Konrad, Anne; Burgard, Jan Pablo; Münnich, Ralf



# A Two-level GREG Estimator for Consistent Estimation in Household Surveys

International Statistical Review

**Provided in Cooperation with:**
John Wiley & Sons

*Suggested Citation:* Konrad, Anne; Burgard, Jan Pablo; Münnich, Ralf (2021) : A Two-level GREG Estimator for Consistent Estimation in Household Surveys, International Statistical Review, ISSN 1751-5823, Wiley, Hoboken, NJ, Vol. 89, Iss. 3, pp. 635-656, https://doi.org/10.1111/insr.12460

This Version is available at:
https://hdl.handle.net/10419/284842

International Statistical Review

*International Statistical Review* (2021), 89, 3, 635–656 doi: 10.1111/insr.12460

# A Two-level GREG Estimator for Consistent Estimation in Household Surveys

**Anne Konrad, Jan Pablo Burgard and Ralf Münnich**

*Economic and Social Statistics Department, University of Trier, Trier, Germany*
*E-mail: konrada@uni-trier.de*

## Summary

**Household surveys provide information on both person-level and household-level characteristics. To ensure consistent estimates between both levels, statistical offices often use integrated weights that are equal for all persons within a household and the household itself. However, these integrated weights ignore the individual patterns of the persons, and the heterogeneity within a household is no longer reflected. As an alternative to integrated weighting, we propose a two-level generalised regression estimator that is capable of both ensuring consistent person and household estimates and allowing for different weights for persons within a household. A Monte Carlo simulation supports the superiority of our two-level generalised regression estimator compared with integrated weighting.**

*Key words*: Integrated weighting; household surveys; generalised regression estimator; calibration; consistent estimation.

## 1 Introduction

Household surveys are an important source of socio-economic and demographic data for scientific researchers as well as for political decision makers. The provided data allow for the estimation of both person and household characteristics. Estimation is characterised by assigning weights to the observed data to infer from the sample to the population (cf. Deville & Särndal, 1992, p. 376). These weights are frequently composed of the design weights and multiple adjustments. The design weight, as the inverse of the inclusion probability, accounts for the sampling design. To improve the efficiency of the resulting estimators, the design weights are often adjusted such that the final estimates of the auxiliary variables agree with their known population totals (cf. Särndal, 2007). A separate weight adjustment at the person and household levels does not necessarily ensure that the estimates of variable that are common to person and household data set coincide between both levels. However, survey users strive to obtain consistent estimates for common variables. Consistent estimates in this context are defined as obtaining the same estimates independent of using the person or the household data set. The vital role of consistency between person-level and household-level estimates is further emphasised as one principle in the European Statistics Code of Practice (cf. Eurostat, 2011, Principle 14).

The current practice to ensure consistent estimates in household surveys is integrated weighting originated by Lemaître & Dufour (1987). The method of integrated weighting produces one single weight for all persons within the same household. This single integrated person weight is then assigned one to one to the household to which the person belongs. Consistency is thereby ensured by using the same weights to estimate person and household characteristics. The strict requirement of equal weights does no longer reflect the heterogeneity of the individual persons within a household, and the individual patterns of the persons are lost. It is intuitive that for very volatile variables, such as income, the resulting estimates might be significantly influenced when the same weights are assigned to all persons within a household, independently of whether they are top earners, children or inactive persons.

As a remedy, we propose a two-level generalised regression (GREG) estimator that ensures consistent estimates while overcoming the strict requirement of equal weights for all persons within the same household and the household itself. The underlying idea is to constrain the consistency requirements to variables that are common to the person and household data set. For this purpose, we adopt the idea of incorporating the common variables as additional auxiliaries known from the literature on multiple independent surveys. In contrast to integrated weighting, our proposed two-level GREG estimator uses the original auxiliary information and herein allows for different weights for the persons within a certain household. As a result, consistency is ensured directly and solely for the relevant variables instead of indirectly by aggregating the individual information per household.

The paper is organised as follows: Section 2 briefly outlines the basic framework of cluster sampling and GREG estimators. Section 3 introduces integrated weighting and discusses consequences of the strict requirement of equal weights. In Section 4, we propose the two-level GREG estimator as alternative to integrated weighting. As a benchmark to our proposed estimator, we discuss in Section 5 the generalised least squares (GLS) adjustment algorithm, which can also be adopted to ensure consistent estimates. A Monte Carlo (MC) simulation study (Section 6) compares our proposed two-level GREG estimator with integrated weighting and the GLS adjustment algorithm. Section 7 contains concluding remarks.

## 2 Basic Framework

Household surveys are often realised by means of single-stage cluster sampling, characterised by sampling all units within a selected cluster. The sampling process of single-stage cluster sampling consists of two stages. At the first stage, from a finite population of households $U_h = \{1, \ldots, g, \ldots, M\}$, a sample $s_h$ is selected according to the sampling design $p(\cdot)$, where $p(s_h)$ is the probability of selecting $s_h$. The size of the sample $s_h$ is denoted by $m$. Subscript $h$ refers to the household level. Let $U_g$ be the population of persons within household $g$ of size $N_g$. The probability sampling design generates for every household a known inclusion probability $\pi_g = Pr(g \in s_h) = \sum_{s_h : g \in s_h} p(s_h)$ with $\pi_g > 0$. At the second stage, all persons within a selected household are sampled. The finite population and the sample of persons are denoted by $U_p = \cup_{g \in U_h} U_g = \{1, \ldots, i, \ldots, N\}$ and $s_p = \cup_{g \in s_h} U_g$, respectively. Subscript $p$ refers to the person level. The sample size is given by $n = \sum_{g \in s_h} N_g$. Because all persons within a selected household are sampled, it follows that $\pi_i = Pr(i \in s_p) = Pr(g \in s_h) = \pi_g$. Let $y_i$ be a variable of interest of person $i$. For simplicity, we assume full response. The objective is to estimate the unknown population total $T_{y,p} = \sum_{i \in U_p} y_i$ at the person level. A basic design unbiased estimator for the population total $T_{y,p}$ is the Horvitz–Thompson (HT) estimator given by $\hat{T}^{\mathrm{HT}}_{y,p} = \sum_{i \in s_p} \pi_i^{-1} y_i$ (cf. Horvitz & Thompson, 1952).

A widely used model-assisted estimator incorporating available auxiliary variables is the GREG estimator (cf. Cassel *et al*., 1997; Särndal, 1980; Isaki & Fuller, 1982, Wright, 1983). The vector containing the auxiliary variables of person $i$ is defined as $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{iq}, \ldots, x_{iQ})^T$ where $x_{i1} = 1$ determines the intercept. The corresponding total vector $\boldsymbol{T}_x = (T_{x_1}, \ldots, T_{x_q}, \ldots, T_{x_Q})^T$ of dimension $Q$ is known from censuses, registers or other reliable sources. The GREG estimator relies on a linear regression model that specifies the relationship between the variable of interest and the auxiliaries. The person-level regression model $\xi$ is given by

$$y_i = \boldsymbol{x}_i^T \boldsymbol{\beta} + \boldsymbol{\epsilon}_i \text{ for all } i \in U_p \tag{1}$$

with $E_\xi(\boldsymbol{\epsilon}_i) = 0$, $V_\xi(\boldsymbol{\epsilon}_i) = v_i\sigma^2$ and $\mathrm{E}_\xi(\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_j) = 0$ for all $i \neq j$. $E_\xi$ and $V_\xi$ denote the expectation and the variance with respect to the model $\xi$. The scale factor $v_i$, with $v_i > 0$, has to be known and describes the residual pattern. Based on the assisting model (1), the linear GREG estimator for the unknown population total $T_{y_p} = \sum_{i \in U_p} y_i$ is obtained from

$$\hat{T}_{y,p}^{\mathrm{GREG}} = \hat{T}_{y,p}^{\mathrm{HT}} + \hat{\boldsymbol{B}}^T (\boldsymbol{T}_x - \hat{\boldsymbol{T}}_x^{\mathrm{HT}}), \tag{2}$$

where $\hat{\boldsymbol{B}} = \left(\sum_{i \in s_p} \pi_i^{-1} \boldsymbol{x}_i \boldsymbol{x}_i^T\right)^{-1} \sum_{i \in s_p} \pi_i^{-1} \boldsymbol{x}_i y_i$ is an HT-type least squares estimate for $\boldsymbol{\beta}$. It is assumed that the matrix $\left(\sum_{i \in s_p} \pi_i^{-1} \boldsymbol{x}_i \boldsymbol{x}_i^T\right)^{-1}$ is nonsingular. According to expression (2), the GREG estimator can be interpreted as an HT estimator plus an adjustment term. This adjustment term is composed by the difference of the known and the estimated totals of the auxiliaries weighted by the magnitude of the relationship between the variable of interest and the auxiliary variables. Alternatively to expression (2), the GREG estimator can be written in linearly weighted form

$$\hat{T}_{y,p}^{\mathrm{GREG}} = \sum_{i \in s_p} w_i y_i$$

with

$$w_i^{\mathrm{GREG}} = \pi_i^{-1} + \pi_i^{-1} v_i^{-1} \boldsymbol{x}_i \left( \sum_{i \in s_p} \pi_i^{-1} v_i^{-1} \boldsymbol{x}_i \boldsymbol{x}_i^T \right)^{-1} \left(\boldsymbol{T}_x - \hat{\boldsymbol{T}}_x^{\mathrm{HT}}\right). \tag{3}$$

An important property of the GREG estimator is that the sums of the auxiliary variables weighted by (3) are consistent with the known totals, that is, $\sum_{i \in s_p} w_i^{\mathrm{GREG}} \boldsymbol{x}_i = \boldsymbol{T}_x$.

Based on the first-order Taylor linear approximation, the GREG estimator is approximately design unbiased under mild design conditions on the assisting model and on the sampling design (cf. Särndal, 2007, p. 103). The design unbiasedness of the GREG estimator does not depend on whether the population is really generated by the model $\xi$. The efficiency is, indeed, influenced by the predictive power of the model (cf. Särndal *et al*., 1992, p. 227, p. 239). The variance estimator of the GREG estimator under single-stage cluster sampling approximated by Taylor linearisation is given in terms of the residuals (cf. Särndal *et al*., 1992, p. 129, p. 235).

$$\hat{V}(\hat{T}_{y,p}^{\mathrm{GREG}}) = \sum_{g \in s_h} \sum_{k \in s_h} \frac{\triangle_{gk}}{\pi_{gk}} \sum_{i \in s_p} \sum_{j \in s_p} w_i^{\mathrm{GREG}} r_i w_j^{\mathrm{GREG}} r_j \tag{4}$$

with residuals $r_i = y_i - \boldsymbol{x}_i^T \hat{\boldsymbol{B}}$, $w_i^{\mathrm{GREG}}$ defined in (3) and $\triangle_{gk} := \pi_{gk} - \pi_g \pi_k$. Further correction methods can be applied in order to stabilise and improve the variance estimates (cf. D'Arrigo & Skinner, 2010; Kott, 2009).

