

## Chapter 20

### **The Validity of Income Measurements in Comparative Perspective**

#### *Non-Responses and Biases*

CHRISTIAN HOLST

#### **1. INTRODUCTION**

Over the last years market research has become significantly more international. Whereas in 1997 German market research companies with membership of ADM (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.) reported that less than one third (31%) of their annual turn-over had been achieved abroad, in 2000 this share had risen to almost half of their turnover (48%) (ADM 2001: 7; see also Chapter 4 in this volume). The globalisation of brands and products requires comparable information on the consumers' likes and dislikes, which spans more than one society or market, respectively. In order to supply this information, market research companies had to develop standardised international tools and instruments which at the same time yield valid, reliable, and comparable results. At this point, market research and academic social research face the same problems when moving in an international environment, and both rely on similar instruments to grasp the 'faits sociaux'.

Thus, both types of research use the same quality standards when it comes to their respective instruments: The instruments used must be at least valid – the extent to which the instruments actually measure what they are supposed to measure – and reliable – the extent to which repeated measurements of one object with the same instrument deliver the same values. However, when it comes to international social and/or market research, instruments must also be comparable, *i.e.* the extent to which an instrument remains equal in its contents, structure, mode of administration, *etc.* and delivers results which, given as a whole, can be interpreted meaningfully.

Thus, it is these three quality criteria concerning the measurement instruments with which we have to deal within international research, and only if all three criteria are met, may we speak of an instrument that fulfils the basic quality requirements of valid and meaningful international research.

In order to meet these objectives, standardisation of instruments has become an indispensable requirement for any international research. By standardisation we mean both, the definition of certain quality requirements which have to be met at any stage during the research process and by any partner involved in international market research, and, as well as, similar survey instruments and identical data collection procedures and analytical methods. Thus, standardisation of international research is a highly ambitious and demanding endeavour, which must be done with care. In order to be careful, one has to know where the traps are. My task here is not to offer a solution to all the problems. What I want to do is to show that even when using a highly standardised survey, such as the Eurobarometer, with a highly standardised instrument, such as the income question, significant problems still remain to be solved.

As mentioned before, if questions are to be compared, they also need to be valid. But how valid are these questions?

- Is a question which asks for something rather simple, such as income, still valid if in some cases half of the respondents refuse to answer this question? It might be valid if those who refuse were distributed in the same way as those who do not refuse to answer.
- Are the questions over the countries comparable? They might be comparable if all countries showed a similar pattern of how the non-responses<sup>1</sup> were distributed.
- Is there a solution to income measurement if both assumptions do not hold true? There might be a solution if it in some way offered a valid proxy instrument for what we wished to measure.

Therefore, with regard to the income variable of the Eurobarometer surveys, in the first step I want to look at the non-responses. Even assuming that the income question fulfils the basic requirements of comparability and reliability, we still need to know how valid the results we obtain are. Non-response, apparently, poses the most important problem for this question. Since we cannot assume that the non-respondents are distributed equally over all income categories, we have to find out how they are biased. This is the second step I want to take.

<sup>1</sup> In the following, I will use the term 'non-response' to denote 'item-non-response', *i.e.* the refusal to answer a particular question. This is of course different from 'unit-non-response', *i.e.* the refusal to participate in a survey in the first place.

If the instrument is to be comparable even taking into account the number of non-responses, then at least the non-respondents should show similar patterns over different countries. If, *e.g.* the probability to refuse to answer the income questions were somehow connected to sex in the northern countries (for whatever reasons) and in southern countries to age, we may not compare these results, even if in both countries we find the same portion of non-respondents. Using income there as an independent variable will inevitably lead to different confounding effects with other variables and will render the results of those analysed doubtful.

In the third step, I will look at an instrument which has been designed by the ESOMAR working group on the standardisation of demographics. This instrument uses consumer durables in households and may turn out to be a proxy instrument for measuring income or, at least, may indicate something of a household's purchasing power. There, I will look both at the correlations between the income variable and this index, and some aspects of its internal reliability.

## 2. THE DATA BASE

I will use the Eurobarometer 46.0 as a data base. This Eurobarometer offers two virtues: First, it has been the standard instrument for comparative social research in Europe for many years. Its sampling and processing procedures are acknowledged to be of very high quality. Furthermore, many of the instruments have remained unchanged and thus give a reliable basis for comparisons, both across countries and over time. What is of particular interest for this paper is the fact that this Eurobarometer comprises of not only the standard income question, but also the ESOMAR proxy scale which measures economic status and which allows one to compare these two scales.

The Eurobarometer 46.0 was conducted on behalf of the GD X by INRA Europe. The fieldwork was in October 1996 and sampled 16,248 interviews. Approximately 1,000 interviews per country were made, with the exceptions of Luxembourg ( $n=600$ ) and Northern Ireland ( $n=300$ ). The data file used for the following analysis was provided by the Central Archive for Empirical Social Research in Cologne (study number 2898).

