Catalogue no. 11-522-XIE

Statistics Canada International Symposium Series - Proceedings

Symposium 2004: Innovative Methods for Surveying Difficult-to-reach Populations

2004

Statistics Canada Statistique Canada

Canada

Proceedings of Statistics Canada Symposium 2004
Innovative Methods for Surveying Difficult-to-reach Populations

A MULTIVARIATE ANALYSIS OF NONRESPONSE AMONG ETHNIC MINORITIES

Remco Feskens, Joop Hox, Gerty Lensvelt-Mulders and Hans Schmeets[1]

ABSTRACT

This article examines nonresponse among ethnic minorities in the Netherlands. We have constructed a structural equation model with different response outcomes in the software package Mplus. We controlled the effect of ethnicity on response outcomes for various socioeconomic and socio demographic variables. The effect of response was almost entirely mediated by degree of urbanisation. We also performed multiple group analyses to examine differences between ethnic groups in response outcome predictors. Here again, we found that urbanisation has a negative effect on response probabilities in all ethnic groups and in particular on contact probabilities. This negative effect, however, is somewhat larger among people with a non-western background.

KEYWORDS: Ethnic Minorities; Nonresponse; Structural Equation Models.

1. INTRODUCTION

Nonresponse rates have increased in almost all western countries in recent years (De Heer & De Leeuw, 2002). Nonresponse is a problem for several reasons. First of all, nonresponse reduces the number of respondents and therefore reduces the precision of estimates. Second, when nonresponse is selective, the survey estimates may be biased and not accurately reflect the "true values" of the target population (Thornberry & Massey, 1988). Nonresponse is selective when nonrespondents are systematically different from respondents in terms which matter to the survey objectives (Groves & Couper, 1998). Third, nonresponse can increase the costs of survey research; for reaching the desired sample size more efforts are needed (Hox & De Leeuw, 1998). The most worrying consequence of nonresponse is bias in point estimators (Groves, 1989). The occurrence of selective nonresponse is more likely when specific groups in society have below average response rates. This makes it more likely that the nonrespondents differ systematically from the respondents, because the nonresponse is not random. So in order to speculate about nonresponse bias, it is useful to look at response rates among various subgroups (Thornberry & Massey, 1988).

For various reasons, including above-average nonresponse, Statistics Netherlands has difficulties in surveying five subgroups, so called difficult-to-reach-populations (Snijkers, 2003). One group that is difficult to reach is the immigrant population (or ethnic minorities), which makes up about 20% of Dutch society (Statistics Netherlands). In the Netherlands the following definition of ethnic minorities is used: "Every person residing in the Netherlands of whom one or both parents is born abroad" (Reep, 2003). Usually, a further distinction is made between people with a western foreign background (one or both parents born in another European country, North America, Australia, Japan or Indonesia) and people with a non-western foreign background (mainly Turks, Moroccans, Surinamese and Antilleans)[2]. Not at least due to this broad definition, almost one in five residents in the Netherlands belong to an ethnic minority group. About half of them are non-western foreigners (10% of the total population), almost 9% of the Dutch residents are western foreigners. Response problems among ethnic minorities are not restricted to the Netherlands. Ethnic minorities have lower response rates in almost all western countries (Feskens, 2005).

R.C.W. Feskens, J. Hox, G. Lensvelt-Mulders (Utrecht University) and J.J.G. Schmeets (Statistics Netherlands), Heidelberglaan 1, 3584 CS Utrecht, The Netherlands (r.c.w.feskens@fss.uu.nl). The views expressed in this paper are those of the authors and do not necessarily reflect the policies of Statistics Netherlands.

[2] For reasons of simplicity we use 'western foreigners' and 'non-western foreigners' in this article.

In particular non-western foreigners tend to have lower response rates than the native population, but they live predominantly in urban areas, they are more often unemployed, and have lower education levels than the native population. These characteristics are negatively correlated with response (Groves & Couper, 1998; Stoop, 2004). A previous study using a loglinear regression-model demonstrated that the high nonresponse rates among ethnic minorities should be attributed to socioeconomic status and urbanisation (Schmeets & Michiels, 2003). In this study we will examine whether this conclusion also holds when using structural equation techniques. In addition, the model will be elaborated by the inclusion of more variables as well as interaction effects. Finally, the effects on noncontacts and refusals will be illustrated.

