The Annals of Applied Statistics
2016, Vol. 10, No. 1, 451–476
DOI: 10.1214/15-AOAS900

EXTENDED STRUCTURE PRESERVING ESTIMATION (ESPREE) FOR UPDATING SMALL AREA ESTIMATES OF POVERTY

BY MARISSA ISIDRO[*,1], STEPHEN HASLETT[*,†] AND GEOFF JONES[*]

Massey University[*] and Australian National University[†]

Small area estimation techniques are now routinely used to generate local-level poverty estimates for aid allocation and poverty monitoring in developing countries. However, the widely implemented World Bank (WB) or Elbers, Lanjouw and Lanjouw [*Econometrica* **71** (2003) 355–364] (ELL) method can only be used when a survey and census are conducted at approximately the same time. The empirical best prediction (EBP) method of Molina and Rao [*Canad. J. Statist.* **38** (2010) 369–385] also requires a new census for updating. Hence, if small area estimation methods that use both survey and census unit record data are required, and the survey is rerun some years after the census, how to update small area estimates becomes an important issue. In this paper, we propose an intercensal updating method for local-level poverty estimates with estimated standard errors which we call Extended Structure PREserving Estimation (ESPREE). This method is a new extension of classical Structure PREserving Estimation (SPREE). We test our approach by applying it to inter-censal municipal-level poverty estimation and carrying out a validation exercise in the Philippines, comparing the estimates generated with an alternative ELL or EBP updating method due to Lanjouw and van der Wiede [Determining changes in welfare distributions at the micro-level: Updating poverty maps. (2006) Powerpoint presentation at the NSCB Workshop for the NSCB/World Bank Intercensal Updating Project] which uses time-invariant variables. The results show that the ESPREE estimates are preferable, generally being unbiased and concurring well with local experts' opinion on poverty levels at the time of the updated survey.

1. Introduction. Effective targeting schemes for aid allocation in Third World countries and monitoring progress toward the Millennium or Sustainable Development Goals (UN website) require reliable information on the poor at the local or community level. This pressing demand for local-level estimates of poverty measures led to the development of various small area estimation techniques for poverty measures in Third World countries. In general, small area estimation uses auxiliary data to improve the precision of subdomain estimates from national survey data [Rao (2003)]. Currently, the most widely used small area estimation technique for poverty measures, proposed by Elbers, Lanjouw and Lanjouw (2003), is the ELL (Elbers, Lanjouw and Lanjouw) method which has been

Received May 2015; revised November 2015.

[1]Supported by Education New Zealand, through the New Zealand Postgraduate Study Abroad Awards (NZPSAA).

Key words and phrases. Small area models, SPREE, poverty mapping, intercensal updating.

extensively used by the World Bank in its poverty mapping projects in collaboration with national statistical agencies in Third World countries. An alternative small area estimation method, Empirical Best Prediction (EBP), which like ELL uses survey and census unit record data, has been developed by Molina and Rao (2010) and Molina, Nandram and Rao (2014). EBP, unlike ELL, includes predictions of area-specific random effects. This extension is useful if small areas are comparatively large so that they contain a sufficient sample, and there are no contextual variables at area level [see Haslett (2016), Namazi-Rad and Steel (2015)], although this is not usually the case for Third World aid allocation.

The focus here is how small area estimates from either method (or any other that uses sample and census unit record data) can be updated when there is a new survey but not a new census.

The poverty measures that are generally included in poverty maps are economic or monetary based, such as poverty incidence, gap and severity. These measures are all generated based on the premise that a household is poor if its income or expenditure falls below a specified monetary standard, known as the poverty line. The three measures mentioned above can be put into a framework proposed by Foster, Greer and Thorbeck (1984), the so-called FGT measures:

$$P_i = \frac{1}{N} \sum_{h=1}^{N} \left(\frac{\ell - Y_h}{\ell} \right)^i \cdot I(Y_h < \ell), \tag{1.1}$$

where N is the size of the population in an area, Y_h is the income or consumption of individual or household h, ℓ is the poverty line (which could vary depending on the area within a particular country) and $I(Y_h < \ell)$ is an indicator function (equal to 1 when income or expenditure is below the poverty line and 0 otherwise). Poverty incidence, gap and severity correspond to $i = 0$, 1 and 2, respectively.

Extensions to under-nutrition measures such as kilocalorie consumption, and stunting, underweight and wasting in children under five years of age are also possible. See, for example, Haslett, Jones and Sefton (2013).

Both EBP and ELL methods combine the sample survey and the census data to produce small area estimates of poverty measures. Both methods assume the two data sets are gathered at the same time period. This assumption is particularly important since the variable of interest (income/consumption or poverty status) is not measured in the census; the model is formulated using the survey data and is then applied in the form of a predictor at unit-record level to the census data, after which predictions are aggregated to small area level. Variations and improvements have been proposed by Haslett, Isidro and Jones (2010), Haslett and Jones (2010), Tarozzi and Deaton (2009) and Haslett (2013, 2016). In most countries, especially in the Third World, a census is only conducted once every ten years. This poses the problem, addressed in this paper, of updating EBP and ELL-type small area estimates in intercensal years, that is, in noncensus years between censuses. The proposed method provides policymakers and other stakeholders with an updated estimate of poverty measures for Third World countries when new survey data becomes available, even if there is no new census. Our intercensal updating method is an extension of the structure preserving estimation (SPREE) method of Purcell and Kish (1980), hence called Extended SPREE (ESPREE).

An alternative updating method for small area poverty estimates, given new survey data, is used for the Small Area and Poverty Estimates (SAIPE) program in the United States. SAIPE produces small area estimates of poverty at the county and at the finer school district level. It uses published and unpublished single year estimates from the American Community Survey (ACS), Internal Revenue Service (IRS) individual-level Federal Income Tax Return information, the US Census-based Current Population estimates, data from the Supplemental Nutrition Assistance Program (SNAP), aggregate personal income estimates from the Bureau of Economic Analysis (BEA) and the number of recipients of Supplemental Security Income (SSI) benefits. For example, among other information, the IRS data gives family size via tax exemptions, and the decennial Census-based Current Population estimates provide information on the number of school-age children. For further details, see http://www.census.gov/did/www/saipe/data/model/info/index.html. The small area estimates are Empirical Bayes, based on weighting via relative estimated precision of the ACS estimates and regression type estimates from the other data sources. The SAIPE method is simple and has been successfully applied to estimate poverty incidences where there may not be survey data available for all the small areas. Considerable research has been involved in developing SAIPE, which utilizes the regular repeated ACS surveys and sound census-type information from administrative and other sources, for example, from the IRS tax data and the census population projections. SAIPE's statistical properties have also been extensively explored [e.g., Bell et al. (2007), Hawala and Lahiri (2012)]. However, sound census population estimate updates, which are required in SAIPE, are seldom available in third world countries. Similarly, in the third world, for example, in countries such as Cambodia, Bangladesh, Nepal, Pakistan, Timor-Leste and Bhutan, limited and variable annual resources stymie regular surveys, so that surveys are not generally repeated as part of a fixed work program. Without sound population projections, good additional data sources and regular surveys, despite their providing a "gold standard," SAIPE methods cannot easily be used.

ESPREE is more suitable for third world application because it has more limited new data requirements for small area estimate updating. ESPREE utilizes only new survey data from a single repeated survey and various fine-level aggregations of the unit-level predictions generated via the small area model developed from the original, contemporaneous survey and census. Fine-level, postcensal population predictions are not explicitly required as in SAIPE for the denominator in revised poverty rates because the ESPREE is based principally on updating percentages.

SPREE itself is a generalization of the synthetic estimation proposed by Gonzalez (1973). It makes better use of direct estimates and uses the method of iterative proportional fitting (IPF) of margins. IPF is used to adjust the census cell counts of a multi-way table called the association structure (census data), such that the adjusted counts satisfy specified margins, called the allocation structure (survey margins). The cell counts are obtained from the last census, while the specified margins represent reliable direct survey estimates of current margins. In this way, SPREE provides intercensal estimates of small area totals of characteristics also measured in the census [Rao (2003)].

