Supplementary material for "Double-calibration estimators accounting for under-coverage and non-response in socio-economic surveys"

A Appendix A - Results of Sections 3 and 4

A.1 An alternative formulation of the calibration estimator (3)

Because the unit constant is adopted as the first auxiliary variable, then $\boldsymbol{e}_1^t \boldsymbol{z}_j = 1, \forall j \in U$, where $\boldsymbol{e}_1^t = [1, 0, ..., 0]$ is the first vector of the standard basis of $\mathbb{R}^M$. Therefore, the difference between equations (3) and (4) is given by

$$
\begin{aligned}
\hat{T}_{Y(B)} - \hat{\boldsymbol{d}}_B^t \hat{\boldsymbol{T}}_{Z(B)} &= \sum_{j\in S} \frac{y_j}{\pi_j} - \hat{\boldsymbol{d}}_B^t \sum_{j\in S} \frac{\boldsymbol{z}_j}{\pi_j} = \\
&= \sum_{j\in S} \frac{y_j}{\pi_j} - \left(\sum_{j\in S} \frac{\boldsymbol{z}_j^t}{\pi_j} \right) \hat{\boldsymbol{d}}_B = \sum_{j\in S} \frac{y_j}{\pi_j} - \left(\sum_{j\in S} \frac{\boldsymbol{e}_1^t \boldsymbol{z}_j \boldsymbol{z}_j^t}{\pi_j} \right) \hat{\boldsymbol{d}}_B = \\
&= \sum_{j\in S} \frac{y_j}{\pi_j} - \boldsymbol{e}_1^t \left(\sum_{j\in S} \frac{\boldsymbol{z}_j \boldsymbol{z}_j^t}{\pi_j} \right) \hat{\boldsymbol{C}}_B^{-1} \hat{\boldsymbol{c}}_B = \sum_{j\in S} \frac{y_j}{\pi_j} - \boldsymbol{e}_1^t \hat{\boldsymbol{C}}_B \hat{\boldsymbol{C}}_B^{-1} \hat{\boldsymbol{c}}_B = \\
&= \sum_{j\in S} \frac{y_j}{\pi_j} - \boldsymbol{e}_1^t \hat{\boldsymbol{c}}_B = \sum_{j\in S} \frac{y_j}{\pi_j} - \boldsymbol{e}_1^t \sum_{j\in S} \frac{y_j \boldsymbol{z}_j}{\pi_j} = \sum_{j\in S} \frac{y_j}{\pi_j} - \sum_{j\in S} \frac{y_j \boldsymbol{e}_1^t \boldsymbol{z}_j}{\pi_j} = \\
&= \sum_{j\in S} \frac{y_j}{\pi_j} - \sum_{j\in S} \frac{y_j}{\pi_j} = 0
\end{aligned}
$$

Hence, (3) and (4) coincide.

A.2 First-order Taylor series approximation of $\hat{T}_{Y(dcal)}$

The double-calibration estimator (5) can be rewritten as

$$\hat{T}_{Y(dcal)} = f\left(\hat{\boldsymbol{a}}_R, \hat{\boldsymbol{A}}_R, \hat{\boldsymbol{c}}_R, \hat{\boldsymbol{C}}_R, \hat{\boldsymbol{T}}_{Z(B)}\right) = \hat{\boldsymbol{a}}_R^t \hat{\boldsymbol{A}}_R^{-1} \boldsymbol{T}_{X(B)} + \hat{\boldsymbol{c}}_R^t \hat{\boldsymbol{C}}_R^{-1} \left(\boldsymbol{T}_Z - \hat{\boldsymbol{T}}_{Z(B)}\right) = \sum_{k=1}^{K} \sum_{k\prime=1}^{K} \hat{a}_{Rk} T_{X(B)k\prime} \hat{a}_R^{kk\prime} + \sum_{m=1}^{M} \sum_{m\prime=1}^{M} \hat{c}_{Rm} (T_{Zm\prime} - \hat{T}_{Z(B)m\prime}) \hat{c}_R^{mm\prime} \quad \text{(A.1)}$$

