any of the contact procedure guidelines to any of its noncontacts; a country with a score of 400 would have applied all the procedure guidelines to all noncontacts (e.g. Belgium). See Table 2 for the information determining the noncontact compliance score.

The increasing contact rate with increasing score at the country level may be empirical evidence that high (or low) contact rates are a consequence of the compliance to the contact procedure guidelines and the amount of effort.

Obvious outliers are Finland and Sweden, probably due to their very low number of noncontacts and Ireland and the United Kingdom with rather high noncontact rates despite a relatively high score. In Ireland the amount of effort is rather high but the impact on the contact rate is not in line with the other countries.

3.4 Refusal rates and refusal conversion

Next to non-contacts, refusals are the main cause for non-participation. Refusals represents in fact a higher percentage of nonresponse, ranging between 8 and 45% compared to noncontacts 0 to 15%. The European Social Survey encourages countries to have a strategy for refusal conversion, typically re-approaching refusals with a more experienced,

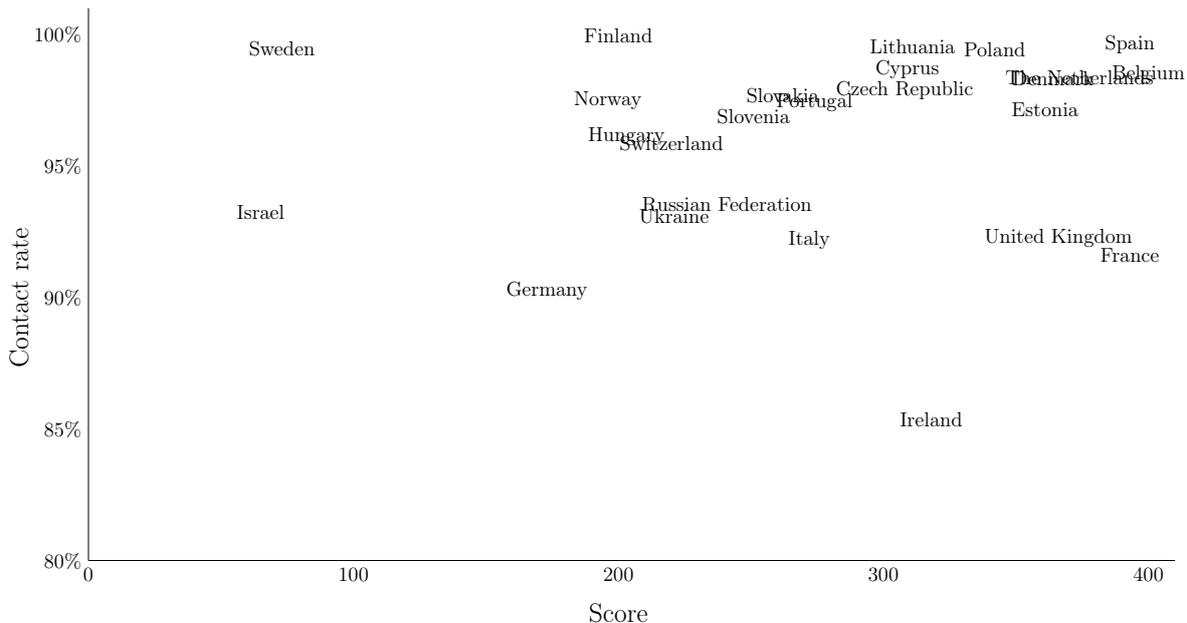


Figure 10: Contact rates by contact procedure scores, ESS6

trained interviewer with the purpose to reduce refusals rates and ultimately nonresponse rates. The refusal rates differ greatly from country to country from 8.3% in Israel to 45.9% in Germany.

We can more or less distinguish three groups of countries. The first group (Germany, Sweden, Netherlands, Italy, Switzerland, Denmark, France, United Kingdom, and Norway) has the highest refusal rates, above 30% with the extreme case of Germany (45.9%). A middle group with somehow lower refusal rates from 20 to 30% (Ukraine, Kosovo, Czech Republic, Belgium, Slovenia, Russian Federation, Finland, Hungary, Slovakia and Lithuania). The last group of countries has refusal rates below 15% (Albania, Poland, Portugal, Bulgaria, Ireland, Cyprus, Spain, Estonia, and Israel).

In some countries, the refusal conversion efforts can have a considerable effect in the eventual response rate, as indicated by Figure 12. The longer lighter bars indicate the response rate without any conversion efforts. The shorter darker bars at the end indicate the response rate increase due to conversion attempts. The darker the bars, the more initially reluctant nonrespondents have been reissued. Apparently, countries showing higher degrees of renewed contact attempts after a refusal, tend to increase their response rates more significantly. In that sense, refusal conversion efforts seem to pay off.

3.5 Indications of nonresponse bias

In order to evaluate the extent to which nonresponse affects survey estimates, the ESS fieldwork protocol has chosen to collect data about observable information that can be recorded by the interviewers at their first visit. The following items were asked to be filled out:

- TYPE: What type of house does the (target) respondent live in? The interviewer could choose between: farm, detached house, semi-detached house, terraced house, the only housing unit in a building with another purpose (commercial property), flat, student apartment, retirement house, House-trailer or boat, or other.