However, the variable of interest under small area estimation for poverty measures in Third World countries is not usually measured in the census. The ESPREE method, as presented in Section 2, extends the SPREE method so that updated estimates can be generated by using "pseudo-census data," as produced by an earlier application of the EBP or ELL method, instead of requiring new census data. Since the pseudo-census data has many replicates, ESPREE also relaxes the assumption of fixed census data under SPREE, by using a superpopulation approach that allows standard error estimation to incorporate both census and sample variation. Data from the Philippines is used to illustrate the ESPREE method, presented in Section 3. An alternative updating method developed by Lanjouw and van der Wiede (2006) (LW), and discussed in Christiaensen et al. (2010), was carried out in the Philippines using the same data sources. A validation study conducted in one of the regions in the Philippines, which allows assessment of the quality of the ESPREE and LW estimates, in comparison with the real poverty situation on the ground, is presented in Section 4.

Our results for the Philippines suggest that the ESPREE method generates updated small area estimates that are unbiased (unlike LW) with standard errors that are smaller than the survey-based estimates for the new period, both at the desired small area (municipal) level and at the more aggregated provincial and regional levels. Moreover, the estimates from the ESPREE method were also found to be closer to the key informants assessment than those from the LW updating method for the municipalities of the region visited in a validation study.

2. Intercensal updating methods.

2.1. *The SPREE method.* The SPREE method can be implemented by fitting a generalized linear model (GLM) with a log link function as follows:

$$\log(\boldsymbol{\mu}) = g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta}, \tag{2.1}$$

where $g()$ is the log function and $\boldsymbol{\mu}$ is the expected value of the vector of the dependent variable. For poverty estimation, $\boldsymbol{\mu}$ could be the number of households or people (which we denote here by the superscript $\mathbf{Y}$ for the new survey period t_1 and $\tilde{\mathbf{Z}}$ for the previous census period t_0) cross-classified by poverty status, province and other related variables. $\mathbf{X}$ is the design or model matrix corresponding to the explanatory variables and is partitioned into $[\mathbf{X}_1 : \mathbf{X}_2]$, so that the first component $\mathbf{X}_1$ contains the elements of the design matrix associated with what are usually the main effects in the parameter vector $\boldsymbol{\beta}$ and the second component $\mathbf{X}_2$ is associated with the higher order effects.

For SPREE applied to poverty estimation, the general model for the census data is

$$g(\boldsymbol{\mu}^{\tilde{Z}}) = \mathbf{X}_{1,t_0}\boldsymbol{\beta}_{1,t_0} + \mathbf{X}_{2,t_0}\boldsymbol{\beta}_{2,t_0}, \tag{2.2}$$

with the subscript t_0 indicating that the data comes from an earlier period, while the survey model

$$g(\boldsymbol{\mu}^{Y}) = \mathbf{X}_{1,t_1}\boldsymbol{\beta}_{1,t_1} + \mathbf{X}_{2,t_1}\boldsymbol{\beta}_{2,t_1}, \tag{2.3}$$

with the subscript t_1 indicating that the data comes from a more recent period. We note specially for SPREE that $\mathbf{X}_{1,t_0} = \mathbf{X}_{1,t_1} = \mathbf{X}_1$ and $\mathbf{X}_{2,t_0} = \mathbf{X}_{2,t_1} = \mathbf{X}_2$ are the partition of the design matrix corresponding to the parameter vector $\boldsymbol{\beta}$ in the GLM model, which is $\boldsymbol{\beta}_{1,t_0}$ and $\boldsymbol{\beta}_{2,t_0}$ for the census model and $\boldsymbol{\beta}_{1,t_1}$ and $\boldsymbol{\beta}_{2,t_1}$ for the survey model. As noted above, the first term in (2.3) generally represents the main effects and/or lower order parameters which can be accurately estimated from the survey data, while the second represents the higher order effects which cannot be accurately re-estimated from the survey.

Considering model (2.2) and model (2.3), SPREE is equivalent to fitting model (2.2), after which some of the lower order parameters in the model (first component in partition) are adjusted or updated in line with the most recent information available from the survey data, while the higher order parameters (second component), for which new information from the survey is not available, remains the same, that is, $\boldsymbol{\beta}_{2,t_0} = \boldsymbol{\beta}_{2,t_1} = \boldsymbol{\beta}_2$ by assumption. In this way SPREE is used to generate updated small area estimates. This process is also equivalent to fitting model (2.3) and then equating the higher order parameters (second partition) to the values generated from the census model (2.2), that is,

$$g(\boldsymbol{\mu}^{Y}) = \mathbf{X}_{1,t_1}\boldsymbol{\beta}_{1,t_1} + \mathbf{X}_2\boldsymbol{\beta}_{2,t_0}. \tag{2.4}$$

2.2. *The ESPREE method.* The SPREE model described above is extended by allowing for a misspecification error or an error in the association structure, γ_{t_1}, that is, $\mathbf{X}_2\boldsymbol{\beta}_2 = \mathbf{X}_2\boldsymbol{\beta}_{2,t_1} = \mathbf{X}_2\boldsymbol{\beta}_{2,t_0} + \gamma_{t_1}$. Using the pseudo-census data, the estimation problem is considered in the context of a superpopulation, that is, we assume that the pseudo-census data which, as indicated by (2.2), applies at the time of the census t_0 are realizations of a superpopulation that produces $\boldsymbol{\mu}^{\tilde{Z}}$, the dependent variable measured at time t_0. Note that $\mathbf{X}$ but not $\boldsymbol{\mu}^{\tilde{Z}}$ is known from the census at t_0, which is why there are sets of pseudo-census values generated for $\boldsymbol{\mu}^{\tilde{Z}}$ rather than census data. In the ELL method there are commonly around 100 pseudo-censuses, that is, 100 predictions of income or expenditure for every census household based on the model fitted to the survey data and for the EBP there is also a set of pseudo-censuses. Under the assumption that the pseudo-census data forms a superpopulation, the coefficients $\boldsymbol{\beta}_{1,t_0}$ and $\boldsymbol{\beta}_{2,t_0}$ of the census model are now considered random with respect to the superpopulation and with expectation $\boldsymbol{\xi}[\boldsymbol{\beta}_{1,t_0}]$ and $\boldsymbol{\xi}[\boldsymbol{\beta}_{2,t_0}]$, respectively. The census model can now be written as

$$g(\boldsymbol{\mu}^{\tilde{Z}}) = \mathbf{X}_1\boldsymbol{\xi}[\boldsymbol{\beta}_{1,t_0}] + \mathbf{X}_2\boldsymbol{\xi}[\boldsymbol{\beta}_{2,t_0}] + u_{t_0}, \tag{2.5}$$

where u_{t_0} is a random variable with respect to the superpopulation assumed for the census data. Note that here $\boldsymbol{\mu}^{\tilde{Z}}$ is a vector of counts, not a continuous variable such as income or expenditure (or the log or some other continuous transformation of them).

For the survey data, we also assume that a superpopulation exists for the population from which the survey data is drawn. As in the census model, the parameters $\boldsymbol{\beta}_{1,t_1}$ and $\boldsymbol{\beta}_{2,t_1}$ of the survey model are now random with respect to the superpopulation and with expectation $\boldsymbol{\xi}E[\boldsymbol{\beta}_{1,t_1}]$ and $\boldsymbol{\xi}E[\boldsymbol{\beta}_{2,t_1}]$, respectively. E is the expectation related to the sampling design of the survey data, while $\boldsymbol{\xi}$ is the expectation related to the superpopulation assumed for the census. However, we could assume that $\boldsymbol{\xi}E[\boldsymbol{\beta}_{1,t_1}] = \boldsymbol{\xi}[\boldsymbol{\beta}_{1,t_1}]$ and $\boldsymbol{\xi}E[\boldsymbol{\beta}_{2,t_1}] = \boldsymbol{\xi}[\boldsymbol{\beta}_{2,t_1}]$ provided that the survey estimation method is unbiased and the sampling design is noninformative. Hence, the survey model will now be

$$g(\boldsymbol{\mu}^{Y}) = \mathbf{X}_1\boldsymbol{\xi}[\boldsymbol{\beta}_{1,t_1}] + \mathbf{X}_2\boldsymbol{\xi}[\boldsymbol{\beta}_{2,t_0}] + u_{t_1} + \gamma_{t_1}, \tag{2.6}$$

where u_{t_1} is a random error term for the survey model and γ_{t_1} is a misspecification error. Hence, our proposed updating model which we call *extended SPREE* (ESPREE) is as follows:

$$g(\boldsymbol{\mu}^{Y}) = \mathbf{X}_1\boldsymbol{\beta}_{1,t_1} + \mathbf{X}_2\boldsymbol{\beta}_{2,t_0} + \varepsilon_{t_1}, \tag{2.7}$$

where $\varepsilon_{t_1} = u_{t_1} + \gamma_{t_1}$.