where $\hat{a}_R^{kk\prime}$ denotes the kk' element of $\hat{\boldsymbol{A}}_R^{-1}$, $\hat{a}_{Rk}$ and $T_{X(B)k}$ are the k components of $\hat{\boldsymbol{a}}_R$ and $\boldsymbol{T}_{X(B)}$, respectively, $c_R^{mm\prime}$denotes the mm' element of $\hat{\boldsymbol{C}}_R^{-1}$, and $\hat{c}_{Rm}$, T_{Zm} and $\hat{T}_{Z(B)m}$ are the m components of $\hat{\boldsymbol{c}}_R$, $\boldsymbol{T}_Z$ and $\hat{\boldsymbol{T}}_{Z(B)}$, respectively. Differentiating A.1 with respect to all the variables involved, it follows that

$$\frac{\partial f}{\partial \hat{a}_{Rk}} = \sum_{k\prime=1}^{K} T_{X(B)k\prime} \hat{a}_R^{kk\prime}, k = 1, ..., K$$
$$\frac{\partial f}{\partial \hat{a}_{Rkk\prime}} = -\hat{\boldsymbol{a}}_R^t \hat{\boldsymbol{A}}_R^{-1} \boldsymbol{E}_{kk\prime} \hat{\boldsymbol{A}}_R^{-1} \boldsymbol{T}_{X(B)}, k, k' = 1, \ldots, M$$
$$\frac{\partial f}{\partial \hat{c}_{Rm}} = \sum_{m'=1}^{K} \left(T_{Zm'} - \hat{T}_{Z(B)m'}\right) \hat{c}_R^{mm'}, m = 1, \ldots, M$$
$$\frac{\partial f}{\partial \hat{a}_{Rmm'}} = -\hat{\boldsymbol{c}}_R^t \hat{\boldsymbol{C}}_R^{-1} \boldsymbol{E}_{mm'} \hat{\boldsymbol{C}}_R^{-1} \left(\boldsymbol{T}_Z - \hat{\boldsymbol{T}}_{Z(B)}\right), m, m' = 1, \ldots, M$$
$$\frac{\partial f}{\partial \hat{T}_{Z(B)m\prime}} = -\sum_{m=1}^{K} \hat{c}_{Rm} \hat{c}_R^{mm'}, m\prime = 1, \ldots, M$$

where $\boldsymbol{E}_{kk\prime}$ is the K-square matrix of 0s, with a 1 in position kk', and $\boldsymbol{E}_{mm\prime}$ is the M-square matrix of 0s, with a 1 in position mm'. Evaluating these partial derivatives at the expected points $\hat{\boldsymbol{a}}_R = \boldsymbol{a}_R$, $\hat{\boldsymbol{A}}_R = \boldsymbol{A}_R$, $\hat{\boldsymbol{c}}_R = \boldsymbol{c}_R$, $\hat{\boldsymbol{C}}_R = \boldsymbol{C}_R$ and $\hat{\boldsymbol{T}}_{Z(B)} = \boldsymbol{T}_{Z(B)}$, the first-order Taylor series approximation of $\hat{T}_{Y(dcal)}$ gives rise to