These considerations resulted into three research questions:

1) How problematic is nonresponse, divided into nonrespondents, noncontacts and refusals among ethnic minorities in the Netherlands?
2) What is the effect of ethnicity on different response outcomes if controlled for other socioeconomic and socio demographic variables?
3) Which variables or predispositions are related to the phenomena of response, contacts and refusals among ethnic groups?

2. DATA

We executed our analyses on the survey files of the Permanent Survey on Living Conditions (POLS). This survey is conducted by Statistics Netherlands. Mostly, about 40,000 interviews are conducted a year. The observation units are individuals. The sample frame is the population register from all Dutch municipal basic administrations. POLS is a cluster sample. First communities are drawn and then people. Big cities are included automatically (Schouten, 2003). Participation is voluntary. Each month a sample of about 3,500 people is drawn. In 1998, there was a two month fieldwork period for the twelve consecutive samples. In the first month Statistics Netherlands collected data with a CAPI mode, nonrespondents with a known telephone line were re-approached with a CATI technique. Nonrespondents without a known telephone line, and sampled units who could not cooperate due to illness, were re-approached with CAPI (Schouten, 2003).

The POLS survey was supplemented with information from the population registration and information about employment and social benefits (Schmeets & Michiels, 2003; Schouten, 2003). By linking these administrative records, we have socio demographic and socioeconomic information available on the nonrespondents at the postal code level. In the POLS 1998 survey 39,431 sampled units were drawn, the number of respondents according to AAPOR response definition three was 23,993 (61.4%) (AAPOR, 2000), which is not unusually low in the Netherlands (see for example De Heer, 1999).

The extra information provided by the links, made it possible to study the nonrespondents. However, there was no extra information available for all units sampled. Data was missing on 1,143 of the units sampled (2.9% of total sample). These were assumed to be missing at random and were deleted from the data file. Because these numbers are quite small, they can be dropped from the sample without significant loss of information. Since we analysed only sampled units aged 15-65, there were 28,542 sampled units left for analysis. Response in this subsample is somewhat lower: 60.4% or 17,123 sampled units responded. To avoid capitalization on chance, we randomly split this new file into an exploration and a validation file. The exploration file consisted of odd case numbers (14,101 cases), the validation file consisted of the even case numbers (14,271 cases).

3. METHODS

First of all, we looked at the bivariate relationships between ethnic groups and several response categories. This provided us with information for answering the first research question. For answering the second research question we made a structural equation model in the software package Mplus. Structural equation modeling allows us to combine latent variables and structural relations between them and other observed variables (Kline, 1998). Mplus enabled us to make structural equation model with a dichotomous dependent variable. As Groves and Couper (1998)

denote: “dissecting the nonresponse phenomenon into one of noncontacts, refusals, and other causes, sensitizes us to considering alternative causes of each outcome.” Therefore we analysed not only the response outcome, but also noncontacts and refusals, resulting in more information. Other causes of nonresponse, had only a minor impact on the response rate, as will be shown in table 1.

We also wanted to know in which variables the two groups (native/western and non-western foreigners) differed in the response phenomenon. In doing so, we evaluated an interactive effect of ethnicity. We did so by making a multiple group analysis where we compared the path coefficients of the groups. Here again we dissected different response outcomes, enabling us to examine which response predispositions exist among subgroups. Since the sample size was large, the assessment of model fits was based on two goodness-of-fit indices that are less sensitive to sample size: Bentler comparative fit index (CFI), and the root mean square error of approximation (RMSEA).

4. RESULTS

4.1 How problematic is nonresponse, divided into nonrespondents, noncontacts and refusals among ethnic minorities in the Netherlands?

The overall response rates for sampled units aged 15-65 in the POLS 1998 survey was 60.4%. Further dissecting the nonresponse in alternative outcomes gives the following picture for the three ethnic groups:

Table 1: Response outcomes among ethnic groups in POLS 1998 in percentages

	Native population	Western foreigners	Non-western foreigners
Response	62.0	57.2	39.7
Noncontacts	12.0	15.9	26.4
Refusals	24.7	23.0	20.1
Language problems	0.0	2.5	13.0
Other	1.3	1.8	1.0

Figures may not add to 100% due to rounding.

The distribution of nonresponse outcomes among western foreigners is more or less similar to that of the native population. The relative numbers of noncontacted sampled units and nonresponse due to difficulties with the survey language are higher among non-western foreigners than among the native population, a result also found in other surveys (Feskens, 2005). These cause the lower response rates among the group of non-western foreigners.

