Assessment of non-response in the health and sexual behaviour survey

1 Introduction

There are two types of errors that can affect a survey: Sample Errors and Non-Sampling Errors. The former can be calculated using statistical procedures, whilst the so-called **non-sampling errors**, which this document refers to, are hard the measure.

These errors appear in the different stages of the statistical process, and can appear before information is garnered (deficient frame, insufficient definitions or in the questionnaires, etc.), during the collection of information (incorrect fieldwork, incorrect statements or non-response of respondents) and, finally, in operations subsequent to fieldwork (errors when coding variables, recording questionnaires, etc.).

As aforementioned, it is very difficult to assess these errors, given the ample variety of causes that could originate them.

Among these causes, one of the most notable is the **non-response of the respondent units**, due to refusal to answer the questionnaire, absence, inability to answer of the whole household composing the respondent unit or dwelling inaccessible when the interview was carried out.

The respondent units considered in the *Health and Sexual Behaviour Survey* are persons aged between 18 and 49 years old.

An **assessment questionnaire** has been designed to analyse non-response in this survey, aiming to obtain information on the basic characteristics of the persons who have not taken part due to any of the aforementioned reasons.

This questionnaire is only completed for **incumbent** persons who, due to a certain incident, did not take part in the survey. It will not be completed for substitutes who were asked to take part in the survey to replace a certain person but were not able to respond either.

The questionnaire comprises four sections. The first section contains the data required to identify the person, i.e. the respondent unit. The second refers to the type of incident and the third indicates if the incumbent person has been replaced or not, entering the reference number of the substitute if applicable. Finally, the fourth section comprises a series of basic data on the person: sex, age, marital status, highest level of education obtained, situation as regards activity, nationality, country of origin if the incumbent person has a foreign nationality.

When an incumbent person has been replaced, the interviewers question as many substitutes as required until finding a person who will collaborate. This person is called the substitute and given a corresponding order number that will be entered on the assessment questionnaire.

Thanks to a computer programme, this order number can be used to obtain data for the substitutes who have replaced the incumbent persons.

If after all substitutes are interviewed, the incumbent person has not yet been supplanted, no replacement is used, thus leading to sample loss and a consequent

decrease of the precision of the estimators. In this case, the order number for the substitute is left blank.

2. Analysis of the data

Table 1 presents distributions by autonomous communities, of the theoretical sample of persons, the total effective sample (total number of persons interviewed) and effective sample of incumbent persons (total number of incumbent persons interviewed).

Table 1. Distribution of the theoretical and effective sample of persons by Autonomous Communities

Autonomous Communities	Theoretical sample		Effective sample			
			Total		Incumbent persons	
	Persons	%	Persons	%	Persons	%
Total	13.600	100,00	10.980	80,74	5.263	38,70
Andalucía	1.800	100,00	1.561	86,72	820	45,56
Aragón	544	100,00	434	79,78	195	35,85
Asturias (Principado de)	488	100,00	450	92,21	219	44,88
Balears (Illes)	472	100,00	356	75,42	168	35,59
Canarias	672	100,00	535	79,61	230	34,23
Cantabria	392	100,00	354	90,31	158	40,31
Castilla y León	768	100,00	661	86,07	333	43,36
Castilla-La Mancha	624	100,00	544	87,18	262	41,99
Cataluña	1.552	100,00	1.071	69,01	475	30,61
Comunidad Valenciana	1.136	100,00	960	84,51	456	40,14
Extremadura	488	100,00	423	86,68	223	45,70
Galicia	784	100,00	692	88,27	325	41,45
Madrid (Comunidad de)	1.520	100,00	1.007	66,25	433	28,49
Murcia (Región de)	544	100,00	525	96,51	298	54,78
Navarra (Comun. Foral de)	400	100,00	306	76,50	151	37,75
País Vasco	736	100,00	594	80,71	271	36,82
Rioja (La)	352	100,00	249	70,74	121	34,38
Ceuta y Melilla	328	100,00	258	78,66	125	38,11

On a national level, the total effective sample represents almost 81 percent of the theoretical sample, whilst the effective sample of incumbent persons only represents 39 percent of the same. These figures indicate that only 39 percent of the total number of incumbent persons were interviewed. A very low figure if compared to that of other surveys, which reflects the complexity of this survey given the content of the questionnaire. Only 42 percent of the incumbent persons who presented incidents were replaced, thus elevating the total effective sample to 81 percent of the aforementioned theoretical sample.

