Component of Statistics Canada
Catalogue no. 12-001-X Business Survey Methods Division

Article

Comparison of survey regression techniques in the context of small area estimation of poverty

by Stephen J. Haslett, Marissa C. Isidro and Geoffrey Jones

December 2010

Statistics Canada Statistique Canada

Canada

Survey Methodology, December 2010
Vol. 36, No. 2, pp. 157-170
Statistics Canada, Catalogue No. 12-001-X

Comparison of survey regression techniques in the context of small area estimation of poverty

Stephen J. Haslett, Marissa C. Isidro and Geoffrey Jones [1]

Abstract

One key to poverty alleviation or eradication in the third world is reliable information on the poor and their location, so that interventions and assistance can be effectively targeted to the neediest people. Small area estimation is one statistical technique that is used to monitor poverty and to decide on aid allocation in pursuit of the Millennium Development Goals. Elbers, Lanjouw and Lanjouw (ELL) (2003) proposed a small area estimation methodology for income-based or expenditure-based poverty measures, which is implemented by the World Bank in its poverty mapping projects via the involvement of the central statistical agencies in many third world countries, including Cambodia, Lao PDR, the Philippines, Thailand and Vietnam, and is incorporated into the World Bank software program PovMap. In this paper, the ELL methodology which consists of first modeling survey data and then applying that model to census information is presented and discussed with strong emphasis on the first phase, *i.e.*, the fitting of regression models and on the estimated standard errors at the second phase. Other regression model fitting procedures such as the General Survey Regression (GSR) (as described in Lohr (1999) Chapter 11) and those used in existing small area estimation techniques: Pseudo-Empirical Best Linear Unbiased Prediction (Pseudo-EBLUP) approach (You and Rao 2002) and Iterative Weighted Estimating Equation (IWEE) method (You, Rao and Kovačević 2003) are presented and compared with the ELL modeling strategy. The most significant difference between the ELL method and the other techniques is in the theoretical underpinning of the ELL model fitting procedure. An example based on the Philippines Family Income and Expenditure Survey is presented to show the differences in both the parameter estimates and their corresponding standard errors, and in the variance components generated from the different methods and the discussion is extended to the effect of these on the estimated accuracy of the final small area estimates themselves. The need for sound estimation of variance components, as well as regression estimates and estimates of their standard errors for small area estimation of poverty is emphasized.

Key Words: Small area models; Nested error regression model; Poverty mapping.

1. Introduction

Poverty is a very complex multidimensional concern: there is no single definition and method of measurement available. In this paper, we adhere to the meaning of poverty that is used by most economists, *i.e.*, households are considered to be in poverty if their income falls below some income threshold called the poverty line. Chambers (2006) described this as income-poverty, and it is the definition adopted by the World Bank in the implementation of their small area poverty mapping projects carried out in conjunction with national statistical agencies and used, for example, for monitoring progress towards the Millennium Development Goals (UN website). Sometimes expenditure-based poverty measures are used instead to assess economic poverty. In public health related contexts, different measures such as standardized weight for age, height for age and weight for height for children (underweight, stunting and wasting, respectively) are used, *e.g.*, in Bangladesh (Haslett and Jones 2004) and Nepal (Haslett and Jones 2006).

Surveys conducted in most third world countries usually allow an acceptable level of precision for reporting poverty statistics at the first and second administrative level or geographical area (*e.g.*, for the Philippines - National and Region respectively). However, for policy makers to properly target assistance and interventions to the neediest communities and households, more disaggregated finer-level poverty statistics are needed. However, survey based poverty statistics at smaller geographical areas or lower administrative level are usually less reliable (have higher standard errors) due to smaller sample sizes, and this is where small area estimation comes into play.

The most common small area estimation methodology used for poverty measures in third world countries proposed by Elbers, Lanjouw and Lanjouw (ELL) (2002, 2003) allows generation of more precise estimates for smaller geographical areas by combining the survey data with information from a recent census. The ELL method consists of two phases: fitting a regression model (or models) to complex survey data and using that model to predict income or expenditure per capita at household level (which is transformed and aggregated to estimate poverty statistics at small area level).

In this paper, we focus specifically on the various algorithms used to fit the phase 1 regression models, and to estimate regression parameter standard errors and variance components from survey data. We emphasise consequences of survey regression modeling decisions rather than the

1. Stephen J. Haslett, Marissa C. Isidro and Geoffrey Jones, Institute of Fundamental Sciences: Statistics, College of Sciences, Massey University, Private Bag 11-222, Palmerston North, New Zealand. E-mail: S.J.Haslett@massey.ac.nz.

entire and rather comprehensive system ELL use to form small area estimates.

The preliminary requirement of the ELL methodology applied to economic measures is to develop an accurate model of per capita income or expenditure of households although this is often used to generate non-linear functions of income or expenditure (*e.g.*, poverty incidence - percentage of households below the poverty line, or poverty gap - sum of relative differences in income or expenditure for households or individuals below the poverty line). The survey-based regression model developed for income or expenditure is critical to accurate poverty statistics, but as we show below the regression model itself is not always the most important element, and other issues such as estimation of variance components deserve emphasis.

Other existing survey-based small area estimation regression techniques - Pseudo-Empirical Best Linear Unbiased Prediction (Pseudo-EBLUP) approach (You and Rao 2002), Iterative Weighted Estimating Equation (IWEE) method (You *et al.* 2003) and the General Survey Regression (GSR) (Skinner, Holt and Smith 1989) method are considered as alternative survey based model-fitting techniques and compared with two variations of the ELL method for fitting regression models to survey data. Our investigation is based on real data from the 2000 Philippine Family Income and Expenditure Survey (FIES), rather than simulated data.

This paper is organized as follows: Section 2 gives relevant background on small area models; the model for income (or expenditure) as presented by Elbers, Lanjouw and Lanjouw is given in Section 3; presented in Section 4 is a summary of the ELL methodology, followed by details on the alternative fitting methods in Section 5, which includes the Pseudo-Empirical Best Linear Unbiased Prediction Approach (5.1), IWEE Method (5.2), and the General Survey Regression Method (5.3). Section 6 discusses differences between the techniques, while Section 7 presents their application to the Philippine FIES 2000 data. This is followed by the conclusion and recommendations (Section 8).

2. Small area models

Ghosh and Rao (1994) classify small area models into two broad categories, area level and unit level models. Area level models refer to sets of models that can be considered when only area-specific auxiliary variables are available. Unit level models, on the other hand, refer to models that can be considered when there are unit-specific auxiliary variables and unit level values of the variable under study can be used. All such models are special cases of a general linear or generalized linear mixed model, and usually involve both fixed and random effects.

For area level models, it is assumed that the population mean $(\bar{Y}_a)$ of the a^{th} small area or some suitable function $\theta_a = g(\bar{Y}_a)$ is related to the area-specific auxiliary variables $\mathbf{x}_a = (x_{a1}, ..., x_{ap})'$ through a linear model

$$\theta_a = \mathbf{x}_a' \boldsymbol{\beta} + c_a v_a \tag{1}$$

where $a = 1, ..., k$, $v_a \sim \text{iid}(0, \sigma_v^2)$, $\boldsymbol{\beta}$ is a vector of regression parameters, c_a are known or estimated positive constants to allow for heteroscedasticity, k is the total number of small areas under study and p is the number of auxiliary variables. It is assumed that a direct design-based estimator, $\hat{\bar{Y}}_a$, of the population mean $\bar{Y}_a$ is available whenever the area sample size $n_a \geq 1$, and that

$$\hat{\theta}_a = \theta_a + e_a \tag{2}$$

where $\hat{\theta}_a = g(\hat{\bar{Y}}_a)$ and the sampling errors e_a are independent $N(0, V_a)$ with known variance V_a. Combining equation (1) and (2) gives the area level linear mixed model:

$$\hat{\theta}_a = \mathbf{x}_a' \boldsymbol{\beta} + c_a v_a + e_a. \tag{3}$$

We note that (3) involves both design-based random variables e_a and model-based random variables v_a (Rao 1999), where design-based variables are due to the sample selection mechanism, and model-based ones to the super-population structure in which the model is embedded.

Area level models have various extensions so they can for example handle correlated sampling errors, spatial dependence of random small area effects, time series and cross-sectional data (see Rao 2003, 1999 and Ghosh and Rao 1994).

The unit level model assumes that the variable of interest Y_{ah} for the h^{th} unit in the a^{th} small area is related to the element-specific auxiliary data $\mathbf{x}_{ah} = (x_{ah1}, ..., x_{ahp})'$ through a nested error regression model:

$$Y_{ah} = \mathbf{x}_{ah}' \boldsymbol{\beta} + v_a + e_{ah} \tag{4}$$

where $a = 1, ..., k$, $h = 1, ..., N_a$, $\boldsymbol{\beta} = (\beta_0, ... \beta_{p-1})$ is $p \times 1$ vector of regression parameters and N_a is the number of population units or households in the a^{th} small area. It is also assumed that the random effects v_a, are iid $N(0, \sigma_v^2)$ and are independent of the unit errors e_{ah} which are assumed to be iid $N(0, \sigma_e^2)$. Extensions that allow errors to be heteroscedastic, with known scaling constant(s) are also possible.

The ELL method uses a unit level model, where the units are households in the case of income or expenditure data, and where the variation is modeled at primary sampling unit, *i.e.*, cluster level and household level. Note that ELL do not include model variation at small area level, only for cluster within small area, and for household within cluster. This is the form of the basic model used for comparisons in this paper since ELL is the standard small area estimation

method for poverty in third world countries. In the real datasets we have studied this additional small area variation has been very small. Despite this empirical evidence however, important questions remain about how best to estimate the small area variance component in the presence of cluster level variation, when there is sample survey weighting, especially where many of the small areas contain only one sampled cluster.