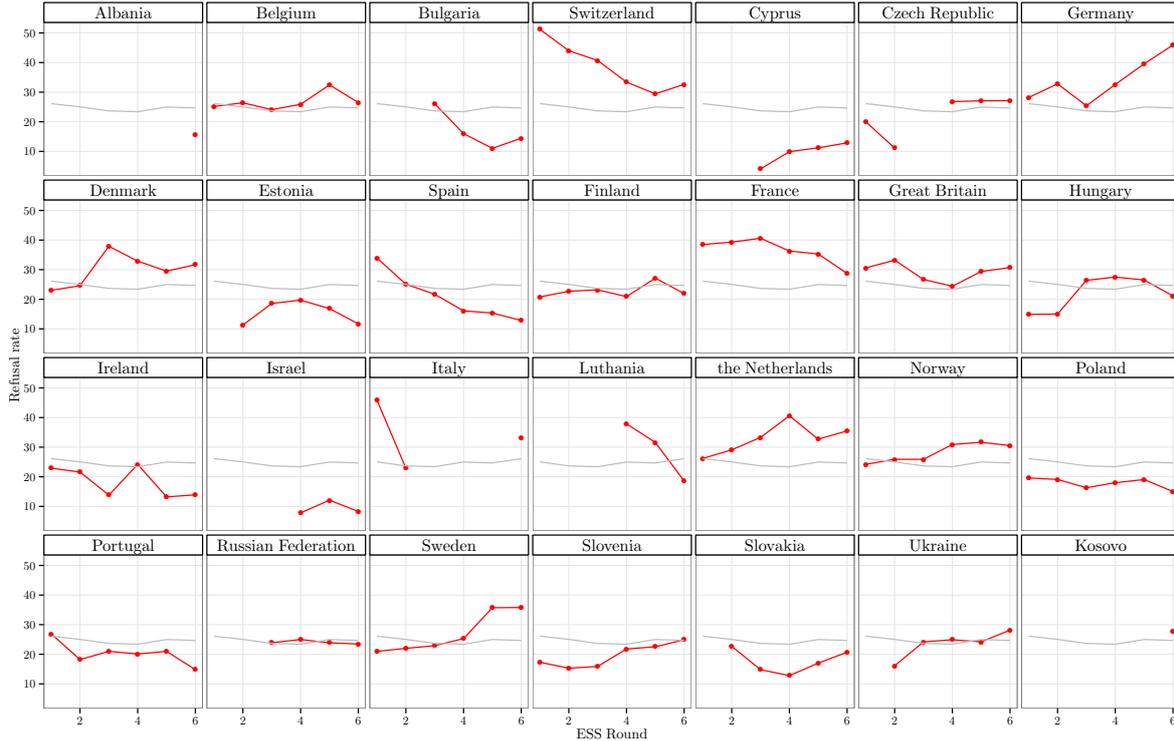


Figure 11: Refusal rates, ineligible excluded, over different ESS rounds

With regard to this variable, we will rather focus on the distinction between apartment dwellers and non-apartment dwellers (indicated as ‘FLAT’).

- GATE/DOOR: Before reaching the (target) respondent’s individual door, is there an entry phone system or locked gate / door? (1: Entry phone system; 2: locked gate/door; 3 : both entry phone system and locked gate/door; 4: neither of these).
- PHYSA: What is your assessment of the overall physical condition of this building/house? (Very good, good, satisfactory, bad, very bad)
- LITTER: In the immediate vicinity, how much litter and rubbish is there? (very large amount, large amount, small amount, none or almost none)
- VANDAA: In the immediate vicinity, how much vandalism and graffiti is there? (very large amount, large amount, small amount, none or almost none)

Apart from these observable data, many countries (particularly individual sample based frame countries) could also provide gender and age information about the sampled individuals. In sum, seven variables are available to make a nonresponse bias assessment.

Figure 13 gives an example of how such a nonresponse bias assessment can be represented. First, a table is provided, comparing the total sample distribution of the German variable ‘FLAT’, to the distribution of respondents only and nonrespondents only. On a total of 8812 sampled cases, only 2993 (about 34%) cooperated in the survey. Of the total sample, 47% is believed to live in an apartment (as recorded by the interviewers). Among the respondents only, 42% are estimated to be apartment dwellers, indicating a

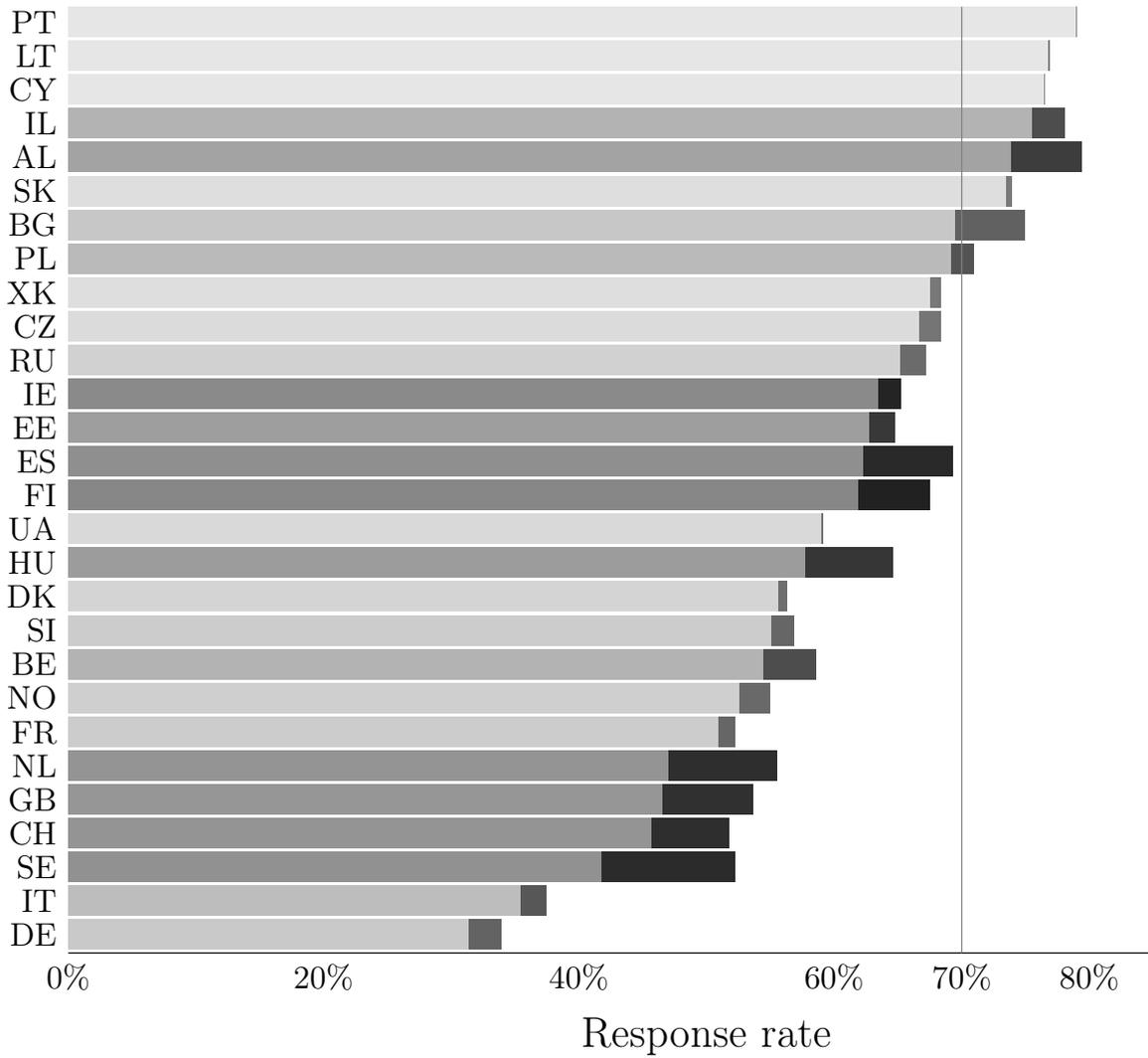
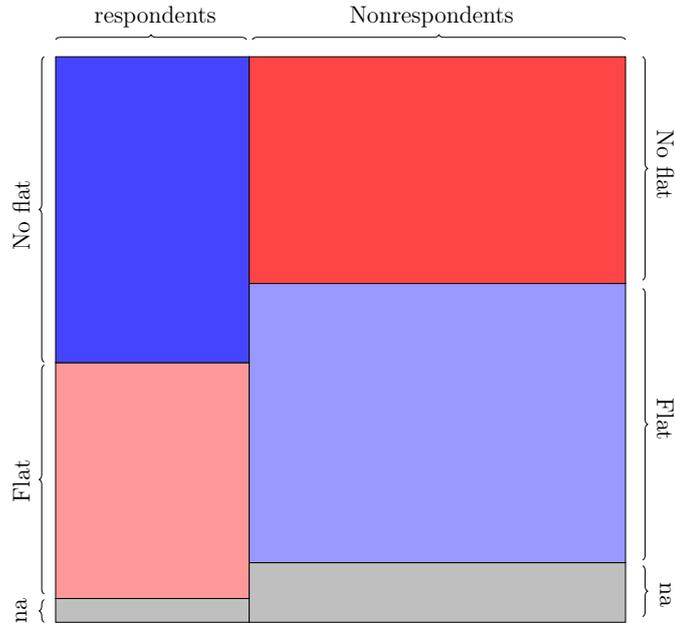