## 3. NON-RESPONSES IN THE INCOME QUESTION

The first question to arise is: are there any meaningful and substantive differences in the response rates of the various EU-countries? The standard

Table 1: Valid Answers, Refusals and Don't Knows by Country (in %)

Country	Valid	Refusals	D.K.	Refusals + D.K.
EU 15	74.5	9.3	16.2	25.5
Northern Ireland	50.1	23.3	26.6	49.9
Ireland	50.7	22.7	26.6	49.3
Belgium	50.7	39.5	9.8	49.3
Italy	60.3	21.8	17.9	39.7
Luxembourg	62.7	26.8	10.5	37.3
Spain	63.9	23.5	12.6	36.1
Great Britain	66.8	17.7	15.5	33.2
Austria	69.6	23.3	7.2	30.4
Greece	73.9	18.1	8.0	26.1
France	81.5	14.2	4.3	18.5
Portugal	82.4	9.1	8.5	17.6
Germany / West	85.5	12.1	2.4	14.5
Netherlands	86.3	8.3	5.4	13.7
Germany / East	87.0	10.9	2.1	13.0
Denmark	89.2	5.1	5.7	10.8
Finland	90.9	4.7	4.4	9.1
Sweden	91.7	3.1	5.2	8.3

question in almost all Eurobarometer surveys asks for the monthly gross household income, including pensions, social benefits *etc.*, by handing the respondent a card with 12 income brackets:

"We also need some information about the income of this household to be able to analyse the survey results for different types of households. Here is a list of income groups. (SHOW CARD) Please count the total wages and salaries PER MONTH of all members of this household; all pensions and social insurance benefits; child allowances and any other income like rents, *etc.* Of course, your answer as all other replies in this interview will be treated confidentially and referring back to your household will be impossible. Please give me the letter of the income group your household falls into before tax and other deductions."

Since questions on income are rather sensitive, the respondent is assured of the confidentiality of his or her answer. To support this, the respondent does not need to mention the exact amount of income to the interviewer, but just has to mention a letter. These measures have been adopted to reduce non-response in this particular item.

Table 1 demonstrates impressively that we do have a problem when measuring income: On (unweighted) average, one out of four respondents refuses to indicate his/her income. We also have large differences between the countries: In Northern Ireland, Ireland and Belgium, every second respondent refuses to indicate his/her household income, whereas all the Scandinavian countries show the lowest portion of non-response. The differ-

ence between the portion in Northern Ireland and Sweden actually amounts to 41.6 percentage points. Thus, portions of non-respondents this large defy any sampling theory, because we must assume that these non-response rates are due to systematic errors. Furthermore, differences this large between the countries may also lead to problems in the comparability over the various countries: Results obtained from only half of the respondents are certainly less useful for analysis than results obtained from nine out of ten respondents.

In all but two countries the refusal rates are higher than the don't know rates, sometimes by a ratio of four (as in Belgium) or even five times (as in both parts of Germany). Thus, it seems safe to assume that for a substantial proportion of the respondents it is not that they simply do not know (as it may be the case if *e.g.* the target person is one of the children and does not have an idea of what the income is), but that most refusals are meant in the way: I do not want to give you an answer to your question. If this behaviour is based on attitudes, then we may assume that the distribution of non-responses within the countries varies. Thus, excluding the non-respondents from analysis renders them problematic, because we may not assume that they are distributed as the respondents are, but that we do have a bias in responding.

#### 4. SOCIAL CHARACTERISTICS OF NON-RESPONDENTS

Having shown above, that we face large differences across Europe in the way the question on household income is answered, in this section I will try to find out whether we are also confronted with different biases across Europe. If this were the case, then the comparison of household income would be even more difficult, because – in a technical sense – we would have to compare samples which were differently structured. It would be as though we had sampled from more or less different universes and had had to compare these different samples. On the other hand, if we can find a similar bias over most of the countries, we would be able to compare these different samples, still taking into account their common bias. Therefore, I will look at some of the socio-demographic characteristics which describe the non-respondents, starting with the basic descriptors sex and age, followed by descriptions of the respondent's role within a household (marital status, household size, division of labour within a household). In order to have a sufficiently large number of cases even for those countries with a relatively small portion of non-response, I will put both "refusals" and "don't knows" into one category.

Table 2: Non-Respondents by Sex and Country (in %)

Country	Male	Female	$\chi^2$
EU 15	25.3	26.0	0.9
Northern Ireland	57.0	42.9	6.1 *
Ireland	49.5	49.0	0.2
Belgium	49.2	49.5	0.0
Italy	37.5	41.7	2.0
Luxembourg	34.0	40.6	2.8
Spain	34.0	38.1	1.8
Great Britain	33.1	33.1	0.0
Austria	31.0	29.9	1.3
Greece	28.3	24.0	2.3
France	17.8	19.3	0.4
Portugal	17.0	18.0	0.2
Germany / West	16.8	12.4	4.0 *
Netherlands	11.0	16.4	6.5 *
Germany / East	13.5	12.6	0.2
Denmark	8.5	13.2	5.7 *
Finland	8.6	9.5	0.3
Sweden	8.2	8.3	0.0

Significance levels: \*  $p \leq 0.05$ .