The ELL model has a number of other characteristics not all of which are standard in a statistical sense (see Haslett and Jones 2005, for example). The intention of this paper is not to discuss differences in the available methods generally, but to focus directly on how methods of fitting regression models to survey data differ when the ELL first phase "base structure" of fitting a survey regression model is used. The focus of this paper therefore is on comparison of the available methods of fitting regression models to survey data on income or expenditure using a specified set of regressors, even though ELL can also be (and is) used relatively routinely to find small area estimates for non-linear functions (*e.g.*, poverty incidence, gap or severity) by applying fitted regression models to a census.

The answer to the 'best regression model fitting' question for survey data on which this paper focuses (as with other matters related to the ELL methodology) is particularly important because there are billions of dollars of aid funding that are (or have the potential to be) allocated based on the regression models used as part of small area estimation of poverty.

3. Income/consumption model

Modeling per capita income or expenditure of households instead of poverty measures themselves (such as poverty incidence and gap) is one of the distinctive features of the ELL method. As mentioned in the previous section, the ELL method involves fitting the income or expenditure model to the survey data and applying it to the census data prior to the generation of the small area estimates of poverty measures. The income/expenditure model is as follows:

$$Y_{bh} = \mathbf{x}'_{bh}\boldsymbol{\beta} + u_{bh} \tag{5}$$

where $b = 1, ..., M$, $h = 1, ..., N_b$; Y_{bh} is the log-transformed per capita income or expenditure of the h^{th} unit or household in the b^{th} cluster, M is the total number of clusters in the population and N_b is the total number of households in the b^{th} cluster in the population. $\mathbf{x}_{bh}$ is a set of the auxiliary variables available in both the survey and the census, which generally need to be contemporaneous; u_{bh} is the random error term representing that part of Y_{bh} that cannot be explained by $\mathbf{x}_{bh}$. Income and expenditure data almost invariably have a skewed distribution, hence a transformation (usually logarithmic) is applied to make the data more symmetrical.

The households for which data on per capita income or expenditure is collected are seldom independent, but have natural groupings or clusters, often defined administratively. Households that are close to each other or in the same cluster, tend to be similar in many respects. In the survey data, the clusters are usually also the primary sampling units (PSUs) for the sample survey design. To account for the clustering of households, the random error term u_{bh} in the regression model is usually assumed to have the following specification:

$$u_{bh} = v_b + e_{bh} \tag{6}$$

where v and e are independent of each other and uncorrelated with $\mathbf{x}_{bh}$, v_b is the error term held in common by the b^{th} group or cluster (*e.g.*, barangay for the Philippines) and e_{bh} is the household level error within the cluster. The importance of each term is measured by their respective variances or variance components, σ_v^2 and σ_e^2. There are various procedures for estimating these variances. This important topic is covered in the sections that follow.

Model (5) can be written as

$$Y_{bh} = \mathbf{x}'_{bh}\boldsymbol{\beta} + v_b + e_{bh} \tag{7}$$

which is similar in form to the unit level model or nested error regression model mentioned in the previous section. However while the form of the model is similar, the group being referred to is different, *e.g.*, Y_{ah} refers to the h^{th} household in the a^{th} small area, while Y_{bh} refers to the h^{th} household in the b^{th} cluster. Clusters, based on the survey design, will typically be much smaller than the areas for which small area estimates are sought, and generally (unlike almost all the small areas) not all clusters are sampled. For example in the Philippines, estimates are sought at the municipal level which is composed of barangays or clusters.

4. The ELL methodology

In the ELL methodology, the estimate of the regression parameter $\boldsymbol{\beta}$ is given, in Elbers *et al.* (2002, page 11 footnote 8) and in the POVMAP software Zhao (2006) developed for the ELL method, as

$$\hat{\boldsymbol{\beta}}_{\text{ELL}} = \left(\sum_{b=1}^{m} \mathbf{X}'_b \mathbf{W}_b \mathbf{V}_b^{-1} \mathbf{X}_b \right)^{-1} \left(\sum_{b=1}^{m} \mathbf{X}'_b \mathbf{W}_b \mathbf{V}_b^{-1} \mathbf{y}_b \right) \tag{8}$$

and the corresponding variance-covariance matrix as

$$\mathbf{V}(\hat{\boldsymbol{\beta}}_{\text{ELL}}) = \mathbf{D} \left[\left(\sum_{b=1}^{m} \mathbf{X}'_b \mathbf{W}_b \mathbf{V}_b^{-1} \mathbf{W}_b \mathbf{X}_b \right)^{-1} \right] \mathbf{D} \tag{9}$$

where $\mathbf{V}_b = (\sigma_e^2 \mathbf{I}_{n_b} + \sigma_v^2 \mathbf{1}_{n_b} \mathbf{1}'_{n_b})$, (σ_υ^2) is the cluster level variance, while (σ_e^2) is the household level variance, $\mathbf{I}_{n_b}$ is an identity matrix, $\mathbf{1}'_{n_b} = (1...1)$ is a constant vector, $\mathbf{D} = (\sum_{b=1}^{m} \mathbf{X}'_b \mathbf{W}_b \mathbf{V}_b^{-1} \mathbf{X}_b)^{-1}$, $\mathbf{X}_b = (\mathbf{x}_{b1}, ..., \mathbf{x}_{bn_b})'$; $\mathbf{y}_b = (y_{b1}, ..., y_{bn_b})'$; $\mathbf{W}_b$ is a diagonal matrix of sampling weights; m is the number of clusters in the sample and n_b is the number of households in each sampled cluster. Equation (8) assumes $\mathbf{V}_b$ is known. In practice we need to estimate σ_e^2 and σ_v^2 to get the estimator $\hat{\mathbf{V}}_b$. We note that the variance expression in (9) is derived under a vaguely specified model assumed for the sample (see Elbers *et al.* 2002). Under the ELL method, fitting the income/expenditure model (7) involves obtaining the initial estimate of $\boldsymbol{\beta}$ through weighted least squares (WLS) method and using the residuals of the initial model to estimate the covariance matrix $\mathbf{V}_b$ needed to obtain $\hat{\boldsymbol{\beta}}_{\mathrm{ELL}}$. The estimate of the cluster level (σ_v^2) and household level (σ_e^2) variances, are derived by Elbers *et al.* (2002) as follows:

$$\hat{\sigma}_v^2 = \max\left(\frac{\sum_b w_b (u_{b.} - u_{..})^2}{\sum_b w_b (1 - w_b)} - \frac{\sum_b w_b (1 - w_b) \tau_b^2}{\sum_b w_b (1 - w_b)}; 0 \right) \quad (10)$$

where $\tau_b^2 = \sum_h (e_{bh} - e_{b.})^2 / (n_b (n_b - 1)); w_b = \sum_h w_{bh} / \sum_b \sum_h w_{bh}$, is the by-cluster transformed sampling weights which sum to one across clusters and w_{bh} is the re-scaled sampling weights which sum to the total sample size. Here $u_{b.} = \sum_h u_{bh}$ and $u_{..} = \sum_b \sum_h u_{bh}$ (which is equal to zero) where u_{bh} is as defined in equation (6).

There are two ways suggested by Elbers *et al.* (2002) to generate the estimate of the household level variance component: "direct" computation which is denoted by $(\hat{\sigma}_e^2)$ or the heteroscedasticity model-based $(\hat{\sigma}_{e,bh}^2)$. Direct computation involves using the difference between the estimated mean square error from the initial WLS regression and the computed estimate of σ_v^2, while the heteroscedasticity model-based computation uses a logistic-type link function to bound the variance as follows:

$$\sigma_{e,bh}^2(z_{bh}, \boldsymbol{\alpha}, A, B) = \left[\frac{A \exp(z'_{bh} \boldsymbol{\alpha}) + B}{1 + \exp(z'_{bh} \boldsymbol{\alpha})} \right] \quad (11)$$

where A and B are the upper and lower bounds respectively, estimated with the parameter vector $\boldsymbol{\alpha}$ using a standard pseudomaximum likelihood procedure (Elbers *et al.* 2003), and where z_{bh} are auxiliary variables. Elbers *et al.* claim that imposing a minimum bound of zero and a maximum bound of $A^* = (1.05) \max\{e_{bh}^2\}$ in general yields similar estimates of the parameters $\boldsymbol{\alpha}$. These restrictions allow one to estimate the simpler form

$$\ln\left[\frac{e_{bh}^2}{A^* - e_{bh}^2} \right] = z'_{bh} \boldsymbol{\alpha} + r_{bh} \quad (12)$$

where r_{bh} is an error term and the other variables are as defined earlier. In most of the World Bank poverty mapping projects, slight modifications are usually made, for example, adding a constant δ to e_{bh}^2 in model (11).

By using model (12), and employing the delta method, $\hat{\sigma}_{e,bh}^2$ is computed as:

$$\hat{\sigma}_{e,bh}^2 = \left[\frac{A^* C_{bh}}{1 + C_{bh}} \right] + \frac{1}{2} \hat{\sigma}_r^2 \left[\frac{A^* C_{bh} (1 - C_{bh})}{(1 + C_{bh})^3} \right] \quad (13)$$

where $C_{bh} = \exp\{z'_{bh} \hat{\boldsymbol{\alpha}}\}$, and $\hat{\sigma}_r^2$ is the estimated variance of the residuals under model (12). If the household level variance component is based on a heteroscedastic model, then, $\mathbf{V}_b = (\sigma_{e,bh}^2 \mathbf{I}_{n_b} + \sigma_v^2 \mathbf{1}_{n_b} \mathbf{1}'_{n_b})$. Heteroscedasticity modeling is conducted on the assumption that variation at the household level depends on some covariates.