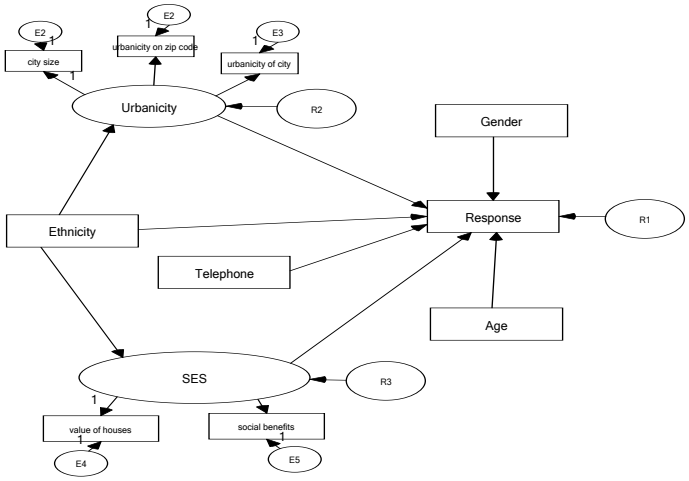
We also considered bivariate relations between ethnicity and various socio demographic and socioeconomic variables. Here again, the characteristics of western foreigners we found were very similar to the native population, whereas those of non-western foreigners differed substantially. Therefore we decided to concentrate our analysis on non-western foreigners.

4.2 What is the effect of ethnicity on response outcomes when controlled for other socioeconomic and socio demographic variables?

4.2.1 Structural Equation Model

We wanted to study if nonresponse is still affected by ethnicity if controlled for socio demographic and socioeconomic variables. First we considered the bivariate relationships between each of the socio demographic and socioeconomic variables and response for all ethnic groups separately. Then we looked at the bivariate relationships between the ethnic groups and the socioeconomic and socio demographic variables. With this information and theoretical considerations we constructed a structural equation model where we could control the effect of ethnicity on different response outcomes for other variables, see figure 1.

Figure I: Structural equation model



The measurement part of the model consisted of the latent variables urbanicity and social economic status (SES). The latent variable urbanicity has been measured by the observed variables urbanisation on postal code level, degree of urbanisation of the city and city size. These variables are sufficiently correlated with each other, but are not correlated to the extent that they measure the same. The latent variable SES was measured by house values and indicators for receiving social benefits. For identification purposes we fixed the factor loadings of the indicators urbanisation of the city and house value to one.

Probit regressions are estimated for the categorical factor indicators, and simple linear regressions are estimated for the continuous factor indicators (Muthén, 1998-2004). We constructed the two latent variables because using the extra information of all indicators made it possible to identify urbanicity and SES more precisely than would be otherwise. Moreover, some indicators are subject to subjective classification. Using latent variables with more indicators reduces this form of measurement error. The relations between the other observed variables and latent variables on response formed the structural part of the model. These observed variables are ethnicity, an indicator for having a known telephone land line (telephone), gender and age. These observed variables were regressed on the binary outcome variable response.

We wanted to examine if the relation between ethnicity and response is mediated by SES and urbanicity. Therefore we also regressed ethnicity on these latent variables. Relations are assumed to be unidirectional, the latent variables SES and urbanicity and also SES and telephone are assumed to covary. The regression coefficients are estimated with unweighted least square estimator and are interpreted as probit regression coefficients. Estimates between brackets are standardized coefficients using the variances of the continuous latent variables as well as the variances of the background and outcome variables (Muthén & Muthén, 1998-2004). The sign of the regression coefficients shows to what degree this characteristic the probability in the direction nonresponse (-) or response (without sign) changes. We assumed that the measurement errors are uncorrelated.

4.2.2 Response

Results of the structural equation model on dependent variable response are presented in table 2.