Figure 12: Effect of refusal conversion on response rates, ESS6

TYPE?	respon- dents		nonres pondents		total sample	
No flat	1619	54%	2333	40%	3952	45%
Flat	1248	42%	2873	49%	4121	47%
na	126	4%	613	11%	739	8%
	2993		5819		8812	

Flat × response table



Flat × response mosaic plot

Figure 13: Association between (non)response and type of housing, Germany, ESS6

nonresponse bias of 5%. However, this estimate of bias may be somewhat blurred because 8% of the total sample (4% among the respondents) has not been coded by the interviewers regarding their housing situation.

The same information as provided by the table in Figure 13 is also shown by the mosaic plot (second panel of Figure 13). The plot consists of two columns, the respondents on the left hand side, the nonrespondents on the right hand side. Within each of the two columns, the distribution of the type of housing is displayed. Surfaces in blue indicate that particular combinations (for example, respondents living in flats) are overrepresented, red surfaces indicate under-representation. The colour intensity reflects the Pearson residual of the table analysis. Grey rectangles represent missing information for the observable variable.

The advantage of the mosaic plot is that it can be interpreted much faster than a contingency table, which is particularly convenient when many variables and many countries need to be compared. Figures 14, 15 and 16 provide mosaic plots for all county × variable combinations. It should thereby be noted that sample sizes of the different counties have been made equal (through weighting on the country level) so that the Pearson residual do not reflect differences in the sample size, but only reflect the degree of association between in the various observable variables on the one hand and the 0-1 response indicator on the other hand.

The three page overview of Figures 14, 15 and 16 clearly shows which countries and variables seem to be affected by nonresponse bias (insofar the available variables are good representatives of the target variables). The ‘closed gate/ entry door’ and the ‘type of housing’ variables seem to be most affected by nonresponse. In nearly all countries where this information is collected, (serious) traces of bias can be found for these variables. For the other three observable variables, bias may be observed, but not as explicit as compared to the ‘closed gate/door’ and the ‘type of housing’ variables. The age and gender variables, that were obtained from the sample frame (only for individual sample frame countries) only indicates rather weak traces of nonresponse bias. In Germany (DE), the total sample has 48.39% men, whereas the respondent only sample counts 50.43% men. Therefore, the blue and red surfaces in the respective mosaic plot are not so intense as compared to the plots about the observable variables (VANDA, LITTER, PHYSA, GATE/DOOR and FLAT).

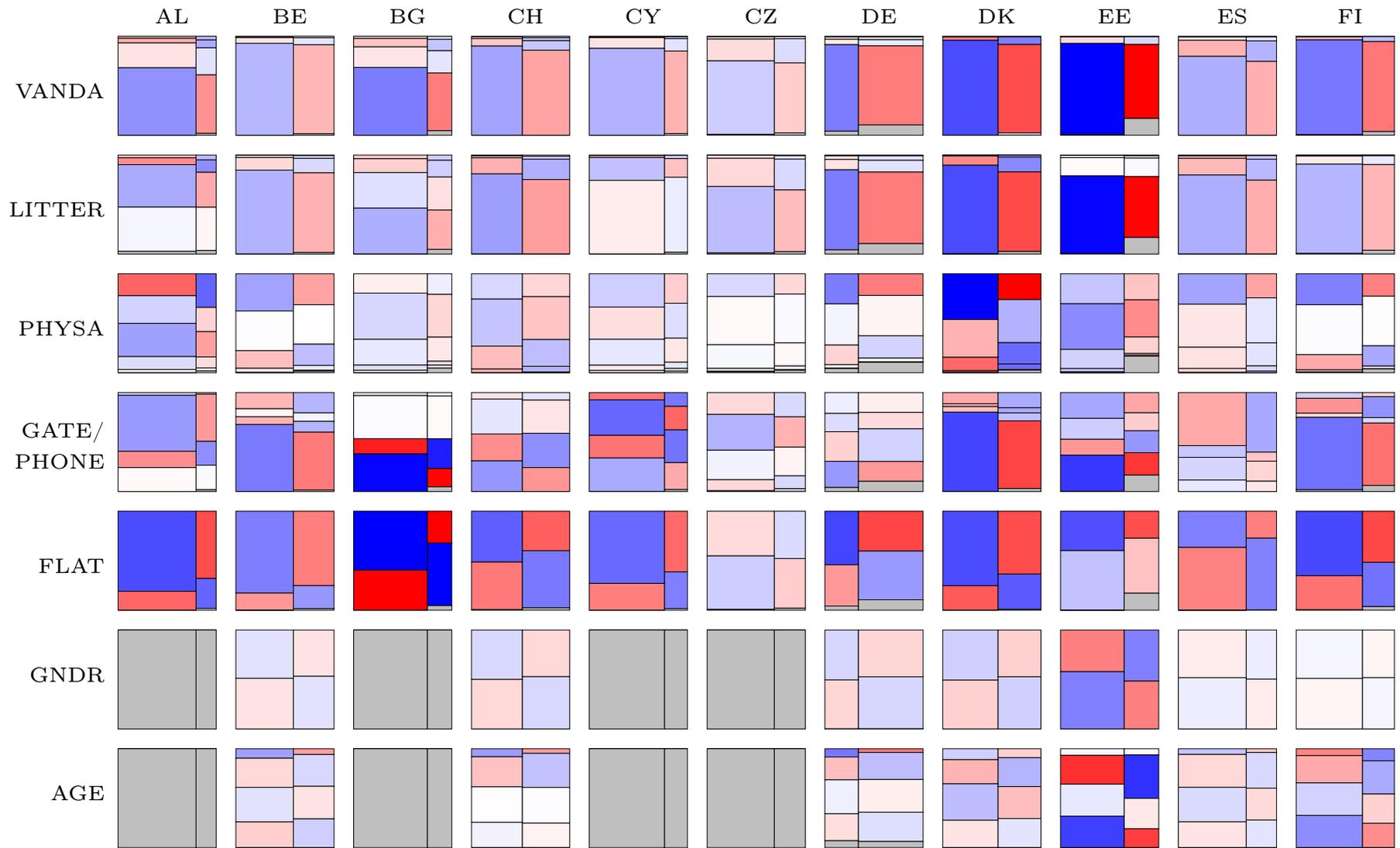


Figure 14: Observable variables (a)

