Measuring response errors in censuses and surveys through reinterview data having a different measurement error variance

Piero Demetrio Falorsi, Alessandro Pallara

National Institute of Statistics (ISTAT), Via Magenta 2, 00184, Rome, Italy
e-mail: falorsi@istat.it; pallara@istat.it

Aldo Russo

Department of Political Institutions and Social Science, Università Roma Tre, Rome, Italy

Keywords: Measurement errors, Reinterview, Response bias, Simple and correlated measurement variance.

1. Introduction

Reinterviews have been extensively used as a tool for estimating and reducing response errors in censuses and sample surveys. By *response error* it is meant any error occurring at the data collection stage for a variety of reasons. Errors may be due to the respondent, to the interviewer, or to both. Such errors can never be completely avoided and, therefore, in the practice of censuses and large surveys some different techniques have been proposed (Forsman and Schreiner, 1991) in order to measure response error components. One such method relies on replicated measurements obtained through reinterview of a subsample of units from the original survey on a set of questions from the original interview. When the responses given by the same units during the reinterview differ from those given in the original interview the differences can be evaluated through *reconciliation*, normally by asking the respondent to determine what is the correct information between the two interviews.

When using the approach based on reinterviews to measure response error components, it is usually assumed (Hansen *et al.*, 1961; Särndal *et al.*, 1992; pp. 614-616) that: (*i*) the two measures are modelled as random variables; (*ii*) the repeated measurements on the same unit are independent; (*iii*) the measurement conditions for the two occasions are identical or as close to identical as possible; this imply that the random variables which are associated to the two measurements are subject to the *same measurement error variance*. Under these assumptions, the standard response error model based on reinterview data, developed at the U.S. Bureau of Census (Hansen *et al.*, 1964), yields unbiased estimates of simple and correlated measurement variances (Särndal *et al.*, 1992; pp. 614-616). Furthermore, an estimate of the measurement bias can be obtained, assuming that the reconciled reinterview process yields the true value. While theoretically sound, the Bureau of Census survey error model seems to be not realistic in many practical situations, because some of the above assumptions are difficult to be met. Frequently, e.g., the reinterview program, because of budget and operational restrictions, are carried out using a different data collection technique with respect to the original interview. In this case, the two measurements, although still independent, cannot be considered as obtained under identical measurement conditions and, therefore, they can be deemed as having different measurement error variances.

In this paper it is shown how to obtain an unbiased estimate of the response variance when the assumption of identical second-order moment of the two measurements fails to hold. Under the standard response error model the variance of the standard estimator of the population total of the true values is decomposed in the sum of the sampling variance and measurement variance and the measurement variance can be split in two terms, both depending from the (unknown) variances and covariance between measurements, but only one depending also from the sampling design. Indeed, in presence of measurement errors the estimator of the total variance is biased and it can be shown (Koch *et al.*; 1975) that the bias is equal to the component of the measurement variance not depending on the sampling design. A method of moments estimation of the simple and correlated measurement variance components under the standard response error model is then proposed, assuming different error variances of the repeated measurements in different surveys.

The paper is organised as follows: in section 2 the parameter of interest and the statistical model for the response error are shortly introduced and then estimation error variance is decomposed in a sampling and a non-sampling component; in section 3 it is shown how to obtain an unbiased estimation of total error variance in presence of measurement errors. Finally, some empirical results are presented in section 4, concerning estimates of measurement variance components for some variables pertaining to land usage such as measured during the 5th Italian Census of Agriculture, and whose accuracy has been evaluated through a reinterview survey. In this case, while the original census responses have been obtained through face-to-face interview, the reinterview survey was based on Computer Assisted Telephone Interviewing (CATI): hence, the

two survey measures have clearly not been collected under the same conditions, which may result in different measurement error variances.