Table 2: Structural equation model with dependent variable response

	Explained variables					
	Estimates on response	Standard errors	Estimates on urbanicity	Standard errors	Estimates on SES	Standard errors
Predictors						
Size of city			1.000 (0.874)fixed	0.000		
Urbanicity of city			0.825 (0.914)**	0.011		
Urbanicity on zip code			0.870 (0.887)**	0.015		
Value of house					1.000 (0.648)fixed	0.000
Social benefits					-0.370 (-0.491)**	0.020
Gender	0.075 (0.037)**	0.021				
Age	-0.002 (-0.033)**	0.001				
Urbanicity	-0.097 (-0.142)**	0.011				
SES	0.047 (0.063)**	0.018				
Telephone	0.341 (0.139)**	0.031				
Ethnicity	-0.285 (-0.072)**	0.048	1.453 (0.250)**	0.054	-1.805 (-0.340)**	0.073

Note: dependent variable coded 1 = response, 0 = nonresponse
**p<0.01
Chi square= 125.432 (df= 4); RMSEA= 0.046; CFI = 0.972, R-square= 0.072
Validation file: Chi square= 91.611 (df= 3); RMSEA= 0.045; CFI = 0.980, R-square=0.082

The relation between ethnicity and response is almost entirely mediated by urbanicity and the social economic status of sampled units. A large amount of the negative effect of ethnicity on response is mediated by urbanicity. The standardized probit regression of urbanicity on response is –0.142, the standardized effect of ethnicity on urbanicity is 0.250. Especially urbanicity and telephone have a large impact on the response probability. The latent variable SES does not much affect the probability of responding.

Ethnic minorities have lower response rates compared to the native population. But ethnic minorities also disproportionally live in urban areas. These results suggest that in particular this urbanicity effect ‘cause’ lower response rates among ethnic minorities, and not ethnicity itself. If controlled for other variables, ethnicity has only a small impact on the response probability, a fairly small standardized coefficient of ethnicity on response of -0.072 remains in this multivariate environment.

The socioeconomic status of sampled units does not much affect response probabilities, suggesting that the negative effect of ethnicity on response is mainly mediated by socio demographic instead of socioeconomic characteristics. This also holds for the native population: urban residents, regardless of their ethnicity and socioeconomic status, have smaller response probabilities, compared to non-urban residents. The availability of a land line proves to be a strong indicator for response. This holds not only for the second part of the field work period – where the re-approaching strategies differed between a CATI mode for nonrespondents with a known phone line, or otherwise CAPI – but also for the first part of the field work period, where no telephone calls are being made. Age and gender only have a small influence on response, older sampled units have slightly smaller response probabilities.

The effect of ethnicity decreases substantially if controlled for other variables. Not only the path coefficient decreased, if the indicator for ethnicity is included in the analysis, the pseudo R square only increases with 1.0% (from 6.2% to 7.2%). Small path coefficients and a low pseudo R square suggest that predicting response is still fairly difficult, also with the availability of rich background information. This suggests that response bias is not as high as some bivariate relationship seems to indicate. However, some selective nonresponse may exist in urban areas. Urban residents, regardless to whether ethnic group they belong, are somewhat underrepresented in this survey.

4.2.3 Contact

Results of the structural equation model on dependent variable contact are presented in table 3.

Table 3: Structural equation model with dependent variable contact

	Explained variables					
	Estimates on contact	Standard errors	Estimates on urbanicity	Standard errors	Estimates on SES	Standard errors
Predictors						
Size of city			1.000 (0.874)fixed	0.000		
Urbanicity of city			0.825 (0.914)**	0.011		
Urbanicity on zip			0.870 (0.887)**	0.015		
Value of house					1.000 (0.649)fixed	0.000
Social benefits					-0.368 (-0.489)**	0.020
Gender	0.060 (0.030)**	0.027				
Age	0.004 (0.056)**	0.001				
Urbanicity	-0.183 (-0.267)**	0.012				
SES	-0.014 (-0.019)ns	0.020				
Telephone	0.386 (0.156)**	0.035				
Ethnicity	-0.215 (-0.054)**	0.052	1.453 (0.250)**	0.054	-1.806 (-0.340)**	0.073

Note: dependent variable coded 1 = contact, 0 = no contact
**p<0.01, ns = non significant
Chi square= 123.290 (df= 4); RMSEA= 0.046; CFI = 0.972. R-square = 0.104
Validation file: Chi square= 90.715 (df= 3); RMSEA= 0.045; CFI = 0.980, R-square = 0.113

In the contact process the roll of urbanicity is even more pronounced. The standardized coefficient of urbanicity on the contact probability is -0.267. The effect of ethnicity on contact has decreased compared to this effect on response rate, indicating that the effect of ethnicity is even stronger mediated by urbanicity. Among others, Groves & Couper (1998) noted that especially in urban areas it is more difficult to establish contact with sampled units compared to non-urban sampled units.

