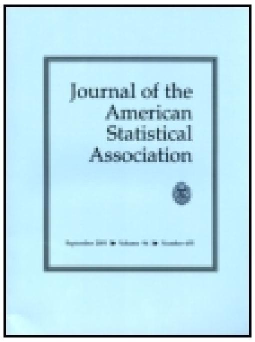

Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

Combining Independent Regression Estimators From Multiple Surveys

Takis Merkouris

To cite this article: Takis Merkouris (2004) Combining Independent Regression Estimators From Multiple Surveys, Journal of the American Statistical Association, 99:468, 1131-1139, DOI: 10.1198/016214504000000601

To link to this article: https://doi.org/10.1198/016214504000000601

Published online: 31 Dec 2011.

Combining Independent Regression Estimators From Multiple Surveys

Takis MERKOURIS

Efficient generalized regression (GREG) procedures are proposed for the combination of comparable information collected independently from multiple surveys of the same population. In particular, for combination through alignment of estimated totals of common target characteristics, an efficiency improvement of Zieschang's original composite GREG method is developed, involving a correction of the GREG estimation for different effective sample sizes in a multiple-sample setting. The proposed method is nearly as efficient as the alternative method of Renssen and Nieuwenbroek for general sampling designs, and considerably more practical. Under broad sampling design conditions, the proposed method produces composite estimators that are equally efficient as the estimators of Renssen and Nieuwenbroek for common characteristics and more efficient for noncommon characteristics. Under some of these sampling design conditions, an equivalence of extended GREG estimation and extended optimal regression estimation is established.

KEY WORDS: Aligned estimates; Composite estimation; Generalized regression estimator; Optimal regression estimator; Weight calibration.

1. INTRODUCTION

Recent years have seen a growing interest within statistical organizations in combining comparable information collected independently from multiple surveys of the same population (prime references are Zieschang 1990; Renssen and Nieuwenbroek 1997; Statistics Canada Symposium 1999). In two or more surveys, the samples of which are drawn independently from the same target population, there is usually some overlap in the collected auxiliary information. It may occur that the samples also contain information on some main target characteristics that are common to these surveys. There is a wide variety of situations involving different sources of comparable information in survey sampling. For instance, there may be some overlap in the collected information in a multicomponent survey, the components of which constitute separate surveys of similar content from the same population. Also, surveys with a split-questionnaire design, intended to reduce response burden, comprise multiple independent samples that are distinguished only by questionnaires of overlapping content.

In this article, we consider generalized regression procedures that combine data from multiple surveys, or from multiple samples in the same survey, for the purpose of producing efficient composite estimators of totals for common target characteristics. Extended regression procedures that incorporate common survey characteristics as additional auxiliary variables can be used to align estimated totals, in the sense that the resulting composite weights of the various surveys produce identical estimates of the population totals for the common characteristics. Besides ensuring consistency between estimates for the common characteristics, the weights of each sample incorporate auxiliary information from the other samples, thereby generating more efficient estimators for common and possibly noncommon characteristics. Such "borrowing of strength" among surveys can be especially beneficial in small-area estimation. Pertaining specifically to noncommon characteristics, an effective application of this methodology may involve a repeated multiple-panel survey in which additional variables are surveyed periodically using a subsample based on one or more panels. A similar use of the methodology can be made in *nonnested* double sampling (Hidiroglou 2001).

A general treatment of the composition problem in a regression context was given in earlier work (Merkouris 2001). The alignment of comparable totals from two separate surveys by means of a regression procedure was first discussed by Zieschang (1990), who introduced a methodology of composite estimation involving an extended regression constraint system and applied it to a two-component consumer expenditure survey. Renssen and Nieuwenbroek (1997) expanded on this idea to propose a class of composite regression estimators, and also suggested the application to split-questionnaire surveys.

In this article, the composition of regression estimators is investigated within an extended framework of optimal regression, with primary focus on the efficiency of derived composite estimators and special emphasis on the practicality of competing composite regression methods. In particular, a modification of Zieschang's method is proposed to account for the differential in effective sample size between two surveys. This modification improves the efficiency of the derived composite regression estimators, while retaining the practical advantages of Zieschang's method.

Prerequisites from the theory of regression in survey sampling are given in Section 2. Two composite regression estimators generated by a separate and a combined regression procedure involving two independent samples are described first in Section 3. The structure of aligned regression estimators generated by a separate or by a combined regression procedure with an extended system of regression constraints is investigated in Section 4. To improve its efficiency, a practicable modification of the combined regression procedure is proposed whereby the differential in effective sample size between two samples is accounted for in the alignment of estimators. Sampling design conditions are given under which the proposed procedure shares with the procedure of Renssen and Nieuwenbroek an optimality property of composite estimators for common characteristics and produces more efficient estimators for noncommon characteristics while being considerably more practical. Importantly, under some of these conditions an equivalence of extended generalized regression (GREG) estimation and extended

Takis Merkouris is Senior Research Methodologist, Household Surveys Methods Division, Statistics Canada, Ottawa, Ontario K1A 0T6, Canada (E-mail: *Takis.Merkouris@statcan.ca*). The author thanks the editor and a referee for constructive suggestions that substantially improved the presentation of the article.

Journal of the American Statistical Association
December 2004, Vol. 99, No. 468, Theory and Methods
DOI 10.1198/016214504000000601

optimal regression estimation is established. The various alignment procedures are compared in terms of statistical efficiency and computational convenience. Summary and discussion are given in Section 5.

2. PRELIMINARIES

Consider a finite population $U = \{1, \dots, k, \dots N\}$, from which a probability sample s of size n is drawn according to a sampling design with known first- and second-order inclusion probabilities π_k and π_{kl} $(k, l \in U)$. Consider the sampling weight vector $\mathbf{w}$ with kth entry defined as $w_k = (1/\pi_k) I(k \in s)$, where I denotes the indicator variable, and let $\mathbf{Y} \in \mathbb{R}^{N \times d}$ denote the population matrix of a d-dimensional survey variable of interest $\mathbf{y}$. The Horvitz–Thompson estimator of the total $\mathbf{t_Y} = \mathbf{Y}'\mathbf{1}$, where $\mathbf{1}$ is the unit N-vector, is given by $\hat{\mathbf{Y}} = \mathbf{Y}'\mathbf{w}$. For the population matrix $\mathbf{X} \in \mathbb{R}^{N \times p}$ of a p-dimensional auxiliary variable $\mathbf{x}$, assume that the total $\mathbf{t_X} = \mathbf{X}'\mathbf{1}$ is known. Then let $\boldsymbol{\Lambda} \in \mathbb{R}^{N \times N}$ be the diagonal "weighting" matrix that has w_k/q_k as the kkth entry, where q_k is a positive constant, and let s designate the subvectors and submatrices corresponding to the sample. A new weight vector, $\mathbf{c}_s \in \mathbb{R}^n$, can be constructed to satisfy the "calibration" constraints $\mathbf{X}_s'\mathbf{c}_s = \mathbf{t_X}$ while minimizing the generalized least squares distance $(\mathbf{c}_s - \mathbf{w}_s)'\boldsymbol{\Lambda}_s^{-1}(\mathbf{c}_s - \mathbf{w}_s)$. This calibration procedure generates a vector of calibrated weights given by

$$\mathbf{c}_s = \mathbf{w}_s + \boldsymbol{\Lambda}_s \mathbf{X}_s (\mathbf{X}_s' \boldsymbol{\Lambda}_s \mathbf{X}_s)^{-1} (\mathbf{t_X} - \mathbf{X}_s'\mathbf{w}_s), \tag{1}$$

assuming that the matrix inverse in (1) exists. The calibration estimator of the total $\mathbf{t_Y}$ is obtained as

$$\mathbf{Y}_s'\mathbf{c}_s = \mathbf{Y}_s'\mathbf{w}_s + \mathbf{Y}_s'\boldsymbol{\Lambda}_s \mathbf{X}_s (\mathbf{X}_s' \boldsymbol{\Lambda}_s \mathbf{X}_s)^{-1} (\mathbf{t_X} - \mathbf{X}_s'\mathbf{w}_s), \tag{2}$$

which can take the form of a GREG estimator,

$$\hat{\mathbf{Y}}^R = \hat{\mathbf{Y}} + \hat{\boldsymbol{\beta}}(\mathbf{t_X} - \hat{\mathbf{X}}) = \hat{\boldsymbol{\beta}}\mathbf{t_X} + (\mathbf{Y}_s - \mathbf{X}_s\hat{\boldsymbol{\beta}}')'\mathbf{w}_s, \tag{3}$$

where $\hat{\mathbf{X}} = \mathbf{X}_s'\mathbf{w}_s$ is the Horvitz–Thompson estimator of $\mathbf{t_X}$ and $\hat{\boldsymbol{\beta}} = \mathbf{Y}_s'\boldsymbol{\Lambda}_s\mathbf{X}_s(\mathbf{X}_s'\boldsymbol{\Lambda}_s\mathbf{X}_s)^{-1}$ is the matrix of sample regression coefficients. The term $(\mathbf{Y}_s - \mathbf{X}_s\hat{\boldsymbol{\beta}}')'\mathbf{w}_s$ in (3) is the sum of weighted sample regression residuals, and $\hat{\boldsymbol{\beta}}$ minimizes the quadratic form $(\mathbf{Y}_s - \mathbf{X}_s\hat{\boldsymbol{\beta}}')'\boldsymbol{\Lambda}_s(\mathbf{Y}_s - \mathbf{X}_s\hat{\boldsymbol{\beta}}')$ in these residuals. By construction, the GREG estimator (3) has the calibration property that $\hat{\mathbf{X}}^R = \mathbf{t_X}$; that is, the GREG estimator of the total for an auxiliary variable is equal to the known population total ("control" total) for that variable. A formulation of the GREG estimator as a calibration estimator was given by Deville and Särndal (1992), and an extensive discussion of it was provided by Särndal, Swensson, and Wretman (1992). When $q_k = 1$ for all k, and $\mathbf{1}_s = \mathbf{X}_s\mathbf{h}$ for a p-vector $\mathbf{h}$, the GREG estimator $\hat{\mathbf{Y}}^R$ reduces to the simple "projection estimator," $\hat{\mathbf{Y}}^R = \hat{\boldsymbol{\beta}}\mathbf{t_X}$.

On the basis of its first-order Taylor linear approximation (see, e.g., Särndal et al. 1992, p. 235), the GREG estimator $\hat{\mathbf{Y}}^R$ is approximately design unbiased with approximate design variance $AV(\hat{\mathbf{Y}}^R) = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}')'\boldsymbol{\Lambda}^\circ(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}')$, where $\boldsymbol{\beta} = \mathbf{Y}'\mathbf{Q}\mathbf{X}(\mathbf{X}'\mathbf{Q}\mathbf{X})^{-1}$, $\mathbf{Q} = \mathrm{diag}(q_1^{-1}, \dots, q_N^{-1})$, and $\boldsymbol{\Lambda}^\circ$ is a semi positive-definite matrix whose klth entry is $(\pi_{kl} - \pi_k\pi_l)/\pi_k\pi_l$, $(\pi_{kk} \equiv \pi_k)$. An estimator of $AV(\hat{\mathbf{Y}}^R)$ is given by $\widehat{AV}(\hat{\mathbf{Y}}^R) = (\mathbf{Y}_s - \mathbf{X}_s\hat{\boldsymbol{\beta}}')'\boldsymbol{\Lambda}_s^\circ(\mathbf{Y}_s - \mathbf{X}_s\hat{\boldsymbol{\beta}}')$, where $\hat{\boldsymbol{\beta}}$ is as in (3) and $\boldsymbol{\Lambda}_s^\circ = \{(\pi_{kl} - \pi_k\pi_l)/\pi_k\pi_l\pi_{kl}\}_{k,l \in s}$.

An alternative regression-type estimator that in general is more efficient than the GREG estimator $\hat{\mathbf{Y}}^R$ in large samples is obtained by choosing from among arbitrary regression coefficients the one that minimizes the variance of the implied estimator. Such an optimal regression coefficient is given by $\boldsymbol{\beta}^\circ = \mathbf{Y}'\boldsymbol{\Lambda}^\circ\mathbf{X}(\mathbf{X}'\boldsymbol{\Lambda}^\circ\mathbf{X})^{-1}$, in which $\mathbf{X}'\boldsymbol{\Lambda}^\circ\mathbf{X}$ and $\mathbf{Y}'\boldsymbol{\Lambda}^\circ\mathbf{X}$ are recognized to be the variance of $\hat{\mathbf{X}}$ and the covariance between $\hat{\mathbf{Y}}$ and $\hat{\mathbf{X}}$. Replacing $\boldsymbol{\beta}^\circ$ by its estimate $\hat{\boldsymbol{\beta}}^\circ = \mathbf{Y}_s'\boldsymbol{\Lambda}_s^\circ\mathbf{X}_s(\mathbf{X}_s'\boldsymbol{\Lambda}_s^\circ\mathbf{X}_s)^{-1}$ gives rise to the "optimal" regression estimator $\hat{\mathbf{Y}}^O = \hat{\mathbf{Y}} + \hat{\boldsymbol{\beta}}^\circ(\mathbf{t_X} - \hat{\mathbf{X}})$. Clearly, $\hat{\mathbf{X}}^O = \mathbf{t_X}$. (For a discussion of various properties of $\hat{\mathbf{Y}}^O$, see Montanari 1987; Rao 1994.)

