# Assessment of non-response in *Survey on Adult Population Involvement in Learning Activities 2007* (AES 2007)

## 1    Introduction

The errors that affect the whole survey can be classified in two large groups: Sample errors and non-sampling errors. The former can be estimated 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 occur before information is garnered (deficient framework, insufficient definitions or questionnaires, etc.), during the collection of information (incorrect fieldwork by the interviewers, incorrect statements or non-response on behalf of respondents) and, lastly, in operations subsequent to the fieldwork (errors when encoding variables, recording questionnaires, etc.).

As stated previously, the evaluation of these errors presents many difficulties, among other reasons, due to the great variety of causes that may lead to them.

Among these causes, one of the most notable is the **non-response of the respondent units**, which may be caused by a refusal to answer the questionnaire, the absence of the same, the inability to answer of the whole household comprising the respondent unit, or the dwelling being inaccessible at the time of the interview.

In the *Survey on Adult Population Involvement in Learning Activities*, the respondent units are those persons 25 to 74 years of age.

An **assessment questionnaire** has been designed to analyse the non-response to the 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 completely only by those **incumbent persons** who, due to a refusal, absence or inability to answer, have not cooperated in the survey. It will not be completed for the reserve persons whose participation has been requested to replace an incumbent person due to a certain event, but who did not complete the questionnaire either.

This questionnaire consists of three sections. In the first, the identification data of the person is entered, this person being the respondent unit. The second enters the type of incidence that has occurred, if the incumbent person has been replaced or not, and the order number of the substitute person if s/he has been substituted. Lastly, the third section serves to enter a series of basic data of the person: sex, age, marital status, highest level of education achieved, situation with regard to activity, nationality, and country of birth in the case of having been born outside of Spain.

When an incumbent person has to be replaced, the interviewers question as many reserve persons as are required to find a person who will collaborate. This person is called the substitute person, and is given a corresponding order number that will be entered in the assessment questionnaire.

If all of the reserve persons have been requested, and it has not been possible to replace the incumbent person with an incidence, a loss in the sample is generated, and therefore, a decrease in the precision of the estimators. In this case, the order number of the substitute person remains blank.

## 2. Analysis of the data

**Table 1** presents the distributions, by Autonomous Community, of the theoretical sample of persons, the total effective sample (total number of persons surveyed) and the effective sample of incumbent persons (total number of incumbent persons surveyed).

### TABLE 1. Distribution of the theoretical and effective sample of persons by Autonomous Community

| Autonomous Communities | Theoretical sample | | Effective sample | | | |
| | | | Total | | Incumbent persons | |
| | Persons | % | Persons | % | Persons | % |
| --- | --- | --- | --- | --- | --- | --- |
| Total | 24,030 | 100.00 | 20,009 | 83.27 | 12,741 | 53.02 |
| Andalucía | 3,250 | 100.00 | 2,614 | 80.43 | 1,691 | 52.03 |
| Aragón | 950 | 100.00 | 783 | 82.42 | 513 | 54.00 |
| Asturias (Principado de) | 890 | 100.00 | 796 | 89.44 | 552 | 62.02 |
| Balears (Illes) | 870 | 100.00 | 627 | 72.07 | 336 | 38.62 |
| Canarias | 1,130 | 100.00 | 929 | 82.21 | 583 | 51.59 |
| Cantabria | 750 | 100.00 | 653 | 87.07 | 422 | 56.27 |
| Castilla y León | 1,270 | 100.00 | 1,136 | 89.45 | 788 | 62.05 |
| Castilla-La Mancha | 1,120 | 100.00 | 987 | 88.13 | 628 | 56.07 |
| Cataluña | 2,910 | 100.00 | 2,349 | 80.72 | 1,418 | 48.73 |
| Comunidad Valenciana | 1,960 | 100.00 | 1,496 | 76.33 | 953 | 48.62 |
| Extremadura | 890 | 100.00 | 743 | 83.48 | 483 | 54.27 |
| Galicia | 1,350 | 100.00 | 1,259 | 93.26 | 897 | 66.44 |
| Madrid (Comunidad de) | 2,480 | 100.00 | 1,948 | 78.55 | 1,104 | 44.52 |
| Murcia (Región de) | 960 | 100.00 | 807 | 84.06 | 524 | 54.58 |
| Navarra (Comun. Foral de) | 760 | 100.00 | 683 | 89.87 | 453 | 59.61 |
| País Vasco | 1,170 | 100.00 | 1,110 | 94.87 | 715 | 61.11 |
| Rioja (La) | 680 | 100.00 | 606 | 89.12 | 420 | 61.76 |
| Ceuta and Melilla | 640 | 100.00 | 483 | 75.47 | 261 | 40.78 |

It can be observed that, on a national level, the total effective sample represents 83% of the theoretical sample, whereas the effective sample of incumbent persons represents only 53% of the same. These figures indicate that, out of the total incumbent persons, it was only possible to survey 53%. Likewise, it was possible to replace 30% of the incumbent persons who presented some incidence, thus elevating the total effective sample to 83% of the previously mentioned theoretical sample.

This percentage of total effective sample (83%) may be considered as a **response rate** in the survey, given that it represents the percentage of persons who have been surveyed in practice (including substitutions), as compared with those that, in theory, should have been surveyed.