$$
\begin{aligned}
\hat{T}_{Y(dcal)} = \boldsymbol{a}_R^t \boldsymbol{A}_R^{-1} \boldsymbol{T}_{X(B)} + \boldsymbol{c}_R^t \boldsymbol{C}_R^{-1} (\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)}) + \sum_{k=1}^{K} \sum_{k'=1}^{K} (\hat{a}_{Rk} - a_{Rk}) T_{X(B)k'} a_R^{kk'} - \\
\sum_{k=1}^{K} \sum_{k'=1}^{K} \boldsymbol{a}_R^t \boldsymbol{A}_R^{-1} \boldsymbol{E}_{kk'} \boldsymbol{A}_R^{-1} \boldsymbol{T}_{X(B)} \left(\hat{a}_R^{kk'} - a_R^{kk'} \right) + \\
\sum_{m=1}^{M} \sum_{m'=1}^{M} (\hat{c}_{Rm} - c_{Rm})(T_{Zm'} - T_{Z(B)m'}) c_R^{mm'} - \\
\sum_{m=1}^{M} \sum_{m'=1}^{M} \boldsymbol{c}_R^t \boldsymbol{C}_R^{-1} \boldsymbol{E}_{mm'} \boldsymbol{C}_R^{-1} \left(\boldsymbol{T}_Z - \boldsymbol{T}_{X(B)} \right) \left(\hat{c}_R^{mm'} - c_R^{mm'} \right) - \\
\sum_{m=1}^{M} \sum_{m'=1}^{M} c_{Rm} c_R^{mm'} (\hat{T}_{Z(B)k} - T_{Z(B)k}) = \\
\boldsymbol{a}_R^t \boldsymbol{A}_R^{-1} \boldsymbol{T}_{X(B)} + \boldsymbol{c}_R^t \boldsymbol{C}_R^{-1} \left(\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)} \right) + (\hat{\boldsymbol{a}}_R - \boldsymbol{a}_R)^t \boldsymbol{A}_R^{-1} \boldsymbol{T}_{X(B)} - \\
\boldsymbol{a}_R^t \boldsymbol{A}_R^{-1} (\hat{\boldsymbol{A}}_R - \boldsymbol{A}_R) \boldsymbol{A}_R^{-1} \boldsymbol{T}_{X(B)} + (\hat{\boldsymbol{c}}_R - \boldsymbol{c}_R)^t \boldsymbol{C}_R^{-1} (\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)}) - \\
\boldsymbol{c}_R^t \boldsymbol{C}_R^{-1} (\hat{\boldsymbol{C}}_R - \boldsymbol{C}_R) \boldsymbol{C}_R^{-1} (\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)}) - \boldsymbol{c}_R^t \boldsymbol{C}_R^{-1} (\hat{\boldsymbol{T}}_{Z(B)} - \boldsymbol{T}_{Z(B)}) = \\
\boldsymbol{a}_R^t \boldsymbol{A}_R^{-1} \boldsymbol{T}_{X(B)} + \hat{\boldsymbol{a}}_R^t \boldsymbol{A}_R^{-1} \boldsymbol{T}_{X(B)} - \boldsymbol{a}_R^t \boldsymbol{A}_R^{-1} \hat{\boldsymbol{A}}_R \boldsymbol{A}_R^{-1} \boldsymbol{T}_{X(B)} + \\
\hat{\boldsymbol{c}}_R^t \boldsymbol{C}_R^{-1} (\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)}) - \boldsymbol{c}_R^t \boldsymbol{C}_R^{-1} \hat{\boldsymbol{C}}_R \boldsymbol{C}_R^{-1} (\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)}) + \\
\boldsymbol{c}_R^t \boldsymbol{C}_R^{-1} \boldsymbol{T}_Z - \boldsymbol{c}_R^t \boldsymbol{C}_R^{-1} \hat{\boldsymbol{T}}_{Z(B)} = c_1 + \hat{\boldsymbol{a}}_R^t \boldsymbol{A}_R^{-1} \boldsymbol{T}_{X(B)} - \boldsymbol{a}_R^t \boldsymbol{A}_R^{-1} \hat{\boldsymbol{A}}_R \boldsymbol{A}_R^{-1} \boldsymbol{T}_{X(B)} + \\
c_2 + \hat{\boldsymbol{c}}_R^t \boldsymbol{C}_R^{-1} (\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)}) - \boldsymbol{c}_R^t \boldsymbol{C}_R^{-1} \hat{\boldsymbol{C}}_R \boldsymbol{C}_R^{-1} (\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)}) - \boldsymbol{c}_R^t \boldsymbol{C}_R^{-1} \hat{\boldsymbol{T}}_{Z(B)}
\end{aligned}
\tag{A.2}
$$

Grouping the constant terms, equation A.2 can be rewritten as

$$\hat{T}_{Y(dcal)} = cost + \left(\sum_{j\in S}\frac{r_j y_j \boldsymbol{x}_j}{\pi_j}\right)^t \boldsymbol{A}_R^{-1}\boldsymbol{T}_{X(B)} -$$
$$\boldsymbol{a}_R^t\boldsymbol{A}_R^{-1}\left(\sum_{j\in S}\frac{r_j \boldsymbol{x}_j \boldsymbol{x}_j^t}{\pi_j}\right)\boldsymbol{A}_R^{-1}\boldsymbol{T}_{X(B)} + \left(\sum_{j\in S}\frac{r_j y_j \boldsymbol{z}_j}{\pi_j}\right)^t \boldsymbol{C}_R^{-1}(\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)}) -$$
$$\boldsymbol{c}_R^t\boldsymbol{C}_R^{-1}\left(\sum_{j\in S}\frac{r_j \boldsymbol{z}_j \boldsymbol{z}_j^t}{\pi_j}\right)\boldsymbol{C}_R^{-1}(\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)}) - \boldsymbol{c}_R^t\boldsymbol{C}_R^{-1}\left(\sum_{j\in S}\frac{\boldsymbol{z}_j}{\pi_j}\right) =$$
$$= cost + \sum_{j\in S}\frac{u_j}{\pi_j} \quad \text{(A.3)}$$