When enquiring about why certain respondents do not answer the income question, one may, among other factors, assume that there are differences in the answering behaviour between the sexes.

However, looking at the data in Table 2, we can clearly see that this is not the case. Only in Northern Ireland, West-Germany, Denmark and the Netherlands can we see some slightly significant differences in the respective distributions: In two of these countries men show a higher portion of non-responses (Northern Ireland and West-Germany), in the Netherlands and Denmark it is women who do. Therefore, it is safe to conclude that sex neither plays a role in the non-response behaviour nor is there a discernible pattern which separates some countries from others.

Table 3 shows a somewhat different finding – in seven out of 17 countries sometimes very large differences appear when the respondent's age groups are compared. This may be due to a life cycle effect in that younger persons who are still at the beginning of their professional career and thus, usually earning less than someone who is already at the peak of his/her career or is 'established', will be less willing to confess their income situation to an interviewer. If this is the underlying reason, then the data support this assumption: Only in four out of 17 countries does the older age group seem to be more reluctant than the younger one, and none of these cases are statistically significant. On the other hand, in all other countries the younger ones show a higher number of non-respondents. Here we can observe differences as large as 15 percentage points between these two age

Table 3: Non-Respondents by Age Groups and Country (in %)

Country	15-39 years	40+ years	$\chi^2$
EU 15	28.4	23.3	54.6 ***
Northern Ireland	57.6	42.2	7.2 **
Ireland	53.6	44.9	7.6 **
Belgium	52.6	46.7	3.5
Italy	48.1	32.8	25.8 ***
Luxembourg	37.6	37.0	0.2
Spain	40.4	31.9	7.9 **
Great Britain	34.1	32.3	0.4
Austria	29.7	31.2	0.3
Greece	29.7	22.8	6.0 *
France	19.9	16.9	1.5
Portugal	20.6	14.7	6.0 *
Germany / West	17.2	12.5	4.5 *
Netherlands	13.6	13.9	0.1
Germany / East	14.9	11.9	2.0
Denmark	9.6	11.7	1.1
Finland	7.9	10.0	1.2
Sweden	10.1	7.0	3.1

Significance levels: \*  $p \leq 0.05$ ; \*\*  $p \leq 0.01$ ; \*\*\*  $p \leq 0.001$ .

groups (Northern Ireland and Italy). Therefore, we may conclude that age does play a role in the respondent's behaviour. However, this is only the case in some of the countries: We find it statistically significant for both Irelands, West-Germany, and all of Southern and South-West Europe. There are no differences in countries with a predominantly protestant population – *i.e.* the Scandinavian countries, Great Britain, Netherlands, East Germany.

Table 4 explores the relationship between response behaviour and marital status. Being either single (single, divorced, widowed or separated) or non single (married, living in a consensual union) tells more about emotional characteristics than about socio-demographic characteristics. Being 'single' does not mean living in a single person household (this we will turn to later). However, single persons still show a characteristic answering behaviour. As Table 4 shows, those who according to this classification are single – with exceptions in two countries – are less willing to offer insights about their income than those who live with a partner. All in all, we find this relation to be statistically significant in 11 out of 17 countries, with differences between the two groups as large as 18 percentage points for Italy. This already comes closer to a pan-European bias in the respondents' behaviour. However, it remains unclear why they behave in such a manner: If they were more or less misanthropic, they would not have consented into an interview in the first place. It appears that for these persons the question on income is a particularly touchy subject.

Table 4: Non-Respondents by Marital Status and Country (in %)

Country	Single	Non Single	$\chi^2$
EU 15	30.0	22.7	109.1 ***
Northern Ireland	57.1	44.7	4.6 *
Ireland	54.5	45.4	8.1 **
Belgium	53.7	46.3	5.5 *
Italy	50.2	31.8	36.8 ***
Luxembourg	44.1	33.0	7.7 **
Spain	42.2	31.0	13.5 ***
Great Britain	40.0	28.5	15.1 ***
Austria	30.3	30.5	0.0
Greece	32.7	22.3	13.0 ***
France	22.7	15.8	7.7 **
Portugal	20.2	15.8	3.2
Germany / West	12.7	15.8	2.0
Netherlands	14.9	13.2	0.6
Germany / East	11.1	14.2	2.1
Denmark	14.0	8.8	6.7 **
Finland	9.2	9.0	0.0
Sweden	12.0	6.0	10.8 ***

Single: Single, divorced, separated, widowed / non single: married, living as married

Significance levels \*  $p \leq 0.05$ ; \*\*  $p \leq 0.01$ ; \*\*\*  $p \leq 0.001$ .

