Catalogue no. 12-001-X
ISSN 1492-0921

Survey Methodology

A few remarks on a small example by Jean-Claude Deville regarding non-ignorable non-response

by Yves Tillé

Release date: December 20, 2016

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

email at STATCAN.infostats-infostats.STATCAN@canada.ca

telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following toll-free numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-877-287-4369

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Standard table symbols

The following symbols are used in Statistics Canada publications:

. not available for any reference period
.. not available for a specific reference period
... not applicable
0 true zero or a value rounded to zero
0^s value rounded to 0 (zero) where there is a meaningful distinction between true zero and the value that was rounded
^{p} preliminary
^{r} revised
x suppressed to meet the confidentiality requirements of the *Statistics Act*
^{E} use with caution
F too unreliable to be published
* significantly different from reference category ($p < 0.05$)

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2016

All rights reserved. Use of this publication is governed by the Statistics Canada Open Licence Agreement.

An HTML version is also available.

Cette publication est aussi disponible en français.

Survey Methodology, December 2016
Vol. 42, No. 2, pp. 307-317
Statistics Canada, Catalogue No. 12-001-X

A few remarks on a small example by Jean-Claude Deville regarding non-ignorable non-response

Yves Tillé[1]

Abstract

An example presented by Jean-Claude Deville in 2005 is subjected to three estimation methods: the method of moments, the maximum likelihood method, and generalized calibration. The three methods yield exactly the same results for the two non-response models. A discussion follows on how to choose the most appropriate model.

Key Words: Calibration; Generalized calibration; Method of moments; Likelihood.

1 Deville's example

During a conference at the University of Neuchâtel, Jean-Claude Deville (2005) presented a simple example to illustrate the value of generalized calibration for dealing with non-ignorable non-response (regarding generalized calibration, see Deville 2000, 2002 and 2004; Kott 2006; Chang and Kott 2008; Kott and Chang 2010; and Lesage and Haziza 2015). The example is reproduced below in its entirety.

> *Adjustments to offset the effects of non-response require very accurate knowledge of the factors that cause it. In particular, if what is to be measured directly influences the response probability, we must take risks with the data. Here is a small fictional example: A group of students is interviewed about their use of drugs. The survey results are as follows:*

Table 1.1
Deville's example

	YES	NO	NON-RESPONSE	COMBINED
Boys	40	80	180	300
Girls	20	160	120	300
Combined	60	240	300	600

> *Naively, we would think that the percentage of drug users is estimated at 60/(240 + 60) = 25%. This estimate is made under the assumption that non-respondents have the same behaviour as respondents. However, we notice that the response rate for girls is greater than the response rate for boys. To correct that, we calculate the rate of drug users among girls, or 1/9, and among boys, or 3/9, and we conclude that the rate of drug users in observed student population is 2/9 = 22.2%. Now, if we think that drug use is causing the non-response, the model has two*

1. Yves Tillé, Institute of Statistics, University of Neuchâtel, Avenue de Bellevaux 51, 2000 Neuchâtel, Switzerland. E-mail: yves.tille@unine.ch.

parameters p_{yes} and p_{no}, the response probabilities of users and non-users, respectively. We find that these probabilities equal 0.2 and 0.8, respectively. The estimated number of users is therefore 200 among boys and 100 among girls, and the estimated overall percentage is 50!

At first glance, the example is simple, and it perfectly explains the usual typology of the three non-response mechanisms. Each of the three estimates proposed in the example corresponds to one of the three categories below:

- Missing completely at random (MCAR): The response probability does not depend on the variable of interest (drug use) or on the auxiliary variable (gender).
- Missing at random (MAR): The response probability does not depend on the variable of interest y after conditioning on the auxiliary variable x (gender). In this case, the response probability would therefore depend on gender only.
- Not missing at random (NMAR): The response probability depends on the variable of interest itself (drug use) even if consideration is given to the auxiliary variable x.

The example shows the value of generalized calibration, which can deal directly with NMAR. Jean-Claude Deville addresses the problem by considering the probabilities p_{yes} and p_{no} as parameters to be estimated. This example can be dealt with in several ways, depending on one's point of view on inference.