For the augmented $n \times (p+q)$ matrix $\boldsymbol{\mathcal{X}}_s = (\mathbf{X}_s\ \mathbf{Z}_s)$, where the columns of $\mathbf{Z}_s$ are linearly independent of the columns of $\mathbf{X}_s$, and for the corresponding vector of totals $\mathbf{t}_{\boldsymbol{\mathcal{X}}} = (\mathbf{t}_\mathbf{X}', \mathbf{t}_\mathbf{Z}')'$, a useful decomposition of the vector of calibrated weights $\mathbf{c}_s = \mathbf{w}_s + \boldsymbol{\Lambda}_s\boldsymbol{\mathcal{X}}_s(\boldsymbol{\mathcal{X}}_s'\boldsymbol{\Lambda}_s\boldsymbol{\mathcal{X}}_s)^{-1}(\mathbf{t}_{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{X}}_s'\mathbf{w}_s)$ is obtained as

$$\mathbf{c}_s = \mathbf{c}_{xs} + \mathbf{L}_s\mathbf{Z}_s(\mathbf{Z}_s'\mathbf{L}_s\mathbf{Z}_s)^{-1}(\mathbf{t_Z} - \mathbf{Z}_s'\mathbf{c}_{xs}), \tag{4}$$

where $\mathbf{c}_{xs} = \mathbf{w}_s + \boldsymbol{\Lambda}_s\mathbf{X}_s(\mathbf{X}_s'\boldsymbol{\Lambda}_s\mathbf{X}_s)^{-1}(\mathbf{t_X} - \mathbf{X}_s'\mathbf{w}_s)$, $\mathbf{L}_s = \boldsymbol{\Lambda}_s \times (\mathbf{I} - \mathbf{P_X})$, and $\mathbf{P_X} = \mathbf{X}_s(\mathbf{X}_s'\boldsymbol{\Lambda}_s\mathbf{X}_s)^{-1}\mathbf{X}_s'\boldsymbol{\Lambda}_s$. Then the extended GREG estimator, $\hat{\mathbf{Y}}^{ER}$, corresponding to the partitioned regression matrix $\boldsymbol{\mathcal{X}}_s$ takes the form

$$\hat{\mathbf{Y}}^{ER} = \hat{\mathbf{Y}}^R + \mathbf{Y}_s'\mathbf{L}_s\mathbf{Z}_s(\mathbf{Z}_s'\mathbf{L}_s\mathbf{Z}_s)^{-1}(\mathbf{t_Z} - \hat{\mathbf{Z}}^R), \tag{5}$$

in terms of GREG estimators based on the regression matrix $\mathbf{X}_s$. (For uses of this form of the GREG estimator, see Singh and Merkouris 1995; Renssen and Nieuwenbroek 1997.)

In the following, the subscript i is used to index vectors and matrices associated with different samples from the same population. It is assumed that the samples are independent, with negligible probability of having common units. Also, the typical convention that $q_k = 1$ for all k is followed.

3. TWO COMPOSITE REGRESSION ESTIMATORS

Consider a general setting involving two independent surveys of the same target population or two independent subsamples of the same survey. Assume next that the associated sampling weight vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ aggregate by design to the same population total. For a single variable z (with population vector $\mathbf{Z}$) that is common to the two surveys, and for regression matrices $\mathbf{X}_1$ and $\mathbf{X}_2$—which may or may not have any columns corresponding to common auxiliary variables—consider the independent GREG estimators $\hat{Z}_i^R = \hat{Z}_i + \hat{\boldsymbol{\beta}}_i(\mathbf{t}_{\mathbf{X}_i} - \hat{\mathbf{X}}_i)$ of the total t_Z, where $\hat{Z}_i$ and $\hat{\mathbf{X}}_i$ are the Horvitz–Thompson estimators of t_Z and $\mathbf{t}_{\mathbf{X}_i}$ based on the sample $s_i, i = 1, 2$, and $\hat{\boldsymbol{\beta}}_i = \mathbf{Z}_{s_i}'\boldsymbol{\Lambda}_{s_i}\mathbf{X}_{s_i}(\mathbf{X}_{s_i}'\boldsymbol{\Lambda}_{s_i}\mathbf{X}_{s_i})^{-1}$. Then a composite GREG estimator of the total t_Z, denoted by $\hat{Z}_s^{CR}$, is given by the weighted average of $\hat{Z}_1^R$ and $\hat{Z}_2^R$,

$$\hat{Z}_s^{CR} = \phi\hat{Z}_1^R + (1 - \phi)\hat{Z}_2^R, \tag{6}$$

where $\phi \in [0, 1]$. The subscript in $\hat{Z}_s^{CR}$ indicates that this estimator is the composite of two separate GREG estimators. This composite estimator can be alternatively constructed by a simultaneous regression for the two samples, using the setup $\mathbf{X}_s = \mathrm{diag}(\mathbf{X}_{s_i})$, $\boldsymbol{\Lambda}_s = \mathrm{diag}(\boldsymbol{\Lambda}_{s_i})$, $\mathbf{w}_s = (\mathbf{w}_{s_1}', \mathbf{w}_{s_2}')'$, and $\mathbf{t_X} = (\mathbf{t}_{\mathbf{X}_1'}, \mathbf{t}_{\mathbf{X}_2'})'$, which generates the vector of calibrated weights

$$\mathbf{c}_s = \begin{pmatrix} \mathbf{w}_{s_1} \\ \mathbf{w}_{s_2} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\Lambda}_{s_1}\mathbf{X}_{s_1}(\mathbf{X}_{s_1}'\boldsymbol{\Lambda}_{s_1}\mathbf{X}_{s_1})^{-1}[\mathbf{t}_{\mathbf{X}_1} - \mathbf{X}_{s_1}'\mathbf{w}_{s_1}] \\ \boldsymbol{\Lambda}_{s_2}\mathbf{X}_{s_2}(\mathbf{X}_{s_2}'\boldsymbol{\Lambda}_{s_2}\mathbf{X}_{s_2})^{-1}[\mathbf{t}_{\mathbf{X}_2} - \mathbf{X}_{s_2}'\mathbf{w}_{s_2}] \end{pmatrix}. \tag{7}$$

Then the estimator (6) is obtained using $\mathbf{Z}_s'\mathbf{c}_s$, where $\mathbf{Z}_s = (\mathbf{Z}_{s_1}', \mathbf{Z}_{s_2}')'$, having multiplied before the regression the first component of $\mathbf{w}_s$ and $\mathbf{t_X}$ by ϕ and the second by $1-\phi$. The estimator (6) can be written as

$$\hat{Z}_s^{CR} = \hat{Z}^C + \hat{\boldsymbol{\beta}}_1\big[\phi\big(\mathbf{t}_{\mathbf{X}_1} - \hat{\mathbf{X}}_1\big)\big] + \hat{\boldsymbol{\beta}}_2\big[(1-\phi)\big(\mathbf{t}_{\mathbf{X}_2} - \hat{\mathbf{X}}_2\big)\big], \tag{8}$$

where $\hat{Z}^C = \phi\hat{Z}_1 + (1-\phi)\hat{Z}_2$ is the composite Horvitz–Thompson estimator of t_Z. It is seen that $\hat{Z}_s^{CR}$ is a GREG estimator derived by using orthogonal regressors for the two samples.

A choice of ϕ based on considerations of efficiency and practicality is

$$\phi = \frac{n_1/d_1}{n_1/d_1 + n_2/d_2} \quad \left(\approx \frac{V(\hat{Z}_2)}{V(\hat{Z}_1) + V(\hat{Z}_2)}\right), \tag{9}$$

where n_1 and n_2 are the sizes of the samples s_1 and s_2 and d_1 and d_2 are the design effects associated with s_1, s_2, and the variable z. This ϕ minimizes the estimated variance of $\hat{Z}^C$ if the ratio of the finite population correction factors $1 - f_i$, $f_i = n_i/N$, is close to 1, but it does not account for the relative amount of regression fit present in $\hat{Z}_1^R$ and $\hat{Z}_2^R$. (See Merkouris 2001 for a discussion of the merits of this choice of ϕ.)

Consider now the case in which all columns of the matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ represent common auxiliary variables, so that $\mathbf{t}_{\mathbf{X}_1} = \mathbf{t}_{\mathbf{X}_2} = \mathbf{t}_{\bar{\mathbf{X}}}$, and define the vectors of adjusted weights $\mathbf{w}_{s_1}^* = \phi\mathbf{w}_{s_1}$ and $\mathbf{w}_{s_2}^* = (1-\phi)\mathbf{w}_{s_2}$, where ϕ is as in (9). Then for $\mathbf{X}_s = (\mathbf{X}_{s_1}', \mathbf{X}_{s_2}')'$ and $\boldsymbol{\Lambda}_s^* = \operatorname{diag}(\boldsymbol{\Lambda}_{s_i}^*)$, with $\boldsymbol{\Lambda}_{s_1}^* = \phi\boldsymbol{\Lambda}_{s_1}$ and $\boldsymbol{\Lambda}_{s_2}^* = (1-\phi)\boldsymbol{\Lambda}_{s_2}$, the regression procedure for the combined sample generates the vector of calibrated weights

$$\mathbf{c}_s = \begin{pmatrix} \mathbf{w}_{s_1}^* \\ \mathbf{w}_{s_2}^* \end{pmatrix} + \begin{pmatrix} \boldsymbol{\Lambda}_{s_1}^*\mathbf{X}_{s_1} \\ \boldsymbol{\Lambda}_{s_2}^*\mathbf{X}_{s_2} \end{pmatrix} \big[\mathbf{X}_{s_1}'\boldsymbol{\Lambda}_{s_1}^*\mathbf{X}_{s_1} + \mathbf{X}_{s_2}'\boldsymbol{\Lambda}_{s_2}^*\mathbf{X}_{s_2}\big]^{-1} \times \big[\mathbf{t}_{\bar{\mathbf{X}}} - \big(\mathbf{X}_{s_1}'\mathbf{w}_{s_1}^* + \mathbf{X}_{s_2}'\mathbf{w}_{s_2}^*\big)\big]. \tag{10}$$

This gives rise to the composite regression estimator of the combined type,

$$\hat{Z}_c^{CR} = \hat{Z}^C + \hat{\boldsymbol{\beta}}^*[\mathbf{t}_{\bar{\mathbf{X}}} - \hat{\mathbf{X}}^C], \tag{11}$$

where $\hat{\boldsymbol{\beta}}^* = [\mathbf{Z}_{s_1}'\boldsymbol{\Lambda}_{s_1}^*\mathbf{X}_{s_1} + \mathbf{Z}_{s_2}'\boldsymbol{\Lambda}_{s_2}^*\mathbf{X}_{s_2}][\mathbf{X}_{s_1}'\boldsymbol{\Lambda}_{s_1}^*\mathbf{X}_{s_1} + \mathbf{X}_{s_2}'\boldsymbol{\Lambda}_{s_2}^* \times \mathbf{X}_{s_2}]^{-1}$ is the common slope for the combined sample. Clearly, $\hat{\mathbf{X}}_c^{CR} = \mathbf{t}_{\bar{\mathbf{X}}}$, but $\hat{\mathbf{X}}_{c1}^{CR} \neq \phi\mathbf{t}_{\bar{\mathbf{X}}}$ and $\hat{\mathbf{X}}_{c2}^{CR} \neq (1-\phi)\mathbf{t}_{\bar{\mathbf{X}}}$, where $\hat{\mathbf{X}}_{ci}^{CR}$ is the ith sample component of $\hat{\mathbf{X}}_c^{CR}$. The regression coefficient $\hat{\boldsymbol{\beta}}^*$ can be written as the weighted average $\hat{\boldsymbol{\beta}}^* = \hat{\boldsymbol{\beta}}_1\mathbf{D} + \hat{\boldsymbol{\beta}}_2(\mathbf{I}-\mathbf{D})$ of the regression coefficients $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ involved in (8), for $\mathbf{t}_{\mathbf{X}_1} = \mathbf{t}_{\mathbf{X}_2} = \mathbf{t}_{\bar{\mathbf{X}}}$, with weighting matrix $\mathbf{D} = \mathbf{X}_{s_1}'\boldsymbol{\Lambda}_{s_1}^*\mathbf{X}_{s_1}[\mathbf{X}_{s_1}'\boldsymbol{\Lambda}_{s_1}^*\mathbf{X}_{s_1} + \mathbf{X}_{s_2}'\boldsymbol{\Lambda}_{s_2}^*\mathbf{X}_{s_2}]^{-1}$. Thus the estimator (11) can be written as

$$\hat{Z}_c^{CR} = \hat{Z}^C + [\hat{\boldsymbol{\beta}}_1\mathbf{D} + \hat{\boldsymbol{\beta}}_2(\mathbf{I}-\mathbf{D})][\mathbf{t}_{\bar{\mathbf{X}}} - \hat{\mathbf{X}}^C]. \tag{12}$$

It will motivate subsequent development to compare the forms of ϕ and $\mathbf{D}$ and point out their complementary roles. Thus, whereas ϕ accounts for the relative effective sample size, the matrix $\mathbf{D}$ accounts in addition for the relative regression effect of the auxiliary information in $\mathbf{X}_{s_1}$ and $\mathbf{X}_{s_2}$. Also, whereas ϕ minimizes the estimated variance of the composite $\phi\hat{Z}_1 + (1-\phi)\hat{Z}_2$, the weighting matrix $\mathbf{D}$ minimizes the quadratic form $(\mathbf{I}-\mathbf{D}, \mathbf{D})\operatorname{diag}(\mathbf{X}_{s_i}'\boldsymbol{\Lambda}_{s_i}^*\mathbf{X}_{s_i})(\mathbf{I}-\mathbf{D}, \mathbf{D})'$, which is the estimated variance of the multivariate composite $(\mathbf{I}-\mathbf{D})\hat{\mathbf{X}}_1 + \mathbf{D}\hat{\mathbf{X}}_2$ when $\Lambda_{s_i}^*$ is replaced by $\boldsymbol{\Lambda}_{s_i}^\circ = \{(\pi_{kl} - \pi_k\pi_l)/\pi_k\pi_l\pi_{kl}\}_{k,l\in s_i}$; the latter follows from the recognition of $\mathbf{X}_{s_i}'\boldsymbol{\Lambda}_{s_i}^\circ\mathbf{X}_{s_i}$ as the estimated variance of $\hat{\mathbf{X}}_i$. The matrices $\mathbf{D}$ and $\mathbf{I}-\mathbf{D}$ may then be thought of as weighting $\hat{\mathbf{X}}_1$ and $\hat{\mathbf{X}}_2$, as well as $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$, approximately in proportion to the estimated variances of $\hat{\mathbf{X}}_1$ and $\hat{\mathbf{X}}_2$. It is shown in the next section that in a framework of optimal composite regression estimation and under certain sampling design conditions, ϕ and $1-\phi$ make two matrices analogous to $\mathbf{D}$ and $\mathbf{I}-\mathbf{D}$ exactly proportional to the estimated variances of two GREG estimators.