As discussed in more detail in the appendix, the way in which the weight matrix $\mathbf{W}_b$ enters the calculation in equation (9) above leads to an asymmetric estimated covariance matrix. A rather better approach based on 'pseudomaximum likelihood' is outlined by Pfeffermann, Skinner, Holmes, Goldstein and Rasbash (1998) and involves splitting $\mathbf{X}'_b \mathbf{V}_b^{-1} \mathbf{X}_b$ into separate sums of squares and cross-product terms, and weighting each appropriately - if we write $\mathbf{V}_b^{-1} = c \mathbf{I}_{n_b} + d \mathbf{1}_{n_b} \mathbf{1}'_{n_b}$ then the appropriate weighting is $c \mathbf{X}'_b \mathbf{W}_b \mathbf{X}_b + d \mathbf{X}'_b \mathbf{W}_b \mathbf{1}_{n_b} \mathbf{1}'_{n_b} \mathbf{W}_b \mathbf{X}_b$.

Since the ELL version, $\mathbf{W}_b \mathbf{V}_b^{-1}$, is not generally symmetric, neither is $\mathbf{D}$ in equation (9). As a consequence the supposed covariance matrix of $\hat{\boldsymbol{\beta}}_{\mathrm{ELL}}$, $\mathbf{V}(\hat{\boldsymbol{\beta}}_{\mathrm{ELL}})$, is also not symmetric. The POVMAP software attempts to solve this problem by taking the average of their $\mathbf{V}(\hat{\boldsymbol{\beta}}_{\mathrm{ELL}})$ and its transpose, thereby forcing the matrix to be symmetric.

Note again that under the ELL method, the regression fit to the survey data and the estimation of variance components is only the first phase. The consequent phase involves prediction at household level based on the entire census data and aggregation to small area level.

The survey fitting methods (derivation of the estimate of $\boldsymbol{\beta}$ and its corresponding variance-covariance matrix) of three alternative regression procedures to ELL are presented in the following sections.

5. Alternative fitting methods

5.1 The pseudo-empirical best linear unbiased prediction approach

You and Rao (2002) proposed an estimator of the small area mean by deriving an estimator of $\boldsymbol{\beta}$ based on the unit level model (4). The process of deriving the estimator of $\boldsymbol{\beta}$ starts with the computation of the best linear unbiased predictor (BLUP) of v_a given the parameters $\boldsymbol{\beta}$, σ_e^2 and

σ_v^2 from the aggregated (survey-weighted) area level model:

$$\bar{Y}_{aw} = \bar{\mathbf{x}}'_{aw}\boldsymbol{\beta} + v_a + \bar{e}_{aw} \quad (14)$$

which proceeds as follows:

$$\hat{v}_{aw}(\boldsymbol{\beta}, \sigma_e^2, \sigma_v^2) = \gamma_{aw}(\bar{y}_{aw} - \bar{\mathbf{x}}'_{aw}\boldsymbol{\beta}) \quad (15)$$

where $\bar{\mathbf{x}}_{aw} = \sum_{h=1}^{n_a} w_{ah}\mathbf{x}_{ah}$, $\bar{y}_{aw} = \sum_{h=1}^{n_a} w_{ah}y_{ah}$, $\gamma_{aw} = \sigma_v^2/(\sigma_v^2 + \sigma_e^2\,\delta_a^2)$, $w_{ah} = \tilde{w}_{ah}/\sum_{h=1}^{n_a}\tilde{w}_{ah}$, $\delta_a^2 = \sum_{h=1}^{n_a} w_{ah}^2$, and $\tilde{w}_{ah}$ are the unit level survey weights; then solving for the survey-weighted estimating equation for $\boldsymbol{\beta}$:

$$\sum_{a=1}^{k}\sum_{h=1}^{n_a}\tilde{w}_{ah}\mathbf{x}_{ah}[y_{ah} - \mathbf{x}'_{ah}\boldsymbol{\beta} - \hat{v}_{aw}(\boldsymbol{\beta}, \sigma_e^2, \sigma_v^2)] = 0 \quad (16)$$

from which the estimator of $\boldsymbol{\beta}$ is obtained as

$$\hat{\boldsymbol{\beta}}_w = \left\{\sum_{a=1}^{k}\sum_{h=1}^{n_a}\mathbf{x}_{ah}\mathbf{z}'_{ah}\right\}^{-1}\left\{\sum_{a=1}^{k}\sum_{h=1}^{n_a}\mathbf{z}_{ah}y_{ah}\right\} \quad (17)$$

where $\mathbf{z}_{ah} = \tilde{w}_{ah}(\mathbf{x}_{ah} - \gamma_{aw}\bar{\mathbf{x}}_{ah})$. The corresponding covariance matrix is then as follows:

$$\begin{aligned}\boldsymbol{\Phi}_w = {} & \sigma_e^2\left(\sum_{a=1}^{k}\sum_{h=1}^{n_a}\mathbf{x}_{ah}\mathbf{z}'_{ah}\right)^{-1} \\ & \left(\sum_{a=1}^{k}\sum_{h=1}^{n_a}\mathbf{z}_{ah}\mathbf{z}'_{ah}\right)\left(\sum_{a=1}^{k}\sum_{h=1}^{n_a}\mathbf{x}_{ah}\mathbf{z}'_{ah}\right)^{-1} \\ & + \sigma_v^2\left(\sum_{a=1}^{k}\sum_{h=1}^{n_a}\mathbf{x}_{ah}\mathbf{z}'_{ah}\right)^{-1} \\ & \left\{\sum_{a=1}^{k}\left(\sum_{h=1}^{n_a}\mathbf{z}_{ah}\right)\left(\sum_{a=1}^{n_a}\mathbf{z}_{ah}\right)'\right\}\left\{\left(\sum_{a=1}^{k}\sum_{h=1}^{n_a}\mathbf{x}_{ah}\mathbf{z}'_{ah}\right)^{-1}\right\}'. \quad (18)\end{aligned}$$

The variance components are estimated using Henderson's Method 3 (Henderson 1953), to generate unbiased estimates even in the presence of correlated elements in the model. The estimators of the variance components are as follows:

$$\hat{\sigma}_{eH}^2 = (n - k - p + 1)^{-1}\sum_{a=1}^{k}\sum_{h=1}^{n_a}\hat{\varepsilon}_{ah}^2 \quad (19)$$

where $\{\hat{\varepsilon}_{ah}^2\}$ are residuals from the OLS regression of $(y_{ah} - \bar{y}_a)$ on $\{x_{ah1} - \bar{x}_{a.1}, ..., x_{ahp} - \bar{x}_{a.p}\}$ and $(\bar{y}_a, \bar{x}_{a.1}, ..., \bar{x}_{a.p})$ are the sample means in the a^{th} group.

$$\hat{\sigma}_{vH}^2 = n_*^{-1}\left[\sum_{a=1}^{k}\sum_{h=1}^{n_a}\hat{u}_{ah}^2 - (n - p)\hat{\sigma}_{eH}^2\right] \quad (20)$$

where $n_* = n - \mathrm{tr}[(\mathbf{X}'\mathbf{X})^{-1}\sum_{a=1}^{k} n_a^2\,\bar{\mathbf{x}}_a\,\bar{\mathbf{x}}'_a]$ with $\mathbf{X} = (\mathbf{x}_1, ..., \mathbf{x}_k)$, and the $\{\hat{u}_{ah}\}$ are the residuals from the OLS regression of y_{ah} on $\{x_{ah1}, ..., x_{ahp}\}$. For the model (7), the subscript a is replaced by b.

However, the Henderson's estimators above do not account for the sampling weights. To address this, an estimation technique has been proposed by You *et al.* (2003) which extends the Pseudo-EBLUP method by incorporating the weights in the estimation of the variance components. This is described in the next section.

5.2 The iterative weighted estimating equation method

The estimator proposed by You *et al.* (2003) is similar to the Pseudo-EBLUP estimator, except that it incorporates the sampling weights in the computation of the variance components, and it generates the parameter estimate $\boldsymbol{\beta}$ and the variance components by using an iterative weighted estimating equation (IWEE) approach. The authors derived the estimator of σ_e^2 and σ_v^2 as follows:

$$\begin{aligned}\hat{\sigma}_{ew}^{2(t)} &= \frac{\sum_{a=1}^{k}\sum_{h=1}^{n_a}\tilde{w}_{ah}[y_{ah} - \bar{y}_{aw} - (\mathbf{x}_{ah} - \bar{\mathbf{x}}_{aw})'\hat{\boldsymbol{\beta}}^{(t-1)}]^2}{\sum_{a=1}^{k}\left[(1 - \delta_a^2)\sum_{h=1}^{n_a}\tilde{w}_{ah}\right]} \\ &\equiv \tilde{\sigma}_{ew}^{2(t)}(\boldsymbol{\beta}) \quad (21)\end{aligned}$$

and

$$\begin{aligned}\hat{\sigma}_{vw}^{2(t)} &= \frac{1}{k}\sum_{a=1}^{k}\tilde{v}_{aw}^2 + \frac{\tilde{\sigma}_{vw}^{2(t-1)}}{k}\sum_{a=1}^{k}(\gamma_{aw} - 1)^2 + \frac{\tilde{\sigma}_{ew}^{2(t)}}{k}\sum_{a=1}^{k}\delta_a^2\gamma_{aw}^2 \\ &\equiv \tilde{\sigma}_{vw}^{2(t)}(\tilde{v}_w, \sigma_e^2, \sigma_v^2). \quad (22)\end{aligned}$$

The survey weighted estimates of $\boldsymbol{\beta}$, σ_e^2, σ_v^2 are obtained simultaneously by following iterative updating steps, t in the equation above stands for the t^{th} iteration. Since the variance components σ_v^2 and σ_e^2 are unknown, initial estimates for the iterative steps are generated by Henderson's method. Again, as for Pseudo-EBLUP, for the ELL regression model formulation (7), the subscript a is replaced by b.