In the following, we will show that there are at least three methods to address the problem, namely the method of moments, the maximum likelihood method and calibration. The maximum likelihood method was not dealt with by Jean-Claude Deville. We develop calculations completely for the first two estimation methods by considering the two models. We also calculate the calibration and generalized calibration results.

We show that the three results obtained are identical. The estimated likelihood function could be used to choose between the two models. Unfortunately, the function has the same value for both models, which does not make it possible to choose the model. However, we propose a way to make a choice.

In Section 2, we present the notation used. Section 3 is devoted to estimation using the method of moments, and Section 4 is devoted to estimation using the maximum likelihood method. In Section 5, we apply the calibration and generalized calibration methods. We close with a discussion on the value of each method in Section 6.

2 Notation

Table 2.1 shows the notation for Table 1.1.

Table 2.1
Notation for Table 1.1

	Drug User	Non-user	Missing	Total
Male	r_{HD}	r_{HS}	m_H	$n_{H.}$
Female	r_{FD}	r_{FS}	m_F	$n_{F.}$
Total	$r_{.D}$	$r_{.S}$	m	n

For simplicity, assume that we are dealing with a census. In other words, the 600 students were not randomly selected. Therefore, the only source of randomness is the non-response mechanism. This assumption is not that restrictive, since it is equivalent to considering that the sample is random, but that the reasoning below is conditional on the random sample. The objective is to estimate the numbers of people in Table 2.2. This table is assumed not to be random. It is therefore a matter of distributing the non-respondents m_H and m_F between drug users and non-users.

Table 2.2
Number of people to be estimated based on Table 1.1

	Drug User	Non-user	Total
Male	n_{HD}	n_{HS}	$n_{H.}$
Female	n_{FD}	n_{FS}	$n_{F.}$
Total	$n_{.D}$	$n_{.S}$	n

As well, it is assumed that the non-response follows a Poisson design, that is, each individual decides whether or not to respond with a probability independent of other individuals. The response probability may vary among individuals.

The two vectors (r_{HD}, r_{HS}, m_H), and (r_{FD}, r_{FS}, m_F) each have a multinomial distribution whose parameters depend on the model used. MCAR cases, which are completely trivial, will not be studied. In Table 2.3, which shows cases of MAR, the response probability depends on gender only (p_H for males, p_F for females). In Table 2.4, which shows cases of NMAR, the response probability depends only on being or not being a drug user (q_D, q_S for the others).

Table 2.3
Case 1: MAR model, non-response depends on gender

	Drug User	Non-user	Missing	Total
Male	$\mathrm{E}(r_{HD}) = n_{HD}p_H$	$\mathrm{E}(r_{HS}) = n_{HS}p_H$	$\mathrm{E}(m_H) = n_{H.}(1-p_H)$	$n_{H.}$
Female	$\mathrm{E}(r_{FD}) = n_{FD}p_F$	$\mathrm{E}(r_{FS}) = n_{FS}p_F$	$\mathrm{E}(m_F) = n_{F.}(1-p_F)$	$n_{F.}$
Total	$\mathrm{E}(r_{.D})$	$\mathrm{E}(r_{.S})$	m	n

Table 2.4
Case 2: NMAR model, non-response depends on being or not being a drug user

	Drug User	Non-user	Missing	Total
Male	$\mathrm{E}(r_{HD}) = n_{HD}q_D$	$\mathrm{E}(r_{HS}) = n_{HS}q_S$	$\mathrm{E}(m_H) = n_{HD}(1-q_D) + n_{HS}(1-q_S)$	$n_{H.}$
Female	$\mathrm{E}(r_{FD}) = n_{FD}q_D$	$\mathrm{E}(r_{FS}) = n_{FS}q_S$	$\mathrm{E}(m_F) = n_{FD}(1-q_D) + n_{FS}(1-q_S)$	$n_{F.}$
Total	$\mathrm{E}(r_{.D})$	$\mathrm{E}(r_{.S})$	m	n