Both composite regression estimators $\hat{Z}_s^{CR}$ and $\hat{Z}_c^{CR}$ are generated conveniently by a single regression procedure based on a combined sample in which the sampling weights of the two component samples are adjusted proportionally to the effective sample sizes n_1/d_1 and n_2/d_2. In both cases, then, a single set of calibrated weights is produced that can be used to generate a composite estimate for any common variable of interest. It should be noted that because ϕ in (9) depends on a specific variable z only through the ratio d_2/d_1, the loss of efficiency of estimators of totals for other common variables should not be substantial when $d_1 \neq d_2$. A detailed comparison of $\hat{Z}_s^{CR}$ and $\hat{Z}_c^{CR}$ in terms of computational and statistical efficiency was given in earlier work (Merkouris 2001). A generalization to more than two samples is straightforward.

4. ALIGNED REGRESSION ESTIMATORS

4.1 A Combined Regression Procedure

4.1.1 Alignment of Regression Estimators. The regression procedures of Section 3 are effective in producing an efficient composite regression estimator for a single common variable of two independent sample surveys. In the most general setting involving a vector $\mathbf{z}$ of q common survey variables and different auxiliary variables for the two samples, we may consider a multivariate composite regression estimator for the total $\mathbf{t_Z}$ of the form $\boldsymbol{\Phi}\hat{\mathbf{Z}}_1^R + (\mathbf{I}-\boldsymbol{\Phi})\hat{\mathbf{Z}}_2^R$, preferably generated by a simultaneous regression procedure on the combined sample. The $q \times q$ matrix coefficient $\boldsymbol{\Phi}$ should preserve the internal weight consistency of the samples, satisfy the calibration constraints of each survey, and, ideally, maximize the efficiency of the composite estimator. In particular, it is desired that the coefficient $\boldsymbol{\Phi}$ incorporate the relative effect of regression fit in $\hat{\mathbf{Z}}_1^R$ and $\hat{\mathbf{Z}}_2^R$.

A suitable approach for determining such a coefficient $\boldsymbol{\Phi}$ is suggested on writing $\boldsymbol{\Phi}\hat{\mathbf{Z}}_1^R + (\mathbf{I}-\boldsymbol{\Phi})\hat{\mathbf{Z}}_2^R$ in the regression forms

$$\begin{aligned}\boldsymbol{\Phi}\hat{\mathbf{Z}}_1^R + (\mathbf{I}-\boldsymbol{\Phi})\hat{\mathbf{Z}}_2^R &= \hat{\mathbf{Z}}_1^R + (\mathbf{I}-\boldsymbol{\Phi})(\hat{\mathbf{Z}}_2^R - \hat{\mathbf{Z}}_1^R) \\ &= \hat{\mathbf{Z}}_2^R + \boldsymbol{\Phi}(\hat{\mathbf{Z}}_1^R - \hat{\mathbf{Z}}_2^R),\end{aligned} \tag{13}$$

where $\boldsymbol{\Phi}$ (or $\mathbf{I}-\boldsymbol{\Phi}$) can be viewed as a matrix of regression coefficients associated with the regression term $\mathbf{0} - (\hat{\mathbf{Z}}_2^R - \hat{\mathbf{Z}}_1^R)$. It is shown later that the regression procedure leading to (13) entails an extension of the calibration constraint system generating (6) that results in equating GREG estimates of totals for the q common variables between the two samples. Consider then the extended regression matrix and the corresponding vector of control totals,

$$\boldsymbol{\mathcal{X}}_s = \begin{pmatrix} \mathbf{X}_{s_1} & \mathbf{0} & \mathbf{Z}_{s_1} \\ \mathbf{0} & \mathbf{X}_{s_2} & -\mathbf{Z}_{s_2} \end{pmatrix}, \qquad \mathbf{t} = \begin{pmatrix} \mathbf{t}_{\mathbf{X}_1} \\ \mathbf{t}_{\mathbf{X}_2} \\ \mathbf{0} \end{pmatrix}, \tag{14}$$

and write $\boldsymbol{\mathcal{X}}_s$ in partition form as $\boldsymbol{\mathcal{X}}_s = (\mathbf{X}_s\ \boldsymbol{\mathcal{Z}}_s)$, where $\mathbf{X}_s$ is of dimension $n \times (p_1 + p_2)$. Now let $\mathbf{c}_{xs}$ denote the vector of calibrated weights (7) corresponding to the regression based on the matrix $\mathbf{X}_s$. Next let $\mathbf{L}_s = \boldsymbol{\Lambda}_s(\mathbf{I} - \mathbf{P}_{\mathbf{X}_s})$, with $\mathbf{P}_{\mathbf{X}_s} = \mathbf{X}_s(\mathbf{X}_s'\boldsymbol{\Lambda}_s\mathbf{X}_s)^{-1}\mathbf{X}_s'\boldsymbol{\Lambda}_s$, and note that $\mathbf{X}_s = \operatorname{diag}(\mathbf{X}_{s_i})$ implies $\mathbf{L}_s = \operatorname{diag}(\mathbf{L}_{s_i})$, where $\mathbf{L}_{s_i} = \boldsymbol{\Lambda}_{s_i}(\mathbf{I} - \mathbf{P}_{\mathbf{X}_{s_i}})$, in obvious notation for $\boldsymbol{\Lambda}_{s_i}$ and $\mathbf{P}_{\mathbf{X}_{s_i}}$. Then, in view of (4), for weight vector $\mathbf{w}_s = (\mathbf{w}_{s_1}', \mathbf{w}_{s_2}')'$ and weighting matrix $\boldsymbol{\Lambda}_s = \operatorname{diag}(\boldsymbol{\Lambda}_{s_i})$ the regression procedure based on the partitioned matrix $\boldsymbol{\mathcal{X}}_s$ generates the vector of calibrated weights

$$\begin{aligned} \mathbf{c}_s &= \mathbf{c}_{xs} + \mathbf{L}_s\boldsymbol{\mathcal{Z}}_s(\boldsymbol{\mathcal{Z}}_s'\mathbf{L}_s\boldsymbol{\mathcal{Z}}_s)^{-1}(\mathbf{0} - \boldsymbol{\mathcal{Z}}_s'\mathbf{c}_{xs}) \\ &= \begin{pmatrix} \mathbf{c}_{xs_1} \\ \mathbf{c}_{xs_2} \end{pmatrix} + \begin{pmatrix} \mathbf{L}_{s_1}\mathbf{Z}_{s_1} \\ -\mathbf{L}_{s_2}\mathbf{Z}_{s_2} \end{pmatrix} \left[\mathbf{Z}_{s_1}'\mathbf{L}_{s_1}\mathbf{Z}_{s_1} + \mathbf{Z}_{s_2}'\mathbf{L}_{s_2}\mathbf{Z}_{s_2}\right]^{-1} \\ &\qquad \times \left[(\mathbf{Z}_{s_2}'\mathbf{c}_{xs_2} - \mathbf{Z}_1'\mathbf{c}_{xs_1})\right]. \end{aligned} \tag{15}$$

It is easy to verify that the vector $\mathbf{c}_s$ satisfies all the calibration constraints, that is, $\mathbf{X}_{s_i}'\mathbf{c}_{s_i} = \mathbf{t}_{\mathbf{X}_i}$ and $\boldsymbol{\mathcal{Z}}_s'\mathbf{c}_s = \mathbf{Z}_{s_1}'\mathbf{c}_{s_1} - \mathbf{Z}_{s_2}'\mathbf{c}_{s_2} = \mathbf{0}$. For any noncommon single variable y_i associated with sample s_i,

$$\begin{aligned} \mathbf{Y}_{s_1}'\mathbf{c}_{s_1} &= \hat{Y}_1^R + \hat{\mathbf{B}}_{y_{s_1}}(\mathbf{I} - \hat{\mathbf{B}})[\hat{\mathbf{Z}}_2^R - \hat{\mathbf{Z}}_1^R] \qquad \text{and} \\ \mathbf{Y}_{s_2}'\mathbf{c}_{s_2} &= \hat{Y}_2^R - \hat{\mathbf{B}}_{y_{s_2}}\hat{\mathbf{B}}[\hat{\mathbf{Z}}_2^R - \hat{\mathbf{Z}}_1^R], \end{aligned} \tag{16}$$

where $\hat{\mathbf{B}}_{y_{s_i}} = \mathbf{Y}_{s_i}'\mathbf{L}_{s_i}\mathbf{Z}_{s_i}[\mathbf{Z}_{s_i}'\mathbf{L}_{s_i}\mathbf{Z}_{s_i}]^{-1}$ and $\hat{\mathbf{B}} = \mathbf{Z}_{s_2}'\mathbf{L}_{s_2}\mathbf{Z}_{s_2} \times [\mathbf{Z}_{s_1}'\mathbf{L}_{s_1}\mathbf{Z}_{s_1} + \mathbf{Z}_{s_2}'\mathbf{L}_{s_2}\mathbf{Z}_{s_2}]^{-1}$, whereas for the q-dimensional common variable $\mathbf{z}$,

$$\mathbf{Z}_{s_1}'\mathbf{c}_{s_1} = \hat{\mathbf{Z}}_1^R + (\mathbf{I} - \hat{\mathbf{B}})[\hat{\mathbf{Z}}_2^R - \hat{\mathbf{Z}}_1^R]$$

and

$$\mathbf{Z}_{s_2}'\mathbf{c}_{s_2} = \hat{\mathbf{Z}}_2^R - \hat{\mathbf{B}}[\hat{\mathbf{Z}}_2^R - \hat{\mathbf{Z}}_1^R],$$

so that

$$\mathbf{Z}_{s_1}'\mathbf{c}_{s_1} = \mathbf{Z}_{s_2}'\mathbf{c}_{s_2} = \hat{\mathbf{B}}\hat{\mathbf{Z}}_1^R + (\mathbf{I} - \hat{\mathbf{B}})\hat{\mathbf{Z}}_2^R, \tag{17}$$

where the matrix regression coefficient $\hat{\mathbf{B}}$ is the requisite matrix weighting coefficient in the multivariate composite GREG estimator (13). The regression setup (14) was used by Zieschang (1990) to align estimates of comparable totals between the Diary and Interview components of the American Consumer Expenditure Survey. Zieschang gave an explicit expression for the resulting estimator of the total $\mathbf{t}_{\mathbf{Z}}$ as the composite of two Horvitz–Thompson estimators in the special case when only $\mathbf{Z}$ is used in the regression, the coefficient $\hat{\mathbf{B}}$ then having the form $\hat{\mathbf{B}} = \mathbf{Z}_{s_2}'\boldsymbol{\Lambda}_{s_2}\mathbf{Z}_{s_2}[\mathbf{Z}_{s_1}'\boldsymbol{\Lambda}_{s_1}\mathbf{Z}_{s_1} + \mathbf{Z}_{s_2}'\boldsymbol{\Lambda}_{s_2}\mathbf{Z}_{s_2}]^{-1}$.

