ISER Working Paper Series

Explaining Cross-Country Differences in Contact Rates

Annelies G. Blom
GESIS - Leibniz Institute for the Social Sciences
ISER, University of Essex

No. 2009-08
March 2009

INSTITUTE FOR SOCIAL
& ECONOMIC RESEARCH

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

www.iser.essex.ac.uk

Explaining Cross-Country Differences in Contact Rates

Non-technical summary

When analysing survey data to infer to the general population researchers assume that each selected person (from a probability sample) was contacted and interviewed. However, in every survey there are persons that can either not be contacted or are contacted but not interviewed, i.e. they are nonrespondents to the survey. The extent to which the sample is contacted by the interviewers (i.e. the contact rate) can differ across surveys. In a cross-national survey this contact rate can also differ across countries. Differences in contact rates can pose a problem for cross-national comparisons if the composition of the contacted sample differs across countries.

Countries in a cross-national survey can differ in (1) population composition, (2) the way fieldwork is carried out and (3) the effect of (1) and (2) on the contact rate. This paper investigates which fieldwork factors are associated with the contact probability of a sampled person in the European Social Survey (ESS). We look at factors associated with contact within seven ESS countries (Belgium, Finland, Greece, Ireland, Portugal, Spain and the UK) and at factors associated with differences in contact rates between two countries. For the cross-national analysis we aim to disentangle which part of the difference in contact rates is due to differences in population composition, which part is due to differences in the way the ESS was implemented and which part is due to differences in the effect of the population composition and fieldwork implementation on the contact rate.

Explaining Cross-Country Differences in Contact Rates

Annelies G. Blom

Abstract

In the European Social Survey (ESS) contact rates differ across countries. These differences are broadly due to (1) differences in survey implementation, (2) differences in population characteristics associated with contact propensity and (3) differences in the association between 1 or 2 and contact propensity. This paper investigates correlates of contact within and across ESS countries by decomposing cross-country differences in predicted mean contact propensities into (population and fieldwork) characteristics effects, coefficients effects and a pseudo-interaction effect. The findings shed light on the cross-national comparability of the manipulable aspects of the contacting process. In addition, we distinguish factors explaining within-country contact propensity from factors explaining cross-country differences.

Keywords: decomposition analysis, counterfactual analysis, survey methods, contact data, comparative research, European Social Survey

JEL codes: C81, C83

Contact: Annelies Blom, GESIS - Leibniz Institute for the Social Sciences,
P.O. Box 122155, 68072 Mannheim, Germany. Email: annelies.blom@gesis.org

Acknowledgements
Part of this paper is based on work carried out during my visit to the European Centre for Analysis in the Social Sciences (ECASS) at the Institute for Social and Economic Research, University of Essex, supported by the Access to Research Infrastructure action under the EU Improving Human Potential Programme. I am most grateful to Annette Jäckle for investigating the usefulness of decompositions for survey methods with me, to Peter Lynn for guidance and encouragement, to Mark Bryan, Cheti Nicoletti and Stephen Jenkins for advice on decomposition methods, and to Mick Couper for feedback on a draft version.

INTRODUCTION

Cross-national comparisons of fieldwork outcomes uncover differences in contact rates across countries (e.g. de Leeuw & de Heer, 2002; Billiet et al., 2007). These differences are broadly due to (1) differences in the way the survey was implemented, (2) differences in the distribution of population characteristics associated with the propensity to be contacted and (3) differences in the association between 1 or 2 and the propensity to be contacted (e.g. making contact attempts on Sundays may be very helpful in one country but have no effect in another country) (Blom et al., 2009). Conventional methods for analysing the differences in contact rates examine the *processes* within each country, *differences in processes* across countries (e.g. by comparing the coefficients of separate within-country models) or look at country-level differences in response outcomes (e.g. Goyder, 1985; Lipps & Benson, 2005). Differences across countries in the *distributions of survey and population characteristics* (1 and 2) are usually ignored.

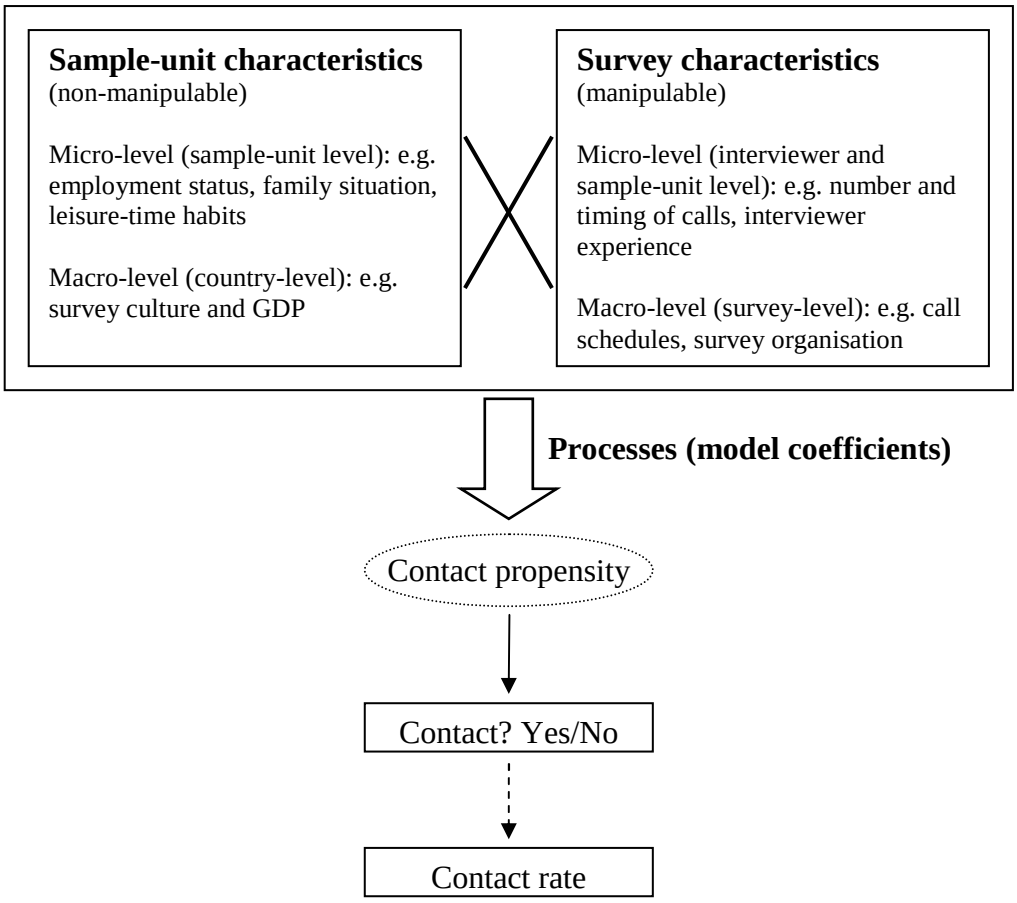
This paper analyses differences in contact rates across seven countries in the first round of the European Social Survey (ESS). The ESS collects detailed contact and neighbourhood information for each sample unit, i.e. respondents and non-respondents. These data are especially well-suited for modelling the contacting process in each ESS country. Yet, the ESS country populations vary in the distribution of these population characteristics. In addition, even though some fieldwork strategies are standardised in the ESS, the survey implementation varies across countries.

We aim to identify characteristics associated with differences in contact propensity across countries. Differences in contact rates between two countries are decomposed into differences due to differential survey characteristics, differential population characteristics, differential coefficients and a pseudo-interaction. The findings shed light on the extent to which the manipulable aspects of the contacting process are comparable across countries and whether standardising fieldwork procedures would lead to equivalent contact rates and sample compositions. At the same time we aim to distinguish factors that explain contact propensity within a country from those that explain differences in contact rates across countries.

CONCEPTUAL MODEL

In a first step we describe the conceptual model used for analysing differences in contact rates across countries (see Figure 1). The contact rate achieved in any particular survey is the result of an underlying (or latent) contact propensity for each sample unit. This contact propensity is a random variable and is a function of each sample unit's characteristics, the way a survey is implemented and the way a sample unit (with given characteristics) reacts to a survey request (of a given implementation).

Figure 1: Determinants for cross-country differences in contact rates



The sample unit characteristics consist of micro-level individual characteristics (such as employment status, family situation, living situation and leisure-time habits) and macro-level country characteristics (such as survey culture and GDP). Both, macro and micro-level sample unit characteristics influence a sample unit's accessible at-home patterns (Groves & Couper, 1998). The survey characteristics can also be divided into micro-level survey characteristics (such as the number, timing and mode of contact attempts and interviewer experience) and macro-level survey characteristics

(such as the issuing of advance letters and the rules, management structure and traditions of a survey organisation). Sample unit characteristics are non-manipulable, i.e. when implementing a survey they cannot be influenced to maximise the contact rate. Survey characteristics, however, are (at least to some extent) manipulable.

Regarding differences in contact rates across countries, survey characteristics can vary across countries depending on choices made by the researcher, survey manager and interviewer. The distribution of sample unit characteristics, whilst not manipulable, can also vary across countries; for example due to differences in population structures, histories and cultures (Lyberg & Dean, 1992). In addition, countries may differ in the extent to which these characteristics are associated with contact. This means that the processes that translate the sample unit and survey characteristics into an underlying contact propensity (and the subsequent contact rate) may differ. Differences in contact rates across countries can therefore be influenced by differences in manipulable and non-manipulable characteristics, as well as differences in processes.

According to Groves and Couper (1998) three factors determine whether sample units in face-to-face household surveys are contacted by the interviewer: "a) whether there are any physical impediments (locked gates, locked apartment entrances) that prevent visiting interviewers to alert the household to their presence, b) when household members are at home, and c) when and how many times the interviewer chooses to visit." (Groves & Couper, 1998, p.80) Of these, the first two factors, the physical impediments and the at-home patterns, are outside the influence of the researcher, i.e. they are not manipulable. The last factor, i.e. the effort made by the interviewer or the fieldwork organisation is manipulable. In analysing differences in contact propensity across countries we distinguish between factors that are naturally different across countries and those that can be influenced. This serves to find out whether standardised fieldwork would lead to comparable contact rates across countries.

There is one caveat though that necessitates mentioning: macro-level manipulable factors, such as advance letters, incentives and interviewer instructions, show no

variation within ESS countries analysed here, though they do vary across countries[1]. Since the macro-level factors are invariant within each country, their effect cannot be identified. (Only if the sample was a random sample of a much larger number of countries would this be possible, e.g. in a multi-level model (see Hox, 2002)). Consequently, the analysis at hand cannot draw conclusions about the effects of the macro-level factors.