3 Estimation using the method of moments

3.1 MAR

The method of moments makes it possible to estimate parameters quickly. For MAR, we obtain the third column of Table 2.3 using the equations

$$\begin{aligned} \mathrm{E}(m_H) &= n_{H.}(1-p_H), \\ \mathrm{E}(m_F) &= n_{F.}(1-p_F), \end{aligned}$$

which yield the estimators

$$\begin{aligned} \hat{p}_H &= 1-\frac{m_H}{n_{H.}}, \\ \hat{p}_F &= 1-\frac{m_F}{n_{F.}}, \end{aligned}$$

and therefore, from the first two columns,

$$\begin{aligned} \hat{n}_{.D} &= \frac{r_{HD}}{\hat{p}_H}+\frac{r_{FD}}{\hat{p}_F} = r_{HD}\frac{n_{H.}}{n_{H.}-m_H}+r_{FD}\frac{n_{F.}}{n_{F.}-m_F}, \\ \hat{n}_{.S} &= \frac{r_{HS}}{\hat{p}_H}+\frac{r_{FS}}{\hat{p}_F} = r_{HS}\frac{n_{H.}}{n_{H.}-m_H}+r_{FS}\frac{n_{F.}}{n_{F.}-m_F}. \end{aligned}$$

The estimated response probabilities are $\hat{p}_H = 0.4$ and $\hat{p}_F = 0.6$. We therefore obtain the estimates shown in Table 3.1.

Table 3.1
Estimates: MAR

	YES	NO	COMBINED
Boys	100.00	200.00	300
Girls	33.33	266.66	300
COMBINED	133.33	466.66	600

3.2 NMAR

For NMAR, we obtain the following equations from Table 2.4:

$$\begin{aligned} \mathrm{E}(m_H) &= E(r_{HD})\frac{1-q_D}{q_D}+\mathrm{E}(r_{HS})\frac{1-q_S}{q_S}, \\ \mathrm{E}(m_F) &= E(r_{FD})\frac{1-q_D}{q_D}+\mathrm{E}(r_{FS})\frac{1-q_S}{q_S}. \end{aligned}$$

After a few calculations, we obtain the following response probability estimators:

$$\hat{q}_D = \frac{r_{HD}r_{FS} - r_{FD}r_{HS}}{(m_H + r_{HD})r_{FS} - (m_F + r_{FD})r_{HS}},$$

$$\hat{q}_S = \frac{r_{HD}r_{FS} - r_{FD}r_{HS}}{(m_F + r_{FS})r_{HD} - (m_H + r_{HS})r_{FD}}.$$

Finally, we obtain

$$\hat{n}_{.D} = \frac{r_{.D}}{\hat{q}_D} = r_{.D}\frac{(m_H + r_{HD})r_{FS} - (m_F + r_{FD})r_{HS}}{r_{HD}r_{FS} - r_{FD}r_{HS}} = r_{.D}\frac{n_{H.}r_{FS} - n_{F.}r_{HS}}{r_{HD}r_{FS} - r_{FD}r_{HS}},$$

$$\hat{n}_{.S} = \frac{r_{.S}}{\hat{q}_S} = r_{.S}\frac{(m_F + r_{FS})r_{HD} - (m_H + r_{HS})r_{FD}}{r_{HD}r_{FS} - r_{FD}r_{HS}} = r_{.S}\frac{n_{F.}r_{HD} - n_{H.}r_{FD}}{r_{HD}r_{FS} - r_{FD}r_{HS}}.$$

As Deville writes, the estimated response probabilities are $\hat{q}_D = 0.2$ and $\hat{q}_S = 0.8$. We therefore obtain the estimates in Table 3.2.

Table 3.2
Estimates: NMAR

	YES	NO	COMBINED
Boys	200	100	300
Girls	100	200	300
COMBINED	300	300	600

4 Estimation using the maximum likelihood method

4.1 MAR

The probability distribution is multinomial. For MAR, the following likelihood function applies:

$$\mathcal{L}(n_{HD}, n_{FD}, p_H, p_F) = \frac{n_{H.}!}{r_{HD}!\, r_{HS}!\, m_H!}\left(\frac{n_{HD}p_H}{n_{H.}}\right)^{r_{HD}}\left(\frac{(n_{H.} - n_{HD})p_H}{n_{H.}}\right)^{r_{HS}}\left(\frac{n_{H.}(1-p_H)}{n_{H.}}\right)^{m_H}$$
$$\times\frac{n_{F.}!}{r_{FD}!\, r_{FS}!\, m_F!}\left(\frac{n_{FD}p_F}{n_{F.}}\right)^{r_{FD}}\left(\frac{(n_{F.} - n_{FD})p_F}{n_{F.}}\right)^{r_{FS}}\left(\frac{n_{F.}(1-p_F)}{n_{F.}}\right)^{m_F}.$$