This approach is similar to the probability-weighted iterative generalized least squares (PIWGLS) method proposed by Pfeffermann *et al.* (1998) for fitting multilevel models where the estimation process considered the unequal selection probabilities at each stage of sampling and involves iterating between the parameter $\boldsymbol{\beta}$ and the variance components until convergence. A model-based approach is also proposed by Pfeffermann, Moura and Silva (2006), which involves deriving the hierarchical model for given sample data as a function of the population model and the selection probabilities, and then fitting the sample model using Bayesian approach by use of Markov Chain Monte Carlo algorithm.

5.3 General survey regression method

Another approach to generate the estimator of the parameter β and its variance is the design-based methodology for fitting regression models (Lohr 1999). This

technique is currently used in the Stata, Sudaan, and WesVar package, for example. The estimator of $\boldsymbol{\beta}$ given below is the sample weighted regression estimator for a model with homoscedastic variance structure and uncorrelated observations in the population.

$$\hat{\boldsymbol{\beta}}_S = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{y}. \qquad (23)$$

This estimator is not derived under the model specified by (7) even under the homoscedastic variances for household errors. The linearized/robust variance estimate for $\hat{\boldsymbol{\beta}}_S$ is based on the design-based variance estimator for a total, given as,

$$\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}_S) = \mathbf{D}\left\{\frac{m}{m-1}\sum_{b=1}^{m}\left(\sum_{h=1}^{n_b} w_{bh}\mathbf{d}_{bh}\right)'\left(\sum_{h=1}^{n_b} w_{bh}\mathbf{d}_{bh}\right)\right\}\mathbf{D} \qquad (24)$$

where $\mathbf{d}_{bh} = \hat{e}_{bh}\mathbf{x}_{bh}$; $\hat{e}_{bh}$ is the residual from WLS regression; $\mathbf{x}_{bh}$ is a vector of the independent variables; w_{bh} is a sampling weight; $\mathbf{D} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}$; and $\mathbf{W}$ is a diagonal matrix of the sampling weights.

The General Survey Regression method differs from the other techniques in the computation of the estimates, and generates the estimates without computing the variance components, σ_v^2 and σ_e^2. As shown above, the equations for the estimator of the parameter $\boldsymbol{\beta}$ and its corresponding estimated covariance matrix only involve the sampling weights matrix $\mathbf{W}$. The estimated covariance matrix in (24) is often referred to as a sandwich estimator.

6. Comparison of the model fitting techniques

The ELL methodology is claimed to be a weighted GLS estimation procedure. However, as pointed out earlier, the sampling weights are not properly incorporated in the estimation process and this leads to non-interpretability of the elements in some matrices involved in the estimation, as well as asymmetry in the estimated covariance matrix. For the ELL method of estimating the variance components, the weights are accounted for only at the cluster level. The two ways (direct computation and heteroscedasticity model-based) that ELL use for generating the household level variance component do not incorporate the sampling weights. For direct computation, the household level variance component is determined from the residual of the survey-weighted (WLS) regression conducted at the preliminary step and the weighted estimate of the cluster level component. The heteroscedasticity based computation is based on modeling the square of the residuals from the WLS regression.

While the ELL methodology follows a GLS-like estimation procedure, the pseudo-EBLUP and IWEE method follow the Generalized Estimating Equation (GEE) procedure (Liang and Zeger 1986) using an exchangeable working correlation matrix, *i.e.*, all the off-diagonal elements of the correlation matrix within clusters are equal, and in Pseudo-EBLUP and IWEE are equal to $\sigma_v^2/(\sigma_v^2 + \sigma_e^2)$. An exchangeable or equicorrelated working correlation matrix is one of the common working correlation matrices presented in the paper of Horton and Lipsitz (1999) when reviewing different software for fitting GEE regression models.

The two procedures, Pseudo-EBLUP and IWEE, both incorporate the sampling weights in the estimation of the parameter $\boldsymbol{\beta}$ and the corresponding standard error, although the Pseudo-EBLUP method uses Henderson's method in the estimation of the variance components. While Henderson's method generates unweighted estimates of the variance components, the IWEE method incorporates the sampling weights iteratively from estimation of variance components for computation of standard error of the estimate of the regression parameter.

There is a very limited published literature on the application to real data sets of the Pseudo-EBLUP and IWEE methods. Those that there are consider the clusters as the small area, and often use the data in Battese, Harter and Fuller (1988), whose data set contains information on hectares of corn and soybeans per segment for counties in North Central Iowa and assumes simple random sampling within areas or clusters. An exception is the recent paper by Militino, Ugarte, Goicoa and Gonzalez-Audicana (2006), which applies Pseudo-EBLUP to estimating the total area occupied by olive trees in Navarra, Spain, where (as in Battese *et al.*) the units are self weighting. Generally for poverty estimation, Pseudo-EBLUP and IWEE techniques must be applied in more complex situations, since sampling clusters and small areas are not identical and the sample is not self weighting. In the example in the next section, the clusters (barangay) are different from the small areas (municipalities), the clusters are sub-units of the small area and the sampling scheme is not self weighting.

The GSR method is one of the least complicated estimation procedures as it employs a weighted least squares procedure using the sandwich estimator for estimating the variance of the estimator of the regression parameter. As mentioned earlier, this method differs from the other techniques in that the estimate of the regression parameters and their corresponding standard errors are generated without computing the variance components.

Based on the discussion above, for all the techniques considered, the survey-based estimation procedure for the parameter $\boldsymbol{\beta}$ and its corresponding standard error are theoretically sound given their assumptions, except for the ELL method where there are some inconsistencies in the estimation of parameters $\boldsymbol{\beta}$ and the covariance of $\hat{\boldsymbol{\beta}}$.

7. Application to real data

In this section, the four different regression techniques (one of which contains two variants of ELL) are compared using the Philippine 2000 Family Income and Expenditure Survey (FIES). The FIES data is a nationwide survey undertaken by the Philippines National Statistics Office (NSO) every three years. The survey gathers details on family income and expenditure as well as information affecting income and expenditure. Selected households are interviewed in two separate operations, each covering a half-year period, in order to allow for seasonal patterns in income and expenditure. For FIES 2000 the interviews were conducted in July 2000, for the period 01 January to 30 June and January 2001 for the period 01 July to 31 December. The sample design for FIES used a multi-stage stratified random sampling technique. Barangays are the primary sampling units (PSUs) and are stratified into urban and rural within each province and selected using systematic sampling with probability proportional to size. Large barangays are further divided into enumeration areas and subjected to further sampling before the final stage in which households are systematically sampled from the 1995 Population Census List of Households. Interview non-response was only 3.4 percent, with 39,615 of the sample households being successfully interviewed in both survey visits. Deterministic imputation was done to address item non-response, *i.e.*, entry for a particular missing item is deduced from other items in the questionnaire.

The auxiliary variables used in this paper are adopted from the variables included in the model formulated by Haslett and Jones (2005) that was fitted without using POVMAP for the small area poverty mapping project in the Philippines. The auxiliary variables included both household characteristics and municipal means (in which the household data used have the same value for every sampled household in a given municipality, *i.e.*, small area). These auxiliary variables are not only derived from the FIES data but also from the Philippine 2000 Labor Force Survey (LFS) and Census of Population and Housing (CPH). The LFS collects socioeconomic characteristics of the population over 15 years old. It is conducted on a quarterly basis by the NSO by personal interview, using previous week as reference period. Being part of the Integrated Survey of Households (NSCB 2000), the July 2000 and January 2001 surveys used the same sample of households as the 2000 FIES. Thus the two data sets can be merged to form a richer set of auxiliary variables. Additional auxiliary variables were also taken from the 2000 CPH in the form of municipal means. Census variables in both the short and long form were averaged at municipal level to create new data sets that could be merged with the set of auxiliary variables from FIES and LFS.

Presented in Tables 1, 2, and 3 are the computed estimates of the parameter $(\boldsymbol{\beta})$ and the corresponding standard errors as well as the estimates of the variance components at the national, regional and provincial levels, respectively. Table 2 is one of the regional models of the 16 models fitted at the regional level (there are 16 regions in the Philippines in the year 2000). Similarly, Table 3 shows one of the provincial models of the 20 models formulated for 20 selected provinces. To standardize comparison, exactly the same set of predictor variables are used for all the different model fitting techniques. (There are five sets of parameter estimates, although there are only four basic methods considered, because ELL is used both with and without heteroscedasticity.) Note that in practice when ELL is applied, the survey data is often subdivided and separate models fitted to each subsample, *e.g.*, to each regionally-based stratum as the 16 regions in the Philippines or even provincial level models. This can lead to overfitted models and downwardly biased standard errors for small area estimates. For the analysis here, a single model (or the national level model) has been fitted. In practice intermediate models with some but not all possible regional effects seem to work best. See for example Haslett and Jones (2005).