Based on this conceptual model our analyses attempt to disentangle the cross-national muddle of differential contact rates at the example of the ESS. The analyses decompose differences in contact rates across countries into differences due to

- differential micro-level manipulable characteristics (i.e. differences in survey implementation) and micro-level non-manipulable characteristics (i.e. differences in sample unit characteristics)
- differential processes and
- an interaction between the differences in characteristics and the differences in processes. (For example making contact attempts on a Sunday may be more effective in country A than in country B, and country A may make more Sunday contact attempts than country B.)

Consequently, the paper not only considers differences in contact rates but also differences in the composition of the contacted sample. Two countries could have equal contact rates, but very different sample compositions (due to different fieldwork processes or population characteristics).

DETERMINANTS OF CONTACT

A major determinant of contact in the literature is the timing of contact attempts. Calls on a weekday evening or at the weekend have been found to be most effective for making contact (see Groves & Couper, 1998; Purdon et al., 1999; Stoop, 2005; Lipps & Benson, 2005). The relationship between the number of calls made and successful contact is less well documented, largely due to measurement problems in the absence

[1] A number of countries have carried out fieldwork experiments in different ESS rounds (e.g. incentive experiments in Poland and the UK and a fieldwork schedule experiment in Switzerland). However, none of these countries/rounds are included in the analysis of this paper. For general information on the use of incentives, advance letters, fieldwork period and other ESS implementation issues please consult the ESS Documentation Report for round 1 (ESS, 2003).

of call-level contact data. Lynn and Clarke (2002) for example urge survey practitioners to record separately the number of contact attempts leading up to first contact and those leading up to a refusal or interview, rather than just a total number of contact attempts.

Interviewer-related variables are also prominently associated with the probability of making contact. O'Muircheartaigh and Campanelli (1999) found that interviewers who are good at gaining cooperation are also good at making contact. Furthermore, the workload of interviewers is generally found to be negatively associated with contact rates. Botman and Thornberry (1992) for example argue that increases in workload mean that interviewers have less time to attempt contact at the productive times of day (e.g. in the evening). Lipps and Benson (2005) report that interviewers' habit of using the phone to attempt contact is related to their ability to make contact with the household.

Non-manipulable factors associated with contact propensity are measures of the sample unit's at-home patterns or impediments to access (Groves & Couper, 1998). If sample units are more likely to be at home at times that interviewers tend to call (e.g. on weekdays during daytime), they are also more likely to be contacted. Contact is found to be easier amongst the elderly and more difficult amongst young people (Groves & Couper, 1998; de Leeuw & de Heer, 2002; Stoop, 2005; Lipps & Benson, 2005). Single persons and those in employment are less likely to be contacted, whilst households with young children have a higher contact propensity (Groves & Couper, 1998; Lynn & Clarke, 2002; Stoop, 2005). Living in multi-unit housing is associated with a lower probability of contact as is living in urban areas (Campanelli et al., 1997; Groves & Couper, 1998; Stoop, 2005). However, Groves and Couper (1998) note that the urbanicity effect disappears once multi-unit housing structures are controlled for. In addition, the effect of multi-unit housing might be due to more single people living in these types of buildings. Lipps and Benson (2005) further find that those living in low quality housing are less likely to be contacted. Impediments to access in face-to-face surveys are typically locked apartment buildings, intercom and locked gates (Groves & Couper, 1998; Lipps & Benson, 2005).

DATA AND OPERATIONALISATION

To illustrate the components of cross-national differences in contact rates, data from the first round of the ESS are used. The ESS is a biennial cross-national face-to-face survey of social and political attitudes across more than twenty countries in Europe, and was first fielded in 2002. In addition to the data collected in the interview, the ESS interviewers use standardised contact forms to collect information on the contacting and cooperation process and on the neighbourhood of all sample units. Seven ESS countries, where contact and neighbourhood data of sufficient quality were available and where the distributions of the dependent variable varied sufficiently, are examined. These are Belgium, Finland, Greece, Ireland, Portugal, Spain and the UK. In a subset of these countries (Belgium, Finland, Portugal and the UK) an interviewer questionnaire was administered to the ESS interviewers. Though the questionnaire primarily gathered information on interviewer strategies for gaining cooperation, it also collected general interviewer characteristics. These characteristics are used as additional variables for modelling contact in these four countries.

Unfortunately, for some countries the round 1 ESS contact data contained a large number of cases with missing items; especially in the neighbourhood variables. In later rounds the quality of the contact data improved. Despite this we decided to use round 1 contact data because the additional interviewer questionnaire was only administered in the first round. The seven countries selected for analysis here were amongst other reasons chosen because their contact data were of sufficient quality; i.e. there were no major inconsistencies or missing cases in the outcome variables and in less than 2% of cases were all values on the neighbourhood variables missing (Appendix A1 lists reasons for excluding countries from our analysis).

The ESS contact data enable us to examine detailed measures of the contacting process in a cross-national context. Given call-level information (i.e. on each contact attempt) is available for each sample unit we can derive many of the manipulable factors that other researchers have found to be relevant in predicting contact. Table 1 lists all the (derived) variables examined.

Table 1: Determinants of contact and their operationalisation

	Variable description
Manipulable factors	• Whether contact was ever attempted on a weekday evening, on a Saturday, on a Sunday • Natural log of the number of in-person contact attempts until contact (for final non-contacts total number of in-person contact attempts) • Interaction terms: ever attempted in the evening * log (number of in-person contact attempts) ever attempted on a Saturday * log (number of in-person contact attempts) ever attempted on a Sunday * log (number of in-person contact attempts) • Interviewer ability: the ESS interviewer cooperation rate • Workload: number of sample units the interviewer worked on • The percentage of sample units that the interviewer attempted by phone • Interviewer contacting strategy: the percentage of contact attempts the interviewer made on a weekday evening, on a Saturday, on a Sunday
Non-manipulable factors	• Whether there was any intercom at the housing unit • Whether there were any security features at the housing unit • Whether housing unit was a farm or single-unit housing (omitted category is multiunit housing) • Physical state of buildings in the area: satisfactory state or bad state (omitted category is good state) • State of the sampled housing unit compared to other housing units in the area: better or worse (omitted category is same) • Urbanicity: percentage of farms or single housing units in the assignment of the interviewer making the first contact attempt

Note on interviewer-related variables: where more than one interviewer worked on a sample unit the derived interviewer variable relates to the interviewer who made contact or, if no contact was achieved, to the last interviewer; unless stated otherwise

A central set of variables in our analysis is the timing and number of contact attempts. The importance of the timing of a contact attempt may depend on the number of contact attempts until contact is made or, in case of a final non-contact outcome, total number of contact attempts. This means that we assume that the association between successful contact and ever calling at the weekend/in the evening differs when one contact attempt is made from when ten contact attempts are made (see also Philippens et al. (2003) on contact probabilities in the ESS and timing of call at the first, second and further calls). Furthermore, given the left-skewed distribution of the number of contact attempts we take the natural logarithm of the number of contact attempts (Goyder, 1985; Olson, 2006).

Regarding interviewers' strategies for achieving contact there are a number of factors that we can account for in the ESS. We derive indicators of interviewer cooperation rates on the ESS, the ESS workload of interviewers and the percentage of sample units an interviewer attempted by phone. In addition, we assume that successful interviewers attempt contact on weekends, Saturdays and Sundays for a larger proportion of their calls. Therefore, we derive indicators of the interviewers' calling strategies with regards the timing of their contact attempts.

With respect to the non-manipulable factors the ESS contact data are a little more limited compared to the manipulable factors. Available measures primarily describe impediments to access or are indirect measures of the sample unit's socio-economic status. Direct measures of the sample unit's age, gender, education, employment etc. are not available. (Though some countries do have information on this from their sampling frame, this is not collected in a standardised way or made available via the data archive.) The lack of direct measures of sample unit characteristics is not ideal; however, it should not be detrimental. While including non-manipulable characteristics improves the accuracy of the models, they merely describe differences between countries, but cannot inform decisions about how to optimise fieldwork efficiency (though they can inform adjustments for non-response bias). One should further note that in their operationalisation it is often unclear whether non-manipulable factors are impediments to access or whether they describe a sample unit's characteristics and therefore also their at-home pattern. For example, an intercom at a housing unit is an impediment. However, since multi-unit houses are more likely to have intercoms than single-unit houses, intercoms are correlated with the sample unit's socio-demographic characteristics, which in turn are related to their at-home pattern. At-home patterns and impediments can therefore be influenced by the same socio-environmental and socio-demographic attributes. We examine two measures of barriers to entry, presence of an intercom and any security features at the house. Furthermore, we look at three non-manipulable factors that were found to be related to contact propensity: the type of housing, the physical state of buildings in the area and the state of the sampled housing unit compared to other housing units in the area. Finally, we control for urbanicity with a derived measure of the percentage of farms and single housing units in an interviewer's assignment.

METHOD

The analyses consist of three components. First, we look at distributional differences across countries in the manipulable and non-manipulable factors that might be associated with people's contact propensities. Second, we model contact propensity in each country in a logit model and discuss differences in the optimal specification across countries. Contact is defined as a dummy variable where 1 denotes that in-person contact with the household was established. Finally, we carry out the central analysis of this paper: a decomposition of differences across countries in contact rates into (1) differences due to differences in manipulable survey and non-manipulable sample unit characteristics (characteristics effect), (2) differences due to the differential effect of these characteristics on the contact propensity (coefficients effect) and (3) differences due to a pseudo-interaction between (1) and (2). In addition, a detailed decomposition of the characteristics effect looks at which characteristics contribute most to the differences in contact propensities across countries. Since the decomposition method (also known as counterfactual analysis) is not a standard analysis technique in survey methodology (cf. Nicoletti & Buck, 2004) it is described in the following.

Aggregate decomposition

The general aim of the aggregate decomposition is to separate out differences in the outcome (i.e. contact propensity) into differences due to distributional differences in the sample characteristics (i.e. the variables included in the model) and those due to differences in the coefficients.

Contact is the binary dependent variable. In a logit model, the predicted contact propensity $\hat{y}_i$ for each sample unit *i* is estimated by

$$\hat{y}_i = F(\beta, X) = \frac{1}{1+e^{-(\beta_1+\beta_2 X_i)}},$$

where β_2 is the vector of all coefficients estimated for the matrix of characteristics X_i.
A country's predicted contact rate then corresponds to the mean predicted contact propensity $\bar{y}$.