2. The measurement error model

Consider the following conditions: (*i*) an original random sample *s* (survey I_1) of size *n* is drawn from the population *P*, constituted by *N* units, with a sampling design *p* (·), with π_k representing the probability that element *k* will be included in the sample and π_{kl} representing the probability that both of the elements *k* and *l* will be included in the sample; (*ii*) for each element $k \in s$, $y_{k,1}$ denotes the observed value of the variable of interest *y*; (*iii*) a second phase sample *r* ($r \subseteq s$) (survey I_2) of size n_r is drawn according to the design $p(\cdot|s)$, such that $p(r|s)$ is the conditional probability of choosing *r*. The inclusion probabilities under this design are denoted $\pi_{k|s}$ and $\pi_{kl|s}$ for elements *k* and $l \in s$; (*iv*) for each element $k \in r$, $y_{k,2}$ denotes the observed value of the variable *y* of interest. The measurement model *m* is specified as follows:

$$y_{k,t} = \mu_k + \varepsilon_{k,t} \qquad (t = 1,2) \tag{1}$$

where μ_k is the *true value* and $\varepsilon_{k,t}$ is the *erratic* component.

Given the general survey conditions of I_1 and of I_2, the expectation values $E_m(\cdot)$ under the model (1) are:

$$E_m(y_{k,t}) = \mu_k \qquad \text{for } k \in s \text{ if } t=1 \text{ or } k \in r \subseteq s \text{ if } t=2$$

$$E_m(y_{k,t} - \mu_k)^2 = \sigma_{k,t}^2 \qquad \text{for } k \in s \text{ if } t=1 \text{ or } k \in r \subseteq s \text{ if } t=2$$

$$E_m((y_{k,t} - \mu_k)(y_{l,t} - \mu_l)) = \sigma_{kl,t} \qquad \text{for } k,l \in s \text{ (with } k \neq l\text{) if } t=1 \text{ or } k,l \in r \subseteq s \text{ (with } k \neq l\text{) if } t=2$$

$$E_m((y_{k,1} - \mu_k)(y_{l,2} - \mu_l)) = 0 \qquad \text{for } k,l \in r \subseteq s \text{ (with } k = l \text{ or } k \neq l\text{).}$$

It is important to note that in the above model both the values μ_k and $\sigma_{k,t}^2$, are related to the specific unit $k \in P$, while $\sigma_{kl,t}$ is a value related to the specific couple $(k,l) \in P$, independently by the selected sample in the I_t (t =1 or 2) survey.

The objective of the survey I_1 is to estimate the population total of the *true values*: $t_y = \sum_P \mu_k$, where we denote with $\sum_P$ the sum over the *N* units of the population *P*. An estimate of t_y may be obtained by the Horvitz-Thompson estimator

$$\tilde{t}_y = \sum_s \frac{y_{k,1}}{\pi_k}. \tag{2}$$

As a measure of the accuracy of the estimator (2), we use the variance of $\tilde{t}_y$ defined with respect to *design* $p(\cdot)$ and *model m* jointly

$$V_{pm}(\tilde{t}_y) = E_{pm}((\tilde{t}_y - t_y)^2) \tag{3}$$

where the indexes *p* and *m* define the operators *E* and *V* with respect to the sample design and the model.

The variance term $V_{pm}(\tilde{t}_y)$ can be decomposed as follows

$$V_{pm}(\tilde{t}_y) = E_p(V_m(\tilde{t}_y|s)) + V_p(E_m(\tilde{t}_y|s)).$$

The *measurement variance*, given by $E_p(V_m(\tilde{t}_y|s)) = V_{A1} + V_{A2}$, may be represented as a sum of two terms

$$V_{A1} = V_{A1,1} + V_{A1,2} = \sum_P \sigma_{k,1}^2 + \sum_k \sum_{l \neq k}{}_P \ \sigma_{kl,1} \tag{4}$$

$$V_{A2} = \sum_P \frac{(1-\pi_k)}{\pi_k} \sigma_{k,1}^2 + \sum_k \sum_{l \neq k}{}_P \ \frac{(\pi_{kl} - \pi_k \pi_l)}{\pi_k \pi_l} \sigma_{kl,1} \tag{5}$$

where V_{A1}, that denotes the measurement variance under complete enumeration, is the sum of the two components $V_{A1,1} = \sum_P \sigma_{k,1}^2$, indicating the *simple measurement variance* and $V_{A1,2} = \sum_k \sum_{l \neq k}{}_P \ \sigma_{kl,1}$, denoting the *correlated measurement variance*. Usually in survey practice the main component in (4) is represented by the $V_{A1,2}$ term, while the simple response variance is typically negligible.