Fitting GLMs for ESPREE to generate small area estimates can be cumbersome and tedious, as it entails fitting two models, one for the survey data and one for the census data, and this is especially so when dealing with many explanatory variables and the large data sets (e.g., national survey and census) common when using ELL-type methods. Hence, we adapted one of the model-fitting procedures for classical SPREE, which basically employs the method called *table standardization* described by Agresti (2002). Table standardization is equivalent to the IPF algorithm.

To illustrate, for a two-way table with counts Y_{ab} such that $\mu_{ab}^{Y} = E(Y_{ab})$, the standardization process corresponds to fitting to the survey data the following model:

$$\log \mu_{ab}^{Y} = \mathbf{X}_1 \Delta\boldsymbol{\beta}_1 + \log \mu_{ab}^{\tilde{Z}} + u_{t_1}. \tag{2.8}$$

This model is fitted with $\log \mu_{ab}^{\tilde{Z}}$ as an offset (equal to the log of the pseudo-census data). We note that $g(\boldsymbol{\mu}^{Y}) = \log \mu_{ab}^{Y}$, u_{t_1} is a random error from the *pseudo-count* replicates for each table cell generated from the survey margins, $\Delta\boldsymbol{\beta}_1 =$

$\boldsymbol{\beta}_{1,t_1} - \boldsymbol{\beta}_{1,t_0}$ and the other parameters are as defined earlier. Fitting model (2.8) is equivalent to scaling the mean of the replicates of the pseudo-census data to agree with the survey estimates of the margin counts in the period t_1.

Note that fitting model (2.8) requires replicates of the pseudo-census and pseudo-counts data. The pseudo-census data here are the cell-level counts generated via the ELL method for the entire census. As metioned earlier, the usual practice is to generate around 100 replicates of these to incorporate the uncertainty in the predictions. The pseudo-counts are cell-level data from the survey which we denote here as $(\hat{Y}_{ab})$. Pseudo-counts are computed from the available survey margins and are generated based on the structure of the design matrix $\mathbf{X}_1$. For example, if we have two sets of reliable margins $y_{a\cdot}$ and $y_{\cdot b}$ for two variables such that one has A categories and the other has B categories, the pseudo-counts will be $\hat{Y}_{ab} = (y_{a\cdot} y_{\cdot b})/y_{\cdot\cdot}$, where, $a = 1, \dots, A$ denotes areas, $b = 1, \dots, B$ denotes categories of the variable of interest and $(\cdot)$ represents a sum over that index. The computed pseudo-counts do not necessarily satisfy the assumption of independence; these values depend on the combination of survey margins available for estimation. Examples of computing pseudo-counts corresponding to specific loglinear models are presented in Table 8.13 of Agresti (2002).

2.3. *Variance estimation.* As indicated in (2.6), there are two sources of variation for the updated small area estimates—the survey data and the pseudo-census data. Hence, the variance of the updated small area estimates under the ESPREE method needs to take into account uncertainty in both the survey margins and in the pseudo-census data, and should be viewed as the sum of the two variances (survey margins variance and pseudo-census variance). Any of the variance estimation methods (i.e., linearization and replication methods) can be used to compute the variances of both the survey and pseudo-census data.

Here we derive the overall variance of the ESPREE estimates using the linearization method. The following notation is used for convenience and simplicity of exposition in presenting the variance estimation procedure: $\hat{\mathbf{p}}$ is the column vector of the required cell estimates, that is, $\hat{\mathbf{p}} = (\hat{p}_{111}, \dots, \hat{p}_{ABC})'$ with dimensions $ABC \times 1$ where ABC is the total number of cells; $\mathbf{p}^*$ is the vector containing the survey margins (elements are $\hat{p}_{\cdot bc}$); and $\boldsymbol{\pi}$ is the vector with dimensions similar to $\hat{\mathbf{p}}$ containing the pseudo-census data (relative cell frequencies) π_{abc} established in the most recent census year.

We let $\boldsymbol{\Pi}$ be the expected value of $\boldsymbol{\pi}$ [i.e., $E(\boldsymbol{\pi}) = \boldsymbol{\Pi}$] and $\mathbf{P}^*$ be the expected value of $\mathbf{p}^*$ [i.e., $E(\mathbf{p}^*) = \mathbf{P}^*$]. Using the first order Taylor series, we will have

$$\begin{aligned}\hat{\mathbf{p}} &= F(\mathbf{p}^*; \boldsymbol{\pi}) \\ &\approx F(\mathbf{P}^*; \boldsymbol{\Pi}) + \frac{\partial F(\mathbf{p}^*; \boldsymbol{\pi})}{\partial \mathbf{p}^*}(\mathbf{p}^* - \mathbf{P}^*) + \frac{\partial F(\mathbf{p}^*; \boldsymbol{\pi})}{\partial \boldsymbol{\pi}}(\boldsymbol{\pi} - \boldsymbol{\Pi}),\end{aligned} \tag{2.9}$$

where F is defined by the generalized linear model. Assuming independence of $\mathbf{p}^*$ and $\boldsymbol{\pi}$, the variance–covariance matrix of $\hat{\mathbf{p}}$ is then approximated by the variance–covariance matrix of the linear function (2.9), that is,

$$
\begin{aligned}
V(\hat{\mathbf{p}}) &\approx E\left\{\left[\frac{\partial F(\mathbf{p}^*;\boldsymbol{\pi})}{\partial \mathbf{p}^*}\right](\mathbf{p}^*-\mathbf{P}^*)(\mathbf{p}^*-\mathbf{P}^*)'\left[\frac{\partial F(\mathbf{p}^*;\boldsymbol{\pi})}{\partial \mathbf{p}^*}\right]'\right\} \\
&\quad + E\left\{\left[\frac{\partial F(\mathbf{p}^*;\boldsymbol{\pi})}{\partial \boldsymbol{\pi}}\right](\boldsymbol{\pi}-\boldsymbol{\Pi})(\boldsymbol{\pi}-\boldsymbol{\Pi})'\left[\frac{\partial F(\mathbf{p}^*;\boldsymbol{\pi})}{\partial \boldsymbol{\pi}}\right]'\right\} \\
&= \left[\frac{\partial F(\mathbf{p}^*;\boldsymbol{\pi})}{\partial \mathbf{p}^*}\right]V(\mathbf{p}^*)\left[\frac{\partial F(\mathbf{p}^*;\boldsymbol{\pi})}{\partial \mathbf{p}^*}\right]' \\
&\quad + \left[\frac{\partial F(\mathbf{p}^*;\boldsymbol{\pi})}{\partial \boldsymbol{\pi}}\right]V(\boldsymbol{\pi})\left[\frac{\partial F(\mathbf{p}^*;\boldsymbol{\pi})}{\partial \boldsymbol{\pi}}\right]',
\end{aligned}
\tag{2.10}
$$

where $V(\mathbf{p}^*)$ and $V(\boldsymbol{\pi})$ are the covariance matrix for $\mathbf{p}^*$ and $\boldsymbol{\pi}$, respectively. That is, $V(\mathbf{p}^*) = E[(\mathbf{p}^*-\mathbf{P}^*)(\mathbf{p}^*-\mathbf{P}^*)']$ and $V(\boldsymbol{\pi}) = E[(\boldsymbol{\pi}-\boldsymbol{\Pi})(\boldsymbol{\pi}-\boldsymbol{\Pi})']$. We note that $V(\mathbf{p}^*)$ under the superpopulation model can be obtained either directly or by using any of the replication methods. From equation (2.10), under ESPREE the estimated variance is the sum of the variability from the pseudo-census and the survey margins.