where

$$u_j = r_j(y_j\boldsymbol{x}_j^t - \boldsymbol{a}_R^t\boldsymbol{A}_R^{-1}\boldsymbol{x}_j\boldsymbol{x}_j^t)\boldsymbol{A}_R^{-1}\boldsymbol{T}_{X(B)} +$$
$$r_j(y_j\boldsymbol{z}_j^t - \boldsymbol{c}_R^t\boldsymbol{C}_R^{-1}\boldsymbol{z}_j\boldsymbol{z}_j^t)\boldsymbol{C}_R^{-1}\left(\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)}\right) - \boldsymbol{c}_R^t\boldsymbol{C}_R^{-1}\boldsymbol{z}_j, \qquad j \in U_B$$

are the influence values (e.g. **?**).

A.3 Approximate unbiasedness of the double calibration estimator

Regarding the first term of equation (6), **?** show that

$$\boldsymbol{b}_R^t\boldsymbol{T}_{X(B)} = T_{Y(B)} + (\boldsymbol{b}_N - \boldsymbol{b}_{NR})^t\,\boldsymbol{T}_{X(NR)}$$

where $\boldsymbol{T}_{X(NR)} = \sum_{j\in U_{B(NR)}}\boldsymbol{x}_j$. Therefore, if condition 1. holds, i.e. if $\boldsymbol{b}_R \approx \boldsymbol{b}_{NR}$, then equation (6) approximately reduces to

$$AE(\hat{T}_{Y(dcal)}) \approx T_{Y(B)} + \boldsymbol{d}_R^t(\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)})$$

in such a way that, if condition 2. holds, i.e. if $\boldsymbol{d}_R \approx \boldsymbol{d}_B$ then

$$AE(\hat{T}_{Y(dcal)}) \approx T_{Y(B)} + \boldsymbol{d}_B^t(\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)}) \quad \text{(A.4)}$$

Therefore, from A.4 it follows that

$$AE\left(\hat{T}_{Y(dcal)}\right) \approx \sum_{j\in U_B} y_j \ + \boldsymbol{d}_B^t \sum_{j\in U-U_B} \boldsymbol{z}_j = \sum_{j\in U_B} y_j \ +$$

$$\boldsymbol{d}_B^t \sum_{j\in U-U_B} \boldsymbol{z}_j + \sum_{j\in U-U_B} y_j - \sum_{j\in U-U_B} y_j = \sum_{j\in U} y_j \ + \boldsymbol{d}_B^t \sum_{j\in U-U_B} \boldsymbol{z}_j -$$

$$\sum_{j\in U-U_B} y_j = T_Y \ + \boldsymbol{d}_B^t \sum_{j\in U-U_B} \boldsymbol{z}_j - \sum_{j\in U-U_B} (\boldsymbol{d}_{NB}^t \boldsymbol{z}_j + e_{NBj})$$

where $\boldsymbol{d}_{NB}$ is the least-square coefficient vector of the regression of Y vs $\boldsymbol{Z}$ performed on the unsampled portion of the population $U - U_B$ and the $e_{NB}j$s are the regression residuals.

Because regression residuals sum to 0, then

$$AE\left(\hat{T}_{Y(dcal)}\right) \approx T_Y + \boldsymbol{d}_B^t \sum_{j\in U-U_B} \boldsymbol{z}_j - \boldsymbol{d}_{NB}^t \sum_{j\in U-U_B} \boldsymbol{z}_j =$$

$$T_Y + (\boldsymbol{d}_B - \boldsymbol{d}_{NB})^t (\boldsymbol{T}_Z - \boldsymbol{T}_{Z(B)}) \quad \text{(A.5)}$$

From equation A.5 it is proved that$\hat{T}_{Y(dcal)}$ is approximately unbiased if condition 3. holds, i.e if $\boldsymbol{d}_B \approx \boldsymbol{d}_{NB}$

