SUPPLEMENT ON THE ESTIMATION OF VARIANCE FOR THE ELFE SURVEY

Simplification of the sampling design, its impact on variance calculations, recommendations to users

This document is intended for users of data on the ELFE cohort. It is intended to complement the document [INED Working Document 226 (WD226), not yet translated: document_travail_2016_226_estimation.de.variance_e.chantillonnage.produit.fr.pdf (ined.fr)].

The reader is **strongly advised to read that document first**.

This document assumes that you are already familiar with the main concepts on sampling and variance presented there, as well as the contextual aspects of the ELFE survey.

After a brief overview of the main conclusions of WD226, we discuss an important aspect of the simplification of the sampling design, explaining why we considered it an appropriate choice. **Contrary to the simplifications proposed and analysed in WD226, the simplification discussed in this document is not a direct simplification of the method of calculating the variance estimator. It is instead a more conceptual simplification of the ELFE survey's sampling design, and of the analysis of its impact in the calculation of variance.**

We analyse how the different steps in the constitution of the sample influence the estimated variance of a given variable under the proposed simplifying hypothesis. We do so by comparing the contributions of different elements to the estimated variance in the ELFE survey to variance under a simple random sampling.

We simulate these calculations on over fifty variables drawn from the survey waves in the maternity units and when the children in the cohort were 2 years old. We show that, with a few precautions, this simplified sampling design can be approached using using classical software procedures associated with a simple random sampling design.

We then perform the same analyses on a selection of variables drawn from the survey wave at the age of 3½ years to check that the analysis is sufficiently constant over time.

As much as possible, this document is based on simulations performed using SAS 9.4 (SAS Institute Inc., 2013).

Contents

1. Main results from INED Working Document 226[1]

The sampling design used for the ELFE survey is not a standard one. It is the product of two independent samplings followed by multiple nonresponse phases and a calibration phase.

The population of inference consists of all infants born in 2011 in metropolitan France: in at most a twin birth; not highly premature; to an adult mother, who was able to give informed consent, in particular in one of the languages offered; in a maternity unit in metropolitan France; to parents who were not temporary residents of metropolitan France. All of the selected families were surveyed shortly after childbirth in a selection of metropolitan maternity units on one of a selection of days in 2011.

The **maternity units** were selected using a **stratified probabilistic sampling design**, with five strata of equal numbers of maternity units. Their sampling rate was proportional to the mean number of births recorded in maternity units in each stratum in 2008.

Twenty-five days were chosen in four multi-day periods, one in each of the four seasons. The days were not selected randomly, but selected by hand (half had to coincide with the days of the French Permanent Demographic Sample, or EDP).

The two samples (maternity units and days) were selected independently.

Strates g	Nb d'accouchements par maternité en 2008	Taille dans la population N_g	Taille de l'échantillon n_g
1	[145-699[	108	28
2	[700-1009[	108	47
3	[1010-1418[	109	66
4	[1422-2187[	108	97
5	[2197-5215[	111	111
TOTAL		544	349

Vague h	Taille dans la population M_h	Taille de l'échantillon m_h
1	90	4
2	91	6
3	92	7
4	92	8
TOTAL	365	25

Figure 1 - Sizes of strata and seasonal survey periods1
Here, 'population' refers respectively to all maternity units in metropolitan France and to all days of the year

Twenty-nine out of the 349 selected maternity units did not participate in the ELFE survey. In addition, out of the other 320 selected maternity units, some did not participate in all seasonal survey periods: 15 did not participate in the first period, 8 in the second, 9 in the third, and 11 in the fourth. With relatively low nonparticipation rates for maternity units (7%) and days (3% on average), these first two nonresponse phases are not included in the variance calculation. They are dealt with by simply adjusting the inclusion probabilities, and thus the weight, for each infant.

Under the hypothesis that this adjustment of inclusion probabilities is sufficient to deal with the first nonresponse phases, Hélène Juillard shows that an unbiased estimator of the ELFE sampling design variance can be decomposed into three elements:

$$\widehat{V_{prod}} = \widehat{V_D} + \widehat{V_M} - \widehat{V_E}, \text{ where}$$

$\widehat{V_D}$ estimates the variance due to the stratified selection of days

$\widehat{V_M}$ estimates the variance due to the stratified selection of maternity units, and

$\widehat{V_E}$ is a "cross effect" due to the fact that days and maternity units are identical (the days surveyed are the same for each selected maternity unit, and vice versa)

[1] This section is entirely based on INED WD226 by Hélène Juillard.

There is also a large nonparticipation phase at the infant level in the ELFE survey: 49% of the 36,000 families who were approached in the maternity units chose not to participate. This must, of course, be considered in calculating variance. To take into account total nonparticipation, the decision to participate is treated as random. This means that under the same conditions (e.g., age, income, nationality, etc.), the participation process will not always yield the same result. Some individuals will agree to participate, others will not. The sample of participants is thus drawn from an (n + 1)-phase selection process. The first *n* phases are selection phases (the infant is selected), and the last is an acceptance phase (the infant's parents agree or refuse to participate). The variance estimator is then:

$$\widehat{V_{ELFE}} = \widehat{V_D} + \widehat{V_M} - \widehat{V_E} + \widehat{V_{NR}} \tag{1}$$

Finally, a calibration is performed. To compute the variance estimator in this context, we conduct a weighted regression of the variable of interest on the calibration variables, and we calculate the variance by applying formula (1), not to the variable of interest, but to the regression residuals.

Note that the preceding parts concern estimating the variance of an estimated total. For other parameters (ratio, mean, etc.), variances can be estimated using linearization. As an example, here and throughout the rest of this document, we will take a ratio $R = \frac{t_Y}{t_X}$ (total of a variable *Y* divided by the total of a variable *X*). To estimate the variance of the estimated ratio $\hat{R}$, we must estimate the totals $\widehat{t_x}$, $\widehat{t_y}$, and then $\hat{R} = \frac{\widehat{t_Y}}{\widehat{t_X}}$, and, finally, for each individual *k*, calculate the linearized parameter, defined by:

$$lin_k = \frac{1}{\widehat{t_X}} (y_k - \hat{R}.x_k). \tag{2}$$

To take the calibration step into account, we then regress lin_k on the calibration variables and use the residuals ε_k of this regression as variables in formula (1).

2. What are we trying to measure?

Most importantly, we must specify what we are trying to estimate. In the case of the ELFE survey, we are not seeking to estimate an average number of individuals or births **per day** (the daily number of infants with a given characteristic for example), but an **annual** total or ratio (the total annual number of infants with a given characteristic, the proportion with a given characteristic, etc.), an average (such as the average height of an infant), a score or a distribution that we want to quantify, make comparisons at a time *t*, or follow a population over time.

We thus only seek to calculate variances on elements whose statistical unit is the infant. To obtain a sufficient sample, surveying had to be performed in maternity units over multiple days, which were distributed throughout the year to facilitate the work of the surveyors. **But the fact that the survey was carried out on different days does not mean that the day should be considered an element of the sampling design.**

The date of birth evidently served to select individuals, but in subsequent survey waves, the date of birth no longer features as an element of the sampling design. All participants are surveyed in the same period (where necessary, distributing survey dates across seasons to ensure that the children's ages are relatively comparable), again on different days in order to facilitate the work of the interviewers. Date of birth thus no longer plays a role in selecting or surveying families.

The principle of stratification, in survey research, is to constitute a priori homogeneous groups of individuals, among which a random selection is made. In this way the estimates obtained for each of the groups (and therefore the final sample) is less likely to be randomly unrepresentative. The more homogeneous the individual groups, and the more heterogeneous the different groups, with respect to the variables of interest, the more stratification improves the precision of the survey.

Another reason for stratification may be to ensure sufficient precision concerning a particular population of interest. In this case, the subpopulation of interest is overrepresented in the selection by applying different sampling rates to different groups.

Neither of these two reasons is relevant to the day of birth in the ELFE survey. The days in the different seasonal periods are similar to each other. No overrepresentation of a given season is needed, and the infants' characteristics do not depend on the seasonal period in which they were surveyed (this does not mean that all days are identical, but that the days in different periods are not more different from each other than days within a given period. The concept of seasonal sampling periods thus has no theoretical meaning for the calculation of variance).

To convince ourselves of this, we plot the proportions of infants with certain characteristics by day of birth. No relationship can be established. The largest fluctuations (in type of childbirth) are due to the type of day (Saturday, Sunday, weekday) and not to the seasonal period.

Figure 2 – Proportion of infants with certain characteristics/mean height and weight of infants by day of birth; 2

We thus consider the day as an aspect of the surveying process that allowed ELFE to obtain a large enough sample by facilitating recruitment and the work of interviewers, not as an element of the sampling design as such. **We thus have a sampling design stratified only on maternity unit size.**

This simplification obviously would not make sense for the selection of maternity units. Not all maternity units could be surveyed (it thus had to be clearly established that they were the result of a random selection), infants differ considerably between maternity units, and weights and sampling rates differ considerably between strata (the stratification is thus needed to improve the estimates and precision calculations). Data on contextual, social, childbirth, and even health characteristics vary considerably among maternity units. All of this gives us reason to concentrate on on this element in our calculations. Here we see that the proportions of infants with certain characteristics vary considerably between maternity units:

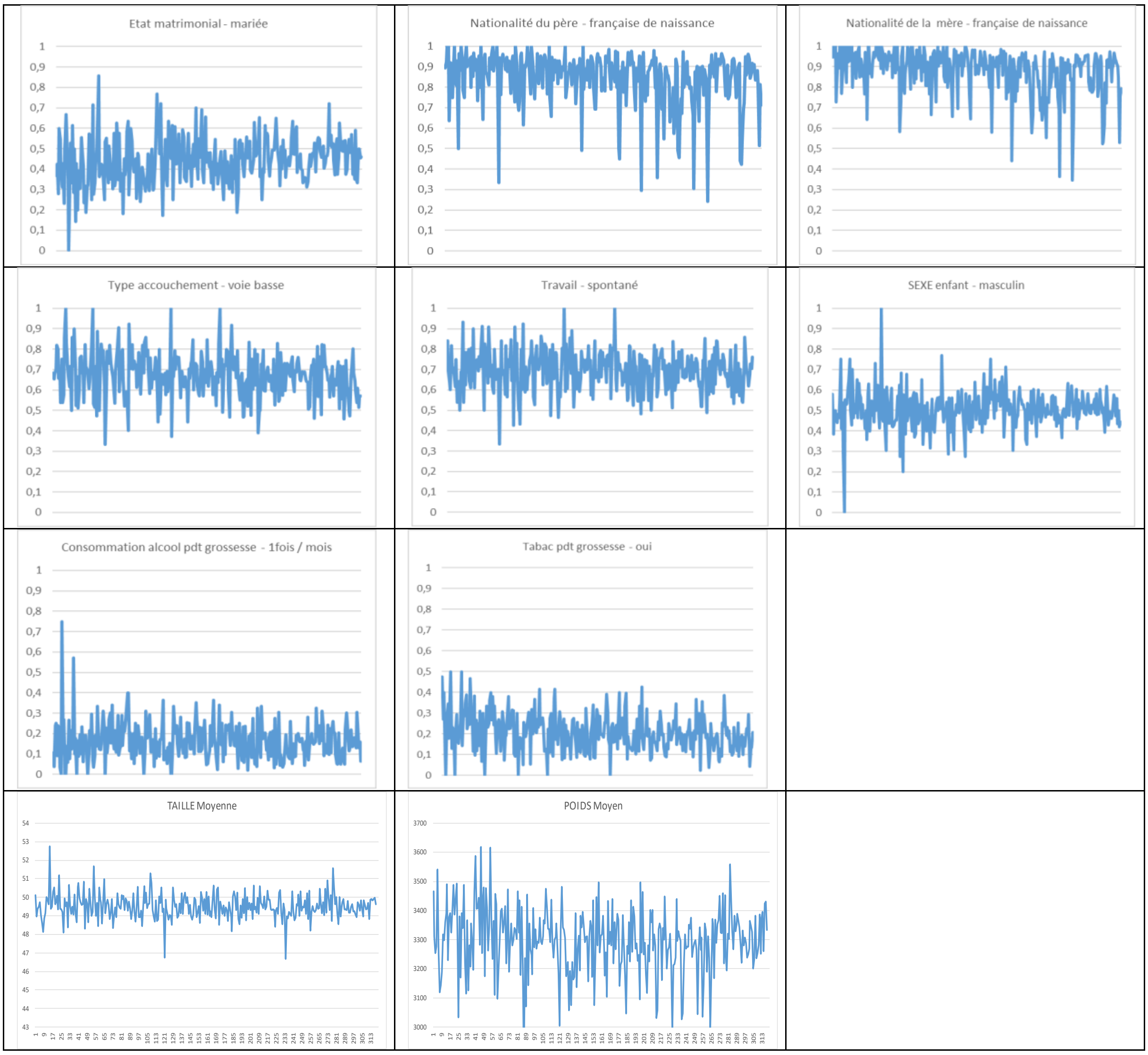

Figure 3 - Proportion of infants with certain characteristics by maternity unit3

Thus, if we accept the proposed simplification, the ELFE survey's sampling design can be schematized as follows:

- Stratification of maternity units: maternity units are divided into 5 strata based on their size;
- Random selection of maternity units within each stratum (allocation proportional to the number of births recorded in 2008);
- The families of all infants born in these maternity units approached;
- A nonresponse process for each of these selections (negligible at the maternity unit level, significant at the infant level).