These results show that this also holds for ethnic minorities. Nonresponse among ethnic minorities is heavily determined by low contact rates. These low contact rates are not unique for ethnic minorities, it is largely mediated by urbanicity. This also provides an explanation why ethnic minorities have lower response rates. Contact difficulties mainly concentrated in urban areas. Nonrespondents with a known landline were re-approached in the second month with a CATI mode. Other nonrespondents were re-approached with again a CAPI mode. However, due to capacities problems in the interviews corps, not all nonrespondents without a known land line were at all re-approached or with less contact attempts compared to nonrespondents in the CATI mode. So, not surprisingly, the regression coefficient from telephone on the contact probability is high. Nonetheless, this positive effect of having a known land phone on the contact probability was also found in the first month of the interview process, where no telephone calls were made. Furthermore, the results show that women and elderly are somewhat, albeit with small probabilities, easier to contact.

4.2.4 Refusals

Nonrespondents who are not contacted do not have a possibility to refuse a request for survey participation (Hox & De Leeuw, 1998). Therefore, we executed our analysis for those sampled units who refused, only on those sampled units who were contacted in the first place. 12,202 Sampled units remained for the exploration file, the validation file contained 12,366 sampled units. Results are presented in table 4.

Table 4: Structural equation model with dependent variable refusals

	Explained variables					
	Estimates on refusal	Standard errors	Estimates on urbanicity	Standard errors	Estimates on SES	Standard errors
Predictors						
Size of city			1.000 (0.860)fixed	0.000		
Urbanicity of city			0.856 (0.911)**	0.012		
Urbanicity on zip			0.896 (0.882)**	0.017		
Value of house					1.000 (0.623)fixed	0.000
Social Benefits					-0.400 (-0.501)**	0.025
Gender	0.070 (0.035)ns	0.024				
Age	-0.003 (-0.046)**	0.001				
Urbanicity	-0.041 (-0.057)**	0.013				
SES	0.042 (0.054)*	0.021				
Telephone.	0.257 (0.102)**	0.034				
Ethnicity	0.188 (0.045)**	0.058	1.324 (0.226) **	0.057	-1.715 (-0.321)**	0.082

Note: dependent variable coded 1 = cooperation, 0 = refusal
*p<0.05, **p<0.01, ns = non significant
Chi square= 177.247 (df= 6); RMSEA= 0.048; CFI = 0.949, R-square= 0.025
Validation file: Chi square= 142.455 (df= 5); RMSEA= 0.047; CFI = 0.956., R square = 0.033

Ethnicity has, perhaps somewhat surprisingly, a positive effect on the cooperation probability. Non-western foreigners tend to refuse a request for survey participation less often than other sampled units. When interpreting these results, precaution should be taken, because of the low contact rates and the high nonresponse due to language problems among ethnic minorities. As already mentioned, sampled units who are not contacted, do not have a chance to refuse a request for a survey (Hox & De Leeuw, 1998). Above that, sampled units can use language problems with the survey language as a friendly way for refusing a request for survey participation. Also the effect of urbanicity, which was large in the response and contact process, is lower, suggesting that there exists in particular a contact- not participation- problem in urban areas. The effect of SES is very small. Again, the indicator 'having a known land phone' proves to be a strong predictor. Older sampled units have a somewhat higher probability of refusing a request for survey participation.

4.3 Which variables or predispositions are related to the phenomena of response, contact and refusals among ethnic groups?

We wanted to examine if predictors for different response outcomes have the same influence for ethnic groups. Therefore we performed a multiple group analysis with again two groups. The native population together with the western foreigners form the first group and are contrasted to the non-western foreigners. The results of the multiple group analysis with outcome variable response are presented in table 5.

Table 5: Multiple group analysis on response

	Explained variables			
	Native population & western foreigners		Non-western foreigners	
Predictors On Response	Estimates	Standard Errors	Estimates	Standard errors
Gender	0.063 (0.032)**	0.016	0.153 (0.076)**	0.058
Age**	-0.002 (-0.022)*	0.001	-0.008 (-0.092)*	0.002
Urbanicity**	-0.085 (-0.127)**	0.006	-0.180 (-0.244)**	0.032
SES	0.021 (0.036)**	0.007	0.016 (0.027)ns	0.032
Telephone*	0.318 (0.129)**	0.020	0.173 (0.080)*	0.074

* p value < 0.05,** p value < 0.01, ns = non significant
* after variable reflects significant difference between groups for this variables at p<0.05 level, ** p value < 0.01
Chi square= 14.512 (df=5) p value= 0.0126; RMSEA= 0.012, CFI=0.698; R-square1 = 0.047, R-square2= 0.097

