

Cross National Contact Strategies

Oliver Lipps¹ and Grant Benson²

Swiss Household Panel, Neuchâtel, oliver.lipps@unine.ch¹

University of Michigan Institute for Social Research, gdbenson@umich.edu²

Abstract

Using paradata gathered from the 11-nation Survey of Health, Ageing and Retirement in Europe (SHARE), this paper examines the impact of the first contact attempt and the first contact properties, respectively, on contact and response efficiency using logistic multilevel models. We find that despite the different sample frames and interviewer compensation structure between countries, there are no considerable country effects with respect to making contact, once interviewer effects are controlled. Moreover, results point to an increased efficiency associated with evenings especially on Sundays, at least on the very first contact attempt. For attempts that result in initial contact, Saturday afternoons are most likely to eventually lead to completed interviews, followed by initial contact on weekdays during the daytime. We hypothesize that this may be due to the SHARE sample being composed of people aged 50 and over.

Keywords: Contact strategies, cooperation rates, cross-national survey, interviewer effects.

1. Introduction

With the ever-increasing cost of conducting surveys in the face of declining contact and response rates, many survey agencies resort to implementing contact strategies. These include directing interviewers on when and how to attempt contact and on how to respond to statements and questions made by the respondent. While we would expect differences in optimal contact strategies for different segments of the population – for example, for the retired compared to the fully employed – it is not clear that we would expect cross-national differences. That is, can we take lessons on contact strategies from the U.S. and U.K. and apply them with similar effects to continental European countries, for example? Or are there significant cultural and demographic differences that will impact the relative efficiency of contact strategies?

This paper analyzes data from the Survey of Health, Ageing, and Retirement in Europe (SHARE), conducted in 11 countries during the second half of 2004. Respondents were selected from among the non-institutionalized, resident persons aged 50 and

over, as well as their spouses. Data was collected at every interviewer attempt, including time and day of the attempt, mode of the attempt, and outcome of the attempt (contact, no contact, resistance, completed interview, etc.). Nine of the countries used a sample management system designed specifically for the project, while two countries (Switzerland and Belgium) used their own sample management systems to track effort and progress.

During centralized trainings for all participating survey agencies (Alcser and Benson 2005), the importance of working in optimal “call windows” was stressed. This was based largely on existing research from the American and British experiences. Several survey agency representatives remarked that the European context was simply different from the American experience. In particular, you could not call on a household after 6 PM without antagonizing household members, and never on Sundays. Taking these concerns seriously, this paper is first concerned with unraveling the effects of *when* attempts were made on successfully obtaining contact with a target respondent. That is, are evening and weekend attempts more likely to yield contact than weekday attempts? In a second step, we examine the impact of time, day, and mode of first contact on ultimately interviewing the respondent. For example, if the survey agency anecdotal evidence is correct, we would expect an annoyance factor to result from Sunday attempts, particularly if it resulted in the very first interaction with a household member. Subsequently, we would expect initial contacts on Sundays to result in a net decrease in final interviews with those sample lines relative to cases where the initial contact was obtained on a different day.

2. Background

Much of the literature for both in-person surveys (e.g., Groves and Couper 1998) and telephone surveys (e.g., Greenberg and Stokes 1990) points to net efficiencies associated with obtaining initial contact by optimizing interviewer contact effort. The data generally show that attempts made on working day evenings and weekends are most likely to yield successful contact with a household member. Presumably, this is the case because traditional work patterns keep respondents out of the home weekday daytime. However, it is not clear

that these patterns would continue to hold either for a sample drawn from an aging population¹ or for countries other than the United States and Britain. That is, with an aging population, we might expect that we would be more likely to reach a respondent during the day. Moreover, if the contact attempt is in-person (face-to-face or F2F), then evening attempts may be *less likely* to obtain contact, potentially because of a perception of respondent fear. Additionally, in the United States and Britain, there is a culture for the general population of working in blocs of eight (or more) continuous hours (typically 9 AM to 5 PM). This is not necessarily the case for other countries, where the work day may be interrupted by a two-hour lunch break allowing employees return to their place of residence. In this case, making contact during the day may increase the chance of reaching a household member at home.

In addition to country effects and the impact of when attempts are made, we are particularly interested in potential interviewer effects on establishing contact. That is, do some interviewers find better ways to obtain contact with respondents, even when controlling for time of day and day of the week that the attempt is made? If this is the case, then it suggests that we must do more thinking about recruiting and training interviewers.