The component $V_p\left(E_m\left(\tilde{t}_y|s\right)\right)=V_B$, denoting the *sampling variance,* is expressed by

$$V_p\left(E_m\left(\tilde{t}_y|s\right)\right)=V_B=\sum_k\sum_l{}_P\ \frac{\pi_{kl}-\pi_k\,\pi_l}{\pi_k\,\pi_l}\mu_k\mu_l\ .$$

3. Estimation of variance in presence of measurement errors

As shown in Särndal *et al*. (1992), in presence of measurement errors, the standard estimator of variance of $\tilde{t}_y$

$$\tilde{V}\left(\tilde{t}_y\right)=\sum_k\sum_l{}_s\ \frac{\pi_{kl}-\pi_k\,\pi_l}{\pi_{kl}}\frac{y_{k,1}}{\pi_k}\frac{y_{l,1}}{\pi_l}$$

is biased, being valid the following result $E_{pm}\left(\tilde{V}\left(\tilde{t}_y\right)\right)=V_{pm}\left(\tilde{t}_y\right)-V_{A1}$.

Therefore, for obtaining an unbiased estimate of $V_{pm}\left(\tilde{t}_y\right)$, it is sufficient to yield an unbiased estimate of the V_{A1} component. In order to find a solution to this problem, it will be introduced in model (1) the simplified assumption that $\sigma_{k,1}^2\cong\sigma_{k,2}^2\cong\sigma_k^2$. This simplification may be justified by noting that the simple response variance $\sigma_{k,t}^2$ (t=1,2) is directly related to the individual response mechanism which can be deemed as substantially not depending from the collection mode, while it is necessary to distinguish between the two correlated components $\sum_k\sum_{l\neq k}{}_P\ \sigma_{kl,1}$ and $\sum_k\sum_{l\neq k}{}_P\ \sigma_{kl,2}$ which are strictly related to the collection mode of the survey. Indeed since $\sum_k\sum_{l\neq k}{}_P\ \sigma_{kl,1}$ represents, as noted above, the leading components of measurement variance, the assumed simplification will have a very small effect on the estimation of V_{A1}.

Under the statistical model presented in section 2 and assuming that $\sigma_{k,1}^2\cong\sigma_{k,2}^2\cong\sigma_k^2$, it can be shown (see Appendix) that an unbiased estimate of the vector of four unknown terms $\mathbf{X}=\left(X_1,...,X_4\right)'$, where

$$X_1=\sum_P{}_P\,\mu_k^2-\frac{1}{N}\left(\sum_P{}_P\,\mu_k\right)^2\ ,\quad X_2=V_{A1,1}=\sum_P{}_P\sigma_{k,1}^2=\sum_P{}_P\sigma_k^2\ ,$$

$$X_3=V_{A1,2}=\sum_k\sum_{l\neq k}{}_P\ \sigma_{kl,1}\ ,\quad X_4=\sum_k\sum_{l\neq k}{}_P\ \sigma_{kl,2}$$

may be found as

$$\hat{\mathbf{X}}=\mathbf{A}^{-1}\tilde{\mathbf{b}} \tag{6}$$

in which $\mathbf{A}$ is a matrix of fixed coefficients

$$\mathbf{A}=\begin{bmatrix}1 & (N-1)/N & -1/N & 0\\ 0 & 1 & 0 & 0\\ 0 & 1 & 0.5 & 0.5\\ 1 & 0 & 0 & 0\end{bmatrix}$$