The GREG estimator can be seen as a special case of a broader class of calibration estimators where the calibrated weights are chosen as close as possible to the original design weights (cf. Deville & Särndal, 1992; Deville *et al*., 1993). Closeness between both weights is measured via a pre-specified distance function. For the requirements on the distance, see, for example, in Haziza and Beaumont (2017, p. 213). The choice of the chi-squared distance function leads to the GREG weights defined in (3). Further distance functions are discussed in Deville & Särndal (1992), Deville *et al*. (1993), Huang & Fuller (1978), Alexander (1987), Singh & Mohl (1996) and Stukel *et al*. (1996). The calibration estimator generated by different distance functions asymptotically equals the GREG estimator (Deville & Särndal, 1992). Thus, for large sample sizes, the choice of the distance function has only a minor impact on the properties of the calibration estimator. Singh & Mohl (1996) and Stukel *et al*. (1996) extended this finding to modest sample sizes. A current review of calibration methods and distance functions can be found in Devaud & Tillé (2019).

## 3 Integrated Weighting

In the integrated weighting approach, introduced by Lemaître & Dufour (1987), consistency between person-level and household-level estimates is ensured by calculating weights at one level and then assigning these weights one to one to the respective other level. As a consequence thereof, it is not necessarily guaranteed that the weights at the person level sum up to the number of persons in the population and that simultaneously the weights at the household level sum up to the total number of households in the population. As it is an important property, we call this compliance at both levels of the sum of the weights with the population values as the *integrated property*. To ensure the integrated property, an additional variable has to be incorporated into the auxiliary variables $\boldsymbol{x}_i$. Therefore, we define at the person level

$$\boldsymbol{x}_i^{\circ} = (x_{i0}, x_{i1}, x_{i2}, \ldots, x_{iQ})^T = (N_g^{-1}, 1, x_{i2}, \ldots, x_{iQ})^T = (N_g^{-1}, \boldsymbol{x}_i)^T$$

as the integrated auxiliary vector of dimension $(Q+1)$, which sums up within each household $g$ to

$$\boldsymbol{x}_g^{\circ} = (x_{g0}, x_{g1}, x_{g2}, \ldots, x_{gQ})^T = (1, N_g, x_{g2}, \ldots, x_{gQ})^T.$$

Superscript $\circ$ indicates the integrated property. The corresponding known and estimated total vectors of dimension $(Q+1)$ are denoted by $\boldsymbol{T}_{x^{\circ}} = (M, \boldsymbol{T}_x)^T$ and $\hat{\boldsymbol{T}}_{x^{\circ}}^{\mathrm{HT}} = (\hat{T}_{x_0}^{\mathrm{HT}}, \hat{\boldsymbol{T}}_x^{\mathrm{HT}})^T$, respectively.

### 3.1 Point and Variance Estimator

To produce the same weights for all persons within the same household, the individual auxiliary variables are replaced by the household mean value. The household mean value is determined by $\bar{\boldsymbol{x}}_g^{\circ} = N_g^{-1} \sum_{i \in U_g} \boldsymbol{x}_i^{\circ}$, which is assigned to all persons within the household. At the

person level, the household mean value is denoted as $\bar{\boldsymbol{x}}_i^{\circ}$. By the substitution of $\bar{\boldsymbol{x}}_i^{\circ}$ as auxiliary variables in the assisting model (1), the integrated GREG estimator results expressed by

$$\hat{T}_{y,p}^{\mathrm{INT}} = \hat{T}_{y,p}^{\mathrm{HT}} + \hat{\boldsymbol{B}}^{\circ T}(\boldsymbol{T}_{x^{\circ}} - \hat{\boldsymbol{T}}_{x^{\circ}}^{\mathrm{HT}}) \tag{5}$$

with $\hat{\boldsymbol{B}}^{\circ} = \left(\sum_{i \in s_p} \pi_i^{-1} v_i^{-1} \bar{\boldsymbol{x}}_i^{\circ} \bar{\boldsymbol{x}}_i^{\circ T}\right)^{-1} \sum_{i \in s_p} \pi_i^{-1} v_i^{-1} \bar{\boldsymbol{x}}_i^{\circ} y_i$. The corresponding integrated person weights are given by

$$w_i^{\mathrm{INT}} = \frac{1}{\pi_i} + \frac{\bar{\boldsymbol{x}}_i^{\circ T}}{\pi_i v_i}\left(\sum_{i \in s_p} \frac{\bar{\boldsymbol{x}}_i^{\circ} \bar{\boldsymbol{x}}_i^{\circ T}}{\pi_g v_g}\right)^{-1} (\boldsymbol{T}_{x^{\circ}} - \hat{\boldsymbol{T}}_{x^{\circ}}^{\mathrm{HT}}), \tag{6}$$

which are equivalent to the household weights

$$\begin{aligned} &= \frac{1}{\pi_g} + \frac{\bar{\boldsymbol{x}}_i^{\circ T}}{\pi_g v_g}\left(\sum_{g \in s_h} \frac{\bar{\boldsymbol{x}}_i^{\circ} \bar{\boldsymbol{x}}_i^{\circ T}}{\pi_g v_g}\right)^{-1} (\boldsymbol{T}_{x^{\circ}} - \hat{\boldsymbol{T}}_{x^{\circ}}^{\mathrm{HT}}) \\ &= w_g^{\mathrm{INT}}. \end{aligned}$$

Note that $v_g = \sum_{i \in U_g} v_i$. Inserting $v_i = 1$ into (6) results in the integrated GREG estimator introduced by Lemaître & Dufour (1987). Inserting $v_i = N_g^{-1}$, in turn, gives the integrated GREG estimator proposed by Nieuwenbroek (1993). The choice $v_i = N_g^{-1}$ can be interpreted that the variance of the variable of interest decreases with the household size (Nieuwenbroek, 1993, p. 9).

As the integrated GREG estimator is a special case of (2), it is also design unbiased. Its variance estimator is approximated by Taylor linearisation that is given by

$$\hat{V}(\hat{T}_{y,p}^{\mathrm{INT}}) = \sum_{g \in s_h} \sum_{k \in s_h} \frac{\triangle_{gk}}{\pi_{gk}} w_g^{\mathrm{INT}} r_g^{\mathrm{INT}} w_k^{\mathrm{INT}} r_k^{\mathrm{INT}} \tag{7}$$

with residuals $r_g^{\mathrm{INT}} = y_g - \boldsymbol{x}_g^{\circ T} \hat{\boldsymbol{B}}^{\circ}$.

As an alternative to integrated weighting, Zieschang (1986), Luery (1986) and Alexander (1987) discussed a calibration estimator using the chi-squared distance function to produce household weights. When these weights are applied for both person and household characteristics, the calibration weights are asymptotically equivalent to the integrated weights defined in (6) using $v_i = N_g^{-1}$. Verma & Clémenceau (1996) proposed to extend the household-level auxiliary variables in the calibration estimator by the person-level information. For this purpose, the sample distribution of the person auxiliaries is inflated by the household size. Branson & Wittenberg (2014) suggested a minimum cross-entropy approach to produce integrated weights. Wittenberg (2010) showed that the minimum cross-entropy approach is equivalent to the raking estimator introduced by Deming & Stephan (1940) and to the calibration estimator with a multiplicative distance function. Compared with the integrated weights defined in (6), the minimum cross-entropy weights are prevented from being negative. Isaki *et al*. (2004) used quadratic programming to produce household weights suitable for the estimation of person and household characteristics. Park & Fuller (2005, p. 8) showed that quadratic programming is equal to the calibration estimator with a truncated linear distance function. When dropping the bounds in the distance function, quadratic programming generates weights that are asymptotically equivalent to the integrated weights defined in (6) with $v_i = N_g^{-1}$. Neethling & Galpin (2006) empirically compared a calibration estimator with integrated weighting using a linear and a

multiplicative distance function. Boonstra *et al*. (2003) and Houbiers (2004) introduced repeated weighting to enforce consistency among contingency tables of survey estimates. We do not pursue repeated weighting because we focus on producing global weights that are suitable for all survey variables.

### *3.2 Consequences of Integrated Weighting*

The requirement of equal weights for all person within and the household itself causes various consequences. First, due to the aggregation of the individual auxiliary variables to constructed household mean values, the outcome values of the original auxiliary variables are redistributed within a household. In explanation, suppose the original auxiliary $x_i$ indicates whether the person is male or female. Whereas the original auxiliary has only two possible outcome values, 0 or 1, the number of possible values in the integrated approach increases with the number of household members: for a single-person household, it is 0 or 1; for a two-person household, 0, 1/2 or 1; for a three-person household, 0, 1/3, 2/3 or 1; for a four-person household, 0, 1/4, 2/4 or 3/4; and so on. The increased number of outcome values in the integrated weighting approach may enlarge the variation of the integrated weights.

Second, through the substitution of the original auxiliary variables by the household mean values, the variances and covariances of the integrated auxiliaries only capture the variances and covariances between the households. Ignoring the within variance and within covariance implies that the heterogeneity of the persons within a household is not taken into account. Thus, the integrated weighting approach does not exploit all available sample information when estimating the coefficient vector in the estimator. The lower the variation of the auxiliaries, the less stable the projection onto the space spanned by the auxiliaries and the higher the variance of the integrated coefficients. Therefore, we expect that the integrated coefficients vary more compared with the coefficients resulting from an ordinary GREG estimator defined in (2), which, in turn, decreases the efficiency of the final estimates. This argumentation contradicts Steel & Clark (2007) who stated that the correlation within households is not relevant to estimate population totals in single-stage cluster sampling because sampled and non-sampled persons live in distinct households. In the simulation study, we examine which argument prevails.

Third, the one-to-one weight assignment in the integrated approach tacitly assumes that the strength of the relationship between the variable of interest and the auxiliary variables is identical at both levels. However, Robinson (1950) showed that the correlations for the same variables can be different at the individual level than at the aggregated level. This phenomenon is known as ecological fallacy. In case of ecological fallacy, the coefficients at both levels differ, and in the extreme case, also their signs differ. The difference between the coefficients at both levels affects the efficiency of the integrated estimates because the GREG estimator is model assisted. Thus, even if its approximately design unbiasedness does not depend on the correctness of the model, its efficiency, in turn, relies on the strength of the relationship between the variable of interest and the auxiliaries.

Fourth, the integrated GREG estimator requires $x_{i0}$ as additional auxiliary variable to ensure the integrated property. The simulation study in Section 6 validates whether these consequences of integrated weighting affect the quality of the final estimates.

## 4 Proposed Two-level Generalised Regression Estimator

As an alternative to integrated weighting, we propose a two-level GREG estimator that is capable of both ensuring consistent person and household estimates and allowing for different weights for persons within a household.

### 4.1 Core Idea

The idea underlying our proposed two-level GREG estimator is to constrain the consistency requirements to variables that are common to the person and household data set. By incorporating the common variables as additional auxiliaries, our two-level GREG estimator produces consistent estimates of these variables. Thus, consistency is ensured more directly and only for the relevant variables, instead of indirectly by aggregating the individual information per household as it is done in the integrated weighting approach. To implement the two-level GREG estimator, we adapt the method given by Renssen & Nieuwenbroek (1997), known from the literature on combining information from multiple independent surveys. However, there are considerable differences between multiple independent surveys and household surveys in terms of the definition of common variables, the dependence of the surveys and differing target populations.

The main advantage of our two-level GREG estimator compared with integrated weighting is that the original individual rather than the constructed aggregated auxiliaries are utilised. Therefore, the individual pattern of the persons is retained. Furthermore, our two-level GREG estimator consists of separate person-level and household-level estimators, providing two further advantages. First, the different calculation levels of person and household characteristics, which prevent the problems caused by ecological fallacy, are considered. Second, the variable selection process is more flexible because different auxiliary variables can be incorporated in the person-level estimator than in the household-level estimator. Finally, no additional auxiliary variable is required to enforce the integrated property.