The regression setup that generates (15) is a mixture of the two regression setups of Section 3, and without the matrices of auxiliary variables $\mathbf{X}_{s_1}$ and $\mathbf{X}_{s_2}$ it is essentially a variation of the setup generating (10). Whereas the procedures of Section 3 were based on an initial composite adjustment of the sampling weights, through ϕ, composition in the present approach is achieved by the additional composite constraint, so that each sample's weights are further adjusted to make composites of the regression estimates $\hat{\mathbf{Z}}_1^R$ and $\hat{\mathbf{Z}}_2^R$. It follows from (15) that the calibration weight vector $\mathbf{c}_s$ takes the "residual" form $\mathbf{c}_s = \mathbf{H}_s\mathbf{c}_{xs}$, where $\mathbf{H}_s = \mathbf{I} - \mathbf{L}_s\boldsymbol{\mathcal{Z}}_s(\boldsymbol{\mathcal{Z}}_s'\mathbf{L}_s\boldsymbol{\mathcal{Z}}_s)^{-1}\boldsymbol{\mathcal{Z}}_s'$ is the multiplicative "composite" weight adjustment due to the additional constraints. In addition to inducing consistency between estimates of totals for the common characteristics, each sample's calibrated weights incorporate auxiliary information from the other sample, and thus they can produce more efficient estimators for both common and noncommon characteristics.

4.1.2 The Efficiency of the Combined Procedure. Regarding the efficiency of estimators of totals for the common characteristics, it is interesting to note that in general the matrix $\hat{\mathbf{B}}$ does not minimize the estimated variance of the composite $\hat{\mathbf{B}}\hat{\mathbf{Z}}_1^R + (\mathbf{I} - \hat{\mathbf{B}})\hat{\mathbf{Z}}_2^R$, in the partial order of nonnegative definite matrices, and thus it is not the optimal choice in that sense. Instead, the matrix $\hat{\mathbf{B}}$ is optimal in the sense of minimizing the quadratic form in the sample residuals corresponding to the regression of $\mathbf{Z}_s$ on $\boldsymbol{\mathcal{X}}_s$ (see App. A for the derivation),

$$4\begin{pmatrix} \mathbf{Z}_{s_1}\hat{\mathbf{B}}' \\ \mathbf{Z}_{s_2}(\mathbf{I} - \hat{\mathbf{B}}') \end{pmatrix}' \begin{pmatrix} \mathbf{L}_{s_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_{s_2} \end{pmatrix} \begin{pmatrix} \mathbf{Z}_{s_1}\hat{\mathbf{B}}' \\ \mathbf{Z}_{s_2}(\mathbf{I} - \hat{\mathbf{B}}') \end{pmatrix}, \tag{18}$$

which is identical to

$$4\begin{pmatrix} \hat{\mathbf{B}}' \\ \mathbf{I} - \hat{\mathbf{B}}' \end{pmatrix}' \begin{pmatrix} \mathbf{Z}_{s_1}'\mathbf{L}_{s_1}\mathbf{Z}_{s_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_{s_2}'\mathbf{L}_{s_2}\mathbf{Z}_{s_2} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{B}}' \\ \mathbf{I} - \hat{\mathbf{B}}' \end{pmatrix}. \tag{19}$$

Considering that for the regression of $\mathbf{Z}_{s_i}$ on $\mathbf{X}_{s_i}$ the regression coefficient is $\hat{\boldsymbol{\beta}}_i = \mathbf{Z}_{s_i}'\boldsymbol{\Lambda}_{s_i}\mathbf{X}_{s_i}(\mathbf{X}_{s_i}'\boldsymbol{\Lambda}_{s_i}\mathbf{X}_{s_i})^{-1}$, it follows readily that $\mathbf{Z}_{s_i}'\mathbf{L}_{s_i}\mathbf{Z}_{s_i}$ can be written as the quadratic form in the regression residuals,

$$\mathbf{Z}_{s_i}'\mathbf{L}_{s_i}\mathbf{Z}_{s_i} = (\mathbf{Z}_{s_i} - \mathbf{X}_{s_i}\hat{\boldsymbol{\beta}}_i')'\boldsymbol{\Lambda}_{s_i}(\mathbf{Z}_{s_i} - \mathbf{X}_{s_i}\hat{\boldsymbol{\beta}}_i'). \tag{20}$$

The quantity $\mathbf{Z}_{s_i}'\mathbf{L}_{s_i}\mathbf{Z}_{s_i}$ with the matrix $\boldsymbol{\Lambda}_{s_i}^\circ = \{(\pi_{kl} - \pi_k\pi_l)/\pi_k\pi_l\pi_{kl}\}_{k,l\in s_i}$ in place of $\boldsymbol{\Lambda}_{s_i}$ in the quadratic form, but retaining $\boldsymbol{\Lambda}_{s_i}$ in $\hat{\boldsymbol{\beta}}_i$, would be the estimated approximate variance $\widehat{AV}(\hat{\mathbf{Z}}_i^R) = \mathbf{Z}_{s_i}'\mathbf{L}_{s_i}^\circ\mathbf{Z}_{s_i} = (\mathbf{Z}_{s_i} - \mathbf{X}_{s_i}\hat{\boldsymbol{\beta}}_i')'\boldsymbol{\Lambda}_{s_i}^\circ(\mathbf{Z}_{s_i} - \mathbf{X}_{s_i}\hat{\boldsymbol{\beta}}_i')$ (see Sec. 2). In that case, $\hat{\mathbf{B}}$ would indeed have the variance-minimizing property, as follows from (19). In the present GREG formulation, the quadratic forms $\mathbf{Z}_{s_i}'\mathbf{L}_{s_i}\mathbf{Z}_{s_i}$ are independent design-consistent estimators of the corresponding population quantities $\mathbf{Z}'\mathbf{L}_i\mathbf{Z}$, where $\mathbf{L}_i = \mathbf{I} - \mathbf{X}_i(\mathbf{X}_i'\mathbf{X}_i)^{-1}\mathbf{X}_i'$ and the index i indicates possibly different auxiliary regression matrices associated with the two samples; this design consistency property entails $\mathbf{Z}_{s_i}'\mathbf{L}_{s_i}\mathbf{Z}_{s_i} = \mathbf{Z}'\mathbf{L}_i\mathbf{Z}$ when $s_1 = s_2 = U$, for fixed sample sizes. When the same auxiliary variables are used in both samples, the quadratic forms $\mathbf{Z}'\mathbf{L}_i\mathbf{Z}$ are identical, which indicates that $\hat{\mathbf{B}}$ is an estimator of $(1/2)\mathbf{I}$ and thus does not account for any difference in effective sample size. Thus, in connection with the relative precision of $\hat{\mathbf{Z}}_1^R$ and $\hat{\mathbf{Z}}_2^R$, the data-dependent composite coefficient $\hat{\mathbf{B}}$ incorporates only the relative regression fit, rendering the composite estimator (17) inefficient.

In general, for any variable y_i and for nonidentical vectors of auxiliary variables for the two samples, the dependence of the efficiency of the composite GREG procedure on the differential in effective sample sizes between the two samples becomes immediately evident on contrasting the generalized regression coefficient $\hat{\mathbf{B}}_{y_{s_1}}(\mathbf{I} - \hat{\mathbf{B}}) = \mathbf{Y}_{s_1}'\mathbf{L}_{s_1}\mathbf{Z}_{s_1}[\mathbf{Z}_{s_1}'\mathbf{L}_{s_1}\mathbf{Z}_{s_1} +$

$\mathbf{Z}_{s_2}'\mathbf{L}_{s_2}\mathbf{Z}_{s_2}]^{-1}$ in (16) with the optimal regression coefficient $\widehat{\text{cov}}(\hat{Y}_1^R, \hat{\mathbf{Z}}_1^R)[\hat{V}(\hat{\mathbf{Z}}_1^R) + \hat{V}(\hat{\mathbf{Z}}_2^R)]^{-1}$ that minimizes the variance of $\hat{Y}_1^R + \mathbf{K}[\hat{\mathbf{Z}}_2^R - \hat{\mathbf{Z}}_1^R]$ for an arbitrary $\mathbf{K}$.

4.2 A Modified Combined Regression Procedure

A modification of the GREG procedure intended to account for the differential in effective sample size between the two samples involves the replacement of the bilinear and quadratic forms $\mathbf{Y}_{s_i}'\mathbf{L}_{s_i}\mathbf{Z}_{s_i}$ and $\mathbf{Z}_{s_i}'\mathbf{L}_{s_i}\mathbf{Z}_{s_i}$ by the respective mean forms $(1/\tilde{n}_i)\mathbf{Y}_{s_i}'\mathbf{L}_{s_i}\mathbf{Z}_{s_i}$ and $(1/\tilde{n}_i)\mathbf{Z}_{s_i}'\mathbf{L}_{s_i}\mathbf{Z}_{s_i}$, where $\tilde{n}_i = n_i/d_i$ are the effective sample sizes. Specifically, the modification entails the scaling adjustment of the entries of the weighting matrix $\mathbf{\Lambda}_{s_i}$ by $1/\tilde{n}_i$. This scaling adjustment is to account for the differential in effective sample size in the weighting of the regression residuals, so that the adjusted coefficients $\hat{\mathbf{B}}$, $\hat{\mathbf{B}}_{y_{s_1}}(\mathbf{I}-\hat{\mathbf{B}})$ and $\hat{\mathbf{B}}_{y_{s_2}}\hat{\mathbf{B}}$ approximate the optimal coefficients in (16) and (17). Clearly, this adjustment does not affect the regression coefficients $\hat{\boldsymbol{\beta}}_i$. In fact, the composite regression estimators (16) and (17) are affected by this adjustment only through the estimated regression coefficient $\hat{\mathbf{B}}$. It should be pointed out that if both n_1 and n_2 are allowed to grow as large as the population size, then the scaling factors $1/\tilde{n}_1$ and $1/\tilde{n}_2$ will be equal and cancel out in the expression of $\hat{\mathbf{B}}$, thus preserving the design consistency of $\hat{\mathbf{B}}$. These scaling factors also preserve the internal weight consistency for each sample, and the adjusted weights $\tilde{\mathbf{c}}_s = \tilde{\mathbf{H}}_s\mathbf{c}_{xs}$ (in obvious notation for $\tilde{\mathbf{H}}_s$) still satisfy the regression constraints while aligning estimates of $\mathbf{t_Z}$ from the two samples.

It is interesting to note that the adjusted $\hat{\mathbf{B}}$, henceforth denoted by $\tilde{\mathbf{B}}$, can be written as $\tilde{\mathbf{B}} = \phi\mathbf{Z}_{s_2}'\mathbf{L}_{s_2}\mathbf{Z}_{s_2}[(1-\phi)\times \mathbf{Z}_{s_1}'\mathbf{L}_{s_1}\mathbf{Z}_{s_1} + \phi\mathbf{Z}_{s_2}'\mathbf{L}_{s_2}\mathbf{Z}_{s_2}]^{-1}$, where ϕ is the same as in (9); note then the analogy of $\tilde{\mathbf{B}}$ with the weighting matrix $\mathbf{D}$ in (12). With $\tilde{n}_i$ relaxed to n_i, this weight adjustment, uniform for each sample, yields the optimal regression coefficient in both (16) and (17) in the setting of the first of the following two theorems. The proof of the theorem is given in Appendix B.

Consider first the following set of conditions:

C1. The ratio of the finite population correction factors for the two surveys is approximately equal to 1.

C2. The ratio $N_{il}/(N_{il}-1)$, where N_{il} is the size of stratum l for survey i, is approximately equal to 1.

C3. The sample size n_{il} in stratum l for survey i is proportional to stratum size.

Theorem 1. (a) Suppose that $\mathbf{1} = \mathbf{X}_i\mathbf{h}_i$, for a constant p_i-vector $\mathbf{h}_i$. Then, under simple random sampling for each survey, and under condition C1,

$$\phi\mathbf{Z}'\mathbf{L}_2\mathbf{Z}[(1-\phi)\mathbf{Z}'\mathbf{L}_1\mathbf{Z} + \phi\mathbf{Z}'\mathbf{L}_2\mathbf{Z}]^{-1} = AV(\hat{\mathbf{Z}}_2^R)[AV(\hat{\mathbf{Z}}_1^R) + AV(\hat{\mathbf{Z}}_2^R)]^{-1}. \quad (21)$$

Furthermore, for large sample sizes n_i,

$$\tilde{\mathbf{B}} \approx \widehat{AV}(\hat{\mathbf{Z}}_2^R)\big[\widehat{AV}(\hat{\mathbf{Z}}_1^R) + \widehat{AV}(\hat{\mathbf{Z}}_2^R)\big]^{-1}. \quad (22)$$

Analogous results hold also for the adjusted coefficients $\hat{\mathbf{B}}_{y_{s_1}}(\mathbf{I}-\tilde{\mathbf{B}})$ and $\hat{\mathbf{B}}_{y_{s_2}}\tilde{\mathbf{B}}$. For instance,

$$\begin{aligned}\hat{\mathbf{B}}_{y_{s_1}}&(\mathbf{I}-\tilde{\mathbf{B}}) \\ &= (1-\phi)\mathbf{Y}_{s_1}'\mathbf{L}_{s_1}\mathbf{Z}_{s_1}\big[(1-\phi)\mathbf{Z}_{s_1}'\mathbf{L}_{s_1}\mathbf{Z}_{s_1} + \phi\mathbf{Z}_{s_2}'\mathbf{L}_{s_2}\mathbf{Z}_{s_2}\big]^{-1} \\ &\approx \widehat{AC}(\hat{Y}_1^R, \hat{\mathbf{Z}}_1^R)\big[\widehat{AV}(\hat{\mathbf{Z}}_1^R) + \widehat{AV}(\hat{\mathbf{Z}}_2^R)\big]^{-1}, \quad (23)\end{aligned}$$

where AC denotes approximate covariance.