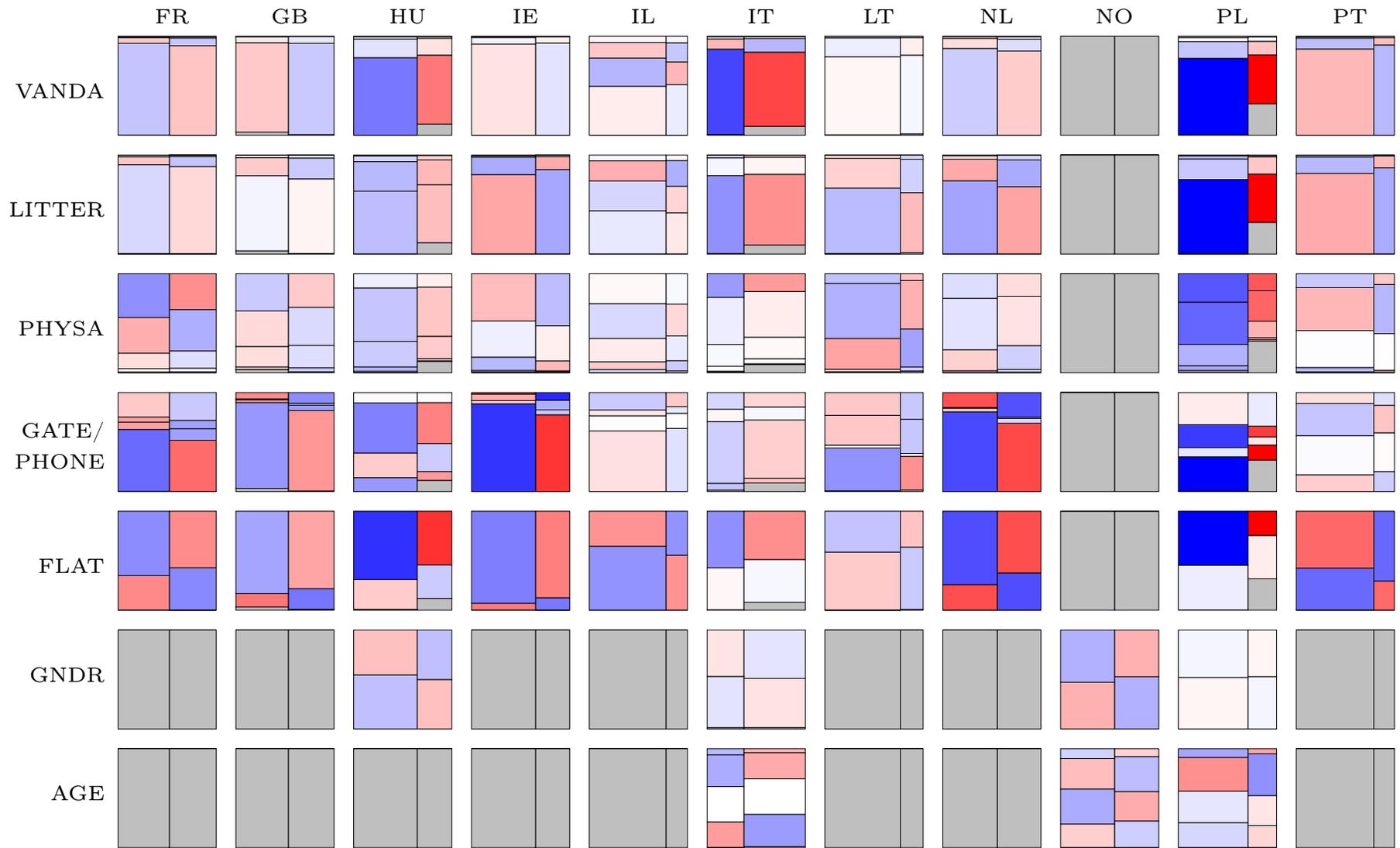


Figure 15: Observable variables (b)

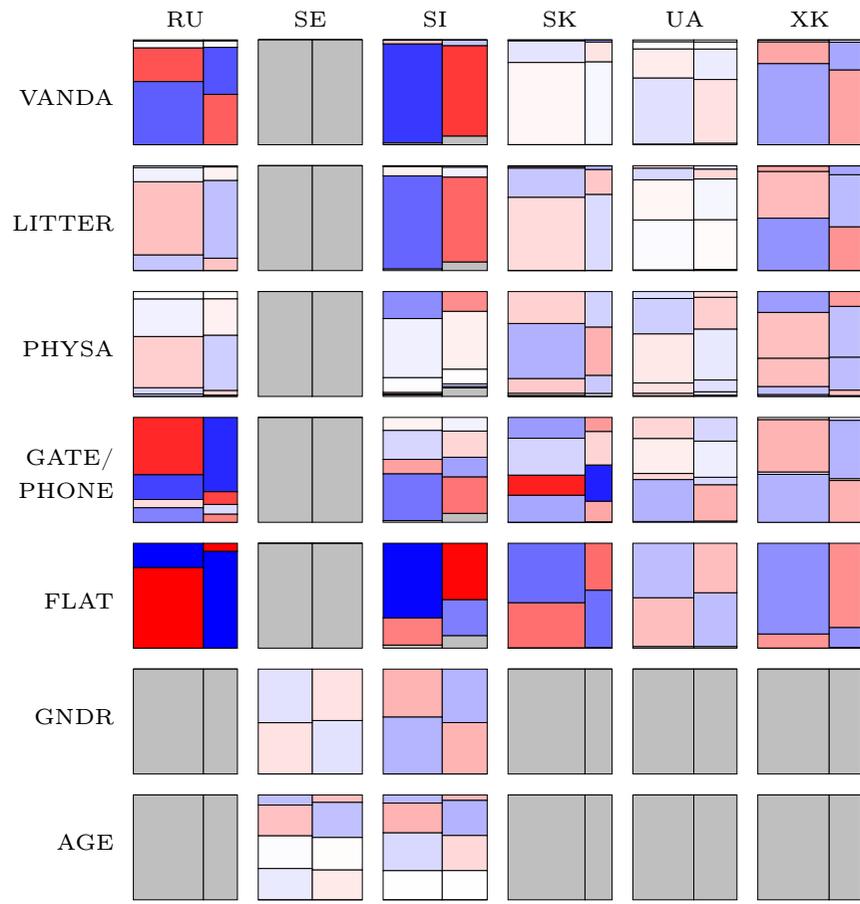


Figure 16: Observable variables (c)

Instead of looking at the rows of the two figures, the columns indicate which countries might be most affected by nonresponse. Particularly Estonia (EE), Poland (PL), the Russian Federation (RU) and Slovenia (SI) show rather deeply coloured mosaic plots.