We propose two different two-level GREG estimators: a naïve and an extended two-level GREG estimator. The difference is given by the estimator for the unknown common variable totals. Whereas the naïve approach is easier to implement, because only the household-level estimator has to be adjusted by the common variables, the extended two-level GREG estimator enables to insert the best available estimate of the unknown common variable totals.

### 4.2 Naïve Two-level Generalised Regression Estimator

In the naïve approach, consistency is ensured by incorporating the common variables in the household-level estimator. The estimator at the person level, in turn, remains unaffected by the consistency requirements between the levels. Let $\boldsymbol{c}_i = (c_{i1}, \ldots, c_{il}, \ldots, c_{iL})^T$ be the person-level common variable vector of dimension $L$, which sums up per household to $\sum_{i \in U_g} \boldsymbol{c}_i = \boldsymbol{c}_g = (c_{g1}, \ldots, c_{gl}, \ldots, c_{gL})^T$. The totals of the common variables are unknown and have to be estimated by $\tilde{\boldsymbol{T}}_c$. The subscript indicating the level of estimation is skipped, because due to the consistency requirements, $\tilde{\boldsymbol{T}}_c$ must be equal at the person and household levels.

At the person level, the naïve two-level GREG estimator (abbreviated with TL1) is given by an ordinary GREG estimator as defined in (2); that is,

$$\hat{T}^{\mathrm{TL1}}_{y,p} = \hat{T}^{\mathrm{GREG}}_{y,p} = \hat{T}^{\mathrm{HT}}_{y,p} + \hat{\boldsymbol{B}}^T (\boldsymbol{T}_x - \hat{\boldsymbol{T}}^{\mathrm{HT}}_x). \tag{8}$$

Accordingly, $\hat{T}^{\mathrm{TL1}}_{y,p}$ contains only the auxiliary variables $\boldsymbol{x}_i$, but not the common variables. The corresponding weights $w_i^{\mathrm{TL1}} = w_i^{\mathrm{GREG}}$ are obtained from (3). For simplicity, we assume $v_i = 1$.

At the household level, a separate estimator is implemented. To emphasise that different auxiliary variables can be included at both levels, we denote $\boldsymbol{a}_g = (a_{g1}, \ldots, a_{gk}, \ldots, a_{gK})^T$ as the $K$ dimensional auxiliary vector of household $g$. The known vector of the totals is given by $\boldsymbol{T}_a$. To ensure consistency between person-level and household-level estimates, the common variables

have to be included into the estimator as additional auxiliaries. To estimate the unknown vector of the common variable totals, we suggest the proposed person-level estimator (8) applied to the common variables, because the common variables are originally person characteristics. Thus,

$$\tilde{\boldsymbol{T}}_c = \hat{\boldsymbol{T}}_{c,p}^{\mathrm{TL1}} = \hat{\boldsymbol{T}}_{c,p}^{\mathrm{GREG}} = \sum_{i \in s_p} w_i^{\mathrm{GREG}} \boldsymbol{c}_i. \tag{9}$$

Given (9), the proposed two-level GREG estimator at the household level is obtained from

$$\hat{T}_{y,h}^{\mathrm{TL1}} = \hat{T}_{y,h}^{\mathrm{HT}} + \hat{\boldsymbol{E}}_a^T \left(\boldsymbol{T}_a - \hat{\boldsymbol{T}}_a^{\mathrm{HT}}\right) + \hat{\boldsymbol{E}}_c^T \left(\hat{\boldsymbol{T}}_{c,p}^{\mathrm{GREG}} - \hat{\boldsymbol{T}}_{c,p}^{\mathrm{HT}}\right), \tag{10}$$

where the coefficients $\hat{\boldsymbol{E}}_a$ and $\hat{\boldsymbol{E}}_c$ are simultaneously estimated by

$$\begin{pmatrix} \hat{\boldsymbol{E}}_a \\ \hat{\boldsymbol{E}}_c \end{pmatrix} = \sum_{g \in s_h} \left[ \pi_g^{-1} \begin{pmatrix} \boldsymbol{a}_g \\ \boldsymbol{c}_g \end{pmatrix} \begin{pmatrix} \boldsymbol{a}_g \\ \boldsymbol{c}_g \end{pmatrix}^T \right]^{-1} \sum_{g \in s_h} \pi_g^{-1} \begin{pmatrix} \boldsymbol{a}_g \\ \boldsymbol{c}_g \end{pmatrix} y_g. \tag{11}$$

It is assumed that the matrix $\sum_{g \in s_h} \pi_g^{-1} \begin{pmatrix} \boldsymbol{a}_g \\ \boldsymbol{c}_g \end{pmatrix} \begin{pmatrix} \boldsymbol{a}_g \\ \boldsymbol{c}_g \end{pmatrix}^T$ is of full rank $K+L$.

To quantify the impact of ensuring consistency on our proposed estimator at the household level, we decompose (10) into an ordinary GREG estimator, which does not ensure consistent estimates, and an adjustment term capturing the effect caused by including the common variables as additional auxiliaries. For the decomposition, an orthogonal decomposition of the coefficients (cf. Seber, 1977) is applied, given by

$$\hat{\boldsymbol{E}}_a = \hat{\boldsymbol{B}}_a - \hat{\boldsymbol{F}}_a \hat{\boldsymbol{E}}_c, \tag{12}$$

where $\hat{\boldsymbol{B}}_a$ arises from

$$\hat{T}_{y,h}^{\mathrm{GREG}} = \hat{T}_{y,h}^{\mathrm{HT}} + \hat{\boldsymbol{B}}_a^{\mathrm{T}} (\boldsymbol{T}_a - \hat{\boldsymbol{T}}_a^{\mathrm{HT}}) \tag{13}$$

as an ordinary GREG estimator solely containing $\boldsymbol{a}_g$ as auxiliaries. Coefficient matrix $\hat{\boldsymbol{F}}_a$ is obtained from $\hat{\boldsymbol{T}}_{c,h}^{\mathrm{GREG}} = \hat{\boldsymbol{T}}_{c,h}^{\mathrm{HT}} + \hat{\boldsymbol{F}}_a^T \left(\boldsymbol{T}_a - \hat{\boldsymbol{T}}_a^{\mathrm{HT}}\right)$. Coefficient vector $\hat{\boldsymbol{E}}_c$ is defined in (11). Hence, the product of $\hat{\boldsymbol{F}}_a$ and $\hat{\boldsymbol{E}}_c$ captures the effects of the common variables on the variable of interest neglected by $\hat{\boldsymbol{B}}_a$.

Inserting the orthogonal decomposition (12) into (10), we obtain

$$\hat{T}_{y,h}^{\mathrm{TL1}} = \hat{T}_{y,h}^{\mathrm{GREG}} + \hat{\boldsymbol{E}}_c^T \left(\hat{\boldsymbol{T}}_{c,p}^{\mathrm{GREG}} - \hat{\boldsymbol{T}}_{c,h}^{\mathrm{GREG}}\right). \tag{14}$$

According to (14), the adjustment term capturing the impact induced by the consistency requirements depends on $\hat{\boldsymbol{E}}_c$ and the difference between the person-level and household-level estimates for the common variable totals. The greater the difference between the two estimates, the greater the adjustment term.

The partial coefficient $\hat{\boldsymbol{E}}_c$ can, alternatively to (11), be expressed in terms of residuals as

$$\hat{\boldsymbol{E}}_c = \left( \sum_{g \in s_h} \boldsymbol{r}_g^{F_a} \boldsymbol{r}_g^{F_a T} \right)^{-1} \sum_{g \in s_h} \boldsymbol{r}_g^{F_a} r_g^{B_a} \tag{15}$$

with $r_g^{B_a} = y_g - \hat{\boldsymbol{B}}_a^T \boldsymbol{a}_g$ and $\boldsymbol{r}_g^{F_a} = \boldsymbol{c}_g - \hat{\boldsymbol{F}}_a^T \boldsymbol{a}_g$ resulting from regressing the variable of interest

or common variables on the auxiliaries. This expression of $\hat{\boldsymbol{E}}_c$ simplifies the expression of the household weights given by

$$w_g^{\mathrm{TL1}} = w_g^{\mathrm{GREG}} + \boldsymbol{r}_g^{F_a T}\left(\sum_{g \in s_h} \boldsymbol{r}_g^{F_a}\boldsymbol{r}_g^{F_a T}\right)^{-1}\left(\hat{\boldsymbol{T}}_{c_p}^{\mathrm{GREG}} - \hat{\boldsymbol{T}}_{c_h}^{\mathrm{GREG}}\right) \tag{16}$$

with $w_g^{\mathrm{GREG}}$ obtained from (13).

The variance estimator of $\hat{T}_{y,p}^{\mathrm{TL1}}$ is determined by (4). At the household level, the variance estimator should account for the additional source of randomness induced by inserting the estimated common variable totals instead of known population totals. Using the Taylor linearisation technique, we get

$$\hat{V}(\hat{T}_{y,h}^{\mathrm{TL1}}) \doteq \hat{V}_1 + \hat{V}_2 + \hat{V}_3 + 2\cdot\hat{V}_{12} - 2\cdot\hat{\mathrm{V}}_{13} - 2\cdot\hat{\mathrm{V}}_{23} \tag{17}$$

with

$$\begin{aligned}
\hat{V}_1 &= \hat{V}(\hat{T}_{y,h}^{\mathrm{GREG}}), & \hat{\mathrm{V}}_{12} &= \hat{\boldsymbol{E}}_c^T\widehat{\mathrm{Cov}}(\hat{T}_{y,h}^{\mathrm{GREG}}, \hat{\boldsymbol{T}}_{c,p}^{\mathrm{GREG}}),\\
\hat{V}_2 &= \hat{\boldsymbol{E}}_c^T\hat{V}(\hat{\boldsymbol{T}}_{c,p}^{\mathrm{GREG}})\hat{\boldsymbol{E}}_c, & \hat{\mathrm{V}}_{13} &= \hat{\boldsymbol{E}}_c^T\widehat{\mathrm{Cov}}(\hat{T}_{y,h}^{\mathrm{GREG}}, \hat{\boldsymbol{T}}_{c_h}^{\mathrm{GREG}}),\\
\hat{V}_3 &= \hat{\boldsymbol{E}}_c^T\hat{V}(\hat{\boldsymbol{T}}_{c,h}^{\mathrm{GREG}})\hat{\boldsymbol{E}}_c, & \hat{\mathrm{V}}_{23} &= \hat{\boldsymbol{E}}_c^T\widehat{\mathrm{Cov}}(\hat{\boldsymbol{T}}_{c,p}^{\mathrm{GREG}}, \hat{\boldsymbol{T}}_{c,h}^{\mathrm{GREG}})\hat{\boldsymbol{E}}_c,
\end{aligned}$$

where $\widehat{\mathrm{Cov}}$ denotes the estimated covariance. Estimated variances and covariances can be obtained from (4) by inserting the appropriate residuals.