Dropping to the Autonomous Community level, we observe that there are only two Communities, País Vasco and Galicia, with percentages of total effective sample that are higher than 90%, the latter worth noting with 95%. At the other end of the spectrum is Illes Balears, with 72%.

On considering now the percentage of effective sample of incumbent persons, the situation is similar, granted that Galicia is the Community with the highest percentage, 66%, and Illes Balears has the lowest percentage, with nearly 39%.

Incidences have been distributed into three groups(**Table 2**), for analytical purposes: framework incidences, incidences of the household resident in the dwelling, and incidences of the selected person. This table shows the distribution of the incidences, both for the incumbent persons and for the reserve persons.

Within the group of framework incidences, this has considered, separately, those that affect the dwelling that appears in the postal address of the selected person, and those that directly affect the person. Unsurveyable dwellings have been considered to be empty dwellings, those dedicated to other purposes, and unlocatable dwellings.

It can be observed that, of the three groups of incidences, the group that carries the greatest weight, in the case of incumbent persons, is that of incidences of the selected person, which represents 38% of the total of the same. The second most important group corresponds to the group of framework incidences, which accounts for nearly 36% of the total, with the incidences in the household being the least quantitatively important, representing only 26%.

Regarding the incidences in the reserve persons, the distribution is similar, though in this case, the percentages of the three groups of incidences are more alike.

On analysing the incidences individually, irregardless of the group to which they belong, we see that, for the incumbent persons, the most numerous is that of the *unlocatable person*, which represents 25.6% of the total incidences. The second and third most important places correspond to the absence from the household, and the refusal of the selected person, with percentages of 20.4% and 19%, respectively. In the reserve persons, the absence from the household corresponds to a percentage that is slightly higher than that of the unlocatable

person (23.5% as compared with 22.8%), with the third most important corresponding, likewise, to the refusal of the selected person, with 17.8%.

## TABLE 2. Distribution of incidences

| Type of incidence | Incumbent persons | | Reserve persons | |
|---|---|---|---|---|
| | Nº | % | Nº | % |
| Total | 24,030 | | 14,708 | |
| Surveyed persons | 12,741 | **53.02** | 7,268 | **49.42** |
| Total incidences | 11,289 | 100.00 | 7,440 | 100.00 |
| **Framework incidences** | 4,026 | **35.66** | 2,461 | **33.08** |
| Unsurveyable dwelling | 1,049 | 9.29 | 725 | 9.74 |
| - Empty dwelling | 534 | 4.73 | 342 | 4.60 |
| - Unlocatable dwelling | 447 | 3.96 | 344 | 4.62 |
| - Dwelling used for other purposes | 68 | 0.60 | 39 | 0.52 |
| Unaccessible dwelling | 36 | 0.32 | 21 | 0.28 |
| Person outside of the study scope | 48 | 0.43 | 16 | 0.22 |
| Persona unlocatable | 2,893 | 25.63 | 1,699 | 22.84 |
| **Incidences in the household** | 2,967 | **26.28** | 2,323 | **31.22** |
| Refusal | 651 | 5.77 | 570 | 7.66 |
| Absence | 2,299 | 20.36 | 1,745 | 23.45 |
| Inability to answer | 17 | 0.15 | 8 | 0.11 |
| **Incidences in the selected person** | 4,296 | **38.05** | 2,656 | **35.70** |
| Refusal | 2,142 | 18.97 | 1,327 | 17.84 |
| Absence | 2,010 | 17.80 | 1,245 | 16.73 |
| Inability to answer | 144 | 1.28 | 84 | 1.13 |

At this point, it is important to state, above all if we compare it with the incidences from other surveys, that the elevated percentage of unlocatable persons is not completely attributable to framework defects. It is due, in part, not to the persons who have truly been unlocatable, but rather to the fact that they are registered in the parents' dwelling, but at the time of the survey, are outside of that domicile due to work, studies, etc. These persons are not registered at a different address, as they do not consider it a definitive location. These cases have been considered unlocatable at the postal address established for the selected person.

**Table 2bis** analyses the distribution of the non-response itself, that is, excluding framework incidences. It can be observed that, overall, both for incumbent persons and for reserve persons, the incidences in the selected person carries a

greater weight than the incidences of the human group, although individually, the most quantitatively important group is that of absence of the household.

Within the incidences of the selected person, the percentages of refusals and absences are very similar, slightly higher in refusals, with the number of inabilities to answer being practically insignificant, as occurs in other surveys.

## TABLE 2bis. Distribution of non-response

| Type of incidence | Incumbent persons | | Reserve persons | |
|---|---|---|---|---|
| | No. | % | No. | % |
| Total | 7,263 | 100.00 | 4,979 | 100.00 |
| **Incidences in the household** | **2,967** | **40.85** | **2,323** | **46.66** |
| Refusal | 651 | 8.96 | 570 | 11.45 |
| Absence | 2,299 | 31.65 | 1,745 | 35.05 |
| Inability to answer | 17 | 0.23 | 8 | 0.16 |
| **Incidences in the selected person** | **4,296** | **59.15** | **2,656** | **53.34** |
| Refusal | 2,142 | 29.49 | 1,327 | 26.65 |
| Absence | 2,010 | 27.67 | 1,245 | 25.01 |
| Inability to answer | 144 | 1.98 | 84 | 1.69 |