(b) Suppose that $\mathbf{1} = \mathbf{X}_i\mathbf{h}_i$. For each survey, assume stratified simple random sampling, not necessarily with the same stratification, with separate regression for each stratum. Then, for large strata samples, the results of (a) hold if the conditions C1–C3 hold.

(c) For each survey, assume stratified simple random sampling with combined regression with an intercept for each stratum; that is, with auxiliary row vector corresponding to unit k for survey i, $\mathbf{x}_{ik}' = (\mathbf{x}_{oik}', d_{1k}, \ldots, d_{H_ik})$, where $\mathbf{x}_{oi}'$ denotes the row p_i-vector of the core auxiliary variables and d_{lk} is the indicator variable of the membership of population unit k to stratum $l = 1, \ldots, H_i$. Suppose that $1 \neq \mathbf{x}_{oik}'\mathbf{h}_i$ for at least one unit k, for any constant p_i-vector $\mathbf{h}_i$, so that the matrix $\mathbf{X}_i'\mathbf{X}_i$ is not singular. Then if conditions C1–C3 hold, the results of (a) also hold. Also, $\hat{\mathbf{Z}}_i^R = \hat{\mathbf{Z}}_i^O$ and $\hat{Y}_i^R = \hat{Y}_i^O$.

A nonuniform weight adjustment for each sample yields the optimal composite regression coefficient in (16) and (17) under the conditions of the following theorem, the proof of which is given in Appendix C.

Theorem 2. (a) Suppose that $\mathbf{1} = \mathbf{X}_i\mathbf{h}_i$. For each survey, assume stratified simple random sampling, not necessarily with the same stratification, with separate regression for each stratum. Then, if finite population correction factors are ignored and condition C2 holds, the scaling adjustment of the entries of the weighting submatrix $\mathbf{\Lambda}_{il}$ by the inverse sampling fraction N_{il}/n_{il} implies that

$$\mathbf{Z}'\mathbf{\Phi}_2\mathbf{L}_2\mathbf{Z}[\mathbf{Z}'\mathbf{\Phi}_1\mathbf{L}_1\mathbf{Z} + \mathbf{Z}'\mathbf{\Phi}_2\mathbf{L}_2\mathbf{Z}]^{-1} = AV(\hat{\mathbf{Z}}_2^R)[AV(\hat{\mathbf{Z}}_1^R) + AV(\hat{\mathbf{Z}}_2^R)]^{-1}, \quad (24)$$

where $\mathbf{\Phi}_i = \text{diag}((N_{il}/n_{il})\mathbf{I}_{N_{il}})$. Furthermore, for large sample sizes n_{il}, the adjusted coefficient $\tilde{\mathbf{B}} = \mathbf{Z}_{s_2}'\mathbf{\Phi}_{s_2}\mathbf{L}_{s_2}\mathbf{Z}_{s_2}[\mathbf{Z}_{s_1}'\mathbf{\Phi}_{s_1}\times \mathbf{L}_{s_1}\mathbf{Z}_{s_1} + \mathbf{Z}_{s_2}'\mathbf{\Phi}_{s_2}\mathbf{L}_{s_2}\mathbf{Z}_{s_2}]^{-1}$, with $\mathbf{\Phi}_{s_i} = \text{diag}((N_{il}/n_{il})\mathbf{I}_{n_{il}})$, satisfies

$$\tilde{\mathbf{B}} \approx \widehat{AV}(\hat{\mathbf{Z}}_2^R)\big[\widehat{AV}(\hat{\mathbf{Z}}_1^R) + \widehat{AV}(\hat{\mathbf{Z}}_2^R)\big]^{-1}. \quad (25)$$

Analogous results hold also for the adjusted coefficients $\tilde{\mathbf{B}}_{y_{s_1}}(\mathbf{I}-\tilde{\mathbf{B}})$ and $\tilde{\mathbf{B}}_{y_{s_2}}\tilde{\mathbf{B}}$.

(b) For each survey, assume stratified simple random sampling with combined regression with an intercept for each stratum, that is, with auxiliary row vector corresponding to unit k for survey i, $(\mathbf{x}_{oik}', d_{1k}, \ldots, d_{H_ik})$, where d_{lk} is the indicator variable of the membership of population unit k to stratum $l = 1, \ldots, H_i$. Suppose that $1 \neq \mathbf{x}_{oik}'\mathbf{h}_i$ for at least one unit k, for any p_i-vector $\mathbf{h}_i$. Then, if finite population correction factors are ignored and condition C2 holds, the scaling adjustment of the entries of the weighting submatrix $\mathbf{\Lambda}_{il}$ by the inverse sampling fraction N_{il}/n_{il} implies that

$$\mathbf{E}_2[\mathbf{E}_1 + \mathbf{E}_2]^{-1} = AV(\hat{\mathbf{Z}}_2^R)[AV(\hat{\mathbf{Z}}_1^R) + AV(\hat{\mathbf{Z}}_2^R)]^{-1}, \quad (26)$$

where $\mathbf{E}_i = \mathbf{Z}'(\mathbf{I}-\mathbf{P}_{\mathbf{X}_i})'\mathbf{\Phi}_i(\mathbf{I}-\mathbf{P}_{\mathbf{X}_i})\mathbf{Z}$ and $\mathbf{P}_{\mathbf{X}_i} = \mathbf{X}_i(\mathbf{X}_i'\mathbf{X}_i)^{-1}\times \mathbf{X}_i'$. Furthermore, for large sample sizes n_{il} the adjusted coefficient $\tilde{\mathbf{B}} = \mathbf{Z}_{s_2}'\mathbf{\Phi}_{s_2}\mathbf{L}_{s_2}\mathbf{Z}_{s_2}[\mathbf{Z}_{s_1}'\mathbf{\Phi}_{s_1}\mathbf{L}_{s_1}\mathbf{Z}_{s_1} + \mathbf{Z}_{s_2}'\mathbf{\Phi}_{s_2}\mathbf{L}_{s_2}\mathbf{Z}_{s_2}]^{-1}$, with $\mathbf{L}_{s_i} = \mathbf{\Lambda}_{s_i}(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_i}})$ and $\mathbf{P}_{\mathbf{X}_{s_i}} = \mathbf{X}_{s_i}(\mathbf{X}_{s_i}'\mathbf{\Phi}_{s_i}\mathbf{\Lambda}_{s_i}\mathbf{X}_{s_i})^{-1}\times \mathbf{X}_{s_i}'\mathbf{\Phi}_{s_i}\mathbf{\Lambda}_{s_i}$, satisfies

$$\tilde{\mathbf{B}} \approx \widehat{AV}(\hat{\mathbf{Z}}_2^R)\big[\widehat{AV}(\hat{\mathbf{Z}}_1^R) + \widehat{AV}(\hat{\mathbf{Z}}_2^R)\big]^{-1}. \quad (27)$$

Analogous results hold also for the adjusted coefficients $\tilde{\mathbf{B}}_{y_{s_1}}(\mathbf{I}-\tilde{\mathbf{B}})$ and $\tilde{\mathbf{B}}_{y_{s_2}}\tilde{\mathbf{B}}$. Furthermore, $\hat{\mathbf{Z}}_i^R = \hat{\mathbf{Z}}_i^O$ and $\hat{Y}_i^R = \hat{Y}_i^O$.

Remark 1. The notation $\tilde{\mathbf{B}}_{y_{s_i}}$ in Theorem 2 indicates that the nonuniform weight adjustment in each of the weighting matrices results also in adjusted coefficients $\hat{\mathbf{B}}_{y_{s_i}}$. The regression coefficients $\hat{\boldsymbol{\beta}}_i$ are not affected by this adjustment in Theorem 2(a), but they are in Theorem 2(b), resulting in $\hat{\mathbf{Z}}_i^R = \hat{\mathbf{Z}}_i^O$ and $\hat{Y}_i^R = \hat{Y}_i^O$. It is noteworthy that in each of the foregoing two theorems, the same adjustment gives the true and estimated approximately optimal composite regression coefficient in (16) and (17). It also preserves the design consistency (in the sense specified in Sec. 4.1) of $\hat{\mathbf{B}}_{y_{s_i}}$ and $\hat{\mathbf{B}}$.

Remark 2. The findings $\hat{\mathbf{Z}}_i^R = \hat{\mathbf{Z}}_i^O$ and $\hat{Y}_i^R = \hat{Y}_i^O$ in Theorem 1(c) and Theorem 2(b) imply that combined regression over strata in each of the two surveys yields optimal composites of optimal regression estimators in (16) and (17). It is to be noted that the particular result $\hat{Y}_i^R = \hat{Y}_i^O$, with the exact regression coefficient in $\hat{Y}_i^O$, has been established also by Särndal (1996) and Montanari (2000) in a general regression context using a different approach based on the specification of a suitable matrix of generalized least squares weights, rather than on the adjustment of the weighting matrix proposed in this article.

Remark 3. The condition $\mathbf{1} = \mathbf{X}_i\mathbf{h}_i$ in Theorems 1 and 2 is customarily satisfied in surveys that use GREG estimation. The condition of an intercept for each stratum in the combined-regression case is also usually satisfied in general sampling designs, on efficiency considerations, when the strata sizes are known.

Remark 4. The results of Theorems 1 and 2 also hold for Bernoulli sampling, with probabilities of inclusion in the two samples being $\pi_i = n_i/N$ ($\pi_{il} = n_{il}/N_{il}$ in the stratification case), where n_i (or n_{il}) is the expected sample size to be used in the scaling adjustment. Notably, because the matrices $\boldsymbol{\Lambda}_i^\circ$ and $\boldsymbol{\Lambda}_{s_i}^\circ$ are diagonal, the conditions on the design matrix $\mathbf{X}_i$ are not necessary, and the proofs are much simpler. Condition 2 is also not necessary.

Remark 5. The consideration of an optimal regression coefficient in the composite regression estimators (16) and (17) broadens the notion of the standard optimal regression estimator Y^O to an estimator of the type defined by (5) in terms of GREG estimators. For the specific designs, Theorems 1 and 2 give the expression for the approximately optimal coefficient $\widehat{AC}(\hat{Y}^R, \hat{\mathbf{Z}}^R)[\widehat{AV}(\hat{\mathbf{Z}}^R)]^{-1}$ in this regression estimator. Moreover, Theorem 1(c) and Theorem 2(b) essentially introduce the extended optimal regression estimator $\hat{\mathbf{Y}}^{EO} = \hat{\mathbf{Y}}^O + \mathbf{Y}_s'\mathbf{L}_s^\circ\mathbf{Z}_s(\mathbf{Z}_s'\mathbf{L}_s^\circ\mathbf{Z}_s)^{-1}(\mathbf{t_Z} - \hat{\mathbf{Z}}^O)$, analogous to the extended GREG $\hat{\mathbf{Y}}^{ER}$ in (5), thus establishing, with a simple adjustment of the weighting matrix in the GREG procedure, the identity $\hat{\mathbf{Y}}^{ER} = \hat{\mathbf{Y}}^{EO}$.

Remark 6. For general complex sampling designs, possibly different for the two samples, the matrix $\tilde{\mathbf{B}}$ as a weighting coefficient in the multivariate composite regression estimator (13) incorporating the effective sample sizes $\tilde{n}_i$ provides a multivariate generalization of ϕ in (9). The coefficient $\tilde{\mathbf{B}}$ also incorporates auxiliary information through the regression on $\mathbf{X}_{s_1}$ and $\mathbf{X}_{s_2}$, as well as the correlations among the q components of $\mathbf{z}$. For such general designs, however, the coefficient $\tilde{\mathbf{B}}$ may be somewhat suboptimal, because it may not precisely reflect the relative interaction of design and regression effects between the two surveys. In any case, the proposed adjustments of the weighting matrices can be easily refined if they are made at the lowest population level at which both surveys are calibrated (e.g., at a subnational-area level).

4.3 A Separate Regression Procedure

For an estimation procedure that generates a multivariate composite regression estimator of the general form $\hat{\mathbf{Z}}^{CR} = \boldsymbol{\Phi}\hat{\mathbf{Z}}_1^R + (\mathbf{I} - \boldsymbol{\Phi})\hat{\mathbf{Z}}_2^R$ to satisfy the internal weight consistency for each sample while aligning the derived common estimates, Renssen and Nieuwenbroek (1997) proposed using the composite $\hat{\mathbf{Z}}^{CR}$ as a control total in an extended regression procedure, separately for each sample, with the common variable $\mathbf{z}$ as an additional regression variable. The resulting regression estimators generated by the two samples for variables y_i can take the partitioned regression form $\hat{Y}_i^R + \mathbf{Y}_{s_i}'\mathbf{L}_{s_i}\mathbf{Z}_{s_i}(\mathbf{Z}_{s_i}'\mathbf{L}_{s_i}\mathbf{Z}_{s_i})^{-1}[\hat{\mathbf{Z}}^{CR} - \hat{\mathbf{Z}}_i^R]$, which, on expanding $\hat{\mathbf{Z}}^{CR}$ and using notation from (16), becomes

$$\begin{aligned} \hat{Y}_1^{CR} &= \hat{Y}_1^R + \hat{\mathbf{B}}_{y_{s_1}}(\mathbf{I} - \boldsymbol{\Phi})[\hat{\mathbf{Z}}_2^R - \hat{\mathbf{Z}}_1^R] \qquad \text{and} \\ \hat{Y}_2^{CR} &= \hat{Y}_2^R - \hat{\mathbf{B}}_{y_{s_2}}\boldsymbol{\Phi}[\hat{\mathbf{Z}}_2^R - \hat{\mathbf{Z}}_1^R]. \end{aligned} \tag{28}$$

One matrix coefficient considered by Renssen and Nieuwenbroek (1997) is $\boldsymbol{\Phi} = \gamma\mathbf{I}$, where γ, $0 \le \gamma \le 1$, is a crude indicator of the confidence in one estimator compared with the other. The choice of γ might depend on indicators for several survey errors, including sampling errors. The special case $\gamma = n_1/(n_1 + n_2)$ accounts for the difference in sample size. On efficiency grounds, these authors suggested the choice $\hat{\boldsymbol{\Phi}}^\circ = \widehat{AV}(\hat{\mathbf{Z}}_2^R)[\widehat{AV}(\hat{\mathbf{Z}}_1^R) + \widehat{AV}(\hat{\mathbf{Z}}_2^R)]^{-1}$, which minimizes the estimated approximate variance of $\hat{\mathbf{Z}}^{CR}$.