To assess the differences of the estimates generated from the different techniques, an informal comparison of the "significance" of the different estimates of β is conducted by subtracting from the estimate by one method the mean of the other methods' estimates, then dividing by the standard error of the one method. At the national level (Table 1), estimates of the regression coefficients generated from the different methods are significantly different from each other for a number of the independent variables. GSR tends to generate estimates of the regression coefficients for the majority of the variables that are significantly different from the other methods. As pointed out earlier, the GSR estimator is the sample weighted regression estimator for a model with homoscedastic variance structure and uncorrelated observations in the population and hence this estimator is not derived under the model specified by (7). However, it is the most conservative as it generates the highest standard error for all the household level characteristics. On the other hand, the IWEE method has the highest estimated standard error for all the municipal means. The ELL_H (ELL with heteroscedasticity) method can be considered to be the least conservative since it produces the lowest standard errors for all the estimated regression coefficients of the household level characteristics as well as for the municipal means, except for two variables where GSR generated the smallest estimates. As to the estimates of the variance components, the ELL method generates the smallest estimated cluster level variance, which is about 92% of the Pseudo-EBLUP method and 86% of the IWEE method. As to the household level variance, the IWEE method generates the smallest estimate.

Table 1
National level estimates of regression parameters with the standard errors and the variance components for the four techniques. *Different value for each household (mean = 0.1576633) **Based from the ELL results

Explanatory	ELL(no hetero)		ELL(w/ hetero)		Pseudo-EBLUP		IWEE		GSR	
Variables	Beta	Std. Error	Beta	Std. Error	Beta	Std. Error	Beta	Std. Error	Beta	Std. Error
famsize	-0.11867	0.00181	-0.12034	0.00165	-0.11875	0.00183	-0.11888	0.00180	-0.11405	0.00216
famsizesqc	0.00937	0.00039	0.00981	0.00036	0.00938	0.00039	0.00939	0.00038	0.00898	0.00044
type_mult	0.03876	0.01697	0.03703	0.01588	0.03699	0.01717	0.03466	0.01692	0.11460	0.02194
per_kids	-0.20342	0.01476	-0.20818	0.01322	-0.20293	0.01491	-0.20216	0.01467	-0.22864	0.01617
roof_light	-0.06314	0.01291	-0.05808	0.01056	-0.06263	0.01306	-0.06175	0.01287	-0.09251	0.01413
per_61up	-0.09402	0.01420	-0.08331	0.01371	-0.09392	0.01435	-0.09389	0.01412	-0.09705	0.01698
roof_strong	0.05882	0.01135	0.05633	0.00962	0.05944	0.01148	0.06030	0.01132	0.03118	0.01293
wall_light	-0.05459	0.01182	-0.04979	0.00975	-0.05426	0.01195	-0.05392	0.01178	-0.06286	0.01353
wall_salvaged	-0.10814	0.02505	-0.11327	0.02058	-0.10748	0.02533	-0.10607	0.02495	-0.15702	0.02925
wall_strong	0.14248	0.01051	0.12964	0.00910	0.14274	0.01063	0.14319	0.01047	0.12662	0.01284
fa_xs	-0.17052	0.00941	-0.16756	0.00782	-0.17144	0.00952	-0.17236	0.00939	-0.14213	0.01110
fa_s	-0.08368	0.00861	-0.08242	0.00725	-0.08403	0.00871	-0.08454	0.00857	-0.06667	0.00964
fa_l	0.09016	0.00908	0.08478	0.00792	0.09065	0.00918	0.09106	0.00904	0.07848	0.01047
fa_xl	0.16959	0.01104	0.15404	0.00992	0.17034	0.01117	0.17121	0.01100	0.14300	0.01334
fa_xxl	0.27072	0.01144	0.24485	0.01094	0.27172	0.01157	0.27274	0.01140	0.23913	0.01457
fa_xxxl	0.36190	0.01371	0.31369	0.01286	0.36270	0.01387	0.36382	0.01367	0.32123	0.02025
all_eled	0.19084	0.01535	0.20497	0.01307	0.19031	0.01551	0.18964	0.01527	0.21344	0.01831
all_hsed	0.42325	0.01250	0.43771	0.01083	0.42192	0.01263	0.42024	0.01244	0.48180	0.01475
all_coed	1.21591	0.01371	1.29368	0.01379	1.21324	0.01386	1.20935	0.01366	1.35022	0.01827
dom_help	0.60207	0.01629	0.61218	0.01886	0.60035	0.01645	0.59733	0.01620	0.70307	0.02656
head_male	-0.05878	0.00988	-0.04581	0.00932	-0.05862	0.00998	-0.05819	0.00982	-0.07410	0.01173
no_spouse	-0.09367	0.00987	-0.07376	0.00917	-0.09361	0.00997	-0.09351	0.00981	-0.09599	0.01123
hou_9600	0.28537	0.07654	0.25643	0.07375	0.28871	0.07911	0.28783	0.08066	0.31956	0.07941
hea_rel_mus	0.09058	0.02645	0.10859	0.02507	0.09753	0.02728	0.09731	0.02782	0.10196	0.02737
Per_eng	0.17273	0.06529	0.14561	0.06298	0.17782	0.06754	0.17799	0.06887	0.17076	0.06407
Hou_coelpg	0.37463	0.04348	0.39784	0.04210	0.37934	0.04494	0.37792	0.04581	0.42682	0.03711
Hou_own_ref	0.17716	0.10497	0.18342	0.10178	0.17189	0.10843	0.17329	0.11055	0.13791	0.09766
Hou_own_tel	1.39287	0.13356	1.42109	0.12987	1.38551	0.13723	1.38974	0.13989	1.23506	0.13019
Per_wor_prh	0.46957	0.15484	0.40302	0.14926	0.47517	0.16006	0.47208	0.16317	0.50814	0.15210
Per_ind_52	-0.76245	0.21708	-0.78120	0.21073	-0.76326	0.22410	-0.76307	0.22849	-0.73294	0.21214
const	9.54013	0.05525	9.54456	0.05290	9.53566	0.05698	9.53594	0.05791	9.52622	0.05613
Variance Components Estimate	HH level	Cluster level	HH level	Cluster level	HH level	Cluster level	HH level	Cluster level	HH** level	Cluster** level
	0.18461	0.04741	NA*	0.04741	0.18820	0.05172	0.18185	0.05498	0.18461	0.04741

Table 2
Regional level estimates of regression parameters with the standard errors and the variance components for the four techniques. *Different value for each household (mean = 0.18930) **Based from the ELL results

Explanatory	ELL(no hetero)		ELL(w/ hetero)		Pseudo-EBLUP		IWEE		GSR	
Variables	Beta	Std. Error	Beta	Std. Error	Beta	Std. Error	Beta	Std. Error	Beta	Std. Error
famsize	-0.12327	0.00760	-0.12934	0.00689	-0.12377	0.00752	-0.12380	0.00749	-0.11786	0.00997
famsizesqc	0.01096	0.00164	0.01190	0.00147	0.01101	0.00163	0.01102	0.00162	0.01030	0.00195
dom_help	0.81037	0.08873	0.75624	0.10986	0.80727	0.08784	0.80708	0.08751	0.84490	0.08911
wall_light	-0.06808	0.04289	-0.06390	0.03743	-0.06020	0.04272	-0.05973	0.04257	-0.14472	0.04226
wall_strong	0.13761	0.03745	0.15212	0.03469	0.14514	0.03737	0.14560	0.03725	0.06116	0.04249
fa_xs	-0.22074	0.04910	-0.22368	0.04518	-0.22723	0.04875	-0.22761	0.04858	-0.14856	0.05665
fa_s	-0.13540	0.03840	-0.12255	0.03344	-0.13775	0.03805	-0.13789	0.03791	-0.11059	0.04538
fa_l	0.09484	0.03709	0.08894	0.03429	0.09590	0.03676	0.09597	0.03663	0.08529	0.04122
fa_xl	0.16627	0.04315	0.15519	0.04072	0.16938	0.04284	0.16958	0.04269	0.13698	0.04897
fa_xxl	0.33706	0.04545	0.31196	0.04829	0.34173	0.04516	0.34201	0.04500	0.29156	0.05148
fa_xxxl	0.33103	0.06185	0.30377	0.06029	0.33762	0.06134	0.33801	0.06111	0.26052	0.06635
all_hsed	0.33987	0.05253	0.35591	0.04783	0.33807	0.05209	0.33796	0.05189	0.35776	0.04843
all_coed	1.21824	0.05734	1.24762	0.05842	1.20787	0.05692	1.20726	0.05671	1.32979	0.06227
per_kids	-0.24699	0.06440	-0.24047	0.05846	-0.24439	0.06371	-0.24424	0.06347	-0.27423	0.07050
per_61up	-0.14609	0.06126	-0.15938	0.05787	-0.14703	0.06063	-0.14708	0.06040	-0.13525	0.07124
hou_9600	1.13985	0.49103	1.27035	0.47888	1.14320	0.52137	1.14357	0.52172	1.07509	0.51937
Hou_own_ref	1.45233	0.24550	1.51020	0.23864	1.44986	0.26072	1.44985	0.26089	1.44779	0.23585
const	9.36877	0.20322	9.32363	0.19660	9.36597	0.21502	9.36569	0.21512	9.41385	0.21430
Variance Components Estimate	HH level	Cluster level	HH level	Cluster level	HH level	Cluster level	HH level	Cluster level	HH** level	Cluster** level
	0.19544	0.03073	NA*	0.03073	0.19052	0.03728	0.18902	0.03748	0.19544	0.03073