The additional effort to compute $\hat{V}(\hat{T}_{y,h}^{\mathrm{TL1}})$ compared with the variance of an ordinary GREG estimator, determined by $\hat{V}_1$, is limited to the calculation of the $L$-dimensional residual vectors $\boldsymbol{r}_g^{F_x} = \sum_{i \in U_g}\boldsymbol{r}_i^{F_x} = \sum_{i \in U_g}(\boldsymbol{c}_i - \boldsymbol{F}_x^T\boldsymbol{x}_i)$ and $\boldsymbol{r}_g^{F_a} = \boldsymbol{c}_g - \boldsymbol{F}_a^T\boldsymbol{x}_g$. Both residual vectors are independent from the variable of interest and thus have to be calculated only once per sample. Variance components $\hat{V}_2$, $\hat{V}_3$, $\hat{V}_{12}$, $\hat{V}_{23}$ and $\hat{V}_{23}$ are computable by an appropriate combination of these residuals. Hence, the additional computational effort depends on the number of variables that are required to be consistent.

### *4.3 Extended Two-level Generalised Regression Estimator*

In the extended approach, we strive to improve the estimates of the unknown common variable totals $\tilde{\boldsymbol{T}}_c$. The underlying idea is that every common variable $c_l$, with $l = 1, \ldots, L$, can be modelled by a separate set of specialised auxiliary variables $\boldsymbol{z}_l$. The specialised auxiliary variable set $\boldsymbol{z}_l$ could be, for example, the auxiliary variable set with the highest explanatory power for the respective common variable $c_l$. Then the unknown totals of the common variables are estimated by

$$\tilde{\boldsymbol{T}}_c = \hat{\boldsymbol{T}}_{c_p^*}^{\mathrm{GREG}} = (\hat{T}_{c_{p,1}^*}^{\mathrm{GREG}}, \ldots, \hat{T}_{c_{p,l}^*}^{\mathrm{GREG}}, \ldots, \hat{T}_{c_{p,L}^*}^{\mathrm{GREG}})^T, \tag{18}$$

where

$$\hat{\boldsymbol{T}}_{c_{p,1}^*}^{\mathrm{GREG}} = \hat{\boldsymbol{T}}_{c_{p,l}}^{\mathrm{HT}} + \hat{\boldsymbol{B}}_{z_l}(\boldsymbol{T}_{z_l} - \hat{\boldsymbol{T}}_{z_l}^{\mathrm{HT}}) \text{for} l = 1, \ldots, L$$

is the estimator for the total of the common variable $c_l$ based on the specialised auxiliary variable set $\boldsymbol{z}_l$ and with $\hat{\boldsymbol{B}}_{z_l} = \left(\sum_{i \in s_p}\pi_i^{-1}\boldsymbol{z}_{il}\boldsymbol{z}_{il}^T\right)^{-1}\sum_{i \in s_p}\pi_i^{-1}\boldsymbol{z}_{il}\boldsymbol{c}_{il}$.

To ensure consistency between person-level and household-level estimates, the same estimates for the unknown common variable totals resulting from (18) have to be inserted into the estimators at both levels. At the person level, the estimator of our extended two-level GREG estimator (abbreviated with TL2) is given by

$$\hat{T}_{y,p}^{\mathrm{TL2}} = \hat{T}_{y,p}^{\mathrm{HT}} + \hat{\boldsymbol{D}}_x^{\,T}(\boldsymbol{T}_x - \hat{\boldsymbol{T}}_x^{\mathrm{HT}}) + \hat{\boldsymbol{D}}_c^{\,T}(\hat{\boldsymbol{T}}_{c*,p}^{\mathrm{GREG}} - \hat{\boldsymbol{T}}_{c,p}^{\mathrm{HT}}), \tag{19}$$

where the coefficients $\hat{\boldsymbol{D}}_x$ and $\hat{\boldsymbol{D}}_c$ are simultaneously estimated by

$$\begin{pmatrix} \hat{\boldsymbol{D}}_x \\ \hat{\boldsymbol{D}}_c \end{pmatrix} = \left[ \sum_{i \in s_p} \pi_i^{-1} \begin{pmatrix} \boldsymbol{x}_i \\ \boldsymbol{c}_i \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_i \\ \boldsymbol{c}_i \end{pmatrix}^T \right]^{-1} \sum_{i \in s_p} \pi_i^{-1} \begin{pmatrix} \boldsymbol{x}_i \\ \boldsymbol{c}_i \end{pmatrix} y_i. \tag{20}$$

It is assumed that the matrix $\sum_{i \in s_p} \pi_i^{-1} \begin{pmatrix} \boldsymbol{x}_i \\ \boldsymbol{c}_i \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_i \\ \boldsymbol{c}_i \end{pmatrix}^T$ is of full rank $Q + L$.

To quantify the impact of the consistency requirements, we once more decompose $\hat{T}_{y,p}^{\mathrm{TL2}}$ into an ordinary GREG estimator and an adjustment term capturing the effect caused by ensuring consistency. By inserting the orthogonal decomposition $\hat{\boldsymbol{D}}_x = \hat{\boldsymbol{B}}_x - \hat{\boldsymbol{F}}_x \hat{\boldsymbol{D}}_c$ into (19), we obtain

$$\hat{T}_{y,p}^{\mathrm{TL2}} = \hat{T}_{y,p}^{\mathrm{GREG}} + \hat{\boldsymbol{D}}_c^{\mathrm{T}} \left( \hat{\boldsymbol{T}}_{c*,p}^{\mathrm{GREG}} - \hat{\boldsymbol{T}}_{c,p}^{\mathrm{GREG}} \right), \tag{21}$$

where

$$\hat{\boldsymbol{D}}_c = \left( \sum_{i \in s_p} \boldsymbol{r}_i^{F_x} \boldsymbol{r}_i^{F_x\,T} \right)^{-1} \sum_{i \in s_p} \boldsymbol{r}_i^{F_x} r_i^{B_x} \tag{22}$$

with residuals $\boldsymbol{r}_i^{F_x} = \boldsymbol{c}_i - \hat{\boldsymbol{F}}_x^{\mathrm{T}} \boldsymbol{x}_i$ and $r_i^{B_x} = y_i - \hat{\boldsymbol{B}}_x^{\mathrm{T}} \boldsymbol{x}_i$ resulting from regressing the common variables and the variable of interest on the auxiliaries. Coefficient matrix $\hat{\boldsymbol{F}}_x$ is obtained from $\hat{\boldsymbol{T}}_{c,p}^{\mathrm{GREG}} = \hat{\boldsymbol{T}}_{c,p}^{\mathrm{HT}} + \hat{\boldsymbol{F}}_x^{\mathrm{T}} (\boldsymbol{T}_x - \hat{\boldsymbol{T}}_x^{\mathrm{HT}})$, a vector containing the person-level GREG estimators for the common variable totals with $\boldsymbol{x}_i$ as auxiliaries.

At the household level, a separate estimator is implemented, which is obtained from

$$\hat{T}_{y,h}^{\mathrm{TL2}} = \hat{T}_{y,h}^{\mathrm{GREG}} + \hat{\boldsymbol{E}}_c^{\,T} \left( \hat{\boldsymbol{T}}_{c*,p}^{\mathrm{GREG}} - \hat{\boldsymbol{T}}_{c,h}^{\mathrm{GREG}} \right) \tag{23}$$

with $\hat{\boldsymbol{E}}_c$ already defined in (15). We learn from (21) and (23) that the higher the difference between the estimated common variable totals, utilising a specialised auxiliary sets $\boldsymbol{z}_{il}$ compared with $\boldsymbol{x}_i$ or $\boldsymbol{a}_g$, the higher the adjustment term. For the special case of $\boldsymbol{z}_{il} = \boldsymbol{z}_i = \boldsymbol{x}_i$ for all $l = 1, \ldots, L$, the estimators of our naïve approach coincide with the estimators of our extended approach.

The corresponding weights are obtained from

$$w_i^{\mathrm{TL2}} = w_i^{\mathrm{GREG}} + \boldsymbol{r}_i^{F_x\,T} \left( \sum_{i \in s_p} \boldsymbol{r}_i^{F_x} \boldsymbol{r}_i^{F_x\,T} \right)^{-1} \left( \hat{\boldsymbol{T}}_{c*,p}^{\mathrm{GREG}} - \hat{\boldsymbol{T}}_{c,p}^{\mathrm{GREG}} \right)$$

and

$$w_g^{\mathrm{TL2}} = w_g^{\mathrm{GREG}} + \boldsymbol{r}_g^{F_a\,T} \left( \sum_{g \in s_h} \boldsymbol{r}_g^{F_a} \boldsymbol{r}_g^{F_a\,T} \right)^{-1} \left( \hat{\boldsymbol{T}}_{c*,p}^{\mathrm{GREG}} - \hat{\boldsymbol{T}}_{c,h}^{\mathrm{GREG}} \right).$$

In the extended approach, both the person-level and household-level estimators are affected by

the common variables inducing an additional source of randomness. Using the Taylor linearisation technique, the variance estimators are given by

$$\hat{V}(\hat{T}^{\mathrm{TL2}}_{y,t}) \doteq \hat{V}_1 + \hat{V}_2 + \hat{V}_3 + 2 \cdot \hat{V}_{12} \; - \; 2 \cdot \hat{V}_{13} \; - \; 2 \cdot \hat{V}_{23} \,\text{for}\, t = \{p,\, h\}. \tag{24}$$

At the person level, the variance components are

$$\begin{aligned} \hat{V}_1 &= \hat{V}(\hat{T}^{\mathrm{GREG}}_{y,p}), & \hat{\mathrm{V}}_{12} &= \hat{\boldsymbol{D}}_c^T \widehat{\mathrm{Cov}}(\hat{T}^{\mathrm{GREG}}_{y,p},\, \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c*,p}), \\ \hat{V}_2 &= \hat{\boldsymbol{D}}_c^{\mathrm{T}} \hat{V}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c*,p}) \hat{\boldsymbol{D}}_c, & \hat{\mathrm{V}}_{13} &= \hat{\boldsymbol{D}}_c^T \widehat{\mathrm{Cov}}(\hat{T}^{\mathrm{GREG}}_{y,p},\, \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,p}), \\ \hat{V}_3 &= \hat{\boldsymbol{D}}_c^{\mathrm{T}} \hat{V}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,p}) \hat{\boldsymbol{D}}_c, & \hat{\mathrm{V}}_{23} &= \hat{\boldsymbol{D}}_c^T \widehat{\mathrm{Cov}}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c*,p},\, \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,p}) \hat{\boldsymbol{D}}_c. \end{aligned} \tag{25}$$

At the household level, the variance components are obtained from

$$\begin{aligned} \hat{V}_1 &= \hat{V}(\hat{T}^{\mathrm{GREG}}_{y,h}), & \hat{\mathrm{V}}_{12} &= \hat{\boldsymbol{E}}_c^T \widehat{\mathrm{Cov}}(\hat{T}^{\mathrm{GREG}}_{y,h},\, \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c*,p}), \\ \hat{V}_2 &= \hat{\boldsymbol{E}}_c^T \hat{V}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c*,p}) \hat{\boldsymbol{E}}_c, & \hat{\mathrm{V}}_{13} &= \hat{\boldsymbol{E}}_c^T \widehat{\mathrm{Cov}}(\hat{T}^{\mathrm{GREG}}_{y,h},\, \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,h}), \\ \hat{V}_3 &= \hat{\boldsymbol{E}}_c^T \hat{V}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,h}) \hat{\boldsymbol{E}}_c, & \hat{\mathrm{V}}_{23} &= \hat{\boldsymbol{E}}_c^T \widehat{\mathrm{Cov}}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c*,p},\, \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,h}) \hat{\boldsymbol{E}}_c. \end{aligned} \tag{26}$$

The variance components in (25) depend solely on the person level, whereas the variance components in (26) are influenced by both person-level and household-level estimates.