3. Quantifying the simplification of the sampling design

The objective of this section is to measure the **effect of the simplification** proposed in the previous section and check the correctness of the assumption that date of birth has a negligible effect compared to maternity unit of birth. After presenting the items used to measure these effects, we calculate them for a large selection of variables.

We define the following elements:

y_k = 1 if infant *k* has a given characteristic, or 0 otherwise (for example y_k = 1 if mother's place of birth is metropolitan France, 0 otherwise).

x_k = 1 if the value of a variable is available for infant *k* (in order to estimate a proportion excluding nonresponse: for example, x_k = 1 if mother's place of birth is known, 0 otherwise).

We are interested in the ratio weighted $\hat{R} = \frac{\widehat{t_Y}}{\widehat{t_X}}$ (the proportion of individuals with a given characteristic out of all those whose place of birth, for example, is known). The linearized variable lin_k is then generated for each individual *k* using formula (2).

We then estimate:[2]

- the proportion of variance due to the selection of maternity units, by calculating the variance in the classical context of a stratified (by maternity unit size) cluster selection (of maternity units).

$effetMAT: \widehat{V_M}$

```
proc surveymeans total= degreM;
cluster maternite; strata strate;
var lin;
weight poids;
run;
```

Where "degreM" represents the number of maternity units per stratum.

- the proportion of variance due to the selection of dates, by calculating the variance in the classical context of a stratified (by season) cluster selection (of dates).

$effetDAY: \widehat{V_D}$

```
proc surveymeans total= degreJ;
cluster jour; strata vague;
var lin;
weight poids;
run;
```

Where "degreJ" represents the number of days in each seasonal period.

- the proportion of the variance due to infant nonparticipation, which amounts to adding a stage to the ELFE survey's selection process. Maternity unit x Day pairs which accepted to participate were selected by some procedure. This new selection stage is a Poisson sampling (among all the infants belonging to the participating Maternity unit x Day pairs, individual infants are selected with probability $\emptyset_k$ and rejected with probability $1 - \emptyset_k$). By decomposing the variance, it can be shown that the unbiased estimate of the variance due to infant nonparticipation is given by:

$effetNR: \widehat{V_{NR}} \quad \sum_{\substack{\text{participating} \\ \text{infants} \\ \text{only}}} \left(\frac{{lin_k}^2}{{\pi_k}^2} \frac{1-\emptyset_k}{{\emptyset^2}_k} \right)$, with π_k the probability of infant *k* being selected, and $\emptyset_k$ the estimated probability that each selected infant will participate.

We can thus quantify:

$$ELFE\ variance^3 = effetMAT + effetDAY + effetNR$$

On order to understand the justification for the simplification proposed above, two other important elements must be evaluated, in addition to these calculations to estimate the variance of a ratio in the selection of infants to participate in the ELFE survey. We calculate, for each maternity unit *i*, the ratio of infants with the measured characteristic, as well as the mean ratio:

$$Ratio_{maternity\ unit\ i} = \sum_{\text{infants k from maternity unit i}} (y_k) \Big/ \sum_{\text{infants k from maternity unit i}} (x_k)$$

[2] Here again, see INED Working Document 226 – Hélène Juillard – May 2015.

[3] As in INED Working Document 226, the "cross effect" is negligible here.

$$\overline{Ratio}_{maternity\ units} = \left(\frac{1}{320}\right) \sum_{\text{maternity unit i}} Ratio_{maternity\ unit\ i}$$

And then:

$$dispersionMAT = \left(\frac{1}{320}\right) \sum_{\text{maternity units}} \left(Ratio_{maternity\ unit\ i} - \overline{Ratio}_{maternity\ units}\right)^2$$

Analogously, for each day *j* we calculate the proportion of infants with the measured characteristic as well as the average ratio, and then:

$$dispersionDAY = \left(\frac{1}{25}\right) \sum_{\text{days}} \left(Ratio_{day\ j} - \overline{Ratio}_{days}\right)^2$$

These different effects are presented below for 53 variables collected from the initial survey in maternity units (prefix M00) to the survey wave when the children were 2 years old (prefix A02). These include variables concerning sociodemographic characteristics (parents' place of birth, nationality, marital status, employment status, etc.), health (smoking, alcohol, etc.), pregnancy and childbirth (diabetes, hypertension, type of childbirth, etc.), and the child's activities (drawing, puzzles, etc.). All 53 variables are listed in Appendix 1.

To estimate these effects, we calculate the proportion of infants with a value of 1 for each of these variables, as well as for all the elements defined above. The data are weighted by the cross-sectional Child weighting of the relevant survey wave for the analysed variable.

The simple quantification of the contributions of different elements to the variance thus shows the preponderant place of the effect of the day in the estimated variance: around 48%, versus 23% for the effect of maternity units and 29% for the NR effect (the precise proportion of each is given in Appendix 2).

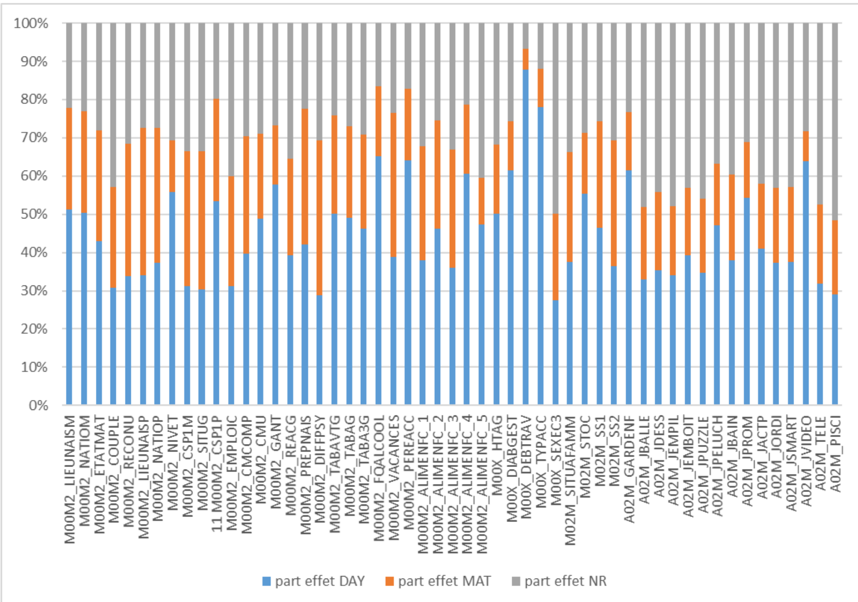

Figure 4 - Share of the effects of day, maternity unit, and NR in theoretical variance[4]

Note: Figure 4 presents, for each indicator, the proportion of the variance due to the effects of the maternity unit (orange), the day (blue) and nonresponse (gray), calculated using the formula: ELFE variance = effetMAT + effetDAY + effetNR. The figure shows the preponderant share of the day (around 40-60%), versus 20-30% for the maternity unit and 20-30% for nonparticipation.

Obviously, the day effect represents the largest proportion of variance for the variables with the largest daily fluctuations (notably for data on childbirth). Conversely, its proportion is lowest for the variables that are the most independent of the day (e.g., sex, employment situation). It is thus clear that the simplification of the sampling design proposed above has a sizeable effect on the theoretically estimated variance in an analysis based on the ELFE survey's complete sampling design.

It is important to be very clear, however: the day and maternity unit effects are based on variance calculations in the case of two stratified surveys. These variances thus depend on two elements: the sampling rate in each stratum, and the mean dispersion of the variables within the strata. But the survey rates are too different for the variability of the data not to be analysed in more detail (the survey took place on between 1 in 11 days and 1 in 22 days, depending on the season, while between a quarter and nine tenths of maternity units were surveyed, depending on the stratum).

The precision of the survey must thus be analysed based on a criterion other than the mere ratio of variances: the mean inter-day dispersion is negligible compared to the mean inter-maternity unit dispersion. While the selection of days is responsible for about 2.5 times more of the total variance than the selection of maternity units, **mean inter-day dispersion is approximately 35 times lower than dispersion between maternity units** (the precise share of each element is given in Appendix 2).

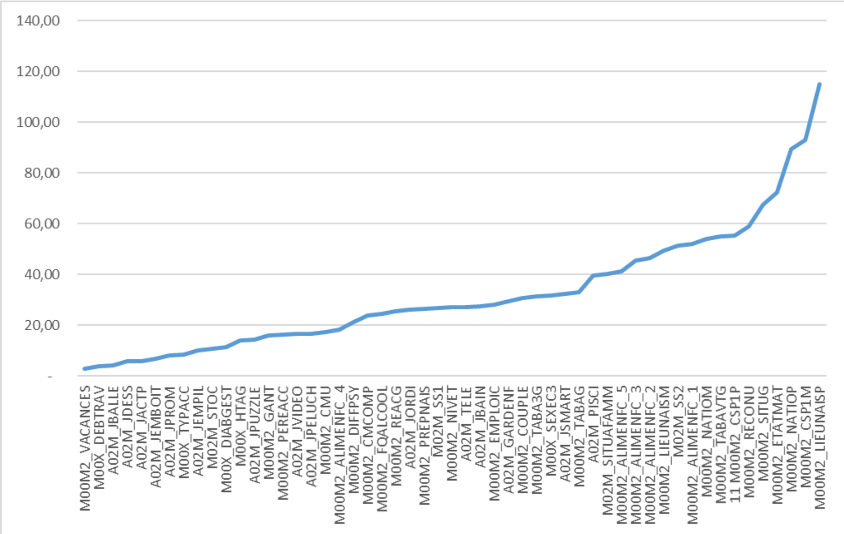

Figure 5 - Ratio of inter-maternity unit dispersion to inter-day dispersion5

It is thus the sizeable difference in the sampling rate and its integration in the calculation of variance, and not fluctuations in the data, which leads to the large imbalance between the effects of the day and the maternity unit. If the survey had been carried out on different days, the results would have been the same or nearly the same. Thus, while survey theory allows for an effect of the survey day to be taken into account, it does not seem illogical to disregard it in calculating the precision of analyses of the ELFE survey data: both because the day of birth can be seen as not constituting an element of the sampling design as such, and because it only slightly affects the precision of the results.

In the remainder of this document, we thus propose to schematize the analysis of variance of the ELFE survey as follows:

- Stratification of maternity units: maternity units are divided into 5 strata by size (stratification effect), and maternity units are selected randomly within each stratum;
- Theoretical surveying of all infants born in these maternity units, but with an infant nonparticipation process – integration of a second selection stage wherein only those who agreed to participate did so (cluster effect due to the second stage + nonparticipation effect);
- Calibration (calibration effect).

4. Definition of the elements of the variance estimate

The aim of this section is to present the elements used to quantify the effects in the simplified sampling design (**cluster effect**, **stratification effect**, **nonresponse effect**, and **calibration effect**) on the unbiased estimate of the variance of the proportion of infants with a given characteristic.