By setting to zero the partial derivatives of the log-likelihood with respect to parameters p_H and p_F, we obtain two equations with two unknowns. The solution yields the estimators

$$\hat{p}_H = 1 - \frac{m_H}{n_{H.}},$$

$$\hat{p}_F = 1 - \frac{m_F}{n_{F.}}.$$

By setting to zero the derivatives with respect to n_{HD} and n_{FD}, we obtain the estimators

$$\hat{n}_{HD} = \frac{r_{HD}}{\hat{p}_H} \text{ and } \hat{n}_{FD} = \frac{r_{FD}}{\hat{p}_F}.$$

Therefore,

$$\hat{n}_{.D} = \hat{n}_{HD} + \hat{n}_{FD} = \frac{r_{HD}}{\hat{p}_H} + \frac{r_{FD}}{\hat{p}_F}.$$

These estimators are exactly the same as those obtained using the method of moments.

4.2 NMAR

For NMAR, the following likelihood function applies:

$$\begin{aligned}\mathcal{L}(n_{HD}, n_{FD}, q_D, p_S) &= \frac{n_{H.}!}{r_{HD}!\, r_{HS}!\, m_H!}\left(\frac{n_{HD} q_D}{n_{H.}}\right)^{r_{HD}} \left(\frac{(n_{H.} - n_{HD}) q_S}{n_{H.}}\right)^{r_{HS}} \left(\frac{n_{HD}(1-q_D) + (n_{H.} - n_{HD})(1-q_S)}{n_{H.}}\right)^{m_H} \\ &\times \frac{n_{F.}!}{r_{FD}!\, r_{FS}!\, m_F!}\left(\frac{n_{FD} q_D}{n_{F.}}\right)^{r_{FD}} \left(\frac{(n_{F.} - n_{FD}) q_S}{n_{F.}}\right)^{r_{FS}} \left(\frac{n_{FD}(1-q_D) + (n_{F.} - n_{FD})(1-q_S)}{n_{F.}}\right)^{m_F}.\end{aligned}$$

By setting to zero the partial derivatives of the log-likelihood with respect to the four parameters q_D, q_S, n_{HD} and n_{FD}, we obtain a system of four rather complicated second-order equations with four unknowns. We used a symbolic computation software program to verify that the solution given by the method of moments is a solution to this system of equations. Obviously, since the system is second-order, there is a second solution. However, for Deville's example, the second solution yields negative values, which are not valid for estimating probabilities and numbers of people.

5 Estimation using calibration and generalized calibration

5.1 Notation

To define calibration, we will establish the following notation. Let $U = \{1, \ldots, k, \ldots, N\}$ be the set of people interviewed (here, $N = 600$) and $R \subset U$ be the set of respondents to the question regarding drug use. As well, we define the following:

$$\mathbf{x}_k = \begin{cases} (1 \quad 0)^{\mathrm{T}} & \text{if individual } k \text{ is male} \\ (0 \quad 1)^{\mathrm{T}} & \text{if individual } k \text{ is female.} \end{cases}$$

and

$$\mathbf{z}_k = \begin{cases} (1 \quad 0)^{\mathrm{T}} & \text{if individual } k \text{ reported using drugs} \\ (0 \quad 1)^{\mathrm{T}} & \text{if individual } k \text{ reported not using drugs.} \end{cases}$$