The household- or address-based sample frame countries do not have any data available gender and age, so that the mosaic plot can not be made. For Norway and Sweden, the observable data are not available since interviewers are legally not allowed to collect such information. Nevertheless, in some countries such Estonia (EE) or Germany (DE) there is a considerable amount of item nonresponse affecting the observable data, particularly among the nonrespondents.

4 The actual interview

Once contact and survey participation have been established with the target person, the actual interview can start. We used to the ESS6 main file to have some first indications of the quality of this questioning and answering process. In the subsequent section, interview quality will be assessed looking at interviewer effects (section 4.1), interview length (section 4.2), non-substantive answers (section 4.3) and straightlining (section 4.4).

4.1 Interviewer variance

In each round of ESS more than 3000 interviewers play a central role in the data collection process. Their tasks are comprehensive and survey researchers are aware of the possible positive and negative impact of interviewers on the data quality. They can help and stimulate the respondent to perform his or her role in an adequate way (positive impact) or they can influence the responses in a systematic way (negative impact). With such a large number of interviewers in a cross-national survey, the implementation of standardized interviewing techniques and the reduction of negative interviewer effects can be considered as a major challenge.

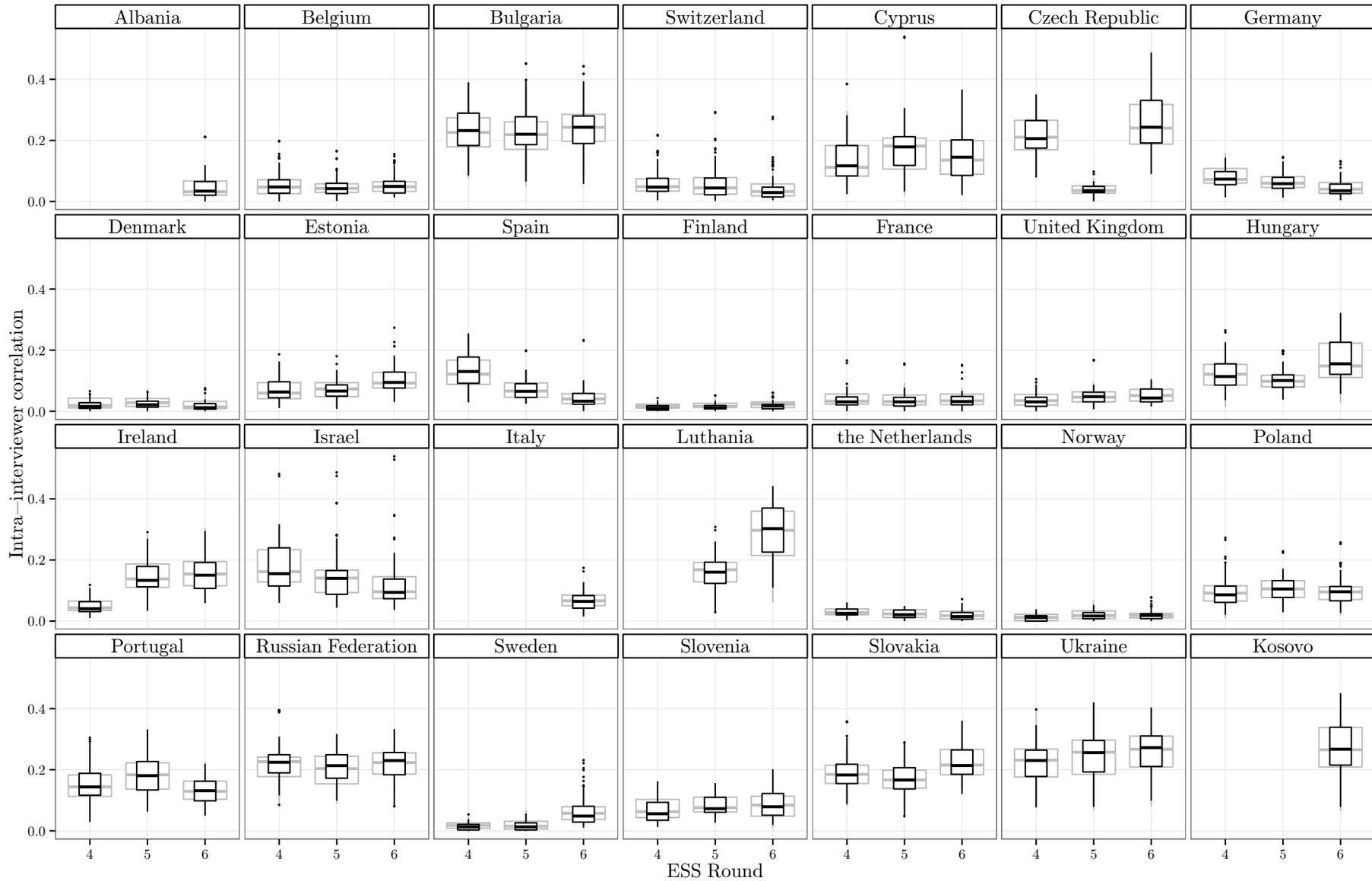


Figure 17: Evolution of intra-interviewer correlation of 51 survey items between ESS4, ESS5 and ESS6, 21 countries

Previous reports provide strong indications of the existence of interviewer effects or interviewer variance on substantial survey questions (ESS-DACE: deliverable 12.2 and 12.10) (Loosveldt & Beullens, 2010; Beullens & Loosveldt, 2013). As a summary of these reports, consider Figure 17, showing the distribution of intra-interviewer correlations over 51 core survey items in 28 ESS countries. For each country, three sets of boxplots are shown, referring respectively to the fourth, fifth and sixth round of ESS. The outer grey boxplots show the distribution of the raw (null model) intra-interviewer correlation. The inner black boxplots indicates the distribution of the same intra-interviewer correlations after controlling for or removing respondent characteristics: age, gender, degree of urbanization, marital status, employment status and level of education. The distinction between raw and controlled intra-interviewer correlation may be relevant as mostly interviewers are assigned to sample cases in their own neighbourhood, trying to reduce travel time and cost. Therefore, area and interviewer effects are hard to disentangle. By controlling for relevant respondent characteristics, these area effects are partially taken into account, making the resulting interviewer-specific samples more comparable.