Regarding the difference between one and more person households, two alternative hypotheses might be examined: The first one states that younger one person households usually are economically better off than families or larger households, because they do not have to support other household members who engage in what is sometimes called 'unproductive' activities like child-rearing or supporting the elderly. Even if tax allowances for families or direct benefits are taken into account, one person households are usually better off. Being better off and knowing it, may on the other hand lead to an uneasiness about disclosing one's income, because – at least in Germany – these differences in income distribution between one and more person households have during the last years become intensely debated and have turned into a somewhat moral issue.

The other alternative picks up the argument put forward in connection with the analyses of sex differences. Here, one may argue that the chance to interview someone who does not know about the family's income simply is larger in a more person household than in a one person household. Thus, there should be higher non-response rates in larger households than in one person households.

Table 5 indicates that the latter proposition has some of the evidence in its favour. Although we find only six of the 17 crosstabulations to show a significant Chi-Square, in 12 out of 17 countries the highest portion of non-

Table 5: Non-Respondents by Size of Household and Country (in %)

Country	1 Person	2-4 Persons	5 and more	$\chi^2$
EU 15	19.0	25.8	34.6	151.1 ***
Northern Ireland	51.2	46.0	61.3	4.5
Ireland	39.0	48.6	54.1	7.3 *
Belgium	47.0	50.1	50.0	0.7
Italy	34.4	39.6	44.0	2.1
Luxembourg	42.1	36.7	36.0	0.9
Spain	33.3	34.7	41.9	4.0
Great Britain	29.5	33.0	39.3	2.8
Austria	23.6	33.3	26.0	9.4 **
Greece	23.7	25.5	30.4	1.9
France	15.6	18.3	25.5	4.7
Portugal	8.2	17.5	22.6	9.1 *
Germany / West	10.6	16.4	11.8	6.0 *
Netherlands	10.1	14.4	16.2	3.2
Germany / East	7.2	14.5	22.9	13.6 **
Denmark	13.1	9.7	12.5	2.5
Finland	7.2	9.1	15.4	5.6
Sweden	3.4	8.4	18.2	18.0 ***

Significance levels: \*  $p \leq 0.05$ ; \*\*  $p \leq 0.01$ ; \*\*\*  $p \leq 0.001$ .

response can be found in households with five and more persons. Only in Luxembourg and Denmark do one person households show a higher portion of non-response, and only in Austria do households with two up to four persons exhibit the highest rate of non-response. Although one has to bear in mind that these relations are somewhat feeble, because in eight of the 17 countries some of the non-respondent cells fall below 30 cases, at least we can conclude that there is an effect due to the size of the household on the incidence of non-response. This is in line with the argument that due to the random key to find the target person in a given household, there is a likelihood to find someone who is simply not able to answer questions of this kind reliably.

The next two tables go deeper into the structure of the respondent's household – i.e. its division of labour. In Table 6 we examine the non-response rate of those mainly responsible for ordinary shopping; in Table 7 the answering behaviours of the main contributors to household income are shown. As this data set shows, more than four out of five respondents responsible for the ordinary shopping are women (83%), on the other hand, more than three out of four of the main contributors to the income are men (76.8%). Thus, this question also picks up the difference between male and female respondents. However, although we have seen above that there are hardly any differences between the respondent's sex and his/her answering behaviour, we find in these tables some of most frequent differences between the two respective groups. In all but four out of the 17 countries we

Table 6: Non-Respondents by Person mainly Responsible for Ordinary Shopping and Country (in %)

Country	Yes	No	$\chi^2$
EU 15	21.8	30.7	164.9 ***
Northern Ireland	40.6	60.7	12.3 ***
Ireland	39.7	58.2	34.4 ***
Belgium	46.1	54.0	6.2 *
Italy	35.3	43.0	6.5 *
Luxembourg	36.5	38.2	0.2
Spain	33.3	38.3	2.7
Great Britain	28.7	39.9	14.2 ***
Austria	28.3	34.3	4.2 *
Greece	23.6	29.1	3.9 *
France	15.9	22.4	6.9 **
Portugal	14.9	20.0	4.6 *
Germany / West	11.8	19.4	10.8 ***
Netherlands	13.1	14.8	0.7
Germany / East	10.1	18.4	14.5 ***
Denmark	11.5	9.5	0.9
Finland	7.6	12.5	6.1 *
Sweden	6.3	10.6	5.9 *

Significance levels: \*  $p \leq 0.05$ ; \*\*  $p \leq 0.01$ ; \*\*\*  $p \leq 0.001$ .

observe significant differences between those who are mainly responsible for ordinary shopping. With the exception of Denmark, it is always the person not responsible for the shopping, who more often refuses to answer. Again, we observe differences as large as 20 percentage points for Northern Ireland and 18.5 points for Ireland, with a difference between the two groups of 8.9 percentage points across all the countries. Thus, it seems that one of the main factors contributing to non-response in regard to the income question is knowledge – those who know or at least have an idea of the income are also more willing to answer than those who do not. This factor can be found throughout Europe, however, it certainly does not account for the high variation of the level of non-response that we can observe between the various countries.