### *4.4 Comparison of the Naïve and Extended Two-level Generalised Regression Estimator*

The implementation expense of the naïve two-level GREG estimator is lower compared with the extended two-level GREG estimator, because to ensure consistency, only the household-level estimator is adjusted by the common variables. As estimator for the unknown totals, our proposed estimator at the person level is applied. The implementation expense of our extended two-level GREG estimator, on the other side, is more demanding, because both the person-level and household-level estimators are affected by the consistency requirements. Moreover, the estimation of $\tilde{\boldsymbol{T}}_c = \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c*,p}$ increases the implementation effort compared with the naïve two-level GREG estimator because for every single common variable, the specialised auxiliary variables $\boldsymbol{z}_l$ have to be determined. However, we expect a precision gain for the estimates of the common variables and for all variables correlated with the common variables. To conclude, the choice between our naïve and extended two-level GREG estimators is characterised by a trade-off between the implementation expense and the quality of the final estimates.

### *4.5 Further Remarks*

Both the naïve and extended two-level GREG estimators can also be applied under two-stage cluster sampling where only a subset of the persons within a household are sampled. The point estimators remain unaffected by the two-stage design. Only the variance estimators (17) and (24) have to be adjusted by a term that captures the additional randomness introduced by the sampling process at the second stage. We refer to Särndal *et al*. (1992, p. 126) and Estevao & Särndal (2006, p. 140) for more details on the variance formula under two-stage cluster sampling. The weights produced by the GREG estimator, and thus also by the proposed two-level GREG estimator, can be very large or negative. The reasons for this might be small sample sizes or a variety of auxiliaries, or both. Even if negative weights do not affect the statistical properties of the estimator, they are undesirable for most survey users. Large weights, in turn, can

cause unstable estimations. Fortunately, considerable literature exists on methods to reduce the range of the weights. Huang & Fuller (1978) first proposed a procedure that prevents extreme weights. Deville *et al*. (1993) introduced generalised raking estimators as a subclass of calibration estimators based on a multiplicative distance function. The subclass of generalised raking estimators contains the classical raking estimator originated by Deming & Stephan (1940). Deville & Särndal (1992) and Deville *et al*. (1993) introduced the truncated linear and logit distance function that produce weights that lie within a given range. Singh & Mohl (1996) compared several bounded distances by means of numerical examples. Husain (1969) and Isaki *et al*. (2004) used quadratic programming as an optimisation method to set the weights boundary within a certain interval. Théberge (2000) deduced conditions under which a solution of the optimisation problem ensuring non-extreme weights exists. Tillé (1998) and Park & Fuller (2005) proposed a procedure that produces weights that are positive for the most samples. Chambers (1996) considered a ridge-type optimisation problem under a certain coefficient matrix to produce non-negative weights. Another possibility for avoiding extreme weights is to relax some of the calibration constraints. Rao & Singh (1997) studied a ridge shrinkage method for range-restricted weights, where the calibration constraints are satisfied within certain tolerances.

Each common variable for which consistency is desired increases the number of auxiliary variables in the two-level GREG estimator. Increasing the number of auxiliary variables might reduce the efficiency improvements compared with integrated weighting. However, in the context of household surveys, it is unlikely that the same item is simultaneously requested in the person and in the household questionnaire. Instead, it is more prevalent that for some person characteristics, their corresponding values at the household level are of interest. For such person characteristics, the per-household aggregated variables are computed and supplementarily added to the household data set. Plausible examples include household income, purchases or the number of employees in a household. The decision for which variable consistency is desired should be well balanced so that the consistency achievements are not counterbalanced by a loss in precision of the estimates of the other variables in the survey.

Benchmarking methods according to Cholette (1984) that are used to agree annual and sub-annual estimates in business surveys can be considered as special case of the two-level GREG estimator.

## 5 Further Weighting Approaches

The GLS weighting adjustment algorithm introduced by Zieschang (1990) can also be adopted to ensure consistent estimates between person-level and household-level estimates. The original intention of the GLS adjustment algorithm was to link the estimates of two independent surveys. The collected information obtained from the different surveys overlaps for some items. To link the estimates of the overlapping variables, the auxiliary information of both surveys is pooled and additional linear constraints are imposed into the weighting adjustment algorithm. Merkouris (2004) modified the GLS adjustment algorithm to account for different sample sizes. The GLS adjustment algorithm is equivalent to a calibration estimator with a chi-squared distance function.

### *5.1 Generalised Least Squares Adjustment Algorithm According to Zieschang (1990)*

To conform the GLS adjustment algorithm to our two-level GREG estimator, we embed the GLS adjustment algorithm into the GREG estimation framework. Then the GLS adjustment algorithm introduced by Zieschang (1990) can be expressed as

$$\hat{T}_{y,p}^{\text{ZIE}} = \hat{T}_{y,p}^{\text{GREG}} - \hat{\boldsymbol{D}}_c^{\text{ZIE}^T}\left(\hat{\boldsymbol{T}}_{c,p}^{\text{GREG}} - \hat{\boldsymbol{T}}_{c,h}^{\text{GREG}}\right) \tag{27}$$

and

$$\hat{T}_{y,h}^{\text{ZIE}} = \hat{T}_{y,h}^{\text{GREG}} + \hat{\boldsymbol{E}}_c^{\text{ZIE}^T}\left(\hat{\boldsymbol{T}}_{c,p}^{\text{GREG}} - \hat{\boldsymbol{T}}_{c,h}^{\text{GREG}}\right). \tag{28}$$

The person-level and household-level coefficients are obtained from

$$\hat{\boldsymbol{D}}_c^{\text{ZIE}} = \left(\sum_{i \in s_p} \boldsymbol{r}_i^{F_x}\boldsymbol{r}_i^{F_x T} + \sum_{g \in s_h} \boldsymbol{r}_g^{F_a}\boldsymbol{r}_g^{F_a T}\right)^{-1} \sum_{i \in s_p} \boldsymbol{r}_i^{F_x} r_i^{B_x} \tag{29}$$

and

$$\hat{\boldsymbol{E}}_c^{\text{ZIE}} = \left(\sum_{i \in s_p} \boldsymbol{r}_i^{F_x}\boldsymbol{r}_i^{F_x T} + \sum_{g \in s_h} \boldsymbol{r}_g^{F_a}\boldsymbol{r}_g^{F_a T}\right)^{-1} \sum_{g \in s_h} \boldsymbol{r}_g^{F_a} r_g^{B_a}, \tag{30}$$

respectively. Superscript ZIE refers to Zieschang. Therefore, according to (29) and (30), both estimators $\hat{T}_{y,p}^{\text{ZIE}}$ and $\hat{T}_{y,h}^{\text{ZIE}}$ use the same pooled auxiliary information from the person and household levels. The only difference between both estimators is given by the variable of interest. The impact of ensuring consistency is quantified by the respective second terms and depends on the difference between the estimated common variable totals.

The weights of the estimators (27) and (28) are obtained from

$$w_i^{\text{ZIE}} = w_i^{\text{GREG}} - \boldsymbol{r}_i^{F_x}\left(\sum_{i \in s_p} \boldsymbol{r}_i^{F_x}\boldsymbol{r}_i^{F_x T} + \sum_{g \in s_h} \boldsymbol{r}_g^{F_a}\boldsymbol{r}_g^{F_a T}\right)^{-1}\left(\hat{\boldsymbol{T}}_{c,p}^{\text{GREG}} - \hat{\boldsymbol{T}}_{c,h}^{\text{GREG}}\right)$$

and

$$w_g^{\text{ZIE}} = w_g^{\text{GREG}} + \boldsymbol{r}_g^{F_a}\left(\sum_{i \in s_p} \boldsymbol{r}_i^{F_x}\boldsymbol{r}_i^{F_x T} + \sum_{g \in s_h} \boldsymbol{r}_g^{F_a}\boldsymbol{r}_g^{F_a T}\right)^{-1}\left(\hat{\boldsymbol{T}}_{c,p}^{\text{GREG}} - \hat{\boldsymbol{T}}_{c,h}^{\text{GREG}}\right).$$

Given $\sum_{i \in s_p} \boldsymbol{x}_i \boldsymbol{r}_i^{F_x T} = \boldsymbol{0}$ as well as $\sum_{g \in s_h} \boldsymbol{a}_g \boldsymbol{r}_g^{F_a T} = \boldsymbol{0}$, it is easy to show that the weights simultaneously satisfy $\sum_{i \in s_p} w_i^{\text{ZIE}} \boldsymbol{x}_i = \boldsymbol{T}_x$ and $\sum_{g \in s_h} w_g^{\text{ZIE}} \boldsymbol{a}_g = \boldsymbol{T}_a$. This implies that the sums of the weighted auxiliaries meet the known totals at both levels. Moreover, Merkouris (2004) showed that

$$\begin{aligned} \hat{\boldsymbol{T}}_{c,p}^{\text{ZIE}} &= \left(1 - \hat{\boldsymbol{D}}_c^{\text{ZIE}^T}\right)\hat{\boldsymbol{T}}_{c,p}^{\text{GREG}} + \hat{\boldsymbol{D}}_c^{\text{ZIE}^T}\hat{\boldsymbol{T}}_{c,h}^{\text{GREG}}, \\ &= \hat{\boldsymbol{E}}_c^{\text{ZIE}^T}\hat{\boldsymbol{T}}_{c,p}^{\text{GREG}} + \left(1 - \hat{\boldsymbol{E}}_c^{\text{ZIE}^T}\right)\hat{\boldsymbol{T}}_{c,h}^{\text{GREG}} \\ &= \hat{\boldsymbol{T}}^{\text{ZIE}}{}_{c,h}. \end{aligned} \tag{31}$$

Accordingly, the GLS adjustment algorithm produces the same estimates for the common variable totals at both levels. The equality $\hat{\boldsymbol{E}}_c^{\text{ZIE}} = \left(1 - \hat{\boldsymbol{D}}_c^{\text{ZIE}}\right)$ is valid because $\sum_{i \in s_p} \boldsymbol{c}_i \boldsymbol{r}_i^{F_x T} = \sum_{i \in s_p} \boldsymbol{r}_i^{F_x}\boldsymbol{r}_i^{F_x T}$. Due to the weighted average form of (31), the GLS adjustment algorithm is also called composite estimator (cf. Merkouris, 2004).