## TABLE 3.1 Distribution of the incumbent persons surveyed and the framework incidences by Autonomous Community

| Autonomous Communities | Incumbent persons | | | | | | | |
| | Total | | Surveyed | | With incidences | | | |
| | | | | | Total | | Framework inciden | |
| | Nº | % | Nº | % | Nº | % | Nº | % |
|---|---|---|---|---|---|---|---|---|
| Total | 24,030 | 100.00 | 12,741 | 53.02 | 11,289 | 46.98 | 4,026 | 16.75 |
| Andalucía | 3,250 | 100.00 | 1,691 | 52.03 | 1,559 | 47.97 | 626 | 19.26 |
| Aragón | 950 | 100.00 | 513 | 54.00 | 437 | 46.00 | 131 | 13.79 |
| Asturias (Principado de) | 890 | 100.00 | 552 | 62.02 | 338 | 37.98 | 133 | 14.94 |
| Balears (Illes) | 870 | 100.00 | 336 | 38.62 | 534 | 61.38 | 226 | 25.98 |
| Canarias | 1,130 | 100.00 | 583 | 51.59 | 547 | 48.41 | 256 | 22.65 |
| Cantabria | 750 | 100.00 | 422 | 56.27 | 328 | 43.73 | 132 | 17.60 |
| Castilla y León | 1,270 | 100.00 | 788 | 62.05 | 482 | 37.95 | 163 | 12.83 |
| Castilla-La Mancha | 1,120 | 100.00 | 628 | 56.07 | 492 | 43.93 | 149 | 13.30 |
| Cataluña | 2,910 | 100.00 | 1,418 | 48.73 | 1,492 | 51.27 | 431 | 14.81 |
| Comunidad Valenciana | 1,960 | 100.00 | 953 | 48.62 | 1,007 | 51.38 | 297 | 15.15 |
| Extremadura | 890 | 100.00 | 483 | 54.27 | 407 | 45.73 | 131 | 14.72 |
| Galicia | 1,350 | 100.00 | 897 | 66.44 | 453 | 33.56 | 194 | 14.37 |
| Madrid (Comunidad de) | 2,480 | 100.00 | 1,104 | 44.52 | 1,376 | 55.48 | 488 | 19.68 |
| Murcia (Región de) | 960 | 100.00 | 524 | 54.58 | 436 | 45.42 | 144 | 15.00 |
| Navarra (Comunidad Foral de) | 760 | 100.00 | 453 | 59.61 | 307 | 40.39 | 133 | 17.50 |
| País Vasco | 1,170 | 100.00 | 715 | 61.11 | 455 | 38.89 | 116 | 9.91 |
| Rioja (La) | 680 | 100.00 | 420 | 61.76 | 260 | 38.24 | 114 | 16.76 |
| Ceuta and Melilla | 640 | 100.00 | 261 | 40.78 | 379 | 59.22 | 162 | 25.31 |

Table 3.1 enables assessing the defects of the framework, whereas Table 3.2 presents the non-response due to refusals, absences or inabilities to answer, either in the selected persons or in the households of which they are a part. Both tables present a breakdown by Autonomous Community.

The **framework incidences** on a national level (**Table 3.1**) represent almost 17% of the theoretical sample. On an Autonomous Community level, the percentages vary between almost 10% for País Vasco and 26% corresponding to Illes Balears.

**TABLE 3.2 Distribution of the surveyable incumbent persons by Autonomous Community**

| Autonomous Communities | Surveyable incumbent persons | | | | | | | |
| | Total | | Surveyed | | Non-response | | | |
| | | | | | Refusals | | | |
| | | | | | of the household | | of the selected person | |
| | No. | % | No. | % | No. | % | No. | % |
| Total | 20,004 | 100.00 | 12,741 | 63.69 | 651 | 3.25 | 2,142 | 10.71 |
| Andalucía | 2,624 | 100.00 | 1,691 | 64.44 | 151 | 5.75 | 319 | 12.16 |
| Aragón | 819 | 100.00 | 513 | 62.64 | 24 | 2.93 | 94 | 11.48 |
| Asturias (Principado de) | 757 | 100.00 | 552 | 72.92 | 10 | 1.32 | 61 | 8.06 |
| Balears (Illes) | 644 | 100.00 | 336 | 52.17 | 15 | 2.33 | 85 | 13.20 |
| Canarias | 874 | 100.00 | 583 | 66.70 | 9 | 1.03 | 71 | 8.12 |
| Cantabria | 618 | 100.00 | 422 | 68.28 | 22 | 3.56 | 37 | 5.99 |
| Castilla y León | 1,107 | 100.00 | 788 | 71.18 | 20 | 1.81 | 124 | 11.20 |
| Castilla-La Mancha | 971 | 100.00 | 628 | 64.68 | 78 | 8.03 | 58 | 5.97 |
| Cataluña | 2,479 | 100.00 | 1,418 | 57.20 | 46 | 1.86 | 278 | 11.21 |
| Comunidad Valenciana | 1,663 | 100.00 | 953 | 57.31 | 37 | 2.22 | 226 | 13.59 |
| Extremadura | 759 | 100.00 | 483 | 63.64 | 11 | 1.45 | 74 | 9.75 |
| Galicia | 1,156 | 100.00 | 897 | 77.60 | 12 | 1.04 | 100 | 8.65 |
| Madrid (Comunidad de) | 1,992 | 100.00 | 1,104 | 55.42 | 140 | 7.03 | 294 | 14.76 |
| Murcia (Región de) | 816 | 100.00 | 524 | 64.22 | 14 | 1.72 | 47 | 5.76 |
| Navarra (Comunidad Foral de) | 627 | 100.00 | 453 | 72.25 | 5 | 0.80 | 85 | 13.56 |
| País Vasco | 1,054 | 100.00 | 715 | 67.84 | 31 | 2.94 | 116 | 11.01 |
| Rioja (La) | 566 | 100.00 | 420 | 74.20 | 8 | 1.41 | 26 | 4.59 |
| Ceuta and Melilla | 478 | 100.00 | 261 | 54.60 | 18 | 3.77 | 47 | 9.83 |