The results from Table 7 corroborate the latter finding: In all countries the persons who are not the main contributor to the household income have a higher propensity to refuse to answer than those who actually are, and only in two countries do we find these differences not to be statistically meaningful. The largest difference can be observed in Italy, here there is almost a 16 percentage point difference between those who are main contributors and those who are not. But again, the high overall levels between the countries still persist, even where there are large differences as is the case for Italy, or differences which are rather negligible, as is the case for Denmark. Thus, again we can observe a constant pattern in the bias of non-response rates for

Table 7: Non-Respondents by Main Contributor to Household Income and Country (in %)

Country	Yes	No	$\chi^2$
EU 15	20.5	32.3	289.2 ***
Northern Ireland	45.2	56.3	3.6 *
Ireland	41.2	56.9	24.8 ***
Belgium	45.8	53.9	6.5 *
Italy	29.7	47.3	33.5 ***
Luxembourg	31.8	43.8	9.3 **
Spain	28.3	42.2	20.5 ***
Great Britain	27.9	39.2	14.6 ***
Austria	28.6	33.6	3.0
Greece	21.2	30.3	10.7 ***
France	16.3	21.7	4.7 *
Portugal	13.5	21.8	12.8 ***
Germany / West	12.9	18.1	4.9 *
Netherlands	10.1	18.8	16.4 ***
Germany / East	11.5	16.3	4.7 *
Denmark	9.8	12.5	1.7
Finland	6.9	12.9	10.0 **
Sweden	3.8	15.4	42.2 ***

Significance levels: \*  $p \leq 0.05$ ; \*\*  $p \leq 0.01$ ; \*\*\*  $p \leq 0.001$ .

all countries, however for some countries, this pattern is more apparent than for other countries.

## 5. PRELIMINARY CONCLUSION

Table 8 sums up the results we have found so far, including some more data which have not been presented yet, such as a grade of schooling, city size, and occupational status. The columns are ranked according to the magnitude of the Chi-square value obtained for the overall crosstabulation of all EU 15 countries – thus, beginning with sex, which showed to have the least relationship with non-response, up to the variables household size, responsibility for doing the ordinary shopping, and being the main contributor to household income. To come back to our initial questions, we asked:

- How are the non-response rates distributed over the various socio-demographic groups, and
- Are there any discernible patterns which differentiate one country or groups of countries from one another?

Table 8:  $\chi^2$ -Values for Various Distributions by Country

Country	Sex	Schooling	City Size	Age Group	Family Status	Occupation	Household Size	Shopping	Main Contributor
EU 15	0.9	13.3 ***	24.0 ***	54.6 ***	109.1 ***	132.0 ***	151.1 ***	164.9 ***	289.2 ***
Ireland	0.2	32.7 ***	7.3 *	7.6 **	8.1 **	9.0 *	7.3 *	34.4 ***	24.8 ***
Italy	2.0	6.5 *	4.9	25.8 ***	36.8 ***	24.8 ***	2.1	6.5 *	33.5 ***
Sweden	0.0	3.4	10.3 **	3.1	10.8 ***	16.8 ***	18.0 ***	5.9 *	42.2 ***
Spain	1.8	15.2 ***	5.4	7.9 **	13.5 ***	4.5	4.0	2.7	20.5 ***
Great Britain	0.0	5.4 *	1.5	0.4	15.1 ***	5.9	2.8	14.2 ***	14.6 ***
Portugal	0.2	15.6 ***	0.6	6.0 *	3.2	3.2	9.1 *	4.6 *	12.8 ***
Germany / West	4.0 *	14.7 ***	3.9	4.5 *	2.0	4.2	6.0 *	10.8 ***	4.9 *
Germany / East	0.2	0.1	17.4 ***	2.0	2.1	0.1	13.6 **	14.5 ***	4.7 *
Greece	2.3	4.8 *	8.7 *	6.0 *	13.0 ***	2.4	1.9	3.9 *	10.7 ***
France	0.4	6.4 *	1.4	1.5	7.7 **	18.3 ***	4.7	6.9 **	4.7 *
Northern Ireland	6.1 *	3.3	3.2	7.2 **	4.6 *	5.1	4.5	12.3 ***	3.6 *
Belgium	0.0	11.3 ***	5.0	3.5	5.5 *	5.9	0.7	6.2 *	6.5 *
Netherlands	6.5 *	0.2	10.5 **	0.1	0.6	4.3	3.2	0.7	16.4 ***
Denmark	5.7 *	4.8 *	0.9	1.1	6.7 **	17.5 ***	2.5	0.9	1.7
Finland	0.3	0.5	10.7 **	1.2	0.0	6.6 *	5.6	6.1 *	10.0 **
Luxembourg	2.8	2.4	3.0	0.2	7.7 **	13.5 **	0.9	0.2	9.3 **
Austria	1.3	2.6	4.0	0.3	0.0	2.5	9.4 **	4.2 *	3.0

Significance levels: \* $p \leq 0.05$ ; \*\* $p \leq 0.01$ ; \*\*\* $p \leq 0.001$ .