The variance estimator of (27) using the Taylor linearisation technique is given by

$$\hat{V}(\hat{T}^{\mathrm{ZIE}}_{y,p}) \doteq \hat{V}_1 + \hat{V}_2 + \hat{V}_3 \; - \; 2\hat{\mathrm{V}}_{12} + 2\hat{\mathrm{V}}_{13} \; - \; 2\hat{\mathrm{V}}_{23}, \tag{32}$$

with

$$\begin{aligned}
\hat{V}_1 &= \hat{V}(\hat{T}^{\mathrm{GREG}}_{y,p}), & \hat{\mathrm{V}}_{12} &= \hat{\boldsymbol{D}}_K^T \widehat{\mathrm{Cov}}(\hat{T}^{\mathrm{GREG}}_{y,p}, \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,p}),\\
\hat{V}_2 &= \hat{\boldsymbol{D}}_K^T \hat{V}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,p})\hat{\boldsymbol{D}}_K, & \hat{\mathrm{V}}_{13} &= \hat{\boldsymbol{D}}_K^T \widehat{\mathrm{Cov}}(\hat{T}^{\mathrm{GREG}}_{y,p}, \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,h}),\\
\hat{V}_3 &= \hat{\boldsymbol{D}}_K^T \hat{V}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,h})\hat{\boldsymbol{D}}_K, & \hat{\mathrm{V}}_{23} &= \hat{\boldsymbol{D}}_K^T \widehat{\mathrm{Cov}}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,p}, \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,h})\hat{\boldsymbol{D}}_K.
\end{aligned}$$

The corresponding variance estimator at the household level is obtained from

$$\hat{V}(\hat{T}^{\mathrm{ZIE}}_{y,h}) \doteq \hat{V}_1 + \hat{V}_2 + \hat{V}_3 + 2\hat{\mathrm{V}}_{12} \; - \; 2\hat{\mathrm{V}}_{13} \; - \; 2\hat{\mathrm{V}}_{23} \tag{33}$$

with

$$\begin{aligned}
\hat{V}_1 &= \hat{V}(\hat{T}^{\mathrm{GREG}}_{y,h}), & \hat{\mathrm{V}}_{12} &= \hat{\boldsymbol{E}}_K^T \widehat{\mathrm{Cov}}(\hat{T}^{\mathrm{GREG}}_{y,h}, \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,p}),\\
\hat{V}_2 &= \hat{\boldsymbol{E}}_K^T \hat{V}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,p})\hat{\boldsymbol{E}}_K, & \hat{\mathrm{V}}_{13} &= \hat{\boldsymbol{E}}_K^T \widehat{\mathrm{Cov}}(\hat{T}^{\mathrm{GREG}}_{y,h}, \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,h}),\\
\hat{V}_3 &= \hat{\boldsymbol{E}}_K^T \hat{V}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,h})\hat{\boldsymbol{E}}_K, & \hat{\mathrm{V}}_{23} &= \hat{\boldsymbol{E}}_K^T \widehat{\mathrm{Cov}}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,p}, \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,h})\hat{\boldsymbol{E}}_K.
\end{aligned}$$

### *5.2 Generalised Least Squares Adjustment Algorithm According to Merkouris (2004)*

Merkouris (2004) modified the GLS adjustment algorithm to account for the effective sample sizes of the independent multiple surveys. Then the estimators are defined as

$$\hat{T}^{\mathrm{MER}}_{y,p} = \hat{T}^{\mathrm{GREG}}_{y,p} \; - \; \hat{\boldsymbol{D}}_c^{\mathrm{MER}^T}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,p} \; - \; \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,h}) \tag{34}$$

and

$$\hat{T}^{\mathrm{MER}}_{y,h} = \hat{T}^{\mathrm{GREG}}_{y,h} + \hat{\boldsymbol{E}}_c^{\mathrm{MER}^T}(\hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,p} \; - \; \hat{\boldsymbol{T}}^{\mathrm{GREG}}_{c,h}). \tag{35}$$

Superscript MER refers to Merkouris. The coefficients are obtained from

$$\hat{\boldsymbol{D}}_c^{\mathrm{MER}} = \left((1 \; - \; q)\sum_{i \in s_p} \boldsymbol{r}_i^{F_x}\boldsymbol{r}_i^{F_x{}^T} + q\sum_{g \in s_h} \boldsymbol{r}_g^{F_a}\boldsymbol{r}_g^{F_a{}^T}\right)^{-1}(1 \; - \; q)\sum_{i \in s_p} \boldsymbol{r}_i^{F_x} r_g^{B_x} \tag{36}$$

and

$$\hat{\boldsymbol{E}}_c^{\mathrm{MER}} = \left((1 \; - \; q)\sum_{i \in s_p} \boldsymbol{r}_i^{F_x}\boldsymbol{r}_i^{F_x{}^T} + q\sum_{g \in s_h} \boldsymbol{r}_g^{F_a}\boldsymbol{r}_g^{F_a T}\right)^{-1} q\sum_{g \in s_h} \boldsymbol{r}_g^{F_a} r_g^{B_a}, \tag{37}$$

where weighting factor

$$q = \frac{V(\hat{T}^{\mathrm{MER}}_{c,h}) - \mathrm{Cov}(\hat{T}^{\mathrm{MER}}_{c,p}, \hat{T}^{\mathrm{MER}}_{c,h})}{V(\hat{T}^{\mathrm{MER}}_{c,h}) + V(\hat{T}^{\mathrm{MER}}_{c,h}) - 2\mathrm{Cov}(\hat{T}^{\mathrm{MER}}_{c,p}, \hat{T}^{\mathrm{MER}}_{c,h})}$$

minimises the variance of the composite estimator for the common variables. If the samples were independent, $q$ approximately equals to $q = \dfrac{n/\mathrm{deff}_p}{n/\mathrm{deff}_p + m/\mathrm{deff}_h}$ with deff as design effect. In this case, $q$ is proportional to the effective sample size. The modified estimators suggested by

Merkouris (2004) differ from the original GLS adjustment algorithm (27) and (28) only with respect to weighting factor $q$. Guandalini & Tillé (2017) showed that (36) and (37) are the optimal coefficients if the samples were independent and the sample size is very large. The variance estimators arise considering the weighting factor $q$ in (32) and (33), as done in (36) and (37).

### 5.3 Comparison of the Two-level Generalised Regression Estimator and the Generalised Least Squares Adjustment Algorithm

The differences between our two-level GREG estimator and the GLS adjustment algorithm arise from how consistency is conceptually ensured and the estimation of the common variable totals. Our two-level GREG estimator uses the same estimated common variable totals in the person-level and household-level estimators. As estimator for the unknown common variable totals, we propose a person-level estimator, because in household surveys, it is more prevalent that the common variables are originally person characteristics that are assigned in aggregated form to household-level data set. In contrast, the GLS adjustment algorithm enforces consistent person-level and household-level estimates more indirectly through the pooling of the auxiliary information. The final estimates of the unknown common variable totals are determined by a weighted average of the single person-level and household-level estimates. Therefore, the same common variable information is used twice, once in its initial form at the person level and once in aggregated form at the household level. However, it is questionable to what extent the aggregated household-level information, supplementary to the person-level information, helps to predict the common variable totals.

Furthermore, when comparing the coefficients in the two-level GREG estimators (15) and (22) with the coefficient in the GLS adjustment algorithm (29), (30), (36) and (37), it becomes evident that the latter coefficients simultaneously use person-level and household-level information by the term $\left(\sum_{i \in s_p} \boldsymbol{r}_i^{F_x} \boldsymbol{r}_i^{F_x{}^T} + \sum_{g \in s_h} \boldsymbol{r}_g^{F_a} \boldsymbol{r}_g^{F_a{}^T}\right)$. However, as mentioned in the previous paragraph, in the context of household surveys, it seems questionable to what extent the household-level auxiliary information helps to predict the person-level variables.

Finally, the variance estimators of the common variables differ. In the two-level GREG estimator, the variance of the common variable totals depends solely on the person-level variance estimator. In contrast, in the GLS adjustment algorithm, the variance estimation for the common variable totals (31) is more demanding, because the variances at both levels and their covariances are required.

## 6 Simulation Study

An MC simulation study compares the performance of the naïve (TL1) and extended two-level GREG estimators (TL2) with the integrated GREG estimator with scale factors $v_i = 1$ (INT1) and $v_i = N_g^{-1}$ (INT2) and with the GLS weighting algorithm according to Zieschang (1990) (ZIE) and according to Merkouris (2004) (MER). The simulation study is based on the synthetic and open accessible data set AMELIA, which is derived from EU-SILC (cf. Burgard *et al*., 2017). To reduce computational burden, we use the data of only one out of four regions. The population consists of approximately 2.6 million persons and 0.9 million households. We draw $R = 1000$ MC samples via simple random sampling. As sampling size, we choose $m = 200$. The MC average sample size of persons is $\sum_{r=1}^{1000} n_r = 577$.

The auxiliaries consist of 18 indicator variables and an intercept: two sex categories, four age categories and four marital status categories plus the cross-classification of sex by age (four

categories) and sex by marital status (four categories). Within the integrated GREG estimators, we also include the additional auxiliary, $N_g^{-1}$, required to ensure the integrated property. The choice of solely categorical variables induces that the GREG estimator is equivalent to a post-stratification estimator (cf. Zhang, 2000). For a fair comparison, we also incorporate the household size as further auxiliary variable into the two-level GREG estimator and into the GLS adjustment algorithm. We chose three types of variables of interest at the person level. Firstly, classical person-level characteristics: personal income `inc` and social income `soc`. Secondly, a variable related with household size: personal income by household size `inc_hs`. In practice, estimates for subgroups or domains are often of as much interest as population totals. Thus, thirdly, we analyse some cross-classifications of age by sex by marital status: `age2_sex1_ms4`(age class 20–39 years, female, widowed) and `age4_sex0_ms2`(male, age class 60 years and older, married) as well as the cross-classifications of employment by sex: `act1_sex0` (at work, male) and `act1_sex1`(at work, female). Note that age, sex and marital status are used as auxiliary variables in the assisting model. The variables of interest at the household level are regular taxes on wealth `taxes`, total disposable household income `disp_inc` and regular inter-household cash transfer received `cash_trans`. We compare two different sets of common variables. In case 1 ($L = 2$), the common variables are `inc` and `soc`. In case 2, there are $L = 10$ common variables including `inc`, `soc`, `dou` as degree of urbanisation with three categories (densely populated area, intermediate area and thinly populated area), basic activity status `act` with three categories (at work, unemployed and inactive persons) and four different income components.

In order to explore the quality of the point estimates, we compare the empirical mean squared error (RRMSE). Suppose $\hat{T}_{y_r}$ as the resulting total estimate for the $r$-th MC replication with $r = 1, \ldots, R$. Define $E^{\mathrm{MC}}(\hat{T}) = R^{-1}\sum_{r=1}^{R}\hat{T}_r$, where the quantity $E^{\mathrm{MC}}(\hat{T})$ denotes the empirical expectation of the estimator $\hat{T}$. Let $T$ be the true value. Then RRMSE is defined as

$$\mathrm{RRMSE}(\hat{T}) = \sqrt{\frac{1}{R}\sum_{r=1}^{R}\frac{\left(\hat{T}_r - T\right)^2}{T^2}}.$$

To examine the spread of the weights, Table 1 outlines the summary statistics for the weights of the competing methods divided by the design weight for all $R = 1\,000$ MC samples (about 577