3. Data

The analyses draw upon the early release of data (“Release 1”) for the Survey of Health, Ageing and Retirement in Europe (SHARE). SHARE was conducted in 9 European countries (Austria, Denmark, Germany, Greece, Italy, Netherlands, Spain, Sweden, Switzerland) between April and October 2004, with another 2 (Belgium and France) starting data collection late (de Luca and Lipps 2005). In general, survey agencies were required to collect interviews from around 1,500 households at an unweighted household response rate of greater than 50 percent. The cross-national aspect of the study presented numerous challenges. The most significant of these were to establish baseline contact rules and flexible sample designs that both accommodated standard practices within each country and retained probability selection

¹ See for example Kirgis et al. (2004). Moreover, the principle underlying a contact window effort is predicated on knowing nothing about the household composition. Thus, Groves and Couper (1998, pp. 89-94) show that there is a reduced effort associated with contacting household members if there are more adults in the household, if there are children in the household, or if there are adults over 70 in the household.

methods. Each survey agency was required to record every contact attempt and make minimally five attempts per household before coding out the case as non-finalized.²

Table1: Household Response Rates and Sampling Frames

Country	Sampling Frame	Household Response Rate
Austria	Telephone (TEL) Directory, Household (HH)	57.3%
Denmark	National (Nat'l) Registry, HH	61.1%
Germany	Regional Registry, Individual (Ind)	60.2%
Greece	TEL Directory, HH	60.2%
Italy	Regional Registry, Ind	54.1%
Netherlands	Regional Registry, Ind	61.6%
Spain	Nat'l Registry, Ind	50.2%
Sweden	Nat'l Registry, Ind	42.1%
Switzerland	TEL Directory, HH	37.6%
Average		53.8%

Source: Adapted from De Luca and Peracchi (2005).

Due to variability in available data, participating countries relied on several different sampling frames.³ For most countries (Denmark, Germany, Italy, the Netherlands, Spain, and Sweden), population registers were available. Sweden selected a reference person, Denmark a reference household randomly from national population registers. The other countries relied on multi-stage designs due to cost concerns and registers being administered at the local level. Three countries (Austria, Greece, and Switzerland) relied upon telephone directories to conduct screenings identifying households with at least one respondent aged 50 or over (Klevmarken 2005).

There were 71,114 in-person household contact attempts (“calls”) recorded for 14,040 first household

² For more details on the fieldwork procedures and SHARE’s sample management system, see De Luca and Lipps (2005).

³ Häder and Gabler (2003) cite Leslie Kish as saying of comparative research that „Sample designs may be chosen flexibly and there is no need for similarity of sample designs. Flexibility of choice is particularly advisable for multinational comparisons, because the sampling resources differ greatly between countries. All this flexibility assumes probability selection methods: known probabilities of selection for all population elements.”

interviews at an average of 5.07 attempts per household interview. An additional 72,945 attempts were recorded for 17,342 non interviews. Almost 29 percent of the households refused to participate, leaving 17.3 percent of households not participating for other reasons, including non-contact, language barrier, and other reasons.

Table 2: Individual Response Rates and Contact Attempts

Country	Individual Response Rate (within household)	Average In-Person Attempts per Household
Austria	87.4%	4.9
Denmark	93.0%	9.9
Germany	86.5%	5.4
Greece	91.8%	6.0
Italy	79.7%	4.7
Netherlands	87.9%	4.3
Spain	73.8%	4.8
Sweden	83.8%	9.6
Switzerland	86.9%	6.1
Average	85.6	6.46

Source: Adapted from De Luca and Peracchi (2005) and own calculations.

4. Issues

In general, we are concerned with accounting for the impact of country and interviewer effects on the likelihood of obtaining contact with a household member. We would expect country effects to impact likelihood of contact for two reasons. First, two countries (Greece and Switzerland) had unknown eligibility on the initial attempt. This would lead us to hypothesize that evening call windows would be more efficient in these countries than in the other countries. Second, Europeans sometimes refer to cultural differences between the “north” and the “south” – in this case, Greece, Italy, and Spain belonging to the south. The southern countries are more likely to have businesses that permit for extended mid-day breaks, while the northern countries are more likely to work continuous 8-hour workdays interrupted only by one-hour lunch breaks. Third, interviewers were compensated using different schemes between countries. Switzerland alone paid interviewers on an hourly basis, while most of the other countries compensated interviewers per completed interview or household listing, and Denmark employed a mix of hourly wages and piecemeal compensation. There were also differences in interviewer recruitment, with some countries using independent contractors as interviewers and other countries using a professional,

in-house staff. Each of these factors could lead to differences in interviewer behavior between countries, due to differential incentive schemes.

With respect to interviewers, some interviewers clearly develop ways of customizing their responses to potential respondents that enable them to achieve higher response rates. But there is no reason to expect interviewer effects in terms of obtaining initial contact, at least once one controls for time of day and day of the week that the attempt is made. If there is an interviewer effect, then it suggests either that some interviewers are doing something different – waiting for longer times at the door, ringing the bell more insistently, or trying multiple entrances – or that the characteristics of particular interviewers (being of a certain age or sex, or wearing clothing appropriate to a particular neighborhood) are more likely to have respondents open the door. If, contrary to our expectations, there are significant interviewer effects with respect to making contact, then we must do a better job of uncovering why that is the case.