Thus, the following elements are defined as in the previous paragraph:

y_k = 1 if infant k possesses the characteristic studied, or 0 otherwise (for example y_k = 1 if mother's place of birth is metropolitan France, 0 otherwise).

x_k = 1 if the analysed variable is available for infant *k* (simply to estimate a proportion excluding nonresponse: for example, x_k = 1 if mother's place of birth is known, 0 otherwise).

The linearized variable lin_k is then generated for each individual *k* using formula (2).

The residuals ε_k resulting from the weighted regression of the variable lin_k on the calibration variables are also calculated.

We then estimate:

$\hat{V}_{SAS}(lin_k)$, which estimates the variance with simple random sampling, where a ratio is estimated in the absence of any other element of the sampling design:

```
proc surveymeans total=764000;
var lin;
weight poids;
run;
```

Note that the `surveymeans` procedure directly offers an option to calculate a ratio. The preceding calculation is thus equivalent to the command:

```
proc surveymeans total=764000;
ratio Y/X;
weight poids;
run;
```

The cluster selection of maternity units is taken into account by calculating $\hat{V}_{GR}(lin_k)$, corresponding to the estimated variance in a classical two-stage selection process (cluster of maternity units):

```
proc surveymeans total= 544;
cluster maternite;
var lin;
weight poids;
run;
```

Here again, this procedure is equivalent to:

```
proc surveymeans total=544;
cluster maternite;
ratio Y/X;
weight poids;
run;
```

The stratified selection of maternity units is taken into account by calculating $\hat{V}_{GR\_ST}(lin_k)$, which represents the estimated variance in a classical two-stage (clusters of maternity units) selection process with stratification (by size of maternity unit).

```
proc surveymeans total= degreM;
cluster maternite; strata strate;
var lin;
weight poids;
run;
```

Here again, this procedure is equivalent to:

```
proc surveymeans total=degreM;
cluster maternite; strata strate;
ratio Y/X;
weight poids;
run;
```

Where "degreM" gives the number of maternity units per stratum.

Note: in all of the preceding procedures, the variable 'TOTAL' represents the size of the population. The value of this variable is thus 764,000 for infants (the total number of children born in mainland France in 2011 who were eligible for the ELFE survey), and 544 for maternity units.

To take into account infant nonresponse (i.e. nonparticipation), we add:

$\hat{V}_{NR}(lin_k) = \sum_{\substack{participating \\ infants \\ only}} \left(\frac{lin_k^2}{\pi_k^2}\frac{1-\emptyset_k}{\emptyset^2{}_k}\right)$, with π_k the probability of an infant *k* being selected, and $\emptyset_k$ the estimated probability that a selected infant will participate.

We can thus estimate the effects of elements in the simplified ELFE sampling design as follows:

Cluster effect: $\hat{V}_{GR}(lin_k) \ / \ \hat{V}_{SAS}(lin_k)$

Stratification effect: $\hat{V}_{GR\_ST}(lin_k) \ / \ \hat{V}_{GR}(lin_k)$

NR effect: $\hat{V}_{GR\_ST}(lin_k) + \hat{V}_{NR}(lin_k) \ / \ \hat{V}_{GR\_ST}(lin_k)$

- ➔ **Sampling design effect:** $\hat{V}_{GR\_ST}(lin_k) + \hat{V}_{NR}(lin_k)/\hat{V}_{SAS}(lin_k)$

The sampling design effect measures the "quality difference" between the ELFE sampling design and a simple random survey. If this coefficient is greater than 1, the sampling design causes a loss of precision. It combines 2 predominantly antagonistic effects: the impact of stratification (precision improves if the analysed variable is fairly homogeneous among similar-sized maternity units), and the impact of nonparticipation (precision decreases since this adds a sampling stage, especially where the variable is linked to NR and where those with a lower estimated probability of responding are atypical).

- ➔ **Calibration effect:** $\hat{V}_{GR_{ST}}(\varepsilon_k) + \frac{\hat{V}_{NR}(\varepsilon_k)}{\hat{V}_{GR_{ST}}(lin_k)} + \hat{V}_{NR}(lin_k)$ (relation of total variance after calibration of total variance in the absence of calibration)

The calibration effect measures the extent to which calibration improves the precision of the estimators for each variable. Generally speaking, calibration always improves precision, limiting certain random effects by controlling the distribution of individuals across categories for a set of variables, and thus reducing uncertainty. If a variable depends at least a little on the calibration variables, the random effect necessarily decreases. If the variable is completely independent of the calibration variables, nothing is improved, but nothing is lost either.

The complete effect of the process implemented for the ELFE survey is thus:

- ➔ $\textbf{ELFE effect} = \textbf{sampling plan effect} \times \textbf{calibration effect}$

As in the previous paragraph, we calculated estimates of these effects on 53 variables collected from the survey in maternity units up to the survey wave when the ELFE children were 2 years old. Recall that all 53 variables are listed in Appendix 1.

The proportion of infants with a value of 1 for each of these variables was calculated, and all of the elements listed above calculated using the procedures as described. The data are weighted by the Child weights for the survey wave from which the analysed variable is drawn.

5. Quantifying the elements of the decomposition of variance

The cluster effect measures how the selection of maternity units and then infants affects the precision of the survey estimates. In most cases, it indicates a loss of precision due to similarities between infants in the same maternity unit. Surveying infants from the same maternity unit thus provides less information than if infants had been surveyed in a totally random fashion. The greater the cluster effect (the greater the ratio of the variance in the cluster selection to the variance from a simple random sample), the greater the loss of precision.

As an example, suppose we measure a cluster effect of 9. This means that the variance calculated with this sampling design is 9 times higher than if we had surveyed the same number of individuals by simple random sampling. To avoid this loss of precision due to the sampling design, then, it would have been necessary to survey three times more infants.

The cluster effect can be expressed more clearly and easily under certain conditions. For example, it can be shown that in case of a simple random survey of maternity units and a simple random survey of infants (although this is not the ELFE sample design), then the cluster effect can simply be written:
$\frac{\hat{V}_{GR}(y_k)}{\hat{V}_{SAS}(y_k)} = 1 + \rho(\bar{n} - 1)$, with $\bar{n}$ the mean number of infants surveyed per maternity unit and ρ proportional to $\sum_{maternity\ units\ i} \sum_{j=1}^{size\ mater} \sum_{\substack{k=1 \\ k \neq j}}^{size\ mater} (y_{ij} - \bar{y})(y_{ik} - \bar{y})$.

This highlights two important elements of the cluster effect: it depends on the average number of infants surveyed per maternity unit (it is obvious that if we surveyed a single child per maternity unit, there would be no cluster effect and the selection would be a simple random sample), and on the dispersion of the studied variable in infants from the same maternity unit with respect to the overall mean of all respondents (if the children in a given maternity unit fall on the same side of the overall mean, ρ will be a sum of positive values, whereas if the values for some infants from the same maternity unit are higher and others lower than the overall mean, ρ will be proportional to a sum of positive and negative values, and will thus be lower, or even negative). Thus, the more similar are the children from the same maternity unit, the less information is provided by surveying a new individual in it. This increases the cluster effect and decreases precision.

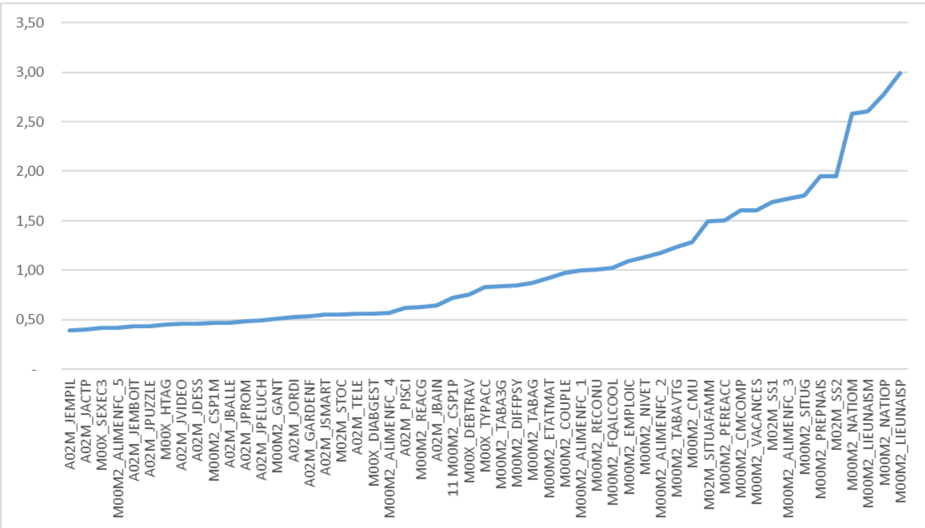

Figure 6 - Cluster effect6

Let us consider some extreme cases: children born to French mothers (M00M2_NATIOM) made up a larger proportion of infants born in some maternity units than in others (see Figure 10 below – the proportion of French mothers reached 100% in many maternity units). Including more children in this type of maternity unit thus does not provide as much information as expected. The cluster effect is high, and precision is decreased.

Conversely, there is no reason for boys (M00X_SEXE3) to be grouped together in certain maternity units (see Figure 8 below). Surveying a new infant in one or another maternity unit thus always provides just as much information. Indeed, as the proportion of boys varies very little between maternity units, the mean proportion of boys in each maternity unit can be estimated thanks to the cluster sampling while sampling less widely than is required for a random sampling of infants in all maternity units. This improves our overall estimate of the proportion of boys. The cluster effect is small (indeed, less than 1), and precision thus increases compared to a simple random sampling.

Note also that the cluster effect depends on the timing of the survey waves. It is obviously much lower for measurements when the children were 2 years old, both because fewer children were surveyed per maternity unit (12,000 in total, compared to 18,000 in the maternity units), and above all because there is less reason for children born in the same maternity unit to resemble each other as time goes on. By its very definition, the cluster effect will certainly be even weaker in future survey waves.

The **stratification effect** measures how stratification improves the precision of the estimates from the ELFE survey. Two elements can allow stratification to decrease the variance of a ratio: using the dispersion of the proportion of infants with a given characteristic in a maternity unit compared to the average rate in maternity units in the same stratum and thus of similar size (and not with the average rate in all maternity units), and using the sampling rate in each stratum.

In a stratified selection, the variance of an estimated mean is estimated as $\hat{V}_{GR\_ST}(\hat{\bar{y}}) = \sum_{stratum\ h} \left(\frac{N_h}{N}\right)^2 (1 - f_h).\ {s_h}^2/n_h$. The more similar the maternity units in each stratum (s_h, the dispersion of our variable of interest calculated within each stratum *h*, will be small if they are similar), or if many maternity units in the strata with the highest dispersion are surveyed (in which case, n_h, the number of maternity units surveyed in stratum *h*, will be high, and in $1/n_h$ this term will be able to compensate for the high dispersion), the more precise will be the estimates based on the stratified selection.

In the case of the ELFE survey, the sampling rate differs widely between strata, ranging from 23% of maternity units surveyed in stratum 1 to more than 80% in stratum 4, and even 90% in stratum 5. The stratification effect will thus be all the larger given the high variability in the large maternity units that make up stratum 5. This variability was thus compensated by a high sampling rate, strongly decreasing intra-stratum variance and thus overall variance. If, on the contrary, variability does not depend on the size of the maternity unit or if it is greater for small maternity units, stratification will have little effect.