A related result for the variance derived from the linearization method could be attained by the approach proposed by Haslett, Green and Zingel (1998). The approach formulated was based on the mean square error formula:

$$
[\widehat{\mathrm{MSE}}(\hat{\mathbf{p}}|\boldsymbol{\pi})] = V(\hat{\mathbf{p}}|\boldsymbol{\pi}) + (\hat{E}(\hat{\mathbf{p}}|\boldsymbol{\pi}) - \hat{\mathbf{P}})(\hat{E}(\hat{\mathbf{p}}|\boldsymbol{\pi}) - \hat{\mathbf{P}})', \tag{2.11}
$$

where $\hat{\mathbf{P}}$ denotes the set of cell estimates based on their long-term averages. The estimated mean square error gives an estimate of the joint design/model (superpopulation) variance by treating both terms in the equation as estimates of conditional variances; see Appendix for detailed derivation and proof. The first term on the right-hand side of the equation can be considered to be the variability from the survey margins and the second term from the pseudo-census as in the variance formula in (2.10).

The variance estimator in (2.10) and (2.11) is based on a large sample. However, a stable estimator of the squared bias in (2.11) can be difficult to obtain [Rao (2003)]. Gonzalez and Waksberg (1973) and Marker (1999) proposed methods, extended to generalized linear models by Noble, Haslett and Arnold (2002, 2006), but these methods have not proved generally satisfactory for bias estimation either. Here, however, there are two important points of difference. First, there are many pseudo-censuses in ELL (the standard is 100 pseudo-census values for each original census unit record and a similar number can be used for EBP), so that estimation of bias is not based on a small number of replicates. Second, and perhaps more importantly, EBP is certainly not a synthetic method, as it includes nonzero small area specific random effects. Further, ELL and ESPREE are not synthetic estimates in the usual sense because, via survey cluster level census averages, they can both make use of contextual effects to supplement the candidate variables for the regression fitted to the original survey data. These contextual variable averages, which are better used in the survey modeling at cluster than the more aggregated small area level, are required in model fitting only for unit records from the small areas that are in the sample. However, they are available from the census for every small area, whether those small areas are sampled or not, and so adjust the small area estimate in every small area using an adjustment specific to the small area, even when that small area is not sampled. Such contextual effects then provide an explicit alternative to the nonzero random effect predictions at small area level used for EBP in Molina and Rao (2010), whose method consequently requires that there be a reasonably substantial sample in every small area. This substantial sample requirement severely reduces the number of small areas for which it is possible to produce reliable small area estimates in Molina and Rao (2010), but not for ELL or necessary for ESPREE. Further details on the central role of such contextual effects in a range of types of small area estimation, and the comparison with Molina and Rao (2010), are given in Namazi-Rad and Steel (2015) and Haslett (2016). For ESPREE updating, whether EBP or ELL has been used at the initial time period, Appendix provides a formal proof of (2.11). The estimation, applied via (2.10), allows for the variation in the new estimate of $\mathbf{X}_1\boldsymbol{\beta}_1$ via the replicates of the new survey margins using balanced repeated replication, and for the variation of the pseudo-censuses (in the estimation of $\mathbf{X}_2\boldsymbol{\beta}_2$) through varying the offset using the bootstrap. See (2.2) and (2.3) for explicit model details. More commonly in the literature on GLMs, an offset is specified as fixed, so what (2.10) is providing via linearization is an estimate of mean square error that is not downwardly biased as would be using a fixed offset. That the mean square error estimation of (2.10) and (2.11) can in practice produce a sufficiently accurate assessment of mean square error for ESPREE is detailed and discussed in the simulation of Section 4.

For the Philippine data presented in the next section, we used the balanced repeated replicates (BRR) method as described fully in Isidro (2010) to compute the variance component from the survey data. Since the survey data has a large number of strata, partial balancing [Wolter (1985)] was used. The variance component from pseudo-census data is computed using the set of bootstrap estimates from the small area estimation project based on 2000 survey and census data, as described in Haslett and Jones (2005), with the 2003 survey data regarded as fixed. The two variance estimates are then added to get an estimate of the variances of the updated small area estimates via ESPREE that allows both for the variation in the pseudo-census and new survey data.

2.4. *The other ELL updating methods.* There are two other types of updating technique for ELL-based poverty measures that have been implemented. One uses panel survey data and the other one (LW) uses “time-invariant” auxiliary variables.

The panel data approach has been used in Uganda [Hoogeveen, Emwanu and Okwi (2006)] and in Thailand [Jitsuchon and Lanjouw (2005)], while the LW approach has been implemented in Vietnam and the Philippines [Lanjouw and van der Wiede (2006)]. The panel data approach requires the availability of a longitudinal data set, while the time-invariant approach requires two cross-sectional surveys.

The implementation of the time-invariant updating method in the Philippines as described by Lanjouw and van der Wiede (2006) takes the per capita income/expenditure data at the household level from the most recent survey (y_{ch,t_1}) and combines it with what are assessed (but not fully tested) to be time-invariant variables, common to the survey and census data, collected in the census year (t_0). The survey model based on the subsample of households that are in both surveys (at times t_0 and t_1) is as follows:

$$y_{ch,t_1} = \mathbf{x}_{ch,t_0}\boldsymbol{\beta} + u_{ch,t_1}, \tag{2.12}$$

where $\mathbf{x}'_{ch,t_0}$ refers to characteristics that are time invariant (i.e., $\mathbf{x}_{ch,t_0} = \mathbf{x}_{ch,t_1}$; in practice, $\mathbf{x}_{ch,t_1}$ is used for fitting the survey regression); $\boldsymbol{\beta}$ is the regression parameter and u_{ch,t_1} is the random error term where c denotes clusters and h households within clusters. The survey-based model is then applied to the old census (at time t_0) since the survey model contains only time invariant variables. This methodology depends on two assumptions: (1) independent or explanatory variables in the survey model are time invariant, that is, household characteristics and municipality/village means do not change from the census period to the most recent survey; and (2) migration (at least among small areas) between the census period and the most recent survey is negligible.

Migration also has the potential to affect ESPREE. The core issue is that, between the original survey and its repeat, internal and/or external migration has not produced structural change in the relevance of variables in $\mathbf{X}_{2,t_0}$ (or, equivalently, $\mathbf{X}_{2,t_1}$) or changes in the estimate of $\boldsymbol{\beta}_2$ in the models of (2.2) and (2.3). Assessment is difficult in practice, as few countries have reliable internal migration statistics. However, it is not migration per se that would affect the ESPREE model, but that migration has induced a change in the relationship between the variable of interest and either the variables in $\mathbf{X}_{2,t_0}$ or the variables that have been excluded from it but have now become important at time t_1. Major disasters or wars may induce such change, for example. Periods of relative structural stability between t_0 and t_1 and shorter time differences between the first survey and the census at t_0 and the second survey at t_1 are advisable. The core assumption is that those who migrate are like in kind to those who stay.

In summary, the LW updating method involves fitting a new survey-based model at t_1 using only time-invariant variables and then, much like the ELL method, applying this model to a census at t_0, and then aggregating. The Hoogeveen, Emwanu and Okwi (2006) updating method is very similar except that, because the second survey is a panel survey, all the units in the second survey were present in the original survey.