On considering data by Autonomous Communities, there are only three, Murcia, Asturias Cantabria, with total effective sample percentages above 90 percent; in fact Murcia stands out with over 96 percent. The opposite occurs in Madrid and Cataluña, with total effective sample percentages below 70 percent, the former reaches 66 percent, whilst the latter amounts to 69 percent.

The situation does not differ greatly when considering the percentage of the effective sample of incumbent persons: Murcia has the highest percentage, almost 55 percent, with Extremadura and Andalucía taking second place, with percentages bordering on 46 percent. There are no variations in the figures for the communities with the lowest percentages either, with Madrid reaching 28.5 percent and Cataluña amounting to 30.6 percent.

Incidents have been divided into three groups (**table 2**) for analytical purposes: framework incidents, incidents pertaining to the human group living in the dwelling and incidents pertaining to the selected person.

Table 2. Distribution of incidents

Type of incident	Incumbent persons		Substitutes	
	No.	%	No.	%
Total	8.332	100,00	11.570	100,00
Framework incidents	2.864	**34,37**	3.595	**31,07**
Non-surveyable dwelling	557	6,69	694	6,00
Person beyond the scope of study	17	0,20	18	0,16
Unreachable person	2.290	27,48	2.883	24,92
Human group incidents	1.165	**13,98**	2.070	**17,89**
Refusal	450	5,40	475	4,11
Absence	706	8,47	1.580	13,66
Inability to respond	9	0,11	15	0,13
Incidents regarding the selected person	4.303	**51,64**	5.905	**51,04**
Total refusal	2.195	26,34	2.830	24,46
Partial refusal	27	0,32	22	0,19
Refusal to use the computer	5	0,06	17	0,15
Absence	1.918	23,02	2.833	24,49
Inability to respond	158	1,90	203	1,75

Two different variables have been considered in the framework incidents classification: incidents that affect the dwelling located at the selected person's postal address and incidents that affect the person directly. Empty dwellings, dwellings used for other purposes, unreachable and inaccessible dwellings have been considered non-surveyable dwellings.

Of the three groups of incidents, the most relevant refers to incidents pertaining to the selected person, which amounts to almost 52 percent of the total. The second most important corresponds to the framework incidents, reaching 34 percent of the total, whilst events pertaining to the human group are the least important from a quantitative viewpoint, since they only account for 14 percent of the total.

Considering incidents regardless of the classification they belong to shows that the most relevant corresponds to *unreachable persons*, representing 27 percent of the total. *Total refusal* and *absence* of the selected person appear in second and third place, with 26 and 23 percent respectively.

At this point, it is important to note, especially if comparing these incidents with those in other surveys, that the high percentage of unreachable persons is not only caused by framework flaws. It is partly due to the fact that many of the persons are not really unreachable, but simply registered in their parents' house whilst not living there when the survey was performed since they were studying, working, etc. in a different location. These persons have not registered at a different address since they do not consider it a permanent dwelling. These cases have been considered unreachable at the postal address established for the selected person. This data shows that all three aforementioned incidents represent 76 percent of the total incidents, i.e. the majority.

Table 2bis analyses incidents corresponding to the selected person, all three types of refusals (total, partial and resistance about using the computer) are considered as *Refusal*. This shows that non-response thus defined is quite balanced between refusals (52 percent) and absences (45 percent), with a practically insignificant number of persons unable to respond, as occurs in other surveys.