Based on Figure 17, the ICC of the different rounds are relatively stable over the three ESS rounds. Only in Czech Republic, particularly the fifth round seems to strongly deviate from the two other rounds. The reason is yet unknown. There is no clear upward or downward trend in the ICC's, although Spain, Germany, Israel and to a lesser extent Switzerland seem to have their interviewer effects reduced over the various rounds. Upward trends can be observed in the United Kingdom or to a lesser extent in Finland. The distributions of the ICC's with and without controlling for respondent background variables does not seem to make much difference. This may indicate that interviewer effects are more dominant than area effects in this regard, supporting the findings as presented in survey literature (O'Muircheartaigh & Campanelli, 1998; Schnell & Kreuter, 2005).

4.2 Interviewer effects on interview length

In the ESS, interviewers are instructed to perform their tasks according to the key principle of standardized interviewing. Standardized interviewing aims to ensure that all respondents answer exactly the same questions under conditions that are as consistent as possible (Groves et al., 2004). It further means that interviewers should apply the same basic task rules during the interaction with each respondent, and that they should spend the same efforts to obtain adequate responses. Standardized interviewing implies that each interviewer's contribution to interview length should be approximately the same for each interview with a similar respondent. Therefore, the overall expectation is that interviewer effects on interview length will be limited.

Table 3 reports the average length in minutes of the interviews (main questionnaire only) per country. Kosovo seems to have recorded the longest average interview length (about 80 minutes), whereas Slovenia has the shortest interviews (about 40 minutes). These differences between the countries can be explained to some extent by acceptable reasons such as language differences or individuals in countries for which systematically other routings apply because of their employment situation, family life, Nevertheless, the relatively large differences may also reflect different interviewing customs, or even reflect different instructions that were given during the interviewer briefing. This may be somewhat worrisome since cross-national research and the comparability of countries relies on standardized interviewing.

Table 3: Interview length in minutes, means and standard deviation per country, ESS6

Country	mean	sd	Country	mean	sd	Country	mean	sd
AL	45.82	12.71	FI	47.75	15.40	NO	53.58	17.84
BE	51.66	12.88	FR	50.22	17.87	PL	65.03	19.46
BG	59.31	10.15	GB	42.62	13.60	PT	45.62	14.40
CH	51.54	21.05	HU	38.34	9.22	RU	53.41	12.81
CY	68.13	20.14	IE	46.72	12.27	SE	54.49	18.47
CZ	65.00	19.35	IL	44.02	15.12	SI	40.97	14.72
DE	62.98	20.20	IS	50.96	17.33	SK	68.75	23.80
DK	54.15	16.10	IT	55.54	18.97	UA	61.95	15.38
EE	53.24	18.10	LT	65.45	17.31	XK	80.62	24.23
ES	49.03	17.32	NL	53.36	17.70			

The differences within the countries regarding interview length may in this respect also be alarming. The within country differences are indicated by the standard deviations (sd) in Table 3. Figure 18 tries to disassemble this within-country variances, taking into account some determinants of interview length. First, the interview length can be explained by the number of questions the respondent needs to answer, provided his/her employment status or family situation. Second, some respondents might need more time to think about the answer they should give. Elderly people, for example, tend to take more time to give answers to survey questions (Loosveldt & Beullens, 2013a, 2013b). Therefore, some background variables will be taken into account such as age, gender, level of urbanization and level of education to make to the respondents more comparable, assessing variability of the interview length. Finally, also interviewer effects will be tested. Unlike to two first sets of variables probably affecting interview length, interviewer effects are not considered appropriate or valid determinants of interview length.

Figure 18 can be read as follows. For every country a null model is determined (standard deviations as reported in Table 3). Then, the standard deviation of the within-country interview length is shown, after the effect of the the number of eligible questions is taken into account. Subsequently, also the effect of the background variables are taken away. Finally, also the interviewer effects are taken into consideration.

The results suggest that raw within-country differences with regard to interview length can be seriously reduced when particularly interviewer effects are taken into account. To a lesser extent, also respondent characteristics may play a role. After controlling for interviewer effects, the country differences seem to be considerable reduced.

4.3 Item nonresponse

Another indicator of survey quality, particularly regarding the obtained answers, is the degree to which respondents give substantive answers to the survey questions. Of course, it should be acknowledged that ‘don’t know’ or ‘refusal’ may be a perfectly reasonable and valid answers, nevertheless, the volume of item nonresponse is preferred to be rather low. Also, difference with regard to item nonresponse between countries or between interviewer may be hard to accept.

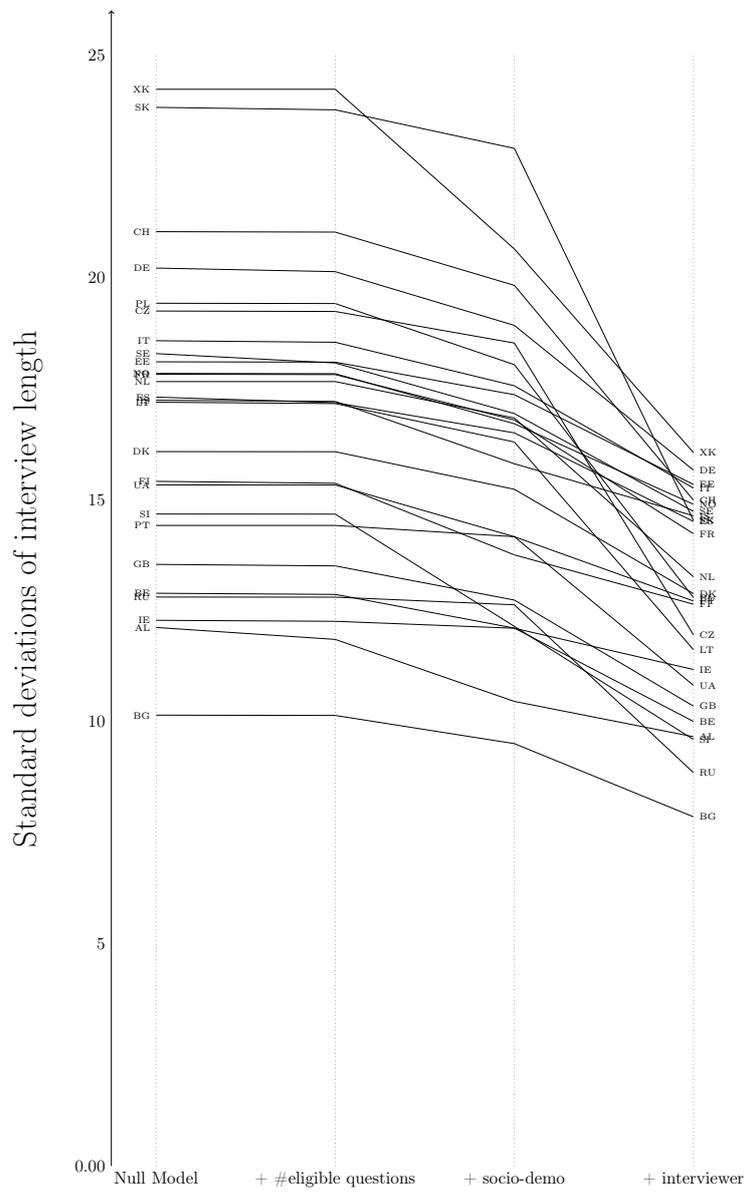


Figure 18: Variety if interview length in different ESS countries under different models

The ESS6 main file counts 557 items to which respondents possibly could give an answer. Many of these items, however, are inapplicable for most respondents because of question routing. Though conditional on previous answers, on average respondents need to provide about 260 answers to survey questions. This number of question to be answered can be determined individually. Then, the number of non-substantive answers is calculated ('don't know', 'no answer', 'refusal') and is divided by the items that should have been answered, provided a percentage of unanswered items per individual. In Figure 19, the ESS6 countries have been ordered according to the average number of these non-substantive answers.