3. Application to the Philippines data.

3.1. *The data.* The ESPREE method is applied here to the Philippines national survey and census data. There are two sets of survey data used—Family Income and Expenditure Survey (FIES) and Labor Force Survey (LFS) which are both conducted by the National Statistics Office (NSO). The FIES collects information on family income and expenditures as well as information affecting them. The LFS, on the other hand, collects data on employment and related information on demographic and socio-economic characteristics of the population over 15 years old. In this study we used the 2000 and 2003 FIES/LFS data. The 2000 survey and census data is used for creating the pseudo-census data via a variation of the ELL method and the 2003 data is used for updating (i.e., as the most recent survey margins). Note that there were no major national disasters leading to mass internal migration in the Philippines during the 2000–2003 period.

The 2000 Census of Population and Housing (CPH) used in this study is also conducted by the NSO. The CPH in the Philippines is done once every ten years with a more limited Census of Population every 5 years. A common questionnaire (short form) for the CPH is given to all households, with an extended questionnaire (long form) completed by a random sample of about 10 percent of the population. The sampling design employed for this 10 percent sample is a systematic cluster design, with the sampled fraction being 100, 20 or 10 percent depending on the size of the municipality. The administrative areas covered by the CPH include 1623 municipalities in 83 provinces which are grouped into 16 regions.

The characteristics of the superpopulation are inferred from a set of 100 bootstrap estimates of poverty status classification of the population generated by the poverty mapping project conducted by the World Bank (WB) in collaboration with the National Statistical Coordination Board (NSCB) in the Philippines employing the ELL method; see Haslett and Jones (2005) and Haslett, Isidro and Jones (2010) for details. We call this set of household-level predictions or replicates from the period in which the earlier survey coincided with the census the pseudo-census.

3.2. *Model formulation.* The variables used to illustrate the ESPREE method are a subset of those used in the joint WB/NSCB analysis mentioned above. See Haslett and Jones (2005) for a complete list and definition of variables. These variables are available in both the 2000 census and the 2000 and 2003 survey data sets. Based on analysis of the survey data, they are strongly correlated with household per capita income, and hence with the poverty status of the household members. Due to limited computer memory capacity for running the program for generating the estimates, we restricted the number of variables used to a maximum of six. With six auxiliary variables, the number of cells in the contingency table is already about 850,000, given that there are 1623 municipalities included in the analysis. An example of a set of 6 variables considered is as follows: urbanity (urban or rural); high school educational attainment (at least one family member 10 years and

TABLE 1
Distribution of marital status by poverty status for Philippines FIES/LFS data

	2000		2003	
Poverty status	**With spouse**	**No spouse**	**With spouse**	**No spouse**
Nonpoor	0.8292	0.1708	0.8390	0.1610
Poor	0.9028	0.0972	0.9158	0.0842
Pearson chi-square	386.7387		433.548	

over with high school education or no high school education); type of house wall materials (strong, light, salvaged or other materials); household head gender (male or female); type of house roof material (strong, light, salvaged or other materials); and presence of household help (present or not). Fitting a regression model to household data for the log per capita income with the above as explanatory variables yielded a multiple correlation coefficient of about 0.67 (or an $R^2 \simeq 0.5$), which is typical of many ELL applications when the model is based on log income or log expenditure. Under the ESPREE method, a loglinear model is fitted directly to counts of poverty status (poor and nonpoor), hence the relationship between the set of variables and poverty status needs to be checked in the loglinear model. Two-way tables (cross-tabulation of poverty status with each of the auxiliary variables) including the chi-square values were constructed and examined. An example is given below (Table 1). We observe that all the chi-square statistics are significant since all the p-values associated with the computed chi-square values were greater than $\chi^2_{\mathrm{df}=1,\alpha=0.001} = 10.83$. Note that each cell in the table contains the within-row relative frequencies for each year. For example, in the year 2000, among those who are nonpoor, 82.92% of the individuals are married (with spouse).

Various models were fitted, some of which are presented in Table 2. Choosing between models is a difficult issue. The loglinear model used in ESPREE is a hybrid with components estimated from both survey and pseudo-census data; a naive Akaike Information Criterion (AIC) can be calculated as

$$\mathrm{AIC} = \frac{-2L + 2\dot{p}}{\tilde{N}},$$

where L is the overall loglikelihood, $\dot{p}$ is the number of covariates in the model (including intercept) and $\tilde{N}$ is the number of cells in the contingency table, but it is hard to see how this can be adapted to account properly for the survey design (note that the offset changes for each model). Alternatively, the model can be chosen by fitting to the survey data alone, in which case the design-adapted AIC of Lumley and Scott (2015) can be used. This is implemented in the function `svyglm()` in the R package `survey` [Lumley (2014)] and can be fitted as a logit model for poverty using the equivalence of the logit and loglinear models. Both approaches are given in Table 2.

TABLE 2

Some of the models fitted with the corresponding naive AIC (AIC_n) for the loglinear model fitted using the pseudo-census counts, and the design-adapted AIC (AIC_d) of Lumley and Scott (2015) for the logit model fitted to the survey data alone

No. of variables	Variable(s)	AIC_n	AIC_d
1	Wall type	9,915.7	45,011.9
2	Wall type and urbanity	16,301.8	43,091.2
4	Wall type, urbanity, head_male, and all_hsed	4,090.7	42,611.5
6	Wall type, roof type, urbanity, head_male, all_hsed and dom_help	748.4	41,750.6

Notice that using six variables for the loglinear model gives the best fit by both criteria. Note that for the loglinear model used in ESPREE, the margins are changed, but the overall model also includes interactions of all orders via the census (or pseudo-census) data in the association structure.

3.3. *Intercensal small area estimates of poverty incidence.* In this section the ESPREE intercensal estimates of poverty incidence, with their estimated standard error (SE) and coefficient of variation (CV), are presented along with those from the LW method obtained from the NSCB report [NSCB (2009)]. These results are compared with the direct survey-based (FIES) estimates at the provincial and regional levels. Because FIES is not a panel survey, the panel data updating method of Hoogeveen, Emwanu and Okwi (2006) and Jitsuchon and Lanjouw (2005) was not used.

The ESPREE-based municipal-level estimates of poverty incidence are mapped in Figure 1. A summary of the statistical properties of the municipal-level estimates is presented in Table 3. The mean of the poverty incidence computed from the ESPREE method is higher than the one generated from the LW method. This is further supported by the scatter plot in Figure 2 which shows that this LW method tends to generate lower values of poverty incidence estimates than the ESPREE method in most of the municipalities or small areas. We can also observe that the average estimated standard errors of poverty incidence estimates from the ESPREE and the LW methods are similar. However, the average CVs computed from the two methods indicate that the ESPREE method generates more precise estimates than the LW method. Note that standard errors for a proportion averaging below 0.05 are generally considered to be acceptable in poverty mapping.

Presented in Table 4 is a statistical comparison of the direct estimates generated from the survey data at a higher level of aggregation using the survey-based estimation procedure in STATA [StataCorp (2007)], ESPREE and LW methods. We note that the direct survey-based estimates generated differ slightly from the official estimates released by the NSCB since our survey-based estimates are based on