We performed our analysis on 26,479 native and western foreigners (forming the first group) and 1,893 non-western foreigners (second group). These results show that the response probability of the native population and western foreigners is more influenced by having a known land phone compared to the non-western foreigners. In terms of an interaction effect: group membership does moderate the relationship between having a known land phone and

response. The negative effect of urbanicity on response probability is only slightly larger for the non-western foreigners. Age has a somewhat larger negative effect on the response rate among non-western foreigners compared to the first group. This reflects the higher nonresponse among this group due to language problems, almost entirely located among older non-western foreigners. The next table shows the results for the two groups on the dependent variable contact:

Table 6: Multiple group analysis on contact

	Explained variables			
	Native population & western foreigners		Non-western foreigners	
Predictors On contact	Estimates	Standard Errors	Estimates	Standard Errors
Gender	0.076 (0.038)**	0.020	0.152 (0.075)*	0.085
Age*	0.006 (0.082)**	0.001	-0.002 (-0.019)ns	0.003
Urbanicity**	-0.138 (-0.206)**	0.007	-0.247 (-0.336)**	0.036
SES	0.002 (0.004)ns	0.007	0.002 (0.004)ns	0.035
Telephone**	0.291 (0.118)**	0.022	0.198 (0.092)*	0.096

* p value < 0.05, ** p value < 0.01, ns = non significant
Chi square= 13.198 (df = 5) p value=0.0215; RMSEA=0.011, CFI=0.739; R-square1 = 0.075, R-square2= 0.143

Again, the parameter indicator for having a known land phone varies across groups. Contact probabilities among the first groups are heavily influenced by this predictor. Among non-western foreigners the negative effect of urbanicity is larger compared to the first group. Nonetheless, urbanicity has also a relatively large negative effect on the contact rate among the native population and western foreigners. Finally, table 7 shows the results of the multiple group analysis with outcome variable refusal.

Table 7: Multiple group analysis on refusals

	Explained variables			
	Native population & western foreigners		Non-western foreigners	
Predictors On refusals	Estimates	Standard Errors	Estimates	Standard Errors
Gender	0.044 (0.022)*	0.018	0.142 (0.071)*	0.072
Age*	-0.005 (-0.062)**	0.001	0.003 (0.043)ns	0.003
Urbanicity*	-0.004 (-0.059)**	0.007	-0.136 (-0.179)**	0.050
SES	0.015 (0.028)*	0.006	-0.022 (-0.039)ns	0.048
Telephone	0.264 (0.104)**	0.022	0.185 (0.088)ns	0.096

* p value < 0.05,** p value < 0.01, ns = non significant
Chi square= 2.804 (df = 2) p value= 0.2429; RMSEA= 0.006, CFI= 0.911; R-square1 = 0.023, R-square2= 0.047

For this multiple group analysis where we looked at the predictors for sampled units who refused a request for survey participation across groups, we again only included sampled units who were contacted. This resulted in 23,210 remaining sampled units for the first group, 1,359 non-western foreigners where contact has been established remained for analysis. Urbanicity has again a somewhat more negative effect for the non-western foreigners on the response outcome, in this case the refusal rate. Having a known land phone again proves to be a strong positive predictor among the first group and does not have much impact on the second group.

5. CONCLUSIONS

Ethnic minorities are a growing part of western societies and are therefore becoming more and more relevant for policy makers. According to the Statistics Netherlands definition almost 20% of the Dutch population are “allochtoneous”. Predictions for the ethnic minorities part of the Dutch population in 2050 top to 35% (De Jong & Hilderink, 2004). Although the need for data about ethnic minorities is increasing, doubts exist about the quality of the data about ethnic minorities, because ethnic minorities have higher nonresponse rates. Nonresponse itself does not automatically imply bias in point estimates. However, nonresponse rates can serve as an indicator for potential bias problems. Nevertheless, fighting nonresponse should concentrate on reducing nonresponse error. Simply trying to increase response rates can actually increase the survey error (Merkle & Edelman, 2002). Therefore it is

important knowing in which societal groups nonresponse is located, so that tailored strategies can be developed in order to reduce nonresponse in these very specific underrepresented subgroups.