and

$\tilde{\mathbf{b}}=\left(\tilde{b}_1,...,\tilde{b}_4\right)'$ is a vector of sampling estimates, where:

$$\tilde{b}_1=\sum_s{}_s\ \frac{y_{k,1}^2}{\pi_k}\left(1-\frac{1}{N}\right)-\frac{1}{N}\sum_k\sum_{l\neq k}{}_s\ \frac{y_{k,1}\,y_{l,1}}{\pi_{kl}}\ ,\quad \tilde{b}_2=\frac{1}{2}\sum_r{}_r\ \frac{1}{\pi_{k,2}}\left(y_{k,1}-y_{k,2}\right)^2\ ,$$

$$\tilde{b}_3=\frac{1}{2}\sum_r{}_r\ \frac{1}{\pi_{k,2}}\left(y_{k,1}^2+y_{k,2}^2-2y_{k,1}\,y_{k,2}\right)+\sum_k\ \sum_{l\neq k}{}_r\ \frac{1}{\pi_{kl,2}}\left(y_{k,1}\,y_{l,1}+y_{k,2}\,y_{l,2}-2y_{k,1}\,y_{l,2}\right)$$

$$\tilde{b}_4 = \sum_r \frac{y_{k,1}\ y_{k,2}}{\pi_{k,2}}\left(1-\frac{1}{N}\right) - \frac{1}{N}\sum_k \sum_{l \neq k}{}_r \frac{y_{k,1}\ y_{l,2}}{\pi_{kl,2}}$$

with $\pi_{k,2} = \pi_k\, \pi_{k|s}$ and $\pi_{kl,2} = \pi_{kl}\, \pi_{kl|s}$.

Using expression (6) on the basis of the estimates $\hat{X}_i$ (i =1,…,4), it is then possible to obtain an estimate of the variance V_{A1} as

$$\hat{V}_{A1} = \hat{X}_2 + \hat{X}_3 = \hat{V}_{A1,1} + \hat{V}_{A1,2}\,, \tag{7}$$

with $\hat{X}_2 = \tilde{b}_2$ and $\hat{X}_3 = (N-1)\tilde{b}_2 + N(\tilde{b}_4 - \tilde{b}_1)$ representing unbiased estimates of, respectively, the simple and correlated components of the measurement variance under complete enumeration in (4).

An estimate of $V_{pm}(\tilde{t}_y)$ may then be given by

$$\tilde{V}_{pm}(\tilde{t}_y) = \tilde{V}(\tilde{t}_y) + \hat{V}_{A1} \tag{8}$$

Finally, we note that in the *standard approach*, where it is assumed: (i) $\sigma^2_{k,1} = \sigma^2_{k,2}$ (for $k \in r$ or $k \in s$); (ii) $\sigma_{kl,1} = \sigma_{kl,2}$ (for $k,l \in r$ or $k,l \in s$); (iii) $E_m\left(\left(y_{k,1} - \mu_k\right)\left(y_{l,2} - \mu_l\right)\right) = 0$ (for $k,l \in r \subseteq s$), an unbiased estimate of the components of measurement variance may be obtained as (Särndal *et al.*, 1992; pp. 614-617):

$$stand\,\hat{V}_{A1} = \tilde{b}_3 \ ;\ stand\,\hat{V}_{A1,1} = \tilde{b}_2 = \hat{V}_{A1,1} \ ;\ stand\,\hat{V}_{A1,2} = stand\,\hat{V}_{A1} - stand\,\hat{V}_{A1,1}\,. \tag{9}$$

4. Empirical results

The proposed methodology is illustrated with an application to the reinterview survey carried out in 2001 to assess accuracy of the 5th Italian Census of Agriculture. The census of agriculture, taken every ten years, collects data and publishes information on land usage, crops and livestock, operator as well as farm characteristics by farms in Italy. Data collection for the 5th Census of Agriculture began at the end of November 2000, through face-to-face interviewing of some 2,75 million names, from individual or family operations to very large corporations as well as publicly managed farms or woodlands included in the census farm list. Field operations were concluded by the end of March 2001, with a few exceptions that were made primarily for small towns that had been affected by floods during the data collection period.