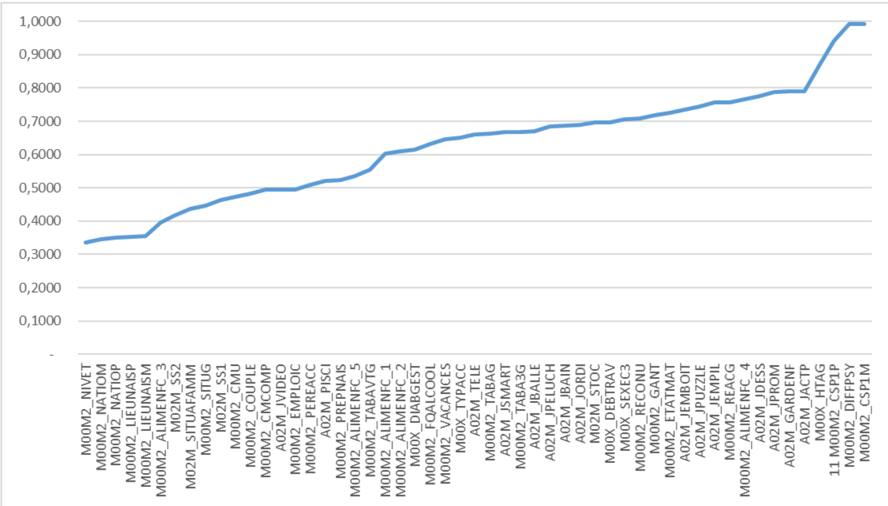

Figure 7 - Stratification effect7

It is important to observe that the stratification effect is relatively constant, especially compared to the other effects studied (between 0.6 and 0.8 for a large majority of the variables analysed). It may simply be noted that the stratification effect is larger for all of the variables concerning the parents' sociodemographic characteristics (place of birth, nationality, level of education, living with partner, etc.) and on access to public healthcare insurance, for example. It is much lower, on the other hand, for activities at the age of 2 years, sex, smoking, or marital status.

It can be verified that these rates do in fact depend on the stratification. In the following graphs, each variable is calculated per maternity unit and presented in order of unit size. Maternity units in the same stratum are represented by the same colour.

For the share of non-smoking mothers or boys born in each maternity unit, it can be seen that there is no link with the stratum. Dispersion even seems to be greater for small maternity units. The use of lower sampling rates where the dispersion of ratios is the highest makes the stratification effect negligible.

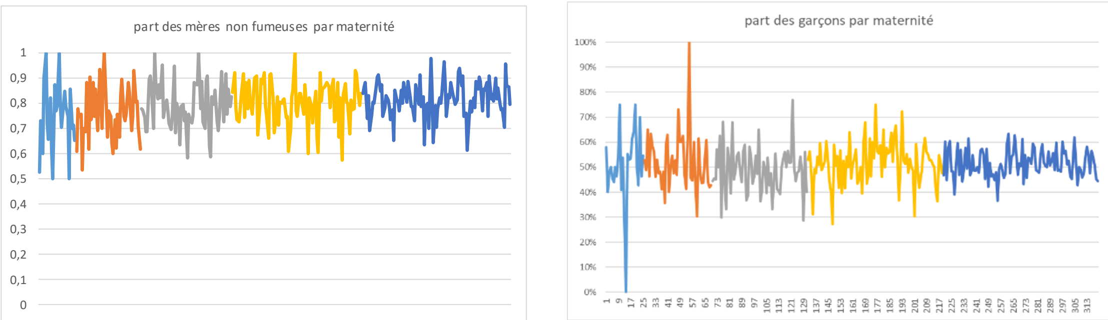

Figure 8 - Some indicators by stratum8

The conclusion is the same for the share of mothers without gestational diabetes, or of children who took a bath every day at the age of 2 years.

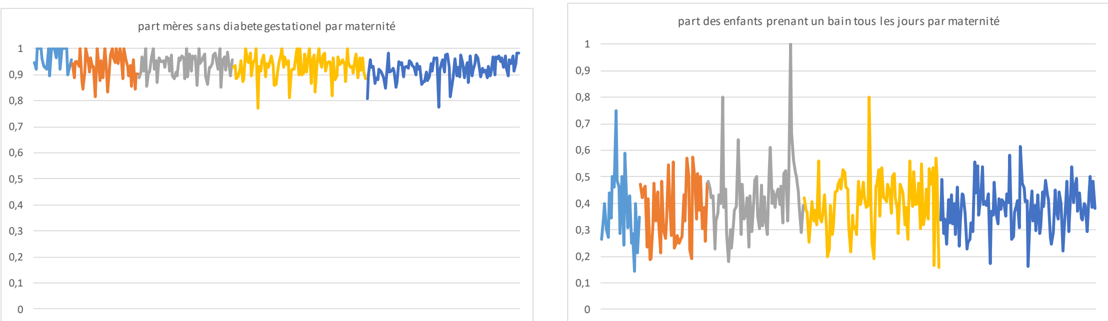

Figure 9 - Some indicators by stratum9

Conversely, the plots of the proportion of French mothers and of mothers in paid employment in each maternity unit show greater dispersion for those in strata 4 and 5. In these cases, then, stratification significantly improves the precision of the analysis.

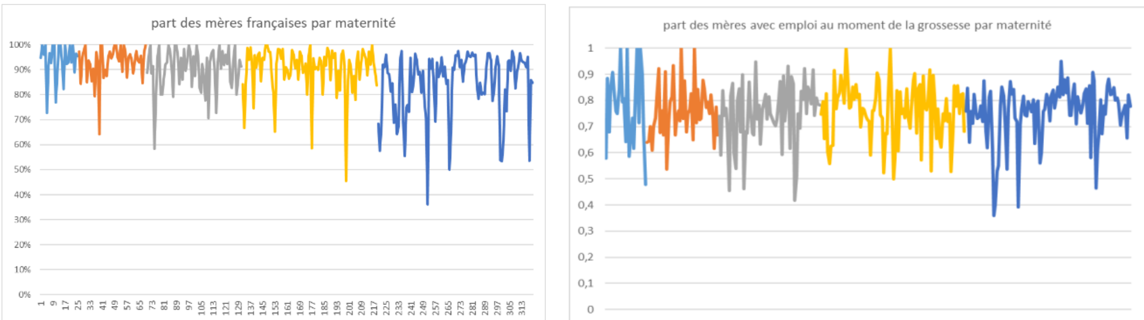

Figure 10 - Some indicators by stratum10

The **nonresponse effect** measures the decrease in the precision of the ELFE survey due to nonresponse. Note that by its very definition, the nonresponse effect amounts to adding a term to the variance that is proportional to $lin_k{}^2$ (and thus to $(y_k - \hat{R})^2$) and inversely proportional to the response probability $\emptyset^2{}_k$.

Thus, if the individuals with the lowest response probabilities are atypical (far from the overall ratio $\hat{R}$), nonresponse adds a significant amount of variance. If, for example, the estimated ratio is less than ½ (there are fewer children with the analysed characteristic than without it), individuals with $y_k = 1$ (who have the characteristic) are atypical, and are the “furthest” from the ratio. If these individuals are especially likely not to respond, the proportion of variance due to NR will be large.

This effect is obviously a little larger for the measurements at the age of 2 years (12,000 respondents vs. 18,000 in maternity units).

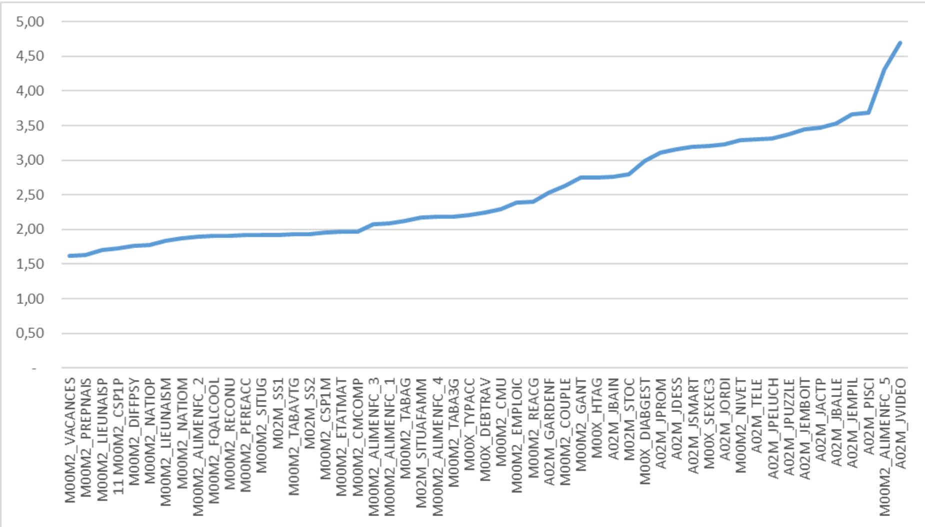

Figure 11 - Nonresponse effect11

The **design effect** measures how estimated variance with the ELFE survey’s sampling design differs from a variance estimated by simple random sampling.

It combines the three previous effects, which individually tend to fluctuate. As expected, comparing the nonresponse and cluster effects shows especially marked and contrasting variation with the number of children surveyed, with the two varying in opposite ways. The fewer people are surveyed, the greater the loss of precision due to nonresponse (as the probability of response decreases and the nonresponse effect grows, precision deteriorates), but the lower the cluster effect (where the average number of children surveyed per maternity unit is lower, the cluster effect is lower, and less precision is lost through clustering).

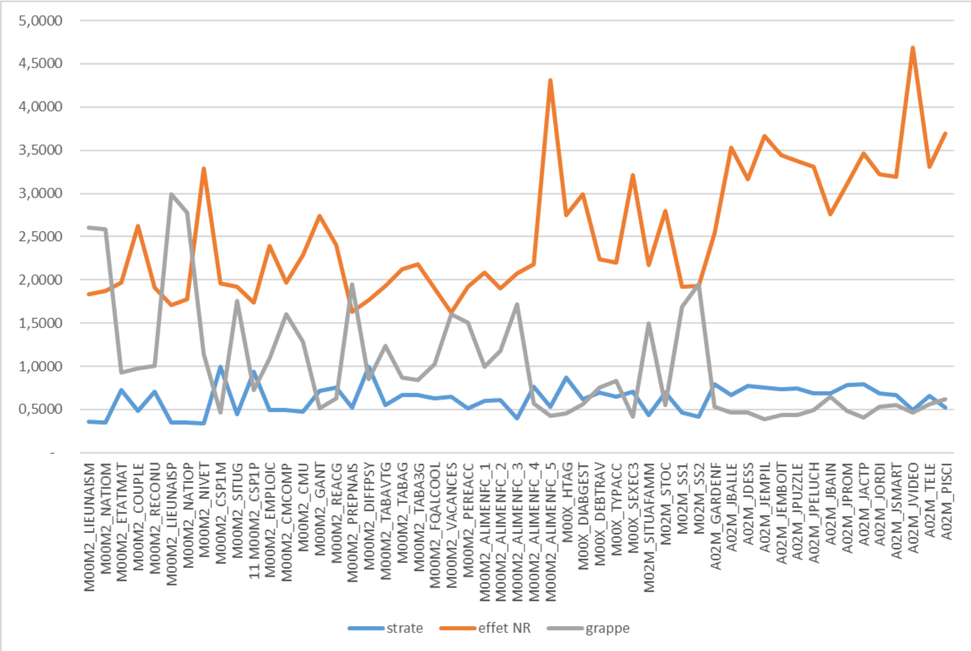

Figure 12 - Analysis of effects included in the calculation of variance with the ELFE sampling design 12

Overall, the design effect is relatively stable in a range between 1 and 1.4. Logically, this effect is slightly larger when nonrespondents are atypical in some way with respect to a given variable (parents' nationality, place of birth, etc.). In this case the NR effect is greater and is not compensated by the other effects (we saw above that the simple impact of a high nonresponse probability led to a large NR effect, which was partially offset by a small cluster effect).

