Survey Research Methods (2008)
Vol.2 , No.3 , pp. 107-121
ISSN 1864-3361
http://www.surveymethods.org

Non-response in a survey among immigrants in Denmark

Mette Deding, Torben Fridberg and Vibeke Jakobsen
Danish National Centre for Social Research

The purpose of this paper is to analyse the character of response bias in a Danish survey among native Danes and immigrants from Turkey, Iran and Pakistan and thus enable the tailoring of future surveys to particular immigrants groups. We find that both contact rates and cooperation rates are lower for immigrants, leading to a significantly lower overall response rate. We also find important differences between groups - immigrants from Pakistan are especially difficult to reach, while refusals are particularly high among those from Turkey. Language is likewise important as a very large share of women could not be interviewed in Danish. We analyse not only the determinants of the probability of contact and the probability of cooperation but also the determinants of overall nonresponse, by looking at individual characteristics and observable interviewer characteristics. We find that the characteristics of the sample persons are important for both contact and cooperation rates, with different factors affecting each. Yet none of the observable interviewer characteristics appear to affect the response rate. Furthermore, after controlling for all the other variables, we find that the lower probability of response among immigrants compared to native Danes persists. The analysis clearly points to the need for tailoring surveys directed to immigrant groups to avoid response bias.
Keywords: Non-response, non-contacts, refusals, immigrants

Introduction

Since the 1970s, the number of immigrants has increased significantly in Denmark as in other European countries, and so has the need for knowledge about these immigrant groups. Consequently, in recent years immigrants have constituted a larger share of national surveys, and special surveys have been conducted among different immigrant groups. The basic lesson from these surveys is that non-response rates are typically relatively large among immigrants, and that interviewing immigrants requires considerations other than those applied to interviewing the majority population. However, only few studies have focused on non-response among immigrants (see e.g. Feskens et al. 2007; Dale and Haraldsen 2000; van den Brakel et al. 2006).

The main reason for concern about high non-response rates is that non-response may generate bias problems. In particular, non-response poses a problem if it is correlated with the variables of interest. Previous research has shown that while non-response bias occurs, the non-response rate of a survey alone is not a good predictor of the magnitude of the bias (Groves 2006). Blind efforts to reduce non-response may increase the bias problems. Instead, efforts at reducing non-response should be guided by knowledge about the character of the non-response bias and about the ways in which groups are affected by the efforts (Groves 2006).

The aim of this paper is to analyse the character of non-response bias in a Danish survey among Danes and immigrants from Turkey, Iran and Pakistan in the interest of allowing the tailoring of future surveys to specific immigrants groups. This Danish survey more than fulfilled the expectation that non-response among the immigrants would be high - on average the response rate among the three immigrant groups were about 20 percentage points lower than for the Danes. However, the survey collecting process also revealed large differences between the immigrant groups and in reasons for non-response - for instance, contacting immigrants from Pakistan was very difficult, while refusals were a greater problem among the immigrants from Turkey.

First, we analyse how contacts as opposed to non-contacts and cooperation as opposed to noncooperation depend on various characteristics for the sample persons and for the interviewers. Different types of non-response are likely to have different causes and different consequences. Nonresponse is in this paper categorised in three groups: non-contacts, refusals and other reasons.

However, non-contacts and refusals constitute the largest components of non-response, and therefore those of greatest concern to survey methodologists (Singer 2006).[1] Separate analyses of contact and cooperation will give us information that researchers can then use to tailor future immigrant surveys. Second, we analyse how overall response, as opposed to non-response, depends on various characteristics for both sample persons and interviewers. Looking at overall response and non-response will give us information about the nature of the non-response bias in the data. For the estimations, we use multilevel models. We use register information for all the individuals selected for interviews and information

Contact information: Vibeke Jakobsen, SFI - The Danish National Centre for Social Research Herluf Trolles Gade 11, DK-1052 Copenhagen K, Denmark, e-mail: VIJ@SFI.DK

[1] That the group of refusals and the group of non-contacts differ with respect to central characteristics is found in Turner (1999) and Campanelli and O'Muircheartaigh (1999), among others.

about the interviewers from the survey organisation.

Theory, hypotheses and previous literature

Groves and Couper (1998) have developed a conceptual model distinguishing between factors influencing survey participation: (a) the societal environment, (b) characteristics of the sample persons, (c) the survey design, (d) the characteristics of the interviewers, and (e) the interaction between the sample persons[2] and the interviewers. The societal environment factors include the socio-demographic and socio-economic composition of a society, as well as attitudes, norms or values among members of a society (e.g. trust, civic duty, fear of crime). Changes in the societal environment over time affect the development in survey response rates. The survey design among other things includes the sampling frame, mode of data collection, number of contact attempts and length of the data collecting-period. Interviewer guidelines whether helping to emphasise a survey sponsor or to stress the potential benefit of a survey to applied groups or individuals - are also survey design features that affect participation. The next section describes the design of the survey that this paper analyses.

However, this paper focuses mainly on the influence of the characteristics of the sample persons and the influence of the interviewers on the contact rate, cooperation rate and overall response. The sample person characteristics most widely studied are socio-demographic characteristics, such as age, marital status, household structure, education, employment status, income and urbanisation (Groves and Couper 1998).[3] Widely studied interviewer characteristics are gender, age and experience as an interviewer (Campanelli and O'Muircheartaigh 1999). These factors may affect the contact rate and cooperation-rate in different ways.

Characteristics of sample persons and contacts

Obviously, different characteristics of the sample persons are correlated with the probability of contacting them. For example, people who are busy with activities outside the home (e.g. work and education) will be more difficult to contact (Abraham et al. 2006). Studies have shown that students and employed persons are more difficult to reach than those outside the labour force and those who are unemployed, just as people who work long hours are more difficult to reach than people who work part-time (Stoop 2004; Abraham et al. 2006). Time spent on activities outside the home may also be related to age, such as engagement in sports. Empirical studies typically find that contact rates are lowest for young people (Groves and Couper 1998; Stoop 2004; Abraham et al. 2006).

The family or household structure may also matter for the contact probability. The larger the number of adults in the household, the larger the probability of contact with someone in the household (Groves and Couper 1998), who can then give information about the sample person (e.g. mobile phone number or information about when the sample person is at home). The implication is that singles are more difficult to contact (Stoop 2004). In addition, the presence of young children may affect the contact probability if households with young children more often have an adult caregiver at home than households without young children. Empirical studies thus show that having children in the household has a positive effect on the contact rate (Groves and Couper 1998; Stoop 2004). Furthermore, since adult caregivers almost always are women, this finding implies that women are easier to contact than men, other things being equal.

Characteristics of the housing conditions and urbanisation may affect the interviewer's probability of getting in contact with the sample person. Special security features, which may limit interviewer access, are typically more widespread in high-crime areas and in blocks of flats (i.e. in urban areas). Empirical studies have found a negative effect of urbanisation on the contact rate (Stoop 2004; Groves and Couper 1998). Living in urban areas may also be correlated with other individual characteristics that affect the probability of contact, e.g. time used outside home may be higher in urban areas because of more entertainment options, longer commutes, difference in employment rates between urban and rural areas, differences in age structure, etc.

From the previous studies we expect that employment, a high level of education (through its positive effect on the employment probability) and young age have negative effects on the contact probability, because individuals with these characteristics are at home less often. Urbanisation is also expected to influence the contact probability negatively. On the other hand, being married and having children are expected to have positive effects on the contact probability. Although we expect these factors to be important across sample groups, as the prevalence of the factors differs, so will their importance. For instance, we know that the immigrants have lower employment rates (especially the women from Turkey and Pakistan) and a lower level of education (especially the immigrants from Turkey) than native Danes (Dahl and Jakobsen 2005). Furthermore, the immigrants from Turkey and Pakistan are married and have children to a higher extent than the Danes, while the immigrants from Iran live as singles without children to a higher extent than the Danes. Finally, the immigrants are much more concentrated in urban areas (Statistics Denmark 2008 and our calculations, see Table 4).