4.1 Issue 1: Contact Efficiencies

A growing body of literature points to the importance of when contact attempts are made on obtaining initial contact (e.g., Groves and Couper 1998; Weeks, Kulka, and Pierson 1987; Purdon et al. 1999). In general, the research points to increased efficiencies associated with working weekday evenings and weekends relative to weekday daytime. Additionally, O’Muircheartaigh and Campanelli (1999) explore the impact of interviewers on non-response in the British Household Panel Survey, using a multilevel approach. By employing multilevel analysis, they disentangle the clustering effects of region from interviewer effects on non-response. They find that there was, after controlling for several household variables and indicators of cooperation and contactability, almost no region effects on refusal and non-contact rates, but still some (although not statistically significant), evidence for an impact of interviewer effects.

For most of the SHARE sample, households had known eligibility, with at least one household member over the age of 50. Subsequently, we would expect that this population might have an increased chance of being contacted during the daytime. However, it is unclear whether there might also be country effects, particularly due to differences in laws or common practice in terms of age of retirement, home maker rates, and other cultural norms. We would expect that interviewer effects would be reduced if we control for time and day of the week that calls were made.

4.2 Issue 2: Contact Effect on Respondent Cooperation

We continue the analysis of contact efficiencies by attempting to determine whether the mode (face-to-face versus telephone) and timing of the contact has any impact on cooperation rate. Thus, if Saturday afternoons are optimal times to reach a respondent at home, does reaching a respondent during the weekend have an impact on their ultimate likelihood to cooperate? The multilevel analysis accounts for the effects of the timing of the first successful contact on obtaining an interview with the household at any point in time.

5. Results

5.1 Contact Efficiencies

Using the the *MLwiN* software (see Rasbash et al. 2004) to conduct the multilevel logistic regression, we create a series of models iteratively building upon each other. We progress by conducting a step-wise (not necessarily hierarchical) inclusion of additional exploratory variables, in each case, increasing specificity depending on the results. The variables used in the analysis are summarized in Table 3.

For the initial model, all contact attempts are included. Random effects of the intercept for country, interviewer and household levels are introduced for all models. These are modeled on either the initial contact attempt or the initial attempt resulting in contact with a household member. Note that when considering only the initial attempt – or the initial contact – with a household, there would be no option of differentiating between respondents within a household.

The dependent variable in the first approach is “contact”, modeled binomially and using a logit link function. Attempts resulting in reaching a person in the household are coded “1”. These would include refusals, completed interviews, as well as appointments. All attempts resulting in no contact, such as no one being home or no one answering the door or the telephone, are coded “0”. The base model equation is given in Equation 1.

The random components $v_{j, \text{ctry}}$, $u_{j, \text{ctry}, \text{iwer}}$ are expected to follow a normal distribution $N(0, \Omega_{\text{var}})$, the random component $e_{j, \text{ctry}, \text{iwer}, \text{hh}}$ is expected to follow an extreme value distribution, $\text{var} \in \{v, u, e\}$. If several independent variables are modeled as random on the same aggregation level (i.e. Ω is a matrix), the non-diagonal elements of Ω are the covariances of the respective coefficients.

Table 3: Variable Descriptions

Dependent Variables

Contact

Binary variable that equals 1 if the contact attempt investigated resulted in some interaction with a household member.

CompletedHH

Binary variable that equals 1 if there was ever a completed interview with at least one household member, for the contact investigated.

Independent Variables

Work

Binary variable that equals 1 if the day of the week of the contact attempt was Monday through Friday.

Sat

Binary variable that equals 1 if the day of the week of the contact attempt was Saturday

Sun

Binary variable that equals 1 if the day of the week of the contact attempt was Sunday.

Morning

Binary variable that equals 1 if the time of day of the contact attempt was between 8 AM and 11:59 AM.

Afternoon

Binary variable that equals 1 if the time of day of the contact attempt was between 12 noon and 5:59 PM.

Evening

Binary variable that equals 1 if the time of day of the contact attempt was between 6 PM and 9:59 PM.

Night

Binary variable that equals 1 if the time of day of the contact attempt was between 10 PM and 7:59 AM.

F2F

Binary variable that equals 1 if the mode of the contact attempt was in-person (face-to-face) rather than by telephone.

Condition

Ordinal variable scoring interviewer evaluation of the physical appearance of the selected housing unit. If 0 then the physical appearance was evaluated as being in “Good” condition. If 1 then it was evaluated as being “Average” and if 2 then it was evaluated as “Poor”.

Environment

Ordinal variable scoring interviewer evaluation of the physical appearance of the neighborhood in which the selected housing unit is located. If 0 then the physical appearance was evaluated as being in “Good” condition. If 1 then it was evaluated as being “Average” and if 2 then it was evaluated as being “Poor”.