It is important to note that for $\boldsymbol{\Phi} = \hat{\boldsymbol{\Phi}}^\circ$, the regression procedure leading to (28) is a hybrid of GREG and optimal regression. This becomes clear on writing the optimal regression coefficient $\hat{\mathbf{B}}_{y_{s_i}}^\circ$ in $\hat{Y}_i^R + \hat{\mathbf{B}}_{y_{s_i}}^\circ[\hat{\mathbf{Z}}_2^R - \hat{\mathbf{Z}}_1^R]$ as $\widehat{\text{cov}}(\hat{Y}_i^R, \hat{\mathbf{Z}}_i^R)[\hat{V}(\hat{\mathbf{Z}}_i^R)]^{-1}\hat{V}(\hat{\mathbf{Z}}_i^R)[\hat{V}(\hat{\mathbf{Z}}_1^R) + \hat{V}(\hat{\mathbf{Z}}_2^R)]^{-1}$.

4.4 Comparison of the Combined and Separate Regression Procedures

In addition to aligning the resulting more efficient estimators for common variables of two sample surveys, the composite regression techniques of Sections 4.1, 4.2, and 4.3 may produce more efficient estimators for noncommon variables, depending on the strength of their correlation with the common ones. For comparison, we view Zieschang's estimator (Z), the modified Zieschang estimator (MZ) proposed in this article, and the estimator of Renssen and Nieuwenbroek (RN) as special cases of the regression estimator (28). Thus for $\boldsymbol{\Phi} = \hat{\mathbf{B}}$, the estimators (28) are identical to Zieschang's composite GREG estimators in (16), as Renssen and Nieuwenbroek (1997) also showed, whereas for $\boldsymbol{\Phi} = \tilde{\mathbf{B}}$ (and for $\hat{\mathbf{B}}_{y_{s_i}}$ replaced by $\tilde{\mathbf{B}}_{y_{s_i}}$ in the setting of Thm. 2), they become the MZ composite GREG estimators.

However, the regression procedures, as well as the properties of the three estimators, are different.

The RN method was proposed as an alternative to the Z method to account for the differences in sample size. The RN method does this implicitly through the coefficients $\mathbf{I}-\hat{\mathbf{\Phi}}^\circ$ and $\hat{\mathbf{\Phi}}^\circ$, which are proportional to the estimated approximate variances of $\hat{\mathbf{Z}}_1^R$ and $\hat{\mathbf{Z}}_2^R$. In contrast, the MZ method accounts for the relative effective sample size explicitly through the adjustment ϕ incorporated in the coefficient $\tilde{\mathbf{B}}$.

Regarding the common variable $\mathbf{z}$, the RN and MZ estimators coincide under the conditions of Theorems 1 and 2. Note, however, that for any noncommon variables y_i, the MZ estimators are optimal, whereas, in view of (28) and the note at the end of Section 4.3, the RN estimators are not. In more general settings than those of Theorems 1 and 2, for the common variable $\mathbf{z}$, the RN estimator sustains the optimality of the coefficient $\hat{\mathbf{\Phi}}^\circ$ and thus is more efficient than the MZ estimator. However, by incorporating the effective sample size, the coefficient $\tilde{\mathbf{B}}$ should be close to $\hat{\mathbf{\Phi}}^\circ$, and thus the difference in efficiency between the two estimators should be small. For any noncommon variables y_i, neither the RN coefficients $\hat{\mathbf{B}}_{y_{s_1}}(\mathbf{I}-\hat{\mathbf{\Phi}}^\circ)$, $\hat{\mathbf{B}}_{y_{s_2}}\hat{\mathbf{\Phi}}^\circ$ nor the MZ coefficients $\hat{\mathbf{B}}_{y_{s_1}}(\mathbf{I}-\tilde{\mathbf{B}})$, $\hat{\mathbf{B}}_{y_{s_2}}\tilde{\mathbf{B}}$ are optimal, but it seems that the MZ (adjusted GREG) coefficients may be closer to the optimal coefficients $\hat{\mathbf{B}}^\circ_{y_{s_i}}$ than the hybrid RN coefficients, the optimality of which is limited to the factor $\hat{\mathbf{\Phi}}^\circ$.

The estimated coefficient $\hat{\mathbf{\Phi}}^\circ$ can be unstable if the degrees of freedom associated with $\widehat{AV}(\hat{\mathbf{Z}}_1^R)$ and $\widehat{AV}(\hat{\mathbf{Z}}_2^R)$ are small. This is akin to the issue of relative stability of estimates of the "optimal" versus the least squares coefficient in a regression estimator (for elaboration on this, see Rao 1994; Montanari 1998). As a consequence, the RN method may result in less-stable composite estimators than the MZ method if the number of auxiliary variables is large relative to the sample size.

Regarding the generation of composite weights for the two samples, in the RN method the formation of the multivariate composite regression estimator $\hat{\mathbf{Z}}^{CR}=\hat{\mathbf{\Phi}}^\circ\hat{\mathbf{Z}}_1^R+(\mathbf{I}-\hat{\mathbf{\Phi}}^\circ)\hat{\mathbf{Z}}_2^R$ requires in a first step the laborious computation of the variance matrices $\widehat{AV}(\hat{\mathbf{Z}}_1^R)$ and $\widehat{AV}(\hat{\mathbf{Z}}_2^R)$—although the estimated covariance terms are not of interest per se. This composite estimator is then used as an additional control total in a separate regression procedure for each sample to satisfy the consistency requirement for all calibration variables. In contrast, the MZ method generates the vector of composite weights, $\tilde{\mathbf{c}}_s$, simultaneously for the two samples in one regression procedure based on the setup (14) and the adjusted weighting matrices $(1/\tilde{n}_i)\mathbf{\Lambda}_{s_i}$. Alternatively, the computations may be simplified by directly calibrating the vector of regression weights, $\mathbf{c}_{xs}$, if it is available from an earlier regression procedure, using $\tilde{\mathbf{c}}_s=\tilde{\mathbf{H}}_s\mathbf{c}_{xs}$. In the MZ procedure the composite coefficient $\tilde{\mathbf{B}}$ is implicitly generated as a regression coefficient [see (17)]. This computational feature becomes relatively more advantageous when composition is carried out at a stratum level. The MZ regression method is thus considerably simpler than the RN method for producing both the aligned and the nonaligned regression estimates.

For the variance of the estimator (28), Renssen and Nieuwenbroek (1997) suggested the Taylor linearization technique, which for the variable y_1 yields the approximate design variance

$$\begin{aligned} AV(\hat{Y}_1^{CR}) &= AV(\hat{Y}_1^R)+\mathbf{B}_{y_1}(\mathbf{I}-\mathbf{\Phi})[AV(\hat{\mathbf{Z}}_1^R)+AV(\hat{\mathbf{Z}}_2^R)](\mathbf{I}-\mathbf{\Phi})'\mathbf{B}'_{y_1} \\ &\quad -2\mathbf{B}_{y_1}(\mathbf{I}-\mathbf{\Phi})[AC(\hat{Y}_1^R,\hat{\mathbf{Z}}_1^R)]. \end{aligned} \tag{29}$$

In addition to the aforementioned difficulties in estimating the approximate variance matrices $AV(\hat{\mathbf{Z}}_1^R)$ and $AV(\hat{\mathbf{Z}}_2^R)$ in (29), estimation of the approximate covariance matrix $AC(\hat{Y}_1^R,\hat{\mathbf{Z}}_1^R)$ is also required. The estimation of the approximate design variance of the MZ estimator (16) is also difficult. It should be noted that any variance estimation procedure based on Taylor linear approximation requires a separate formula for each nonlinear statistic that may be of interest, and its validity relies on large-sample assumptions.

Moreover, the RN method does not lend itself to an easy variance estimation procedure by resampling techniques. For example, in the jackknife variance procedure, typically used in surveys with complex designs, the composite total $\hat{\mathbf{Z}}^{CR}=\hat{\mathbf{\Phi}}\hat{\mathbf{Z}}_1^R+(\mathbf{I}-\hat{\mathbf{\Phi}})\hat{\mathbf{Z}}_2^R$, where $\hat{\mathbf{\Phi}}=\hat{V}(\hat{\mathbf{Z}}_2^R)[\hat{V}(\hat{\mathbf{Z}}_1^R)+\hat{V}(\hat{\mathbf{Z}}_2^R)]^{-1}$, used as an additional control in the RN method, is random and thus must be replicated. This requires first saving the replicates of $\hat{\mathbf{Z}}_1^R$ and $\hat{\mathbf{Z}}_2^R$ from the initial step of computing jackknife estimates of $V(\hat{\mathbf{Z}}_1^R)$ and $V(\hat{\mathbf{Z}}_2^R)$ and then using the replicates of $\hat{\mathbf{Z}}^{CR}$ as controls in the separate jackknife variance estimation procedure for each sample. In contrast, the jackknife variance estimation for the MZ method is straightforward. The weighting procedure involving the combined sample is replicated, and variance estimates for each sample are obtained by treating the samples as domains; replication of the regression estimators involved is implicit. In particular, in the MZ method, the variability of $\tilde{\mathbf{B}}$ is accounted for in variance estimation.

The relative performance of the various composite regression procedures discussed earlier was tested in earlier work (Merkouris 2001) in an application involving two annual household expenditure surveys with stratified multistage design and with considerable difference in effective sample size. In one comparison, each of two common characteristics was used as the common composite variable in separate extended regression procedures, readily yielding the variance of the RN estimator (using the jackknife replication method) for these two characteristics without the difficulties mentioned in the previous section. The MZ estimators were nearly as efficient as the RN estimators relative to the GREG estimators from the two samples for both common characteristics, and, as expected, were substantially more efficient than the Z estimators. In a second comparison that focused on the relative merits of the MZ and Z regression procedures, two univariate and one bivariate composite procedures were used; each of the two univariate ones used as a single composite variable one of the common characteristics, whereas the bivariate used as composite variables both of these characteristics. Regarding the aligned composite estimators for the two common characteristics, the MZ estimators displayed larger efficiency gains over the Z estimators relative to the GREG estimators from the two samples. For the nonaligned estimators of a few noncommon target characteristics, in general the MZ estimators were more efficient than the Z estimators relative to the GREG estimators. Details and more comparative features of this empirical study can be found in the aforementioned article.

5. SUMMARY AND DISCUSSION

For the wide variety of multiple independent surveys from the same population that have some target characteristics in common, we have proposed an efficient and practical composite estimation method for the alignment of regression estimators of the totals of these characteristics. Specifically, an adjustment of the sampling weights based on the effective sample size is incorporated in the weighting matrix of a suitably extended regression procedure involving the combined sample. This leads to aligned composite estimators for common survey characteristics that are more efficient than the Zieschang estimators and nearly as efficient as (and considerably more usable than) the estimators of Renssen and Nieuwenbroek in general sampling settings. Under stratified simple random sampling or stratified Bernoulli sampling, with either separate or combined regression in each sample, a modification of Zieschang's original composite GREG procedure involving relative sampling fractions produces composite estimators that are of identical efficiency to the estimators of Renssen and Nieuwenbroek for common characteristics and more efficient for noncommon characteristics. Notably, in the particular stratification setting involving combined regression in each sample, the proposed method establishes an equivalence of GREG estimation and optimal regression estimation.

For complex sampling designs, a modification of the GREG procedure intended to generate optimal composite regression coefficients might require a very extensive system of regression constraints and a complicated adjustment of the weighting matrix, and for small sample sizes it might result in unstable estimates of regression coefficients. Then again, the motivation for the proposed modification is a practicable correction of the GREG estimation for different effective sample sizes in a multiple-sample setting, which should provide a good approximation to the optimal regression setup. In this respect, the results of the aforementioned limited empirical study have been positive.

By its formulation, the MZ method affords the flexibility of using an adjustment ϕ that is not motivated solely by efficiency considerations, while retaining all of its other attractive properties. For instance, if a ϕ that accounts for several relative survey errors can be determined, then it can be used as a shrinkage factor giving more weight to the more reliable of the two components of the composite estimator.