Table 1. *Summary statistics for the weights.*

| | $L = 2$ *common variables* | | | | | $L = 10$ *common variables* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *Mean* | *SD* | *Min* | *Max* | *Range* | *Mean* | *SD* | *Min* | *Max* | *Range* |
| Person level | | | | | | | | | | |
| INT1 | 1.00 | 0.28 | −1.30 | 8.24 | 9.54 | 1.00 | 0.28 | −1.30 | 8.24 | 9.54 |
| INT2 | 1.00 | 0.32 | −0.85 | 5.32 | 6.17 | 1.00 | 0.32 | −0.85 | 5.32 | 6.17 |
| TL1 | 1.00 | 0.22 | 0.01 | 5.09 | 5.07 | 1.00 | 0.22 | 0.01 | 5.09 | 5.07 |
| TL2 | 1.00 | 0.23 | −0.22 | 5.12 | 5.34 | 1.00 | 0.24 | −1.95 | 5.11 | 7.05 |
| ZIE | 1.00 | 0.22 | 0.03 | 5.11 | 5.08 | 1.00 | 0.73 | −1.15 | 5.74 | 6.88 |
| MER | 1.00 | 0.22 | 0.03 | 5.11 | 5.08 | 1.00 | 0.73 | −1.15 | 5.74 | 6.88 |
| Household level | | | | | | | | | | |
| INT1 | 1.00 | 0.33 | −1.30 | 8.24 | 9.54 | 1.00 | 0.33 | −1.30 | 8.24 | 9.54 |
| INT2 | 1.00 | 0.29 | −0.85 | 5.32 | 6.17 | 1.00 | 0.29 | −0.85 | 5.32 | 6.17 |
| TL1 | 1.00 | 0.32 | −1.16 | 5.19 | 6.36 | 1.00 | 2.79 | −8.38 | 8.92 | 17.31 |
| TL2 | 1.00 | 0.33 | −1.20 | 5.21 | 6.41 | 1.00 | 2.79 | −8.66 | 9.23 | 17.89 |
| ZIE | 1.00 | 0.32 | −1.16 | 5.18 | 6.35 | 1.00 | 0.75 | −1.94 | 6.07 | 8.02 |
| MER | 1.00 | 0.32 | −1.16 | 5.18 | 6.35 | 1.00 | 0.75 | −1.94 | 6.07 | 8.02 |

000 values). It becomes apparent that at the person level, both integrated GREG estimators INT1 and INT2 have the highest ranges. Note that INT1 and INT2 show the same minimum and maximum at the person and household levels because of the one-to-one weight assignment between the levels. Moreover, both integrated approaches remain unaffected by the increase of the number of common variables. The same is true for TL1, which is at the person level equal to an ordinary GREG estimator. For $L = 2$, the weights of both proposed two-level GREG estimators TL1 and TL2 have a considerably smaller range than the integrated weights. For TL1 as ordinary GREG estimator, this observation confirms with the results found by Lemaître & Dufour (1987) and Rottach & Hall (2005). Moreover, the standard deviation is lower for TL1 and TL2 than for both integrated approaches. The higher variation of the integrated weights is caused by the increased number of outcome values of the auxiliary variables as described in Section 7. However, even if the weights of TL2 vary less, their range increases with the number of common variables. This means that in some MC samples, extreme weights emerge. The household weights vary more than the person weights except for INT2, which weights are modelled at the household level. At the household level, TL1 and TL2 suffer from the increase of the common variables. As mentioned in Section 4.4, there is a considerable literature on methods to reduce the range of the weights. Accordingly, all estimators under consideration can be adjusted using other calibration distance functions than the chi-squared distance function. However, because negative weights do not affect the statistical properties of the resulting estimators, we refrain focusing on these methods in this paper.

Figures 1 and 2 plot the ratios of the MSEs of integrated weighting relative to our two-level GREG estimators and relative to the GLS adjustment algorithm. Both plots give almost the

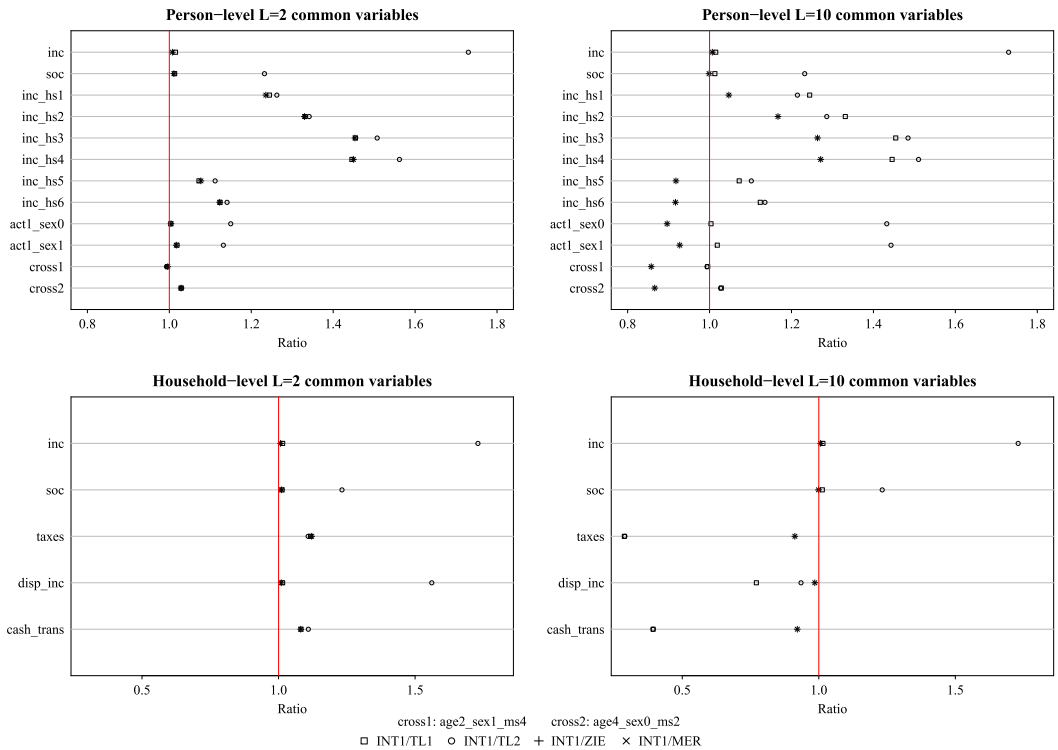


**Figure 1.** *Ratios of INT1 relative to the two-level generalised regression estimator and the generalised least squares adjustment algorithm.*

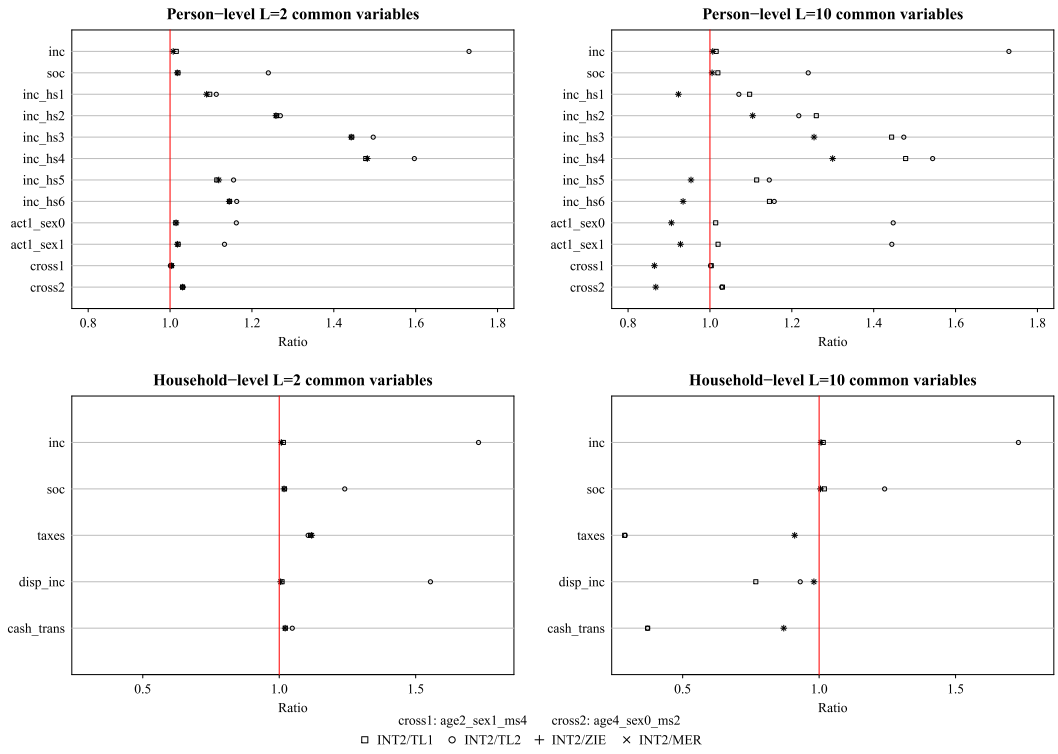


**Figure 2.** *Ratios of INT2 relative to the two-level generalised regression estimator and the generalised least squares adjustment algorithm.*

similar pictures on the efficiency of the estimators relative to INT1 and INT2. With respect to the person level, it becomes evident that all squares and circles indicating TL1 and TL2 lie nearby or to the right of the red line. Therefore, TL1 and TL2 perform at least as well as integrated weighting independent from the number of common variables. The greatest gains in precision are realised for TL2. Here, all variables benefit from inserting improved estimates for the common variable totals estimated by a specialised auxiliary variable set for each common variable. Actually, for the common variable `inc`, the variance of integrated weighting exceeds the variance of the two-level GREG estimator up to 73%. There are also significant efficiency gains realised for the subdomains of `inc` and `hs`, because of the correlation structure with the common variables in the assisting model. Interestingly, these precision gains are less pronounced for smaller and larger households. This observation confirms with the different explanatory power of the assisting models, which ranges from 0.19 to 0.22 for smaller and larger households and from 0.39 to 0.44 for the medium-size households. The high efficiency gains with respect to income and income-related variables are a hint that the proposed two-level GREG estimators better capture the heterogeneity within a household for variables that are differently distributed within a household. The efficiency improvements of the two-level GREG estimator compared with integrated weighting are a hint that the within variance of the households should not be ignored.

Because the person-level TL1 estimator is equivalent with an ordinary GREG estimator, the ratios of INT1/TL1 or INT2/TL1 are comparable with the results given in the literature on integrated weighting. Except for the cross-classification of `inc` and `hs`, our results are in line with the statement given by Lemaître & Dufour (1987), Steel & Clark (2007) and van den

Brakel (2013,2016) that the inefficiencies of integrated weighting would be limited compared with an ordinary GREG estimator.

MER and ZIE perform in all scenarios nearly identical. Even if the common variables are included as additional auxiliaries, no considerable efficiency gains are realised with respect to `inc` and `soc`. For the cross-classifications of `inc` and `hs`, ZIE and MER perform more efficient than integrated weighting. Therefore, the explanatory power of household size outweighs the explanatory power of income with the variable of interest in the assisting model. However, the efficiency of ZIE and MER suffers with increasing number of common variables.

At the household level and for $L = 2$, our two-level GREG estimators outperform integrated weighting; however, there are considerably efficiency losses for $L = 10$ and `taxes`, `disp_inc` and `cash_trans`. Therefore, in order to avoid an efficiency loss for the estimation of household characteristics, the number of common variables should be chosen with caution. Nevertheless, the efficiency gains at the person level of our proposed two-level GREG estimator are convincing, in particular for variables related to the common variables.

## 7 Conclusion

Consistent person-level and household-level estimates are a major concern of statistical offices. Integrated weighting produces consistent estimates by using the same weights for all persons within a household and the household itself. The consequences of equal weights are an increased number of factor values of the auxiliaries and the ignorance of heterogeneity of the persons within households. As remedy, we proposed a two-level GREG estimator to achieve consistency without the strict requirement of equal weights for all household members. To ensure consistent estimates, we restrain the consistency requirements to the variables common to the person-level and household-level data sets. By using the same estimated total for the common variables at both levels, consistency is ensured more directly and only for the relevant variables. In many household surveys, only few variables are provided on both the person and household levels such that the number of additional auxiliary variables is limited.