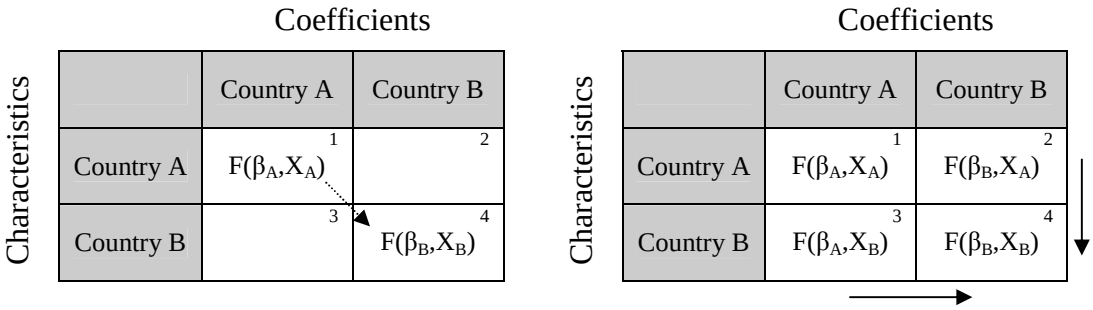
$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} \hat{y}_i$$

The difference in mean predicted contact propensity between two countries A and B is a function of the coefficients and the characteristics in the country models.

$$\bar{y}_A - \bar{y}_B = F(\beta_A, X_A) - F(\beta_B, X_B). \quad (1)$$

Figure 2 below illustrates the country difference in mean predicted contact propensity that we decompose. We aim to describe how the mean predicted contact propensity would change, if both the characteristics *X* and the coefficients β changed from those of country A to those of country B (dotted arrow in panel 1). We investigate the extent to which the country difference in contact rate is due to a difference in characteristics and to what extent it is due to a difference in coefficients. For this we estimate the mean predicted propensity associated with the characteristics of country A evaluated at the coefficients of country B (box 2 in panel 2) and the contact propensity associated with the characteristics of country B evaluated at the coefficients of country A (box 3 in panel 2). Changing the coefficients from those of country A to those of country B (box 1 – box 2 or box 3 – box 4) describes coefficients effects. Changing the characteristics from those of country A to those of country B (box 1 – box 3 or box 2 – box 4) describes characteristics effects.

Figure 2: Illustration of differences in mean predicted contact propensities



In mathematical notation we expand the right-hand side of equation (1) by adding and subtracting $F(\beta_B,X_A)$, $F(\beta_A,X_B)$ and $F(\beta_B,X_B)$.

$$\begin{aligned}&\bar{y}_A - \bar{y}_B \\ &= [F(\beta_A, X_A) - F(\beta_B, X_B)] + [F(\beta_B, X_A) - F(\beta_B, X_A)] + [F(\beta_A, X_B) - F(\beta_A, X_B)] + [F(\beta_B, X_B) - F(\beta_B, X_B)]\end{aligned} \quad (2)$$

By rearranging the addends in (2) we get (Even & Macpherson, 1993)

$$\begin{aligned}&\bar{y}_A - \bar{y}_B \\ &= \underbrace{[F(\beta_B, X_A) - F(\beta_B, X_B)]}_{\text{characteristics effect}} + \underbrace{[F(\beta_A, X_B) - F(\beta_B, X_B)]}_{\text{coefficients effect}} + \underbrace{[F(\beta_A, X_A) - F(\beta_B, X_A) + F(\beta_B, X_B) - F(\beta_A, X_B)]}_{\text{pseudo-interaction effect}}\end{aligned} \quad (3)$$

The difference in the mean predicted contact rate between countries A and B is therefore composed of three parts. The first addend describes the difference in predicted probabilities arising from a change in sample characteristics assuming that the coefficients remain unchanged (*characteristics effect*). This characteristics effect indicates how many percentage points the predicted contact rate in country B would have been higher (lower) if country B had had the manipulable and non-manipulable characteristics of country A. The second addend describes the difference in predicted probabilities arising from a change in coefficients assuming that the characteristics remain unchanged (*coefficients effect*). This coefficients effect indicates how many percentage points the predicted contact rate in country B would have been higher (lower) if country B had had the coefficients of country A. The third addend reflects the extent to which these two assumptions are simultaneously broken; i.e. that some variables simultaneously have different coefficients and different distributions. This is similar to an interaction effect in regression analyses. Note that the predicted characteristics and coefficients effects differ depending on whether $\bar{y}_A - \bar{y}_B$ or $\bar{y}_B - \bar{y}_A$ is estimated. In the former the characteristics effect expresses that we evaluate a change in characteristics from country A to country B at the coefficients of country B. In contrast, in the latter comparison the characteristics effect expresses that we evaluate a change in characteristics from country B to country A at the coefficients of country A. Equivalently, a change in coefficients between country A and B is evaluated at the characteristics of country B in the first and at the characteristics of country A in the second case. The choice of reference country is therefore central to decomposition analyses.

Choosing a reference country

The comparison of two subgroups as described in the aggregate decomposition above is the standard type of comparison. This standard decomposition makes sense when comparing subgroups that are inherently binary, e.g. men and women. However, comparing contact rates across countries is not a binary problem. Though one can compare all possible pairs of countries with each other, this is rather tedious and unintuitive when more than three countries are involved. Comparing seven ESS countries with each other (plus the comparisons with the opposite reference group) yields 42 comparisons, with each a characteristics effect, a coefficients effect and a

pseudo-interaction effect. In addition to the sheer number of results that need to be interpreted, one would also need to make sense of the possibly different results depending on which two countries are involved in the comparison.

In this paper we use the UK as the reference country in all the decompositions. The rationale behind this is that the ESS was in many ways modelled on the UK-way of designing and implementing surveys: the questionnaire is developed in British English, the maximum number of sample units to be assigned to an interviewer is 48, which corresponds to the norm in the UK's largest social survey organisation; interviewers' contacting schedules need to contain at least one evening and one weekend call to each sample unit, which is a requirement based on research findings from UK studies (Campanelli et al., 1997; Purdon et al., 1999; Groves and Couper (1998) found similar associations for the US). Based on equation (3) and with the UK as the reference country the decomposition of the difference in contact rates between the UK and another country *C* is

$$\begin{aligned}&\bar{y}_{UK}-\bar{y}_C \\ &=[F(\beta_C,X_{UK})-F(\beta_C,X_C)]+[F(\beta_{UK},X_C)-F(\beta_C,X_C)]+[F(\beta_{UK},X_{UK})-F(\beta_C,X_{UK})+F(\beta_C,X_C)-F(\beta_{UK},X_C)]\end{aligned} \quad (4)$$

Detailed decomposition of the characteristics effect

The aggregate decomposition decomposes the difference in contact rates between the UK and each of the six other countries into a characteristics effect, a coefficients effect and a pseudo-interaction effect. We can further decompose the characteristics effect by looking at the contribution of individual characteristics (as a part of the whole characteristics effect). While there are different ways of looking at the contribution of separate characteristics, the by now most common and simplest is to evaluate the contribution of each characteristic at their sample mean and weight the characteristics effect by this contribution. The proportion of the characteristics effect due to the r_{th} variable then is (Even & Macpherson, 1993; Yun, 2004)

$$F(\beta_C,X_{UK})-F(\beta_C,X_C)=\left[\sum_{r=1}^{r=K}\frac{\beta_C^r\left(\bar{X}_{UK}^r-\bar{X}_C^r\right)}{\beta_C\left(\bar{X}_{UK}-\bar{X}_C\right)}\right]\left[\overline{F(\beta_C,X_{UK})}-\overline{F(\beta_C,X_C)}\right]. \quad (5)$$

The total characteristics effect is the sum of the effects of each separate characteristic *r*. The contributions of the dummies of a categorical variable need to be examined together, because their contribution will depend on the choice of omitted category.

The sum of the characteristics effects of the categories (dummies), however, is invariant to the choice of omitted category (Yun, 2005).

Interpreting the decompositions

The results from decompositions can yield new insights into explanations for differences in contact rates across countries. The aim in this paper is to find significant characteristics effects of manipulable survey characteristics that have an influence on the differential contact propensity. The decompositions then show what the contact rate in Belgium / Finland / Greece / Ireland / Portugal / Spain would have been, had these countries adopted similar fieldwork strategies as the UK.

Where the aggregate decompositions yield large coefficients effects the interpretation is a little intricate. The coefficients effects indicate that the characteristics included in the model either have a different effect on contact propensity across countries or that these characteristics do not explain the difference in contact rate, i.e. at least one relevant variable was omitted from the model. In either case, the manipulable characteristics that were measured then do not explain the difference in contact rate.

Interpreting decompositions with large pseudo-interaction effects is also difficult. The interaction term reflects the extent to which the two assumptions of the characteristics and coefficients effects are simultaneously broken i.e. some variables simultaneously have different coefficients and different distributions. As a result, in models where a large proportion of the difference in contact rate is due to a pseudo-interaction the characteristics and the coefficients effects are difficult to interpret.

RESULTS

Our data allow us to fit logit models of contact in seven countries (Belgium, Finland, Greece, Ireland, Portugal, Spain and the UK) and decompose the differences in contact rates between the UK and the other six countries. Only cases where at least one contact attempt was made in person were included in the analyses. The reason for this is that we define contact as in-person contact with the household. Cases where no in-person contact attempts were made, therefore, had no chance of being contacted. Furthermore, if all neighbourhood characteristics were missing for a sample unit the

case was dropped from the analyses.[2] In total, 4.5% of cases (960 out of 21,433) had to be dropped because of no in-person contact attempts to the case or missing neighbourhood characteristics. The largest proportion of cases had to be excluded in Finland, where 442 cases (16.5%) were never attempted in person. In the other countries the proportion of cases excluded from the analysis ranged from 0.2% in Spain to 6.4% in Belgium.

In the following we present the results of our analysis. The descriptive statistics and the logit models are only briefly presented (Table 2 and Table 3). They are discussed in detail together with the results from the decomposition analysis. Only variables that were significant at the 5-percent level in the logit model of at least one country were included in the analyses. The following variables were not statistically significant.

- Type of housing: whether the sampled person lived in a single-unit house (including farms) or a multi-unit house. The urbanicity of the area, i.e. the percentage single-unit houses and farms in an interviewer's assignment (excluding re-issues), however, was significantly associated with probability of contact. Once the type of housing was excluded from the models, the indicator for whether there was an intercom at the housing unit showed a significant association with the probability of contact.
- Security features: whether there were any security features at the house, such as security lights, bars on the windows or alarm systems.
- Contact attempts by phone: whether contact with the sample unit was ever attempted by phone. A variable indicating the percentage of cases that the interviewer attempted by phone, however, was found to be significant in Finland.

[2] In total 763 cases had to be excluded from the analysis because no in-person contact attempt had been made. An additional 197 cases were excluded due to missing neighbourhood information.