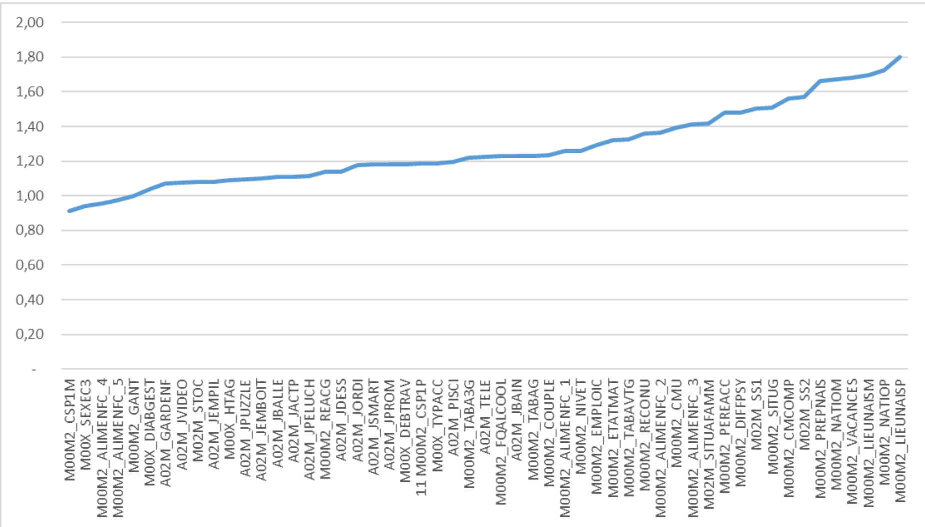

Figure 13 - Sampling design effect13

Finally, the **calibration effect** measures the extent to which calibration improves the precision of the estimators. Generally speaking, calibration always improves precision, as it limits certain random effects by controlling the distribution of individuals across categories for a set of variables, and thus reduces uncertainty. . If the studied variable depends (even a little) on the calibration variables, the random effect due to the random selection of individuals necessarily decreases. In the extreme case, when measuring the proportion of infants with a characteristic that depends directly on the calibration variables, the data is obviously no longer subject to any random effects, as the proportion of infants with a given characteristic is "fixed" in advance. This is why the variances for variables such as place of birth and nationality drop to zero after calibration, for example.

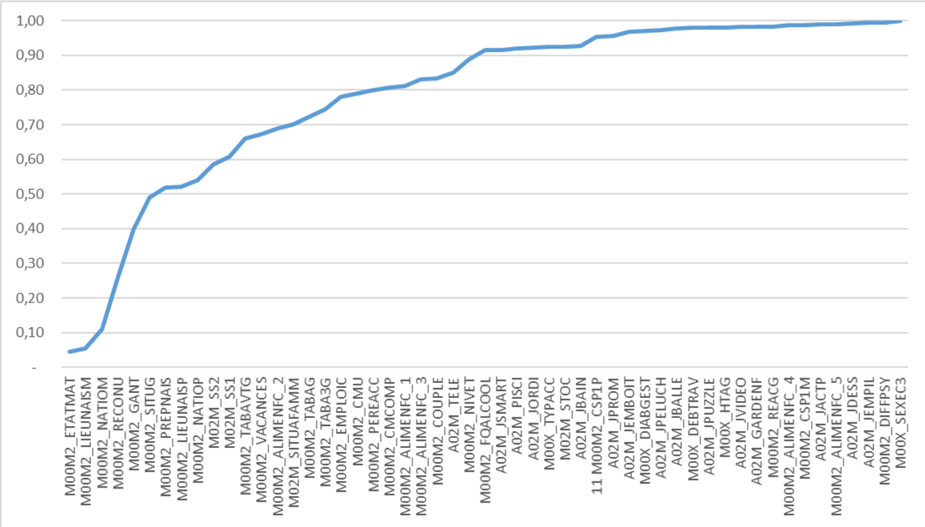

Figure 14 - Calibration effect14

It is interesting to compare the design and calibration effects. Recall that the calibration variables are strongly linked to the variables that explain nonresponse (this is true until the survey wave at the age of 2 years, and even more so from age 3½[4]). For example, during the survey in maternity units, the variables explaining nonresponse include notably mother's age, gestational age, region of residence, mother's SPC, mother's activity status at the time of pregnancy, twin indicator, and primiparity. The calibration variables are age, region, marital status, immigrant status, level of education, and primiparity, to which the twin indicator can be added, as there are two different calibrations, one for families (including only one child per family in the case of twins) and one for all children (including each twin). The two sets thus include common or extremely correlated variables (level of education explains SCP and activity status, etc.).

If a variable is not related to the calibration variables, there is no reason for respondents to differ from nonrespondents, or for that particular characteristic to be under- or overrepresented. The sampling design has little effect on the precision of the estimated variable, and the calibration effect is limited. If, on the other hand, a variable is more strongly correlated with the calibration variables, then the variable is at least partially correlated to the variables that explain nonresponse. In this case, then, the design effect will be relatively large. But by the mechanism of the calibration, which decreases the random effect on the proportion of nonresponse explained by the calibration variables, the calibration effect will be much larger.

[4] Cf. ELFE Survey: Weighting national survey data: https://www.elfe-france.fr/fichier/rte/178/Cot%C3%A9%20recherche/Weighting-Elfe-surveys-general-document.pdf.

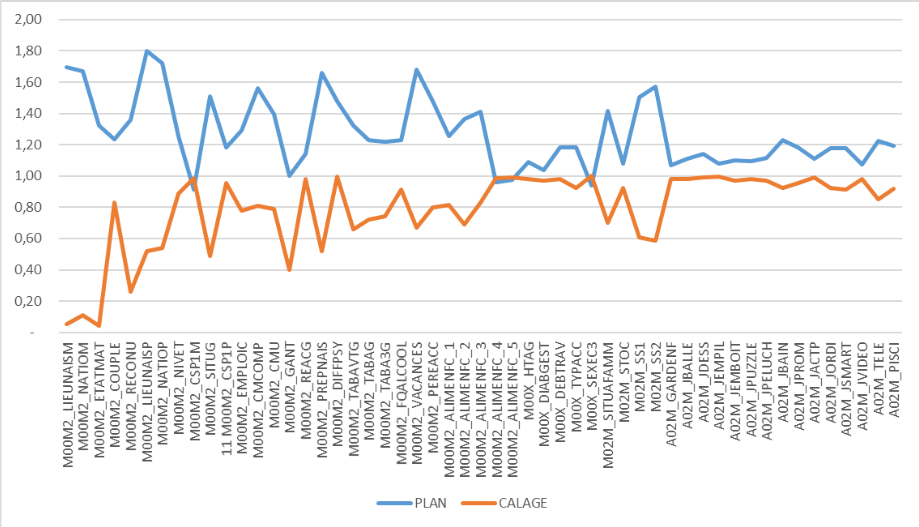

Figure 15 - Comparison between design effect and calibration effect

Finally, the **ELFE effect** compares the variance estimated according to the simplified sampling design compared to simple random sampling. Note that this effect is below 1.2 in almost all cases.

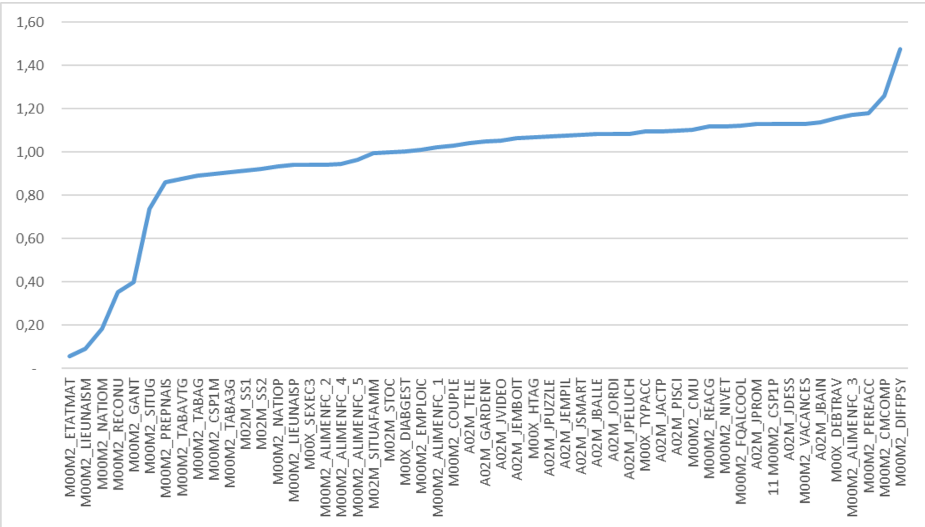

Figure 16 - Overall ELFE effect: ratio of simplified ELFE variance to simple random sampling variance15

When we measure this ratio with standard deviations rather than variances, and thus with the size of confidence intervals, we obtain ratios below 1.1.

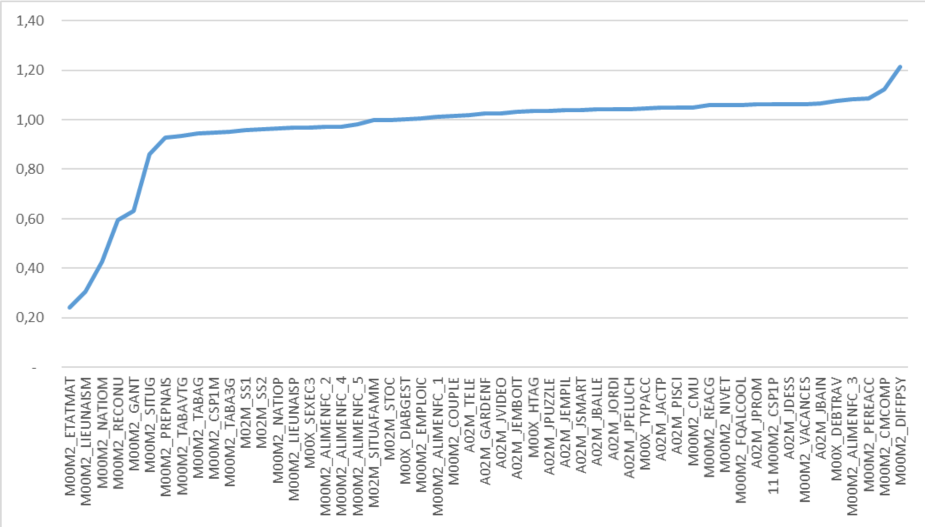

Figure 17 - Ratio between standard deviation with the ELFE simplified sampling design and with a simple random sample16

6. Some observations on the survey at the age of 3½ years

To conclude this analysis of the elements included in the calculation of variance with the ELFE sampling design, we performed these same calculations on some data collected when the participating children were 3½ years old. These variables concern the children's activities, the care they received, and their use of electronic devices.

To perform these calculations, we used the cross-sectional Child weighting produced using the new simultaneous calibration method. In its application to variance calculations, the main difference between the new and old weighting methods is that the new method involves more calibration variables (the 6 previously used variables + 7 more), and that these new variables consist precisely of the variables that explained nonresponse/nonparticipation. Strictly speaking, in the simultaneous method there is no calculation of the probability of an infant participating in the survey: this probability is estimated "a posteriori" by comparing the final weight after calibration with the weight drawn from the sampling design.[5]

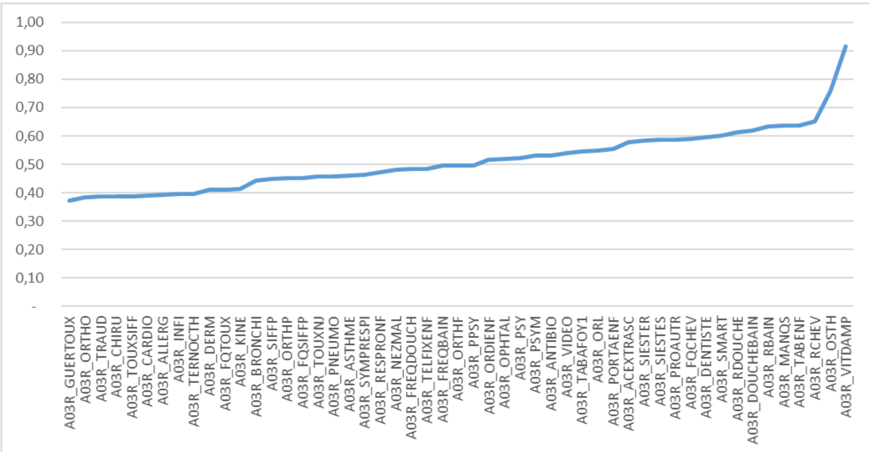

Figure 18 – Cluster effect - survey at age 3½17

As in the survey wave when the members of the ELFE cohort were 2 years old, the analyses for the wave at age 3½ yield cluster effects concentrated between 0.4 and 0.6. The number of respondents per maternity unit (with 11,700 respondents in total) is equivalent to the number in the previous wave at age 2, while at this later age children born in the same maternity unit have even less reason to resemble each other. There is now no cluster

---

[5] Cf. here again, ELFE Survey: Weighting national survey data: https://www.elfe-france.fr/fichier/rte/178/Cot%C3%A9%20recherche/Weighting-Elfe-surveys-general-document.pdf.

effect above 1. For all analysed variables, the variability between maternity units in the proportion of infants with a given characteristic is necessarily lower than the variability of the same characteristic measured directly among the infants.