The analyses in this article are based on the results of the survey of living conditions in the Netherlands in 1998. Bivariate tables of response and ethnicity show large differences in the response rates between ethnic groups. One of the interesting findings here are the high noncontact rates among ethnic minorities and more specifically, among non-western foreigners. This is not at least because of some fieldwork problems in this year, due to insufficient interviewer capacity. Since then a lot of things have changed at Statistics Netherlands, resulting in better response rates. Contact and persuasion strategies at Statistics Netherlands have been standardized in order to increase response rates (Snijkers & Kockelkoren, 2004). Remarkably, the cooperation rate among ethnic minorities is higher than for the native population. However, if sampled units are not contacted, it is of course impossible for them to refuse a request for a survey (Hox & De Leeuw, 1998). Nonetheless, an increase in the minimal number of contact attempts by earlier noncontact at Statistics Netherlands in March 2004 showed a substantial increase in the contact and response rate among non-western foreigners, but not in the refusal rate (Schmeets, 2005).

The results of the structural equation models show that the negative effect of ethnicity on response partially disappears if we controll for other variables. Especially urbanicity has a large impact on nonresponse. However, ethnicity still affects the response probability. Nevertheless, knowing that nonresponse among ethnic minorities is especially found in highly urban areas enables the researcher to concentrate its efforts on this very specific group. Of course, there exist a high correlation between urban areas and ethnic minorities, but for example in the Netherlands about 70% of the allochtoneous population (and 70% of the non-western foreigners) is *not* living in one of the four major cities. Urbanicity has been related to survey nonresponse for many years (Groves & Couper, 1998). Indeed, one of the main reasons for higher nonresponse rates in the Netherlands compared to most other countries could be the urban character of in particular the western part of the country. Response remains a process that is greatly influenced by chance. Some societal groups, like ethnic minorities, do have lower response rates. But, fortunately, also with the inclusion of many background information it is extremely difficult to predict whether a person will respond or not. Although rich administrative information is available, a low pseudo R square value (0.072) suggests that the nonresponse is not much systematic. Although regression coefficients and R square are small, one should concentrate in enlarging response in urban areas, where in particular a contact problem exists. Tailoring data collection strategies on higher contact rates will also lead to increased response rates among ethnic minorities.

The results of the multiple group analysis show that the ethnic groups are not homogeneous in their response processes. Age and urbanicity have a more negative impact on the response and contact probability of non-western foreigners compared to the combined native population and the western foreigners. On the other hand, the indicator ‘having a known land phone' proves to be a very strong indicator for responding among the native population and western foreigners. This negative effect is much smaller among non-western foreigners. Of course, the results of this study can only partially be generalized to other countries, because the results are based on a Dutch survey. Nevertheless, the outcomes of our analysis are suggesting that although ethnic minorities have lower response rates, one should concentrate in enlarging response in urban areas. Extra efforts should be made in increasing the contact rate in urban areas. A possible solution could be to approach sampled units with special tailored strategies, where for example the minimal number of contact attempts are higher than for the non-urban sampled units.

REFERENCES

The American Association for Public Opinion Research. (2000), *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*, Ann Arbor, Michigan: AAPOR.

Feskens, R.C.W. (2005), “Het Afbakenen en Benaderen van Allochtonen in internationaal Perspectief”, in Schmeets, H. and Bie, van der, R. (eds.) *Over het omgaan met Allochtonen in Enquêteonderzoek*, Heerlen: Centraal Bureau voor de Statistiek (in progress).

Groves, R.M. (1989), *Survey Errors and Survey Costs*, New York: Wiley.

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

De Heer, W. (1999), "De Heer, W. (1999), International Response Trends: Results of an International Survey", *Journal of Official Statistics*, 15, 129-142.

De Heer, W. and De Leeuw, E. (2002), "Trends in Household Survey Nonresponse: A Longitudinal and International Comparison", in Groves R. et al. (eds.) *Survey Nonresponse*, New York: John Wiley, pp. 41-54.

Hox, J.J. and De Leeuw, E. (1998), "Nonrespons in Surveys: Een Overzicht", *Kwantitatieve Methoden*, 19, pp. 31-53.

De Jong, A. and Hilderink, H. (2004), "Lange termijn allochtonenscenario's voor Nederland", *Bevolkingstrends*, 52 (1), Voorburg/Heerlen: Centraal Bureau voor de Statistiek, pp. 77-82.

Kline, R.B. (1998), *Principles and Practice of Structural Equation Modeling*, New York: Guilford Press.