	UK	Belgium	Finland	Greece	Ireland	Portugal	Spain	Total
No in-person contact attempt	2	171	442	7	77	60	4	763
Neighbourhood vars missing	77	32	9	0	78	0	1	197

Table 2: Descriptive statistics

	UK	Belgium	Finland	Greece	Ireland	Portugal	Spain	Total
Contact rate (%)	95.0	95.3	95.7	98.0	93.3	96.4	89.7	94.6
Manipulable factors								
Timing of contact attempts								
% of cases ever attempted on a weekday daytime (before 5pm)	88.8	63.8	82.2	62.8	75.0	54.7	77.4	72.9
% of cases ever attempted on a weekday evening (after 5pm)	23.0	29.6	17.0	28.5	23.9	31.5	35.7	27.2
% of cases ever attempted on Saturday	25.2	22.8	2.9	21.7	20.4	50.0	19.1	22.7
% of cases ever attempted on Sunday	9.0	3.0	1.6	20.6	4.3	32.7	12.2	11.4
Mean number of in-person contact attempts per sample unit	2.4	1.4	1.1	1.6	1.7	2.4	2.1	1.8
Standard deviation	2.2	1.0	0.5	1.2	1.2	1.8	1.7	1.5
Mean interviewer cooperation rate (%)*	47.5	59.6	72.6	77.6	70.3	72.4	52.4	63.5
Standard deviation (%)	14.8	19.7	11.3	21.0	16.3	18.9	20.0	21.0
Missing (n)	*2*			*1*	*1*			*4*
Mean interviewer workload (cases worked on)*	29.1	21.2	23.4	25.1	32.2	47.3	47.4	32.0
Standard deviation	11.7	10.8	8.9	12.7	14.1	20.1	39.8	22.3
Mean % of cases in workload phoned by the interviewer*	0.4	8.9	90.7	0.7	1.9	4.4	0.5	12.4
Standard deviation (%)	1.2	14.4	8.8	2.3	4.2	7.7	1.7	28.7
Interviewer calling strategy								
Mean % of an interviewer's in-person calls on a weekday daytime (before 5pm)*	67.8	55.1	79.1	51.3	61.2	33.8	58.0	58.5
Standard deviation (%)	18.3	24.8	12.9	22.1	20.9	15.7	20.4	23.1
Mean % of an interviewer's in-person calls on a weekday evening (after 5pm)*	14.2	23.9	16.5	20.4	16.9	17.8	22.6	19.0
Standard deviation (%)	11.5	18.9	11.5	13.5	15.2	10.9	14.7	14.5
Mean % of an interviewer's in-person calls on a Saturday*	13.2	17.6	2.9	14.3	13.3	28.4	12.1	14.3
Standard deviation (%)	11.8	21.1	5.7	14.3	11.8	13.2	13.5	15.2
Mean % of an interviewer's in-person calls on a Sunday*	4.7	2.4	1.4	13.9	2.6	19.9	7.3	7.1
Standard deviation (%)	7.6	7.1	2.8	15.2	4.0	12.3	9.8	11.1

Non-manipulable factors	**UK**	**Belgium**	**Finland**	**Greece**	**Ireland**	**Portugal**	**Spain**	**Total**
Intercom/entry phone at the address (%)	9.9	21.9	10.3	38.1	4.4	42.3	68.1	28.0
State of buildings in the area (%)								
Good state	63.0	67.0	72.9	40.2	73.5	48.8	51.3	59.3
Satisfactory state	33.6	28.2	24.4	50.3	22.9	44.1	40.5	35.1
Bad state	3.4	4.8	2.7	9.5	3.6	7.1	8.2	5.7
Missing (n)	*1*	*26*	*18*		*30*	*3*	*62*	*140*
State of sampled address compared to the area (%)								
Better	11.0	17.6	15.7	20.9	21.2	16.7	16.2	17.0
Same	82.2	74.1	77.2	68.1	70.7	75.4	77.6	75.1
Worse	6.7	8.3	7.2	11.0	8.1	7.9	6.2	7.9
Missing (n)	*4*	*29*	*29*		*33*	*3*	*69*	*167*
Urbanicity								
Mean % cases of multi-unit housing in interviewer assignment	16.6	14.6	37.2	47.2	2.1	41.4	67.6	32.0
Standard deviation (%)	17.3	17.6	20.9	30.8	5.2	24.1	29.1	31.0
Mean % cases of single-unit housing in interviewer assignment	80.8	80.0	55.5	52.2	81.7	53.6	29.4	62.4
Standard deviation (%)	18.2	17.5	18.7	30.8	15.0	22.6	26.7	29.2
Mean % cases of farms in interviewer assignment	0.5	2.3	6.1	0.4	11.5	0.9	2.9	3.4
Standard deviation (%)	2.0	4.4	7.7	2.6	14.4	2.6	10.7	8.6
Base	*3,592*	*2,984*	*2,276*	*3,219*	*2,995*	*2,132*	*3,288*	*20,486*

Note: * weighted for all sample units

Table 3: Logit models of contact

	UK	**Belgium**	**Finland**	**Greece**	**Ireland**	**Portugal**	**Spain**
	b	b	b	b	b	b	b
Ever called in the evening	1.747 *	0.563	0.345	-0.377	0.328	-5.128 ***	-0.231
Ever called on a Saturday	4.635 ***	-0.039	-0.989	-2.072	-0.142	-0.857	0.156
Ever called on a Sunday	-0.423	0.466	1.327	-2.516	-1.504 *	-5.412 ***	0.798
Log (# in-person contact attempts)	-0.619 **	-1.624 ***	-1.597 ***	-4.483 ***	-2.479 ***	-6.821 ***	-1.193 ***
Log(attempts)*evening	-1.446 **	-0.969 **	-1.792 **	-0.008	-0.465	3.108 ***	-0.451
Log(attempts)*Saturday	-2.698 ***	-0.123	-0.175	1.266	0.262	0.573	-1.122 ***
Log(attempts)*Sunday	-0.092	-0.694	-2.215	1.392	1.398 *	3.147 ***	-1.28 **
Interviewer workload	0.001	-0.021 *	-0.005	0.005	0.01	0.008	0.001
Interviewer ability (cooperation rate)	0.005	0.033 ***	0.053 ***	0.024 ***	0.009	-0.02 *	-0.007
Interviewer calling strategy:							
Evening calling	0.034 **	0.006	0.016	0.004	0.019 **	-0.04 *	0.011
Saturday calling	-0.006	0.001	-0.026	0.006	0.02 *	-0.018	0.005
Sunday calling	0.01	-0.011	0.081	-0.003	0.012	-0.034	0.007
Interviewer phone contacting	-0.031	0.01	-0.04 *	-0.022	0.024	-0.002	0.005
Any intercom	-0.433	-0.233	0.158	0.555	0.127	-0.903 **	0.346
Physical state of building:							
Satisfactory	-0.046	-0.189	-0.142	-0.192	-0.055	-0.316	-0.047
Bad	-1.165 **	-0.035	-1.408 *	0.843	-0.742	0.098	-0.21
Comparative state of building:							
Better	0.191	1.258 ***	-0.147	0.641	0.976 ***	0.185	0.001
Worse	0.465	-0.071	0.162	-0.235	0.485	-0.876	-0.591
Urbanicity (farms, single-unit housing)	0.001	0.003	0.013 *	-0.01	0.021 **	-0.015 *	0.003
Constant	3.457 ***	2.199 ***	2.846	7.619 ***	0.798	17.657 ***	4.087 ***
Chi^2	512.499	293.206	195.093	290.437	437.913	239.01	930.042
Pseudo R^2	0.362	0.258	0.243	0.468	0.299	0.361	0.425
N	3590	2984	2276	3218	2994	2132	3288

Legend: *p<0.05; **p<0.01; ***p<0.001

Logit models

In a next step we fitted the same logit models in all countries to compare the coefficients of these models across countries. Modelling decisions were based on the literature discussed above. Categorical variables were included as dummies and the category that was modal across most countries was omitted. A small amount of missing values was captured in the modal (omitted) category. Other item-missing cases were list-wise deleted; since their number is small (seven cases across the seven countries) this is unlikely to have a significant impact on the analysis. In the case of the interviewer calling strategy and urbanicity the variables are not strictly speaking categorical; nevertheless, to prevent multicollinearity one of the composite variables was omitted. In Table 3 the logit coefficients of seven models of contact are displayed; one model for each country included in the analysis.

The χ^2 tests and McFadden's R^2 show that the model fit is good yet varies across countries. In addition, Wald tests and Likelihood Ratio tests were performed to confirm the significance of the variables in the model. While there are some similarities across countries in the relevant variables, there are also many differences. Overall, the logit models give insight into manipulable and non-manipulable factors associated with contact propensity in the seven countries. However, they cannot explain which factors are associated with *differences* in contact propensity *across* countries. The interpretation of the logit models is limited since they only reveal differences in the processes correlated with contact but not differences in the characteristics. Finding out whether differences in contact rates are due to differences in characteristics, due to differences in coefficients or both is the purpose of the decomposition analysis that follows.

Decompositions

The results from the decompositions between the UK model and each of the other six countries are listed in Table 4. The first part of the table displays the mean predicted probability of contact for the UK model (row 1), for the country models (row 2) and the mean predicted contact probability when evaluating each country's characteristics at UK coefficients (row 3) and when evaluating UK characteristics at the coefficients of each country (row 4). The fifth row displays the difference in mean predicted contact probabilities (i.e. contact rates) between the UK model and the model in each of the other countries. These are the differences in contact rates that the decompositions aim to disentangle.

Table 4: Results from the aggregate and detailed decompositions $\bar{y}_{UK} - \bar{y}_{C}$

	Belgium	**Finland**	**Greece**	**Ireland**	**Portugal**	**Spain**
Predicted probabilities	%	%	%	%	%	%
(1) (β_{UK} ,X_{UK})	95.0	95.0	95.0	95.0	95.0	95.0
(2) (β_{C}, X_{C})	95.3	95.7	98.0	93.4	96.4	89.7
(3) (β_{UK} ,X_{C})	98.2	83.2	98.3	98.3	96.1	96.5
(4) (β_{C}, X_{UK})	82.6	90.1	91.3	84.1	98.4	87.6
(5) (β_{UK}, X_{UK}) – (β_{C}, X_{C})	-0.23	-0.70	-3.00	1.69	-1.35	5.38
Aggregate decomposition	% points	% points	% points	% points	% points	% points
Characteristics: (β_{C},X_{UK})-(β_{C},X_{C})	-12.63	-5.63	-6.77	-9.22	2.00	-2.11
Coefficients: (β_{UK},X_{C})-(β_{C},X_{C})	2.96	-12.52	0.24	4.96	-0.25	6.80
Pseudo-interaction: (β_{UK},X_{UK})-(β_{C},X_{UK})+(β_{C},X_{C})-(β_{UK},X_{C})	9.44	17.45	3.53	5.95	-3.10	0.69
Detailed decomposition (characteristics)	% points	% points	% points	% points	% points	% points
Timing of calls * # of calls	-6.42	-5.55	-6.19	-3.91	6.94	-1.04
Interviewer workload	-1.29	0.04	-0.02	-0.27	0.07	-0.10
Interviewer ability (cooperation rate)	-3.12	0.00	0.00	-1.61	0.00	0.25
Interviewer calling strategy (% cases called weekday evenings / weekend)	-0.65	-0.16	0.00	-0.25	-0.25	-0.76
Interviewer phone contacting	-0.65	-0.06	-0.01	-0.30	-0.01	0.00
Any intercom	0.22	0.14	-0.54	0.06	-4.70	-1.54
Physical state of building	-0.08	-0.03	-0.01	-0.04	-0.03	0.10
Comparative physical state of building	-0.63	0.00	0.00	-0.84	-0.01	-0.03
Urbanicity	-0.02	0.00	0.00	-2.04	0.00	1.02

Below this, the aggregate decomposition is shown. The measurement unit here is the percentage point difference in the contact rates explained by the composite parts. For example, the Greek aggregate decomposition (third column) shows that the majority of the 3.00 percentage points difference in contact rate between the UK (95.0%) and Greece (98.0%) is due to a difference in characteristics between countries. In fact, had Greece had UK characteristics, the Greek contact rate would have been $(\beta_C, X_{UK}) - (\beta_C, X_C)$, i.e. 6.77 percentage points lower (row 4 – row 2).