| Autonomous Communities | Surveyable incumbent persons | | | | | | | | | |
| | Non-response | | | | | | | | | |
| | Absences | | | | Inabilities to answer | | | | Total | |
| | of the household | | of the selected person | | | of the household | | of the selected person | | |
| | No. | % | No. | % | | No. | % | No. | % | No. | % |
| Total | 2,299 | 11.49 | 2,010 | 10.05 | 551 | 17 | 0.08 | 144 | 0.72 | 7,263 | 36.31 |
| Andalucía | 242 | 9.22 | 196 | 7.47 | 117 | 5 | 0.19 | 20 | 0.76 | 933 | 35.56 |
| Aragón | 129 | 15.75 | 56 | 6.84 | 29 | 0 | 0.00 | 3 | 0.37 | 306 | 37.36 |
| Asturias (Principado de) | 78 | 10.30 | 47 | 6.21 | 29 | 0 | 0.00 | 9 | 1.19 | 205 | 27.08 |
| Balears (Illes) | 134 | 20.81 | 64 | 9.94 | 23 | 1 | 0.16 | 9 | 1.40 | 308 | 47.83 |
| Canarias | 82 | 9.38 | 119 | 13.62 | 11 | 2 | 0.23 | 8 | 0.92 | 291 | 33.30 |
| Cantabria | 101 | 16.34 | 32 | 5.18 | 10 | 0 | 0.00 | 4 | 0.65 | 196 | 31.72 |
| Castilla y León | 98 | 8.85 | 66 | 5.96 | 15 | 0 | 0.00 | 11 | 0.99 | 319 | 28.82 |
| Castilla-La Mancha | 118 | 12.15 | 86 | 8.86 | 13 | 0 | 0.00 | 3 | 0.31 | 343 | 35.32 |
| Cataluña | 440 | 17.75 | 274 | 11.05 | 47 | 1 | 0.04 | 22 | 0.89 | 1,061 | 42.80 |
| Comunidad Valenciana | 170 | 10.22 | 259 | 15.57 | 53 | 5 | 0.30 | 13 | 0.78 | 710 | 42.69 |
| Extremadura | 61 | 8.04 | 124 | 16.34 | 23 | 1 | 0.13 | 5 | 0.66 | 276 | 36.36 |
| Galicia | 61 | 5.28 | 79 | 6.83 | 6 | 0 | 0.00 | 7 | 0.61 | 259 | 22.40 |
| Madrid (Comunidad de) | 190 | 9.54 | 249 | 12.50 | 87 | 1 | 0.05 | 14 | 0.70 | 888 | 44.58 |
| Murcia (Región de) | 150 | 18.38 | 77 | 9.44 | 8 | 0 | 0.00 | 4 | 0.49 | 292 | 35.78 |
| Navarra (Comunidad Foral de) | 14 | 2.23 | 67 | 10.69 | 8 | 1 | 0.16 | 2 | 0.32 | 174 | 27.75 |
| País Vasco | 117 | 11.10 | 72 | 6.83 | 55 | 0 | 0.00 | 3 | 0.28 | 339 | 32.16 |
| Rioja (La) | 51 | 9.01 | 58 | 10.25 | 6 | 0 | 0.00 | 3 | 0.53 | 146 | 25.80 |
| Ceuta and Melilla | 63 | 13.18 | 85 | 17.78 | 11 | 0 | 0.00 | 4 | 0.84 | 217 | 45.40 |

In **Table 3.2**, the percentages are calculated referring to the surveyable persons, that is, excluding the framework incidences; thus, the percentages of the **surveyable persons** are different (greater, as they are now calculated with regard to a lesser quantity) from those that appear in Table 3.1. On a national level, this percentage reached a value of 64%, and by Autonomous Community, worth noting with the highest percentage was Galicia, with almost 78%, and on the other extreme, Illes Balears, with 52%.

Consistent with the above, and continuing with the analysis of Table 3.2, it may be observed that the highest percentage of **total non-response** corresponds to Illes Balears, with almost 48%, whereas Galicia is the Community with the lowest percentage, 22.4%.

After breaking down non-response into its components, the highest percentage of **refusals of the selected person**, the most numerous of the two existing types, reached a value of 14.8% in Comunidad de Madrid, while La Rioja recorded the lowest percentage, with 4.6%. As per **absences**, the highest percentage was in those **of the household** corresponded to Illes Balears, which stood at almost 21%, whilst the lowest percentage corresponded to Galicia, with 5.3%. In turn, in the absences **of the selected person**, the highest percentage was from Ceuta and Melilla, nearly 18%, and the lowest percentage corresponded to Cantabria, with somewhat more than 5%.