Impediment

Binary variable indicating whether the interviewer observed any impediments to access to the housing unit. If 1, then the interviewer observed barriers to access, such as a locked entrance to an apartment complex.

Equation 1: Base Model

$$\text{logit}(\text{call_success})_{\text{ctry, iwer, hh}} = \sum \text{coefficient}_i * (\text{fixed}) \text{variable}_i + \sum \text{coefficient}_{j, \text{ctry, iwer, hh}} * (\text{random}) \text{variable}_j$$

with

$$\text{coefficient}_{j, \text{ctry, iwer, hh}} = c_j + v_{j, \text{ctry}} + u_{j, \text{ctry, iwer}} + e_{j, \text{ctry, iwer, hh}}$$

We test for potential under- or overdispersion (Rabash et al. 2004) by relaxing the assumption of binomial variation, but found only slight deviations from a binomial variation. For parsimony, we opted to keep the lowest level variation binomial. This choice also facilitated the comparison of coefficients across models.

Table 4: Null model using all contact attempts (N=83,150 Release 1: 106,469) Model 0

Dependent variable: logit Contact	Fixed Effects	Random Effects (between ⁴)		
		Country	Interviewer	Household
Constant (Intercept) (std err.)	1.203 (0.150)	0.174 (0.090)	0.369 (0.027)	1.268 (0.025)

The first (null model) includes only the intercept, modeled both as a fixed and a random effect on each of the other hierarchical levels, being country, the interviewer, and the household level. Such a base model is primarily designed to calculate the amount of (disjoint) variation of the dependent variable on the different levels. Thus, we expect significant random variation on the dependent variable “contact” on all three (nested) levels.

Model 0 (Table 4) shows that the likelihood of obtaining contact is quite high. Overall, interviewers have a 77 percent likelihood of obtaining contact on any given attempt.⁵ There is small significant random variation on the country level. However, the variance on the interviewer and household levels are highly significant, well above the p<0.05 level. The interviewer level variation more than doubles that on the county level, while the former is in turn less than a third of that on the household level. Thus, most of the variation on contact is due to differences of the households.

⁴ Here and in the following tables, we omit the random effects within the lowest level (in this case: call level), which - due to binomial variation assumed - always equals 1.

⁵ Calculated as the antilogit of 1.203.

Due to the compounding effects of contact with multiple respondents, we drop analysis at the single respondent level and consider only the first contact attempt for the remaining analyses. Subsequently, the lowest level of analysis is the household⁶. We would expect that this would significantly decrease the likelihood of contact, as the interviewers would have no previous knowledge about the household.

Table 5: Contact Achieved -- Results of Multilevel Logit Models, First Contact Attempt (N=22,447), Models 1 and 2

Dependent variable: logit Contact	Model 1		Model 2	
	β	std. err.	β	std. err.
Fixed Effects				
const	.554	.129	.562	.129
Sat			-.144	.056
Sun			.054	.071
Random Effects: Country				
Constant	.123	.067	.122	.066
Random Effects: Interviewer				
Constant	.603	.044	.603	.044

Comparing the two-level variance components in Model 1 with the all-call model above (see Table 4), on the country level, we find only a slightly smaller variance (and significance). This suggests that the binary variable contact success varies between countries in the same way, irrespective whether all calls are considered or only first calls. As suspected, the chances of a successful contact drop considerably if only first contact attempts are considered.

The interviewer level variation is much lower when all calls are taken into account. This indicates that for first calls only, the interviewer’s behavior has a greater impact on the call result. In particular, we would suspect that an interviewer’s decision to make initial effort on particular days of the week or during particular times of day would have a significant impact on the likelihood to obtain contact.

In Model 2, we begin to attempt to decrease variations through a non-hierarchical stepwise inclusion of further explanatory variables. Here we included the day of the week, in a first step modeled as fixed effects. Model 2 generates a significantly negative Saturday effect on obtaining contact with a household member. There are (statistically insignificant) negative Sunday effects as well. Both of the weekend effects

⁶ The random effect on the household level is in the following standardized at 1.

are measured against traditional working days (Monday through Friday). The coefficients thus indicate a somewhat smaller contact success rate on weekends relative to weekdays. The country and interviewer specific variation basically remain unchanged by the inclusion of the day of the week.

In Model 3 (Table 6), we extend the model and take the *random* effects of Saturday and Sunday on the interviewer level into account. The hypothesis behind the model is that interviewers gaining contact on the first attempt during the weekend may be different from other interviewers. First, the fixed effects changed somewhat. More interestingly, the variation of Sunday and Saturday (vs. weekdays) successes on the interviewer level are highly significant (.558 and .618); with a variance much higher than the fixed coefficient (-.051 and .076). This indicates that around half of the interviewers being more efficient on the weekends than during the weekdays at an interviewer-specific level.