Using the notation defined above,

$$\sum_{k\in U}\mathbf{x}_k = \begin{pmatrix} n_{H.} \\ n_{F.} \end{pmatrix}, \quad \sum_{k\in R}\mathbf{x}_k = \begin{pmatrix} n_{H.}-m_H \\ n_{F.}-m_F \end{pmatrix}, \quad \sum_{k\in R}\mathbf{z}_k = \begin{pmatrix} r_{.D} \\ r_{.S} \end{pmatrix},$$

$$\sum_{k\in R}\mathbf{x}_k\mathbf{x}_k^{\mathrm{T}} = \begin{pmatrix} n_{H.}-m_H & 0 \\ 0 & n_{F.}-m_F \end{pmatrix}, \quad \sum_{k\in R}\mathbf{x}_k\mathbf{z}_k^{\mathrm{T}} = \begin{pmatrix} r_{HD} & r_{HS} \\ r_{FD} & r_{FS} \end{pmatrix},$$

and

$$\sum_{k\in R}\mathbf{z}_k\mathbf{z}_k^{\mathrm{T}} = \begin{pmatrix} r_{.D} & 0 \\ 0 & r_{.S} \end{pmatrix}.$$

5.2 Estimation using simple calibration

Using simple calibration as described in Deville and Särndal (1992), we seek a weight that is expressed as

$$w_k = F\left(\mathbf{x}_k^{\mathrm{T}}\boldsymbol{\lambda}\right),$$

where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ is a parameter vector and $F(.)$ is a calibration function, that is, a strictly increasing function such that $F(0) = 1$ and whose derivative $F'(.)$ is such that $F'(0) = 1$.

Vector $\boldsymbol{\lambda}$ is determined by using the Newton method to solve the system of equations

$$\sum_{k\in R}F\left(\mathbf{x}_k^{\mathrm{T}}\boldsymbol{\lambda}\right)\mathbf{x}_k = \sum_{k\in U}\mathbf{x}_k. \tag{5.1}$$

Finally, the calibration estimator is given by

$$\begin{pmatrix} \hat{n}_{.D} \\ \hat{n}_{.S} \end{pmatrix} = \sum_{k\in R} w_k\mathbf{z}_k.$$

In our application, equation (5.1) becomes

$$\sum_{k\in R}F\left(\mathbf{x}_k^{\mathrm{T}}\boldsymbol{\lambda}\right)\mathbf{x}_k = \begin{pmatrix} (n_{H.}-m_H)F(\lambda_1) \\ (n_{F.}-m_F)F(\lambda_2) \end{pmatrix} = \sum_{k\in U}\mathbf{x}_k = \begin{pmatrix} n_{H.} \\ n_{F.} \end{pmatrix}.$$

We directly obtain the following:

$$w_k = F\left(\mathbf{x}_k^{\mathrm{T}}\boldsymbol{\lambda}\right) = \begin{cases} n_{H.}/(n_{H.}-m_H) & \text{if individual } k \text{ is male} \\ n_{F.}/(n_{F.}-m_F) & \text{if individual } k \text{ is female.} \end{cases}$$

Therefore, the calibrated estimators are

$$\hat{n}_{.D} = r_{HD}\frac{n_{H.}}{n_{H.}-m_H} + r_{FD}\frac{n_{F.}}{n_{F.}-m_F}$$

$$\hat{n}_{.S} = r_{HS}\frac{n_{H.}}{n_{H.}-m_H} + r_{FS}\frac{n_{F.}}{n_{F.}-m_F},$$

which is exactly the same result as that yielded by the method of moments and the maximum likelihood method. In this case, the solution does not depend on the calibration function used. Obviously, the example is especially simple. In more complex cases where the category definitions do not overlap, the result depends on the calibration function used.

5.3 Generalized calibration

For generalized calibration as defined in (Deville 2000, 2002, 2004; Kott 2006), the weights are expressed as

$$w_k = F\left(\mathbf{z}_k^{\mathrm{T}}\boldsymbol{\lambda}\right).$$

Vector $\boldsymbol{\lambda}$ is determined by solving the system of equations

$$\sum_{k\in R} F\left(\mathbf{z}_k^{\mathrm{T}}\boldsymbol{\lambda}\right)\mathbf{x}_k = \sum_{k\in U}\mathbf{x}_k. \tag{5.2}$$

Finally, the generalized calibration estimator is given by

$$\begin{pmatrix}\hat{n}_{.D}\\ \hat{n}_{.S}\end{pmatrix} = \sum_{k\in R} w_k\mathbf{z}_k.$$