In addition to the factors that are relevant both for Danes and immigrants, we look at some specific immigrant factors: an increase in years since migration and having a Danish citizenship are expected to have a negative effect on the contact probability for two reasons. One is that both factors are positively related to economic assimilation (e.g. employment probability and wage rate) (Ekberg 1994, Chiswick et al. 1997, Husted et al. 2001). The other is that both factors may be positively related to participation in cultural and sports activities in the local community, thereby suggesting lower contact probabilities after controlling for the employ-

[2] Sample persons are individuals selected to participate in the survey (Groves and Heeringa 2006).

[3] Groves and Couper categorise urbanisation under the category 'societal environment'.

Table 1: Overview of hypotheses concerning respondent characteristics and contact and cooperation

	Contact	Cooperation
Gender (men opposed to women)	(-)	(-)
Age	(+)	(-)
Couple (as opposed to single)	+	+
Children	+	+
Urban (as opposed to rural)	-	-
Education	-	+
Employment	-	+
Years since migration	-	+
Danish citizenship	-	+

ment situation.[4] Another very important factor for the immigrants is language problems that may make contact more difficult - because not only communicating with the sample person, but also getting information about the sample person (e.g. from a spouse) are both more difficualt. Hence, the effect of marriage on contact probability may be smaller for immigrants than native Danes.

Characteristics of sample persons and survey cooperation

Different hypotheses exist about the relationship between characteristics of the sample person and cooperation in the survey. We base our hypotheses on the social isolation hypothesis, which some have used to explain lower cooperation response rates among racial and ethnic subgroups (see Groves and Couper 1998). According to this hypothesis, social isolates are out of touch with mainstream culture behaving in accordance either with sub-cultural norms or in rejection of the dominant norms. The hypothesis is that socially isolated persons will be less likely to cooperate with a survey request that represents the broader society (e.g. government agencies). According to the social isolation hypothesis, individuals with lower socio-economic status are expected to have lower cooperation rates, as they are likely to be alienated from central social institutions. Conversely, individuals with higher socio-economic status may perceive themselves as occupying an important social place and consequently either have a higher sense of civic duty or recognize the value of survey data as a common good. However, the empirical evidence on the relationship between income and education on the one hand and cooperation rates on the other is mixed. For instance, some studies find a positive relationship between the cooperation rate and education, while other studies find the opposite result (Groves and Couper 1998). Although according to the social isolation hypothesis, race and ethnicity also influence the cooperation rate, Groves and Couper do not find evidence of an effect of these variables after controlling for socioeconomic variables.

Sample persons age may affect the cooperation rate in different ways - one hypothesis is that the elderly are expected to have lower cooperation rates because of disengagement; another hypothesis is that the elderly have a higher sense of civic duty, leading to higher cooperation. The empirical finding in the literature appears consistent with lower cooperation among the elderly (Groves and Couper 1998; DeMaio 1980). With respect to gender, most studies find either no gender effect on the cooperation rate or lower cooperation rates for men. The explanations for the latter can also be related to the social isolation hypothesis, if it holds true that women take more care of social relations than men (Groves and Couper 1998).

The social isolation hypothesis also predicts that household indicators can affect cooperation: people living in single-person households are expected to have lower cooperation rates (tendency to social isolation); households with children, to have higher cooperation rates (through schools and networks of friends); and those living in large blocks of flats, to have lower cooperation rates (less contact with neighbours, greater transience). Thus sample persons in urban areas are likely to have lower cooperation rates than those in rural areas, because large blocks of flats are urban phenomena. Empirical studies show without exception that the presence of children increases the cooperation rate, while the evidence with respect to single-person households is mixed (Groves and Couper 1998).

Given the social isolation hypothesis, we would expect being employed, having a high level of education, living in rural areas, being married and having children to have positive effects on the cooperation rate. The number of years since migration and having a Danish citizenship are also expected to have a positive effect on the cooperation rate, as these factors are related to the assimilation of immigrants into Danish society.

Table 1 summarises our hypotheses with respect to sample persons characteristics and the contact and the cooperation phases. Some factors are expected to have the same effect on contact and cooperation, while others are expected to work in opposite directions. While we also expect the different factors to be important both for Danes and immigrants, as the prevalence of the factors differs (e.g. the employment rate and the share having children), so will their importance.

Characteristics of interviewers and non-response

No matter whether interviews are carried out by telephone or face-to-face, they involve both a interviewee and an interviewer, and expecting that the interviewer may unintentionally affect the response-rate is natural. For instance, the interviewer's expectations of and attitudes towards nonresponse may influence the response rate (Campanelli and O'Muircheartaigh 1999). Although we have no information on the expectations and the attitudes of the interviewer, socio-demographic characteristics of the interviewer and interviewer experience may affect the interviewer's expectations and behaviour and therefore the response rate (Groves and Couper 1998). One study has thus found that female interviewers are more likely than male interviewers to be perceived as friendly. However, little empirical evidence sup-

[4] Other things being equal, we expect a person's network to be larger the longer the person has lived in the local community.

Table 2: Response and non-response, percentages

	Denmark		Iran		Pakistan		Turkey	
	Men	Women	Men	Women	Men	Women	Men	Women
Response (completed and partial interviews)	76.87	82.03	62.65	64.87	41.37	42.32	55.04	55.28
Non-response	23.13	17.97	37.35	35.13	58.63	57.68	44.96	44.72
Causes of non-response:								
Non-contact	8.21	4.09	16.75	14.36	32.75	29.68	16.09	12.63
Refusals	13.62	12.81	13.44	14.36	16.86	17.05	22.87	25.47
Other reasons	1.31	1.07	5.24	4.36	5.88	5.89	3.10	1.04
Total	100	100	100	100	100	100	100	100
n observations	536	562	573	390	510	475	516	483

ports the view that female interviewers in general have higher response rates. The results for the age of interviewers are also mixed (Campanelli and O'Muircheartaigh 1999). Interviewer experience is supposed to have a positive affect on the response rate (Groves and Couper 1998). Pickery and Loosveldt find that both the chances of refusals and non-contacts are subject to interviewer effects and that the experience of the interviewer is particularly important (Pickery and Loosveldt 2002).

Previous surveys on immigrants and non-response

The experience of previous surveys among or including immigrants is that interviewing immigrants involves specific problems and higher non-response rates (see e.g. Feskens et al. 2006; Feskens et al. 2007; Dale and Haraldsen 2000). A study based on survey data from six European countries shows that the higher non-response rate among immigrants in these countries is driven by higher non-contact rates and higher non-response e.g. due to inadequate proficiency in the host country language. Yet the cooperation rate is higher for the immigrants in the same study (Feskens et al. 2006).

Furthermore, the non-response rate typically varies according to country of origin. Some European studies find a significant difference between immigrants from Western countries and those from non-Western countries. For instance, Feskens et al. (2007) find for the Netherlands that the response pattern of immigrants with a Western background is very similar to the response pattern of the native population, but that response among immigrants with a non-Western background is considerably lower. Higher response rates for Western immigrants than non-Western immigrants are also found in a Danish survey (Mørkeberg 2000).[5]

Two Danish studies based on survey data collected among immigrants show bias in the samples as a consequence of non-response. Persons with a low level of education and persons without employment have a relatively high non-response (Nielsen and Pedersen 2000; Jakobsen 2004). However, Feskens et al. (2007) find for immigrants in the Netherlands that the single most important factor for getting a response is the degree of urbanisation, a facet even more important for establishing contact with the sample persons. Feskens et al. thus claim that much of the apparent negative effect of ethnicity on response rates is attributable to immigrants being more urbanised than the majority population. Furthermore, ethnicity is found to have a positive effect on participation for sample persons who are actually contacted. All of these studies point to immigrants as a difficult group to survey.