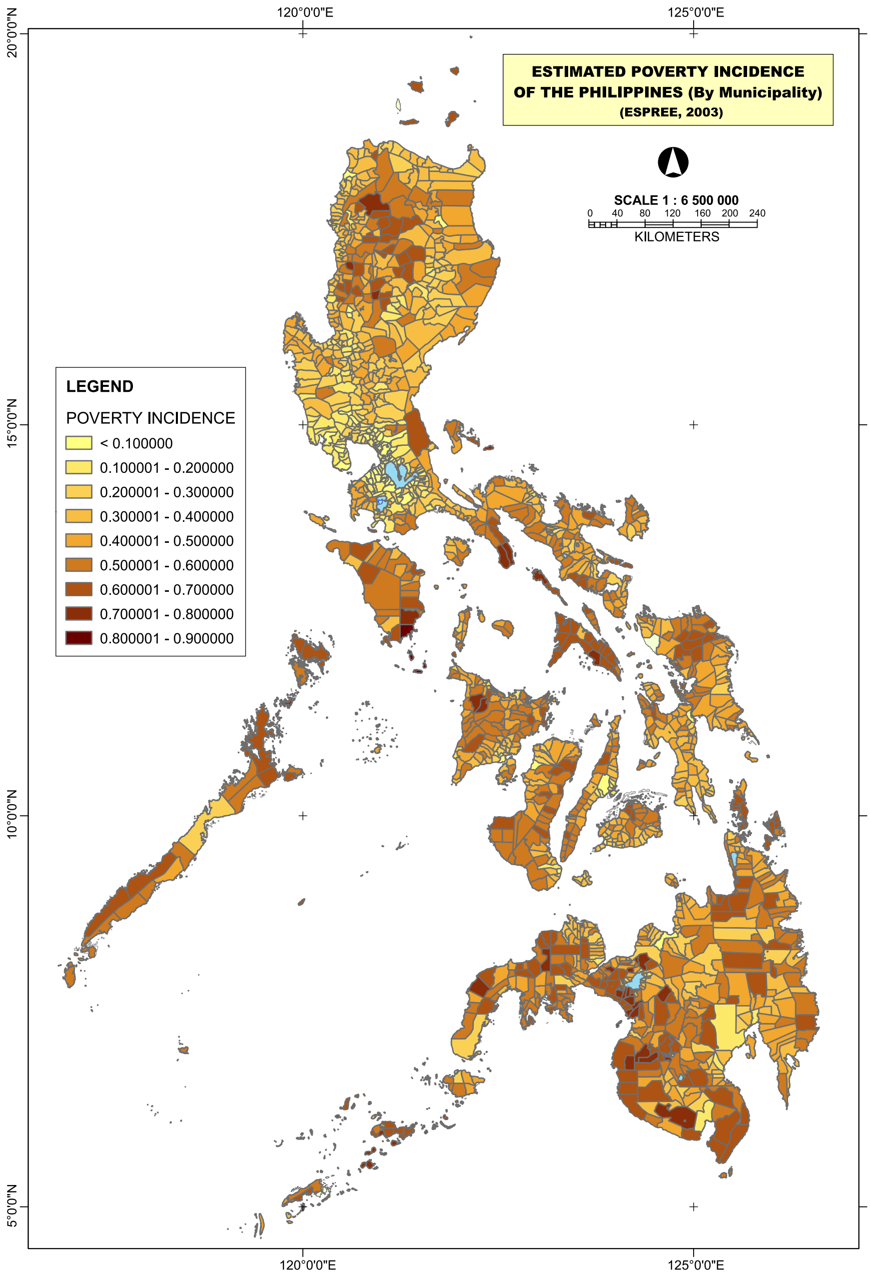

FIG. 1. 2003 *municipal level updated estimates of poverty incidence using ESPREE.*

TABLE 3
Comparison of 2003 *municipal-level updated estimates via ESPREE and LW methods*

	ESPREE			LW		
	Incidence	**SE**	**CV**	**Incidence**	**SE**	**CV**
Mean	0.4200	0.0418	0.1177	0.3755	0.0413	0.1371
Std. dev.	0.1713	0.0144	0.0620	0.1843	0.0194	0.0892
Min	0.0204	0.0038	0.0358	0.0114	0.0044	0.0140
Max	0.8937	0.1725	0.5787	0.9746	0.1812	0.8600

the combined FIES/LFS data using the Philippine standard geographic codes for the year 2000.

It appears that the average poverty incidence estimates computed using the ESPREE method are closer to the survey-based estimates than the LW estimates. This is seen more clearly in the scatter plots (Figure 3) of the ESPREE and LW estimates versus the survey-based poverty incidence estimates for all the provinces.

It can be observed from Table 4 that ESPREE generated estimated standard errors and coefficient of variation are on average slightly smaller than those generated from the LW method. It is also clear that the two methods (ESPREE and LW) generated much lower estimated SE and CV than the survey-based estimates, the average CVs being around 6% and 12% respectively. The large values of estimated CVs for the survey-based estimates are due to the sample sizes at the provincial level being too small for accurate estimation.

Although the aim is to produce small area estimates at levels lower than regional, for comparison purposes the ESPREE estimates were also accumulated to generate estimates at the regional level. Note that the direct survey estimates are regarded as having acceptably small standard errors at this level. Our regional-

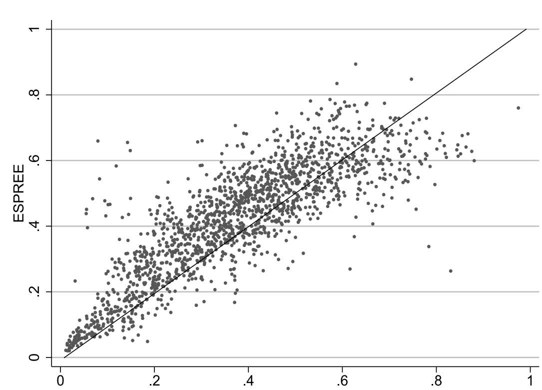

FIG. 2. *Scatter plot of* 2003 *updated municipal-level estimates by LW and ESPREE methods.*

TABLE 4
Comparison of provincial-level 2003 *direct survey-based estimates with updated estimates via ESPREE and LW methods*

	Survey-based			**LW**			**ESPREE**		
	Incidence	**SE**	**CV**	**Incidence**	**SE**	**CV**	**Incidence**	**SE**	**CV**
Mean	0.3708	0.0417	0.1241	0.3316	0.0181	0.0660	0.3677	0.0173	0.0535
Std. dev.	0.1526	0.0237	0.0655	0.1490	0.0081	0.0398	0.1440	0.0059	0.0241
Min	0.0530	0.0098	0.0523	0.0302	0.0055	0.0198	0.0457	0.0058	0.0261
Max	0.6851	0.1792	0.4875	0.6804	0.0413	0.2318	0.6408	0.0312	0.1473

level estimates were compared with the estimates from the LW method and the direct survey-based (combined FIES/LFS) estimates (Table 5). The differences between the survey-based and LW estimates, as well as between survey-based and ESPREE, are summarized by Z scores which represent the standardized distance between the two sets of estimates. The Z-scores for the ESPREE estimates are computed as follows:

$$Z = \frac{\text{ESPREE estimate} - \text{FIES estimate}}{\sqrt{(\text{ESPREE standard error})^2 + (\text{FIES standard error})^2}}.$$

In the formula for the Z-scores, although the components of the ESPREE (census-based) and FIES (survey-based) estimates may appear to be correlated, ELL and ESPREE are both based on updates of predictions of the full census information rather than on the survey data, per se. The ELL and ESPREE small area estimates are based on a model, even for households in the sample. In practice, linking respondent IDs between survey and census is almost never feasible. However, even

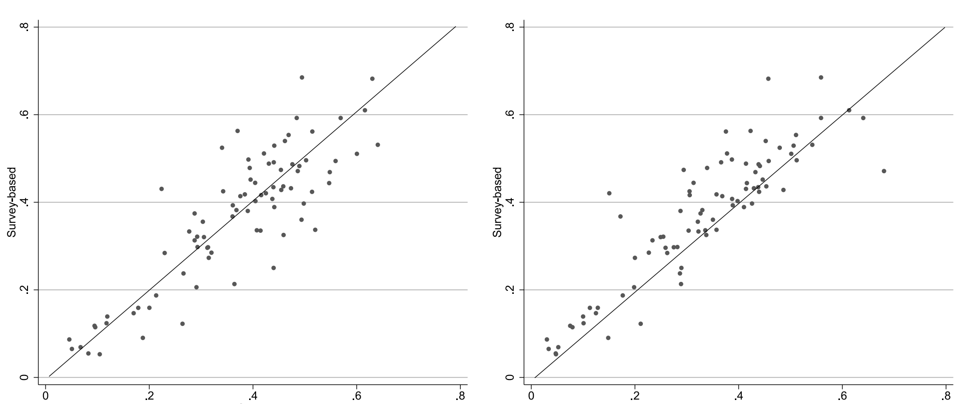

FIG. 3. *Scatter plots of provincial-level* 2003 *updated estimates by ESPREE and LW versus direct survey-based estimates.*