TABLE 2bis. Distribution of non-response

Type of incident	Incumbent persons		Substitutes	
	No.	%	No.	%
Total	4.303	100,00	5.905	100,00
Refusal	2.227	51,75	2.869	48,59
Absence	1.918	44,57	2.833	47,98
Inability to respond	158	3,67	203	3,44

Before moving on to analyse the remaining tables, it is important to note that the total in table 2bis does not coincide with that of the subsequent tables (tables 3 to 7), since some assessment questionnaires have not been obtained. The latter tables have been prepared using information contained in the assessment questionnaires.

Likewise, it is important to note the existence of a high non-response percentage in the actual assessment questionnaire, given the aspects it aims to garner, especially in the case of the *economic activity* and *educational level*, as occurs in all surveys. In the tables this results in the fact that the total number of persons classified depending on the different characteristics researched is much lower than the quantity theoretically possible.

Table 3 analyses incidents according to the age and sex of the person selected. The non-response percentage is similar among men and women, albeit slightly superior for the former (52 and 48 percent, respectively). Data from the 2001 Census indicate that among the population aged between 18 and 49 years old,

50.6 are men and 49.4 are women. Thus, in view of these figures, the non-response rate is slightly concentrated among men.

Table 3. Distribution of non-response by sex and age of the selected person

Sex/age	Total		Type of incident						2001
			Refusal		Absence		Inab. to respond		Census
	No.	%	No.	%	No.	%	No.	%	%
Total	3.987	-	2.054	-	1.783	-	150	-	
No data recorded for sex	811	20,34	364	17,72	408	22,88	39	26,00	
Total classified	3.176	-	1.690	-	1.375	-	111	-	100,0
Men	1.651	51,98	789	46,69	802	58,33	60	54,05	50,6
No data recorded for age	394	12,41	165	9,76	220	16,00	9	8,11	-
Men classified by age	1.257	100,00	624	100,00	582	100,00	51	100,00	100,0
18 to 29 years old	455	36,20	170	27,24	267	45,88	18	35,29	39,1
30 to 39 years old	390	31,03	210	33,65	165	28,35	15	29,41	33,2
40 to 49 years old	412	32,78	244	39,10	150	25,77	18	35,29	27,7
Women	1.525	48,02	901	53,31	573	41,67	51	45,95	49,4
No data recorded for age	384	12,09	213	12,60	163	11,85	8	7,21	-
Women classified by age	1.141	100,00	688	100,00	410	100,00	43	100,00	100,0
18 to 29 years old	437	38,30	196	28,49	229	55,85	12	27,91	38,3
30 to 39 years old	352	30,85	237	34,45	102	24,88	13	30,23	33,2
40 to 49 years old	352	30,85	255	37,06	79	19,27	18	41,86	28,5

Analysing each type of incident shows that, whilst persons who are absent and unable to respond are mostly men, refusals predominate among women, with the greatest percentage difference appearing in the variable analysing absences (58 percent of the men and 42 percent of the women). Comparing these values with the aforementioned census data, it is possible to say that refusals tend to appear among women and absences or inabilities among men, with a greater rate of absences.

On analysing non-response in terms of the type of incident and age of the person, both sexes present the highest percentages of refusals in the *40 to 49 years old* age group, with more noticeable differences among the males.

If comparing percentages of refusals of males and females by age groups with the data registered in the 2001 Census, this incident is, in fact, concentrated in the section ranging *between 40 and 49 years old*, with greater intensity among the males.

As regards absences, the highest percentages appear for both sexes in the *18 to 29 years old* modality, with females ten points above males. On comparing these figures with the census values, the absences are, indeed, concentrated in this age group, especially among females.

Finally, given the low number of persons unable to respond, the data is not at all significant. Nevertheless, in males the highest percentage (35 percent) is shared by the following age groups: *18 to 29 years old* and *40 to 49 years old*. For females, the highest percentage (42 percent) appears in the *40 to 49 years old* age group. When comparing these figures and those included in the Census, this incident is concentrated in the *40 to 49 years old* age group, with greater intensity among females.

Non-response according to sex and marital status of the selected person is analysed in **table 4**.