The survey

Data used for the analysis is a Danish survey collected in 2006 including 18-to-45-year-old immigrants from Turkey, Iran and Pakistan and 18-to-45-year-old Danes. The immigrants came to Denmark before 2006 (December 2005 at the latest). The selection of sample persons into immigrants and native Danes is based on Statistics Denmark's classification of the population into three groups: immigrants, descendants of immigrants and native Danes (Pedersen 1991). Statistics Denmark defines immigrants as persons who are foreign born, whose parents are foreign born or who hold a foreign citizenship. Descendants of immigrants are defined as persons born in Denmark, and whose parents either are foreign born or hold a foreign citizenship. Native Danes are defined as persons who have at least one parent, who is a Danish citizen and who was born in Denmark.[6] For simplicity, we will continue referring 'native Danes' as 'Danes' throughout the paper. We selected 4,045 individuals for interviewing - nearly 1,000 from each of the immigrant groups and about 1,100 Danes. The sample was drawn as a simple random sample of individuals living in private households in Denmark from each of the four groups in the Danish Central Person Register (CPR). The CPR has approximately 99.9% cov-

[5] The exact definition of Western vs. non-Western countries can differ. In Denmark, Western countries are typically defined as countries in the European Union, Iceland, Norway, Switzerland, North America, Australia and New Zealand. All other countries are defined as non-Western (Tranæs and Zimmermann 2004). In a Dutch study Western countries are defined as Europe, the USA, Canada, Australia, Japan and Indonesia (Feskens et al. 2007).

[6] Descendants are not included in this analysis.

Table 3: Response, contact, refusal and cooperation rates

	Denmark		Iran		Pakistan		Turkey	
	Men	Women	Men	Women	Men	Women	Men	Women
Response rate	76.87	82.03	62.65	64.87	41.37	42.32	55.04	55.28
Contact rate	91.79	95.91	83.25	85.64	67.25	70.32	83.91	87.37
Refusal rate	13.62	12.81	13.44	14.36	16.86	17.05	22.87	25.47
Cooperation rate 2	83.74	85.53	75.26	75.75	61.52	60.18	65.59	63.27
Cooperation rate 4	84.95	86.49	82.34	81.88	71.04	71.28	70.65	68.46

Note: The calculations of the response, contact and cooperation rates are based on the standard definitions issued by the American Association for Public Opinion Research (AAPOR, 2006).
Response rate = (I+P)/(I+P+R+NC+O)
Contact rate = (I+P+R+O)/(I+P+R+O+NC)
Refusal rate = R/(I+P+R+NC+O)
Cooperation rate 2 = (I+P)/(I+P+R+O)
Cooperation rate 4 = (I+P)/(I+P+R)
I=Completed interview
P=Partial interview
R=Refusal and break-off
NC=Non-contact
O=Other

erage of persons living in Denmark and includes all persons who expect to stay in Denmark for at least 3 months.

We conducted the survey as part of a study on the integration of young first-generation immigrants into the Danish educational system and labour market. The questionnaire included questions about family structure, years since migration, education (in both Denmark and the country of origin), employment, working hours, job search, working experience, proficiency in Danish, social networks, housework, religion, and attitudes towards employment and gender roles.

The survey data has subsequently been merged to administrative register data from Statistics Denmark. This register data includes information for all the individuals selected for interview, such as gender, age, family situation, region, citizenship, education obtained in Denmark, employment history in Denmark and years since migration. All of these background variables are from 2006, except for the variables on education and employment, which are from 2005 and 2003, respectively (the latest available information).

In addition, we have some information about the interviewers assigned to the survey, including age, gender, experience as an interviewer (tenure in the survey organisation), and the number of interviews the interviewer was assigned. The information about the interviewers comes from the survey organisation (SFI Survey). Because the interviewing period was very long (see the following section), in some cases sample persons were reassigned to a different interviewer. Unfortunately, we have information only about the last interviewer assigned to the sample person. Likewise, we do not know how many different interviewers each sample person was assigned to during the interviewing period. Thus, we do not know the number of interviews originally assigned to each interviewer but only the number assigned after the transfers between interviewers.

The data collection process

The interviewing was carried out either by visits or by telephone. An introductory letter announced that an interviewer would contact the sample person by telephone or by visit to make an appointment for the interview. The immigrants received two letters: one in Danish and one in Turkish, Farsi or Urdu. Although the interview was to be carried out in Danish if possible, the questionnaire was translated into the relevant languages (Turkish, Farsi and Urdu) and interviewers speaking the relevant language were available for assignment to the interview. Each interview was approximately 40 minutes.

The mode of data collection was interview by telephone (CATI) supplemented with face-to-face interviews (CAPI). The interviewer was to try to make contact by telephone at least six times at different hours of the day and on different days of the week. If the interview could not be done by telephone, the interviewer was to try to make an appointment for a visit. If telephone contact did not work, the interviewer was to visit the address at least three times. However, as the survey was expected to be difficult, the interviewers had some flexibility for making contact. If the interviewer had the impression that an interview was more likely by visit than by telephone, he or she could simply visit the address instead.[7] Furthermore, if language problems made conducting the interview in Danish impossible, an interviewer speaking the relevant language was available. Finally, if neither mode of interviewing proved possible, the interviewer hold the option of offering the potential interviewee a chance for self-completion while the interviewer sat waiting or, as a last resort, of leaving the questionnaire.

The interviewers did the majority of interviews by CATI, 93% for the Danes and between 67 and 76% among the immigrant groups. For all three immigrant groups more, the

[7] If no phone number was available, the interviewer could also start with a visit.

CAPI-mode was used more often for women than for men. The Turkish group proved to be most difficult to interview by telephone: 34% (29%) of the Turkish women (men) were interviewed by CAPI, and 43% (37%) of the Turkish women (men) were interviewed by a Turkish-speaking interviewer. The figures are almost as high for the Pakistani group: 32% (27%) of the Pakistani women (men) were interviewed by CAPI, and 38% (33%) of the Pakistani women (men) were interviewed by an Urdu-speaking interviewer. For the Iranians 27% (21%) of the women (men) were interviewed by CAPI, and 24% (16%) of the women (men) were interviewed by a Farsi-speaking interviewer. These figures show the necessity of using interviewers who speak the relevant languages doing surveys among immigrant groups.

Because the survey was expected to be difficult, the survey period was unusually long initially planned from February to June 2006. However, by June the response rate was so low - especially among the Pakistanis - that the interviewing period was extended to November 2006. Moreover, in June many interviews were reassigned to different interviewers.

Response and non-response

Of the 4,045 individuals selected for interviewing, 2,448 individuals were interviewed, corresponding to an overall response rate of 60.5 (table 2).[8] However, the response rate varies greatly across countries of origin: roughly speaking, approximately 40% for the Pakistanis, 55% for the Turks, 60% for the Iranians, and 80% for the Danes. Thus, the survey clearly demonstrates that the immigrant groups are more difficult to survey than the native population. In addition, the survey shows the differences between the various ethnic groups. Yet, gender does not appear to be very important. The only group with a significant gender difference in response rates is the Danes, where women have a higher response rate than men. Although language problems appear larger for immigrant women than immigrant men, this difference is not reflected in the overall response rate.

Table 2 also shows the distribution of the non-response in three categories: 1) Non-contacts, including the subcategories Moved, Not met, and Fictitious address[9]; 2) Refusals, including the subcategories Refusal - lack of time, Refusal - suspicious, Refusal by parent, Refusal by spouse, Refusal due to the gender of the interviewer, and Refusal other reasons; 3) Other reasons, including the subcategories Illness, Hospitalised/away from home, Handicapped, Moved out of the country, Dead, and Language problems. The distribution of all sub-categories appears in appendix table 1.