TABLE 5
Comparison of regional-level 2003 *direct survey-based estimates with updated estimates via ESPREE and LW methods*

	Survey-based		LW			ESPREE		
Region	**Incidence**	**SE**	**Incidence**	**SE**	Z	**Incidence**	**SE**	Z
Region I	0.3030	0.0170	0.2579	0.0144	−2.0260	0.2963	0.0126	0.3156
Region II	0.2430	0.0134	0.2639	0.0125	1.1402	0.3226	0.0147	−3.9984
Region III*	0.1720	0.0105	0.1386	0.0073	−2.6040	0.1854	0.0073	−1.0472
Region IV*	0.2443	0.0089				0.2433	0.0067	0.0958
Region V	0.4845	0.0152	0.3899	0.0119	−4.9032	0.4533	0.0139	1.5145
Region VI	0.3894	0.0154	0.3243	0.0099	−3.5593	0.3978	0.0102	−0.4563
Region VII	0.2778	0.0149	0.2717	0.0101	−0.3419	0.3481	0.0128	−3.5841
Region VIII	0.4303	0.0179	0.4199	0.0125	−0.4760	0.4133	0.017	0.6899
Region IX	0.4958	0.0205	0.4631	0.0154	−1.2743	0.4320	0.0197	2.2458
Region X	0.4137	0.0231	0.4212	0.0148	0.2730	0.3369	0.0153	2.7733
Region XI	0.3490	0.0141	0.3191	0.0155	−1.4303	0.3295	0.0146	0.9651
Region XII	0.4379	0.0291	0.3600	0.0165	−2.3296	0.4731	0.0137	−1.0948
NCR	0.0697	0.0063	0.0388	0.0053	−3.7406	0.0579	0.0044	1.5211
CAR	0.3290	0.0199	0.271	0.0133	−2.4231	0.3430	0.0174	−0.5316
ARMM	0.5520	0.0278	0.4601	0.0274	−2.3535	0.5947	0.0182	−1.2853
Region XVI	0.5300	0.0200	0.5244	0.0134	−0.2329	0.4712	0.0170	2.2422

*Using the 2000 Philippine standard geographic codes. In 2002 one of the provinces in Region IV was moved to Region III and the remaining provinces were divided into Regions IVA and IVB.

if feasible, the census is several orders of magnitude larger in size than the survey, so linking would make essentially no difference to the ELL or ESPREE based small area estimates because they are based on aggregates of a set of predictions for every census respondent. Unlike the small area estimation methods detailed in Rao (2003) then, which do not incorporate unit record census data or use it for prediction, the correlation between the survey data and the small area estimates (here the updated census predictions from ESPREE) is essentially negligible, both at small area level and also when aggregated further to regional level to control the otherwise high standard errors of the survey-based estimates. In essence, the whole census, not just the survey respondents' data, is predicted using the original ELL model and this is used in ESPREE. The very low correlation between the direct and the small area estimates is what makes the Z-scores a useful diagnostic when aggregated to a higher level. The same low correlation would apply to ESPREE updating of EBP small area estimates.

For aggregated small area estimates, it is noticeable from Table 5 that some regional-level estimates from both the ESPREE and LW methods are more than two standard errors away from the corresponding survey-based estimates, and that this is more common with the LW estimates. In addition, the average of the absolute values of the Z-scores is higher for the LW method, which means that in

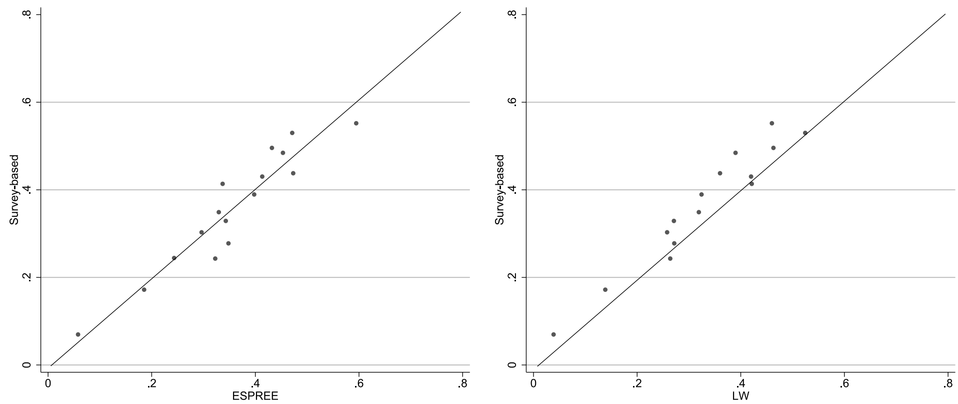

FIG. 4. *Scatter plots of regional-level* 2003 *updated estimates by ESPREE and LW versus direct survey-based estimates.*

general the ESPREE method generates regional-level poverty incidence estimates closer to the survey-based ones.

The high level of aggregation necessary for these comparisons with direct survey estimates loses much of the advantage of the small area methods, so in some cases at the higher levels of aggregation, direct estimates may be more accurate. Considering the estimated standard errors computed for the different methods, the ESPREE method tends to have lower estimated standard error compared to the survey-based except for two regions (Region II and Region XI). The LW method tends to have the lowest estimated standard error among the three methods. The lower values of the estimated standard error suggest better precision, however, the LW method appears to generate biased estimates which can be observed from the scatter plots of the regional-level estimates shown in Figure 4. The same pattern can be observed from the scatter plot of the survey-based versus the ELL updating regional-level estimates (using the revised geographic codes, i.e., the 2002 changes in regional boundaries incorporated) presented in the WB/NSCB report [see NSCB (2009), page 58]. We can see that the LW poverty incidence estimates tend to be lower than the survey-based estimates. Hence, the LW method estimates may be more precise, but they are not necessarily accurate. In any event, all estimated standard errors are conditional on the relevant model being correct.

The ESPREE poverty incidence estimates at the provincial and regional levels are evidently close to the survey-based estimates with no evidence of bias, suggesting that this property is also retained at the small area level. Thus, we can claim that the ESPREE method appears to be unbiased and performs better than the LW method. In this application, the ESPREE method is only using a model with six auxiliary variables; nevertheless, it is able to incorporate the new information from the most recent survey in an optimal manner. It appears then that the ESPREE method can be a very useful updating method for poverty estimates based on ELL and EBP given new survey but no new census data, and would be even more useful given additional computation power that would allow inclusion of more variables in the underlying multiway table. A software template for implementation of ESPREE in Stata and SAS is available from the authors.

4. Simulation study. To assess the accuracy of the variance estimates generated by the ESPREE method, a simulation study based on the Philippines data was conducted to generate empirical standard errors to be compared with the ESPREE standard error estimates. We used a model with two explanatory variables—wall type and urbanity, and selected one province with 23 municipalities for the simulation study.

We used the full survey data to estimate $\boldsymbol{\beta}_{1,t_1}$ and its associated covariance, $\mathrm{Cov}(\boldsymbol{\beta}_{1,t_1})$, and the mean of the corresponding set of pseudo-census data as an estimate of $\boldsymbol{\beta}_{2,t_0}$. Then we drew $\hat{\boldsymbol{\beta}}_{1,t_1}$ from a multivariate normal distribution with parameters $\boldsymbol{\beta}_{1,t_1}$ and $\mathrm{Cov}(\boldsymbol{\beta}_{1,t_1})$ and randomly chose a value from the set of 100 pseudo-census bootstrap estimates for $\hat{\boldsymbol{\beta}}_{2,t_0}$. A total of 1000 draws was conducted and an ESPREE estimate was calculated for each of these draws. The standard deviation of these simulated values (empirical standard error) was compared with the estimated standard errors derived from the ESPREE method in Table 6. Results show that the values of the estimated standard errors from both methods are very close to each other.