Our MC simulation study supports the superiority of our proposed two-level GREG estimator compared with integrated weighting in particular with respect to the common variables and variables correlated with them. However, the number of variables for which consistency is required should be chosen carefully in order to avoid inefficiencies for some household estimates. The choice between the naïve and extended two-level GREG estimators is characterised by a trade-off between of the implementation expenses and quality of the final estimates. In case of the extended two-level GREG estimator, specialised variables have to be specified and the variance estimation is more demanding, but there are significant efficiency gains possible. These efficiency gains depend on the strength of the relation between the common variables and variables of interest. With this two-level GREG estimator, we contradict the assumption prevailing in the literature that equal weights of persons within the same household and the household itself are indispensable to ensure consistent person and household estimates (cf. Lavallée, 1995, p. 27 Estevao & Särndal, 2006, p. 139; Steel & Clark, 2007, p. 51; Branson & Wittenberg, 2014, p. 20; van den Brakel, 2016, p. 149).

Another application field for our proposed two-level GREG estimator is integrated surveys where several smaller surveys are integrated in a core sample. In this context, consistent estimates between the core sample and the subsample are desired. However, in integrated surveys, the number of common variables might be large. This enlarges the range of the weights and results in less stable estimates. Therefore, the two-level GREG estimator has to be extended to handle with a variety of auxiliary variables. The combination with generalised calibration using soft constraints may additionally be considered in this case (Burgard *et al*., 2019).

For most survey practitioner, negative weights may be undesirable. Therefore, further research should focus on methods that prevent negative weights within the two-level GREG estimator framework. A possible starting point would be the substitution of the chi-squared distance function by a linear truncated or multiplicative distance function. However, an extension to various other calibration functions seems straightforward but urges using adequate numerical iteration methods.

Further future research should address the effects of consistency requirements on non-response adjustment. In general, methods to prevent a non-response bias proceed at the person level. Hence, the adjusted person weights are no longer necessarily equal within a household. In order to still guarantee consistency, Eurostat (cf. European Commission, 2014, p. 40) recommends averaging the adjusted person weights within a household and assigning this average weight to all household members. In contrast, our two-level GREG estimator allows a non-response adjustment at the person level without the need for a subsequent averaging process of the resulting weights. The incorporation of the common variables guarantees consistency even in the case of non-response adjustment. Therefore, individual response patterns are retained. This flexibility reinforces the superiority of our two-level GREG estimator compared with integrated weighting.

## ACKNOWLEDGEMENTS

The research was conducted within the RIFOSS research project, which is funded by the German Federal Statistical Office. We profoundly thank the editor and the anonymous reviewers for the valuable, detailed and helpful reviews.

## References

Alexander, C.H. 1987. A class of methods for using person controls in household weighting. *Surv. Methodol.*, **13**(2), 183–198.

Boonstra, H.J.H., van den Brakel, J.A., Knotterus, P., Nieuwenbroek, N.J. & Renssen, R.H. 2003. A strategy to obtain consistency among tables of survey estimates. In DACSEIS report D7.2.

Branson, N. & Wittenberg, M. 2014. Reweighting South African national household survey data to create a consistent series over time: a cross-entropy estimation approach. *S. Afr. J. Econ.*, **82**(1), 19–38.

Burgard, J.P., Kolb, J. -P., Merkle, H. & Müunnich, R. 2017. Synthetic data for open and reproducible methodological research in social sciences and official statistics. *AStA Wirtsch. Sozialstat. Arch.*, **11**(3-4), 233–244.

Burgard, J.P., Münnich, R.T. & Rupp, M. 2019. A generalized calibration approach ensuring coherent estimates with small area constraints. Technical report, Research Papers in Economics.

Cassel, C.-M., Särndal, C. & Wretman, J. 1997. Foundations of inference in survey sampling. New York: Wiley.

Chambers, R.L. 1996. Robust case-weighting for multipurpose establishment surveys. *J. Off. Stat.*, **12**(1), 3.

Cholette, P. 1984. Adjusting subannual series to yearly benchmarks. *Surv. Methodol.*, **10**, 35–49.

D'Arrigo, J. & Skinner, C.J. 2010. Linearization variance estimation for generalized raking estimators in the presence of nonresponse. *Surv. Methodol.*, **36**(2), 181–192.

Deming, W.E. & Stephan, F.F. 1940. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Ann. Math. Stat.*, **11**(4), 427–444.

Devaud, D. & Tillé, Y. 2019. Deville and Särndal's calibration: Revisiting a 25-years-old successful optimization problem. *TEST*, **28**, 1033–1065.

Deville, J.-C. & Särndal, C.-E. 1992. Calibration estimators in survey sampling. *J. Am. Stat. Assoc.*, **87**(418), 376–382.

Deville, J.-C., Särndal, C.-E. & Sautory, O. 1993. Generalized raking procedures in survey sampling. *J. Am. Stat. Assoc.*, **88**(423), 1013–1020.

Estevao, V.M. & Särndal, C.-E. 2006. Survey estimates by calibration on complex auxiliary information. *Int. Stat. Rev.*, **74**(2), 127–147.

European Commission. 2014. Methodological guidelines and description of EU-SILC target variables. Technical Report Doc-SILC065, Eurostat, Directorate F: Social Statistics.

Eurostat. 2011. European statistics code of practice. Retrieved from https://ec.europa.eu/eurostat/documents/3859598/5921861/KS-32-11-955-EN.PDF/5fa1ebc6-90bb-43fa-888f-dde032471e15. visited 23/10/2018.
Guandalini, A. & Tillé, Y. 2017. Design-based estimators calibrated on estimated totals from multiple surveys. *Int. Stat. Rev.*, **85**(2), 250–269.
Haziza, D. & Beaumont, J.F. 2017. Construction of weights in surveys: a review. *Stat. Sci.*, **32**(2), 206–226.
Horvitz, D.G. & Thompson, D.J. 1952. A generalization of sampling without replacement from a finite universe. *J. Am. Stat. Assoc.*, **47**(260), 663–685.
Houbiers, M. 2004. Towards a social statistical database and unified estimates at statistics Netherlands. *J. Off. Stat.*, **20**(1), 55.
Huang, E. & Fuller, W.A. 1978. Non-negative regression estimation in sample survey data. *Proc. Soc. Statitics Sect. Am. Stat. Assoc.*, 300–305.
Husain, M. 1969. Construction of regression weights for estimation in sample surveys. unpublished M.S. Thesis, Ames, Iowa.
Isaki, C.T. & Fuller, W.A. 1982. Survey design under the regression superpopulation model. *J. Am. Stat. Assoc.*, **77**(377), 89–96.
Isaki, C.T., Tsay, J.H. & Fuller, W.A. 2004. Weighting sample data subject to independent controls. *Surv. Methodol.*, **30**(1), 35–44.
Kott, P.S. 2009. Calibration weighting: combining probability samples and linear prediction models. *In Handbook of Statistics*, Vol. **29**. Elsevier, pp. 55–82.
Lavallée, P. 1995. Cross-sectional weighting of longitudinal surveys of individuals and households using the weight share method. *Surv. Methodol.*, **21**(1), 25–32.
Lemaître, G. & Dufour, J. 1987. An integrated method for weighting persons and families. *Surv. Methodol.*, **13**, 199–207.
Luery, D. 1986. Weighting survey data under linear constraints on the weights. In *Proceedings of the Survey Research Methods Section, American Statistical Association*, pp. 325–330. Chicago, Illinois.
Merkouris, T. 2004. Combining independent regression estimators from multiple surveys. *J. Am. Stat. Assoc. Univ. Wollongong*, **99**(468), 1131–1139.
Neethling, A. & Galpin, J.S. 2006. Weighting of household survey data: A comparison of various calibration, integrated and cosmetic estimators: Theory and methods. *S. Afr. Stat. J.*, **40**(2), 123–150.
Nieuwenbroek, N. 1993. An integrated method for weighting characteristics of persons and households using the linear regression estimator. Netherlands Central Bureau of Statistics, Department of Statistical Methods.
Park, M. & Fuller, W.A. 2005. Towards nonnegative regression weights for survey samples. *Surv. Methodol.*, **31**(1), 85–93.
Rao, J.N.K. & Singh, A.C. 1997. A ridge-shrinkage method for range-restricted weight calibration in survey sampling. *Proc. Sect. Surv. Res. Methodes*, 57–65. American Statistical Association.
Renssen, R.H. & Nieuwenbroek, N.J. 1997. Aligning estimates for common variables in two or more sample surveys. *J. Am. Stat. Assoc.*, **92**(437), 368–374.
Robinson, W.S. 1950. Ecological correlations and the behavior of individuals. *Am. Sociol. Rev.*, **15**(3), 351–357.
Rottach, R.A. & Hall, D.W. 2005. Using equalization constraints to find optimal calibration weights. American Statistical Association, Section on Survey Research Methods.
Särndal, C.E. 1980. On -inverse weighting versus best linear unbiased weighting in probability sampling. *Biometrika*, **67**(3), 639–650.
Särndal, C.-E. 2007. The calibration approach in survey theory and practice. *Surv. Methodol.*, **33**(2), 99.
Särndal, C.-E., Swensson, B. & Wretman, J. 1992. Model assisted survey sampling. Springer Science & Business Media.
Seber, G.A.F. 1977. Linear Regression Analysis. Wiley series in probability and mathematical statistics. John Wiley & Sons.
Singh, A.C. & Mohl, C.A. 1996. Understanding calibration estimators in survey sampling. *Surv. Methodol.*, **22**(2), 107–115.
Steel, D.G. & Clark, R.G. 2007. Person-level and household-level regression estimation in household surveys. *Surv. Methodol.*, **33**(1), 55–60.
Stukel, D.M., Hidiroglou, M.A. & Särndal, C.-E. 1996. Variance estimation for calibration estimators: A comparison of jackknifing versus Taylor linearization. *Surv. Methodol.*, **22**(2), 117–126.
Théberge, A. 2000. Calibration and restricted weights. *Surv. Methodol.*, **26**(1), 99–108.
Tillé, Y. 1998. Estimation in surveys using conditional inclusion probabilities: Simple random sampling. *Int. Stat. Rev.*, **66**(3), 303–322.
van den Brakel, J. 2013. Sampling and estimation techniques for household panels, Statistics Netherlands, Heerlen.
van den Brakel, J. 2016. Register-based sampling for household panels. *Surv. Methodol.*, **42**(1), 137–159.

Verma, V. & Clémenceau, A. 1996. Methodology of the european community household panel. *Stat. Transit.*, **2**(7), 1023–1062.
Wittenberg, M. 2010. An introduction to maximum entropy and minimum cross-entropy estimation using Stata. *Stata J.*, **10**(3), 315–330.
Wright, R.L. 1983. Finite population sampling with multivariate auxiliary information. *J. Am. Stat. Assoc.*, **78**(384), 879–884.
Zhang, L.C. 2000. Post-stratication and calibration - a synthesis. *Am. Stat.*, **54**(3), 178–184.
Zieschang, K.D. 1986. A generalized least squares weighting system for the consumer expenditure survey. In *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 64–71. Chicago, Illinois.
Zieschang, K.D. 1990. Sample weighting methods and estimation of totals in the consumer expenditure survey. *J. Am. Stat. Assoc.*, **85**(412), 986–1001.

[*Received May 2019; accepted May 2021*]