



Arnaud RÉGNIER-LOILIER*

Inconsistencies in the Number of Children Reported in Successive Waves of the French Generations and Gender Survey

What is the best way to ask survey respondents about the number of children they have? Arnaud RÉGNIER-LOILIER compares the number of children reported by the people who completed the three waves of the Étude des relations familiales et intergénérationnelles (ERFI), the French version of the European Generations and Gender Survey. While there was no single question in the first two waves about the number of children born or adopted, this figure was determined by a succession of questions about the number of children alive or deceased, resident or non-resident in the household. This procedure introduced errors which are measured in this article by a further question contained in the third wave which asked respondents to give their total number of children. It shows that some respondents tend not to report non-resident children. The article draws survey designers' and users' attention to these collection biases and highlights the value of including a question about the total number of children.

From 2005 to 2011, the French National Institute of Demographic Studies (Institut national d'études démographiques, INED) and the National Institute for Statistics and Economic Studies (Institut national de la statistique et des études économiques, INSEE) carried out the French component of the international programme of Generations and Gender Surveys (GGS), known in French as the Étude des relations familiales et intergénérationnelles (Study of family and intergenerational relationships, ERFI).⁽¹⁾ As in any survey, it is important to assess the quality of the data before they are analysed. Since this was a multi-round survey (the same people were questioned on three occasions),

(1) The name of the survey was changed in French because of the vagueness of the word "gender", which was considered by respondents in the test surveys to be too scientific and uninformative.

* Institut national d'études démographiques, Paris.

Correspondence: Arnaud Régnier-Loilier, Institut national d'études démographiques, 133 Boulevard Davout, 75980 Paris Cedex 20, France, Tel: +33 (0)1 56 06 20 71, email: arnaud.regnier-loilier@ined.fr

this point is particularly crucial: in addition to the possible biases linked to the response rates in successive waves, the consistency of the information collected in each wave must also be verified. The present article focuses on this latter point, and more specifically on the number of children reported by respondents in the various survey waves.

The question of data quality

The reliability of statistics taken from sample surveys depends on many things. The response rate is the first factor affecting data representativeness. Some categories of people are harder to contact, and others more often refuse to participate. It is essential, therefore, to correct for non-response and calculate weighting variables in order to make the sample a representative one. The quality of the statistics also depends on the reliability of the information recorded. And this is closely related to the wording of the questions (which must not be ambiguous) and the ability of respondents to give an exact answer. Retrospective surveys tend to delve further and further into people's past and this increases the chance of faulty recall or reinterpretation of past events (Auriat, 1996). On the specific point of retrospective data about family events (births, marriages, etc.), which is our main concern here, various studies have concluded that the dates reported by women are more reliable (Poulain et al., 1992) and note that men more often fail to report some of their children,⁽²⁾ particularly when they are not living under the same roof (Toulemon, 1997) and are therefore less involved in their daily upbringing (Joyner et al., 2012). This is especially the case when the father is not married (Rendall et al., 1999).

The mode of data collection and survey administration also significantly affects the responses. The presence of a partner, for example, leads to under-reporting of multi-partner relationships and drug use (Firdion, 1993), while the presence of small children during the interview causes greater inaccuracy in the dates ascribed to past events: the children may demand the respondent's attention and reduce their concentration. Conversely, notably when men are interviewed, the partner's presence increases the accuracy of dates (the partner helps to jog the respondent's memory) and the quality of information about the division of domestic and parental tasks. The partner acts as a monitor of the information supplied to the interviewer, preventing the man from overstating his participation in daily activities (Régnier-Loilier, 2007, 2014).

Longitudinal surveys, in which the same people are interviewed more than once, raise further data quality issues. First, sample attrition from one wave to the next may undermine the robustness of statistical tests. Second,

(2) Here we examine developed countries. With respect to Mali, Véronique Hertrich (1991) has shown that that society is organized around men, who remember certain events as clearly as women (omissions from men's reproductive histories are no more frequent than from women's), indeed more clearly for some events, such as marriage.

since attrition is rarely random but rather selective (for the ERFI-GGS survey, see Régnier-Loilier and Guisse, 2012), it distorts the structure of the initial sample and may affect the survey's variables of interest, biasing the results and interpretation (Razafindratsima and Kishimba, 2004; Mazuy et al., 2005). Furthermore, questioning the same people in successive waves raises the issue of consistency in their responses. In addition to the increasing risk of inaccurate recall over time, changes in a respondent's situation may also alter their perception of things and their way of telling their story (Mazuy and Lelièvre, 2005; Molinié and Laville, 2000). For example, when a couple breaks up, the children usually go and live with their mother and sometimes see their father infrequently or not at all. In such situations, the father may no longer report those children in the following survey wave. Conversely, if the children divide their time between both parents' homes, the father may consider that they live with him and elsewhere too, leading to double counts.

Study objectives and scope

The aim of this paper is not to assess the quality of reproductive histories collected in the French ERFI-GGS survey by comparing them with external sources.⁽³⁾ It is rather to compare a given respondent's reported number of children in successive waves to see if the response is consistent from one wave to the next and to better understand any observed discrepancies. The scope of the study is therefore restricted to people who took part in all three survey waves, a total of 5,436 observations.

What is the degree of inconsistency between these various indicators? Are the observed discrepancies due to under-reporting of children in Wave 2 and/or Wave 3, or double counting in Wave 1? Which children disappear between waves: those resident in the household in the previous wave, those living elsewhere, those deceased? Do the respondents who under-reported one or more children in Wave 2 also under-report children in Wave 3? Do the people who provide inconsistent responses have a specific profile?

Since the extent of under-reporting is considerable and the data are now available to the research community, the purpose of this study is to alert users of the French ERFI-GGS survey to the issue of data quality. More generally, it aims to encourage other countries participating in the GGS programme to undertake a similar assessment and to warn designers of future surveys of the risk of only using indirect questions to determine a respondent's number of children, especially when this is a central focus of the study.

(3) On this point, a study was carried out in 13 countries that performed the first GGS survey wave (Neels et al., 2011).

I. The ERFI-GGS survey and the counting of children

The data: the Generations and Gender Survey

As part of the Generations and Gender programme of longitudinal, comparative surveys involving some twenty countries (www.ggp-i.org), France held the first wave of its Étude des relations familiales et intergénérationnelles (ERFI-GGS) in the autumn of 2005. ERFI covered the family in general. One respondent aged 18-79 was drawn at random from the members of each participating household. They were asked about their childhood, past and present marital status, their children, their employment status and that of their partner where relevant, their well-being, their opinions, etc. (for more details of the international questionnaire, see Vikat et al., 2007; for more on the French survey, see Régnier-Loilier, 2012). The same people were contacted again three and six years later (autumn 2008 and 2011) to answer a further questionnaire. No proxy respondents were allowed in subsequent waves: if the original respondent refused to participate, had died, etc., no other household member was interviewed.

From one survey wave to the next, the questionnaires were very similar in structure, content and question wording so as to ensure comparability between waves (Régnier-Loilier, 2006, 2010, 2012). Similarly, the mode of collection was identical in the three waves held in France: a CAPI⁽⁴⁾ questionnaire administered by INSEE interviewers, face-to-face in the home of respondents living throughout metropolitan France (mainland France and Corsica). As far as possible, the same interviewers were used for the successive waves, although unfortunately it is not possible to know if a given person was questioned by the same interviewer in the various waves.