In addition, we find a weakly significant positive covariance between the interviewer’s successes on Saturday and Sunday on the one hand, and negative covariances of these weekend days with the interviewer’s overall performance (intercept) (-.119 and -.192). Thus, interviewers who are successful on Saturdays are more likely to be successful on Sundays as well, and vice-versa, and interviewers, whose first calls are more successful overall, tend to be less successful on weekends, and vice-versa, respectively. We subsequently added time of day to the analysis, and abandoned the randomness of the weekend days on the interviewer level. The literature generally indicates that weekday evenings are best for obtaining initial contact, followed by weekend afternoons. We follow Kulka and Weeks (1988) for definitions of time of day, extending the afternoon time slot from 5 PM to 6 PM, following Groves and Couper (1998). The results are controlled against a “night” time slot (10:00 PM – 7:59 AM, all days).⁷

⁷ Given that this characterizes first attempts, we consider this to be a curious choice for first attempts in any case. We believe that there are three acceptable interpretations for attempts registered during the night time slot. First, it can simply be very poor judgment on the part of the interviewer. He or she might be attempting a first look at the neighborhood, but this would not be particularly optimal due to the lack of daylight in most cases. Second, it is possible that the interviewer simply miscoded the time of day. We have no reason to believe that this miscoding would systematically have come from one of the other time slots. Third, there is the possibility that interviewers recorded the time that they made the note, rather than the time that they attempted the contact.

Table 6: Contact Achieved -- Results of Multilevel Logit Models, First Contact Attempt, Models 3 and 4

Dependent variable:logit Contact	Model 3		Model 4	
	β	std. err.	β	std. err.
Fixed Effects				
const	.560	.130	.833	.166
Sat	-.051	.076		
Sun	.108	.090		
Work Morning			-.411	.105
Work Afternoon			-.302	.101
Work Evening				
Sat Morning			-.116	.103
Sat Afternoon			-.609	.143
Sat Evening			-.395	.121
Sun Morning			-.236	.155
Sun Afternoon			-.615	.172
Sun Evening			-.233	.137
			.139	.167
Random Effects: Country level				
Constant	.125	.067	.136	.073
Random Effects: Interviewer level				
Constant	.617	.047	.597	.044
Sat	.558	.130		
Sun	.618	.173		
Covariances: Interviewer level				
Constant, Sat.	-.119	.068		
Constant, Sun.	-.192	.084		
Saturday, Sun.	.272	.138		

In Model 4, only Sunday evenings have a higher probability of obtaining contact on the first attempt than the control group (late nights), which is not surprising. Weekend mornings perform worst, followed by Saturday afternoons.

We are surprised by the lack of country effects. That is, there is no decrease in the random constant coefficient on the country level when we include the day and time of day variables. The given differences in interviewer compensation and sample frames suggest some more explanation of the country level variation. In particular, with three countries using the telephone for initial household listing, and other countries selectively allowing interviewers to attempt initial contact by telephone, we hypothesize that country level explanation could be increased by including call mode (face-to-face or telephone) as a dummy variable in Model 5 (Table 7).

However, there is little evidence of any interviewer systematically doing so.

Table 7: Contact Achieved -- Results of Multilevel Logit Models, First Contact Attempt, Models 5 and 6

Dependent variable: logit Contact	Model 5		Model 6	
	β	<i>std. err.</i>	β	<i>std. err.</i>
Fixed Effects				
const	.325	.287	.429	.294
Work Morning	-.277	.052	-.276	.052
Work Afternoon	-.180	.041	-.181	.041
Sat Morning	-.446	.118	-.444	.118
Sat Afternoon	-.248	.085	-.237	.085
Sun Morning	-.509	.156	-.510	.156
Sun Afternoon	-.172	.108	-.168	.108
Sun Evening	.254	.144	.260	.144
F2F	1.120	.410	1.131	.410
Impediment			-.196	.046
House average			-.112	.036
House poor			-.079	.077
Random Effects: Country level				
Constant	.597	.325	.625	.339
F2F dummy	1.255	.672	1.254	.672
Random Effects: Interviewer level				
Constant	.673	.068	.676	.068
F2F dummy	1.076	.167	1.075	.167
Covariances: Country level				
Constant, F2F	-.605	.404	-.624	.414
Covariances: Interviewer level				
Constant, F2F	-.500	.096	-.504	.096

With the introduction of a fixed and random effect mode dummy for the initial contact attempt being in-person, the daytime fixed effects are either reduced or made insignificant. Even the intercept becomes insignificant by controlling for mode. However, making the attempt in person (F2F) produces a positive impact on the likelihood of obtaining contact on the first call attempt.

The random effect of F2F on the country level is slightly significant (1.255), while the effect on the interviewer level is highly significant (1.076). Interestingly, the covariances with the intercept on both levels are negative even if statistically significant on the interviewer level. This points to an interesting counter-intuitive observation. In particular, it suggests that interviewers develop specialized skills such that if they have good skills contacting respondents by telephone, then those skills do not transfer to in-person attempts, and vice-versa.