We find marked differences in the distribution of non-response across the four sub-groups. Generally, the interviewers found it much more difficult to contact the immigrants, especially those from Pakistan: The share of non-contacts for Pakistan immigrants is about twice the share of each of the other two immigrant groups. Refusals, however, are the greatest problem among the Turkish immigrants. Indeed, the share of refusals is relatively similar across Danes, Iranians and Pakistanis, thus underlining that the low response rate for the Pakistani immigrants largely results from contact problems rather than refusals. We cannot know, however, what the cooperation rate for the noncontacts would be.

As previously mentioned the category 'refusals' comprises six sub-categories. Of these, the dominant ones are 'refusal due to lack of time' (especially for the men) and 'refusal for other reasons' (see appendix table 1). However, for 4-5% of the Turkish and Pakistani women, their husbands refused on their behalf (whereas no woman refused on behalf of her husband). This situation poses a special problem for the interviewers, who have to convince another person to allow the interview before they get the chance to convince the sample person herself.

Non-response due to 'other reasons' is a greater problem among the immigrants than among the Danes: the immigrants to a higher extent either left the country or cannot participate because of language problems (appendix table 1). Whereas language problems were only minor for the non-response among immigrants from Iran and men from Pakistan and Turkey, about 5% of the women from Pakistan and Turkey could not be interviewed because of language problems. In principle, the language problem category should not exist, because of the availability of bilingual interviewers. However, even though sample persons agreed to having a bilingual interviewer, the contact was not always successful, especially in the case of Pakistani and Turkish women. In addition, a very large share of the Pakistani and Turkish female interviewees were not interviewed in Danish. Again, this finding shows the importance of taking language problems into account in surveys of specific immigrant groups.

Next we define the contact and cooperation rates. These rates, calculated according to the Standard Definitions of Outcome Rates issued by the American Association for Public Opinion Research (AAPOR 2006), appear in Table 3. For these calculations, by definition we do not have any sample persons with unknown eligibility, because sample persons are drawn as a representative sample from the CPR register.[10] Furthermore, we do not distinguish between fully completed and partially completed interviews. Thus, that the calculated rates in Table 3 constitute the full set of rates to be calculated with the AAPOR standards (while all response rates, contact rates and refusal rates are identical, there are two different cooperation rates). As in other European studies (see Feskens et al. 2006) we find that the contact rate is lower for immigrants than for the majority population. However, unlike Feskens et al. (2006), we do not find that the cooperation rate is higher for immigrants. On the contrary, when we use cooperation rates number 2 and number 4 we find from the

[8] As very few interviews (about 1%) were only partially completed, we pool these with the completed interviews. Likewise, there were only very few item non-responses, probably because the interviewers had very clear instructions about getting answers to all questions ('don't know' is a category).

[9] A fictitious address occurs when a sample person gives an address at which he or she is not living. In Denmark, the main reason for a fictitious address is tax-evasion.

[10] Although some respondents are dead or have left the country by the time the interviewer tries to contact them, they remain part of the representative sample.

Table 4: Means of respondent variables pooled sample

	N	Mean	Std. dev.
Woman	3836	0.474	0.499
Age group 18-29	3836	0.327	0.469
Age group 30-39	3836	0.397	0.489
Age group 40-45	3836	0.276	0.447
Couple	3836	0.649	0.477
Single	3836	0.351	0.477
No children	3836	0.413	0.492
Children	3836	0.587	0.492
Copenhagen	3836	0.579	0.494
Urban area	3836	0.313	0.464
Rural area	3836	0.108	0.311
Danish education (in years)[1]	2397	11.914	2.581
No Danish education[1]	3836	0.375	0.484
Employed[2]	3836	0.587	0.492
Non-employed[2]	3836	0.387	0.487
Employment status unknown[2]	3836	0.026	0.159
Years since migration	2771	15.029	8.271
Years since migration unknown	2771	0.009	0.096
Danish citizen[3]	2771	0.484	0.500
Danish sub sample	3836	0.278	0.448
Iranian sub sample	3836	0.242	0.428
Pakistani sub sample	3836	0.226	0.418
Turkish sub sample	3836	0.254	0.435

[1] Information from 2005
[2] Information from 2003
[3] Only relevant for immigrants

Danish survey that the cooperation rate is lower for immigrants than for Danes (although cooperation rate 4 for the Iranian immigrants is close to the Danish rate). The refusal rate is especially high among the immigrants from Turkey, while the refusal rate for the Iranians is very similar to that for the Danes. We thus learn that, first, surveying immigrants are a challenge both for contact and for cooperation and that, second, large variation exists across groups. That country-specific experiences should not be generalised to all immigrant groups is very clear.

Strategy of analysis

In the empirical analysis, we focus on the different determinants of non-contact and cooperation and on the determinants of overall non-response. Therefore, we estimate three different models: (1) the probability of contact, (2) the probability of cooperation and (3) the probability of overall response. We estimate the first two models to learn more about the causes of non-response and estimate the third model to learn more about the bias that results from non-response.

As a definition of the contact rate we use 'Contact Rate 2' according to the AAPOR standard definitions, i.e. the number of interviewees, refusals and non-response for other reasons as a share of the total sample drawn from the population register (see the note to table 3). The cooperation rate is defined as 'Cooperation Rate 4' in the standard definitions, i.e. the share of those interviewed out of the total number of those who were interviewed and those who refused to be interviewed (see note to table 3).[11] The response rate, as previously mentioned, is defined as the number of completed and partially completed interviews as a share of the total random sample drawn from the CPR register.

Using both sample person-specific variables and interviewer-specific variables we estimate the three models for the pooled sample. However, to study potential differences between the ethnic groups, we also estimate separate models for the four groups.[12] In addition, because previous studies have shown that obtaining response is especially difficult in urban areas, we estimate separate models for Copenhagen.

All analyses apply logistic random multilevel models - more precisely a logistic random intercept model.[13] Multilevel models have become quite standard in the analysis of survey non-response because this type of data very often includes clustered information for instance on interviewers (Pickery and Loosveldt 2002).

Explanatory variables

The explanatory variables for the sample persons include information on gender, age, family situation (couple/single, no children/children), and region (Copenhagen/other urban/rural). These register variables are from 2006. Education is from 2005 and consists of the official duration of education obtained in Denmark. While some information is in the registers about education obtained outside Denmark, the quality and coverage of this information is poor. Instead, we include a dummyvariable for everybody without Danish education. Employment status is from November 2003 (the latest available information). For a minor share of the sample, employment information is not available - primarily because some immigrants were not in Denmark in 2003 but have immigrated or re-immigrated in 2004 or 2005.

For the immigrants, we include two specific variables: years since migration and having Danish citizenship (having Danish citizenship/not having Danish citizenship). The latest information in the registers regarding time of immigration is from 2004; however, using other register information we can identify individuals who immigrated in 2005. Consequently, 'years since immigration' is only unknown for about 1% of the immigrant sample. Information on citizenship is from 2006. The means of the sample person-specific variables appear in Table 4 for the pooled sample and in Table 5 for each of the subgroups.

Table 5 reveals significant differences between the countries especially in terms of family situation, region, educa-

[11] We have also estimated the models using cooperation rate 2 rather than cooperation rate 4. The differences in the estimation results are only minor.

[12] We do not estimate a separate model for gender, because of the relative small number of observations and minor gender differences in the response, contact and cooperation rates that appear Table 1.