A reinterview survey was planned at the end of data collection to evaluate accuracy of the information recorded during census field work. The sample design is two-stage with farms selected within municipalities (primary sampling units) to control geographical spread of the sample. The choice of this sampling plan was motivated essentially from budget and time constraints, since survey design involved a number of preliminary operations related to special processing of census questionnaires of farms selected in the reinterview sample. The selected sample included some 8200 farms in about 200 municipalities from all regions of the country.

As mentioned previously, interviewing was conducted through CATI. Using CATI had a number of advantages with respect to field reinterview for obtaining uncontaminated estimates of response errors (cf. Forsman and Schreiner, 1991), although in this case it was the only option available due to limited budget allocated to the reinterview program. Primarily, for the measurement of response bias there were specific advantages concerning reconciliation of differences because: *a*) the reinterviewer had no access to the original interview data, unless the difference between responses to the original interview and the reinterview for a given question did not cross certain tolerance limits which had been programmed in the software for the reinterviews; b) it was not possible for the reinterviewer to alter the reinterview response once the reinterview had been completed.

Because of the complexity of the original census questionnaire, it was decided that for successful realization of telephone interviewing only a subset of the original questions was included in the reinterview, pertaining to: *i*) major crops, flowers or tree-growing and vineyards, *ii*) cattle and poultry raising and *iii*) family and other personnel employed in the farm. The reinterview included some 50 questions, 30 of which involving reconciliation as part of the reinterview process. In order to replicate as much as possible the original interview scenario, reinterviewers were instructed to ask speaking to the person that completed census questionnaire, assuming that this was the most knowledgeable person in the farm. This was deemed to avoid spurious effects in estimating response error components.

For the application here considered, since the first survey is a census, only the measurement variance under complete enumeration V_{A1} and its components $V_{A1,1}$ and $V_{A1,2}$ are different from zero.

In table 1 estimates of the measurement variance obtained with the proposed methodology (expressions (6) and (7)), are compared with the estimates obtained with the standard approach (expression (9)), for some of the main items constituting land usage in census of agriculture. In order to have a better understanding, the results are presented as the percent ratio of

the root square of the estimate of each component of the measurement variance over the estimated total surface for the selected census item

$$RV_c = \frac{\sqrt{\widetilde{V}_c}}{\widetilde{t}_y} 100 \qquad (10)$$

with $\widetilde{V}_c$ denoting alternatively the estimate of each component V_{A1}, $V_{A1,1}$, $V_{A1,2}$, through (6)-(7) for the proposed methodology and (9) for the standard approach.

Furthermore, the table shows the estimates of the *correlation between response errors*, ρ, (Cochran, 1977, pp.387) estimated by

$$\widetilde{\rho} = \frac{\hat{V}_{A1}}{(N-1)\hat{V}_{A1,1}} - \frac{1}{N-1} \quad \text{and} \quad {}_{stand}\widetilde{\rho} = \frac{{}_{stand}\hat{V}_{A1}}{(N-1)\,{}_{stand}\hat{V}_{A1,1}} - \frac{1}{N-1} \qquad (11)$$

for the proposed and the standard approach, respectively.

Table 1: Empirical comparison of the estimates of the components of measurement variance with different methodologies