Merkle, D.M. and Edelman, M. (2002), "Nonresponse in Exit Poll: A Comprehensive Analysis", in Groves R. et al. (eds.) *Survey Nonresponse*, New York: John Wiley, pp. 243-257.

Muthén, B.O. (1998-2004), *Mplus Technical Appendices*, Los Angeles, CA: Muthén & Muthén

Muthén, L.K. and Muthén, B.O. (1998-2004), *Mplus Users's Guide*, Third Edition. Los Angeles, CA: Muthén & Muthén.

Reep, C. (2003), "Moeilijk Waarneembare Groepen. Een inventarisatie", CBS-report H1568-03-SOO, Heerlen: Centraal Bureau voor de Statistiek.

Schmeets, H. and Michiels, J. (2003), "Het Effect van Non-respons onder Allochtonen", *Bevolkingstrends*, 51 (4), Voorburg/Heerlen: Centraal Bureau voor de Statistiek, pp. 52-57.

Schmeets, H., Reep, C. and Snijkers, G. (2003), "Afbakenen van moeilijk waarneembare bevolkingsgroepen", *Bevolkingstrends*, 51 (3), Voorburg/Heerlen: Centraal Bureau voor de Statistiek, pp. 62-68.

Schmeets, H. (2005), "De invloed van het aantal contactpogingen op de respons van allochtonen", in Schmeets, H. and Bie, van der, R. (eds.) *Over het omgaan met Allochtonen in Enquêteonderzoek*, Heerlen: Centraal Bureau voor de Statistiek. (in progress).

Schouten, B. (2003), "De verschillende Typen Non-respons in POLS 1998", CBS-report 2319-03-TMO. Heerlen: Centraal Bureau voor de Statistiek.

Snijkers, G. (2003), "Programmaplan Strategisch Onderzoek 'Non-respons en Moeilijk Waarneembare Groepen': Naar een optimale Meetstrategie", CBS-report H3318-03-SOO, Heerlen: Centraal Bureau voor de Statistiek.

Snijkers, G. and Kockelkoren. S. (2004), "De gestandaardiseerde Benaderingsstrategie 2003: Evaluatie en Aanbevelingen", CBS-report H415-04-SOO, Heerlen: Centraal Bureau voor de Statistiek.

Stoop, I.A.L. (2004), "Surveying Nonrespondents", *Field Methods*, Vol. 16, No.1, February 2004, pp. 23-54.

Thornberry, O.T. ,Jr., and. Massey, J.T. (1988), "Trends in United States Telephone Coverage across Time and Subgroups"; in Groves R. et al. (eds.) *Telephone Survey Methodology*, New York: John Wiley & Sons, pp: 25-49.

APPENDIX

Variables:

Gender; gender of sampled unit:

0 = man
1 = women

Age; age of sampled unit:

15-65 years

Size of city; size of community where sampled unit is registered:

1 = small, 8 = large

Urbanicity of community; urbanicity of community where sampled unit is registered:

1	non urban
2	little urban
3	slightly urban
4	urban
5	highly urban

Urbanicity on zip code level; urbanicity on zip code level of address where sampled unit is registered:

1	<500 addresses per square km
2	500 -< 1000 addresses per square km
3	1000-< 1500 addresses per square km
4	1500-< 2500 addresses per square km
5	>2500 addresses per square km

Indicator for known registered land phone (Telephone); Does sampled unit have a known registered land phone:

0 = no known registered land phone
1 = known registered land phone

Response; Did sampled unit respond (partially):

0 = nonresponse
1 = response

Value of houses; value of the house (in Dutch Guilders) where sampled unit is registered:

1	< 50 thousand
2	50 – 75 thousand
3	75 – 100 thousand
4	100 – 125 thousand
5	125 – 150 thousand
6	150 – 200 thousand
7	200 – 250 thousand
8	250 – 300 thousand
9	300 – 350 thousand
10	350 – 400 thousand
11	400 – 500 thousand
12	> 500 thousand

Indicator for having some form of social benefits (Indicator social benefits); Does sampled unit receive some form of social benefits:

0 = no
1 = yes

Ethnicity; Is sampled unit a non-western foreigner:

0 = no
1 = yes

Contact; Is contact established with sampled unit during field work period:

0 = no, no contact
1 = yes, contact

Refusals; Did sampled units refuses to participate in survey:

0 = sampled unit refused to deliver information asked for
1 = sampled unit cooperates