[13] For the estimations, we use the Stata Program GLLAMM (see e.g. Rabe-Hesketh et al. 2005)

Table 5: Means of respondent variables - by country

	Denmark		Iran		Pakistan		Turkey	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Woman	0.512	0.500	0.412*	0.493	0.482	0.500	0.485	0.500
Age group 18-29	0.367	0.482	0.322*	0.467	0.310*	0.463	0.304*	0.460
Age group 30-39	0.382	0.486	0.271*	0.445	0.475*	0.500	0.465*	0.499
Age group 40-45	0.251	0.434	0.407*	0.492	0.215*	0.411	0.232	0.422
Couple	0.607	0.489	0.512*	0.500	0.712*	0.453	0.768*	0.422
Single	0.393	0.489	0.488*	0.500	0.288*	0.453	0.232*	0.422
No children	0.490	0.500	0.583*	0.493	0.321*	0.467	0.248*	0.432
Children	0.510	0.500	0.417*	0.493	0.679*	0.467	0.752*	0.432
Copenhagen	0.331	0.471	0.474*	0.500	0.911*	0.285	0.656*	0.475
Urban area	0.359	0.480	0.463*	0.499	0.076*	0.265	0.329	0.470
Rural area	0.311	0.463	0.064*	0.244	0.013*	0.112	0.014*	0.119
Education (in years)[1]	12.518	2.398	12.489	2.588	11.055*	2.363	10.365*	2.328
No Danish education[1]	0.000	0.000	0.396*	0.489	0.619*	0.486	0.548*	0.498
Employed[2]	0.796	0.403	0.480*	0.500	0.467*	0.499	0.567*	0.496
Non-employed[2]	0.204	0.403	0.481*	0.500	0.479*	0.500	0.416*	0.493
Employment status unknown[2]	-	-	0.039	0.193	0.054	0.227	0.016	0.127
Years since migration	-	-	13.949	6.924	14.113	9.049	16.872	8.414
Years since migration unknown	-	-	0.010	0.098	0.016	0.126	0.003	0.055
Danish citizen	-	-	0.677	0.468	0.373	0.484	0.398	0.490
n observations	1065		929		867		975	

[1] Information from 2005
[2] Information from 2003
* Significantly different from the Danish sample at a 5 pct. level

tion, and employment status. Immigrants from Pakistan and Turkey more often live in couples and more often have children than Danes and Iranians.[14] Almost all the immigrants from Pakistan live in Copenhagen. Immigrants from Pakistan and (especially) Turkey have less Danish education than Danes, while Iranians have the same educational level as Danes. Employment rates are much lower among the immigrant groups than among the Danes. As to years since immigration, although the three immigrant groups are very similar, the large standard deviation reflects great variation within the immigrant groups: Some immigrants have only been in Denmark for a few years, while others have been in Denmark almost all their lives. All these factors potentially influence non-response.

Table 6 presents the interviewer-specific variables. These variables include gender, age, seniority in the survey-organisation and number of interviews per interviewer. The interviewers are relatively old - with a mean age of 58 years, indicating that the survey organisation does not employ students (who in Denmark are often in their twenties) as interviewers, but rather people past their twenties, as the organisation has better experience with relatively older interviewers.[15] In addition, the interviewers are quite experienced - about 5 years each on average. Such experience should suggest better chances for positive responses. Number of interviews, the average number of sample persons assigned to the specific interviewer, is very large (and with a large standard deviation), underlining the need for taking into account the clustered observations in the empirical framework.

Table 6: Means of interviewer-specific variables

	Mean	Std. dev.
Woman	0.564	0.501
Age	58.200	9.952
Age group 30-39	0.073	0.262
Age group 40-59	0.364	0.485
Age group 60+	0.564	0.501
Seniority (years)	4.945	3.955
Seniority 1 year	0.164	0.373
Seniority 2-5 years	0.455	0.503
Seniority 6+ years	0.382	0.490
Number of interviews	70.036	71.702
n observations	55	

[14] According to the hypotheses it would also be highly relevant to include information about whether the respondents live in extended families. But unfortunately data does not include such information.

[15] In practice, the survey institute employs many senior citizens.

Estimation results

Contact

We begin the analysis by looking at the probability of contact for the pooled sample shown in the first column of Table 7. First, we find that contact is more likely for the age group 18-29 years compared to the left-out category, 30-39 years. Although young people are typically considered to be more difficult to contact, this assumption is not confirmed here. Gender is significant. After controlling for the other variables in the model, the contact probability is higher for women than men. In accordance with our hypotheses we also find, that individuals who live in couples and have children are easier to contact.

Several studies have shown that urbanisation has a negative effect on the probability of contacting the sample persons. We likewise find that the probability of contact is significantly lower in Copenhagen than in other areas. Yet no significant difference exists between the response in other urban and rural areas. This result indicates that the negative effect on contact probability from living in the Copenhagen area persists after controlling for all other factors (including the country dummies). Of course the problem of lower contact rates in a highly urbanised area like Copenhagen is especially problematic when a relatively large share of the sample lives in this specific area, and this factor calls for special attention by the survey organisation. In this case, we know that 92% of the immigrants from Pakistan live in Copenhagen (table 5), thus partly explaining the low contact rate among the Pakistanis.

For the pooled sample, we do not find any effect of 'years since migration' and socio-economic status - measured by education and employment variables - on the contact probability.[16] Thus our expectation that employed people for example would be more difficult to contact; because they spend more hours away form home than unemployed people, is not confirmed. We do, however, find a positive effect on contact probability from having Danish citizenship, a factor that is related to economic assimilation.

The country dummies show that, after controlling for all other factors, the contact probability is lower for immigrants from Pakistan and Iran than for Danes. The contact rate is lower in all the three immigrant groups than for Danes (table 3). However, for the immigrants from Turkey this result is attributable to the characteristics included in the model. That the lower contact rate for the immigrant groups not solely is attributable to a higher concentration of immigrants in Copenhagen is confirmed by the separate estimation for Copenhagen (table 7, column 6) - to a high extent the same variables are significant in columns 1 and 6.

The interviewer variables do not indicate any effect of the gender of the interviewer or of the number of interviews per interviewer, but rather that elderly interviewers (age 60+) have higher contact probability relative to the left-out category (interviewer 40-59 years). We also find that interviewers with very short tenure in the survey institute are less likely to make contact.

With the sample divided into four subgroups, the findings are highly similar to the pooled sample (see columns 2-5 in table 7). Which variables are significant differs from group to group (in many cases due to the small sample size). However, there are a few differences compared to the pooled sample. First, for the Danes, years of education have a positive impact on the contact probability, while education has no effect for the pooled sample or any of the immigrants groups. Second, for the Iranians living in rural areas has a positive effect on the contact probability compared to living in urban areas, while this variable is insignificant in the other estimations in Table 7.

Cooperation

Next we turn to the analysis of cooperation (table 8). A comparison of the first columns of Tables 7 and 8 reveals that the probability of contact and of cooperation for the pooled sample to a high extent is influenced by different characteristics of the sample persons. First, gender, age, whether the sample person is single or has children, has no significant influence on cooperation. As factors all influence the contact probability, we had expected that these factors would influence the cooperation probability. Second, education and employment influence cooperation, while these socio-economic factors did not influence the contact probability. Years of education and employment have - in agreement with the isolation hypothesis - a positive influence on cooperation.

Besides a negative influence on the probability of contact, living in Copenhagen also has a negative effect on the probability of cooperation. Thus, as expected, living in a metropolitan area has a strong negative effect on the response. The country dummies show that, after controlling for all other factors, the probability of cooperation is lower for immigrants from Turkey. That immigrants for some countries have a lower cooperation rate is in accordance with the isolation hypothesis. Contrary to the model of contact rates, none of the interviewer-specific variables are significant in this model. Although we believe that the personality traits of the interviewers affect the probability of cooperation, the observable interviewer characteristics appear to capture none of these effects.