A total of 10,079 people were questioned in the first survey wave (2005), 6,534 in the second (2008), and 5,781 in the third (2011).⁽⁵⁾ Of the respondents in 2011, 5,436 had participated in both 2005 and 2008, while 345 had not been questioned in Wave 2. Although a weighting variable is available to correct the sample distortion due to attrition, it is not used in this article, since representativeness is not an aim here.

Determining total number of children in the ERFI-GGS survey

The questionnaire was generally designed to avoid asking in Waves 2 and 3 for information already collected in Wave 1, and respondents were not shown the responses they had given in previous waves.⁽⁶⁾ However, since family

(4) Computer-assisted personal interviewing with automatic filtering.

(5) See Régnier-Loilier and Guisse (2012) for a detailed description of attrition and associated factors.

(6) For example, a comparative question such as “Last time, you said you were living with n resident/non-resident children. Is this still the case?” was not included in the international GGS reference questionnaire, because GGS was designed to be available in paper and pencil form too, with no retrieval of data from previous waves.

situations were liable to change between interviews, items concerning residents in the dwelling, non-resident children and deceased children were repeated in identical terms in each survey wave. These questions were intended to identify any births or deaths of children between waves, children who had left the parental home or returned to the household.

The questionnaire did not, therefore, contain a direct question about the respondent's total number of children. This figure was determined *ex post* by summing the numbers of children reported in three separate modules:

- 1) "Other household members" module. This provided precise details of the relationship between each of the people reported in the dwelling and the respondent: respondent's spouse/partner, child born to current partner, child born to previous partner, stepchild, adopted child, foster-child (temporary), respondent's sibling, partner's sibling, respondent's parent, partner's parent, respondent's grandparent, grandchild, great-grandchild or other relative, partner's other relative, non-family member;
- 2) "Non-resident children" module, introduced by a question which might take two forms:

If the respondent had reported one or more resident children: "We already talked about those children who currently live in your household. In addition to them, have you given birth to/ fathered any other children or have you ever adopted any other children? Do not include stepchildren, that is, children belonging to your current or prior partner/spouse. I will ask you about those children later: Yes/No", and if Yes, "How many?"

If the respondent had not reported any resident children: "Have you given birth to/ fathered any children or have you ever adopted any children? Do not include stepchildren, that is, children belonging to your current or prior partner/spouse. I will ask you about those children later: Yes/No", and if Yes, "How many?"

Each child was then recorded in detail: sex, date of birth, relationship with current partner, date of leaving parental home, frequency of contact with respondent, satisfaction rating of relationship, child's parental situation;

- 3) "Deceased children" module, introduced by "Sometimes it happens that one loses a child. Have you had any other children who are now dead? Yes/No" and if Yes, "How many?"

Each deceased child was then recorded in less detail than a living child: sex, date of birth, date of death, relationship with current partner, child's parental situation.

Although in theory it should be possible to sum the children recorded in these different sections of the questionnaire and obtain the total number of children a person has had, the matching of data from the first two survey waves (2005 and 2008) revealed that for a not insignificant proportion of respondents,

the total number of children differed in 2005 and 2008 (Régnier-Loilier et al., 2011), despite identical questions in the two waves. Although the total number of children should logically be higher in 2008 than in 2005 (as a result of births), in many cases the respondent reported fewer children in Wave 2 than in Wave 1 (8% of all Wave 2 respondents). The extent of this phenomenon might suggest technical error (inaccurate matching between waves, miscalculation of number of children, filtering errors in the questionnaire) but verification ruled out that hypothesis.⁽⁷⁾

When these inconsistencies were examined more closely, the discrepancy appeared to come from people who reported non-resident children in Wave 1 and then said they had no children in Wave 2, but there was no convincing explanation for this. Was it an avoidance strategy by respondents who remembered from Wave 1 that each reported child led to further questioning (ten or so questions for each child)? Was it unintended under-reporting by respondents who omitted their children because their situation had not changed since Wave 1? Or had there been over-reporting of children in Wave 1, with some, say, reported as resident and also as living elsewhere (a possibility in Wave 1 since the person chosen at random from the household was not necessarily the person who had completed the household roster)?⁽⁸⁾

No further investigation was possible before Wave 3 of the survey in 2011. However, when Wave 3 was being set up, particular care was taken in this area. First, during training sessions the interviewers were alerted to the problem and specific instructions were added to the questionnaire to draw their attention to this point in the sections for recording children living outside the household: “Very important: record all the respondent’s (living) non-resident children, even if they were already non-resident in the previous wave”. In addition, to establish whether the observed discrepancies were due to omissions in one of the survey waves or rather to double counting, a more direct summary question was added in Wave 3 about how many children the respondent had had during his or her life: “To sum up, how many children have you yourself had in all, whether your own or adopted, with your current partner or a previous one?”

(7) A number of comparisons between waves were made to ensure that there were no matching errors: number of siblings, parents’ characteristics, first names of residents in the dwelling (taken from the raw data collected). In addition, since the total number of children was deduced from the count of resident, non-resident and deceased children, the construction of the counting procedure was verified for each wave. We also checked the software programs to ensure that the questions intended to count children had indeed been systematically asked. Finally, there would appear to be no “interviewer effect” (with some interviewers, say, recording that a person had no children without asking the question in order to save time, since they were paid per questionnaire rather than by the hour), because under-reporting occurred with most interviewers and never with all the address-forms of a particular interviewer.

(8) The random draw was made after one of the residents had described the household (Régnier-Loilier, 2006). So a child who was resident only some of the time might have been recorded as living in the household by the person completing the household grid, but considered as non-resident by the survey respondent. In practice, some cases of double counting were found, but they were very few in number.

Please include any children who are no longer living with you or who are deceased”.

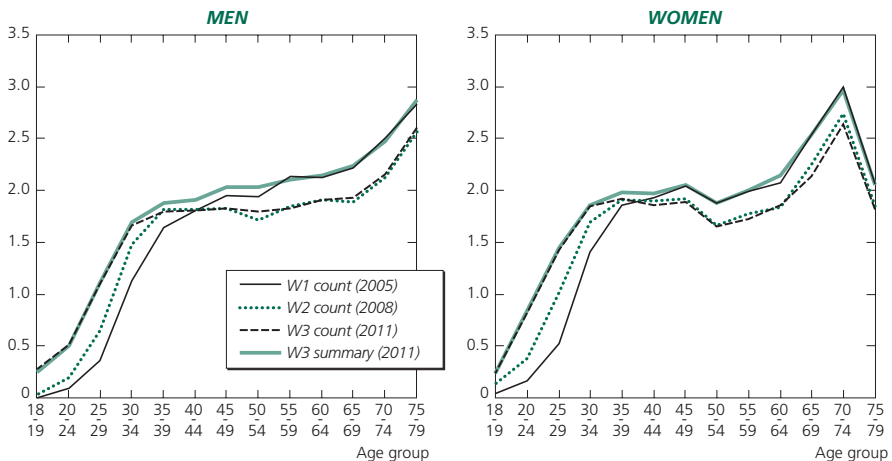
Following Wave 3, for each respondent we now have four indicators of their total number of children: three taken from the counts of children who were resident in the household, non-resident and deceased in each wave (2005, 2008 and 2011) and a fourth from the summary question asked in the final wave. These are the four indicators compared in this article.