We test this hypothesis by means of a simple logit model, with first contact attempts resulting in contact as the dependent variable, proportion of face-to-face

first calls and its square as independent variables (not shown). We find significant positive effects on first contacts from the proportion of in-person attempts, and significant negative effects from the square of the proportion of in-person attempts, and a nonsignificant intercept. This means that we have an inverse U-shaped relation between in-person proportion of calls and success of the first call, with the maximum contact rates for those interviewers, who realize a F2F proportion somewhere between 0 and 100%. This finding supports a hypothesis that well performing interviewers are able to apply the most appropriate mode of initial contact, depending on the special situation required.

We are curious whether interviewer evaluations of the “state” or condition of the selected household or neighborhood impacted their willingness to make contact on a household. In fact, we asked the interviewers to rate both the neighborhood and the housing unit in terms of being “good,” “average” or “bad”. The neighborhood evaluation does not show any significance. However, in Model 6 we see the highest impacts of evaluations of the housing unit being in an “average” state. Barriers to entry have an expected negative impact on likelihood to obtain contact with the household on the first attempt. The other coefficients only change to a very minor extend, once these housing observations are included in the model.

Finally, we are interested in whether we can determine anything about the relationship between the initial contact attempt and completing an interview with at least one household member. In short, are there interviewer actions that are particularly “off-putting” to respondents? Thus, we introduce the variable completedHH in Model 7 (Table 8), which equals 1 for those first calls on households that eventually deliver at least one individual interview. This variable is modeled with a fixed and a random effect on the interviewer level. We also keep the F2F variable, with a fixed effect, and a random effect on the interviewer level in the model.

Again, the time of day effects further decrease, as does the neighborhood effects. Not surprisingly, the fixed effect of the variable completedHH is positive and highly significant on all levels. Households that eventually have a completed interview are more likely to have had contact on the first attempt, as opposed to not having contact on that attempt. Reversing the logic of the equation, an initial contact attempt resulting in contact is more likely to result in a completed interview than a call that does not obtain contact on the first attempt. Moreover, there is a highly significant

random effect between interviewers (0.325), which is nevertheless smaller than the fixed effect (0.501). The fixed effect suggests that – everything else equal - even some of the relatively unsuccessful interviewers do better if the initial call yields contact (again, reversing the logic of the equation).

Table 8: Contact Achieved -- Results of Multilevel Logit Models, First Contact Attempt, Model 7

Dependent variable: logit Contact	β	Model 7 std. err.
Fixed Effects		
const	-.104	.310
Work Morning	-.217	.051
Work Afternoon	-.137	.040
Sat Afternoon	-.161	.085
Sun Morning	-.466	.158
Sun Evening	.343	.147
F2F dummy	1.157	.411
Impediment	-.166	.047
House average	-.083	.037
House poor	-.011	.078
CompletedHH	.501	.136
Random Effects: Country level		
Const.	.696	.377
F2F	1.257	.674
Compl.HH	.133	.074
Random Effects: Interviewer		
Constant	.654	.084
F2F dummy	1.026	.166
Compl.HH	.325	.057
Covariances: Country level		
Const, F2F	-.621	.428
Constant, F2F	-.039	.159
Const, Compl.HH	-.110	.125
Covariances: Interviewer level		
Constant, F2F	-.471	.101
Constant, Compl.HH	-.048	.056
F2F, Compl.HH	-.065	.067

5.2 Issue 2: Contact Effect on Respondent Cooperation

Of course, the ultimate objective of any contact attempt is to obtain a completed interview with at least one eligible respondent, not just to obtain contact with a household member. We are particularly interested in whether there are aspects of the initial attempt yielding contact that would have an influence on the final outcome of respondent participation.

The dependent variable is thus set to the binary variable *completedHH*, equaling 1 if the household approached ultimately has at least one respondent agreeing to be interviewed, else 0. The database is

reconfigured to include only attempts that result in the initial contact. Subsequently, the number of observations decreases from total initial contact attempts on all sample lines.

Again, we start with the base model, listed below as Model 8 (Table 9), but proceed with the complexity models more quickly than in the first seven models. The fixed effect of the intercept is significant, as is the random effect on the interviewer levels, with the country level effect still only barely significant at the $p < 0.05$ level.

Table 9: Null model on completed interview with at least one household member using first call attempt yielding contact, (N=20,486), Model 8.

Dependent variable: CompletedHH	Fixed Effects	Random Effects	
		Country	Interviewer
Constant (standard error)	-.838 (.097)	.064 (.038)	.604 (.046)

Following the iterative steps listed above, we include interaction variables for day (weekday, Saturday, Sunday) and mode of the attempt (either in-person (F2F) or not), dropping statistically insignificant terms and re-estimating the model until all terms have coefficients greater than their standard errors. The excluded combinations automatically serve as base category (aggregated).