Table 4. Distribution of non-response by sex and marital status of the selected person

Sex/marital status	Total		Type of incident					
			Refusal		Absence		Inab. to respond	
	No.	%	No.	%	No.	%	No.	%
Total	3.987	-	2.054	-	1.783	-	150	-
No data recorded for sex	811	20,34	364	17,72	408	22,88	39	26,00
Total classified	3.176	-	1.690	-	1.375	-	111	-
Men	1.651	51,98	789	46,69	802	58,33	60	54,05
No data recorded for marital status	691	21,76	318	18,82	353	25,67	20	18,02
Men classified by marital status	960	100,00	471	100,00	449	100,00	40	100,00
-Single	486	50,63	193	40,98	265	59,02	28	70,00
-Married	461	48,02	271	57,54	178	39,64	12	30,00
-Widowed	-	-	-	-	-	-	-	-
-Separated	10	1,04	6	1,27	4	0,89	-	-
-Divorced	3	0,31	1	0,21	2	0,45	-	-
Women	1.525	48,02	901	53,31	573	41,67	51	45,95
No data recorded for marital status	704	22,17	416	24,62	274	19,93	14	12,61
Women classified by marital status	821	100,00	485	100,00	299	100,00	37	100,00
-Single	369	44,95	161	33,20	194	64,88	14	37,84
-Married	440	53,59	318	65,57	102	34,11	20	54,05
-Widowed	3	0,37	1	0,21	-	-	2	5,41
-Separated	7	0,85	3	0,62	3	1,00	1	-
-Divorced	2	0,24	2	0,41	-	-	-	-

As regards refusals, the highest percentages appear among *married* persons, with the percentage for females eight points above that of males.

As regards absences, the highest percentages appear among single persons, also with a higher percentage (six points) among females.

In terms of inability to respond, despite the low relevance, the highest percentages appear among *married females* and *single males*.

Table 4.bis shows the percentage distribution of the three components of non-response according to the marital status of the selected person, compared to the distribution of persons aged 18 to 49 years old taken from the 2001 Census. Comparing these figures shows that refusals appear mainly among married persons, absences among single persons and inability to respond mainly appear among single persons and, to a lesser degree, among widowers.

Table 4.bis. Percentage distribution of non-response by marital status of the person selected. Comparison with 2001 Census

Marital status	Type of incident			2001 Census
	Refusal	Absence	Inability to respond	
Single	36,8	61,5	54,4	46,2
Married	61,6	37,3	40,5	49,2
Widowed	0,1	0,0	3,8	0,7
Separated	1,0	0,9	1,3	2,4
Divorced	0,5	0,3	0,0	1,4

As regards the economic activity of the selected person (**table 5.1**), most incidents affect persons who are *employed*. The other modalities appear a long way behind in the cases of refusals and absences. As regards inability to respond, the difference between the first and second modality (*another situation*) is quite a lot lower.

This table includes the percentage distribution of the population aged between 18 and 49 years old, according to economic activity, in line with data from the 2001 Census. Comparing these figures with those corresponding to incidents shows that refusals and absences appear among *employed* persons, whilst inabilities appear in the modality *another situation*.

Table 5.1 Distribution of non-response by economic activity of the selected person

Economic activity	Total		Type of incident						Censo 2001
			Refusal		Absence		Inab. to respond		
	No.	%	No.	%	No.	%	No.	%	%
Total	3.987	-	2.054	-	1.783	-	150	-	
No data recorded econ. Activ.	2.232	55,98	1.128	54,92	1.023	57,38	81	54,00	
Total classified	1.755	100,00	926	100,00	760	100,00	69	100,00	100,0
Employed	1.325	75,50	726	78,40	569	74,87	30	43,48	65,3
Unemployed	92	5,24	46	4,97	37	4,87	9	13,04	11,1
Retired or pensioner	10	0,57	6	0,65	3	0,39	1	1,45	1,7
Housework	99	5,64	81	8,75	9	1,18	9	13,04	10,5
Another situation	229	13,05	67	7,24	142	18,68	20	28,99	11,4

Table 5.2 presents a different perspective on these results, showing the weight of the different incidents in each of the modalities of the characteristic *economic activity*. Consequently, in all except *another situation*, the main incident is refusal, especially in the modality *Housework*, where it exceeds 80 percent. The *another situation* modality shows a greater incidence of absence, reaching 62 percent.