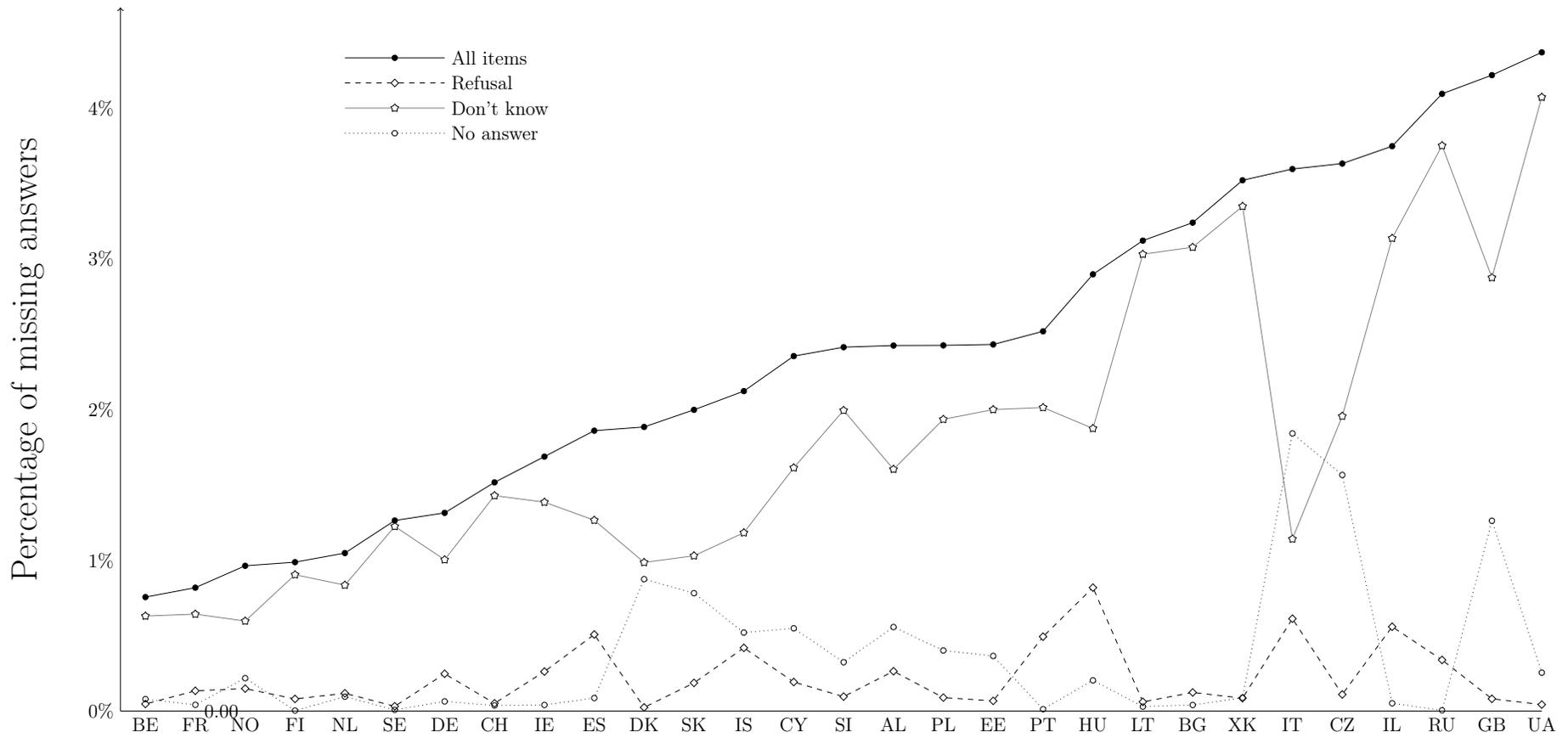


Figure 19: Item nonresponse in different ESS countries

The black solid line describes the percentage of all non-substantive answers to survey questions. This total can subsequently be split up in refusals (dashed curve), grey solid line indicates the ‘DK’-answers and the ‘no answers’ are indicated by a dotted line.

The total percentage of missing answers varies between less than 1% in Belgium and more than 4% in Ukraine. ‘DK’-answers seem to be the largest subcategory of the missing answers. Some countries, however, seem to have rather exceptionally high percentages of non-substantive answers coded as ‘no answer’. These countries are Denmark, Slovakia, Czech Republic and the United Kingdom. Countries such as Spain, Iceland, Portugal, Israel and the Russian Federation seem to have a relatively large amount of item-refusals.

4.4 Straightlining

Straightlining can be defined as providing the same answer to an item as compared to the previous item. In this regard, consider the illustrative example of how two respondents could have answered a series of 6 sets of likert scale set of items.

As such, there is of course nothing wrong with straightlining, as long as the respondents’ scores consistently reflect the true underlying attitude. However, it is hard to explain why differences can be observed between countries or interviewers, or why straightlining occurs more frequently as interviewer becomes more acquainted with the questionnaire (Beullens & Loosveldt, 2013). Straightlining can be considered as a kind of satisficing that can be provoked by the interviewer. Therefore, interviewer variances should be avoided, unless it reflects differences between the respondents in the small interviewer samples. Otherwise, some interviewers might be more inclined than others to facilitate straightlining.

For the forthcoming analysis, five sets of likert type questions have been used (ESS6). They deal with social trust, resistance against immigrants, perceived threat from immigrants, trust in political institutions and satisfaction about social institutions.