Table 3
Provincial level estimates of regression parameters with the standard errors and the variance components for the four techniques. *Different value for each household (mean = 0.23749) **Based from the ELL results

Explanatory	ELL(no hetero)		ELL(w/ hetero)		Pseudo-EBLUP		IWEE		GSR	
Variables	Beta	Std. Error	Beta	Std. Error	Beta	Std. Error	Beta	Std. Error	Beta	Std. Error
famsize	-0.1450	0.0175	-0.1489	0.0156	-0.1452	0.0179	-0.1449	0.0171	-0.1413	0.0097
famsizesqc	0.0090	0.0063	0.0124	0.0067	0.0091	0.0065	0.0090	0.0062	0.0085	0.0055
fa_xs	-0.4549	0.1126	-0.3816	0.1010	-0.4552	0.1149	-0.4546	0.1095	-0.4479	0.0718
fa_s	-0.2550	0.0976	-0.2653	0.0794	-0.2545	0.0995	-0.2555	0.0951	-0.2693	0.1198
wall_light	-0.2055	0.0945	-0.1474	0.0778	-0.2057	0.0965	-0.2058	0.0919	-0.2063	0.1070
all_hsed	0.4007	0.1643	0.3531	0.1448	0.4015	0.1673	0.4006	0.1601	0.3891	0.1585
all_coed	1.5411	0.1677	1.8202	0.1769	1.5429	0.1709	1.5429	0.1635	1.5439	0.2326
Hou_own_tel	3.4373	1.0270	3.2630	1.0582	3.4265	1.0622	3.4274	0.9871	3.4392	0.5733
Per_wor_prh	-1.1075	1.1933	-1.5801	1.2008	-1.1049	1.2327	-1.1056	1.1483	-1.1150	0.8729
const	10.0976	0.1480	10.0798	0.1279	10.0988	0.1517	10.0981	0.1435	10.0872	0.1373
Variance Components Estimate	HH level	Cluster level	HH level	Cluster level	HH level	Cluster level	HH level	Cluster level	HH** level	Cluster** level
	0.25753	0.01871	NA*	0.25753	0.26682	0.02079	0.24498	0.01671	0.25753	0.01871

At the regional level, estimates of the regression coefficients are generally similar for all the different estimation methods, except that the GSR and/or ELL_H methods generated estimates for a few variables which were significantly different from the other methods. Similar to the national level estimated standard errors, GSR also tends to be the most conservative method for the majority of the regional level models - it generated the highest estimated standard errors for most of the regression coefficients of the household characteristics. IWEE has the highest estimated standard error for most of the coefficients of the municipal means. The ELL_H method produces the lowest standard errors for the majority of the regression coefficients of the household characteristics and municipal means. The ELL method tends to generate the smallest estimated cluster level variance with ratios to Pseudo-EBLUP and IWEE ranging from around 82% to 100%. The IWEE method still has the smallest household level variance.

Similar to the regional level estimates, the regression coefficients' estimates at the provincial level are similar except for some discrepancies from the GSR and ELL_H estimates. For the estimated standard errors of the regression coefficients, the ELL_H still produces the lowest estimates for the majority of the coefficients of the household characteristics; however, the GSR method (instead of the ELL_H method) now produces the lowest estimated standard error for the majority of the municipal means. The ELL method still tends to generate the smallest estimated cluster level variance for most provinces with the smallest ratio to Pseudo-EBLUP about 53% and to IWEE about 48%. For a number of provinces, IWEE tends to generate the smallest estimated cluster level variance. For the household level variance, IWEE still generated the smallest estimate. Generally, estimates of the cluster level variance tend to be more variable at the provincial level which is due to smaller sample sizes.

For small area estimates of poverty, after the regression model is applied to census data, estimated standard errors in the regression are only one part of the small area estimates' standard errors. There is also variation at the cluster level in (7) that needs to be considered (to different degrees depending on the level of aggregation used to construct the small areas) and there is variation at household level too. These additional sources of variation can be assessed via the estimated variance components. As shown above, regardless of the level (national, regional and provincial) at which the model is formulated, the IWEE method generates the smallest household level variance, while the ELL method generates the smallest cluster level variance. Since the cluster level variation usually makes a much larger contribution to the estimated standard error at the small area level, ELL is again the least conservative. We note that the household level variance under the ELL method with heteroscedasticity model varies from one unit to another, hence, the mean value is reported, and that the estimated R^2 for the heteroscedasticity model is negligible, $R^2 =$ 0.03 even at the national level, so that in terms of regression model fit at least it may offer few advantages for this data set. In our experience with applying the ELL method we have found that heteroscedasticity modeling is unnecessary.

Returning to the regression (*i.e.*, the estimates generated for $\boldsymbol{\beta}$ and the estimated standard error for the different techniques), IWEE is the method that best incorporates the sampling weights from the computation of the variance components necessary for the generation of small area estimates and their estimated standard errors. In terms of implementation, the GSR method would generally be the simplest option as it is available for example in packages such as Stata, Sudaan or WesVar. The ELL method combines sampling weights and covariance structure in a way that is non-standard in that it uses an estimate of $\mathbf{W}_b\mathbf{V}_b^{-1}$ in (8) and (9) to produce an asymmetric estimated covariance matrix for the estimates of $\boldsymbol{\beta}$ and for estimating $\boldsymbol{\beta}$ itself. For estimating $\boldsymbol{\beta}$ this would be acceptable if the asymmetric matrix were a generalized inverse of the correct covariance matrix. It is however clearly not acceptable as an

estimated covariance matrix, a problem ELL attempt to circumvent (*e.g.*, in the World Bank's POVMAP software) by averaging each of the relevant pairs of off-diagonal elements to meet the necessary condition that a covariance matrix be symmetric.

Generally in the ELL method of poverty estimation only variables matching in terms of average and standard deviation in both survey and census plus census averages can be used. This is because, after the regression model has been fitted to the survey data, in the second phase it is applied to the census data as a predictor at household level, *i.e.*, the regression equation (however it has been estimated) is used to find predicted values of per capita income or expenditure for each census household, generated via

$$\hat{Y}_{bh} = \mathbf{x}'_{bh}\hat{\boldsymbol{\beta}} + \hat{v}_b + \hat{e}_{bh} \qquad (25)$$

using imputed values of v_b and e_{bh} (based for example on bootstrap sampling from their survey estimates). Here $\mathbf{x}_{bh}$ are auxiliary variables from the census. Poverty indices are typically based on non-linear functions of log-income or log-expenditure, so the predictions from (25) are transformed appropriately before averaging over each small area. Note that in practice v_b can be estimated for the sampled clusters, but the sample and census codes usually do not match so these cannot be identified in the census, and it is the bootstrap (by selecting from the sampled barangays, *i.e.*, PSUs) that provides imputed values for all barangays; a parallel comment applies to $\hat{e}_{bh}$ for households within clusters. The general benefit of using census data in this way (as ELL does) is that the predictor variables can be used for all census households (of which there are many) not just those in the survey, thereby increasing accuracy of the small area estimates (conditional on the model being correct). Note that the estimates in (25) remain unbiased even if v_b and e_{bh} are not included in the prediction itself, but the variance estimate for small area a needs to be computed based on equation (25) so that it incorporates the necessary additional variation at cluster and household levels.

In poverty estimation, we are interested in area-level summaries of non-linear functions of $\hat{Y}_{bh}$, for example, whether it is below the poverty line (poverty incidence) and poverty gap rather than the regression fitting per se. It is instructive here to examine the effects of model uncertainty on area mean estimates

$$\bar{y}_a = \bar{\mathbf{x}}'_a \hat{\boldsymbol{\beta}} \qquad (26)$$

where $\bar{\mathbf{x}}_a$ is the population (*i.e.*, census) mean for area a of the covariates including the constant 1, after the regression model has been applied to the census data as in phase 2 of ELL. By similarly averaging (7) to get the true mean $\bar{Y}_a$, subtracting from (26), and applying the variance operator, we get the prediction error variance equation:

$$V(\bar{y}_a - \bar{Y}_a) = \bar{\mathbf{X}}_a \boldsymbol{\Phi}_w \bar{\mathbf{X}}'_a + \frac{1}{N_a^2}\sum_{b=1}^{m} N_b^2\, \sigma_v^2 + \frac{1}{N_a}\sigma_e^2 \qquad (27)$$

where N_a is the population size at a particular level of aggregation, N_b is the population size in each cluster, $\boldsymbol{\Phi}_w$ is the variance-covariance matrix of the regression coefficient estimates, and (σ_v^2, σ_e^2) are the cluster and household level variance components, respectively. Note that estimating this prediction error variance requires estimates of the variance components, but any bias caused by uncertainty in these would be a second order effect (see Prasad and Rao 1990).

Based on (27), the extent of the influence of the survey based regression model and other variance components (cluster and household level) on the accuracy of the final small area estimates can be compared for any fitting technique and/or levels of aggregation. Generally, it is either the regression model (via the estimate of the regression parameters) or the cluster effect that dominates the estimated accuracy of the computed small area estimate. Using the national level model in Table 1 and the survey data (instead of the census) auxiliary variables to estimate the first term in (27), shows that the extent to which the regression model effect contributes to small area estimate variance increases markedly as household data are more aggregated - about 0.25% at the municipal level, 20% at the provincial level and 70% at the regional level. In other words, the more aggregated the data into larger areas, the greater the dominance of the regression model parameter uncertainty, regardless of the regression fitting method. This is as expected because even at high levels of aggregation, the contribution to the overall variance from the model effect depends on the average covariate values, not on the population size. This is the reason that, at the most aggregated regional level, small area techniques usually offer little improvement over direct estimates. This is also why it is important (as this paper has done) to examine in detail the regression fitting procedures applied in small area estimation of third world poverty.