Finally, the detailed decomposition displays the percentage points of the characteristics effect explained by each characteristic. In the UK-Greece comparison, for example, the characteristics effect is largely due to the timing and number of calls, which are associated with a higher contact rate in Greece. According to the detailed decompositions Greece would have had a 6.19 percentage points lower contact rate, if the Greek interviewers had followed similar calling patterns as the UK interviewers. The descriptive statistics show that in Greece a larger proportion of the sample units was contacted on weekday evenings and on Sundays than in the UK. The evening and Sunday calls are possibly the reason for this characteristics effect.

One should however be cautious interpreting decompositions, where large pseudo-interaction effects are involved, since we do not know which characteristics contribute to the interaction. In the Greek-UK comparison, for example, it might well be that the interaction effect of 3.53 percentage points offsets some of the effect of timing and number of calls. The decompositions in Table 3 show especially large coefficients and interaction effects for the model comparing the UK with Finland. Looking at the descriptive statistics one finds that these are likely to be due to the large number of cases that were attempted by phone in Finland. Here 90.7% of cases were attempted by phone compared to 0.4 in the UK. Earlier we also mentioned that in Finland 16.5% of cases were never attempted in person and had to be excluded from the models in this paper. Apparently, fieldwork in Finland was conducted very differently from that in the UK. Possibly the UK-Finland decomposition is a comparison of apples and oranges.

For the UK-Belgium and UK-Portugal comparisons the aggregate decompositions also show large interaction effects (relative to the size of their characteristics effects). While the logit models of these countries can be interpreted, the detailed decompositions of the characteristics effect are less meaningful. For better decompositions of these comparisons, the models need to be fit such that like is compared with like. Additional information on sample units and the fieldwork process, if available, can improve these decomposition models if these additional variables aid explaining the cross-national difference in contact rates. For the UK-Spain comparison (and to a lesser extent also for the UK-Greece and UK-Portugal comparisons) interpreting the detailed decomposition of the characteristics effect is more meaningful due to a relatively smaller interaction effect; even though the difference in contact rate of 5.38 percent between the two countries is largely explained by a difference in processes (i.e. the coefficients effect) rather than a difference in characteristics.

FINDINGS

In the following we discuss the findings from the decompositions, the logit models and the descriptive statistics across all countries. However, to set the scene, we first describe the fieldwork and neighbourhood characteristics in the reference country and the contribution of each of these variables in the UK logit model.

The UK as reference country

The UK achieved an average in-person contact rate (95.0%). While Finland, Greece and Portugal achieved higher contact rates, Belgium's contact rate was similar to the UK's, and Ireland and Spain had lower contact rates.

The mean number of contact attempts in the UK (2.4) was amongst the highest, though with a very large standard deviation. Most cases were attempted at least once during the day on a weekday (88.8%), which was the highest proportion of cases across these seven countries. 23.0% of cases were attempted on a weekday evening and 9.0% on a Sunday, which was less than average. However, the UK interviewers attempted 25.2% of cases at least once on a Saturday; only Portuguese interviewers attempted more cases on Saturdays. The interviewer calling strategies (with respect to the percentage of calls an interviewer carried out on weekday days, weekday evenings, Saturdays and Sundays) showed very similar patterns across countries.

The mean interviewer cooperation rate in the UK (47.5%) was much lower than in the other countries. The interviewer workload, however, was average (29.1 cases per interviewer). The proportion of cases interviewers attempted by phone was very low in the UK (0.4), which is likely to be due to its sampling frame of addresses and interviewer instructions not to use the phone to attempt contact.

In the UK the percentage of addresses with an intercom (9.9%) was low compared to the other countries. Only in Ireland there were fewer sample units with an intercom. Differences in the presence of an intercom at the house were large across the countries analysed here. In addition, the urbanicity indicator shows that the number of sample units situated in areas of predominantly single-unit houses was similarly high in the UK, Belgium and Ireland. In the other countries (Finland, Greece, Portugal and Spain) a much larger proportion of sample units lived in multi-unit housing areas. This is correlated with the proportion of cases with an intercom (especially in the UK, Belgium, Greece, Portugal and Spain). Apparently, the ESS samples of the UK and Belgium were both less urban and less likely to have intercoms than the samples of Spain, Greece and Portugal.

The state of the buildings in the UK sample was average. However, the state of the sampled address when compared to the neighbourhood showed less variation in the UK than in the other ESS countries. Apparently, countries differed in how homogenous the neighbourhoods were. In addition,

interviewers might record the state of a house or neighbourhood differently across countries; because these items are inherently subjective and normative, and norms might well differ systematically between countries.

Finally, the logit model of contact for the UK shows that especially the number and timing of calls made to a sample unit and their interaction were highly significant. There was a negative association between contact propensity and both the number of contact attempts and the attempting of contact at a certain times. However, interviewers that attempted contact in the evening for most of their sample units were more likely to make contact. Furthermore, the logit model estimated that the fact that a building was in a bad state was associated with a lower contact propensity. The proportion of cases an interviewer called upon during weekday evenings was positively associated with the probability of making contact in the UK.

Correlates of contact

Next we examine which factors were consistent predictors of non-contact across countries and which were associated with non-contact in some countries but not in others. For this we jointly analyse the descriptive statistics, the logit models and the decompositions.

Number and timing of calls

To assess the impact of number and timing of contact attempts four variables were included in the models: the natural logarithm of the number of contact attempts and whether there were any contact attempts in the evening, on Saturday and on Sunday. Furthermore, we assumed that the association between successful contact and ever calling at the weekend / in the evening differs when one contact attempt was made from when ten contact attempts were made. To account for this we included interaction terms.

The logit models and the decomposition confirm our suspicion that the timing and number of contact attempts play an important role when predicting contact. Furthermore, the detailed decompositions with Greece, Ireland and Portugal show some of the largest composite effects in these variables. However, interpreting interactions in non-linear models is not straightforward. As Norton et al. (2004) demonstrate, for non-linear models the output of standard statistical software does not display the correct coefficients and standard errors of interaction effects (and thereby also of the main effects). Therefore, the interpretation of the effect of the number and timing of contact attempts necessitates further analyses.

We carried out additional analyses of the interaction effect by means of Norton et al.'s Stata programme *inteff* for calculating the marginal effects of the interactions. To reduce the complexity of our analyses and interpretations we examined each interaction effect in turn, i.e. separately for weekday evening, Saturday and Sunday calls. In non-linear models marginal effects depend on the values of all other variables in the model. Table A2 in the appendix displays the mean interaction effects, i.e. the mean change in the marginal effect of making an evening, Saturday or Sunday call associated with making an additional call; together with their standard errors and z-statistics[3]. All other variables in the models were kept at their actual value. For each interaction effect in each country, we further examined plots of the interaction effect with the predicted probability of contact and of the associated z-statistic with the predicted probability of contact (not displayed). In the UK and in Spain we found significant interaction effects. In the UK, this was the case for evening calls and Saturday calls (both interacted with the natural logarithm of the total number of calls). In Spain the interactions with evening calls, Saturday calls and Sunday calls were largely significant. In the other countries there was no or little significance in the interaction effects. Here only the main effects of timing and number of contact attempts were interpreted.

To understand the interaction effects in the UK and Spain Figure A1 in the Appendix plots the predicted probability of contact against the number and timing of calls. The plots show that in both countries there was a negative association of the number of contact attempts and contact propensity. This means that the more contact attempts were undertaken, the lower was the probability of making contact with the sample unit. In addition, if a sample unit was ever attempted in the evening, on Saturday or on Sunday, the probability of contact decreased more steeply with each additional contact attempt. Finally, for weekday evening and Saturday calls this interaction effect was more pronounced in the UK than in Spain.

For those countries not affected by the interaction effects, we estimated predicted probabilities of contact associated with different timings and numbers of calls (Table A3 in the Appendix). All other variables were kept at their actual value. These predicted probabilities, as well as the coefficients of the logit models (Table 3) show that the number of contact attempts has a highly significant negative association with contact propensity across all countries analysed. In Greece, for

[3] The interaction effect is the discrete difference (in terms of timing of call) of the single derivative (in terms of log(total number of calls)). Norton et al.'s interaction effects are therefore similar to calculating the marginal effect of ever making an evening/Saturday/Sunday call associated with a unit change in the number of calls (i.e. with an additional call).
The standard errors and z-statistics take into account that the marginal effect varies with the sample's combinations of values of other variables in the model.

example, the predicted probability of contact for one call was 99.9%, while the predicted probability of contact for five or more calls was 86.7%. In the other countries these differences are even more pronounced. The most extreme difference was found in Ireland, where the predicted contact rate was 98.6% for cases with one call and 62.6% for cases with five or more calls. The main effects of ever calling a sample unit on a weekday evening, Saturday or Sunday were more heterogeneous across countries. In Ireland and Portugal, ever attempting contact on a Sunday was negatively associated with probability of contact. In Portugal attempting contact on weekday evenings had a negative effect. In the other countries, no significant effect of the timing of contact attempts was found.

In summary, interviewers face a decrease in the marginal effect of making additional contact attempts. This finding is consistent (though of different magnitude) across all countries. The effect of ever calling in the evening, on Saturdays and on Sundays varies across countries and no consistent pattern was found. In countries where there was such an effect it was negative. In Spain and the UK we found interaction effects; the change in the marginal effect of calling in the evening, on Saturday and (for Spain) on Sunday associated with making an additional call was negative. This means that in Spain and the UK the probability of making contact was lower the more contact attempts were carried out; and for sample units that had ever been attempted in the evening or at the weekend this decrease in contact probability with each additional call was steeper than for sample units never called at these times. Across countries, the decompositions showed that the timing and number of contact attempts were strongly associated with differences in contact rates.