Just to remind ourselves, we noted at the beginning that in order for the different results to be comparable, at least, the biases should be somewhat equally distributed over the countries.

The answer to the first question ("How are non-responses distributed over the various socio-demographic groups?") is that the classical factors, such as: sex, schooling or age group play only minor roles. It does not make a difference whether you are male or female, well or badly educated, or young or old to how you react when asked for your income. Thus, using these categories to analyse income distribution seems to be fairly safe, because respondents and non-respondents are spread rather evenly across these groups. However, it does make a difference whom you ask in the household: Those who administer the money (*i.e.* usually either the father or the mother) are the ones who are more able and willing to answer this question. Therefore, if this question has to be posed, it should be addressed to the person which is in this case the most knowledgeable. This may, in some instances, lower the non-response rate significantly – *e.g.* in Ireland and Northern Ireland around 10 percentage points from a high of almost 50% non-respondents. For practitioners, this could also mean switching from the initially identified target person to the head of household when it comes to hard facts and information which cannot be expected to be reliably supplied by, let's say the sixteen year old son or daughter.

The next question ("Are there any discernible patterns across countries or groups of countries?") is somewhat harder to answer. What we can observe is that not all countries show the same strong or weak relationships between the socio-demographic variables and non-response. Strong effects can be observed for Ireland (where only the difference between the sexes does not prove to be statistically significant), Italy and Sweden. At the opposite end, we find countries, such as Austria, Luxembourg or the Netherlands, where only two or three effects are significant. Furthermore, with respect to some socio-demographic variables (*e.g.* schooling) we find that in some of the countries they do make a difference (such as in Ireland, Spain, Portugal or West Germany), whereas in other countries they do not have an effect (such as in the Netherlands or East Germany). This can be observed for almost any variable. Even with the most promising variables related to the respondent's role in the household, strong relationships appear for Sweden, Italy, Ireland and Spain, but they are not significant for Finland and Austria. Thus, all countries react differently to the variables included here, a common pattern can not be found.

We are left with three results so far:

- Countries do show large differences in their respective non-response rates.

- Most countries do have a bias in non-response rates which is connected to the interviewed person's knowledge or position in the household.
- Otherwise the response rates do not show a discernible pattern.

These findings in turn cast a shadow on the comparability of income data.

## 6. A PROXY SCALE FOR INCOME

Having come this far, I would like to address the last question which I raised in my introduction: Is there a solution to income measurement if both assumptions (*i.e.* an even distribution of the non-response over the various sub-samples within one country and a discernible pattern how non-respondents are distributed across the countries) do not hold true? There could be if this solution in some way offers a valid proxy instrument for what we want to measure.

Although to my knowledge there exists no easy ready-to-use solution for this problem yet, I want to indicate that there may be more than one way to measure income in comparative perspective, and, given time and sociological creativity, what can be called 'proxy measurements' may come of use to solve this problem which concerns us here.

Of course, this problem of comparability is not new and some efforts have been made to overcome this. One of these efforts has been carried out by ESOMAR via several working groups since the beginning of the 1980s. ESOMAR has come up with a Social Grade variable which allows a socio-economic/-demographic classification relevant to all countries (ESOMAR 1997; reprinted as Chapter 6 in this volume). This grading grid is based on three variables, they are: the occupation of the main income earner, the educational level of the main income earner, and the economic status of the household. Since this grading scheme incorporates at least three different variables (occupation, education, economic status), I will select only one, which is the scale on economic status.

This scale is based on 10 long-lasting consumer goods. The underlying idea apparently is that the possession of these goods is an indicator of the household's economic purchasing power and the accumulation of these goods can be interpreted in terms of relative distances. As such it is related to the economist's measurement instrument of 'consumer goods baskets' and the derived 'purchase power parities', used for the international comparison of two countries' currencies or national GNP-data. The contents of this particular ESOMAR-basket is shown in Table 9. As with any other basket, the definition of its contents always remains subject for debate.