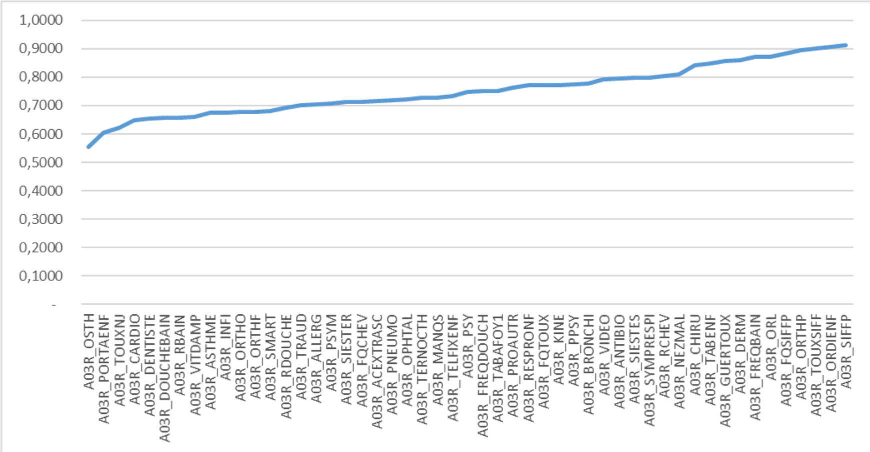

Figure 19 - Stratification effect - survey at age 3½18

The stratification effect is becoming stable. As time goes on, there is less and less reason for the size of the maternity unit to play a role in the analyses. We thus end up with an effect simply due to the high sampling rate in larger maternity units, and where as a result more participants were recruited.

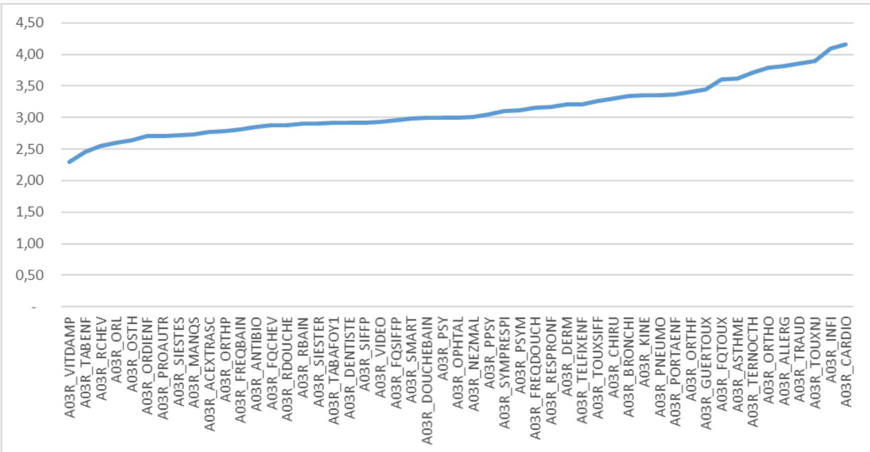

Figure 20 - Nonresponse effect – survey at age 3½19

As expected, there is a sizeable nonresponse effect. It is equivalent to the nonresponse effect obtained in the survey wave at age 2 (around 3, or even higher), which yields a design effect that is constant at around 1.2. This effect ultimately only results from effects that have been stabilizing over time, as the sampling design for the maternity units – which concentrates on the fact that the infants were selected in maternity units which themselves were divided into strata – has a decreasing overall impact on the precision of the results, with a cluster effect around 0.5, a stratum effect around 0.7, and an NR effect around 3.5, so the overall design effect = 0.5 x 0.7 x 3.5 = 1.2.

Here again, however, the design and calibration effects must be combined to obtain the overall effect of the sampling design implemented in the ELFE survey. Here, we deliberately chose variables that depended little or not at all on the calibration variables (we showed above that if the analysed variables depend on the calibration variables, by construction the calibration effect is very large and makes the precision of the analysis very high and the variance much lower than under simple random sampling). The calibration effect is thus relatively low (between 0.8 and 1).

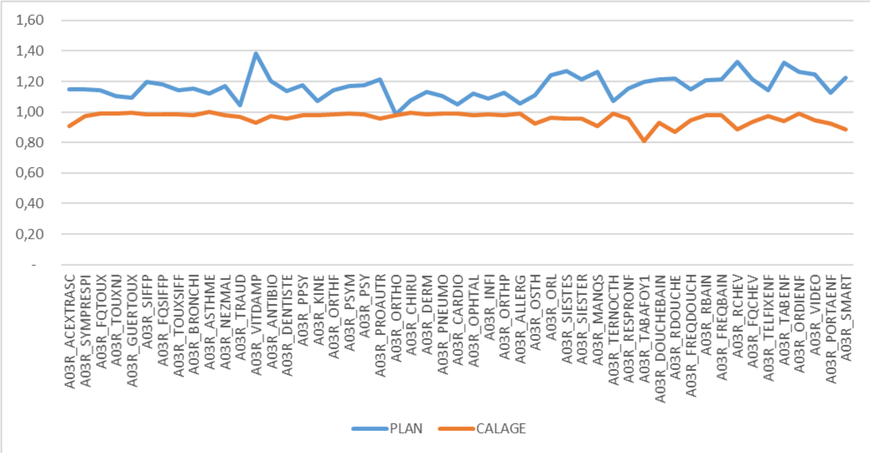

Figure 21 - Comparison between design effect and calibration effect – survey at age 3½20

Finally, the **ELFE effect** comparing the estimated variance under the ELFE simplified sampling design and a simple random sampling again falls between 1 and 1.2. When this ratio is measured on standard deviations rather than variances, and thus on the size of confidence intervals, the ratio is below 1.1 in every case.

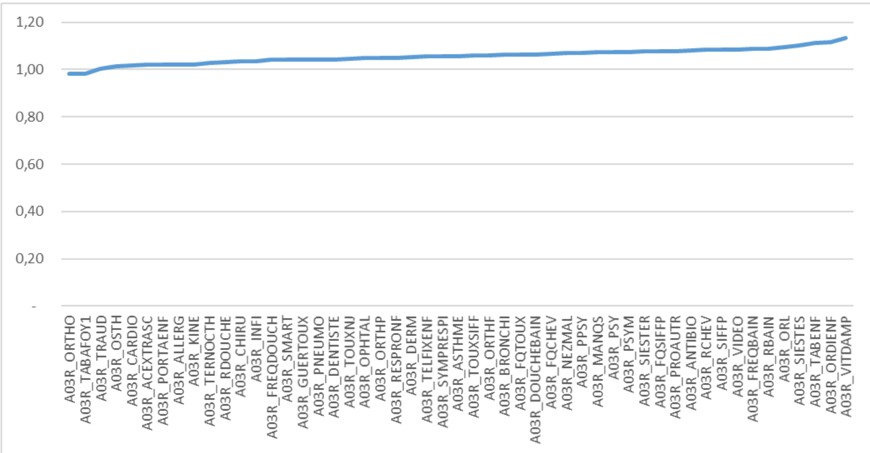

Figure 22 - Ratio of standard deviation (SD) under the ELFE simplified sampling design and SD under simple random sampling - survey at age 3½21

7. Recommendations for users of the ELFE survey

In this document, we have shown that the complete sampling design implemented in the ELFE survey can be simplified by disregarding an element of the design (the initial survey day in maternity units) without compromising the precision of the analyses, thanks to its negligible contribution to the variability of the data in comparison to the other elements of the sampling design.

We then showed that, under the proposed simplification hypothesis, the precision of the simplified sampling design compared to simple random sampling depends mainly on the sample size.

A significant decrease in the number of respondents (and thus in children's probability of participation) automatically decreases precision, due to the increase in the nonresponse effect. But this effect has been compensated both by the change in the cluster effect with the decrease in the number of respondents per maternity unit, and by the very principle of the calibration, in which the variables were chosen to ensure that the loss of precision due to nonresponse would be partly offset by the calibration effect.

We also showed that these effects have been stabilizing over time. The impact of the method of selection in the maternity units on data measured when the children were first 2, and then 3½ years old, decreased considerably.

In the end, we verified with around a hundred variables that the precision obtained with the simplified sampling design is comparable to the precision with simple random sampling (standard deviations underestimated by 10% or – often much – lower)**.**

ELFE survey users may thus use the classical SAS procedures, potentially being slightly conservative when choosing test thresholds (for example, choosing a significance level of 3% rather than the usual 5%), namely:

To estimate **mean**, **proportion**, **frequency**

- PROC SURVEYMEANS – continuous variables
- PROC SURVEYFREQ – discrete variables

For **linear regression**

- PROC SURVEYREG – linear regression, equality test.

For **logistic regression**

- PROC SURVEYLOGISTIC – logistic regression

Finally, be careful to in fact use the SAS Survey… procedures, and not *proc means*, *freq*, *reg*, for example, which very strongly underestimate variances by estimating the elements necessary for the tests as if the table used to estimate the parameters contained the whole population, and not a survey sample. The estimated means, totals, and ratios will be identical, but the variances, and thus the significance of the results, may be very different.

Appendix 1: List of analysed variables

nom	
M00M2_LIEUNAISM	Lieu de naissance mère
M00M2_NATIOM	Nationalité mère
M00M2_ETATMAT	Etat matrimonial mère
M00M2_COUPLE	La mère vit en couple
M00M2_RECONU	Le père a reconnu l'enfant
M00M2_LIEUNAISP	Lieu de naissance père
M00M2_NATIOP	Nationalité père
M00M2_NIVET	Niveau d'études mère
M00M2_CSP1M	Recodage : profession et catégorie sociale de la mère
M00M2_SITUG	Emploi mère au moment de la grossesse
M00M2_CSP1P	Recodage : profession et catégorie sociale du père
M00M2_EMPLOIC	Situation professionnelle père
M00M2_CMCOMP	Couverture maladie complémentaire
M00M2_CMU	CMU complémentaire
M00M2_GANT	Grossesse(s) antérieure(s)
M00M2_REACG	Réaction à la découverte de la grossesse
M00M2_PREPNAIS	Séances de préparation à la naissance
M00M2_DIFFPSY	Difficultés psy pendant la grossesse
M00M2_TABAVTG	Tabagisme avant la grossesse
M00M2_TABAG	Tabagisme pendant la grossesse
M00M2_TABA3G	Tabagisme pendant le 3e trimestre
M00M2_FQALCOOL	Consommation d'alcool
M00M2_VACANCES	Vacances pendant la grossesse
M00M2_PEREACC	Le père a assisté à l'accouchement
M00M2_ALIMENFC_1	Alimentation de l'enfant : lait maternel uniquement
M00M2_ALIMENFC_2	Alimentation de l'enfant : lait 1er âge uniquement
M00M2_ALIMENFC_3	Alimentation de l'enfant : allaitement mixte
M00M2_ALIMENFC_4	Alimentation de l'enfant : NSP
M00M2_ALIMENFC_5	Alimentation de l'enfant : autre
M00X_HTAG	Hypertension artérielle pendant la grossesse
M00X_DIABGEST	Diabète gestationnel
M00X_DEBTRAV	Début du travail
M00X_TYPACC	Accouchement
M00X_SEXEC3	Sexe
M02M_SITUAFAMM	Situation familiale de la mère
M02M_STOC	Situation du ménage par rapport au logement
M02M_SS1	Régime de sécurité sociale
M02M_SS2	Couverture maladie complémentaire
A02M_GARDENF	Mode de garde principal semaine
A02M_JBALLE	Balle
A02M_JDESS	Dessin ou peinture
A02M_JEMPIL	Empiler
A02M_JEMBOIT	Emboîter
A02M_JPUZZLE	Puzzles
A02M_JPELUCH	Peluches
A02M_JBAIN	Jeux de bains ou jeux d'eau
A02M_JPROM	Promenades
A02M_JACTP	Jeux / activités physiques
A02M_JORDI	Jm_ordi
A02M_JSMART	Smartphone
A02M_JVIDEO	Jeux vidéo
A02M_TELE	Télévision
A02M_PISCI	Piscine