The effect of cluster level variation is different: at lower levels of aggregation (*e.g.*, municipality) the computed variance of the small area estimates are dominated by the cluster component of variance or cluster level effect, *i.e.*, for small areas (other than regional estimates) the variance component, not the regression model, has the greatest impact on the value of the standard error of the small area estimates. Consequently, the accuracy of estimates of variance components especially at cluster level can be crucial to accurate estimation of standard error of small area estimates at the aggregation level at which they are most useful (for example at municipal level in the Philippines). Again, this is why the method used for phase 1 fitting for variance components as discussed in this paper, are critical to small area estimation of poverty.

Presented in Tables 4-6 are Kruskal-Wallis (KW) tests (Siegel 1956) for the various fitting methods conducted on the estimated variances at the municipal (Table 4), provincial (Table 5) and regional (Tables 6) levels. In Table 4 significant differences exist among the variance estimates generated by the various small area techniques, as shown by the p-values of the Kruskal-Wallis statistics. Multiple comparison of mean ranks shows the Pseudo-EBLUP and IWEE methods have variance estimates at cluster level that are significantly higher than the other methods, but not significantly different from each other (although for the IWEE method the Z-value for the difference from average rank is in general rather higher than all the others).

The ELL method and the GSR method generate signifycantly lower and similar variance component estimates. This is principally because we used the ELL variance components estimation technique in generating variance components for the GSR method (because GSR does not usually estimate variance components), although the residuals we used were not identical for the two regression fitting methods. As expected, at the municipal level for which small area estimates were used in practice, the cluster effect (rather than regression coefficient uncertainty) is generally the dominant part of the small area variance estimates. Since the ELL and GSR methods have similar cluster level variance, their corresponding variance estimates at small area also tend to be similar. Explicitly, observe from Table 4 that the ranking of the variance estimates generally conforms with the ranking of the cluster effects.

In poverty estimation, estimates at higher levels of aggregation, such as those in Table 5 and 6, are generally carried out for comparison with direct survey estimates at these more aggregated levels, even though they are not particularly useful for aid allocation. The results do however, support those indicated for lower level of aggregation. In Table 5 and Table 6, the estimated variances for the poverty estimates generated by the different techniques are not significantly different from each other at the provincial and regional level, an effect that is partially due to the small number of provinces and even smaller number of regions. The variances and hence the standard errors may not be significantly different from each other, but it is worth noting that the GSR method tends to generate the smallest estimated standard error for the regression model and in turn the smallest variance estimate for poverty at the regional level, even though GSR generates higher standard errors for the individual regression coefficients (corresponding to the diagonal elements only in the estimated covariance matrix of $\hat{\boldsymbol{\beta}}$). As expected, at an even higher level of aggregation for all methods, the relative effect of the regression component is more pronounced.

The general conclusion is that, whether fitting survey data alone or using survey based regression parameter estimates in conjunction with census data, it is crucial not only to find a suitable model (*i.e.*, set of regressors) based on an adequate sample size, but also to get sound estimates of the regression parameters and their standard errors under this model as well as good estimates of the variance components at all relevant levels of aggregation. Usually the relevant levels of aggregation are determined via the survey design, rather than simply through the level at which small area estimates are sought, although the number of levels need not be limited to two (*e.g.*, to cluster-level and household-level).

Survey data, whether used for poverty estimation or in other context, also introduces problems involving survey weights that can be important not only for regression parameter estimation (and their estimated standard errors) but also for estimating variance components. Incorporating survey weights into regression models with correlated data introduces problems because it is the population correlation as it applies to the weighted survey data that needs to be properly modeled, so that weighting correlation matrices using matrix multiplication (as ELL do) is not technically adequate (see Appendix).

For the Philippine data and for the specified list of regressors, regardless of which of the four methods are used, parameter estimates were very similar, which suggests that the more important issue is possible underestimation of standard errors of parameter estimates and of variance components particularly at cluster level. ELL is the least conservative in that it gave the lowest estimates of both variance measures, and in this respect (as with its use of asymmetric estimated covariance matrices) some caution may be warranted with the regression and variance component aspects of the ELL technique. GSR gave similar estimates of standard errors for the small area estimates to ELL when using the same technique for variance components, despite having higher standard errors (and using a sound covariance matrix) for regression parameters. This is because when there is less aggregation, the level at which most small area estimates are actually used, variance components dominate.

The Pseudo-EBLUP and IWEE methods incorporate survey weights correctly (given a suitable choice of pseudo-likelihood and hence GEE) and gave larger (*i.e.*, more conservative) estimates of cluster level variance components. This suggests that these two methods and particularly IWEE are among the best of the currently available methods, not necessarily for estimating regression equations (where availability of standard software may give GSR an advantage), but for estimating the crucial variance components.

Table 4
Kruskal-Wallis test for estimated variances at the municipal level (N = 1,243)

SAE	Cluster Effect			Beta Effect			Variance		
Techniques	Median	Mean Rank	Z	Median	Mean Rank	Z	Median	Mean Rank	Z
ELL(no hetero)	0.002843	2,961.2(a)	-3.22	0.0002311	3,067.3(ab)	-0.89	0.00318	2,963.4(a)	-3.18
ELL(w/ hetero)	0.002843	2,961.2(a)	-3.22	0.0002128	2,802.0(c)	-6.72	0.00316	2,930.8(a)	-3.89
Pseudo-EBLUP	0.003094	3,229.4(b)	2.67	0.0002449	3,257.5(ad)	3.28	0.00346	3,241.3(b)	2.93
IWEE	0.003294	3,426.9(b)	7.01	0.0002529	3,364.5(d)	5.64	0.00366	3,441.3(b)	7.32
GSR(Stata)	0.002843	2,961.2(a)	-3.22	0.0002311	3,048.7(b)	-1.3	0.00317	2,963.1(a)	-3.18
Overall		3,108			3,108			3,108	
KW Statistic	H = 69.92	(P = 0.000)		H = 72.19	(P = 0.000)		H = 78.06	(P = 0.000)	

Table 5
Kruskal-Wallis test for estimated variances at the provincial level (N = 83)

SAE	Cluster Effect			Beta Effect			Variance		
Techniques	Median	Mean Rank	Z	Median	Mean Rank	Z	Median	Mean Rank	Z
ELL(no hetero)	0.0002518	200.3	-0.65	0.0001162	207.7	-0.03	0.00039	202.3	-0.48
ELL(w/ hetero)	0.0002518	200.3	-0.65	0.0001095	190.1	-1.52	0.00038	196.3	-0.99
Pseudo-EBLUP	0.000274	214.9	0.59	0.0001239	224.2	1.37	0.00042	217.1	0.78
IWEE	0.0002916	224.2	1.38	0.0001287	234.1	2.22	0.00045	227.8	1.68
GSR (Stata)	0.0002517	200.3	-0.65	0.00010	184	-2.04	0.00037	196.4	-0.98
Overall		208			208			208	
KW Statistic	H = 2.82	(P = 0.589)		H = 10.61	(P = 0.031)		H = 4.48	(P = 0.344)	

Table 6
Kruskal-Wallis test for estimated variances at the regional level (N = 16)

SAE	Cluster Effect			Beta Effect			Variance		
Techniques	Median	Mean Rank	Z	Median	Mean Rank	Z	Median	Mean Rank	Z
ELL(no hetero)	0.000050	38.2	-0.45	0.000077	40.9	0.08	0.00013	39.3	-0.23
ELL(w/ hetero)	0.000050	38.2	-0.45	0.000073	35.1	-1.05	0.00012	37	-0.67
Pseudo-EBLUP	0.000055	42.6	0.4	0.000082	46.9	1.23	0.00014	44	0.67
IWEE	0.000058	45.3	0.93	0.000085	50.1	1.85	0.00015	46.6	1.17
GSR(Stata)	0.000050	38.2	-0.45	0.000070	29.6	-2.1	0.00013	35.6	-0.94
Overall		40.5			40.5			40.5	
KW Statistic	H = 1.30	(P = 0.861)		H = 8.36	(P = 0.079)		H = 2.58	(P = 0.630)	

Of course, such considerations (while central) need to be predicated by adequate data cleaning, sound matching of possible regressor variables (in terms of mean, variance, and meaning) between survey and census where census data is also being used. Also needed are the proper, time consuming consideration of a wide range of possible regressor variables and recognition of the limits placed on subdividing survey data by small sample sizes, since all estimated standard errors for both regression parameter and small area estimates (whatever method is used for fitting the variance component estimate) are conditional on the regression model being correct.

8. Conclusion and recommendation

There is a great need for sound poverty statistics in order to effectively monitor interventions and assistance to various impoverished localities. Small area estimation techniques are one methodology that is being used to provide such statistics. In this sense the issues raised in this paper concerning the accuracy of the small area estimates are not simply an academic issue but are central to the Millennium Development Goals and to aid allocation in what is a multi-billion dollar industry.