1. Social trust (11-point scale)
 - (a) Most people can be trusted or you can’t be too careful
 - (b) Most people try to take advantage of you, or try to be fair
 - (c) Most of the time people helpful or mostly looking out for themselves
2. Resistance against immigrants (4-point scale)
 - (a) Allow many/few immigrants of same race/ethnic group as majority
 - (b) Allow many/few immigrants of different race/ethnic group from majority
 - (c) Allow many/few immigrants from poorer countries outside Europe
3. Perceived thread from immigrants (11-point scale)
 - (a) Immigration bad or good for country’s economy
 - (b) Country’s cultural life undermined or enriched by immigrants
 - (c) Immigrants make country worse or better place to live
4. Trust in political institutions (11-point scale)
 - (a) Country’s parliament
 - (b) Legal system

Table 4: Example of straightlining

		Respondent 1	straightline?	Respondent 2	straightline?
Set 1	Item 1	7	na.	6	na.
	Item 2	6	0	6	1
	Item 3	6	1	4	0
Set 2	Item 1	4	na.	2	na.
	Item 2	3	0	1	0
	Item 3	2	0	2	0
Set 3	Item 1	8	na.	5	na.
	Item 2	6	0	5	1
	Item 3	6	1	2	0
Set 4	Item 1	6	na.	6	na.
	Item 2	6	1	3	0
	Item 3	6	1	2	0
	Item 4	6	1	1	0
	Item 5	6	1	2	0
Set 5	Item 1	6	na.	6	na.
	Item 2	5	0	4	0
	Item 3	6	0	7	0
	Item 4	5	0	8	0
	Item 5	5	1	8	1
Set 6	Item 1	4	na.	2	na.
	Item 2	4	1	1	0
	Item 3	2	0	1	1
	Item 3	4	0	2	0
		7/17=41%		4/17=24%	

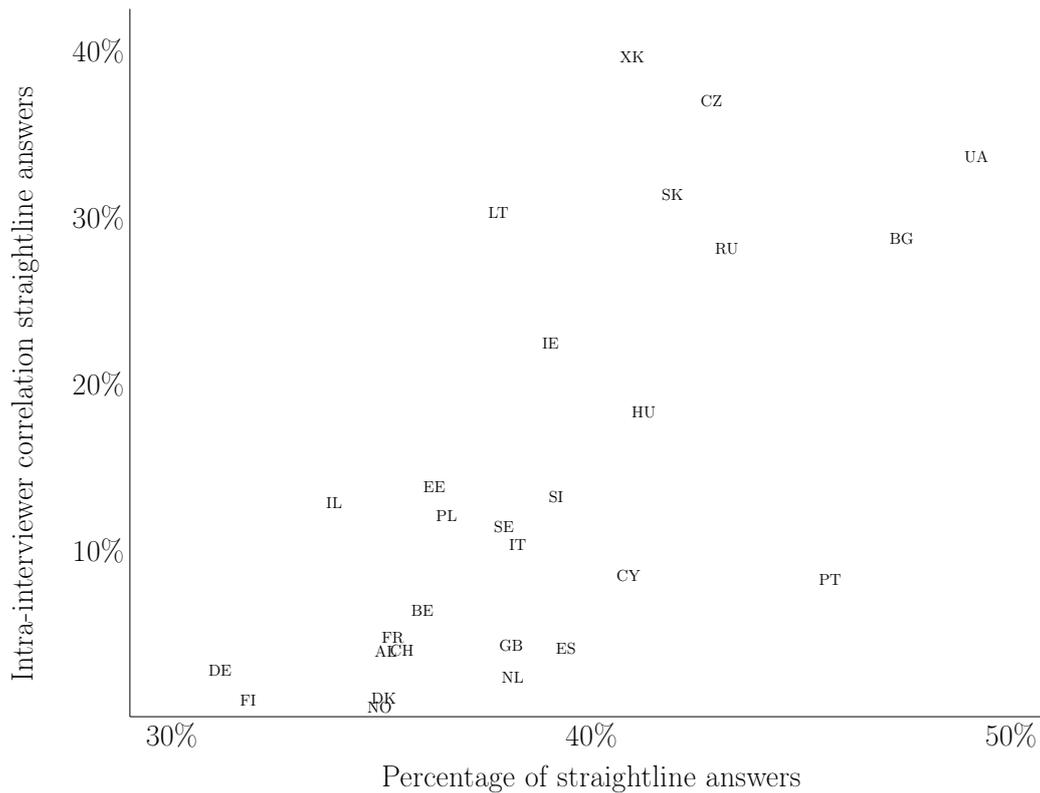


Figure 20: Overall level and intra-class correlation of straightlining per country

- (c) The police
- (d) Politicians
- (e) Political parties

5. Satisfaction about social institutions (11-point scale)

- (a) How satisfied with present state of economy in country
- (b) How satisfied with the national government
- (c) How satisfied with the way democracy works in country
- (d) State of education in country nowadays
- (e) State of health services in country nowadays

Figure 20 shows per country the overall percentage of straightline answers (x-axis). Germany and Finland have the lowest degree of straightlining (30-35%), whereas countries such as Bulgaria, Kosovo, Czech Republic, Slovakia, Portugal and the Russian Federation record the highest levels (40-50%). At the same time, there seems to be a close relationship between the overall level of straightlining in a countries and the degree to which straightlining is attributable to interviewer effects.

A possible hypothesis explaining this relationship may be the fact that some interviewers resort to straightlining in order to short-cut their interviewing efforts. As a result, not only the intra-interviewer correlation increases, also the overall country-specific level of straightlining increases. Countries reporting high percentages and/or interviewer correlation regarding straightlining may be asked to more closely monitor their interviewer force to assess whether some interviewer are responsible for these sizeable figures.

References

- Beullens, K., & Loosveldt, G. (2013). Assessing Interviewer Variance and its Implication for Data Collection, Deliverable 12.10 (Tech. Rep.). KU Leuven, ESS-DACE.
- Biemer, P., & Lyberg, L. (2003). Introduction to Survey Quality. Hoboken (N.J.): Wiley.
- Groves, R., Fowler, F., Couper, M., Lepkowski, J., Singer, E., & Tourangeau, R. (2004). Survey Methodology. New York: Wiley.
- Loosveldt, G., & Beullens, K. (2010). Evaluation of Interviewer Effects in Round Four of the European Social Survey, Deliverable 12.2 (Tech. Rep.). KU Leuven, ESS-DACE.
- Loosveldt, G., & Beullens, K. (2013a). How Long Will it Take? An Analysis of Interview Length in the Fifth Round of the European Social Survey. Survey Research Methods, *7*(2), 69–78.
- Loosveldt, G., & Beullens, K. (2013b). The Impact of Respondents and Interviewers on Interview Speed in Face-to-Face Interviews. Social Science Research, *42*(6), 1422–1430.
- O’Muircheartaigh, C., & Campanelli, P. (1998). The Relative Impact of Interviewer Effects and Sample Design Effects on Survey Precision. Journal of the Royal Statistical Society: Series A (Statistics in Society), *161*(1), 63–77.
- Schnell, R., & Kreuter, F. (2005). Separating Interviewer and Sampling-Point Effects. Journal of Official Statistics, *21*(3), 389–410.