The **inabilities to answer** carried so little weight in non-response that they deserve no mention at all.

Regarding the analysis of non-response according to the characteristics of the selected persons, worth noting is that the extremely high non-response in the assessment questionnaires themselves has made it impossible to use them for such purpose, having had to turn to the Municipal Register to obtain said characteristics. Given that this administrative register does not garner the relationship of persons with activity, it has not been possible to compile a table on the distribution of non-response according to this characteristic.

**Table 4** analyses the distribution of non-response and of the persons surveyed in the theoretical survey, according to the sex and age of the selected person. The information in this table has been taken directly from the selected sample, granted that it was included in the same upon making the selection from the Register. This table and the following two (Tables 5 and 6) consider the theoretical sample to be the sum of surveyed persons and non-response, that is, they do not include the framework incidences in the same. The percentages in these three tables are calculated in comparison with the total of the theoretical sample thus considered.

Firstly, non-response represents 36% of the total of the theoretical sample. Analysing it separately for the two sexes, it may be observed that it is slightly higher in men than in women (38% as compared with 34%). The percentages of refusals are practically the same for the totals of men and women (around 14%), and distinguishing by age bracket, the highest are reached, for both sexes, in the modality from *66 to 74 years of age*, standing near 21%, and being slightly higher in the case of men. In the case of absences, the percentages are slightly

higher for men than for women (four points on average), reaching the highest value among men, 26%, in the modality from *56 to 65 years of age*, and among women in the modality from *66 to 74 years of age* (almost 22%). The inabilities to answer, with such a small minority, deserve no mention at all.

**TABLE 4. Distribution of non-response and surveyed persons in the theoretical sample, by sex and age of the selected person**

| Sex/Age | Theoretical sample | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total | | Surveyed persons | | Non-response | | | | | | | |
| | | | | | Total | | Refusals | | Absences | | Inability to answer | |
| | No. | % | No. | % | No. | % | No. | % | No. | % | No. | % |
| Total classified | 20,004 | 100.00 | 12,741 | 63.69 | 7,263 | 36.31 | 2,793 | 13.96 | 4,309 | 21.54 | 161 | 0.80 |
| Men | 9,904 | 100.00 | 6,104 | 61.63 | 3,800 | 38.37 | 1,381 | 13.94 | 2,333 | 23.56 | 86 | 0.87 |
| 25 to 35 years of age | 2,300 | 100.00 | 1,535 | 66.74 | 765 | 33.26 | 237 | 10.30 | 524 | 22.78 | 4 | 0.17 |
| 36 to 45 years of age | 2,241 | 100.00 | 1,508 | 67.29 | 733 | 32.71 | 222 | 9.91 | 497 | 22.18 | 14 | 0.62 |
| 46 to 55 years of age | 2,082 | 100.00 | 1,278 | 61.38 | 804 | 38.62 | 304 | 14.60 | 485 | 23.29 | 15 | 0.72 |
| 56 to 65 years of age | 1,842 | 100.00 | 1,041 | 56.51 | 801 | 43.49 | 308 | 16.72 | 478 | 25.95 | 15 | 0.81 |
| 66 to 74 years of age | 1,439 | 100.00 | 742 | 51.56 | 697 | 48.44 | 310 | 21.54 | 349 | 24.25 | 38 | 2.64 |
| Women | 10,100 | 100.00 | 6,637 | 65.71 | 3,463 | 34.29 | 1,412 | 13.98 | 1,976 | 19.56 | 75 | 0.74 |
| 25 to 35 years of age | 2,325 | 100.00 | 1,660 | 71.40 | 665 | 28.60 | 202 | 8.69 | 453 | 19.48 | 10 | 0.43 |
| 36 to 45 years of age | 2,261 | 100.00 | 1,611 | 71.25 | 650 | 28.75 | 237 | 10.48 | 406 | 17.96 | 7 | 0.31 |
| 46 to 55 years of age | 2,021 | 100.00 | 1,351 | 66.85 | 670 | 33.15 | 282 | 13.95 | 378 | 18.70 | 10 | 0.49 |
| 56 to 65 years of age | 1,833 | 100.00 | 1,097 | 59.85 | 736 | 40.15 | 344 | 18.77 | 376 | 20.51 | 16 | 0.87 |
| 66 to 74 years of age | 1,660 | 100.00 | 918 | 55.30 | 742 | 44.70 | 347 | 20.90 | 363 | 21.87 | 32 | 1.93 |

The information on non-response in Tables 5 and 6 (educational level and nationality) has had to be obtained from the Register, through a cross in files with the same, given that it was not introduced in the sample on carrying out the selection of the same. Due to the difficulty in performing these crosses, there has been a certain number of persons who have not been locatable in the Register, thus the difference between the totals from Table 4 and the totals from Tables 5 and 6 as regards non-response, and therefore, as regards the total theoretical sample.

The analysis of non-response in the theoretical sample, according to the highest level of education achieved by the selected person, may be carried out from the data in **Table 5**. In this table, the persons have been classified into the four large groups of educational level used in the Register, as doing so on a greater breakdown level could put at risk the comparison of data from the Register with data from the survey, given the encoding system used in this administrative register, which in many cases does not allow for ascertaining the specific educational level that corresponds to each person, being a very general classification. The distribution of the persons surveyed has been obtained from the information in the survey itself, as with Table 4.