Model 9 (Table 10) is suggestive of the importance of completing an initial phone contact on Saturday evenings, a contact on working day evenings with either mode, or on working day afternoons by F2F, or on Sunday afternoons by phone in order to have this household interviewed. A poor choice is weekend morning in-person, which is not very surprising.

Finally we are interested in determining whether the time of day and day of the week of the first contact is different for different interviewers or different countries. Forty-four-and-a-half (44.5) percent of successful first contacts were performed on workday afternoons. Subsequently, we only consider this combination and investigate the random effects on both levels, after controlling the condition of the environment, the state of the house, and whether there are potential impediments:

Table 10: Time of day by type of day and F2F included (with fixed effects), first successful contact, (N=20,486), Model 9.

Dependent variable: Logit Completed HH	Fixed Effects	Random Effects	
		Country	Interviewer
Constant	.318	0.307	0.199
(std. err.)	(0.199)	(0.156)	(0.020)
Sun. Morn.*F2F	-.424		
(std. err.)	(0.181)		
Sun. Aftern*. phone	.259		
(std. err.)	(0.154)		
Sun.Even.*F2F	-.231		
(std. err.)	(0.172)		
WorkMorn*F2F	-.078		
(std. err.)	(0.057)		
WorkAftern*F2F	.137		
(std. err.)	(0.046)		
WorkEve.*phone	.176		
(std. err.)	(0.057)		
WorkEven.*F2F	.124		
(std. err.)	(0.054)		
Sat.Morn.*F2F	-.324		
(std. err.)	(0.141)		
Sat.Even.*phone	.560		
(std. err.)	(0.213)		
Sat.Even.*F2F	0.19:		
(std. err.)	(0.154)		

Table 11: Working day afternoon fixed and random effects and F2F included (with fixed effects), first successful contact (N=20,486), Model 10.

Dependent variable: Logit CompletedHH	Fixed Effects	Random Effects	
		Country	Inter-viewer
Constant	.334	0.327	0.223
(std. err.)	(0.205)	(0.166)	(0.027)
WorkAftern.*phone	-.094	0	.186
(std. err.)	(0.056)		(.077)
WorkAftern * F2F	.113	.007	.099
(std. err.)	(0.051)	(.009)	(.039)
Conditions good	.112		
(std. err.)	(0.051)		
Conditions bad	-.073		
(std. err.)	(0.079)		
House good	.218		
(std. err.)	(0.036)		
House bad	-.249		
(std. err.)	(0.074)		
Impediments	-.228		
(std. err.)	(0.043)		

Contacting a household in-person on a working day afternoon increases the probability to realize an interview with this household, the more the environments and especially the house is subjectively evaluated to be in good condition and without

impediments. More importantly, there are virtually no random effects of the choice of this time and mode on the country level, whereas there are significant random effects on the interviewer level. That is – controlling for physical appearance of the housing – time and mode does not have a variance across countries as regards to the dependent variable considered. All variance (more for phone calls) are therefore on the interviewer level.

6. Discussion and Conclusions

Surprisingly, we do not find any noteworthy significant country effects throughout our analyses, once the interviewer effects are controlled. That is, it simply did not appear to be the case that either the circumstances or the difficulties to obtain contact in some countries than in others are different. However, the interviewer effects within each country are quite significant on obtaining contact with a household, even as we introduce additional explanatory variables including day of the week and time of the day that contact was attempted.

In addition, the results point to an increased efficiency associated with evenings especially on Sundays, at least on the very first contact attempt. This is not surprising to us, as it confirms similar studies in the United States and Britain.

Sunday mornings, however, appear to be a call window with decreased initial contact efficiency. For attempts that result in initial contact, Saturday afternoons are most likely to eventually lead to completed interviews, followed by initial contact on weekdays during the daytime. However, Sunday and Saturday morning and partially Sunday evening attempts by F2F appear to generate an annoyance factor with household members, decreasing the willingness of respondents to ultimately participate.

Throughout the analysis, interviewer effects persist. That is, there are interviewers who more likely find ways to obtain initial contact than other interviewers, even when controlling for when and how the attempt is made. Clearly, more research needs to be conducted to determine what accounts for these differences.

Acknowledgements

We gratefully acknowledge funding for the SHARE project. The SHARE data collection has been primarily funded by the European Commission through the 5th framework program (project QLK6-CT-2001-00360 in the thematic program Quality of Life). Additional funding came from the US National

Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, Y1-AG-4553-01 and OGHA 04-064). Data collection in Austria (through the Austrian Science Foundation, FWF), Belgium (through the Belgian Science Policy Administration) and Switzerland (through BBW/OFES/UFES) was nationally funded. The SHARE data set is introduced in Börsch-Supan et al. (2005); methodological details are contained in Börsch-Supan and Jürges (2005). We also wish to acknowledge support of the Swiss Household Panel at Neuchâtel and the University of Michigan Institute for Social Research, Survey Research Operations in Ann Arbor, Michigan.