With the sample divided into four subgroups, the findings are very similar to the pooled sample (see columns 2-5 in table 8). The significant variables differ from group to group (in many cases due to the small sample size), with only a few differences compared to the pooled sample. Being young has a positive influence on cooperation for the Danes, being old has a positive influence for the Pakistanis, and having children has a positive influence on cooperation for the Iranians. For some of the subgroups the interviewer variable 'number of interviews per interviewer' is significant: hence, we find a positive relationship between the number of interviews and the cooperation. One plausible explanation for this relationship may be that the survey institute has assigned the

[16] The special immigrant variables (years since migration and citizenship) are interacted with a dummy for being an immigrant.

Table 7: Estimation of contact rate

	Pooled sample		Denmark		Iran		Pakistan		Turkey		Copenhagen	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Women	0.283	0.116**	0.297	0.342	0.311	0.253	0.279	0.199	0.149	0.232	0.250	0.131*
Age group: 18-29	0.507	0.140***	0.775	0.402*	0.221	0.324	0.629	0.236***	0.915	0.303***	0.437	0.159***
Age group: 40-45	0.218	0.142	0.742	0.497	0.255	0.302	0.317	0.250	-0.224	0.273	0.018	0.159
Single	-0.583	0.146***	-0.852	0.431**	-0.428	0.268	-0.491	0.263*	-0.861	0.311***	-0.411	0.167**
Children	0.337	0.155**	0.884	0.491*	0.364	0.308	0.112	0.269	0.318	0.323	0.343	0.174**
Copenhagen	-1.375	0.204***	-1.737	0.565***	-0.923	0.336***	-1.917	0.512***	-1.222	0.373***		
Rural area	0.232	0.325	-0.446	0.521	1.335	0.803*			-0.136	1.193		
Danish education (in years)	0.049	0.032	0.268	0.082***	0.014	0.066	0.028	0.066	-0.107	0.069	0.032	0.037
No Danish education	0.520	0.381			0.048	0.881	0.403	0.762	-0.917	0.770	0.379	0.438
Non-employed	-0.184	0.118	0.016	0.385	-0.201	0.251	-0.144	0.197	-0.083	0.229	-0.054	0.134
Employment unknown	-0.121	0.360			-0.448	0.732	-0.213	0.484	-1.027	0.729	0.041	0.411
Years since migration	0.005	0.008			0.015	0.021	0.012	0.012	0.008	0.016	0.007	0.009
Years since migration unknown	0.426	0.650			0.913	1.213	1.021	1.002			0.704	0.774
Danish citizen	0.266	0.139*			-0.143	0.298	0.510	0.226**	0.502	0.252**	0.313	0.153**
Iran	-0.643	0.243***									-0.685	0.309**
Pakistan	-1.139	0.250***									-1.257	0.300***
Turkey	-0.361	0.270									-0.412	0.327
Interviewer women	-0.048	0.377	-0.715	0.646	-0.360	0.627	-0.139	0.559	-0.195	0.474	-0.223	0.455
Interviewer age: 30-39	0.703	0.694	-0.119	1.197	0.810	1.107	0.182	0.946	1.009	0.948	0.165	0.754
Interviewer age: 60+	0.741	0.401*	1.613	0.669**	0.737	0.638	-0.220	0.588	0.448	0.521	-0.169	0.494
Interviewer seniority 1 year	-0.881	0.505*	-1.198	0.978	-2.487	0.829***	-0.924	0.711	-0.980	0.638	-1.057	0.553*
Interviewer seniority 6+ years	-0.234	0.394	-0.252	0.644	-0.748	0.640	0.115	0.557	-0.449	0.502	0.088	0.474
Number of interviews per interviewer	-0.001	0.002	0.002	0.004	-0.004	0.003	-0.004	0.003	0.000	0.003	-0.003	0.003
Constant	2.733	0.731***	-0.036	1.465	3.618	1.393***	2.962	1.301**	4.058	1.150***	2.353	0.891***
Number of level 1 units	3836		1065		929		867		975		2222	
Number of level 2 units	55		47		49		44		52		43	
Log L	-1218		-169		-301		-425		-315		-919	

* significant at 10%, ** significant at 5%, *** significant at 1%

Table 8: Estimation of cooperation rate

	Pooled sample		Denmark		Iran		Pakistan		Turkey		Copenhagen	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Women	-0.001	0.098	0.111	0.194	0.053	0.225	0.121	0.225	-0.099	0.174	-0.012	0.124
Age group: 18-29	0.188	0.125	0.519	0.261**	-0.022	0.315	0.035	0.276	0.134	0.224	0.147	0.157
Age group: 40-45	0.080	0.119	-0.146	0.236	-0.032	0.274	0.540	0.268**	-0.023	0.222	0.240	0.152
Single	-0.064	0.129	-0.333	0.248	0.118	0.245	0.033	0.322	-0.070	0.267	0.030	0.169
Children	0.122	0.134	0.070	0.266	0.599	0.268**	0.008	0.324	0.063	0.274	0.110	0.172
Copenhagen	-0.591	0.146***	-0.417	0.241*	-0.751	0.271***	-0.832	0.417**	-0.556	0.220**		
Rural area	-0.043	0.197	0.134	0.253	-0.593	0.448			1.360	1.084		
Danish education (in years)	0.071	0.026**	0.114	0.043***	-0.013	0.062	0.036	0.071	0.118	0.058**	0.046	0.034
No Danish education	0.456	0.315			-0.309	0.849	-0.111	0.818	0.831	0.612	0.046	0.405
Non-employed	-0.186	0.103***	-0.233	0.241	-0.489	0.236**	-0.161	0.230	-0.096	0.174	-0.125	0.131
Employment unknown	0.919	0.498			0.638	0.830	0.362	0.695	1.730	1.117	0.937	0.575
Years since migration	0.001	0.008			0.026	0.020	-0.027	0.014*	0.012	0.012	0.001	0.009
Years since migration unknown	-2.111	0.759***			-1.669	1.274	-2.028	1.046*			-1.951	0.940**
Danish citizen	-0.036	0.126			-0.549	0.289*	0.161	0.240	0.057	0.190	-0.007	0.149
Iran	0.007	0.206									-0.259	0.283
Pakistan	-0.186	0.223									-0.324	0.281
Turkey	-0.494	0.217**									-0.596	0.287**
Interviewer women	0.122	0.189	0.533	0.211**	-0.310	0.301	0.260	0.351	0.064	0.260	0.128	0.234
Interviewer age: 30-39	-0.082	0.352	0.737	0.852	0.030	0.518	-0.170	0.575	-0.952	0.495*	-0.029	0.367
Interviewer age: 60+	-0.002	0.217	0.052	0.276	0.454	0.357	-0.413	0.398	-0.323	0.295	-0.058	0.263
Interviewer seniority 1 year	0.102	0.289	-0.536	0.543	0.025	0.548	0.402	0.563	0.138	0.366	0.222	0.296
Interviewer seniority 6+ years	-0.036	0.198	0.011	0.218	-0.024	0.309	-0.263	0.331	0.005	0.273	-0.022	0.237
Number of interviews per interviewer	0.002	0.001	0.002	0.001*	0.003	0.002**	0.001	0.002	0.001	0.001	0.003	0.001**
Constant	0.791	0.489	-0.085	0.716	1.602	1.063	1.804	1.170	0.019	0.823	0.437	0.627
Number of level 1 units	3098		998		744		568		788		1616	
Number of level 2 units	53		43		49		44		51		41	
Log L	-1495		-380		-319		-319		-455		-906	