II. Substantial under-reporting of children by respondents aged 45 and over in Waves 2 and 3

First, we will compare the aggregated average number of children by age group and sex as shown by the four child counts (Figure 1): children resident in the household, non-resident and deceased in 2005 (W1 count), 2008 (W2 count), 2011 (W3 count) and the total number of children recorded from the summary question in 2011 (W3 summary), the indicator that can be considered most reliable. Remember that the data are not weighted, since the aim is not to measure numbers of children by age group but to compare indicators for a given group of people.

At ages 18-44 (respondent's age in Wave 1), the differences between the W1, W2 and W3 counts correspond to births: at each age the average number of children is higher in W3 than in W2, and higher in W2 than in W1. But

Figure 1. Mean number of children by respondent's age in 2005 according to the four child count indicators



Interpretation: According to the separate child counts (resident in the household, non-resident and deceased), men aged 55-59 reported 2.14 children in 2005 (W1 count), 1.82 in 2008 (W2 count), 1.82 in 2011 (W3 count), and according to the summary question, 2.10 in 2011 (W3 summary).

Note: Unweighted data; respondent's age in 2005.

Coverage: Respondents to all three survey waves (N = 5,436).

Source: INED-INSEE, ERFI-GGS123, 2005-2011.

from age 45, the curves of the W2 and W3 counts are superimposed and drop below that of the W1 count. At ages 45-79, therefore, respondents reported fewer children on average in 2008 and 2011 than in 2005 (-0.25 children). The fact that the 2008 and 2011 curves are perfectly superimposed might well suggest an over-reporting of children in Wave 1. However, the level of the W3 summary indicator argues rather for an under-reporting of children in Waves 2 and 3. This is because the W3 summary curve is almost exactly superimposed on the W3 count curve up to age 40, quite logically since the two indicators were observed in the same year, and almost exactly on the W1 count curve after age 45 (with a slight deviation for men at ages 40-54). This comparison shows that summing the counts of resident, non-resident and deceased children is a fairly reliable method in Wave 1 after ages 45-50⁽⁹⁾ and in Wave 3 in the childbearing age groups. It appears, however, to reveal significant under-reporting in Waves 2 and 3 (of 0.25 children on average), among respondents aged 45-79 in 2005.

The value of this comparative analysis is limited, however, by the fact that it is done not at individual but at aggregate level (comparing averages). To validate the reliability of the counting method used in Wave 1, we can compare a given individual's total number of children (biological or adopted) obtained with the counting method and with the summary indicator in W3 (Figure 2). We thus have three possible configurations for an individual *i*:

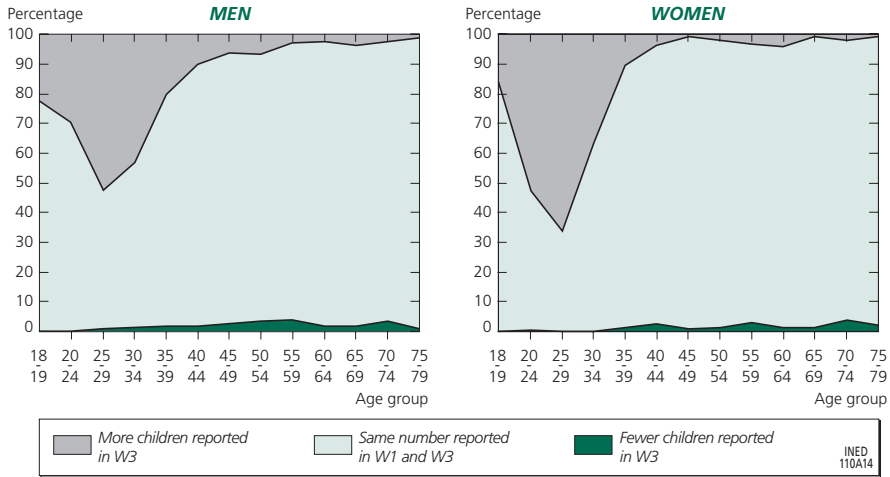
- the number of children is identical in W1 count and W3 summary;
- the number of children in W1 count is lower than in W3 summary (more children in 2011 than in 2005 due to births, under-reporting in W1, over-reporting or double counting in W3);
- the number of children in W1 count is higher than in W3 summary (fewer children in 2011 than 2005 due to over-reporting in W1 or under-reporting in W3).

Logically, the proportion of women who report more children in W3 is substantial until about the age of 40 (peaking among those aged 25-29 in 2005), since children are born to women of those ages, and then falls to almost nothing (2% for the 45-79 groups). The same holds for men, but over different age groups, reflecting fertility after 45-49 and also less consistent responses from one wave to the next (see below). The proportion of women and men reporting fewer children in W3 than W1 is tiny (1.4% for all women aged 18-79; 1.9% for men). For the over-45s, there is consequently a high degree of consistency (over 95%) between W1 count and W3 summary, showing that the total number of resident, non-resident and deceased children given by the 2005 count is generally reliable.

We may therefore disregard the W3 summary indicator and make the comparative analysis at individual level between the W1, W2 and W3 count

(9) And probably for younger groups too, but comparing the number of children in 2005 and 2011 only makes sense for those who have reached the end of their reproductive lives.

Figure 2. Proportion of respondents reporting more, fewer or the same number of children in W3 summary compared with W1 count



Interpretation: 2% of men aged 60-64 in 2005 reported fewer children in the Wave 3 summary question about total number of children than in the Wave 1 count, 95% reported the same number and 3% reported more.

Note: Unweighted data; respondent's age in 2005.

Coverage: Respondents to all three survey waves (N = 5,436).

Source: INED-INSEE, ERFI-GGS123, 2005-2011.

indicators. However, in order to focus on the under-reporting of children in Wave 1 while ensuring that the comparisons between waves are not distorted by any further births during the childbearing years, we shall consider only the number of children born up to September 2005, i.e. before the Wave 1 survey began. Although on average (Figure 1) under-reporting appears to be similar in Waves 2 and 3, the aim here is to see whether the same people under-reported in Waves 2 and 3 or whether some under-reported in one wave but not the other. Using our remaining three indicators, a large number of configurations are possible.⁽¹⁰⁾ To simplify matters, five situations will be compared for an individual i (Figure 3):

- the number of children born before October 2005 is identical in the three waves: W1 count = W2 count = W3 count;
- the number of children is lower in W3 only: W3 count < W1 count, and W3 count < W2 count, while W2 count = W1 count;
- the number of children is lower in W2 only: W2 count < W1 count, and W3 count = W1 count;
- the number of children is lower in W2, and the same between W2 and W3: W2 count < W1 count, and W2 count = W3 count. This would mean that under-reporting was identical in Waves 2 and 3, or there was over-

(10) $W1 = W2 = W3$; $W1 > W2 = W3$; $W1 < W2 = W3$; $W1 = W2 > W3$; $W1 = W2 < W3$; $W1 > W2 > W3$; $W1 < W2 < W3$; $W1 > W2 < W3$ and $W1 = W3$; $W1 > W2 < W3$ and $W1 < W3$; $W1 > W2 < W3$ and $W1 > W3$; $W1 < W2 > W3$ and $W1 = W3$; $W1 < W2 > W3$ and $W1 < W3$; $W1 < W2 > W3$ and $W1 > W3$.