In our application, equation (5.2) becomes

$$\sum_{k\in R} F\left(\mathbf{z}_k^{\mathrm{T}}\boldsymbol{\lambda}\right)\mathbf{x}_k = \begin{pmatrix} r_{HD}F\left(\lambda_1\right) + r_{HS}F\left(\lambda_2\right)\\ r_{FD}F\left(\lambda_1\right) + r_{FS}F\left(\lambda_2\right)\end{pmatrix} = \sum_{k\in U}\mathbf{x}_k = \begin{pmatrix} n_{H.}\\ n_{F.}\end{pmatrix},$$

Which can be written as a matrix

$$\begin{pmatrix} r_{HD} & r_{HS}\\ r_{FD} & r_{FS}\end{pmatrix}\begin{pmatrix} F\left(\lambda_1\right)\\ F\left(\lambda_2\right)\end{pmatrix} = \begin{pmatrix} n_{H.}\\ n_{F.}\end{pmatrix}.$$

We simply solve the linear system

$$\begin{pmatrix} F\left(\lambda_1\right)\\ F\left(\lambda_2\right)\end{pmatrix} = \begin{pmatrix} r_{HD} & r_{HS}\\ r_{FD} & r_{FS}\end{pmatrix}^{-1}\begin{pmatrix} n_{H.}\\ n_{F.}\end{pmatrix} = \begin{pmatrix} \dfrac{n_{H.}r_{FS} - n_{F.}r_{HS}}{r_{FS}r_{HD} - r_{FD}r_{HS}}\\ \dfrac{n_{H.}r_{FD} - n_{F.}r_{HD}}{r_{FD}r_{HS} - r_{FS}r_{HD}}\end{pmatrix}.$$

The estimators are therefore

$$\hat{n}_{.D} = r_{.D}\frac{n_{H.}r_{FS} - n_{F.}r_{HS}}{r_{FS}r_{HD} - r_{FD}r_{HS}}$$

$$\hat{n}_{.S} = r_{.S}\frac{n_{H.}r_{FD} - n_{F.}r_{HD}}{r_{FD}r_{HS} - r_{FS}r_{HD}}.$$

Again, the solution does not depend on the calibration function used. The solution is identical to the solution obtained using the method of moments and the maximum likelihood method. Here, too, this property results from the simplicity of the example. In more complex cases, the result depends on the calibration function used.

6 Discussion

Deville's example is especially welcome since, for both models, the three estimation methods provide exactly the same estimators. Obviously, if the model is more complicated, using the maximum likelihood method becomes cumbersome, if not impossible. The calibration and generalized calibration method works in all cases as long as the number of calibration variables whose totals are known is sufficient and the matrix

$$\sum_{k \in R} \mathbf{x}_k \mathbf{z}_k^{\mathrm{T}}$$

is invertible. In this example, the determinant of this matrix appears in the denominator of the estimators. Therefore, a small determinant makes the estimates especially risky. Lesage and Haziza (2015) recommend verifying that the correlations between variables $\mathbf{x}_k$ and $\mathbf{z}_k$ are great enough to avoid potentially amplifying the bias.

If the variables are quantitative, the solutions will depend on the calibration function used $F(.)$. The use of the calibration function $F\left(\mathbf{z}_k^{\mathrm{T}}\boldsymbol{\lambda}\right) = 1 + \exp\left(\mathbf{z}_k^{\mathrm{T}}\boldsymbol{\lambda}\right)$ is recommended, since it has the advantage of providing weights greater than 1. The inverse of the weights can now be interpreted as a response probability estimated using a logistic model.

The main difficulty is obviously choosing between the two proposed models. In Deville's example, it may seem more "logical" to see the non-response depend rather on drug use than on gender. However, we are not well equipped to make a choice between the two models. The values of the two likelihood functions for the estimated parameters are equal. Is it possible to choose the model based on more than a strong conviction? As suggested in Haziza and Lesage (2016), we recommend always calculating both weightings and comparing the weights and estimates obtained with each of them.