Appendix 2: Proportion of effects of survey day, maternity unit, and nonresponse in the theoretical variance estimate before calibration and analysis of mean dispersions

nom	part effet DAY	part effet MAT	part effet NR
M00M2_LIEUNAISM	51%	27%	22%
M00M2_NATIOM	50%	27%	23%
M00M2_ETATMAT	43%	29%	28%
M00M2_COUPLE	31%	26%	43%
M00M2_RECONU	34%	35%	32%
M00M2_LIEUNAISP	34%	39%	27%
M00M2_NATIOP	37%	35%	27%
M00M2_NIVET	56%	13%	31%
M00M2_CSP1M	31%	35%	34%
M00M2_SITUG	30%	36%	33%
11 M00M2_CSP1P	53%	27%	20%
M00M2_EMPLOIC	31%	29%	40%
M00M2_CMCOMP	40%	31%	30%
M00M2_CMU	49%	22%	29%
M00M2_GANT	58%	15%	27%
M00M2_REACG	39%	25%	35%
M00M2_PREPNAIS	42%	36%	22%
M00M2_DIFFPSY	29%	40%	31%
M00M2_TABAVTG	50%	26%	24%
M00M2_TABAG	49%	24%	27%
M00M2_TABA3G	46%	25%	29%
M00M2_FQALCOOL	65%	18%	17%
M00M2_VACANCES	39%	38%	23%
M00M2_PEREACC	64%	19%	17%
M00M2_ALIMENFC_1	38%	30%	32%
M00M2_ALIMENFC_2	46%	28%	25%
M00M2_ALIMENFC_3	36%	31%	33%
M00M2_ALIMENFC_4	61%	18%	21%
M00M2_ALIMENFC_5	47%	12%	41%
M00X_HTAG	50%	18%	32%
M00X_DIABGEST	61%	13%	26%
M00X_DEBTRAV	88%	5%	7%
M00X_TYPACC	78%	10%	12%
M00X_SEXEC3	28%	23%	50%
M02M_SITUAFAMM	37%	29%	34%
M02M_STOC	55%	16%	29%
M02M_SS1	46%	28%	26%
M02M_SS2	36%	33%	31%
A02M_GARDENF	62%	15%	23%
A02M_JBALLE	33%	19%	48%
A02M_JDESS	35%	20%	44%
A02M_JEMPIL	34%	18%	48%
A02M_JEMBOIT	39%	18%	43%
A02M_JPUZZLE	35%	19%	46%
A02M_JPELUCH	47%	16%	37%
A02M_JBAIN	38%	23%	40%
A02M_JPROM	54%	15%	31%
A02M_JACTP	41%	17%	42%
A02M_JORDI	37%	19%	43%
A02M_JSMART	37%	20%	43%
A02M_JVIDEO	64%	8%	28%
A02M_TELE	32%	21%	48%
A02M_PISCI	29%	19%	52%

nom	rapport effet DAY / effet MAT	rapport dispersion MAT / dispersion DAY
M00M2_LIEUNAISM	1,92	49,54
M00M2_NATIOM	1,89	53,96
M00M2_ETATMAT	1,48	72,37
M00M2_COUPLE	1,16	30,75
M00M2_RECONU	0,97	58,91
M00M2_LIEUNAISP	0,88	115,04
M00M2_NATIOP	1,06	89,39
M00M2_NIVET	4,17	27,00
M00M2_CSP1M	0,89	92,83
M00M2_SITUG	0,84	67,29
11 M00M2_CSP1P	1,99	55,37
M00M2_EMPLOIC	1,09	28,15
M00M2_CMCOMP	1,29	23,77
M00M2_CMU	2,18	17,24
M00M2_GANT	3,76	15,94
M00M2_REACG	1,56	25,33
M00M2_PREPNAIS	1,18	26,55
M00M2_DIFFPSY	0,72	21,09
M00M2_TABAVTG	1,94	55,00
M00M2_TABAG	2,05	32,87
M00M2_TABA3G	1,87	31,51
M00M2_FQALCOOL	3,57	24,59
M00M2_VACANCES	1,03	2,80
M00M2_PEREACC	3,43	16,38
M00M2_ALIMENFC_1	1,28	51,91
M00M2_ALIMENFC_2	1,64	46,51
M00M2_ALIMENFC_3	1,17	45,53
M00M2_ALIMENFC_4	3,36	18,18
M00M2_ALIMENFC_5	3,86	41,37
M00X_HTAG	2,77	14,02
M00X_DIABGEST	4,78	11,41
M00X_DEBTRAV	16,16	3,70
M00X_TYPACC	7,86	8,31
M00X_SEXEC3	1,22	31,80
M02M_SITUAFAMM	1,30	40,18
M02M_STOC	3,48	10,62
M02M_SS1	1,67	26,78
M02M_SS2	1,11	51,45
A02M_GARDENF	4,05	29,35
A02M_JBALLE	1,73	4,27
A02M_JDESS	1,72	5,94
A02M_JEMPIL	1,89	10,07
A02M_JEMBOIT	2,22	6,79
A02M_JPUZZLE	1,80	14,24
A02M_JPELUCH	2,96	16,60
A02M_JBAIN	1,68	27,48
A02M_JPROM	3,68	8,14
A02M_JACTP	2,41	5,97
A02M_JORDI	1,93	26,26
A02M_JSMART	1,91	32,28
A02M_JVIDEO	8,31	16,53
A02M_TELE	1,54	27,11
A02M_PISCI	1,51	39,42

Appendix 3: Quantifying the components of the estimated variance

nom	ratio	$\hat{V}_{SAS}(lin_k)$	$\hat{V}_{GR}(lin_k)$	$\hat{V}_{GR\_ST}(lin_k)$	$\hat{V}_{NR}(lin_k)$	$\hat{V}_{GR\_ST}(lin_k)+\hat{V}_{NR}(lin_k)$	$\hat{V}_{GR\_ST}(\varepsilon_k)+\hat{V}_{NR}(\varepsilon_k)$
M00M2_LIEUNAISM	80,2%	1,830E-05	4,768E-05	1,690E-05	1,410E-05	3,099E-05	1,693E-06
M00M2_NATIOM	81,8%	1,744E-05	4,504E-05	1,556E-05	1,355E-05	2,911E-05	3,189E-06
M00M2_ETATMAT	43,9%	2,115E-05	1,960E-05	1,423E-05	1,373E-05	2,796E-05	1,236E-06
M00M2_COUPLE	91,3%	9,623E-06	9,391E-06	4,524E-06	7,353E-06	1,188E-05	9,889E-06
M00M2_RECONU	48,0%	2,144E-05	2,159E-05	1,528E-05	1,385E-05	2,913E-05	7,557E-06
M00M2_LIEUNAISP	77,5%	1,880E-05	5,628E-05	1,980E-05	1,405E-05	3,385E-05	1,767E-05
M00M2_NATIOP	79,1%	1,802E-05	5,004E-05	1,750E-05	1,353E-05	3,103E-05	1,679E-05
M00M2_NIVET	1,3%	2,216E-06	2,518E-06	8,480E-07	1,939E-06	2,787E-06	2,479E-06
M00M2_CSP1M	0,3%	2,083E-07	9,770E-08	9,696E-08	9,280E-08	1,898E-07	1,874E-07
M00M2_SITUG	64,0%	2,443E-05	4,291E-05	1,914E-05	1,767E-05	3,681E-05	1,805E-05
11 M00M2_CSP1P	1,5%	1,136E-06	8,244E-07	7,759E-07	5,694E-07	1,345E-06	1,281E-06
M00M2_EMPLOIC	84,3%	1,551E-05	1,692E-05	8,385E-06	1,165E-05	2,004E-05	1,565E-05
M00M2_CMCOMP	84,8%	1,542E-05	2,477E-05	1,223E-05	1,185E-05	2,408E-05	1,943E-05
M00M2_CMU	7,5%	9,034E-06	1,161E-05	5,503E-06	7,089E-06	1,259E-05	9,947E-06
M00M2_GANT	67,8%	1,814E-05	9,225E-06	6,616E-06	1,154E-05	1,815E-05	7,234E-06
M00M2_REACG	73,9%	1,731E-05	1,087E-05	8,208E-06	1,151E-05	1,972E-05	1,936E-05
M00M2_PREPNAIS	48,2%	2,118E-05	4,120E-05	2,158E-05	1,360E-05	3,518E-05	1,825E-05
M00M2_DIFFPSY	12,9%	9,967E-06	8,454E-06	8,374E-06	6,381E-06	1,476E-05	1,469E-05
M00M2_TABAVTG	41,4%	2,065E-05	2,558E-05	1,418E-05	1,320E-05	2,737E-05	1,809E-05
M00M2_TABAG	21,3%	1,524E-05	1,329E-05	8,822E-06	9,945E-06	1,877E-05	1,358E-05
M00M2_TABA3G	17,3%	1,326E-05	1,110E-05	7,416E-06	8,752E-06	1,617E-05	1,202E-05
M00M2_FQALCOOL	15,0%	9,686E-06	9,892E-06	6,239E-06	5,645E-06	1,188E-05	1,086E-05
M00M2_VACANCES	47,8%	2,121E-05	3,409E-05	2,200E-05	1,366E-05	3,566E-05	2,396E-05
M00M2_PEREACC	76,6%	1,839E-05	2,772E-05	1,413E-05	1,302E-05	2,715E-05	2,170E-05
M00M2_ALIMENFC_1	57,5%	2,148E-05	2,146E-05	1,292E-05	1,405E-05	2,698E-05	2,192E-05
M00M2_ALIMENFC_2	32,0%	1,906E-05	2,243E-05	1,369E-05	1,232E-05	2,601E-05	1,795E-05
M00M2_ALIMENFC_3	9,2%	8,367E-06	1,440E-05	5,698E-06	6,122E-06	1,182E-05	9,812E-06
M00M2_ALIMENFC_4	0,3%	2,138E-07	1,227E-07	9,395E-08	1,108E-07	2,047E-07	2,019E-07
M00M2_ALIMENFC_5	0,4%	4,237E-07	1,792E-07	9,590E-08	3,175E-07	4,134E-07	4,092E-07
M00X_HTAG	1,7%	1,813E-06	8,272E-07	7,196E-07	1,259E-06	1,978E-06	1,938E-06
M00X_DIABGEST	7,3%	6,198E-06	3,485E-06	2,143E-06	4,273E-06	6,416E-06	6,217E-06
M00X_DEBTRAV	69,0%	1,868E-05	1,416E-05	9,857E-06	1,221E-05	2,207E-05	2,160E-05
M00X_TYPACC	67,2%	1,873E-05	1,550E-05	1,008E-05	1,212E-05	2,221E-05	2,052E-05
M00X_SEXEC3	51,0%	2,151E-05	8,943E-06	6,303E-06	1,394E-05	2,024E-05	2,022E-05
M02M_SITUAFAMM	76,9%	2,125E-05	3,179E-05	1,390E-05	1,621E-05	3,011E-05	2,114E-05
M02M_STOC	7,3%	6,598E-06	3,660E-06	2,547E-06	4,576E-06	7,123E-06	6,586E-06
M02M_SS1	66,5%	2,371E-05	4,009E-05	1,854E-05	1,715E-05	3,569E-05	2,170E-05
M02M_SS2	72,6%	2,345E-05	4,564E-05	1,903E-05	1,779E-05	3,682E-05	2,160E-05
A02M_GARDENF	18,9%	1,333E-05	7,131E-06	5,629E-06	8,629E-06	1,426E-05	1,399E-05
A02M_JBALLE	27,8%	2,891E-05	1,357E-05	9,085E-06	2,296E-05	3,204E-05	3,135E-05
A02M_JDESS	23,0%	2,355E-05	1,095E-05	8,488E-06	1,835E-05	2,683E-05	2,659E-05
A02M_JEMPIL	18,7%	2,068E-05	8,077E-06	6,099E-06	1,626E-05	2,236E-05	2,225E-05
A02M_JEMBOIT	21,6%	2,218E-05	9,617E-06	7,072E-06	1,728E-05	2,435E-05	2,358E-05
A02M_JPUZZLE	7,5%	8,691E-06	3,775E-06	2,813E-06	6,682E-06	9,495E-06	9,300E-06
A02M_JPELUCH	36,0%	3,058E-05	1,503E-05	1,030E-05	2,381E-05	3,411E-05	3,318E-05
A02M_JBAIN	52,6%	3,365E-05	2,187E-05	1,500E-05	2,636E-05	4,136E-05	3,829E-05
A02M_JPROM	27,3%	2,875E-05	1,390E-05	1,094E-05	2,302E-05	3,396E-05	3,242E-05
A02M_JACTP	21,9%	2,419E-05	9,809E-06	7,744E-06	1,908E-05	2,683E-05	2,653E-05
A02M_JORDI	11,4%	1,505E-05	7,964E-06	5,487E-06	1,221E-05	1,770E-05	1,632E-05
A02M_JSMART	9,6%	1,284E-05	7,101E-06	4,741E-06	1,040E-05	1,514E-05	1,383E-05
A02M_JVIDEO	0,5%	9,126E-07	4,229E-07	2,090E-07	7,714E-07	9,804E-07	9,622E-07
A02M_TELE	17,1%	2,655E-05	1,491E-05	9,826E-06	2,266E-05	3,249E-05	2,762E-05
A02M_PISCI	9,4%	1,641E-05	1,021E-05	5,310E-06	1,429E-05	1,960E-05	1,804E-05