Figure 3. Proportion of respondents reporting more, fewer or the same number of children born before 2005 in the different waves according to the indicators W1 count, W2 count and W3 count



Interpretation: 76% of men aged 60-64 in 2005 reported the same number of children (born before 2005) in all three waves, 8% reported fewer in Wave 3 (but the same number in Waves 1 and 2), 7% reported fewer in Wave 2 (but the same number in Waves 1 and 3) 5% reported fewer in Wave 2 but the same number in Waves 2 and 3, and 4% reported more in Wave 2 and/or Wave 3.

Note: Unweighted data; respondent's age in 2005.

Coverage: Respondents to all three survey waves (N = 5,436).

Source: INED-INSEE, ERFI-GGS123, 2005-2011.

reporting in W1 (possible double counting, for example, of children reported both on the household roster and as living elsewhere, such as students or children in alternating custody);

- the number of children is higher in W2 and/or W3: W1 count < W2 count and/or W1 count < W3 count (more children in 2008 and/or 2011 than 2005: birth, adoption, child omitted in Wave 1).

Up to age 40, the comparison of the three indicators produces relatively consistent results. On the other hand, after age 40, more than one in five respondents of both sexes (approximately 22%) under-report children in Wave 2 and/or Wave 3. And, significantly, the under-reporting is relatively inconsistent: most respondents who under-report do so only once (in either Wave 2 or Wave 3).

Although the interviewers were specifically alerted to this phenomenon for Wave 3, (interviewer training, instructions displayed on CAPI screens), the degree of under-reporting of children in Wave 3 was similar to that observed in Wave 2 (9% and 8% respectively, Table 1). Furthermore, even if reporting fewer children in Wave 2 than Wave 1 more often led to reporting fewer in Wave 3 than Wave 1 (39% who under-reported in W2 also did so in W3,

Table 1. Percentages and numbers of respondents reporting fewer children in Wave 2 and/or Wave 3 than in Wave 1

	No under-reporting in W3		Under-reporting in W3		Total	
	%	N	%	N	%	N
No under-reporting in W2	93.4	4,666	6.6	327	91.8	4,993
Under-reporting in W2	61.4	272	38.6	171	8.2	443
Total	90.8	4,938	9.2	498	100.0	5,436

Interpretation: 38.6% of those who reported fewer children (born before 2005) in Wave 2 than Wave 1 also reported fewer children in Wave 3 than Wave 1.
Note: Unweighted data.
Coverage: Respondents to all three survey waves (N = 5,436).
Source: INED-INSEE, ERFI-GGS123, 2005-2011.

compared with just 7% of those who did not under-report in W2),⁽¹¹⁾ this under-reporting did not generally concern the same people: among “under-reporters”, 42% under-reported in Wave 3 only,⁽¹²⁾ 35% in Wave 2 only, and 22% under-reported in both Waves 2 and 3 (with respect to Wave 1).

III. Omissions mainly concern non-resident children

The indicators record three categories of children: resident in the household, non-resident, and deceased. Does the under-reporting observed affect all children reported in Wave 1 or only a particular category? It is known, for example, that deceased children tend to be under-reported in surveys, and that men sometimes under-report children they are no longer living with, especially after a separation (Toulemon, 2005). Under-reporting might also be related to children who were resident in Wave 1 but then left the parental home. However, given that under-reporting was particularly marked among respondents aged 55 and above in 2005, and therefore aged at least 58 in Wave 2 and 61 in Wave 3, ages when fewer people are still living with their children,⁽¹³⁾ this suggests that under-reporting mainly affects non-resident children.

If we compare respondents reporting fewer children in Wave 2 and/or Wave 3 than Wave 1 with those reporting the same number in all three waves,⁽¹⁴⁾

(11) The comparison between these two proportions might be open to question because coverage in this case is not restricted to only those in a position to have “under-reported” one or more children, i.e. who had at least one child in wave 1. However, a discrepancy of similar size remains even when observation is limited to parents who had at least one child in wave 1 (39% compared with 9%).

(12) 327 out of 770 under-reporters in one of the waves (327 + 171 + 272). The same calculation method was used for the other proportions given.

(13) In 2005, 28% of respondents aged 55-59 were still living with at least one child in the household, but only 16% of those aged 60-64.

(14) Coverage is restricted to persons who reported at least one child in wave 1, i.e. the population exposed to the risk of under-reporting in subsequent waves.

the under-reporters include a much higher proportion of parents with at least one non-resident child: 84% of Wave 2 under-reporters and 76% of Wave 3 under-reporters were in that situation in 2005, compared with 51% of those who reported the same number in all three waves. Many of them, too, were parents of a child reported deceased in 2005 (12% compared with 4% of non-under-reporters). Conversely, proportionately fewer under-reporters were living with one or more children in 2005. Under-reporting appears, therefore, to be related to non-resident children and those reported deceased in Wave 1.

In order to refine this observation and identify which children were under-reported, the total number of children born before Wave 1 is broken down into three categories: resident in the household, non-resident and deceased. Observation is restricted to the respondents who under-reported one or more children in Wave 2 (N = 443 respondents) and Wave 3 (N = 498).⁽¹⁵⁾ We compare the number of non-resident children reported in Wave 2 with those in Wave 1, and then those reported in Wave 3 with those in Wave 1 (Table 2A). We do the same for resident (Table 2B) and deceased (Table 2C) children.

First, under-reporting appears in most cases to be related to reporting one or more non-resident children in 2005. Nine in ten under-reporters reported no deceased child in 2005. Their under-reporting could, therefore, only concern the resident and non-resident children in 2005. Similarly six in ten under-reporters reported no resident child (Tables 2B and 2C, Column 1). Conversely, eight in ten under-reporters had at least one non-resident child.

The diagonals in Table 2 indicate that the reported number of children born before Wave 1 is identical in both compared waves (2005 and 2008; 2005 and 2011); values above the diagonal indicate fewer children reported in 2008 (or 2011) than 2005; values below indicate more children reported in 2008 (or 2011) than 2005. In both 2008 and 2011, the under-reporters rarely reported fewer deceased children (Table 2C) than in 2005: most were on the diagonal, indeed at the intersection of the “0 children” row and column. There were 35 children reported deceased in 2005 who were not reported in 2008, and 34 who were not reported in 2011. Although these numbers of unreported children in Waves 2 and 3 represent over half of the total number of children reported deceased by the under-reporters in Wave 1, they do not provide an explanation for under-reporting as a whole: these thirty-odd cases account for less than 10% of the total number of under-reporters.

In fact, 80% of under-reporters reported fewer non-resident children in 2008 and 2011 than in 2005 (Table 2A): most of them are above the diagonal. Indeed, there is a particularly high concentration in the top line of the table, whatever the number of children reported in 2005. This means that under-reporting does not affect one non-resident child among the respondent’s children but all their non-resident children. For example, of the 443 respondents who

(15) Of these, 134 under-reported in both wave 2 and wave 3.