Table 5.2. Distribution of economic activity of selected person by type of incident

(Continúa)

Type of incident	Total		No data recorded		Total class.		Economic activity	
							Employed	
	No.	%	No.	%	No.	%	No.	%
Total	3.987	100,00	2.232	100,00	1.755	100,00	1.325	100,00
Refusal	2.054	51,52	1.128	50,54	926	52,76	726	54,79
Absence	1.783	44,72	1.023	45,83	760	43,30	569	42,94
Inab. to respond	150	3,76	81	3,63	69	3,93	30	2,26

(Conclusión)

Type of incident	Economic activity							
	Unemployed		Retired		Housework		Another sit.	
	No.	%	No.	%	No.	%	No.	%
Total	92	100,00	10	100,00	99	100,00	229	100,00
Refusal	46	50,00	6	60,00	81	81,82	67	29,26
Absence	37	40,22	3	30,00	9	9,09	142	62,01
Inab. to respond	9	9,78	1	10,00	9	9,09	20	8,73

Nevertheless, it is important to consider the different percentages of the persons classified in each modality, which vary, for non-response as a whole, from 0.6 for retired persons to 75 percent of employed persons. This can have a negative bearing on the validity of the comparisons.

Table 6 analyses the distribution of non-response considering the highest level of education achieved by the person selected. In terms of refusals, the highest percentages appear among persons whose educational level is *primary education*, as occurs in others surveys (EAPS, Household Budgets Survey, Inability Survey, Use of Time Survey). *First cycle of secondary education* appears in second place, not far behind. Conversely, for absences, the highest percentage appears among the latter educational level, although *primary education* and *second cycle of secondary education* are also relevant groups. As regards inabilities to respond, the highest percentage corresponds to *primary education*, although there is also a high percentage of *illiterates*. This seems logical considering that illiteracy is one of the grounds stated when *unable to answer*.

Table 6. Distribution of non-response by educational level of selected person

Educational level	Total		Type of incident						2001 Census
			Refusal		Absence		Inab. to respond		
	No.	%	No.	%	No.	%	No.	%	%
Total	3.987	-	2.054	-	1.783	-	150	-	
No data recorded for ed. Level	2.751	69,00	1.356	66,02	1.287	72,18	108	72,00	
Total classified	1.236	100,00	698	100,00	496	100,00	42	100,00	100,0
Illiterate	39	3,16	18	2,58	10	2,02	11	26,19	0,7
Primary education	319	25,81	193	27,65	109	21,98	17	40,48	18,8
1st cycle secondary education	309	25,00	172	24,64	132	26,61	5	11,90	31,4
Intermediate vocational training	119	9,63	71	10,17	46	9,27	2	4,76	6,5
2nd cycle secondary education	224	18,12	115	16,48	106	21,37	3	7,14	16,4
Advanced vocational training	90	7,28	55	7,88	32	6,45	3	7,14	7,4
University education	136	11,00	74	10,60	61	12,30	1	2,38	18,8

Considering the 2001 Census data as reference values shows the refusals appear mainly among persons with *primary education*; absences are more balanced, mainly among persons with *second cycle secondary education, primary education* and *intermediate vocational training*. The inability to respond prevails among *illiterate persons* and people with *primary education*.

Non-response by nationality of the person selected can be analysed using **table 7**, which shows that –as well as a considerable amount of *no data recorded*–, the majority have a Spanish nationality; foreigners only amount to 5 percent of the sample. Consequently, both refusals and absences mainly appear among persons with a Spanish nationality, with percentages above 96 percent. On the other hand, 36 percent of the inabilities to respond appear among persons with a foreign nationality. This could be explained by the fact that they do not understand Spanish, which is another of the grounds causing *inability to respond*.