Item	Simple Measurement Variance	Proposed Methodology			Standard Methodology		
	$RV_{A1,1}$	$RV_{A1,2}$	RV_{A1}	$\widetilde{\rho}$	$RV_{A1,2}$	RV_{A1}	${}_{stand}\widetilde{\rho}$
Field crops (including greenhouse and fluriculture)	0.071	6.941	6.941	0.00762	1.734	1.735	0.00048
of which: Wheat	0.167	3.646	3.650	0.00047	2.846	2.851	0.00029
Corn for grains	0.091	5.578	5.578	0.00370	1.608	1.611	0.00031
Fruit trees	0.057	1.375	1.376	0.00047	1.195	1.197	0.00035
of which: Vineyards	0.063	2.009	2.010	0.00082	1.115	1.117	0.00025
Olive trees	0.031	4.199	4.199	0.01619	0.764	0.764	0.00054
Apple trees	0.134	4.027	4.029	0.00074	2.319	2.323	0.00024
Haylage, silage and permanent pasture and graceland	0.113	3.221	3.223	0.00065	1.251	1.256	0.00010
All cropland	0.055	5.410	5.410	0.00784	1.297	1.299	0.00045
Woodlands	0.043	3.865	3.865	0.00656	1.112	1.113	0.00054
All other lands	0.153	2.972	2.976	0.00030	7.040	7.042	0.00170
Total land use	0.074	5.886	5.886	0.00505	2.238	2.240	0.00073

The results of the application suggest the following considerations:

- the estimate of the correlated component of the measurement variance, $V_{A1,2}$, obtained with the proposed methodology is always larger than that resulting from the standard approach; since $V_{A1,2}$ is the dominant component of the total measurement error, the estimate of the total measurement error V_{A1} resulting from the proposed methodology is larger with respect to the standard approach, which seems to underestimate the true values of the measurement errors, in the example herein considered; consequently, the resulting estimate of the *correlation between response errors*, ρ, which also depends from V_{A1}, is larger in the case of the proposed methodology;
- the results confirm that the simple response variance is a small component of total response error and then the most relevant issue when analyzing measurement error is to yield accurate estimation of the correlated component of response error;
- the differences in the results obtained with the two approaches suggest that when the reinterview uses a different collection mode, the standard approach - which assumes the two surveys having the same response variance - fails to hold, and one should consider an estimation which takes into account explicitly the difference between components of response errors in the two surveys;

– the results with the proposed methodology indicates that the measurement error under complete enumeration may be a significant component of total error variance, that is important to account for when assessing the accuracy of the reported results of surveys and censuses.

References

Cochran W.G. (1977), *Sampling Techniques*, Wiley, New York.

Forsman G., Schreiner I. (1991), The Design and Analysis of Reinterview: an Overview, in: *Measurement Errors in Surveys*, Biemer P., Groves, R.M., Lyberg L., Mathiowetz N. and Sudman S. (Eds.), Wiley, New York, 279-302.

Hansen M., Hurwitz W., Bershad M. A. (1961), Measurements Errors in Census and Surveys, *Bullettin of the International Statistical Institute*, 38:2, 359-374.

Hansen M. H., Hurwitz W.N., Pritzker L. (1964), The Estimation and Interpretation of Gross Differences and the Simple Response Variance, in: *Contribution to Statistics*, Rao C.R. (Eds.), Calcutta, Statistical Publishing Society, 111-136.

Koch G. G., Freeman D. H., Jr., Freeman J. L. (1975), Strategies in the Multivariate Analysis of Data from Complex Surveys, *International Statistical Review*, 43: 59-78.

Särndal C. E., Swensson B., Wretman J. (1992), *Model Assisted Survey Sampling*, Springer-Verlag, New York.

APPENDIX

It is trivial to demonstrate that

$$E_p\left(\tilde{\mathbf{b}}\right) = \mathbf{b} = \left(b_1,...,b_4\right)', \tag{A.1}$$

being

$$b_1 = \sum_P y_{k,1}^2 - \frac{1}{N}\left[\sum_P y_{k,1}\right]^2 ; \; b_2 = \frac{1}{2}\sum_P \left(y_{k,1} - y_{k,2}\right)^2$$

$$b_3 = \frac{1}{2}\left[\sum_P y_{k,1} - \sum_P y_{k,2}\right]^2 ; \; b_4 = \sum_P y_{k,1} y_{k,2} - \frac{1}{N}\sum_P y_{k,1} \sum_P y_{k,2} ,$$

indeed $\tilde{b}_1$ is the Horvitz-Thompson estimate of the corresponding population parameter b_1; while the estimates $\left(\tilde{b}_2, \tilde{b}_3, \tilde{b}_4\right)$ correspond to the π^* estimators for two phase sample (Särndal *et al.*, 1992; pp. 347-350) of the corresponding population parameters $\left(b_2, b_3, b_4\right)$.