* significant at 10%, ** significant at 5%, *** significant at 1%

Table 9: Probability of contact, cooperation and response, pooled sample

	Contact		Cooperation		Response	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Women	0.283	0.116**	-0.001	0.098	0.102	0.080
Age group: 18-29	0.507	0.140***	0.188	0.125	0.274	0.099***
Age group: 40-45	0.218	0.142	0.080	0.119	0.099	0.098
Single	-0.583	0.146***	-0.064	0.129	-0.331	0.103***
Children	0.337	0.155**	0.122	0.134	0.230	0.108**
Copenhagen	-1.375	0.204***	-0.591	0.146***	-0.873	0.123***
Rural area	0.232	0.325	-0.043	0.197	0.018	0.170
Danish education (in years)	0.049	0.032	0.071	0.026**	0.061	0.021***
No Danish education	0.520	0.381	0.456	0.315	0.374	0.258
Non-employed	-0.184	0.118	-0.186	0.103***	-0.278	0.083***
Employment unknown	-0.121	0.360	0.919	0.498	0.261	0.282
Years since migration	0.005	0.008	0.001	0.008	0.006	0.006
Years since migration unknown	0.426	0.650	-2.111	0.759***	-1.505	0.563***
Danish citizen	0.266	0.139*	-0.036	0.126	0.232	0.101**
Iran	-0.643	0.243***	0.007	0.206	-0.530	0.162***
Pakistan	-1.139	0.250***	-0.186	0.223	-0.877	0.174***
Turkey	-0.361	0.270	-0.494	0.217**	-0.818	0.176***
Interviewer women	-0.048	0.377	0.122	0.189	0.081	0.231
Interviewer age: 30-39	0.703	0.694	-0.082	0.352	0.241	0.435
Interviewer age: 60+	0.741	0.401*	-0.002	0.217	0.303	0.255
Interviewer seniority 1 year	-0.881	0.505*	0.102	0.289	-0.202	0.333
Interviewer seniority 6+ years	-0.234	0.394	-0.036	0.198	-0.043	0.243
Number of interviews per interviewer	-0.001	0.002	0.002	0.001	0.001	0.002
Constant	2.733	0.731***	0.791	0.489	0.610	0.469
Number of level 1 units	3836		3098		3836	
Number of level 2 units	55		53		55	
Log L	-1218		-1495		-2172	

* significant at 10%, ** significant at 5%, *** significant at 1%

highest number of interviews to the interviewers with highest rate of previous success.

Response

Overall response is a weighted average of contact and cooperation. If the bias in contact and cooperation point in different directions, the result may be that the overall response is unbiased. To investigate this possibility, we estimate the probability of response (see column 3 in table 9, where for comparison we also reshow the results for the contact and cooperation models in the first and second columns - the same results as in the first columns of tables 7 and 8). The estimation results for overall response for the four country subgroups and the Copenhagen subgroup appear in appendix table 2.

Table 9 shows that nearly all sample person characteristics that significantly affect either contact or cooperation also significantly affect overall response. Living in Copenhagen influences the probability of contact, cooperation and overall response. Being young, living in couples, having children and having been Danish citizenship influence the probability of contact and the overall probability of response, but not the probability of cooperation. Years since migration and socioeconomic status - measured by education and employment variables - influence the probability of cooperation and overall response, but not the probability of contact. Thus, an effort to reduce the bias in overall response should focus on the contact phase as well as on the cooperation phase.

The bias in overall response we can summarise as follows: being young, living in couples and having children, a high number of years of education, being employed and a Danish citizenship increase the probability of response, while living in Copenhagen has a negative influence on the probability of response. Furthermore, the country dummies for Iran, Turkey and Pakistan are also significant - immigrants from the three countries experience lower response rates than Danes, including after controlling for all other variables. However, the interviewer variables have no affect on the probability of response. Although we expect interviewer characteristics to be important, the interviewer-specific variables that we are able to include in the analysis do not capture this effect.

Conclusion

In this survey, both contact rates and cooperation rates are lower for immigrants than for Danes, leading to a significantly lower overall response rate. Furthermore, we find important differences between groups - the immigrants from Pakistan are especially difficult to contact, while refusals are particularly high among those from Turkey. In addition, language is extremely important as a very large share of particularly women from Turkey and Pakistan could not be interviewed in Danish.

To learn more about the causes of non-response, we analyse the determinants of the probability of contact and the probability of cooperation, respectively; and to learn more about the bias resulting from non-response, we analyse the determinants of overall non-response. We find that the characteristics of the sample persons are important for both the contact rate and the cooperation rate, and that different factors highly affect contact and cooperation. While gender, age and family structure significantly affect contact, education and employment - in accordance with the isolation hypothesis - affect cooperation. Living in Copenhagen has a negative effect on both contact and cooperation. The sample person characteristics that significantly affect either contact or cooperation also significantly affect overall response. Furthermore, we find that the lower probability of response among immigrants when compared to Danes persists after controlling for all the other variables. Thus, the results clearly show bias in the overall response rate with respect to characteristics of the sample persons, along with a 'country-of-origin' factor beyond the factors that we can include in the model.

Finally, another result is the lacking impact of the interviewer-specific variables. Unquestionably, the individual interviewer is very important for the data collecting process, but this impact cannot be measured by the type of very aggregate information that is available about the interviewers in this survey.

The analysis clearly points to the need of tailoring surveys directed to immigrant groups to avoid response bias, and that an effort to reduce the bias in overall response should focus on the contact phase as well as the persuasion phase effect. It is also important to handle the language problems - as described above is a high number of the immigrants - especially the women from Turkey and Pakistan - interviewed by a bilingual interviewer.

References

Abraham, K. G., Maitland, A., & Bianchi, S. M. (2006). Nonresponse in the American time use survey. *Public Opinion Quarterly*, *70*, 676-703.

American Association for Public Opinion Research. (2006). *Standard definitions: final dispositions of case codes and outcome rates for surveys* (4th ed.). Lenexa, Kansas: AAPOR.

Campanelli, P., & O'Muircheartaigh, C. (1999). Interviewers, interviewer continuity and panel survey nonresponse. *Quality and Quantity*, *33*, 59-76.

Chiswick, B. R., Cohen, Y., & Zach, T. (1997). The labor market status of immigrants: Effects of the unemployment rate at arrival and duration of residence. *Industrial Relations Review*, *50*, 289-303.

Dahl, K. M., & Jakobsen, V. (2005). *Køn, etnicitet og barrierer for integration*. Copenhagen: The Danish National Centre for Social Research.

Dale, T., & Haraldsen, G. (2000). *Turning nonresponse into response in surveys among immigrants.* (presented at the 2000 Workshop for Household Survey Nonresponse in Budapest)

DeMaio, T. J. (1980). Refusals: Who, Where and Why. *The Public Opinion Quarterly*, *44*, 223-233.

Ekberg, J. (1994). Economic progress of immigrants in Sweden from 1970 to 1990: a longitudinal study. *Scandinavian Journal of Social Welfare*, *3*, 148-157.

Feskens, R., Hox, J., Lensvelt-Mulders, G., & Schmeets, H. (2006). Collecting data among ethnic minorities in an international perspective. *Fields Methods*, *18*, 284-304.

Feskens, R., Hox, J., Lensvelt-Mulders, G., & Schmeets, H. (2007). Nonresponse Among Ethnic Minorities: A Multivariate Analysis. *Journal of Official Statistics*, *23*, 387-408.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, *70*, 646-675.

Groves, R. M., & Couper, M. P. (1998). *Nonresponse in household interview surveys*. New York: John Wiley and Sons Inc.

Groves, R. M., & Heeringa, S. (2006). Responsive design for household surveys: tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society A*, *169*, 439-457.

Husted, L., Nielsen, H. S., Rosholm, M., & Smith, N. (2001). Employment and Wage assimilation of male first generation immigrants in Denmark. *International Journal of Manpower*, *22*, 39-68.

Jakobsen, V. (2004). *Young immigrants from the former Yugoslavia, Turkey and Pakistan: Educational attainment, wages and employment*. Aarhus: Department of Economics, Aarhus School of Business. (PhD thesis)

Mørkeberg, H. (2000). *Indvandrernes uddannelse*. Copenhagen: Statistics Denmark.