In this paper, we have considered four estimation techniques for fitting regression models using survey data and related them to small area poverty estimation. We have shown that although differences in estimates are insufficient to invalidate the published national studies, the most frequently implemented survey data fitting technique, ELL with heteroscedasticity, recommended by the World Bank, has some limitations since (like its homoscedastic version) it lacks sound theoretical underpinning. Replacing the survey fitting part of the ELL method is recommended. For the other methodologies considered (the Pseudo-EBLUP, IWEE, and the GSR method), all have valid theoretical basis mathematically and the results generated can be clearly interpreted once the assumptions have been checked. The different methodologies when applied to complex weighted survey data from the Philippines indicate that for variance component estimation from survey data and hence for small area estimation at a fine level, Pseudo-EBLUP and particularly IWEE are likely to be better than the GSR or the ELL methods, although GSR is sound and easy to use because it is available in off-the-shelf software.

We have also shown that at the level where small area estimation is actually used for aid allocation, the variance estimate of the small area tends to be dominated by the cluster level variance rather than by the accuracy of the regression parameter estimates. Hence, it is particularly important that the cluster-level component of variance (and, if fitted as recommended, any small area level variance component) is properly estimated. It is also important that the regression model used in the generation of small area estimates (including choice of suitable regressors) is appropriate. Essentially, at lower levels of aggregation it is the variance components that dominate the standard error of the small area estimates, so that the estimation of the variance components is critical whatever the choice of aggregation level. Sound survey-based regression method, good choice of regression variables, and care with sample size (especially if separate regression models are fitted to subsets of survey data), also remain central to sound small area estimation of third world poverty.

Acknowledgements

The authors would like thank the referees and the Associate Editor for their careful reading of the manuscript and for their helpful suggestions.

Appendix

In footnote 8 of the Elbers *et al.* (2002) World Bank working paper and implicitly in Elbers *et al.* (2003) in *Econometrica*, the covariance of the error process is denoted $\mathbf{\Omega}$ and it is stated that $\mathbf{W\Omega}^{-1} = \mathbf{P}^T\mathbf{P}$ where $\mathbf{W}$ is 'a weighting matrix of expansion factors'. In the notation of Section 4 above, $\mathbf{W}$ is block diagonal with or diagonal with diagonal blocks $\mathbf{W}_b$, and $\mathbf{\Omega}$ is block diagonal with diagonal blocks $\mathbf{V}_b$.

However, either $\mathbf{W}$ and $\mathbf{\Omega}$ (or $\mathbf{\Omega}^{-1}$) are non-conformable (with weighting factors in $\mathbf{W}$ at cluster level and the observations and hence $\mathbf{\Omega}^{-1}$ at individual level), or if conformable $\mathbf{W\Omega}^{-1}$ is generally asymmetric (even if W is diagonal) unless $\mathbf{W}$ is a simple multiple of the identity matrix, *i.e.*, $\mathbf{W} = \sigma^2\mathbf{I}$.

Hence, $\mathbf{W\Omega}^{-1}$ does not equal $\mathbf{P}^T\mathbf{P}$ as has been claimed since $\mathbf{P}^T\mathbf{P}$ is symmetric in general and $\mathbf{W\Omega}^{-1}$ is not. Making $\mathbf{W\Omega}^{-1}$ symmetric by adding it to its transpose and dividing by two, as is done in the World Bank PovMap software, is not a technically adequate solution to this problem. (Note that even in the simple case where $\mathbf{W}$ and $\mathbf{\Omega}^{-1}$ are conformable, and $\mathbf{W}$ is diagonal but not all diagonal elements are equal, $\mathbf{W\Omega}^{-1}$ is not diagonal because it has every element of row i of $\mathbf{\Omega}^{-1}$ multiplied by w_i (where w_i is the i^{th} diagonal element of $\mathbf{W}$) but the i^{th} column does *not* have every element multiplied by an identical weight.)

Putting this issue of symmetry to one side, and using $\mathbf{P}^T\mathbf{P}$ in place of $\mathbf{W\Omega}^{-1}$, ELL seem to be claiming that comparing their 'sample survey adjusted weighted GLS estimator' to the 'unadjusted GLS' estimator implies that instead of using $\mathbf{\Omega}^{-1}$ as the underlying metric (*i.e.*, the inverse of the relevant covariance matrix), a weighted version namely $\mathbf{W\Omega}^{-1}\mathbf{W}^T$ should be used. This creates no asymmetry issue in itself (provided $\mathbf{P}^T\mathbf{P}$ were used in place of $\mathbf{W\Omega}^{-1}$). However, even if $\mathbf{W}$ were diagonal and $\mathbf{P}^T\mathbf{P}$ used, the weight matrix $\mathbf{W}$ cannot use even unequal diagonal weights corresponding to the sampled units, *i.e.*, w_i say, because the ij^{th} element of $\mathbf{\Omega}^{-1}$ (unlike the ij^{th} element of $\mathbf{\Omega}$) does *not* correspond to the i^{th} and j^{th} unit in the sample (or in the population), so it is rather unclear what $\mathbf{W}$ is or how $\mathbf{W}$ can be sensibly defined as 'a weighting matrix of expansion factors'.

This argument still applies when $\mathbf{V}_b$ is replaced by its estimator $\hat{\mathbf{V}}_b$ which uses estimates in place of σ_e^2 and σ_v^2.

References

Battese, G.E., Harter, R.M. and Fuller, W.A. (1988). An error components model for prediction of county crop area using survey and satellite data. *Journal of the American Statistical Association*, 83, 28-36.

Chambers, R. (2006). What is poverty? Who asks? Who answers? *Poverty in Focus*, UNDP, December 2006, 3-4.

Elbers, C., Lanjouw, J. and Lanjouw, P. (2003). Micro-level estimation of poverty and inequality. *Econometrica*, 71, 355-364.

Elbers, C., Lanjouw, J. and Lanjouw, P. (2002). *Micro-level Estimation of Welfare.* Research Working Paper 2911, World Bank, Development Research Group, Washington, D.C.

Ghosh, M., and Rao J.N.K. (1994). Small area estimation: An appraisal. *Statistical Science*, 9, 55-93.

Haslett, S., and Jones, G. (2004). *Local Estimation of Poverty and Malnutrition in Bangladesh*, Bangladesh Bureau of Statistics and United Nations World Food Programme.

Haslett, S., and Jones, G. (2005). *Local Estimation of Poverty in the Philippines*, Philippine National Statistics Co-ordination Board/World Bank Report. http://siteresources.worldbank.org/INTPGI/Resources/342674-1092157888460/Local_Estimation_of_Poverty_Philippines.pdf.

Haslett, S., and Jones, G. (2005). Small area estimation using surveys and censuses: Some practical and statistical issues. *Statistics in Transition*, 7, 541-556.

Haslett, S., and Jones, G. (2006). *Small Area Estimation of Poverty, Caloric Intake and Malnutrition in Nepal.* Published: Nepal Central Bureau of Statistics/World Food Programme, United Nations/World Bank, September 2006, 184pp, ISBN 999337018-5.

Henderson, C.R. (1953). Estimation of variance and covariance components. *Biometrics*, 9, 226-252.

Horton, N.J., and Lipsitz, S.R. (1999). Review of software to fit Generalized Estimating Equation regression models. *The American Statistician*, 53, 160-169.

Liang, K.L., and Zeger, S. (1986). Longitudinal data analysis using Generalized Linear Models. *Biometrika*, 73, 13-22.

Lohr, S.L. (1999). *Sampling*: *Design and Analysis.* Duxbury Press, Brooks/Cole Publishing Company.

Militino, A.F., Ugarte, M.D., Goicoa, T. and Gonzalez-Audicana, M. (2006). Using small area models to estimate the total area occupied by olive trees. *Journal of Agricultural*, *Biological*, *and Environmental Statistics*, 11, 450-461.

Pfeffermann, D., Skinner, C.J., Holmes, D.J., Goldstein, H. and Rasbash, J. (1998). Weighting for unequal selection probabilities in multilevel models. *Journal of the Royal Statistical Society* B, 60, 23-40.

Pfeffermann, D., Moura, F.A. and Silva, P.L. (2006). Multi-level modelling under informative sampling. *Biometrika*, 93, 949-959.

Prasad, N.G.N., and Rao, J.N.K. (1990). The estimation of the mean squared error of small-area estimators. *Journal of the American Statistical Association*, 85, 163-171.

Rao, J.N.K. (1999). Some recent advances in model-based small area estimation. *Survey Methodology*, 25, 175-186.

Rao, J.N.K. (2003). *Small Area Estimation*, Wiley Series in Survey Methodology. Wiley-Interscience, John Wiley & Sons, Inc.

NSCB (2000). *Profile of Censuses and Surveys*. National Statistical Coordination Board, Philippines.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences.* McGraw-Hill Series in Psychology. New York: McGraw-Hill.

Skinner, C.J., Holt, D. and Smith, T.M.F. (1989). *Analysis of Complex Surveys*. Chichester: John Wiley & Sons.

UN website. *http://www.un.org/milleniumgoals/*.

You, Y., and Rao, J.N.K. (2002). A pseudo-empirical best linear unbiased prediction approach to small area estimation using survey weights. *The Canadian Journal of Statistics*, 30, 431-439.

You, Y., Rao, J.N.K. and Kovačević, M. (2003). Estimating fixed effects and variance components in a random intercept model using survey data. *Proceedings*: *Symposium 2003*, *Challenges in Survey Taking for the Next Decade*. Statistics Canada.

Zhao, Q. (2006). User manual for PovMap, The World Bank. http://siteresources.worldbank.org/INTPGI/Resources/342674-1092157888460/Zhao_ManualPovMap.pdf.