Appendix 4: List of variables analysed in the survey wave at the age of 3 ½ years

nom	
A03R_ACEXTRASC	Cette année, [enfant elfe] pratique-t-il/elle régulièrement une activité de loisir dans un club ou association, comme par exemple du judo, du dessin ou de la musique (en dehors de l'école et du centre de loisir) ?
A03R_SYMPRESPI	[enfant elfe] a-t-il/elle déjà eu, dans les 12 derniers mois une toux, une gêne respiratoire ou un épisode de sifflements ?
A03R_FQTOUX	Ces épisodes de toux surviennent-ils ?
A03R_TOUXNJ	[enfant elfe] tousse-t-il/elle ?
A03R_GUERTOUX	Entre les épisodes de toux, [enfant elfe] est-il/elle complètement guéri(e) ?
A03R_SIFFP	[enfant elfe] a-t-il/elle déjà eu, au cours des 12 derniers mois, au moins un épisode de sifflements dans la poitrine ?
A03R_FQSIFFP	Ces épisodes de sifflements surviennent-ils ?
A03R_TOUXSIFF	Ces sifflements accompagnent-ils toujours les épisodes de toux ?
A03R_BRONCHI	[enfant elfe] a-t-il/elle fait une bronchiolite depuis l'âge de 2 ans ?
A03R_ASTHME	[enfant elfe] a-t-il/elle eu des crises d'asthme au cours des 12 derniers mois ?
A03R_NEZMAL	Considérez-vous que [enfant elfe] a souvent le nez bouché ou le nez qui coule ?
A03R_TRAUD	[enfant elfe] est-il/elle suivi pour un trouble de l'audition ?
A03R_VITDAMP	Au cours des 12 derniers mois, [enfant elfe] a-t-il/elle pris de la vitamine d sous forme d'ampoule (zymad, vitamine d3 bon, uvedose) ou en doses quotidiennes (zymad, zymaduo, uvesterol…) ?
A03R_ANTIBIO	Au cours des 12 derniers mois [enfant elfe] a-t-il/elle reçu un traitement antibiotique (clamoxyl, hiconcil, agram, amoxicilline, augmentin, ciblor, orelox, penicilline g, oroken, bristopen, bactrim, rocephine, josacine, zythromax, pediazole, pyostacine)
A03R_DENTISTE	Dentiste
A03R_PPSY	Pédopsychiatre
A03R_KINE	Kinésithérapeute
A03R_ORTHF	Orthophoniste
A03R_PSYM	Psychomotricien
A03R_PSY	Psychologue
A03R_PROAUTR	Un ou d'autres professionnels de santé spécialisés
A03R_ORTHO	Orthopédiste
A03R_CHIRU	Un chirurgien autre qu'orthopédiste
A03R_DERM	Dermatologue
A03R_PNEUMO	Pneumologue
A03R_CARDIO	Cardiologue
A03R_OPHTAL	Ophtalmologiste
A03R_INFI	Infirmière
A03R_ORTHP	Orthoptiste
A03R_ALLERG	Allergologue
A03R_OSTH	Ostéopathe
A03R_ORL	Orl
A03R_SIESTES	[enfant elfe] fait-il/elle la sieste en semaine ?
A03R_SIESTER	Fait-il/elle la sieste le week-end, en vacances ?
A03R_MANQS	Selon vous [enfant elfe] manque-t-il/elle de sommeil ?
A03R_TERNOCTH	Lui arrive-t-il de se réveiller la nuit en criant, en étant confus(e), impossible à approcher, sans s'en souvenir le matin ?
A03R_RESPRONF	Lorsque [enfant elfe] n'est pas enrhumé(e), à quelle fréquence ronfle-t-il/elle ?
A03R_TABAFOY1	Oui, un fumeur
A03R_DOUCHEBAIN	[enfant elfe] prend-il/elle :
A03R_RDOUCHE	Indiquer un nombre de fois
A03R_FREQDOUCH	A quelle fréquence [enfant elfe] prend-t-il/elle une douche sans compter les bains ?
A03R_RBAIN	Indiquer un nombre de fois
A03R_FREQBAIN	Freqbain
A03R_RCHEV	Indiquer un nombre de fois
A03R_FQCHEV	En général à quel rythme [enfant elfe] a-t-il/elle les cheveux lavés ?
A03R_TELFIXENF	Arrive-t-il à [enfant elfe] de parle-t-il/elle au téléphone sans fil au moins une fois par semaine ?
A03R_TABENF	[enfant elfe] utilise-t-il/elle une tablette au domicile au moins une fois par semaine ?
A03R_ORDIENF	[enfant elfe] utilise-t-il/elle un ordinateur au domicile au moins une fois par semaine ?
A03R_VIDEO	[enfant elfe] joue-t-il/elle actuellement à des jeux vidéo sur une console (wii, psp, xbox, ds, …) au moins une fois par semaine ?
A03R_PORTAENF	Arrive-t-il à [enfant elfe] de parle-t-il/elle au téléphone portable au moins une fois par semaine ?
A03R_SMART	[enfant elfe] joue-t-il/elle actuellement sur un téléphone portable au moins une fois par semaine ?

Appendix 5: SAS code used to generate the components of the variance estimate

A table containing at least the following fields:

```
Table = name of the table containing the data.

-       The calibration variables CS_1 to CS_13
-       identifiant = field identifying the records ID...
-       strate = field identifying the stratum of the maternity unit. By default
M00M1_MATSTRATEC1
-       vague = field identifying the seasonal period in maternity units. By default
M00M1_VAGUE
-       mater = field identifying the maternity unit. By default M00M1_IDGROUPNAMEALEAC1
-       jour = field identifying the date of birth. By default M00M2_JNAISSEALEA

-       poids = name of the weighting variable. For example M00E_PONDVALC2

-       variable = name of the variable whose TOTAL is to be calculated
        OR
-       variable1 = name of the variable used to NUMERATOR is to be calculated
-       variable2 = name of the variable used to DENOMINATOR is to be calculated

-       méthode = 1 if survey wave BEFORE the age of 3½ years, méthode = 2 if survey wave AT OR
AFTER the age of 3½ years

data degreM;
input strateN _TOTAL_;
datalines;
1 108
2 108
3 109
4 108
5 111
;
data degreJ;
input vagueN _TOTAL_;
datalines;
1 90
2 91
3 92
4 92
;

%if &methode=1 %then %let listeCALAGE= CS_1 CS_2 CS_3 CS_4 CS_5 CS_6 ;
%if &methode=2 %then %let listeCALAGE= CS_1 CS_2 CS_3 CS_4 CS_5 CS_6 CS_7 CS_8 CS_9 CS_10
CS_11 CS_12 CS_13;
```

For a TOTAL:

```
/* régression pondérée par poids de la variable d'intérêt sur les variables de calage*/
proc glm data=table noprint ;
        class &listeCALAGE ;
        model &variable = &listeCALAGE ;
        weight &poids;
        output out=residus RESIDUAL = res;
run;
```

For a RATIO:

```
/* estimation du total de la variable1 pour le numérateur, total variable2 pour le
dénominateur et du ratio */
proc sql;
create table ESTIMATION_TOTAL as
select N as estim_NUM, D as estim_DEN, N/D as estimateur from
(select sum (&variable1 * &poids) as N, sum (&variable2 * &poids) as D
from  table) as t;quit;

data _null_;
set ESTIMATION_TOTAL;
CALL SYMPUT('numerateur',estim_NUM);
CALL SYMPUT('denominateur',estim_DEN);
```

```
CALL SYMPUT('ratio',estimateur);run;

/*création de lin*/
data table;
set table;
Ratio_i = (1/&denominateur)*(&variable1 - &ratio*&variable2);run;

/* régression pondérée par poids de la linéarisée sur les variables de calage*/
proc glm data=table noprint ;
       class &listeCALAGE ;
       model Ratio_i = &listeCALAGE ;
       weight &poids;
       output out=residus RESIDUAL = res;
run;
```

And then calculate:

```
data residus (keep= &identifiant res); set residus; run;
proc sort data= table; by &identifiant; run;
proc sort data= residus; by &identifiant; run;

data tableRES; merge table residus; by &identifiant; run;

/* effet MAT*/
proc surveymeans data=tableRES total=degreM mean clm stderr var sum clsum std varsum;
weight &poids;
cluster &mater;
strata & strateN;
var res;
ods output Statistics=StatRATMAT;run;
/* effet JOUR*/
proc surveymeans data=tableRES total=degreJ mean clm stderr var sum clsum std varsum;
weight &poids;
cluster &jour;
strata &vagueN;
var res;
ods output Statistics=StatRATJOUR;run;
/* effet NR*/
proc sql;
create table NR_calage as
select sum (res*res*(1-probaR)/(probaR*probaR*(1/pondAVANT_calage)*(1/pondAVANT_calage))) as
effetNR from tableRES;run;quit;

data _null_;
set  StatRATMAT;
CALL SYMPUT('var_calage_effetMAT',varsum);run;
data _null_;
set  StatRATJOUR;
CALL SYMPUT('var_calage_effetJOUR',varsum);run;
data _null_;
set  NR_calage;
CALL SYMPUT('NR_calage',effetNR);run;
```

To calculate the exact variance, we sum var_calage_effetMAT + var_calage_effetJOUR + NR_calage

To calculate the variance under the simplified sampling design, we sum var_calage_effetMAT + NR_calage

To estimate variance under simple random sampling

```
* SAS TOTAL;
proc surveymeans data=table total=764000;
var &variable;
weight & poids;
ods output Statistics=SAS;run;

data _null_; set  SAS;
CALL SYMPUT('V_SAS',varsum);run;


* SAS RATIO;
proc surveymeans data= table total=764000;
ratio &variable1/&variable2;
weight & poids;
ods output Statistics=SAS;run;

data _null_; set  SAS;
```

```
CALL SYMPUT('V_SAS',var);run;
```

REFERENCES

Hélène Juillard (2016) : Méthodes d'estimation et d'estimation de variance pour une enquête longitudinale - Application aux données de l'Etude Longitudinale Française depuis l'Enfance (Elfe) – Université de Toulouse – Thèse

Hélène Juillard (2016) : Estimation de la variance pour l'enquête ELFE - Collection : Documents de travail INED n° 226

Fabien DELL, Xavier d'HAULTFOEUILLE, Philippe FÉVRIER, Emmanuel MASSÉ (**) (2002) – Mise en œuvre du calcul de variance par linéarisation - Insee-Méthodes : Actes des Journées de Méthodologie Statistique.

Jean-Claude Deville (1999): Variance estimation for complex statistics and estimators: Linearization and residual techniques.Survey Methodology

Jean-Claude Deville, Carl-Erik Sarndal (1992): Calibration Estimators in Survey Sampling - Journal of the American Statistical Association, Vol. 87, No. 418, pp.376-382

Skinner, C. J. (2015): Cross-classified sampling: some estimation theory. Statistics and Probability Letters. ISSN 0167-7152