Table 7. Distribution of non-response by nationality of the selected person

Nationality	Total		Type of incident						2001
			Refusal		Absence		Inab. to respond		Census
	No.	%	No.	%	No.	%	No.	%	%
Total	3.987	-	2.054	-	1.783	-	150	-	
No data recorded for natior	1.541	38,65	688	33,50	800	44,87	53	35,33	
Total classified	2.446	100,00	1.366	100,00	983	100,00	97	100,00	100,0
Spaniards	2.334	95,42	1.328	97,22	944	96,03	62	63,92	94,7
Foreigners	112	4,58	38	2,78	39	3,97	35	36,08	5,3

In the light of the 2001 Census data, it is possible to state that inabilities predominate among persons with a foreign nationality, most probably due to the fact that a great many of them do not speak Spanish, which leads to them being unable to respond.

Given the low number of persons with a foreign nationality, it was not considered relevant to analyse their country of origin. Furthermore, among these people, there is also a considerable number of *no data recorded*.

Tables 8 and 9 compare percentage distributions by level of education and economic activity, respectively, of the incumbent persons with non-response, of the substitutes and of the persons in the total effective sample.

In the first place, it is important to note that in **table 8** the total number of persons in the effective sample, 10,830, does not coincide with the number included in table 1, 10,980. This difference can be explained by the fact that for some of the people interviewed there is *no data recorded for educational level*. The fact that the number of substitutes is also greater than the number of incumbents with non-response is also notable. This can be explained by the fact that among the latter there is a very high percentage of *no data recorded for educational level* (69 percent), which does not apply to substitute households.

Table 8. Distribution of persons with non-response and substitutes by educational level

Educational level	Incumbent persons with non-response		Substitutes		Persons in the total effective sample	
	No.	%	No.	%	No.	%
Total classified	1.247	100,00	5.640	100,00	10.830	100,00
Illiterate	39	3,13	-	-	-	-
Primary education	323	25,90	1.327	23,51	2.603	24,04
1st cycle secondary education	311	24,94	920	16,30	1.830	16,90
Intermediate vocational training	119	9,54	680	12,05	1.274	11,76
2nd cycle secondary education	226	18,12	952	16,87	1.784	16,47
Advanced vocational training	90	7,22	533	9,44	1.047	9,67
University education	139	11,15	1.228	21,76	2.292	21,16

After clarifying this point, although the differences between the three distributions are not too large, they are significant, with the main differences, around 10 percent, appearing in *university education* and *first cycle secondary education* between the distribution of incumbent households and the other two. In practice, persons with a low/average educational level are replaced by persons with a high educational level, especially persons with *first cycle secondary education* are replaced by persons with *university education*. Furthermore, illiterate persons have been removed from the effective sample, since illiteracy, as aforementioned, causes *inability to respond.*

The lowest differences, for all modalities, appear in the distribution of substitutes and the distribution of the total effective sample, probably due to the fact that the former are a subgroup of the latter, with a very similar behaviour.

Finally, comparing the distribution of persons by educational of level in the 2001 Census (table 6) with the data from the total effective sample of the survey (table 8) shows that both offer slightly discordant figures. The greatest difference appears in the *first cycle secondary education* modality: the Census presents a percentage fourteen points higher than the survey. This would mean that, considering the census data as the correct information, the survey would

underestimate persons with said educational level and slightly overestimate persons with *primary education* and *intermediate vocational training*.

An analysis of the economic activity (**table 9**) shows that, as in table 8, the number of persons in the total effective sample does not coincide with the figure in table 1. This can be explained as before, when referring specifically to said table. The difference between the number of persons with non-response and the number of substitutes can also be explained by the reasons given for table 8.