**TABLE 5. Distribution of non-response and surveyed persons in the theoretical sample, by educational level of the selected person**

| Educational level | Theoretical sample | | | | | | | | | | | |
| | Total | | Surveyed persons | | Non-response | | | | | | | |
| | | | | | Total | | Refusals | | Absences | | Inability to answer | |
| | No. | % | No. | % | No. | % | No. | % | No. | % | No. | % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Total | 19,965 | 100.00 | 12,741 | 63.82 | 7,224 | 36.18 | 2,777 | 13.91 | 4,286 | 21.47 | 161 | 0.81 |
| No data recorded | 68 | 100.00 | - | - | 68 | 100.00 | 23 | 33.82 | 45 | 66.18 | - | - |
| Total classified | 19,897 | 100.00 | 12,741 | 64.03 | 7,156 | 35.97 | 2,754 | 13.84 | 4,241 | 21.31 | 161 | 0.81 |
| Cannot read or write | 485 | 100.00 | 340 | 70.10 | 145 | 29.90 | 56 | 11.55 | 69 | 14.23 | 20 | 4.12 |
| Qualification lower than school graduate | 6,359 | 100.00 | 4,211 | 66.22 | 2,148 | 33.78 | 912 | 14.34 | 1,163 | 18.29 | 73 | 1.15 |
| School graduate or equivalent | 5,194 | 100.00 | 2,682 | 51.64 | 2,512 | 48.36 | 928 | 17.87 | 1,543 | 29.71 | 41 | 0.79 |
| High school graduate or 2nd degree Vocational Training, or equival | 7,859 | 100.00 | 5,508 | 70.09 | 2,351 | 29.91 | 858 | 10.92 | 1,466 | 18.65 | 27 | 0.34 |

Total non-response accounts for 36% of the theoretical sample, the greatest weight within this corresponding to absences, with 21%. By educational level, *School Graduate or equivalent* is that which recorded the highest percentage of non-response (48%), as well as the highest percentages of refusals, and above all, absences (almost 30%), which may be taken as a reference, we observe that both refusals and absences are mainly concentrated in the modality of *School Graduate or equivalent*. Regarding the inabilities to answer, despite their scarce importance in total non-response, it can be observed that its highest percentage is reached in the modality of *Cannot read or write*.

The analysis of non-response considering the nationality of the selected person may be done from **Table 6**.

We can see that non-response is greater in the group of persons with a foreign nationality, where it represents 45% of the theoretical sample. However, the refusals reach their highest percentage (14%) in persons with Spanish nationality, as compared with 10% for persons with a foreign nationality. In turn, the percentage of absences is much greater in the group of persons with a foreign nationality (30.4% as compared with 21% in persons with Spanish nationality). Lastly, the inabilities to answer are very concentrated in those persons with a foreign nationality, surely due to the fact that a good number of them do not know the Spanish language, which constitutes one of the causes of the inability to answer.

**TABLE 6. Distribution of non-response in the theoretical sample and of the effective sample, by nationality of the selected person**

| Nationality | Theoretical sample | | | | | | | | | | | |
| | Total | | Surveyed persons | | Non-response | | | | | | | |
| | | | | | Total | | Refusals | | Absences | | Inability to answer | |
| | No. | % | No. | % | No. | % | No. | % | No. | % | No. | % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Total classified | 19,965 | 100.00 | 12,741 | 63.82 | 7,224 | 36.18 | 2,777 | 13.91 | 4,286 | 21.47 | 161 | 0.81 |
| Spanish | 18,709 | 100.00 | 12,049 | 64.40 | 6,660 | 35.60 | 2,651 | 14.17 | 3,904 | 20.87 | 105 | 0.56 |
| Foreign | 1,256 | 100.00 | 692 | 55.10 | 564 | 44.90 | 126 | 10.03 | 382 | 30.41 | 56 | 4.46 |

**Table 7** shows the percent distributions, according to the educational level fo the selected person, in the groups of incumbent persons surveyed and with non-response, of the substitute persons and of the persons of the total effective sample.

It can be observed that the number of substitute persons is greater than the number of incumbent persons with non-response; this is due to the fact that the survey replaces framework incidences, in addition to the incidences corresponding to the household and the selected person.

**TABLE 7. Percent distributions, by educational level, of the incumbent persons surveyed and with non-response, of the substitute persons and of the total effective sample**

| Educational level | Incumbent persons surveyed | | Incumbent persons with non-response | | Substitute persons (reserve persons) | | Persons in the effective sample total | |
| | No. | % | No. | % | No. | % | No. | % |
|---|---|---|---|---|---|---|---|---|
| Total | 12,741 | - | 7,224 | - | 7,268 | - | 20,009 | - |
| No data recorded | - | - | 68 | | - | - | - | - |
| Total classified | 12,741 | 100.00 | 7,156 | 100.00 | 7,268 | 100.00 | 20,009 | 100.00 |
| Cannot read or write | 340 | 2.67 | 145 | 2.03 | 184 | 2.53 | 524 | 2.62 |
| Qualification lower than school graduate | 4,211 | 33.05 | 2,148 | 30.02 | 2,203 | 30.31 | 6,414 | 32.06 |
| School graduate or equivalent | 2,682 | 21.05 | 2,512 | 35.10 | 1,587 | 21.84 | 4,269 | 21.34 |
| High school graduate or 2nd degree Vocational Training, or equivalent or | 5,508 | 43.23 | 2,351 | 32.85 | 3,294 | 45.32 | 8,802 | 43.99 |