An extension of the MZ method to more than two surveys is straightforward. For example, to combine GREG estimators from three surveys, the partitioned matrix $\mathcal{X}_s$ in the regression setup (14) is augmented in an obvious way to include the auxiliary matrix $\mathbf{X}_{s_3}$ for the third survey, whereas the matrix $\mathcal{Z}_s$ pertaining to the composite constraints comprises only two sets of columns, $(\mathbf{Z}_{s_1}, -\mathbf{Z}_{s_2}, \mathbf{0})'$ and $(\mathbf{Z}_{s_1}, \mathbf{0}, -\mathbf{Z}_{s_3})'$, a third one being redundant. Three scalar adjustment factors given in terms of effective sample sizes, and adding up to 1, can be easily determined in the same manner leading to ϕ in (9); in the case of identical designs for the three surveys, they take the simple form $\phi_i = n_i/(n_1 + n_2 + n_3)$, $i = 1, 2, 3$. Then one regression procedure for the three samples will generate the composite weights, and implicitly the three matrix coefficients of the composite estimator $\tilde{\mathbf{Z}}^{CR} = \tilde{\mathbf{B}}_1\hat{\mathbf{Z}}_1^R + \tilde{\mathbf{B}}_2\hat{\mathbf{Z}}_2^R + \tilde{\mathbf{B}}_3\hat{\mathbf{Z}}_3^R$, with $\tilde{\mathbf{B}}_1 + \tilde{\mathbf{B}}_2 + \tilde{\mathbf{B}}_3 = \mathbf{I}$.

APPENDIX A: DERIVATION OF EXPRESSION (18)

For simplicity, the subscript s indicating sample quantities is dropped here. It follows from (15) that

$$\begin{aligned}
\mathbf{Z}'\mathbf{c} &= \hat{\mathbf{Z}}^R - \mathbf{Z}'\mathbf{L}\mathcal{Z}(\mathcal{Z}'\mathbf{L}\mathcal{Z})^{-1}\hat{\mathcal{Z}}^R \\
&= \hat{\mathbf{Z}}^R - (\mathbf{I} - 2\hat{\mathbf{B}})\hat{\mathcal{Z}}^R \\
&= \hat{\mathbf{Z}} + \mathbf{Z}'\boldsymbol{\Lambda}\mathbf{X}(\mathbf{X}'\boldsymbol{\Lambda}\mathbf{X})^{-1}[\mathbf{t_X} - \hat{\mathbf{X}}] \\
&\quad - (\mathbf{I} - 2\hat{\mathbf{B}})[\hat{\mathcal{Z}} + \mathcal{Z}'\boldsymbol{\Lambda}\mathbf{X}(\mathbf{X}'\boldsymbol{\Lambda}\mathbf{X})^{-1}[\mathbf{t_X} - \hat{\mathbf{X}}]] \\
&= \hat{\mathbf{Z}} - (\mathbf{I} - 2\hat{\mathbf{B}})\hat{\mathcal{Z}} \\
&\quad + \big[\mathbf{Z}'\boldsymbol{\Lambda}\mathbf{X}(\mathbf{X}'\boldsymbol{\Lambda}\mathbf{X})^{-1} - (\mathbf{I} - 2\hat{\mathbf{B}})\mathcal{Z}'\boldsymbol{\Lambda}\mathbf{X}(\mathbf{X}'\boldsymbol{\Lambda}\mathbf{X})^{-1}\big][\mathbf{t_X} - \hat{\mathbf{X}}].
\end{aligned}$$

The corresponding matrix of sample residuals is

$$\begin{aligned}
&\mathbf{Z} - \mathcal{Z}(\mathbf{I} - 2\hat{\mathbf{B}}') \\
&\qquad - \big[\mathbf{X}(\mathbf{X}'\boldsymbol{\Lambda}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Lambda}\mathbf{Z} + \mathbf{X}(\mathbf{X}'\boldsymbol{\Lambda}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Lambda}\mathcal{Z}(\mathbf{I} - 2\hat{\mathbf{B}}')\big] \\
&\quad = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\boldsymbol{\Lambda}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Lambda}][\mathbf{Z} - \mathcal{Z}(\mathbf{I} - 2\hat{\mathbf{B}}')] \\
&\quad = (\mathbf{I} - \mathbf{P_X})[\mathbf{Z} - \mathcal{Z}(\mathbf{I} - 2\hat{\mathbf{B}}')].
\end{aligned}$$

Then the quadratic form in these residuals, with matrix $\boldsymbol{\Lambda}$, is

$$\begin{aligned}
&[\mathbf{Z} - \mathcal{Z}(\mathbf{I} - 2\hat{\mathbf{B}}')]'(\mathbf{I} - \mathbf{P}_\mathbf{X}')\boldsymbol{\Lambda}(\mathbf{I} - \mathbf{P_X})[\mathbf{Z} - \mathcal{Z}(\mathbf{I} - 2\hat{\mathbf{B}}')] \\
&\quad = [\mathbf{Z} - \mathcal{Z}(\mathbf{I} - 2\hat{\mathbf{B}}')]'\mathbf{L}[\mathbf{Z} - \mathcal{Z}(\mathbf{I} - 2\hat{\mathbf{B}}')] \\
&\quad = \begin{pmatrix} \mathbf{Z}_1 - \mathbf{Z}_1(\mathbf{I} - 2\hat{\mathbf{B}}') \\ \mathbf{Z}_2 + \mathbf{Z}_2(\mathbf{I} - 2\hat{\mathbf{B}}') \end{pmatrix}' \begin{pmatrix} \mathbf{L}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_2 \end{pmatrix} \begin{pmatrix} \mathbf{Z}_1 - \mathbf{Z}_1(\mathbf{I} - 2\hat{\mathbf{B}}') \\ \mathbf{Z}_2 + \mathbf{Z}_2(\mathbf{I} - 2\hat{\mathbf{B}}') \end{pmatrix} \\
&\quad = 4\begin{pmatrix} \mathbf{Z}_1\hat{\mathbf{B}}' \\ \mathbf{Z}_2(\mathbf{I} - \hat{\mathbf{B}}') \end{pmatrix}' \begin{pmatrix} \mathbf{L}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_2 \end{pmatrix} \begin{pmatrix} \mathbf{Z}_1\hat{\mathbf{B}}' \\ \mathbf{Z}_2(\mathbf{I} - \hat{\mathbf{B}}') \end{pmatrix}.
\end{aligned}$$

APPENDIX B: PROOF OF THEOREM 1

(a) The approximate variance of $\hat{\mathbf{Z}}_i^R$ is given by $AV(\hat{\mathbf{Z}}_i^R) = (\mathbf{Z} - \mathbf{X}_i\boldsymbol{\beta}_i')'\boldsymbol{\Lambda}_i^\circ(\mathbf{Z} - \mathbf{X}_i\boldsymbol{\beta}_i')$, as the quadratic form in the population residuals $\mathbf{Z} - \mathbf{X}_i\boldsymbol{\beta}_i'$, with the matrix $\boldsymbol{\Lambda}_i^\circ = \{(\pi_{kl} - \pi_k\pi_l)/\pi_k\pi_l\}$ associated with sample s_i, and with $\boldsymbol{\beta}_i = \mathbf{Z}'\mathbf{X}_i(\mathbf{X}_i'\mathbf{X}_i)^{-1}$. It can be easily shown that for simple random sampling with sampling fraction $f_i = n_i/N$, $\boldsymbol{\Lambda}_i^\circ = N^2(1 - f_i)/[n_i(N - 1)](\mathbf{I} - \mathbf{P}_{\mathbf{1}_N})$, where $\mathbf{P}_{\mathbf{1}_N} = \mathbf{1}_N(\mathbf{1}_N'\mathbf{1}_N)^{-1}\mathbf{1}_N'$. Writing $\mathbf{Z} - \mathbf{X}_i\boldsymbol{\beta}_i' = (\mathbf{I} - \mathbf{P}_{\mathbf{X}_i})\mathbf{Z}$ for $\mathbf{P}_{\mathbf{X}_i} = \mathbf{X}_i(\mathbf{X}_i'\mathbf{X}_i)^{-1}\mathbf{X}_i'$, we get

$$\begin{aligned}
AV(\hat{\mathbf{Z}}_i^R) &= N^2(1 - f_i)/[n_i(N - 1)] \\
&\qquad \times \mathbf{Z}'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_i})'(\mathbf{I} - \mathbf{P}_{\mathbf{1}_N})(\mathbf{I} - \mathbf{P}_{\mathbf{X}_i})\mathbf{Z} \\
&= N^2(1 - f_i)/[n_i(N - 1)] \\
&\qquad \times \big[\mathbf{Z}'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_i})\mathbf{Z} - \mathbf{Z}'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_i})\mathbf{P}_{\mathbf{1}_N}(\mathbf{I} - \mathbf{P}_{\mathbf{X}_i})\mathbf{Z}\big],
\end{aligned}$$

using the property that the projection matrix $\mathbf{I} - \mathbf{P}_{\mathbf{X}_i}$ is idempotent. Now, by assumption, $\mathbf{1}_N = \mathbf{X}_i\mathbf{h}_i$, so that $\mathbf{1}_N'\mathbf{P}_{\mathbf{X}_i} = \mathbf{1}_N'$ and hence $\mathbf{P}_{\mathbf{1}_N}(\mathbf{I} - \mathbf{P}_{\mathbf{X}_i}) = \mathbf{0}$. Thus $AV(\hat{\mathbf{Z}}_i^R) = N^2(1 - f_i)/[n_i(N - 1)] \times \mathbf{Z}'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_i})\mathbf{Z}$. Then under condition C1, (21) follows immediately.

A similar derivation gives $AC(\hat{Y}_i^R, \hat{\mathbf{Z}}_i^R) = N^2(1 - f_i)/[n_i(N - 1)]\mathbf{Y}'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_i})\mathbf{Z}$.

The estimated approximate variance of $\hat{\mathbf{Z}}_i^R$ is given by $\widehat{AV}(\hat{\mathbf{Z}}_i^R) = (\mathbf{Z}_{s_i} - \mathbf{X}_{s_i}\hat{\boldsymbol{\beta}}_i')'\boldsymbol{\Lambda}_{s_i}^\circ(\mathbf{Z}_{s_i} - \mathbf{X}_{s_i}\hat{\boldsymbol{\beta}}_i')$, as the quadratic form in the sample residuals $\mathbf{Z}_{s_i} - \mathbf{X}_{s_i}\hat{\boldsymbol{\beta}}_i'$, with the matrix $\boldsymbol{\Lambda}_{s_i}^\circ = \{(\pi_{kl} - \pi_k\pi_l)/\pi_k\pi_l \times \pi_{kl}\}_{k,l\in s_i}$. Similar derivations as before give

$$\begin{aligned}
\widehat{AV}(\hat{\mathbf{Z}}_i^R) &= N^2(1 - f_i)/[n_i(n_i - 1)] \\
&\qquad \times \mathbf{Z}_{s_i}'(\mathbf{I} - \mathbf{P}_{\mathbf{X}_{s_i}})(\mathbf{I} - \mathbf{P}_{\mathbf{1}_{n_i}})(\mathbf{I} - \mathbf{P}_{\mathbf{X}_{s_i}})\mathbf{Z}_{s_i},
\end{aligned}$$

where $\mathbf{P}_{\mathbf{X}_{s_i}} = \mathbf{X}_{s_i}(\mathbf{X}_{s_i}'\boldsymbol{\Lambda}_{s_i}\mathbf{X}_{s_i})^{-1}\mathbf{X}_{s_i}'\boldsymbol{\Lambda}_{s_i} = \mathbf{X}_{s_i}(\mathbf{X}_{s_i}'\mathbf{X}_{s_i})^{-1}\mathbf{X}_{s_i}'$, because $\boldsymbol{\Lambda}_{s_i} = (N/n_i)\mathbf{I}$, and $\mathbf{P}_{\mathbf{1}_{n_i}} = \mathbf{1}_{n_i}(\mathbf{1}_{n_i}'\mathbf{1}_{n_i})^{-1}\mathbf{1}_{n_i}'$. Here also

$\mathbf{P}_{\mathbf{1}_{n_i}}(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_i}})=\mathbf{0}$, and thus $\widehat{AV}(\hat{\mathbf{Z}}_i^R)=N^2(1-f_i)/[n_i(n_i-1)]\times \mathbf{Z}'_{s_i}(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_i}})\mathbf{Z}_{s_i}$. With condition C1 and for large sample sizes n_i, we obtain the approximation

$$\begin{aligned}&\widehat{AV}(\hat{\mathbf{Z}}_2^R)\big[\widehat{AV}(\hat{\mathbf{Z}}_1^R)+\widehat{AV}(\hat{\mathbf{Z}}_2^R)\big]^{-1}\\ &\quad\approx n_1^2\mathbf{Z}'_{s_2}\big(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_2}}\big)\mathbf{Z}_{s_2}\\ &\qquad\times\big[n_2^2\mathbf{Z}'_{s_1}\big(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_1}}\big)\mathbf{Z}_{s_1}+n_1^2\mathbf{Z}'_{s_2}\big(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_2}}\big)\mathbf{Z}_{s_2}\big]^{-1}. \quad (B.1)\end{aligned}$$

On the other hand, it easy to see that the expression $\tilde{\mathbf{B}}=n_1\mathbf{Z}'_{s_2}\mathbf{L}_{s_2}\times \mathbf{Z}_{s_2}[n_2\mathbf{Z}'_{s_1}\mathbf{L}_{s_1}\mathbf{Z}_{s_1}+n_1\mathbf{Z}'_{s_2}\mathbf{L}_{s_2}\mathbf{Z}_{s_2}]^{-1}$, with $\mathbf{L}_{s_i}=\mathbf{\Lambda}_{s_i}(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_i}})$, reduces to the right side of (B.1). Furthermore, $\widehat{AC}(\hat{Y}_i^R,\hat{\mathbf{Z}}_i^R)=N^2\times(1-f_i)/[n_i(n_i-1)]\mathbf{Y}'_{s_i}(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_i}})\mathbf{Z}_{s_i}$, from which follows (23).