Table 9: Availability of Long-Lasting Consumer Goods by Country (in %)

Country	Mean Number of Goods	Colour TV	Still Camera	Video-recorder	Radio Clock	Electric Drill	Electric Deep Fryer	PC	2 or more Cars	Video Camera	Second/Holiday Home
EU 15	4.98	97.8	78.0	70.9	68.7	60.4	38.7	29.4	25.2	17.4	11.6
Greece	3.65	97.2	66.4	47.1	49.2	29.0	30.6	9.3	11.7	6.4	18.1
Portugal	4.09	99.3	56.7	51.7	63.7	45.5	43.1	12.6	18.0	9.0	9.6
Finland	4.43	95.7	80.9	67.8	54.7	56.1	1.6	28.9	22.6	10.5	24.6
Germany / East	4.58	98.6	81.8	61.9	66.5	68.8	18.0	22.1	19.3	18.9	2.5
Ireland	4.62	96.8	60.7	78.2	64.4	48.7	56.5	19.6	22.2	11.4	3.0
Spain	4.83	99.1	74.4	73.8	66.7	46.4	41.9	24.5	21.8	15.7	19.1
Germany / West	4.97	99.3	81.0	69.5	68.6	62.6	37.9	29.4	22.5	21.2	4.8
Northern Ireland	4.99	96.4	62.5	84.9	66.1	57.6	59.2	26.3	27.0	14.1	5.3
Denmark	5.01	97.4	87.2	74.0	64.5	72.3	16.4	49.6	13.3	13.9	12.2
Austria	5.07	98.6	82.1	71.9	67.8	61.5	34.0	28.4	26.4	26.2	9.7
France	5.14	93.8	79.8	72.5	81.9	64.2	38.9	21.1	35.7	17.8	8.7
Sweden	5.19	98.3	83.9	73.5	78.8	67.3	10.9	34.0	26.5	22.3	23.1
Great Britain	5.36	98.3	83.2	87.7	71.1	68.0	37.3	37.5	28.9	18.0	6.4
Italy	5.37	99.0	81.7	73.9	65.9	61.3	28.2	30.8	52.9	20.8	22.5
Belgium	5.56	97.1	74.0	73.5	81.8	62.5	84.5	32.4	23.9	16.6	9.9
Netherlands	5.78	98.3	90.3	77.5	75.2	79.5	63.1	53.0	16.3	22.6	2.4
Luxembourg	6.53	99.7	93.1	76.8	84.6	77.9	87.1	41.1	47.8	31.8	12.9

Table 10: Correlation Between Household Income and Proxy-Scale, Cronbach's Alpha, Scale Mean and Variance

Country	Correlation	Cronbach's Alpha	Mean	Variance
Portugal	0.677 ***	0.757	4.09	5.20
Finland	0.600	0.631	4.43	3.67
Spain	0.599 ***	0.716	4.83	4.92
Denmark	0.580 ***	0.602	5.01	3.25
Ireland	0.580 ***	0.647	4.62	3.90
Germany / West	0.559 ***	0.714	4.97	4.62
Greece	0.551 ***	0.627	3.65	3.59
Germany / East	0.549 ***	0.670	4.58	3.75
Great Britain	0.542 ***	0.658	5.36	3.85
Sweden	0.533 ***	0.601	5.19	3.49
Italy	0.521 ***	0.708	5.37	5.01
EU 15	0.513 ***	0.677	4.98	4.40
Luxembourg	0.512 ***	0.630	6.53	3.43
Belgium	0.505 ***	0.672	5.56	3.94
France	0.499 ***	0.658	5.14	4.06
Netherlands	0.484 ***	0.610	5.78	3.21
Austria	0.451 ***	0.697	5.07	4.62
Northern Ireland	0.398 ***	0.713	4.99	4.67

Significance levels: \*  $p \leq 0.05$ ; \*\*  $p \leq 0.01$ ; \*\*\*  $p \leq 0.001$ .

It seems reasonable that the possession of scarce goods, such as a second car, is an indicator for wealth and economic status. However, does the scarcity of electric deep fat fryers in Finland imply that the possession of this particular good is an indicator of a high economic status? Furthermore, does it take into account how these durables have been acquired? Someone who is completely in debt may, according to this scale, be attributed to an economic status higher than he/she actually has. Yet, what is more important here is how well this scale behaves compared to the income question. I do not want to scrutinise this scale in all details, but in order to give an impression I want to display some of the key indicators used for evaluating such a scale.

The first question to be asked, of course, is: do we have the same problem here regarding non-response, as with the income scale? No, not at all: Out of 16,248 cases, only 86 cases (= 0.5%) have missing answers. Thus, the worst problems which are the high level plus the large differences in response rates across the countries, do not play a role here.