Table 9. Distribution of persons with non-response and substitutes by economic activity

Economic activity	Incumbent persons with non-response		Substitutes		Persons in the total effective sample	
	No.	%	No.	%	No.	%
Total classified	1.771	100,00	5.632	100,00	10.807	100,00
Employed	1.331	75,16	3.847	68,31	7.436	68,81
Unemployed	92	5,19	649	11,52	1.242	11,49
Retired or pensioner	12	0,68	116	2,06	215	1,99
Housework	99	5,59	482	8,56	878	8,12
Another situation	237	13,38	538	9,55	1.036	9,59

In table 9, distributions of substitutes and persons in the total effective sample are very similar (as occurs with the educational level), whilst the differences between both distributions and that of incumbent households with non-response are greater, although fewer than in the educational level section. The most noticeable differences appear among employed and unemployed persons, with values reaching 6 percent.

In the effective sample, the percentage of retired persons, housework and, especially, the unemployed, has increased at the expense of persons in another situation and employed persons.

Comparing the distribution of persons according to the economic activity in the 2001 Census (see table 5.1) with that obtained from the total effective sample of the survey (table 9) shows that both offer similar figures. The greatest difference appears in the *employed persons* modality, in which the percentage appearing in the Census is only 3.5 points below that considered in the survey. Thus, one could say that the survey offers quite good estimates.

3 Estimate of the correction coefficient for differential non-response due to nationality

The correction coefficient for differential non-response measures the different behaviour of the groups of sample elements in terms of non-response. Specifically, it is the quotient of the reverse of the probability of response in each of the groups. If it is close to one, both groups have a similar behaviour. Values above one represent higher non-response in the numerator group, and values below one indicate greater non-response among the denominator group.

In order to perform the estimate, the theoretical sample of persons has been broken down to indicate persons interviewed (effective sample) and incidents. The latter have been structured in terms of framework incidents and incidents concerning people. The latter includes those linked to the human group living in the dwelling and to the person selected. Only incumbent persons have been considered in the effective sample and in the incidents, substitutes have been disregarded.

The initial approach was to separate persons, both interviewed and affected by an incident, into two groups:

- Extra-community citizens, in other words persons from outside the EU.
- Non extra-community citizens, which were, in turn, divided into two subgroups:
  - Spaniards
  - Persons from the EC who are not Spanish, henceforth Community citizens.

The previous breakdown has been prepared using the *country of nationality* stated in the Register. This information has not been included for one of the persons considered in the survey, since it was impossible to locate him/her in the Register.

Horizontal percentages (regarding the total number of persons in the theoretical sample with nationality, the total for each type of incident and the effective sample) and **vertical percentages** (regarding the theoretical sample with nationality in each group of persons) have been calculated both for extra-Community and non-Community citizens.

The estimate of the differential non-response correction coefficient has been calculated considering the theoretical sample in four different manners:

- Including all data: theoretical sample = effective sample + all incidents
- With refusals: theoretical sample = effective sample + refusals
- With absences: theoretical sample = effective sample + absences
- With refusals and absences: theoretical sample = effective sample + refusals + absences

Table 10 garners the results obtained showing that, in the first place, extra-Community citizens only represent 7.5 percent of the total number of persons in the theoretical sample which have stated their nationality. The percentage is even lower for Community citizens, positioned at 1.3 percent.

It is also worth noting that:

- The percentage of empty dwellings is more prominent among dwellings inhabited by Community citizens (11.7 percent) than among dwellings accommodating extra-Community citizens (4.3 percent) and Spaniards (2.2 percent).
- The same occurs with *unreachable dwelling*, although percentages are somewhat lower.
- As regards *unreachable persons*, the highest percentages appear among extra-Community citizens (41.1 percent) than among Community citizens (28.7 percent) and Spaniards (14.7 percent).
- The percentage of refusals is a lot higher among Spaniards (21 percent) than among Community citizens (8.8 percent) and extra-Community citizens (6.1 percent).
- Absences are more prevalent among Community citizens (20.5 percent) and Spaniards (19.7 percent) than among extra-Community citizens (14.2 percent).
- The highest inability to respond percentages appear among Community citizens and extra-Community citizens (5.8 percent in each case), whilst the figure is irrelevant in the case of the Spaniards (0.8 percent). These differences are probably produced by language differences between people with foreign nationalities.
- The high level of incidents results in low percentages of persons interviewed, especially in terms of Community citizens, which hardly amount to 16 percent. The highest percentage has been obtained among Spaniards, which only amount to 40 percent. The number of extra-Community citizens amounts to 25 percent.
- The refusals ratios calculated show significant differences, which is not the case for refusals and absences considered jointly.
- Regarding the estimate of the differential non-response correction coefficient, the further it strays from the unit, the more incidents are considered, attaining a value of 1.60 percent. This is due to the tremendous relevance of the *unreachable person* incident among extra-Community citizens.