Consider now the following results.

$$E_m\left(b_1\right) = E_m\left[\sum_P y_{k,1}^2 - \frac{1}{N}\left[\sum_P y_{k,1}\right]^2\right] = \frac{N-1}{N}\sum_P E_m\left(y_{k,1}^2\right) - \frac{1}{N}\sum_k \sum_{l \neq k}{}_P \; E_m\left(y_{k,1}\, y_{l,1}\right) =$$

$$= \frac{N-1}{N}\sum_P \left(\sigma_{k,1}^2 + \mu_k^2\right) - \frac{1}{N}\sum_k \sum_{l \neq k}{}_P \; \left(\sigma_{kl,1} + \mu_k\, \mu_l\right) =$$

$$= \sum_P \mu_k^2 - \frac{1}{N}\left(\sum_P \mu_k\right)^2 + \frac{N-1}{N}\sum_P \sigma_{k,1}^2 - \frac{1}{N}\sum_k \sum_{l \neq k}{}_P \; \sigma_{kl,1} = X_1 + \frac{N-1}{N} X_2 - \frac{1}{N} X_3 . \tag{A.2}$$

$$E_m\left(b_2\right) = \frac{1}{2}\sum_P E_m\left(y_{k,1}^2\right) + E_m\left(y_{k,2}^2\right) - 2E_m\left(y_{k,1}\, y_{k,2}\right) = \frac{1}{2}\sum_P \sigma_k^2 + \mu_k^2 + \sigma_k^2 + \mu_k^2 - 2\mu_k^2 = X_2 . \tag{A.3}$$

$$E_m(b_3) = \frac{1}{2}\left[\sum_P E_m\left(y_{k,1}^2\right) + \sum_k \sum_{l \neq k}{}_P\, E_m\left(y_{k,1}\, y_{l,1}\right) + \sum_P E_m\left(y_{k,2}^2\right) + \sum_k \sum_{l \neq k}{}_P\, E_m\left(y_{k,2}\, y_{l,2}\right) + \right.$$

$$\left. -2\sum_k \sum_l{}_P\, E_m\left(y_{k,2}\, y_{l,2}\right)\right] =$$

$$= \frac{1}{2}\left[\sum_k{}_P\, \sigma_k^2 + \sum_k \sum_{l \neq k}{}_P\, \sigma_{kl,1} + \sum_k{}_P\, \sigma_k^2 + \sum_k \sum_{l \neq k}{}_P\, \sigma_{kl,2}\right] = X_2 + 0.5X_3 + 0.5X_4 \tag{A.4}$$

$$E_m(b_4) = \sum_P E_m\left(y_{k,1}\, y_{k,2}\right) - \frac{1}{N}\sum_P y_{k,1} \sum_P y_{k,2} = \sum_P \mu_k^2 - \frac{1}{N}\left[\sum_P \mu_k\right]^2 = X_1 . \tag{A.5}$$

From (A.1),…, (A.5) we derive the following result

$$E_m\left[E_p\left(\tilde{\mathbf{b}}\right)\right] = \mathbf{AX} . \tag{A.6}$$

From the above it is possible demonstrate that $\hat{\mathbf{X}}$, as defined by (6) is an unbiased estimator of the vector **X**, i.e.

$$E_m\left[E_p\left(\hat{\mathbf{X}}\right)\right] = E_m\left[E_p\left(\mathbf{A}^{-1}\tilde{\mathbf{b}}\right)\right] = \mathbf{A}^{-1} E_m\left[E_p\left(\tilde{\mathbf{b}}\right)\right] = \mathbf{A}^{-1}\,\mathbf{AX} = \mathbf{X} .$$