One option may be to calculate an indicator of the dispersion of the response probabilities, such as the variance. For example, if the variance is great, it means that the model has made it possible to calculate response probabilities with greater contrast between individuals and that the model has therefore taken better account of the non-response. Validation through a search for contrasting weights is the basis for identifying response homogeneity groups (RHGs) for all segmentation methods, for example with the chi-square automatic interaction detector (CHAID) algorithm developed by Kass (1980). For example, with CHAID, in each step the RHGs are split based on categories that result in response probabilities with the greatest contrast. By using the same principle in choosing the model, we can select the model that provides the weights with the greatest contrast. For example, if the variance is small, it means that the non-response model could not highlight the differences in non-response probabilities between individuals. Incidentally, the variance in response probabilities is the square of the R-indicator defined by Schouten, Cobben and Bethlehem (2009), used here to choose a non-response model.

In both cases, the average response probability equals 0.5. Specifically,

$$\overline{p} = n_{H.}\frac{n_{H.}\hat{p}_H + n_{F.}\hat{p}_F}{n} = \frac{300 \times 0.4 + 300 \times 0.6}{600} = 0.5$$

and

$$\overline{q} = \hat{n}_{.D}\frac{n_{.D}\hat{q}_D + \hat{n}_{.S}\hat{q}_S}{n} = \frac{300\times 0.2 + 300\times 0.8}{600} = 0.5.$$

For the MAR model, the variance is

$$V_{MAR} = \frac{n_{H.}\left(\hat{p}_H - \overline{p}\right)^2 + n_{F.}\left(\hat{p}_F - \overline{p}\right)^2}{n} = \frac{300(0.4-0.5)^2 + 300(0.6-0.5)^2}{600} = 0.01.$$

For the NMAR model, the variance is

$$V_{NMAR} = \frac{\hat{n}_{.D}\left(\hat{q}_D - \overline{q}\right)^2 + \hat{n}_{.S}\left(\hat{q}_S - \overline{q}\right)^2}{n} = \frac{300(0.2-0.5)^2 + 300(0.8-0.5)^2}{600} = 0.09.$$

The greater variance of the NMAR model is an argument in its favour. In fact, the response probabilities show much greater contrast.

Acknowledgements

The author thanks Audrey-Anne Vallée for her meticulous proofreading of an earlier version of this text and an anonymous referee for their especially pertinent comments.

References

Chang, T., and Kott, P.S. (2008). Using calibration weighting to adjust for nonresponse under a plausible model. *Biometrika*, 95, 555-571.

Deville, J.-C. (2000). Generalized calibration and application to weighting for non-response. In *Compstat - Proceedings in Computational Statistics:* 14th Symposium held in Utrecht, Netherlands, pages 65-76, New York: Springer.

Deville, J.-C. (2002). La correction de la nonréponse par calage généralisé. In the *Actes des Journées de Méthodologie Statistique*, Paris. Insee-Méthodes.

Deville, J.-C. (2004). Calage, calage généralisé et hypercalage. Technical report, internal document, INSEE, Paris.

Deville, J.-C. (2005). Calibration, past, present and future? Presentation at the conference: *Calibration Tools for Survey Statisticians*, Neuchâtel.

Deville, J.-C., and Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87, 376-382.

Haziza, D., and Lesage, E. (2016). A discussion of weighting procedures for unit nonresponse. Will appear in the *Journal of Official Statistics*.

Kass, G.V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, 119-127.

Kott, P.S. (2006). Using calibration weighting to adjust for nonresponse and coverage errors. *Survey Methodology*, 32, 2, 133-142. Paper available at http://www.statcan.gc.ca/pub/12-001-x/2006002/article/9547-eng.pdf.

Kott, P.S., and Chang, T. (2010). Using calibration weighting to adjust for nonignorable unit nonresponse. *Journal of the American Statistical Association*, 105(491), 1265-1275.

Lesage, E., and Haziza, D. (2015). On the problem of bias and variance amplification of the instrumental calibration estimator in the presence of unit nonresponse. Under revision for *Journal of Survey Statistics and Methodology*.

Schouten, B., Cobben, F. and Bethlehem, J. (2009). Indicators for the representativeness of survey response. *Survey Methodology*, 35, 1, 101-113. Paper available at http://www.statcan.gc.ca/pub/12-001-x/2009001/article/10887-eng.pdf.