The second question is whether this scale is at least statistically reliable. The most commonly used indicator to determine this is Cronbach's Alpha. This is a model of internal consistency, which is based on the average inter-item correlation. Usually values above 0.7 are regarded as satisfactory for assuming the internal consistency of a scale. The average of this scale for all of the EU 15 is 0.677, which is somewhat short of the required benchmark (see Table 10). Only in Northern Ireland, Italy, West Germany, Spain and

Table 11: Means for Proxy-Scale by Household Income Non-Respondents and Respondents by Country (in %)

Country	Response	Non-response
EU 15	4.93	5.13 ***
Northern Ireland	5.54	4.70 *
Ireland	5.06	4.75 *
Belgium	4.94	5.59
Italy	5.29	5.50
Luxembourg	6.54	6.51
Spain	4.73	5.03 *
Great Britain	5.35	5.39
Austria	4.96	5.31 *
Greece	3.73	3.42 *
France	5.10	5.34
Portugal	4.03	4.39
Germany / West	4.94	5.12
Netherlands	5.75	6.01
Germany / East	4.59	4.52
Denmark	5.06	4.61 *
Finland	4.40	4.76
Sweden	5.14	5.67 *

Significance levels: \*  $p \leq 0.05$ ; \*\*  $p \leq 0.01$ ; \*\*\*  $p \leq 0.001$ .

Portugal does the scale pass this convention. The lowest values are to be observed in Denmark, Sweden, and the Netherlands with values only slightly higher than 0.6. However, since none of the observed countries departs much from the EU 15 average, for our purpose this scale may be regarded as reliable.

The third question regards its face validity: Do the results we gain from it correspond with what seems reasonable? Thus, are countries which are ranked by the mean number of durables in a relative position to each other which seems reliable? At least for the lower end of the table this holds true: Spain, Ireland, East Germany, Portugal, and Greece are usually the countries which also rank at the lower end of GNP-scale within the EU. However, Finland, which usually is found in the middle section of these countries, is also to be found at the bottom of this scale. Furthermore, both Denmark and West Germany find themselves in the middle of the countries, although one would expect to find them on the top of the list, along with Luxembourg. Thus, one can conclude that this scale (which of course is only a sub-scale of the ESOMAR Social Grade System) somewhat correctly discriminates against the poorer countries, yet it is not very reliable when it comes to ranking the rich countries in order.

The fourth question considers the correlation between the income scale and this proxy scale. If we want to use this scale as another measurement of income, then it should at least satisfactorily correlate with the income scale.

We can observe correlations ranging from 0.677 for Portugal down to 0.451 for Austria and 0.398 for Northern Ireland. The overall correlation for the EU is 0.513, which can be regarded as fairly good, considering the large differences of income and wealth over the different countries.

Thus, if we accept this proxy scale as giving us, at least, somewhat of an idea of each country's income distributions, then we might also get an idea of how non-response is distributed over household income (Table 11). The results, again, form anything but a pattern. For the EU overall, we observe a higher economic status as measured by the proxy scale for the non-respondents than for the respondents. This lends credibility to the thesis that non-response is also affected by household income such that the higher income classes are either reluctant to disclose their true income or that they underreport their true income. This reluctance is seen in Spain, Austria, and Sweden, where we find significant differences between the means of both groups, but this can also be seen in most of the other countries. However, the opposite holds true for Northern Ireland, Ireland, Greece, and Denmark: Here it is the case that the respondent's households are more well off than the non-respondent's households. Further analysis might explore how these different biases can be explained.

## 7. CONCLUSION

The aim of this paper was to draw attention to some of the inherent problems when doing international research. I have started with some comments on the instruments' quality needed, especially on the problem of comparability, which goes beyond keeping instruments constant. What I have tried to show is that one encounters problems even when using highly standardised instruments, such as the income question. Standardisation is the first, indispensable step towards comparability. However, even then we may obtain results which are hard to compare. This I have tried to show by analysing the different biases which underlie the non-response rates in different countries. Although I have found a somewhat pan-European bias which is related to the target person's knowledge, I have also found that apart from this there is no single pattern which underlies non-response rates and which can be taken into account when we compare the results of such surveys. I have suggested that one of the possibilities to circumvent this problem may lie in using proxy scales. As I have also shown, the proxy scale used here has other disadvantages, e.g. concerning the countries' relative ranking order. To conclude, when we undertake international research we are in some sense in a position like Ulysses was: We are left between Scilla and Charybdis, being aware of the problems both instruments have. However, this should not keep

us from starting this venture, because, as Ulysses' example shows, with care and creativity we will find satisfactory solutions for these problems, too.

## REFERENCES

- ESOMAR. 1997. *Standard Demographic Classification. A System of International Socio-Economic Classification of Respondents to Survey Research*. Amsterdam: ESOMAR.
- ADM (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.). 2001. *Jahresbericht 2000*. Frankfurt/M.: ADM.

# Advances in Cross- National Comparison

## A European Working Book for Demographic and Socio-Economic Variables

Edited by

Jürgen H. P. Hoffmeyer-Zlotnik

*Centre for Survey Research and Methodology  
Mannheim, Germany*

and

Christof Wolf

*University of Cologne  
Cologne, Germany*

Kluwer Academic / Plenum Publishers  
New York, Boston, Dordrecht, London, Moscow