These findings appear to contrast with the literature, which generally reports a positive effect of attempting contact in the evening, on Saturdays and on Sundays (see Groves & Couper, 1998; Purdon et al., 1999; Stoop, 2005; Lipps & Benson, 2005). However, one should be careful with a causal interpretation. While other authors looked at the probability of contact at each call conditional on the outcome of the previous call, we examined the marginal effects of the total number of calls (however cf. Olson, 2006). Since contact attempts in the ESS were not randomly assigned, interviewers chose to attempt contact at times and days that they felt might be most productive and that suited them. Therefore, it is likely that only those sample units were contacted repeatedly and in the evening / at the weekend that by their very nature were more difficult to contact. The frequency distributions show that, across all countries, most of the sample units were attempted at least once during daytime on a weekday. But much fewer cases were contacted on weekday evenings, Saturdays and Sundays. Apparently, only sample units that are difficult to

contact were attempted during the evening and weekend hours. In Spain and the UK, apparently, difficult-to-contact sample units were more likely to be attempted repeatedly and were more likely to ever be attempted in the evening and at weekends.

Workload

Interviewer workload was negatively associated with contact propensity in Belgium, however, not in any of the other countries. The descriptive statistics show that the Belgian sample units were contacted by interviewers with comparatively low workloads. That Botman and Thornberry's (1992) expectation about the negative effect of high interviewer workloads is only confirmed in Belgium, might be due to many countries allocating more sample units to their best interviewers thereby counteracting this workload effect.

Interviewer ability (cooperation rate)

The overall ability of an interviewer (proxied by the interviewer cooperation rate) was a significant predictor in four of the seven logit models. In Belgium, Finland and Greece a higher cooperation rate was associated with a higher contact rate. O'Muircheartaigh and Campanelli's (1999) finding that interviewers who are good at gaining cooperation are also good at making contact therefore found support in these countries. However, in Portugal higher interviewer cooperation rates were significantly associated with lower contact rates. The UK-Greece and UK-Portugal decompositions (these are the decompositions with small pseudo-interaction effects), however, show little contribution of the distribution of this characteristic on the difference in contact rates across countries. Apparently, interviewer abilities are not easily compared across countries and cross-national differences in interviewer cooperation rates do not explain differences mean contact propensities.

Calling strategies (timing of calls)

Having found that the timing of calls was associated with a lower contact propensity in some countries, we examine the effect of interviewer strategies with respect to timing of calls. It is reasonable to assume that each interviewer's assignment contained a mix of sample units with respect to their at-home patterns and contactability. Therefore, examining timing of call at this aggregate level accounts for the variation in contactability across sample units. The fact that the model controls for urbanicity, furthermore, accounts for differences in contactability between urban and rural areas. The logit models show that the proportion of cases an interviewer attempted in the evening was positively associated with contact propensity in the UK and Ireland and negatively

associated with contact propensity in Portugal. Additionally, in Ireland the proportion of cases an interviewer called on a Saturday was positively associated with contact propensity. In the decompositions, however, interviewer calling strategies play a less important role for these countries. Only the UK-Spain comparison shows that Spain would have had a 0.76 percentage points lower contact rate, had Spanish interviewers had the UK interviewers' calling strategies. The descriptive statistics show that interviewers in the UK made a proportionally fewer calls on weekday evenings and Sundays and more during weekday daytime than interviewers in Spain. However, again a causal interpretation would be misleading here. Figure A1 shows that the difference in marginal effect of additional contact attempts given between no evening call and at least one evening (no Saturday call and at least one Saturday call) is much more pronounced in the UK than in Spain. This indicates that UK interviewers were more likely to only call difficult cases in the evening (on Saturday), while Spanish interviewers also called easier cases at these times. Consequently, interviewers' calling strategies play a role in achieving contact. However, cross-country comparisons are hindered by cross-national differences in the use of evening and weekend calls for difficult-to-contact sample units.

Contact by phone

Only in Finland was the proportion of cases an interviewer attempted by phone associated with contact propensity. As discussed previously, the Finnish fieldwork was carried out quite differently from fieldwork in the other countries examined here. Whether a sample unit was contacted by phone did not explain differences in contact rates across countries.

Intercom

The type of housing a sample unit lived in was highly correlated with whether the house was equipped with an intercom. In fact, while type of housing was not a significant predictor of contact in the logit models the intercom was significant in Portugal, once type of housing was excluded from the model. Here the presence of an intercom was negatively associated with contact, which emphasises its function as a barrier to contact. Across countries, the UK-Spain decomposition showed that had Spain had the intercom distribution of the UK, its contact rate would have been 1.54 percentage points lower. The descriptive statistics showed that the Spanish sample had a very large proportion of housing units with intercoms (68.1% compared to 9.9% in the UK). However, in contrast with the literature (Groves & Couper, 1998; Lipps & Benson, 2005) this high proportion of intercoms in Spain was beneficial for its contact rate. Possibly, the presence of intercoms measures impediments to access differently across countries. Whereas in Spain having no intercom may

indicate a tendency to have no direct access at all to the building (i.e. a locked gate with no intercom is a bigger barrier than a locked gate with an intercom) in UK no intercom may indicate direct access to the housing unit.

State of building

A bad state of the sample unit's building had a negative impact of contact propensity in the UK and Finland. Furthermore, if the building was comparatively better than the rest of the buildings in the area this was positively associated with contact propensity in Belgium and Ireland. In the decompositions, however, the state of the building and the comparative state of the building contributed little to explain cross-national differences. (There was a, relatively small, effect in the UK-Ireland detailed decomposition. However, the pseudo-interaction effect in the aggregate decomposition might well cancel out this effect.)

Urbanicity

The effect of urbanicity on the probability of contact varies across countries. While in the UK, Belgium, Greece and Spain no within-country effect is found, more rural areas are associated with higher a higher probability of contact in Finland and Ireland, but with a lower probability of contact in Portugal. In the UK-Ireland decomposition urbanicity has a relatively large effect; however, again this effect might be cancelled out by the pseudo-interaction effect. Consequently, the findings are mixed.

One should remember that urbanicity was measured by looking at the proportion of multi-unit houses in the interviewer's initial assignment. Countries differ with respect to how common multi-unit structures are in large cities and villages. Usual indicators of urbanicity also implicitly capture the busy life-style associated with living in a big city. Due to differences in building styles across Europe, the urbanicity indicator included in these models, might not capture this 'busyness' in the cross-national decompositions.

Additional interviewer characteristics

In addition to the general models based on information from the ESS contact data, we were able to include information from an interviewer survey in a second set of analyses. The interviewer survey was not carried out in all ESS countries. Of the countries included in this paper data for the UK, Belgium, Finland and Portugal were available. Furthermore, not all interviewers in these countries filled in the interviewer questionnaire. Therefore, for some sample units no additional interviewer

data existed. As a result, when estimating the logit models with additional interviewer variables the base was smaller in these reduced models. Table A4 in the Appendix displays the base of the reduced sample (i.e. the number of sample units that were matched with interviewer information), as well as the number of sample units that could not be matched with the interviewer data. While in the UK about 86% of cases and in Finland and Portugal about 94% of cases had matched interviewer data, in Belgium less than half of the cases were matched with additional interviewer information. Table A4 further shows the contact rates of these new samples. In Finland and Portugal the in-person contact rate remained approximately the same in the reduced sample. In the UK and Belgium, however, the interviewers that provided information in the interviewer questionnaire achieved a higher contact rate than the full sample (0.5 percentage points higher in the UK and 2.0 percentage points higher in Belgium).

How did the additional variables impact on the logit models and decompositions? To compare the general logit models with the logit models with additional interviewer variables we re-estimated the general model on the reduced data set. The general logit models of the reduced data set and the models with the additional interviewer variables are displayed in Table A5. The additional interviewer variables changed the strength of the effects of the other variables in the logit models only slightly and the direction of the effects remained the same.

The effects of the additional interviewer variables varied across countries. Being contacted by a female interviewer or an interviewer with upper secondary or tertiary education was associated with a higher contact propensity in Belgium, but not in the other countries. The interviewers' age was not associated with contact propensity when work experience was controlled for. Work experience is only significantly associated with contact propensity in Finland, where it has a negative association. Sample units contacted by interviewers that usually leave a card or message upon non-contact had higher contact propensities in the UK but lower contact propensities in Portugal. Finally, whether sample units were contacted by interviewers that ask neighbours for information was negatively associated with contact propensity in Belgium.

Table A6 compares the decompositions of the general models on the reduced data set with the models with additional interviewer variables for each UK-country comparison. Blom (2008) noted that additional covariates can reduce the coefficients and pseudo-interaction effects, if these covariates add additional explanatory power to the decompositions. In the decomposition in Table 6, this is the case for the UK-Finland decomposition. The decomposition of the general model with

the reduced datasets showed immense coefficients and interactions effects, which were considerably smaller for the extended models with additional interviewer information; albeit still substantial. To a lesser extent the additional interviewer variables reduced the pseudo-interaction effect in the UK-Portugal decomposition; though its coefficients effect was slightly increased. The additional interviewer variables added no explanatory power to the UK-Belgium decomposition. Apparently, the variables measured in the interviewer questionnaires did not explain differences in contact rates between Belgium and the UK. While within Belgium interviewer gender, age and strategy with regards to asking neighbours for information were associated with contact propensity, differences in contact propensity between Belgium and the UK were not explained by them.

DISCUSSION AND CONCLUSION

The paper set out to identify correlates of differences in contact rates across countries. We aimed to distinguish differences in the distributions of *characteristics* associated with contact propensity from differences in the *effect* of these characteristics (i.e. the coefficients or processes) on contact propensity. In addition, we tested the suitability of decomposition (or counterfactual) analyses for investigating cross-national differences in contact rates. The analyses yielded four main findings: (1) the main characteristics associated with cross-national differences in contact rates are the number and timing of contact attempts; (2) in most country comparisons differences in processes play a major role; (3) the meaning of some of the variables might well differ across countries; and (4) a cross-national comparison between the UK and Finland was hindered by fundamental differences fieldwork approaches.

The number of in-person contact attempts made to a sample unit and their timing were found correlated with contact propensity within countries and with differences in contact propensities across countries. Across all countries, sample units that received more calls were less likely to have been contacted by the end of fieldwork. (One should note however that unlike most previous studies, we only modelled unconditional probabilities. No conclusions about the contact propensity at each additional contact attempt can be drawn.) Number and timing of contact attempts were the only variables consistently correlated with both within-country contact propensities and cross-national differences in contact rates.

In all decompositions we found sizable coefficients or pseudo-interaction effects. This can mean three things: (1) processes differ across countries (i.e. the characteristics have different effects), (2)

the indicators included in the models measure different things across countries or (3) the models are missing key variables for explaining cross-country differences in contact rates. Though differences in processes are to be expected, the analyses have also found some evidence for differences in measurement and missing variables. Models with additional interviewer variables showed that pseudo-interaction and coefficients effects were somewhat reduced when additional interviewer characteristics were controlled for. Furthermore, the decompositions detected variables with different meanings across countries. Notably the urbanicity indicator (i.e. proportion of single-unit housing and farms in an interviewer assignment) might have a different meaning across countries. If urbanicity is to distinguish the tranquillity of country life from the bustle of the city, then the proportion of single-unit housing might not accurately reflect this in a cross-country comparison. Building styles differ across countries and so might the proportion of single-unit housing in busy areas differ across countries. However, other standard urbanicity indicators (e.g. city size) bear similar problems concerning their cross-country implications for busyness. To improve analyses of cross-national nonresponse we therefore need to continue searching for truly comparative indicators.