Nielsen, N. K., & Pedersen, S. (2000). Databeskrivelse. In G. V. Mogensen & P. C. Matthiesen (Eds.), *Integration i danmark omkring årtusindeskiftet.* Aarhus: University of Aarhus Publisher. (in Danish)

Pedersen, L. (1991). *Indvandrere og efterkommere i Danmark*. Copenhagen: Statistics Denmark. (in Danish)

Pickery, J., & Loosveldt, G. (2002). A Multilevel Multinomial Analysis of Interviewer Effects on various Components of Unit Nonresponse. *Quality and Quantity*, *36*, 427-437.

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2005). Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. *Journal of Econometrics*, *128*, 301-323.

Singer, E. (2006). Introduction. Nonresponse bias in household surveys. *Public Opinion Quarterly*, *70*, 637-645.

Statistics Denmark. (2008). *www.statistikbanken.dk*.

Stoop, I. A. L. (2004). Surveying nonrespondents. *Field Methods*, *16*, 24-54.

Tranæs, T., & Zimmermann, K. F. (2004). *Migrants, work and the welfare state*. Odense: University Press of Southern Denmark.

Turner, H. A. (1999). Participation bias in AIDS-Related telephone surveys: Result from the National AIDS Behavioural Survey (NABS) Non-response study. *The Journal of Sex Research*, *36*, 52-58.

van den Brakel, J. A., Vis-Visschers, R., & Schmeets, J. J. G. (2006). An experiment with data collection modes and incentives in the Dutch family and Fertility Survey for Young Moroccans and Turks. *Fields Methods*, *18*, 321-334.

Appendix

Table 1: Response and non-response in the survey, percentages

	Denmark		Iran		Pakistan		Turkey	
	Men	Women	Men	Women	Men	Women	Men	Women
Response:								
I: Complete interviws	*75.75*	*82.03*	*61.78*	*63.08*	*40.00*	*41.68*	*54.07*	*53.83*
P: Partial Interviews	*1.12*	*0.00*	*0.87*	*1.79*	*1.37*	*0.63*	*0.97*	*1.45*
Total interviews	76.87	82.03	62.65	64.87	41.37	42.32	55.04	55.28
Non-response:								
f. refusal lack of time	6.90	6.76	7.16	5.38	8.04	4.21	12.60	10.77
g. refusal suspicious	0.75	0.36	1.40	1.28	1.18	1.89	1.94	2.07
h. refusal by parent	0.00	0.00	0.00	0.00	0.20	0.42	0.00	0.41
i. refusal by spouse	0.37	0.36	0.00	1.28	1.37	4.21	0.97	4.97
j. refusal due to the gender of the interviewer	0.00	0.00	0.00	0.26	0.00	0.42	0.00	0.21
k. refusal other reasons	5.60	5.34	4.89	6.15	6.08	5.89	7.36	7.04
R: Total Refusals	*13.62*	*12.81*	*13.44*	*14.36*	*16.86*	*17.05*	*22.87*	*25.47*
c. moved	0.56	0.71	0.52	1.03	3.14	1.89	1.74	0.41
d. not met	7.46	3.38	15.36	12.56	28.43	27.58	13.76	11.80
e. fictive address	0.19	0.00	0.87	0.77	1.18	0.21	0.58	0.41
NC: Non-contact	*8.21*	*4.09*	*16.75*	*14.36*	*32.75*	*29.68*	*16.09*	*12.63*
m. illness	0.00	0.00	0.35	1.03	0.00	0.84	0.19	0.21
n. hospitalised/away from home	0.37	0.53	0.70	1.28	1.18	2.74	0.58	0.21
o. handicapped	0.56	0.18	1.05	0.77	0.39	0.00	0.39	0.21
p. moved out of the country	0.37	0.36	2.97	1.28	4.12	2.32	1.74	0.41
q. dead	0.00	0.00	0.17	0.00	0.20	0.00	0.19	0.00
l. language problems	0.00	0.00	1.92	2.05	3.14	5.05	2.91	5.59
0: Other reasons	*1.31*	*1.07*	*7.16*	*6.41*	*9.02*	*10.95*	*6.01*	*6.63*
Total	100	100	100	100	100	100	100	100
N	536	562	573	390	510	475	516	483

Table 2: Estimation of response rate

	Pooled sample		Denmark		Iran		Pakistan		Turkey		Copenhagen	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Women	0,102	0,080	0,193	0,172	0.145	0.175	0.150	0.169	-0.005	0.147	0,061	0,100
Age group: 18-29	0.274	0.099***	0.590	0.221***	-0.074	0.232	0.406	0.202**	0.303	0.187	0.261	0.123**
Age group: 40-45	0.099	0.098	0.036	0.218	0.061	0.215	0.331	0.205	-0.145	0.186	0.147	0.123
Single	-0.331	0.103***	-0.529	0.220**	-0.210	0.189	-0.265	0.231	-0.357	0.218	-0.192	0.132
Children	0.230	0.108**	0.208	0.237	0.447	0.210**	0.264	0.234	0.090	0.225	0.278	0.136**
Copenhagen	-0.873	0.123***	-0.659	0.242***	-0.740	0.225***	-1.438	0.352***	-0.778	0.193***		
Rural area	0.018	0.170	-0.020	0.231	0.211	0.382			0.336	0.643		
Danish education (in years)	0.061	0.021***	0.167	0.040***	-0.018	0.047	-0.006	0.054	0.050	0.046	0.033	0.027
No Danish education	0.374	0.258			-0.472	0.633	-0.306	0.624	0.174	0.500	-0.043	0.326
Non-employed	-0.278	0.083***	-0.277	0.207	-0.418	0.180**	-0.312	0.171*	-0.154	0.147	-0.181	0.105*
Employment unknown	0.261	0.282			0.073	0.527	-0.151	0.427	0.398	0.568	0.216	0.338
Years since migration	0.006	0.006			0.025	0.015	-0.002	0.010	0.013	0.010	0.006	0.007
Years since migration unknown	-1.505	0.563***			-0.784	0.957	-1.890	0.857**			-1.285	0.744*
Danish citizen	0.232	0.101**			-0.106	0.214	0.551	0.187***	0.280	0.161*	0.307	0.120**
Iran	-0.530	0.162***									-0.770	0.226***
Pakistan	-0.877	0.174***									-1.020	0.221***
Turkey	-0.818	0.176***									-0.943	0.233***
Interviewer women	0.081	0.231	0.327	0.245	-0.301	0.347	0.067	0.431	0.116	0.245	-0.071	0.338
Interviewer age: 30-39	0.241	0.435	0.253	0.645	0.325	0.640	0.059	0.738	-0.540	0.481	0.005	0.564
Interviewer age: 60+	0.303	0.255	0.627	0.288**	0.549	0.388	-0.458	0.460	-0.106	0.278	-0.239	0.377
Interviewer seniority 1 year	-0.202	0.333	-0.559	0.508	-1.134	0.550**	-0.142	0.582	-0.272	0.348	-0.183	0.424
Interviewer seniority 6+ years	-0.043	0.243	-0.065	0.250	-0.102	0.361	0.137	0.426	-0.092	0.262	0.070	0.351
Number of interviews per interviewer	0.001	0.002	0.003	0.001**	0.001	0.002	-0.002	0.002	0.000	0.001	-0.001	0.002
Constant	0.610	0.469	-1.374	0.683**	1.130	0.888	1.483	1.012	0.246	0.697	0.551	0.664
Number of level 1 units	3836		1065		929		867		975		2222	
Number of level 2 units	55		47		49		44		52		43	
Log L	-2172		-474		-515		-544		-624		-1376	

* significant at 10%, ** significant at 5%, *** significant at 1%