References

- Alcser, Kirsten and Grant Benson (2005) "The SHARE Train-The-Trainer Program" in *The Survey of Health, Ageing and Retirement in Europe – Methodology*. Börsch-Supan and Jürges (eds.) Mannheim: Mannheim Research Institute for the Economics of Aging.
- Börsch-Supan, Axel and Hendrik Jürges (eds.) (2005) *Health, Ageing and Retirement in Europe - Methodology*, Mannheim: MEA, 2005.
- Campbell, P, Patrick Sturgis and Susan Purdon (1997) "Can you hear me knocking: An investigation into the Impact of Interviewers on Survey Response Rates." The Survey Methods Centre at the Social and Community Planning Research, London. URL: <http://joni.soc.surrey.ac.uk/~scs1ps/papers/can%20you%20hear%20me%20knocking.pdf>
- De Luca, Giuseppe and Oliver Lipps (2005) "Fieldwork and Survey Management in SHARE" in *The Survey of Health, Ageing and Retirement in Europe - Methodology*. Börsch-Supan and Jürges (eds.) Mannheim: Mannheim Research Institute for the Economics of Aging.
- De Luca, Giuseppe and Franco Peracchi (2004) "Survey participation in the first wave of SHARE," unpublished manuscript, Mannheim 2004.
- Greenberg, Betsy S. and S. Lynne Stokes (1990) "Developing an Optimal Call Scheduling Strategy for a Telephone Survey." *Journal of Official Statistics*, Vol. 6 (4), pp. 421-435.
- Groves, Robert M. and Mick P. Couper (1998) *Nonresponse in Household Interview Surveys*. Wiley Series in Probability and Statistics. New York: Wiley.
- Groves, Robert M. John Van Hoewyk, Grant Benson, Paul Schulz, M. Patricia Maher, Lynette Hoelter, William Mosher, Joyce Abma, and Anjani Chandra (2003) "Using Process Data from Computer Assisted Face to Face Surveys to Make Survey Management Decisions." Paper presented at the 2003 AAPOR Conference, Nashville, TN.
- Häder, Sabine and Siegfried Gabler (2003) "Sampling and Estimation." Chapter 8 in Harkness, Van de Vijver, and Mohler (eds.) *Cross-Cultural Survey Methods*. Wiley Series in Survey Methodology (Hoboken: Wiley & Sons).
- Kalsbeek, William D., Steven L. Bottman, James T. Massey, and Pao-Wen Liu (1994) "Cost-Efficiency and the Number of Allowable Call Attempts in the National Health Interview Survey" *Journal of Official Statistics*, Vol. 10 (2), pp. 133-152.
- Kirgis, Nicole et al. (2004) "Health and Retirement Study 2004 Study Manual." Unpublished interviewer training manual.
- Klevmarken, Anders (April 2005) "Sample Design" in *Health, Ageing and Retirement in Europe: First Results from the Survey of Health, Ageing and Retirement in Europe*. Börsch-Supan et al. (eds.) Mannheim: Mannheim Research Institute for the Economics of Aging.
- Kulka, Richard A. and Michael F. Weeks (1988) "Toward the Development of Optimal Calling Protocols for Telephone Surveys." *Journal of Official Statistics*, Vol. 4 (4), pp. 319-332.
- O'Muirheartaigh, Colm and Pamela Campanelli (1998) "The Relative Impact of Interviewer Effects and Sample Design Effects on Survey Precision." *Journal of the Royal Statistical Societies*, Vol. 161 (1), pp. 63-77.
- O'Muirheartaigh, Colm and Pamela Campanelli (1999) "A Multilevel Exploration of the Role of Interviewers in Survey Non-Response." *Journal of the Royal Statistical Society*, Vol. 162 (3), pp. 437-446.
- Purdon, Susan et al. (1999) "Interviewers' Calling Strategies on Face-to-Face Interview Surveys." *Journal of Official Statistics*, Vol. 15 (2), pp. 199-216.
- Rasbash, J., F. Steele, W. Browne, and B. Prosser (2004). *A user's guide to MLwiN version 2.0*. London, Institute of Education. URL: <http://multilevel.ioe.ac.uk/>.
- SRC (2004) "Survey of Health, Ageing and Retirement in Europe. Interviewer Project Manual. Main Test 2004." Unpublished manuscript, Survey Research Center, University of Michigan. URL: <http://www.share-project.org/index.php?page=Documentation&menue=4&sub=>
- Weeks, Michael F., Richard A. Kulka, and Stephanie A. Pierson (1987) "Optimal Call Scheduling for a Telephone Survey." *Public Opinion Quarterly*, Vol. 51, pp. 540-549.