B Appendix B - Results of simulation study presented in Section 5.1

Table B.1: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.3$ and $\rho_{ZY} = 0.6$. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	-0.9	75	-0.8	17.9	18.8 (11.5)	18.4	92.9
2250	-0.9	250	-0.8	9.7	9.8 (6.2)	9.8	95.1
2250	-0.9	500	-1.0	6.7	6.9 (4.4)	6.8	95.1
4500	-1.0	75	-1.0	12.2	12.5 (11.5)	12.5	94.1
4500	-1.0	250	-1.1	6.6	6.7 (6.2)	6.7	94.9
4500	-1.0	500	-1.1	4.6	4.7 (4.4)	4.7	95.0
6750	-1.4	75	-1.5	9.3	9.6 (11.5)	9.6	94.7
6750	-1.4	250	-1.4	5.1	5.3 (6.2)	5.1	94.3
6750	-1.4	500	-1.4	3.5	3.8 (4.4)	3.6	93.6

Table B.2: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.3$ and $\rho_{ZY} = 0.9$. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	0.6	75	0.6	13.9	14.7 (11.5)	14.2	94.1
2250	0.6	250	0.6	7.5	7.6 (6.2)	7.5	95.2
2250	0.6	500	0.5	5.2	5.3 (4.4)	5.2	95.2
4500	-0.9	75	-0.9	8.1	8.5 (11.5)	8.4	95.3
4500	-0.9	250	-0.9	4.4	4.5 (6.2)	4.5	95.3
4500	-0.9	500	-0.9	3.0	3.2 (4.4)	3.1	94.3
6750	-0.9	75	-1.0	4.8	5.0 (11.5)	5.0	95.2
6750	-0.9	250	-1.0	2.6	2.8 (6.2)	2.6	94.1
6750	-0.9	500	-0.9	1.8	2.0 (4.4)	1.8	92.6

Table B.3: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.6$ and $\rho_{ZY} = 0.3$. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	-1.8	75	-1.8	16.8	17.7 (11.5)	17.3	92.5
2250	-1.8	250	-1.8	9.1	9.4 (6.2)	9.3	94.6
2250	-1.8	500	-1.9	6.3	6.6 (4.4)	6.4	94.1
4500	-1.3	75	-1.3	11.8	12.1 (11.5)	12.1	94.2
4500	-1.3	250	-1.3	6.4	6.5 (6.2)	6.5	94.6
4500	-1.3	500	-1.3	4.4	4.6 (4.4)	4.5	94.7
6750	-1.6	75	-1.7	9.4	9.6 (11.5)	9.6	94.4
6750	-1.6	250	-1.7	5.1	5.4 (6.2)	5.2	93.9
6750	-1.6	500	-1.6	3.5	3.9 (4.4)	3.6	92.8

Table B.4: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.6$ and $\rho_{ZY} = 0.6$. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	-0.8	75	-1.7	9.4	9.6 (11.5)	9.6	94.4
2250	-0.8	250	-1.7	5.1	5.4 (6.2)	5.2	93.9
2250	-0.8	500	-1.6	3.5	3.9 (4.4)	3.6	92.8
4500	-1.3	75	-1.3	9.5	9.8 (11.5)	9.7	94.3
4500	-1.3	250	-1.3	5.1	5.3 (6.2)	5.2	94.7
4500	-1.3	500	-1.3	3.6	3.8 (4.4)	3.6	93.9
6750	-1.4	75	-1.5	7.0	7.3 (11.5)	7.2	94.7
6750	-1.4	250	-1.5	3.8	4.1 (6.2)	3.9	93.8
6750	-1.4	500	-1.4	2.6	3.0 (4.4)	2.7	91.8

Table B.5: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.6$ and $\rho_{ZY} = 0.9$. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	0.6	75	0.6	10.5	11.4 (11.5)	11.0	94.6
2250	0.6	250	0.6	5.7	5.8 (6.2)	5.8	95.3
2250	0.6	500	0.5	4.0	4.1 (4.4)	4.0	95.0
4500	-1.0	75	-1.0	5.7	6.1 (11.5)	6.2	95.8
4500	-1.0	250	-1.0	3.1	3.3 (6.2)	3.2	94.8
4500	-1.0	500	-1.0	2.2	2.4 (4.4)	2.2	93.2
6750	-0.8	75	-0.8	2.6	3.0 (11.5)	3.0	96.6
6750	-0.8	250	-0.8	1.4	1.7 (6.2)	1.5	92.7
6750	-0.8	500	-0.8	1.0	1.3 (4.4)	1.0	88.5