Once this point is clarified, it is possible to see, firstly, that the first and last distributions that appear in the table (effective sample of incumbent persons and total effective sample) are very much alike, which seems logical, if we consider that the effective sample of incumbent persons is an appreciable subgroup of the total effective sample. There is also a great deal of similarity between the distributions of the substitute persons and the total effective sample. The most significant differences are in the second distribution (incumbent persons with non-response) and the three remaining distributions, yielding the greatest differences in the modalities of *School Graduate or equivalent* and *High school graduate or 2nd degree Vocational Training, or equivalent or higher degrees*. It

may be observed that, in practice (comparing the second and third distributions), persons with an average educational level have been replaced by others with a higher educational level, and especially, persons with *School Graduate or equivalent* by persons with *High school graduate or 2nd degree Vocational Training, or equivalent or higher degrees*.

## 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 opposite of the probability of response in each of the groups. If it approaches one, both groups have a similar behaviour. Values above one represent greater non-response in the numerator group, and values below one indicate greater non-response in the denominator group.

In order to perform the estimate, the theoretical sample of persons has been broken down to indicate persons surveyed (effective sample) and incidences. Within the latter, we have distinguished between framework incidences and incidences in persons, including in the latter both those of the selected person and those of the household of which s/he is a part. Only incumbent persons have been considered, both in the effective sample and in the incidences, whereas reserve persons have been disregarded.

The initial idea consisted of separating the persons, both those surveyed and those who had some type of incidence, into two groups:
- Extra-community citizens, in other words, persons from outside the EU
- Non extra-community citizens, who were, in turn, divided into two subgroups:
  - Spaniards
  - Non-Spaniard community persons, henceforth referred to as "community citizens"

The previous breakdown has been prepared using the *country of nationality* stated in the Register. There were one hundred eight persons whom it was not possible to find in this administrative register, and another two for which no country of nationality was recorded.

**Horizontal percentages** (compared with the total number of persons in the theoretical sample with nationality, compared with the total for each type of incidence and compared with the effective sample) and **vertical percentages** (compared with the theoretical sample with nationality in each group of persons), have been calculated, both for extra-community and non-extra-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 incidences

- 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 8** garners the results obtained, showing that, in the first place, extra-community citizens represent only 6.9% of the total number of persons in the theoretical sample for whom it has been possible to assign their nationality. This percentage is even lower for community citizens, standing at 2%.

It is also worth noting that:

- The percentage of **empty dwellings** is approximately double in the dwellings in which the selected person is theoretically an extra-community or community citizen (nearly 4%) than in those dwellings in which the same has Spanish nationality (2%).

- In the **unlocatable dwellings**, the percentage is much higher in the dwellings in which the selected person is a community citizen according to the Register (nearly 9%) than in the other two types.

- Worth noting is the difference in percentage existing in **unlocatable persons**, given that the percentage of the same when they are extra-community or community citizens (31%) is triple the percentage when they are Spaniards (10%).

- In the percentages of **absences from the household**, there are no significant differences among the three types.

- The **refusals of the household** are relatively more numerous in those households in which the selected person is a Spaniard (2.8%) or extra-community citizen (1.9%) than in those households in which the selected person is a community citizen (0.8%).

- Regarding the **absences of persons**, these carry more weight among Spaniards (8.5%) and extra-community citizens (7.6%) than among community citizens (5.5%).

- In the **refusals of persons**, there are significant differences, with the highest percentage corresponding to those persons of Spanish nationality (9.4%),a nd the lowest percentage to extra-community citizens (3.7%), while the intermediate value corresponds to the community citizens (6.1%).

- The **inabilities to answer** carry more weight among community and extra-community citizens (2.5% and 1.8%, respectively) than among Spaniards, where it is irrelevant (0.5%). These differences can surely be explained by language differences of persons with foreign nationalities.

- As a result of the large number of incidences recorded, the percentages of **persons surveyed** are quite low, especially in the case of community citizens, which reaches a value of 24.6%. The highest percentage has been obtained among Spaniards, which only amount to 55%, while the percentage for extra-community citizens stands at 35.5%.

- In the ratios of refusals which have been calculated, significant differences are observed, though this is not the case for the refusals and absences considered together.

- Regarding the estimate of the differential non-response correction coefficient, it is when all of the incidents are considered that it strays the most from one, reaching a value of 1.53. This is due to the greater weight that the framework incidences carry in the group of extra-community citizens, fundamentally in the *unlocatable person* incidence.