Table 2. Comparison of numbers of non-resident (A), resident (B) and deceased (C) children reported in 2005 and 2008 (Waves 1 and 2) and 2005 and 2011 (Waves 1 and 3)

A. Waves 1 - 2		Number of non-resident children in 2005					
		0	1	2	3	4+	Total
Number of non-resident children in 2008	0	71	115	132	60	27	405
	1	0	8	6	2	1	17
	2	0	1	2	3	1	7
	3	0	1	0	4	2	7
	4+	0	0	0	0	7	7
	Total	71	125	140	69	38	443
A. Waves 1 - 3		Number of non-resident children in 2005					
		0	1	2	3	4+	Total
Number of non-resident children in 2011	0	119	128	132	53	28	460
	1	3	5	6	0	1	15
	2	0	1	4	2	2	9
	3	0	0	1	5	2	8
	4+	0	0	2	0	4	6
	Total	122	134	145	60	37	498
B. Waves 1 - 2		Number of resident children in 2005					
		0	1	2	3	4+	Total
Number of resident children in 2008	0	278	35	4	4	1	322
	1	8	36	23	6	0	73
	2	0	1	16	16	1	34
	3	0	0	2	5	2	9
	4+	0	0	1	0	4	5
	Total	286	72	46	31	8	443
B. Waves 1 - 3		Number of resident children in 2005					
		0	1	2	3	4+	Total
Number of resident children in 2011	0	284	57	24	7	2	374
	1	5	22	38	12	0	77
	2	0	0	6	27	1	34
	3	0	0	0	6	2	8
	4+	0	0	0	1	4	5
	Total	289	79	68	53	9	498
C. Waves 1 - 2		Number of deceased children in 2005					
		0	1	2	3	4+	Total
Number of deceased children in 2008	0	388	30	1	0	0	419
	1	2	16	3	0	0	21
	2	0	0	2	1	0	3
	3	0	0	0	0	0	0
	4+	0	0	0	0	0	0
	Total	390	46	6	1	0	443
C. Waves 1 - 3		Number of deceased children in 2005					
		0	1	2	3	4+	Total
Number of deceased children in 2011	0	435	32	0	0	0	467
	1	3	23	0	1	0	27
	2	0	1	2	1	0	4
	3	0	0	0	0	0	0
	4+	0	0	0	0	0	0
	Total	438	56	2	2	0	498

Interpretation: Of the 443 Wave 2 under-reporters, 132 reported having two non-resident children in 2005 and no non-resident children in 2008 (Table 2A).

Note: Unweighted data.

Coverage: Respondents to all three survey waves, parents of at least one child in 2005 (resident, non-resident or deceased) and reporting fewer children born before 2005 in 2008 than 2005 (N = 443) / fewer children born before 2005 in 2011 than in 2005 (N = 498).

Source: INED-INSEE, ERFI-GGS123, 2005-2011.

reported fewer children in 2008 than 2005, half (219) reported at least two non-resident children in 2005 and none at all in 2008. Nearly eight in ten under-reporters said No to the question “We already talked about those children who currently live in your household” (if such was the case). “In addition to them, have you given birth to/ fathered any other children or have you ever adopted any other children? Do not include stepchildren, that is, children belonging to your current or prior partner/spouse. I will ask you about those children later”. The same phenomenon is observed for Wave 3 under-reporters.

However, when we compare the number of resident children reported in the various waves, things seem to operate rather differently (Table 2B). Where the number of resident children reported in Wave 2 and/or Wave 3 is lower than in Wave 1, the numbers are not concentrated in the top line. For example, of the 56 Wave 2 (2008) under-reporters living with two children in 2005, 23 were still living with one child in 2008. In these cases, under-reporting may be due to one of the children leaving the parental home and being omitted from the number reported as non-resident in 2008.

IV. No particular profile among under-reporters

To shed more light on the observed omissions, we used three logistic regressions to establish a “typical” portrait of respondents who under-reported their number of children in survey Wave 2 and/or 3, (Table 3): Model A estimates the probability of reporting fewer children in Wave 2 than Wave 1, Model B the probability of reporting fewer in Wave 3 than Wave 1, and Model C the probability of reporting fewer in Wave 2 and/or 3 than Wave 1.

Various factors were included, most of them relating to the situation observed in Wave 1:⁽¹⁶⁾

- Respondents’ individual characteristics in 2005: sex, age and educational level (this last taken as a proxy for ability to understand the meaning of a question);
- Separation from partner since Wave 1: being separated may lead to a loss of contact with children;
- Number of resident, non-resident and deceased children reported in 2005, in order to assess the effect of family setups that may affect the likelihood of under-reporting;
- Existence of one or more estranged non-resident children in 2005 (who never see the respondent): estrangement and the passage of time may cause the respondent to no longer report these children. Also, since the

(16) Unless otherwise mentioned in Table 3. Note that since the under-reporting may have occurred in either wave 2 or wave 3, the same regression was run with the characteristics observed in the wave before the under-reporting (2005 in the case of under-reporting in 2008; 2008 if in 2011). The results were identical, so only the regression relating to 2005 characteristics is given here.

Table 3. Probability of reporting fewer children in 2008 and/or 2011 than in 2005 (logit model)