(b) The proof is based on the proof of (a), and is omitted for the sake of brevity.

(c) Here $AV(\hat{\mathbf{Z}}_i^R)=(\mathbf{Z}-\mathbf{X}_i\boldsymbol{\beta}'_i)'\mathbf{\Lambda}_i^\circ(\mathbf{Z}-\mathbf{X}_i\boldsymbol{\beta}'_i)$, where $\mathbf{\Lambda}_i^\circ=\operatorname{diag}(\mathbf{\Lambda}_{il}^\circ)$, $\mathbf{\Lambda}_{il}^\circ=N_{il}^2(1-f_{il})/[n_{il}(N_{il}-1)](\mathbf{I}-\mathbf{P}_{\mathbf{1}_{N_{il}}})$, and $\mathbf{X}_i=(\mathbf{X}_{\circ i}\ \operatorname{diag}(\mathbf{1}_{N_{il}}))$. Then, under conditions C2 and C3,

$$\begin{aligned}AV(\hat{\mathbf{Z}}_i^R)&=(N/n_i)(1-f_i)\mathbf{Z}'\big(\mathbf{I}-\mathbf{P}_{\mathbf{X}_i}\big)'\big(\mathbf{I}-\operatorname{diag}(\mathbf{P}_{\mathbf{1}_{N_{il}}})\big)(\mathbf{I}-\mathbf{P}_{\mathbf{X}_i})\mathbf{Z}\\ &=(N/n_i)(1-f_i)\\ &\quad\times\big[\mathbf{Z}'\big(\mathbf{I}-\mathbf{P}_{\mathbf{X}_i}\big)\mathbf{Z}-\mathbf{Z}'\big(\mathbf{I}-\mathbf{P}_{\mathbf{X}_i}\big)\operatorname{diag}\big(\mathbf{P}_{\mathbf{1}_{N_{il}}}\big)\big(\mathbf{I}-\mathbf{P}_{\mathbf{X}_i}\big)\mathbf{Z}\big].\end{aligned}$$

Denote $\mathbf{D}_i=\operatorname{diag}(\mathbf{1}_{N_{il}})$. It can be easily checked that $\mathbf{P}_{\mathbf{D}_i}(=\mathbf{D}_i\times(\mathbf{D}'_i\mathbf{D}_i)^{-1}\mathbf{D}'_i)=\operatorname{diag}(\mathbf{P}_{\mathbf{1}_{N_{il}}})$. Using algebra of partitioned matrices and noting that $\mathbf{P}_{\mathbf{X}_i}=\mathbf{P}'_{\mathbf{X}_i}$, it can be shown that

$$\begin{aligned}\mathbf{P}_{\mathbf{X}_i}&=\mathbf{X}_{\circ i}\big[\mathbf{X}'_{\circ i}\big(\mathbf{I}-\mathbf{P}_{\mathbf{D}_i}\big)\mathbf{X}_{\circ i}\big]^{-1}\mathbf{X}'_{\circ i}\big(\mathbf{I}-\mathbf{P}_{\mathbf{D}_i}\big)\\ &\quad+\mathbf{D}_i\big[\mathbf{D}'_i\big(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{\circ i}}\big)\mathbf{D}_i\big]^{-1}\mathbf{D}'_i\big(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{\circ i}}\big).\end{aligned}$$

Noting that $(\mathbf{I}-\mathbf{P}_{\mathbf{D}_i})\mathbf{P}_{\mathbf{D}_i}=\mathbf{0}$, it can be easily verified that $\mathbf{P}_{\mathbf{X}_i}\mathbf{P}_{\mathbf{D}_i}=\mathbf{P}_{\mathbf{D}_i}$. It follows then that $(\mathbf{I}-\mathbf{P}_{\mathbf{X}_i})\operatorname{diag}(\mathbf{P}_{\mathbf{1}_{N_{il}}})=\mathbf{0}$, and, therefore,

$$AV(\hat{\mathbf{Z}}_i^R)=(N/n_i)(1-f_i)\mathbf{Z}'\big(\mathbf{I}-\mathbf{P}_{\mathbf{X}_i}\big)\mathbf{Z},$$

which implies (21) under condition C1.

An estimate of $AV(\hat{\mathbf{Z}}_i^R)$ is given in terms of sample quantities as $\widehat{AV}(\hat{\mathbf{Z}}_i^R)=\mathbf{Z}'_{s_i}(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_i}})'\operatorname{diag}(\mathbf{\Lambda}_{sil}^\circ)(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_i}})\mathbf{Z}_{s_i}$, where $\mathbf{\Lambda}_{sil}^\circ=N_{il}^2(1-f_{il})/[n_{il}(n_{il}-1)](\mathbf{I}-\mathbf{P}_{\mathbf{1}_{n_{il}}})$ and $\mathbf{P}_{\mathbf{X}_{s_i}}=\mathbf{X}_{s_i}(\mathbf{X}'_{s_i}\mathbf{X}_{s_i})^{-1}\mathbf{X}'_{s_i}$, because $\mathbf{\Lambda}_{s_i}=(N/n_i)\mathbf{I}$ under proportional allocation. Then, following the derivation of $AV(\hat{\mathbf{Z}}_i^R)$, we get $\widehat{AV}(\hat{\mathbf{Z}}_i^R)\approx N^2(1-f_i)/n_i^2\mathbf{Z}'_{s_i}\times(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_i}})\mathbf{Z}_{s_i}$ for large n_{il} and proportional allocation. Next, using the adjustment $(1/n_i)\mathbf{\Lambda}_{s_i}$, we get $(N/n_i^2)\mathbf{Z}'_{s_i}\mathbf{L}_{s_i}\mathbf{Z}_{s_i}=(N/n_i^2)\mathbf{Z}'_{s_i}(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_i}})\mathbf{Z}_{s_i}$. Then (22) follows immediately under condition C1. Similarly for (23).

The proof of $\hat{Y}_i^R=\hat{Y}_i^O$ and $\hat{\mathbf{Z}}_i^R=\hat{\mathbf{Z}}_i^O$ is given next under the broader condition of Theorem 2.

APPENDIX C: PROOF OF THEOREM 2

(a) The proof is similar to the proof of Theorem 1(b), but without the assumption of proportional allocation of the stratum sample sizes.

(b) The proof closely follows the proof of Theorem 1(c). In particular, again without the assumption of proportional allocation, the estimated variance $\widehat{AV}(\hat{\mathbf{Z}}_i^R)$ reduces to $\mathbf{Z}'_{s_i}(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_i}})'\mathbf{\Phi}_{s_i}\mathbf{\Lambda}_{s_i}(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_i}})\mathbf{Z}_{s_i}$, where $\mathbf{P}_{\mathbf{X}_{s_i}}=\mathbf{X}_{s_i}(\mathbf{X}'_{s_i}\mathbf{\Phi}_{s_i}\Lambda_{s_i}\mathbf{X}_{s_i})^{-1}\mathbf{X}'_{s_i}\mathbf{\Phi}_{s_i}\mathbf{\Lambda}_{s_i}$, and further reduces to $\mathbf{Z}'_{s_i}\mathbf{\Phi}_{s_i}\mathbf{\Lambda}_{s_i}(\mathbf{I}-\mathbf{P}_{\mathbf{X}_{s_i}})\mathbf{Z}_{s_i}$.

The row vector of control totals corresponding to $\mathbf{X}_{s_i}=(\mathbf{X}_{\circ s_i}\ \operatorname{diag}(\mathbf{1}_{n_{il}}))$ is $(\mathbf{t}'_{\mathbf{X}_{\circ i}},\mathbf{N}'_i)$, where $\mathbf{N}_i$ is the vector of stratum totals. Noting that $\hat{\mathbf{N}}_i=\mathbf{N}_i$, the GREG estimator $\hat{Y}_i^R$ based on the partitioned matrix $\mathbf{X}_{s_i}$ and the adjusted weighting matrix $\mathbf{\Phi}_{s_i}\mathbf{\Lambda}_{s_i}$ takes the form $\hat{Y}_i^R=\hat{Y}_i+\mathbf{Y}'_{s_i}\mathbf{\Phi}_{s_i}\mathbf{\Lambda}_{s_i}(\mathbf{I}-\mathbf{P}_{\mathbf{D}_{s_i}})\mathbf{X}_{\circ s_i}[\mathbf{X}'_{\circ s_i}\mathbf{\Phi}_{s_i}\mathbf{\Lambda}_{s_i}(\mathbf{I}-\mathbf{P}_{\mathbf{D}_{s_i}})\mathbf{X}_{\circ s_i}]^{-1}[\mathbf{t}_{\mathbf{X}_{\circ i}}-\hat{\mathbf{X}}_{\circ i}]$, where $\mathbf{P}_{\mathbf{D}_{s_i}}=\mathbf{D}_{s_i}(\mathbf{D}'_{s_i}\mathbf{\Phi}_{s_i}\mathbf{\Lambda}_{s_i}\mathbf{D}_{s_i})^{-1}\times\mathbf{D}'_{s_i}\mathbf{\Phi}_{s_i}\mathbf{\Lambda}_{s_i}$ and $\mathbf{D}_{s_i}=\operatorname{diag}(\mathbf{1}_{n_{il}})$. Replacing $\mathbf{\Phi}_{s_i}\mathbf{\Lambda}_{s_i}(\mathbf{I}-\mathbf{P}_{\mathbf{D}_{s_i}})$ by $\mathbf{\Lambda}_{s_i}^\circ=\mathbf{\Phi}_{s_i}^\circ(\mathbf{I}-\operatorname{diag}(\mathbf{P}_{\mathbf{1}_{n_{il}}}))$, where $\mathbf{\Phi}_{s_i}^\circ=\operatorname{diag}((N_{il}^2(1-f_{il})/[n_{il}\times(n_{il}-1)])\mathbf{I}_{n_{il}})$, gives the optimal regression estimator $\hat{Y}_i^O$. It is easy to show that $\mathbf{I}-\mathbf{P}_{\mathbf{D}_{s_i}}=\mathbf{I}-\operatorname{diag}(\mathbf{P}_{\mathbf{1}_{n_{il}}})$. Also, under the current assumptions, $\mathbf{\Phi}_{s_i}^\circ=\operatorname{diag}((N_{il}^2/n_{il}^2)\mathbf{I}_{n_{il}})=\mathbf{\Phi}_{s_i}\mathbf{\Lambda}_{s_i}$. It then follows that $\hat{Y}_i^R=\hat{Y}_i^O$. Similarly, $\hat{\mathbf{Z}}_i^R=\hat{\mathbf{Z}}_i^O$.

[Received April 2002. Revised February 2004.]

REFERENCES

Deville, J. C., and Särndal, C. E. (1992), "Calibration Estimators in Survey Sampling," *Journal of the American Statistical Association*, 87, 376–382.

Hidiroglou, M. A. (2001), "Double Sampling," *Survey Methodology*, 27, 143–154.

Merkouris, T. (2001), "Combining Regression Estimators From Independent Samples From the Same Finite Population," Working Paper HSMD-2001-008E, Statistics Canada.

Montanari, G. E. (1987), "Post-Sampling Efficient QR-Prediction in Large-Scale Surveys," *International Statistics Review*, 55, 191–202.

——— (1998), "On Regression Estimation of Finite Population Mean," *Survey Methodology*, 24, 69–77.

——— (2000), "Conditioning on Auxiliary Variable Means in Finite Population Inference," *Australia and New Zealand Journal Statistics*, 42, 407–421.

Rao, J. N. K. (1994), "Estimating Totals and Distribution Functions Using Auxiliary Information at the Estimation Stage," *Journal of Official Statistics*, 10, 153–165.

Renssen, R. H., and Nieuwenbroek, N. J. (1997), "Aligning Estimates for Common Variables in Two or More Sample Surveys," *Journal of the American Statistical Association*, 92, 368–375.

Särndal, C. E. (1996), "Efficient Estimators With Simple Variance in Unequal Probability Sampling," *Journal of the American Statistical Association*, 91, 1289–1300.

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

Singh, A., and Merkouris, T. (1995), "Composite Estimation by Modified Regression for Repeated Surveys," in *Proceedings of the Section on Survey Research Methods*, American Statistical Association, pp. 420–425.

Statistics Canada Symposium (1999), *Proceedings of the XVIth Annual International Methodology Symposium on Combining Data From Different Sources*.

Zieschang, K. D. (1990), "Sample Weighting Methods and Estimation of Totals in the Consumer Expenditures Survey," *Journal of the American Statistical Association*, 85, 986–1001.