4 Conclusion

Summarising the previous paragraphs, the population group that presented the greatest number of incidents comprised employed persons aged from 40 to 49 years old, who are single or married, and have received *primary education.*

Table 10. Survey on health and sexual behaviour

(**Nationality** has been obtained from the **Register for the whole sample)**

Persons	Total	Extracommunity citizens			Non-extracommunity citizens								
					Total			Community citizens			Spaniards		
Theoretical sample (incumbent persons)	13.600		% hor.	% vert.		% hor.	% vert.		% hor.	% vert.		% hor.	% vert.
-Do not appear in Register	1												
Theoretical sample with nationality	13.599	1.018	7,5		12.581	92,5		171	1,3		12.410	91,3	
-Incidents	8.336	764			7.572			144			7.428		
Framework:- Empty dwellings	334	44	13,2	4,3	290	86,8	2,3	20	6,0	11,7	270	80,8	2,2
- Dwell. used other purposes	34	8	23,5	0,8	26	76,5	0,2	1	2,9	0,6	25	73,5	0,2
- Inaccessible dwelling	17	5	29,4	0,5	12	70,6	0,1	0	0,0	0,0	12	70,6	0,1
- Unreachable dwelling	172	21	12,2	2,1	151	87,8	1,2	14	8,1	8,2	137	79,7	1,1
- Persons beyond scope of study	17	2	11,8	0,2	15	88,2	0,1	0	0,0	0,0	15	88,2	0,1
- Unreachable persons	2.293	418	18,2	41,1	1.875	81,8	14,9	49	2,1	28,7	1.826	79,6	14,7
Persons:													
-Absences	2.624	145	5,5	14,2	2.479	94,5	19,7	35	1,3	20,5	2.444	93,1	19,7
-Refusals	2.678	62	2,3	6,1	2.616	97,7	20,8	15	0,6	8,8	2.601	97,1	21,0
-Inability to respond	167	59	35,3	5,8	108	64,7	0,9	10	6,0	5,8	98	58,7	0,8
-Interviewed (effective sample)	5.263	254	4,8	25,0	5.009	95,2	39,8	27	0,5	15,8	4.982	94,7	40,1

		%			
Ref. extracom. theoretical sample	62	19,6	Ref. +Abs. extracom. theoretical sample	207	44,9
Total extracom. (eff. sample +ref.)	316		Total extracom. (eff. sample +ref. +abs.)	461	
Ref. rest theoretical sample	2.616	34,3	Ref. +Abs. rest theoretical sample	5.095	50,4
Total rest (eff. sample +ref.)	7.625		Total rest (eff. sample +ref. +abs.)	10.104	

COEFFICIENT ESTIMATE (NON-RESPONSE DIFFERENTIAL TENDENCY)

	Including all data	Including refusals	Including absences	Including refusals and absences
Pxt (extracom. citiz. theor. sample)	1.018	316	399	461
Pxr(extracom. citiz. eff. sample)	254	254	254	254
Pyt(non extracom. citiz. theor. sample)	12.581	7.625	7.488	10.104
Pyr(non extracom. citiz. eff. sample)	5.009	5.009	5.009	5.009
Value estimate(Pxt/Pxr)/(Pyt/Pyr)	1,60	0,82	1,05	0,90