		Model A	Model B	Model C	Number
		Under-reporting in Wave 2 vs. Wave 1	Under-reporting in Wave 3 vs. Wave 1	Under-reporting in Wave 2 or Wave 3 vs. Wave 1	
Constant		-1.48 ***	-2.24 ***	-0.93 ***	-
Respondent's sex	Male	0.00 Ref.	0.00 Ref.	0.00 Ref.	1,670
	Female	-0.03 -	-0.03 -	-0.05 -	2,371
Age group in 2005	18-29	-0.87 *	-1.12 **	-1.31 ***	151
	30-34	-1.48 ***	-0.99 ***	-1.55 ***	366
	35-39	-1.01 ***	-0.44 *	-0.96 ***	530
	40-44	-0.51 **	0.08 -	-0.31 *	475
	45-49	-0.31 -	0.14 -	-0.21 -	442
	50-54	0.00 Ref.	0.00 Ref.	0.00 Ref.	462
	55-59	-0.17 -	0.38 **	0.15 -	494
	60-64	-0.32 -	0.10 -	-0.19 -	364
	65-69	-0.41 *	0.26 -	-0.09 -	335
70-74	-0.13 -	-0.11 -	-0.03 -	239	
75-79	-0.40 -	-0.21 -	-0.35 -	183	
Educational level in 2005	None	-0.16 -	0.22 -	-0.02 -	478
	Primary	-0.35 -	0.13 -	-0.23 -	521
	Lower secondary	0.18 -	0.09 -	0.05 -	316
	Lower secondary, vocational	-0.01 -	0.05 -	-0.08 -	1,141
	Upper secondary, vocational	-0.44 -	0.29 -	-0.06 -	283
	Upper secondary, general	0.00 Ref.	0.00 Ref.	0.00 Ref.	252
	2 years in higher education	-0.22 -	-0.27 -	-0.20 -	387
University degree or higher	0.01 -	-0.18 -	-0.15 -	663	
Change in partnership status since 2005	No change	0.00 Ref.	0.00 Ref.	0.00 Ref.	3,601
	Separation	-0.34 -	0.64 ***	0.18 -	265
	New partnership	-0.09 -	0.24 -	0.03 -	175
Number of resident children in 2005	0	0.00 Ref.	0.00 Ref.	0.00 Ref.	1,855
	≥ 1	0.24 -	0.20 -	0.25 *	2,186
Number of non-resident children in 2005	0	-1.20 ***	-0.46 **	-0.86 ***	1,768
	1	0.07 -	0.07 -	0.02 -	709
	2	0.00 Ref.	0.00 Ref.	0.00 Ref.	874
	≥ 3	-0.07 -	-0.31 **	-0.25 **	690
Number of deceased children in 2005	0	0.00 Ref.	0.00 Ref.	0.00 Ref.	3,825
	≥ 1	0.88 ***	0.75 ***	0.87 -	216
Number of estranged children in 2005	0	0.00 -	0.00 Ref.	0.00 Ref.	3,859
	≥ 1	0.96 ***	0.66 ***	0.99 ***	182
Interaction sex (female) * estranged children in 2005 (1 or more)		-1.24 ***	-0.23 -	-0.86 **	-
Wished to receive results (2005)	Yes	0.00 Ref.	0.00 Ref.	0.00 Ref.	3,930
	No	-0.20 -	-0.01 -	-0.19 -	111
Answered income question in 2005	Answered	0.45 -	-0.67 -	0.31 -	3,995
	Refused	0.00 Ref.	0.00 Ref.	0.00 Ref.	46
Agreed to recording of sensitive questions	Agreed	0.00 Ref.	0.00 Ref.	0.00 Ref.	3,927
	Refused	0.77 ***	-0.25 -	0.31 -	114
Time taken for questionnaire in 2005	20-44 minutes	0.02 -	0.17 -	0.08 -	423
	45-59 minutes	0.00 Ref.	0.00 Ref.	0.00 Ref.	1,301
	60-74 minutes	-0.16 -	-0.07 -	-0.16 -	1,126
	75+ minutes	0.01 -	-0.22 -	-0.06 -	949
	Not known	0.04 -	0.18 -	0.11 -	242
Change in interview conditions	Other person present in W1, none in W2/W3	0.30 **	-0.04 -	0.20 **	832
	Other situations	0.00 Ref.	0.00 Ref.	0.00 Ref.	3,209
Under-reported in 2008 (Wave 2)	No		0.00 Ref.		
	Yes		1.62 ***		
R ²		0.06	0.09	0.08	
Number and %	Under-reported	443 11.0	498 12.3	770 19.1	4,041
	Same number of children	3,598 89.0	3,543 87.7	3,271 80.9	

Interpretation: A statistically significant positive (negative) coefficient indicates that this is a factor that increases (decreases) the probability of reporting fewer children than in 2005.

Note: unweighted data.

Significance levels: * 10%; ** 5%; *** 1%; – non significant.

Coverage: respondents aged 18-79 in 2005 who took part in all three survey waves and reported at least one child in 2005.

Source: INED-INSEE, ERFI-GGS123, 2005-2011.

parent-child relationship after separation is different for fathers and mothers (Régnier-Loilier, 2013), an interaction with the respondent's sex was included in the models;

- Respondent's attitude to the survey in 2005: refusal to answer certain questions, such as those on household income, or to allow recording of answers to "sensitive" questions (civil partnership and religion) as defined by the French data protection authority, refusal to receive initial results (sign of less interest in the study topic), duration of interview (the respondent may seek to shorten the interview time in subsequent waves if he or she has memories of a lengthy earlier questionnaire);
- Different interview conditions (presence or not of another person) with respect to Wave 1: the partner present at the interview may reduce the risk of misunderstanding a question and may be able to correct a wrong answer, mention an omission, etc.

Coverage was restricted to respondents who took part in all three survey waves and reported at least one child in 2005 (otherwise no under-reporting would be possible).

Overall, the explanatory power of Models A, B and C is extremely weak; the factors analysed have little effect on under-reporting. Any effect comes mainly from the characteristics of the "types" of children reported in 2005. In all three models, the probability of under-reporting is significantly lower if the respondent had no non-resident children in 2005, which confirms our initial observations: under-reporting the total number of children is closely related to the existence of non-resident children in Wave 1. However, the actual number of non-resident children (1, 2, 3 or more) has no effect. Under-reporting also appears to be more probable if the respondent mentioned having one or more deceased children in Wave 1. These children tend to be not mentioned again in subsequent waves. The existence of resident children in 2005 has little effect on the probability of under-reporting children in subsequent waves (only in Model C does any effect become visible).⁽¹⁷⁾ Finally, reporting one or more estranged children in 2005 (who never see the respondent) increases the probability of reporting fewer children in one or more following waves. The factor of interaction with the respondent's sex is also significant (except in Model B) and shows a much less pronounced effect of under-reporting among women. This result supports the hypothesis that men fail to report some children after separation, a finding that was observed in the 1999 family survey (*Enquête famille*) (Toulemon, 2005).

Changes in interview conditions after Wave 1 also significantly affect Models A and C: where another person was present in 2005 but the respondent was interviewed alone in a subsequent wave, the probability of under-reporting is higher. The presence of another person, usually the partner, may therefore

(17) An interaction between numbers of resident and non-resident children was tested. No significant effect was found.

have a regulating effect on the answers given by the respondent, limiting the risk of omitting details or misunderstanding a question.

We noted above that the under-reporters were to some extent the same in Waves 2 and 3. In order to identify any cumulative under-reporting from one wave to the next, we included in Model B an indicator of under-reporting in Wave 2. The effect is significant and this is the variable whose estimated parameter is highest in the model (1.62), revealing a much higher probability that Wave 2 under-reporters again omitted to report some children in Wave 3, all other things being equal.

Apart from the effect of these two variables and the respondent's parental situation, hardly any of the other characteristics included in the models affect the probability of under-reporting: educational level does not, and neither do sex, or age among the over-50s (younger people under-report less, cf. notes to Figure 3). Attitude to the survey, interview conditions, and even interview duration are equally devoid of effects, so there is little evidence to support the hypothesis of an "avoidance strategy" on the part of the respondents.

Conclusion and discussion

The quality of the data collected on children in the successive waves of the ERFI-GGS survey is open to question. Indeed, this analysis not only confirms the under-reporting of deceased children in the surveys, but also reveals major inconsistencies in the number of non-resident children reported by respondents of both sexes in the three successive survey waves. On average, respondents aged 50 and over in 2005 reported 0.25 children fewer in Wave 2 and Wave 3 than in Wave 1. While double-counting of children in Wave 1 might be suspected, this explanation can be ruled out. The inclusion of a summary question on the respondent's total number of children in the last wave has enabled us to validate the data collected in Wave 1 and supports the hypothesis of under-reporting in Waves 2 and 3.

A recent study based on the UK Continuous Household Survey showed that the proportion of childless respondents tends to increase across successive waves, and suggested a number of explanatory hypotheses: changes in interviewers' instructions, survey reorganization, changes in data collection method, and the substantial lengthening of the questionnaire probably all contributed to less accurate reporting of children (Ní Bhrolcháin et al., 2011). However, for the French ERFI-GGS survey, none of these factors would appear to explain the omissions observed: the questionnaire architecture was identical in the three waves, question wording was similar, and the data collection method unchanged. Only a change in interview conditions, with another person present in Wave 1 and the respondent being interviewed alone in Wave 2, increases the probability of under-reporting.