Table B.6: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.9$ and $\rho_{ZY} = 0.3$. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	-1.3	75	-1.3	9.3	10.1 (11.5)	9.7	93.9
2250	-1.3	250	-1.3	5.0	5.3 (6.2)	5.1	94.2
2250	-1.3	500	-1.4	3.5	3.8 (4.4)	3.6	93.3
4500	-1.4	75	-1.4	6.3	6.8 (11.5)	6.7	94.4
4500	-1.4	250	-1.4	3.4	3.7 (6.2)	3.5	93.7
4500	-1.4	500	-1.4	2.4	2.8 (4.4)	2.4	91.4
6750	-1.4	75	-1.4	4.9	5.2 (11.5)	5.2	94.5
6750	-1.4	250	-1.4	2.6	3.0 (6.2)	2.7	92.5
6750	-1.4	500	-1.4	1.8	2.3 (4.4)	1.9	88.4

Table B.7: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.9$ and $\rho_{ZY} = 0.6$. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	-0.7	75	-0.7	6.7	7.6 (11.5)	7.5	95.4
2250	-0.7	250	-0.7	3.6	3.8 (6.2)	3.8	95.1
2250	-0.7	500	-0.7	2.5	2.7 (4.4)	2.6	94.1
4500	-1.3	75	-1.4	4.1	4.7 (11.5)	4.7	94.9
4500	-1.3	250	-1.3	2.2	2.6 (6.2)	2.3	92.2
4500	-1.3	500	-1.3	1.5	2.1 (4.4)	1.6	86.8
6750	-1.9	75	-1.1	2.7	3.2 (11.5)	3.2	95.9
6750	-1.9	250	-1.2	1.5	1.9 (6.2)	1.6	89.2
6750	-1.9	500	-1.2	1.0	1.6 (4.4)	1.1	81.2

C Appendix C - Results of simulation study presented in Section 5.2

Table C.1: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.2$ in U_B, $\rho_{XY} = 0.3$ in the respondent stratum, $\rho_{ZY} = 0.6$ in $U - U_B$, and $\rho_{ZY} = 0.5$ in U_B. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	-1.3	75	-1.0	18.7	19.7 (11.6)	19.4	92.8
2250	-1.3	250	-1.2	10.3	10.4 (6.3)	10.4	94.6
2250	-1.3	500	-1.3	7.2	7.3 (4.4)	7.2	94.7
4500	0.0	75	0.2	12.9	13.3 (11.6)	13.0	94.13
4500	0.0	250	0.1	7.0	7.0 (6.3)	7.0	95.3
4500	0.0	500	0.0	4.9	4.9 (4.4)	4.9	95.2
6750	0.3	75	0.2	10.2	10.3 (11.6)	10.2	94.74
6750	0.3	250	0.3	5.5	5.5 (6.3)	5.5	95.2
6750	0.3	500	0.3	3.8	3.9 (4.4)	3.8	95.6

Table C.2: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.4$ in U_B, $\rho_{XY} = 0.3$ in the respondent stratum, $\rho_{ZY} = 0.6$ in $U - U_B$, and $\rho_{ZY} = 0.7$ in U_B. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	-0.7	75	-1.2	17.0	17.9 (11.6)	17.6	93.0
2250	-0.7	250	-0.8	9.2	9.4 (6.3)	9.3	94.5
2250	-0.7	500	-1.0	6.5	6.5 (4.4)	6.5	95.2
4500	0.1	75	0.1	11.2	11.3 (11.6)	11.3	94.9
4500	0.1	250	0.2	6.0	6.1 (6.3)	6.0	95.2
4500	0.1	500	0.2	4.2	4.2 (4.4)	4.2	95.5
6750	0.3	75	0.2	8.4	8.5 (11.6)	8.4	94.8
6750	0.3	250	0.4	4.5	4.6 (6.3)	4.5	95.7
6750	0.3	500	0.3	3.1	3.1 (4.4)	3.1	95.6