Finally, the analyses raised awareness to the problem of comparing apples and oranges, i.e. of comparing fieldwork strategies across countries. In the UK-Finland decomposition we found very large coefficients and pseudo-interaction effects indicating that the fieldwork processes differed strongly between the two countries. In Finland fieldwork in the ESS relied heavily on contact attempts by phone, while this is hardly practiced in the UK. Consequently, the Finnish and UK fieldwork strategies were too dissimilar to warrant a sensible comparison.

The novel application of the decomposition method has allowed us to reach some important conclusions about the correlates of cross-national differences in contact rates. The analyses showed that the composition of the contacted sample is primarily influenced by the number and timing of contact attempts. Consequently, the countries differ with respect to the contactability of the contacted sample, which is likely to be correlated with other sample unit characteristics such as employment and family situation. However, there is a caveat. Due to a lack of direct measures of sample-unit demographics we cannot ascertain whether differences in contacted sample composition is due to differences in the population (as regards the distribution of employment and family characteristics) or due differences in the way the ESS was carried out across countries. Future research will need to address this problem.

REFERENCES

Billiet, J., Koch, A., & Philippens, M. (2007). Understanding and improving response rates. In R. Jowell, C. Roberts, R. Fitzgerald, & G. Eva (Eds.), *Measuring attitudes cross-nationally: Lessons from the European Social Survey* (pp. 113-137). London: Sage Publications.

Blom, A. G. (2008). Explaining differences across countries in survey contact rates. In *Proceedings of the Survey Research Methods Section*. Alexandria, VA: American Statistical Association.

Blom, A. G., Jäckle, A., & Lynn, P. (2009). Understanding Cross-National Differences in Unit Non-Response: The Role of Contact Data. In J. Harkness, M. Braun, B. Edwards, T. Johnson, L. Lyberg, & P. Ph. Mohler (Eds.), *Survey methods in multinational, multiregional, and multicultural contexts*. New York: Wiley & Sons.

Botman, S. L. & Thornberry, O. T. (1992). Survey Design Features Correlates of Nonresponse. In *Proceedings of the Survey Research Methods Section* (pp. 309-314). Alexandria, VA: American Statistical Association.

Campanelli, P., Sturgis, P., & Purdon, S. (1997). *Can you hear me knocking? An investigation into the impact of interviewers on survey response rates*. London: National Centre for Social Research.

de Leeuw, E. & de Heer, W. (2002). Trends in household survey nonresponse: A longitudinal and international comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey Nonresponse* (pp. 41-54). New York: John Wiley & Sons.

Even, W. E. & Macpherson, D. A. (1993). The Decline of Private-Sector Unionism and the Gender Wage Gap. *The Journal of Human Resources*, 28(2), 279-296.

Goyder, J. (1985). Nonresponse on surveys: a Canada - United States comparison. *Canadian Journal of Sociology*, 10(3), 231-251.

Groves, R. M. & Couper, M. P. (1998). *Nonresponse in Household Interview Surveys*. New York: John Wiley & Sons, Inc.

Hox, J. J. (2002). *Multilevel Analysis: Techniques and Applications*. London: Lawrence Erlbaum Associates.

Lipps, O. & Benson, G. (2005). Cross-national contact strategies. In *Proceedings of the Survey Research Methods Section* (pp. 3905-3914). Alexandria, VA: American Statistical Association.

Lyberg, L. & Dean, P. (1992). *Methods for reducing nonresponse rates: A review*.

Lynn, P. & Clarke, P. (2002). Separating Refusal Bias and Non-Contact Bias: Evidence from UK National Surveys. *The Statistician*, 51(3), 319-333.

Nicoletti, C. & Buck, N. H. (2004). Explaining interviewee contact and co-operation in the British and German Household Panels. In M. Ehling & U. Rendtel (Eds.), *Harmonisation of panel surveys and data quality*. Wiesbaden: Federal Statistical Office.

Norton, E. C., Wang, H., & Ai, C. (2004). Computing interaction effects and standard errors in logit and probit models. *The Stata Journal*, 4(2), 154-167.

O'Muircheartaigh, C. & Campanelli, P. (1999). A Multilevel Exploration of the Role of Interviewers in Survey Non-Response. *Journal of the Royal Statistical Society.Series A (Statistics in Society)*, 162(3), 437-446.

Olson, K. (2006). Survey Participation, Nonresponse Bias, Measurement Error Bias, and Total Bias. *Public Opinion Quarterly*, 70(5), 737-758.

Philippens, M., Loosveldt, G., Stoop, I., & Billiet, J. (2003). Noncontact Rates and Interviewer Calling Strategies in the ESS. Paper presented at the International Workshop on Household Nonresponse in Leuven, Belgium.

Purdon, S., Campanelli, P., & Sturgis, P. (1999). Interviewers' calling strategies on face-to-face interview surveys. *Journal of Official Statistics*, 15(2), 199-216.

Stoop, I. (2005). *The hunt for the last respondent*. The Hague: SCP.

Yun, M.-S. (2004). Decomposing Differences in the First Moment. *Economics Letters*, 82(2), 275-280.

Yun, M.-S. (2005). A simple solution to the identification problem in detailed wage decompositions. *Economic Inquiry*, 43, 766-772.

APPENDIX

Table A1: Reasons for excluding countries from the analysis

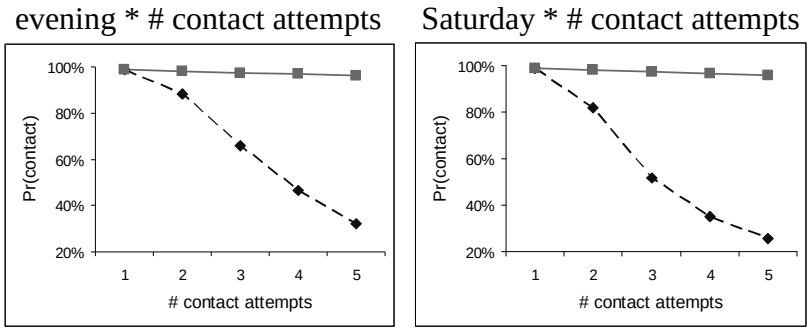
Country	Problems
Austria	No interviewer identification numbers in the contact data.
Czech Republic	Incomplete outcome codes; in 10% of the contact attempts the result code is missing; the physical state of the building was coded 'same' for all cases
Denmark	The first contact attempt was missing and neighbourhood data were very incomplete.
France	France did not deposit any contact data.
Germany	In 30% of cases all neighbourhood variables were missing.
Hungary	A large number of interviewer identification numbers was missing.
Israel	A large discrepancy between the number of cases in the main data file and the number of interview outcomes in contact data
Italy	A large number of interviewer identification numbers was missing for the last contact attempt.
Luxembourg	All security features were missing.
Netherlands	A large number of interviewer identification numbers was missing for the last contact attempt.
Norway	For data protection reasons Norway could not deposit any data on nonrespondents.
Poland	After deleting cases where all neighbourhood variables were missing and where there were no in-person contact attempts, Poland had a 99.93 contact rate and therefore too little variation in the dependent variable.
Sweden	For data protection reasons Sweden could not deposit any contact data.
Switzerland	Many contact attempts were carried out on the phone and registered in the contact data. While the Swiss data are of good quality, their fieldwork strategy was not comparable to that in other countries, because for some cases more than 100 contact attempts were made. The combined contact data file registers only 10 attempts.
Slovenia	661 sample persons without sample unit identification number; many item missings on other variables

Table A2: Mean interaction effects of timing and number of calls

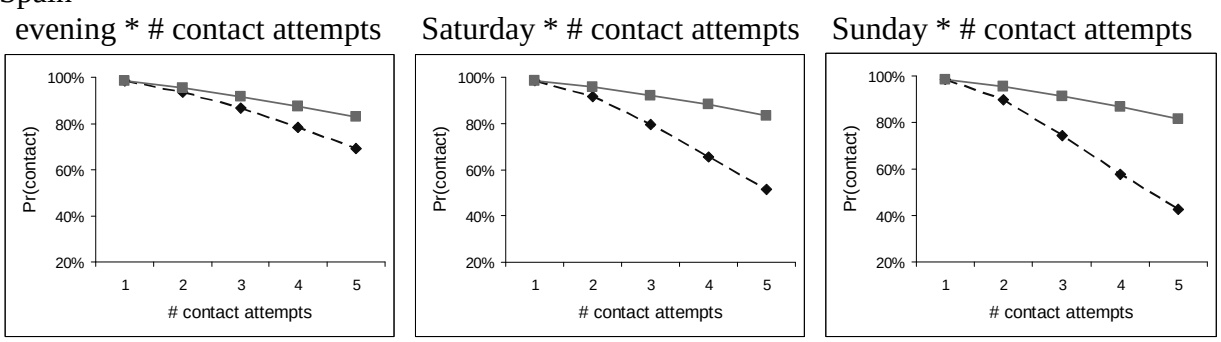
	Evening calls		Saturday calls		Sunday calls	
		Mean		Mean		Mean
UK	interaction effect	-0.075	interaction effect	-0.065	interaction effect	-0.035
	standard error	0.022	standard error	0.023	standard error	0.023
	z-statistic	0.647	z-statistic	1.452	z-statistic	0.380
Belgium	interaction effect	-0.031	interaction effect	-0.014	interaction effect	-0.015
	standard error	0.022	standard error	0.022	standard error	0.042
	z-statistic	-0.218	z-statistic	-0.595	z-statistic	0.024
Finland	interaction effect	-0.050	interaction effect	-0.077	interaction effect	0.018
	standard error	0.034	standard error	0.063	standard error	0.106
	z-statistic	-1.038	z-statistic	-1.321	z-statistic	0.438
Greece	interaction effect	-0.011	interaction effect	0.006	interaction effect	-0.010
	standard error	0.018	standard error	0.021	standard error	0.021
	z-statistic	-0.778	z-statistic	-0.863	z-statistic	-1.045
Ireland	interaction effect	-0.014	interaction effect	0.016	interaction effect	0.041
	standard error	0.026	standard error	0.025	standard error	0.037
	z-statistic	0.164	z-statistic	0.232	z-statistic	-0.302
Portugal	interaction effect	0.040	interaction effect	0.017	interaction effect	0.022
	standard error	0.036	standard error	0.032	standard error	0.038
	z-statistic	-0.736	z-statistic	-0.269	z-statistic	-1.000
Spain	interaction effect	-0.056	interaction effect	-0.094	interaction effect	-0.072
	standard error	0.025	standard error	0.035	standard error	0.039
	z-statistic	-2.133	z-statistic	-1.989	z-statistic	-0.619