This suggests the existence of a “learning effect” or “panel effect”: in successive waves the respondents might get used to the questionnaire filtering system and anticipate that answering Yes to the question “Have you had any children?” will then lead on to a series of questions about each child. Omissions might thus result from respondents’ avoidance strategies. But this hypothesis seems a weak one for several reasons. First, the results show that the interview time in Wave 1 does not affect the probability of under-reporting in subsequent waves. Second, since the children’s details were recorded towards the start of the questionnaire, it seems unlikely that the respondent would already be fed up at that early stage. Third, most of the omissions noted in Wave 2 do not reoccur in Wave 3. Fourth, in a study of characteristics linked to attrition between Waves 2 and 3 (Régnier-Loilier and Guisse, 2012), we included as an explanatory variable the fact of having reported fewer children in 2008 than 2005, on the assumption that under-reporting might reflect lassitude on the part of the respondent and therefore be associated with greater attrition, but no significant effect was found.

The absence of any effect of individual characteristics on under-reporting (apart from the number of resident, non-resident and deceased children in Wave 1) and the volatility of under-reporting from one wave to another (under-reporters in Wave 2 and Wave 3 are not always the same people) reveal a somewhat random phenomenon that is hard to understand. By a process of elimination, the most convincing explanation to date is the vagueness of the question used to record non-resident children when this occurs in a later wave. In the Wave 2 and 3 questionnaires, the respondents were asked whether they had “adopted or had children” (or “other children” in the case of those who had reported at least one resident child) but no explicit mention was made of having “children not living in the household” or “living somewhere else”. Some respondents may have considered that as they had already reported their children in the earlier wave(s), or, more simply, that they had not had any children since the last wave, there was no point in repeating the information. This hypothesis is all the more plausible given that the omissions mainly concern people who were no longer living with their children at the time of the survey and who were asked “Have you given birth to/ fathered any children or have you ever adopted any children?” after first being told “We already talked about those children who currently live in your household” for those who had reported at least one resident child. Indeed, this question followed immediately on from a question referring to the previous 12 months, “Over the last 12 months, have you given regular help with childcare to other people?” Some respondents may well have understood the later question as meaning “Over the last 12 months... Have you given birth to/ fathered any children or have you adopted any children?” However, this explanation is a tenuous one, because the sequence of questions remained the same in all waves. How then are we to explain why these omissions did not occur in Wave 1?

Although this study does not provide a clear explanation for the under-reporting of children in the ERFI-GGS survey, it does serve two purposes. First, it alerts data users to the problem so that they can allow for it in their use and analysis of the datasets. To date, no imputation has been performed on the disseminated data, either nationally or in harmonized international databases.⁽¹⁸⁾ It is therefore important to retrieve the “missing” children from survey Wave 1 for any studies using the data from Waves 2 and 3. Similarly, the reporting of grandchildren (a section introduced in Wave 3) is also biased, since this question was only asked of respondents who had at least one child aged 16 or over. Second, it draws the attention of the designers of future longitudinal surveys to this risk of under-reporting when information about the number of children is obtained indirectly from various points in the questionnaire rather than from a direct question. The method of summing resident, non-resident and deceased children to establish the total number of children appears to have been reliable in the French survey, but only in Wave 1. Since the French survey used CAPI techniques, it would perhaps have been prudent to include in Waves 2 and 3 a “validation/correction” procedure for data collected in the previous wave (for example, recalling the previous situation of the household and asking if it had changed).⁽¹⁹⁾ It will be interesting to see whether this problem has occurred in other surveys, particularly the GGS surveys in other countries, once the data become available. A methodological study after the first wave of the German GGS questioned the reliability of the reproductive histories recorded. The authors concluded that the process of recording non-resident children was too complex and regretted that a simple summary question had not been included (Kreyenfeld et al., 2013), an opinion we support here.

Acknowledgements: The author warmly thanks the referees and the editorial committee for their careful revision of the text and for their constructive suggestions. The ERFI-GGS survey was funded by INED, INSEE, ANR, COR, CNAF, CNAV, DARES, DREES and iPOPs.

(18) Indeed, the very purpose of this investigation was to understand the nature of inconsistencies in answers between waves, in order to make any necessary imputations or corrections. However, no directive on this point has been included in the instructions for international data harmonization and dissemination.

(19) This type of procedure is only possible with computer-assisted data collection. It can also have unintended effects. Some respondents may doubt whether the survey is really confidential if the interviewer recalls information recorded three years earlier, and may consequently not wish to take part in the following wave. Others may dispute the data collected in the previous wave.

REFERENCES

- AURIAT N., 1996, *Les défaillances de la mémoire humaine. Aspects cognitifs des enquêtes rétrospectives*, Ined/PUF, Cahier no. 136, 204 p.
- FIRDION J.-M., 1993, "L'effet du rang d'appel et de la présence du conjoint dans une enquête par téléphone", *Population*, 48(5), pp. 1281-1314.
- HERTRICH V., 1997, "Les réponses des hommes valent-elles celles des femmes ? Une double collecte sur les questions génésiques et matrimoniales dans une population au Mali", *Population*, 52(1), pp. 45-62.
- JOYNER K., PETERS H. E., HYNES K., SIKORA A., RUBENSTEIN TABER J., RENDALL M. S., 2012, "The quality of male fertility data in major U.S. Surveys", *Demography*, 49(1), pp. 101-124.
- KREYENFELD M., HORNING A., KUBISCH K., 2013, "The German Generations and Gender Survey: Some critical reflections on the validity of fertility histories", *Comparative Population Studies*, 38(1), pp. 3-28
- MAZUY M., LELIÈVRE É., 2005, "Déclarer ses enfants, déclarer ses conjoints : rationalité des locuteurs et mode de questionnement", in Lefèvre C., Filhon A. (eds.), *Histoires de familles, histoires familiales. Les résultats de l'enquête Famille de 1999*, INED, Cahier no. 156, pp. 573-588.
- MAZUY M., RAZAFINDRATSIMA N., LA ROCHEBROCHARD É. DE, 2005, *Déperdition dans l'enquête "Intentions de fécondité"*, INED, Document de travail no. 129, 36 p.
- NEELS K., DE WINTER T., VERMANT G., 2011, "The quality of demographic data in GGS Wave 1", paper presented at the First User Conference of the GGP "Demographic and social challenges in an aging Europe", 23-24 May, Budapest, Hungary.
- NÍ BHROLCHÁIN M., BEAUJOUAN É., MURPHY M., 2011, "Sources of error in reported childlessness in a continuous British household survey", *Population Studies*, 65(3), pp. 305-318.
- POULAIN M., RIANDEY B., FIRDION J.-M., 1992, "Data from a life history survey and the Belgian Population Register", *Population, An English Selection*, 4, pp. 77-96.
- RAZAFINDRATSIMA N., KISHIMBA N., 2004, "Attrition in the COCON Cohort between 2000 and 2002", *Population, English Edition*, 59(3-4), pp. 357-386.
- RÉGNIER-LOILIER A., 2006, *Présentation, questionnaire et documentation de l'Étude des relations familiales et intergénérationnelles (Erfi). Version française de l'enquête Generations and Gender Survey (GGS)*, INED, Document de travail no. 133, 232 p.
- RÉGNIER-LOILIER A., 2007, "Conditions de passation et biais occasionnés par la présence d'un tiers sur les réponses obtenues à l'enquête Erfi", *Économie et statistique*, 407, pp. 27-50.
- RÉGNIER-LOILIER A., 2010, *Présentation, questionnaire et documentation de l'Étude des relations familiales et intergénérationnelles (Erfi-GGS2)*, INED, Document de travail no. 165, 212 p.
- RÉGNIER-LOILIER A., 2012, *Présentation, questionnaire et documentation de la troisième vague de l'Étude des relations familiales et intergénérationnelles (Erfi-GGS 2011)*, INED, Document de travail no. 187, 328 p.