Table C.3: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.4$ in U_B, $\rho_{XY} = 0.3$ in the respondent stratum, $\rho_{ZY} = 0.9$ in $U - U_B$, and $\rho_{ZY} = 0.95$ in U_B. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	0.2	75	-0.9	13.7	13.9 (11.6)	13.9	94.5
2250	0.2	250	-0.8	7.3	7.4 (6.3)	7.3	94.9
2250	0.2	500	-0.8	5.1	5.1 (4.4)	5.1	95.0
4500	0.2	75	0.1	7.8	7.7 (11.6)	7.7	95.3
4500	0.2	250	0.1	4.1	4.1 (6.3)	4.1	95.2
4500	0.2	500	0.1	2.8	2.9 (4.4)	2.8	95.0
6750	0.1	75	0.1	4.2	4.1 (11.6)	4.2	95.9
6750	0.1	250	0.2	2.2	2.2 (6.3)	2.2	95.6
6750	0.1	500	0.1	1.5	1.5 (4.4)	1.5	95.3

Table C.4: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.5$ in U_B, $\rho_{XY} = 0.6$ in the respondent stratum, $\rho_{ZY} = 0.3$ in $U - U_B$, and $\rho_{ZY} = 0.2$ in U_B. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	-1.3	75	-1.6	17.2	18.1 (11.5)	17.7	92.7
2250	-1.3	250	-1.2	9.4	9.4 (6.3)	9.5	94.6
2250	-1.3	500	-1.3	6.6	6.6 (4.4)	6.6	94.6
4500	-0.2	75	-0.4	12.1	12.4 (11.5)	12.2	94.31
4500	-0.2	250	-0.3	6.5	6.5 (6.3)	6.6	95.4
4500	-0.2	500	-0.3	4.5	4.6 (4.4)	4.6	95.2
6750	0.1	75	0.1	9.8	10.0 (11.5)	9.9	94.8
6750	0.1	250	0.1	5.3	5.3 (6.3)	5.3	95.3
6750	0.1	500	0.1	3.7	3.7 (4.4)	3.7	95.3

Table C.5: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.7$ in U_B, $\rho_{XY} = 0.6$ in the respondent stratum, $\rho_{ZY} = 0.3$ in $U - U_B$, and $\rho_{ZY} = 0.4$ in U_B. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	-0.9	75	-0.7	16.2	17.3 (11.6)	16.6	92.5
2250	-0.9	250	-0.9	8.8	8.9 (6.3)	8.9	95.1
2250	-0.9	500	-0.9	6.2	6.2 (4.4)	6.2	95.2
4500	0.0	75	0.2	11.2	11.5 (11.6)	11.3	94.4
4500	0.0	250	0.1	6.1	6.0 (6.3)	6.1	95.1
4500	0.0	500	0.1	4.2	4.2 (4.4)	4.2	95.6
6750	0.2	75	0.2	8.9	9.0 (11.6)	8.9	94.9
6750	0.2	250	0.1	4.8	4.8 (6.3)	4.8	95.1
6750	0.2	500	0.2	3.3	3.3 (4.4)	3.3	95.2

Table C.6: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.5$ in U_B, $\rho_{XY} = 0.6$ in the respondent stratum, $\rho_{ZY} = 0.9$ in $U - U_B$, and $\rho_{ZY} = 0.8$ in U_B. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	-1.1	75	-1.3	11.6	11.9 (11.6)	11.7	94.7
2250	-1.1	250	-1.3	6.2	6.2 (6.3)	6.1	95.1
2250	-1.1	500	-1.3	4.4	4.4 (4.4)	4.2	94.5
4500	-0.6	75	-0.5	6.35	6.5 (11.6)	6.5	95.9
4500	-0.6	250	-0.4	3.4	3.4 (6.3)	3.4	95.5
4500	-0.6	500	-0.4	2.4	2.4 (4.4)	2.3	95.1
6750	-0.3	75	-0.3	3.4	3.2 (11.6)	3.3	97.4
6750	-0.3	250	-0.3	1.7	1.7 (6.3)	1.6	95.9
6750	-0.3	500	-0.3	1.2	1.2 (4.4)	1.1	95.2