Figure A1: Interactions of number and timing of contact attempts in the UK and Spain

UK



Spain



— no evening/Saturday/Sunday call --- at least one evening/Saturday/Sunday call

Table A3: Main effects of timing and number of calls on predicted contact rates across countries

	UK	Belgium	Finland	Greece	Ireland	Portugal
Δ evening call (keeping number of contact attempts at actual value)						
No sample unit called ever in the evening		95.9	96.4	98.4	93.8	96.1
Actual situation in contact data		95.3	95.7	98.0	93.4	96.4
All sample units ever called in the evening		95.3	95.8	97.8	93.2	94.6
Δ predicted contact rate none vs. all called in the evening		-0.6	-0.6	-0.5	-0.6	-1.5 ***
Δ Saturday call (keeping number of contact attempts at actual value)						
No sample unit ever called on Saturday		95.5	96.0	98.1	93.1	96.2
Actual situation in contact data		95.3	95.7	98.0	93.4	96.4
All sample units ever called on Saturday		95.0	91.2	97.7	93.6	96.3
Δ predicted contact rate none vs. all called on Saturday		-0.5	-4.8	-0.5	0.5	0.2
Δ Sunday call (keeping number of contact attempts at actual value)						
No sample unit ever called on Sunday	95.5	95.3	95.7	98.3	93.3	96.6
Actual situation in contact data	95.0	95.3	95.7	98.0	93.4	96.4
All sample units ever called on Sunday	93.6	95.1	96.9	97.4	91.3	94.5
Δ predicted contact rate none vs. all called on Sunday	-1.9	-0.3	1.1	-0.9	-2.0 *	-2.1 ***
Δ # contact attempts (keeping evening, Saturday and Sunday calls at actual value)						
Actual situation in contact data		95.3	95.7	98.0	93.4	96.4
1 contact attempt		98.2	97.2	99.9	98.6	98.4
2 contact attempts		93.9	90.9	99.3	93.6	98.0
3 contact attempts		87.9	82.6	97.6	84.8	96.9
4 contact attempts		81.1	74.5	93.9	73.9	94.0
5 or more contact attempts		74.3	67.5	86.7	62.6	82.6
Δ predicted contact rate 1 vs. 5 or more contact attempts		-23.9 ***	-29.7 ***	-13.2 ***	-36.0 ***	-15.8 ***

Legend: * p<0.05; ** p<0.01; *** p<0.001

Table A4: Descriptive statistics - Contact data with interviewer characteristics

Interviewer variables	**UK**	**Belgium**	**Finland**	**Portugal**	**Total**
Contact rate (%)	95.1	97.3	95.7	96.3	
Sex of interviewer (%)*					
Male	50.3	39.0	97.2	38.2	57.3
Female	49.7	61.0	2.8	61.8	42.7
Missing (n)	*1*	*20*	*9*	-	*30*
Interviewers' level of education (%)*					
Primary	3.8	1.0	11.4	2.7	5.0
Secondary, vocational	20.9	33.1	41.8	29.3	30.0
Secondary, higher	30.3	44.4	40.0	47.0	38.8
Tertiary	45.0	21.6	6.9	20.9	26.2
Mean interviewer age*	55.8	49.7	51.2	38.4	49.6
Standard deviation	9.0	12.1	6.4	10.7	11.5
Missing (n)	*20*	*31*	*24*	-	*75*
Interviewer leaves message / card upon non-contact (%)*					
Always	9.9	8.6	1.8	24.6	11.1
Often	26.3	29.8	10.9	10.1	19.3
Sometimes	44.5	21.0	22.4	24.0	30.4
Rarely	14.0	15.8	40.8	8.1	19.5
Never	5.4	24.8	24.1	33.2	19.7
Missing (n)	*8*	-	*16*	-	*24*
Interviewer asks neighbours upon non-contact (%)*					
Always	3.7	18.1	10.0	9.9	9.1
Often	12.9	25.9	48.4	10.8	23.4
Sometimes	60.0	30.1	35.3	32.8	42.6
Rarely	19.0	20.2	6.2	25.9	17.6
Never	4.3	5.8	0.0	20.6	7.3
*Base (n)**	*3,082*	*1,426*	*2,153*	*1,996*	*8,657*
Cases without matching interviewer data (n)	*510*	*1,558*	*123*	*136*	*2,327*

Note: * weighted for all sample units

Table A5: Logit models of contact – general model with reduced data set and model with interviewer characteristics

	UK	**UK**	**Belgium**	**Belgium**	**Finland**	**Finland**	**Portugal**	**Portugal**
	b	b	b	b	b	b	b	b
Ever called in the evening	0.514	0.426	-0.182	-0.467	0.314	0.329	-4.834 **	-5.372 ***
Ever called on a Saturday	5.987 ***	6.198 ***	0.666	0.714	-1.343 *	-1.313 *	-1.182	-1.228
Ever called on a Sunday	-0.656	-0.690	-2.192	-3.493	0.750	0.973	-5.542 **	-5.173 **
Log (# in-person contact attempts)	-0.757 **	-0.731 **	-2.210 **	-2.780 **	-1.390 **	-1.462 **	-7.560 ***	-8.126 ***
Log(attempts)*evening	-0.691	-0.646	0.054	0.046	-1.865 **	-1.883 **	2.973 **	3.503 ***
Log(attempts)*Saturday	-3.340 ***	-3.517 ***	-1.306	-1.220	0.338	0.329	0.585	0.561
Log(attempts)*Sunday	-0.001	-0.036	0.875	1.455	-1.944	-1.970	3.285 **	3.087 **
Interviewer workload	0.001	0.003	-0.065 *	-0.080 *	-0.004	0.008	0.020 **	0.014
Interviewer ability (cooperation rate)	0.007	0.006	-0.038 *	-0.049 *	0.049 ***	0.043 ***	-0.044 ***	-0.030 *
Interviewer calling strategy:								
Evening calling	0.032 **	0.037 **	-0.044 **	-0.045 *	0.013	0.010	-0.076 ***	-0.096 ***
Saturday calling	-0.012	-0.013	-0.042 **	-0.011	-0.018	-0.028	-0.027	-0.042 *
Sunday calling	0.017	0.018	0.099	0.137	0.054	0.072	-0.035	-0.056 *
Interviewer phone contacting	-0.070	-0.033	-0.020	-0.023	-0.049 *	-0.053 *	0.005	0.046
Any intercom	-0.566	-0.544	-0.814	-0.770	-0.005	-0.008	-1.008 **	-1.162 **
Physical state of building:								
Satisfactory	-0.080	-0.126	-0.767	-0.833	-0.132	-0.131	-0.254	-0.233
Bad	-1.443 **	-1.402 **	0.377	0.472	-1.529 *	-1.414 *	0.457	0.628
Comparative state of building:								
Better	0.134	0.114	1.227	0.857	-0.056	-0.047	-0.004	-0.083
Worse	0.496	0.498	-1.172	-1.615 *	0.210	0.142	-1.014	-1.184
Urbanicity (farms, single-unit housing)	0.000	0.003	-0.015	-0.009	0.016 **	0.013 *	-0.021 **	-0.029 **
Interviewer sex (female)		-0.021		1.410 *		1.321		-0.747
Interviewer age		0.006		0.065		0.047		-0.031
Interviewer education (higher secondary/tertiary)		0.097		1.258 *		-0.047		-0.560
Interviewer work experience (in years)		-0.008		0.099		-0.043 *		-0.045
Interviewer leaves card (inverse coding)		-0.333 **		-0.440		-0.185		0.520 ***
Interviewer asks neighbours (inverse coding)		-0.111		0.718 *		0.131		0.144
Constant	3.753 ***	4.516 ***	13.674 ***	8.964 *	3.766	3.079	21.296 ***	23.065 ***
Chi^2	394.0	405.9	134.1	174.2	164.9	173.9	268.7	293.3
Pseudo R^2	0.353	0.363	0.405	0.526	0.234	0.247	0.431	0.47
N	3039	3039	1337	1337	1941	1941	1968	1968

Legend: *p<0.005; **p<0.01; ***p<0.001

Table A6: Results from the aggregate and detailed decompositions $\bar{y}_{UK} - \bar{y}_{C}$ with interviewer characteristics

	Belgium		**Finland**		**Portugal**	
	General model (reduced dataset)	Interviewer model	General model (reduced dataset)	Interviewer model	General model (reduced dataset)	Interviewer model
Predicted probabilities	%	%	%	%	%	%
(1) (β_{UK} ,X_{UK})	95.1	95.5	95.1	95.5	95.1	95.5
(2) (β_{C}, X_{C})	97.3	97.3	95.7	95.6	96.3	96.3
(3) (β_{UK} ,X_{C})	96.6	97.9	39.4	79.4	95.8	95.5
(4) (β_{C}, X_{UK})	95.3	97.0	90.8	96.5	98.1	97.0
(5) (β_{UK}, X_{UK}) – (β_{C}, X_{C})	-2.21	-1.77	-0.60	-0.11	-1.17	-0.80
Aggregate decomposition	% points	% points	% points	% points	% points	% points
Characteristics: (β_{C},X_{UK})-(β_{C},X_{C})	-1.99	-0.30	-4.97	0.92	1.82	0.68
Coefficients: (β_{UK},X_{C})-(β_{C},X_{C})	-0.78	0.62	-56.29	-16.15	-0.51	-0.79
Pseudo-interaction: (β_{UK},X_{UK})-(β_{C},X_{UK})+(β_{C},X_{C})-(β_{UK},X_{C})	0.56	-2.09	60.65	15.13	-2.48	-0.69
Detailed decomposition (characteristics)	% points	% points	% points	% points	% points	% points
Timing of calls * # of calls	3.25	0.31	7.27	-0.41	0.80	0.17
Interviewer workload	1.15	0.12	0.08	0.01	-0.16	-0.02
Interviewer ability (cooperation rate)	-1.94	-0.19	5.58	-0.28	0.75	-0.04
Interviewer calling strategy (% cases called weekday evenings / weekend)	-3.91	-0.28	0.04	-0.01	0.51	0.07
Interviewer phone contacting	-0.71	-0.06	-16.75	1.19	0.01	-0.04
Any intercom	-0.30	-0.03	0.00	0.00	0.19	0.01
Physical state of building	0.22	0.02	0.09	0.00	0.00	0.00
Comparative physical state of building	0.32	0.02	-0.01	0.00	0.01	0.00
Urbanicity	-0.08	0.00	-1.27	0.07	-0.30	0.00
Interviewer sex		0.04		0.16		0.57
Interviewer age		-0.11		0.06		0.00
Interviewer education level		-0.04		0.00		0.00
Interviewer work experience		0.01		0.07		-0.04
Interviewer leaves card		-0.04		0.04		-0.02
Interviewer asks neighbours		-0.06		0.02		0.00