- RÉGNIER-LOILIER A., 2013, "When fathers lose touch with their children after a separation", *Population and Societies*, 500, 4 p.
- RÉGNIER-LOILIER A., 2014, "How are household chores divided? Responses vary with the respondent's gender and the partner's presence or absence during the interview", in Régnier-Loilier A. (eds.), *The Contemporary Family in France: Partnership Trajectories and Domestic Organization*, Springer, INED Population Studies, (forthcoming).
- RÉGNIER-LOILIER A., GUISSÉ N., 2012, "Attrition entre les trois vagues d'enquête", in Régnier-Loilier A. (ed.), *Présentation, questionnaire et documentation de la troisième vague de l'Étude des relations familiales et intergénérationnelles (Erfi-GGS 2011)*, Document de travail no. 187, pp. 29-43.
- RÉGNIER-LOILIER A., SABONI L., VALDES B., 2011, *Presentation and Modifications to the GGS Questionnaire in France (Wave 2)*, INED, Document de travail no. 173, 146 p.
- RENDALL M. S., CLARKE L., PETERS H. E., RANJIT N., VERROPOULOU G., 1999, "Incomplete reporting of men's fertility in the United States and Britain: A research note", *Demography*, 36(1), pp. 135-144.
- TOULEMON L., 1997, "The fertility of step-families: The impact of childbearing before the current union", Paper presented at the annual meeting of the Population Association of America, 29-31 March, New York.
- TOULEMON L., 2005, "Enfants et beaux-enfants des hommes et des femmes", in Lefèvre C. and Filhon A. (ed.), *Histoires de familles, histoires familiales. Les résultats de l'enquête Famille de 1999*, Cahier no. 156, pp. 59-77.
- VIKAT A., SPÉDER Z., BEETS G., BILLARI F., BÜHLER C. et al., 2007, "Generations and Gender Survey (GGS): Towards a better understanding of relationships and processes in the life course", *Demographic Research*, 17(14), pp. 389-440.

Arnaud RÉGNIER-LOILIER • INCONSISTENCIES IN THE NUMBER OF CHILDREN REPORTED IN SUCCESSIVE WAVES OF THE FRENCH GENERATIONS AND GENDER SURVEY

The ERFI survey (*Étude des relations familiales et intergénérationnelles*) is the French version of the international Generations and Gender Survey. The same respondents were interviewed three times, in 2005, 2008 and 2011. Although the survey was designed to avoid redundant questions from one wave to the next, a respondent's family situation is likely to change over time. In each wave respondents were therefore asked about their children living in the dwelling, those living elsewhere and those deceased. The results of these separate questionnaire modules were then summed to give the total number of children (no direct question about the total was included). Substantial under-reporting of children in Waves 2 and 3 of the French survey was observed. This article aims to measure the extent of the omissions in order to alert potential users of the data, to identify which children "disappear" and establish whether this corresponds to particular respondent profiles. More broadly, the study calls into question the method of summing the children recorded in various parts of the same questionnaire in order to determine the respondent's total number of children, and suggests that similar verifications should be made in other GGS surveys.

Arnaud RÉGNIER-LOILIER • INCOHÉRENCE DU NOMBRE D'ENFANTS DÉCLARÉS ENTRE LES VAGUES DE L'ENQUÊTE FRANÇAISE GENERATIONS AND GENDER SURVEY

L'enquête Erfi (*Étude des relations familiales et intergénérationnelles*) est la déclinaison française de l'enquête internationale et longitudinale GGS (*Generations and Gender Survey*). Les mêmes personnes ont été interrogées à trois reprises, en 2005, 2008 et 2011. Si l'enquête est conçue de façon à éviter une redondance du questionnement d'une vague à l'autre, la situation familiale de la personne est susceptible d'évoluer au cours du temps. On décrit ainsi à chaque vague les enfants du répondant vivant dans son logement, ceux qui résident ailleurs et les enfants décédés. De ces différents modules du questionnaire est déduit par addition le nombre total d'enfants (aucune question directe sur le nombre total d'enfants du répondant n'est prévue). Un important phénomène de sous-déclaration des enfants est mis à jour lors des vagues 2 et 3 de l'enquête française. L'objectif de cet article est de mesurer l'ampleur des omissions afin d'avertir les utilisateurs potentiels des données, de repérer quels enfants « disparaissent » et de voir si cela correspond à des profils de répondants spécifiques. Cette étude interroge plus généralement la méthode du décompte des enfants décrit dans différentes parties d'un même questionnaire pour déterminer le nombre total d'enfants d'une personne et invite à procéder à des vérifications analogues dans les autres enquêtes GGS.

Arnaud RÉGNIER-LOILIER • INCOHERENCIA DEL NÚMERO DE NIÑOS DECLARADOS EN LAS TRES OLEADAS DE LA ENCUESTA FRANCESA GENERATIONS AND GENDER SURVEY

La encuesta Erfi (*Étude des relations familiales et intergénérationnelles*) es la declinación francesa de la encuesta internacional y longitudinal GGS (*Generation and Gender Survey*). Las mismas personas han sido interrogadas tres veces consecutivas, en 2005, 2008 y 2011. Bien que la encuesta esté concebida para evitar una redundancia en las preguntas de una oleada a otra, la situación familiar de la persona puede variar en el curso del tiempo. Se pregunta así en cada oleada sobre los hijos del encuestado que viven en su alojamiento, los que viven en otro domicilio y los que han muerto. De estos diferentes módulos del cuestionario se deduce por adición el número total de hijos (en ningún momento se hace una pregunta directa sobre dicho número). En las dos últimas oleadas de la encuesta francesa se ha producido un número importante de omisiones en la declaración de los hijos. El objetivo de este artículo es medir la importancia de estas omisiones a fin de advertir a los usuarios potenciales de los datos, de identificar qué niños "desaparecen" y ver si ello corresponde a perfiles específicos de encuestados. De manera más general, este estudio cuestiona el método de cómputo de los hijos descrito en las diferentes partes de un mismo cuestionario para determinar el número total de hijos de persona e invita a proceder a verificaciones similares en las otras encuestas GGS.

Keywords: Generations and Gender Survey (GGS), *Étude des relations familiales et intergénérationnelles* (ERFI), survey methodology, longitudinal survey, data quality, reported number of children, data collection.

Translated by Roger Depledge.