# TABLE 8. Incidences in the theoretical sample, by nationality of the selected person

| Persons | Total | Extra-community citizens | horizontal % | vertical % | Non-extra-community citizens Total | horizontal % | vertical % |
|---|---|---|---|---|---|---|---|
| **Theoretical sample (incumbent persons)** | 24,030 | | | | | | |
| -Not found in the Register | 108 | | | | | | |
| **Incumbent persons found in the Register** | 23,922 | | | | | | |
| -Nationality not recorded | 2 | | | | | | |
| **Theoretical sample with nationality** | 23,920 | 1,652 | 6.9 | | 22,268 | 93.1 | |
| -Incidences | 11,214 | 1,066 | | | 10,148 | | |
| In the framework:- Empty dwellings | 532 | 69 | 13.0 | 4.2 | 463 | 87.0 | 2.1 |
| - Dwellings used for other purposes | 68 | 15 | 22.1 | 0.9 | 53 | 77.9 | 0.2 |
| - Inaccessible dwellings | 36 | 1 | 2.8 | 0.1 | 35 | 97.2 | 0.2 |
| - Unlocatable dwellings | 445 | 44 | 9.9 | 2.7 | 401 | 90.1 | 1.8 |
| - Persons outside of the study scope | 48 | 2 | 4.2 | 0.1 | 46 | 95.8 | 0.2 |
| - Unlocatable persons | 2,862 | 513 | 17.9 | 31.1 | 2,349 | 82.1 | 10.5 |
| In households: | | | | | | | |
| -Absences | 2,286 | 169 | 7.4 | 10.2 | 2,117 | 92.6 | 9.5 |
| -Refusals | 647 | 32 | 4.9 | 1.9 | 615 | 95.1 | 2.8 |
| -Inability to answer | 17 | 4 | 23.5 | 0.2 | 13 | 76.5 | 0.1 |
| In persons: | | | | | | | |
| -Absences | 1,999 | 126 | 6.3 | 7.6 | 1,873 | 93.7 | 8.4 |
| -Refusalss | 2,130 | 61 | 2.9 | 3.7 | 2,069 | 97.1 | 9.3 |
| -Inability to answer | 144 | 30 | 20.8 | 1.8 | 114 | 79.2 | 0.5 |
| -Surveyed (**effective sample**) | 12,706 | 586 | 4.6 | 35.5 | 12,120 | 95.4 | 54.4 |

| Persons | Non-extra-community Community citizens | horizontal % | vertical % | Spaniards | horizontal % | vertical % |
|---|---|---|---|---|---|---|
| **Theoretical sample (incumbent persons)** | | | | | | |
| -Not found in the Register | | | | | | |
| **Incumbent persons found in the Register** | | | | | | |
| -Nationality not recorded | | | | | | |
| **Theoretical sample with nationality** | 475 | 2.0 | | 21,793 | 91.1 | |
| -Incidences | 358 | | | 9,790 | | |
| In the framework:- Empty dwellings | 21 | 3.9 | 4.4 | 442 | 83.1 | 2.0 |
| - Dwellings used for other purposes | 4 | 5.9 | 0.8 | 49 | 72.1 | 0.2 |
| - Inaccessible dwellings | 1 | 2.8 | 0.2 | 34 | 94.4 | 0.2 |
| - Unlocatable dwellings | 42 | 9.4 | 8.8 | 359 | 80.7 | 1.6 |
| - Persons outside of the study scope | 1 | 2.1 | 0.2 | 45 | 93.8 | 0.2 |
| - Unlocatable persons | 148 | 5.2 | 31.2 | 2,201 | 76.9 | 10.1 |
| In households: | | | | | | |
| -Absences | 60 | 2.6 | 12.6 | 2,057 | 90.0 | 9.4 |
| -Refusals | 4 | 0.6 | 0.8 | 611 | 94.4 | 2.8 |
| -Inability to answer | 10 | 58.8 | 2.1 | 3 | 17.6 | 0.0 |
| In persons: | | | | | | |
| -Absences | 26 | 1.3 | 5.5 | 1,847 | 92.4 | 8.5 |
| -Refusalss | 29 | 1.4 | 6.1 | 2,040 | 95.8 | 9.4 |
| -Inability to answer | 12 | 8.3 | 2.5 | 102 | 70.8 | 0.5 |
| -Surveyed (**effective sample**) | 117 | 0.9 | 24.6 | 12,003 | 94.5 | 55.1 |

14

|  |  | % |
|---|---|---|
| Refusals extra-com., theoretical sample | 61 | 9.4 |
| Total extra-com.(effective sample + refusals) | 647 |  |
|  |  |  |
| Refusals rest of theoretical sample | 2,069 | 14.6 |
| Total rest (effective sample + refusals) | 14,189 |  |
|  |  |  |
| Refusals + absences extra-com, theoretical sample | 187 | 24.2 |
| Total extra-com.(effective sample + refusals + absences) | 773 |  |
|  |  |  |
| Refusals + absences rest of theoretical sample | 3,942 | 24.5 |
| Total rest (effective sample + refusals + absences) | 16,062 |  |

## ESTIMATION OF THE COEFFICIENT (DIFFERENTIAL PROPENSITY OF NON-RESPONSE)

|  | With everything | With refusals | With absences | With refusals and absences |
|---|---|---|---|---|
| Pxt (extra-com. persons, theoretical sample) | 1,652 | 647 | 712 | 773 |
| Pxr (extra-com. persons, effective sample) | 586 | 586 | 586 | 586 |
| Pyt (non-extra-com. persons, theoretical sample) | 22,268 | 14,189 | 13,993 | 16,062 |
| Pyr (non-extra-com. persons, effective sample) | 12,120 | 12,120 | 12,120 | 12,120 |
|  |  |  |  |  |
| Estimation value (Pxt/Pxr)/(Pyt/Pyr) | 1.53 | 0.94 | 1.05 | 1.00 |