Table C.7: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.7$ in U_B, $\rho_{XY} = 0.6$ in the respondent stratum, $\rho_{ZY} = 0.9$ in $U - U_B$, and $\rho_{ZY} = 0.95$ in U_B. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	0.7	75	-0.4	12.0	12.3 (11.6)	12.1	94.3
2250	0.7	250	-0.2	6.4	6.4 (6.3)	6.4	95.4
2250	0.7	500	-0.3	4.4	4.4 (4.4)	4.4	95.1
4500	0.2	75	0.2	7.5	7.5 (11.6)	7.5	95.1
4500	0.2	250	0.2	4.0	4.0 (6.3)	4.0	95.4
4500	0.2	500	0.2	2.8	2.8 (4.4)	2.7	95.6
6750	0.0	75	0.1	5.2	5.2 (11.6)	5.2	96.0
6750	0.0	250	0.1	2.7	2.7 (6.3)	2.7	95.7
6750	0.0	500	0.1	1.9	1.9 (4.4)	1.9	95.5

Table C.8: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.95$ in U_B, $\rho_{XY} = 0.9$ in the respondent stratum, $\rho_{ZY} = 0.3$ in $U - U_B$, and $\rho_{ZY} = 0.4$ in U_B. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	-0.1	75	-0.1	9.4	9.6 (11.7)	9.3	94.2
2250	-0.1	250	-0.2	4.9	4.9 (6.3)	4.9	95.2
2250	-0.1	500	-0.2	3.4	3.4 (4.4)	3.4	95.0
4500	0.1	75	0.1	6.3	6.3 (11.7)	6.3	95.3
4500	0.1	250	0.1	3.3	3.3 (6.3)	3.3	95.5
4500	0.1	500	0.1	2.3	2.3 (4.4)	2.3	95.2
6750	0.0	75	0.1	4.9	4.9 (11.7)	4.9	95.6
6750	0.0	250	0.1	2.5	2.5 (6.3)	2.5	95.4
6750	0.0	500	0.0	1.7	1.7 (4.4)	1.7	95.5

Table C.9: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.8$ in U_B, $\rho_{XY} = 0.9$ in the respondent stratum, $\rho_{ZY} = 0.6$ in $U - U_B$, and $\rho_{ZY} = 0.5$ in U_B. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	-1.3	75	-1.1	8.1	8.1 (11.7)	8.0	95.2
2250	-1.3	250	-1.2	4.2	4.2 (6.3)	4.1	94.9
2250	-1.3	500	-1.3	3.1	3.0 (4.4)	2.8	93.4
4500	-0.7	75	-0.5	4.9	4.8 (11.7)	4.9	96.3
4500	-0.7	250	-0.6	2.5	2.4 (6.3)	2.4	95.3
4500	-0.7	500	-0.6	1.8	2.8 (4.4)	1.7	93.9
6750	-0.3	75	-0.3	3.3	3.1 (11.7)	3.2	96.2
6750	-0.3	250	-0.2	1.6	1.6 (6.3)	1.6	95.8
6750	-0.3	500	-0.3	1.1	1.1 (4.4)	1.1	94.8

Table C.10: Percentage values of RB, $ARRMSE$, $RRMSE$, $ERRMSEE$, $COV95$ and first order approximation of relative bias (ARB) achieved from a population of 10000 units, a sampled sub-population of 7500 units with 2250, 4500 and 6750 respondent units, sample sizes $n = 75; 250; 500$ selected by means of simple random sampling without replacement. Correlations $\rho_{XY} = 0.95$ in U_B, $\rho_{XY} = 0.9$ in the respondent stratum, $\rho_{ZY} = 0.6$ in $U - U_B$, and $\rho_{ZY} = 0.7$ in U_B. Values in parentheses are the $RRMSE$s of the Horvitz-Thompson estimator in the absence of nonresponse and under-coverage.

$N_{B(R)}$	ARB	n	RB	$ARRSME$	$RRMSE$	$ERRMSEE$	$COV95$
2250	0.4	75	0.1	8.8	8.8 (11.7)	8.7	95.1
2250	0.4	250	0.0	4.5	4.5 (6.3)	4.5	95.3
2250	0.4	500	0.0	3.1	3.0 (4.4)	3.1	96.0
4500	0.2	75	0.3	6.4	6.2 (11.7)	6.3	95.6
4500	0.2	250	0.2	3.3	3.3 (6.3)	3.3	95.8
4500	0.2	500	0.2	2.3	2.3 (4.4)	2.3	95.7
6750	0.0	75	0.1	5.3	5.3 (11.7)	5.3	95.8
6750	0.0	250	0.1	2.8	2.8 (6.3)	2.8	95.4
6750	0.0	500	0.1	1.9	1.9 (4.4)	1.9	95.8