5. Validation study. Acceptability and consistency of the estimates were assessed during a field visit for validation purposes by comparing the ESPREE estimates with the expert opinion of key informants and their perception of available poverty-related indicators at the small area or municipal level. These validation activities are adapted from the validation exercises conducted for the results (small area estimates of poverty measures) of the collaborative poverty mapping project of the World Bank and National Statistical Coordination Board in the Philippines [NSCB (2005)].

The validation exercise was carried out by having a one-on-one interview with each of the identified participants or key informants (representatives from local government units such as the Municipal and Provincial Planning and Development Office, Provincial Social Welfare and Development Office, Provincial Health Office, City Planning and Development Office, and offices of National Statistics Office and National Police) using a validation form or questionnaire, an adaptation of the validation form used in the WB/NSCB poverty mapping project [NSCB (2005)]. This validation form contained the poverty related variables used in formulating the ESPREE model, other correlates of poverty and indicators of the Millennium Development Goals (MDGs). A total of thirty participants from the four provinces of Region 1 were interviewed.

One of the advantages of doing the validation study in this region is that it has comparatively stable administrative boundaries at the small area (municipality)

TABLE 6
Simulation results comparing empirical standard errors of ESPREE estimates with estimates calculated using (2.10)

Municipality	**ESPREE SE**	**Empirical SE ($n = 1000$)**
1	0.1202	0.1223
2	0.0330	0.0352
3	0.0320	0.0372
4	0.0658	0.0680
5	0.0350	0.0414
6	0.0556	0.0598
7	0.0931	0.0890
8	0.0453	0.0455
9	0.0349	0.0421
10	0.2026	0.2100
11	0.0397	0.0428
12	0.0160	0.0171
13	0.0682	0.0681
14	0.0465	0.0481
15	0.0549	0.0569
16	0.0412	0.0417
17	0.0378	0.0426
18	0.0453	0.0474
19	0.0375	0.0443
20	0.0290	0.0310
21	0.0401	0.0421
22	0.0488	0.0556
23	0.0325	0.0376

level and provinces have not moved from one region to another in the last five years. Moreover, its Regional Development Council (RDC) was one of the earliest groups to respond to the call of the Philippine Government for improvement of the implementation of poverty alleviation programs in the country. The RDC has created a masterlist of municipalities in the four provinces which are now the beneficiaries for the various poverty alleviation programs in the region [RDC-I (2008)].

There were two sets of municipal rankings gathered from the participants: (1) the indicator-based and (2) the overall level of poverty assessment. The indicator-based rank is computed as an average of the participants' ranking of municipalities in a particular province based on the indicators included in the validation form. The overall rank, on the other hand, is a single question in the validation form asking the participants to rank the municipalities in their province based on their perception of the poverty situation of the said areas. The two sets of munici-

TABLE 7
Rank correlation between participants assessment and the 2003 *updated small area estimates* (*ESPREE and LW methods*)

	Ilocos Norte		**Ilocos Sur**		**La Union**		**Pangasinan**	
	Rs	***p*-value**	**Rs**	***p*-value**	**Rs**	***p*-value**	**Rs**	***p*-value**
Indicator-based vs ESPREE	0.629	0.001	0.1614	0.3619	0.744	0.000	0.587	0.000
Overall rank vs ESPREE	0.610	0.002	0.3702	0.0312	0.837	0.000	0.062	0.677
Indicator-based vs LW	0.476	0.022	0.2046	0.2457	0.599	0.005	0.491	0.000
Overall rank vs LW	0.320	0.136	0.3106	0.0738	0.711	0.000	0.073	0.624

pal ranks were then compared with the ranking of the updated small area estimates generated from the ESPREE and LW methods.

The rank correlations (Rs) of the participants' assessment and the ESPREE and LW estimates are presented in Table 7. The ranking generated from the ESPREE method tends to be in agreement with at least one (indicator-based or overall level of poverty) of the participants' ranking in all the provinces. In addition, among those provinces in which both the ESPREE-based and LW-based ranks are significantly correlated with the participants' assessment, the ESPREE-based ranking tends to have a higher correlation coefficient estimate, signifying that the participants' assessment generally agree with the estimates generated from the ESPREE method more than the estimates generated from the LW method.

Based on the table above, we can deduce that the ESPREE estimates are in general concurring with the participants' perception of the real poverty situation of the municipalities in the provinces visited for the validation study, although there seem to be differences in Ilocos Sur and the overall ranking for Pangasinan that warrant further investigation. Details of the results per province are presented in Isidro (2009).

6. Conclusions and recommendations. In this paper a novel updating method for small area estimates of poverty measures in Third World countries has been developed. The problem of updating or generating small area estimates of poverty measures during noncensus years is an offshoot of the small area estimation procedures for poverty measures in Third World countries that use census data, such as the ELL method, which requires a survey and a census assumed to have been conducted at the same time period. Parallel considerations apply to updating EBP-based estimates from Molina and Rao (2010). Such survey-plus-census unit record-based methods cannot be used for generating small area estimates when we have a new survey but no new census. The ESPREE method provides a means of circumventing the fact that, in Third World countries, a census is usually conducted only once every ten years, while a national survey is conducted more often (once every three years in the Philippines) and updates are required.

A case study in the Philippines suggests that the method can provide adequately precise estimates which are approximately unbiased and agree well with experts' assessments. ESPREE should therefore be a valuable tool for providing timely information on the spatial distribution of poverty in the years between censuses. This applies whether the original survey-plus-census estimation was by ELL or EBP.

Classical SPREE can be used to generate updated small area estimates, but the ESPREE method has the advantage of allowing for a stochastic association structure (census or pseudo-census data) and a superpopulation model which incorporate the uncertainty in the association structure that is ignored under SPREE. Comparison of the small area estimates of poverty measures in the Philippines generated from the ESPREE and LW methods with the survey-based estimates at the provincial and regional levels showed that ESPREE estimates are unbiased and more accurate than those from the LW method based on time-invariant variables. The validation study results suggest that the ESPREE method is able to generate estimates reflecting the real poverty situation on the ground.

The current model used for ESPREE application to the Philippines data might be further improved by including area-level effects to account for the correlation between households within small areas (i.e., fitting a GLMM). It might be appropriate to include both structured and unstructured random effects, perhaps with a conditional autoregressive (CAR) or simultaneous autoregressive (SAR) model for the former accounting for the spatial correlation between small areas (although CAR models can obscure model deficiencies if not checked properly). Alternatively, geographically weighted regression (GWR), which accounts for spatial heterogeneity in model coefficients, could be considered. In addition, research would be useful on effective ways to combine ESPREE-based intercensal estimates with estimates from other data sources or estimation techniques that could lead to improvement of small area updated estimates of poverty measures.

Poverty gap and severity cannot be treated the same way as poverty incidence via a contingency table where margins and internal structure are adjusted as in ESPREE, essentially because half the multiway table would contain zeros (representing the poverty gap or severity for those people above the poverty line). It would be possible in principle to carry out an alternative analysis at household rather than aggregate level for poverty gap and severity, but this would be multiplicative computationally rather than additive [i.e., if there are r pseudo-censuses and q survey replicates, ESPREE produces $(q + r)$ sets of small area estimates, but this method would require qr sets], so would be a different method to ESPREE and generally would not be computationally feasible. Nevertheless, ESPREE can be used for a wide variety of variables of common interest that are based on proportions, for example, incidence of food poverty, proportion of people below the minimum kilocalorie requirement, and child stunting, underweight and wasting.

Further investigation is also needed on various diagnostic or evaluation methods for assessing competing small area updating methods or small area models. This might include exploration of some of the diagnostics employed by Brown et al.

(2001) and more recently by Inglese, Russo and Russo (2008). A standardized methodology for conducting a validation study for assessing acceptability and consistency of small area poverty measures estimates in Third World countries would also be useful.

APPENDIX: PROOF OF EQUATION (2.11)

The accuracy of estimation using SPREE has traditionally been formulated by defining an estimate of mean square error with respect to the survey design [Ghosh and Rao (1994), Purcell (1979)]:

$$[\widehat{\mathrm{MSE}}(\hat{\mathbf{p}}|\boldsymbol{\pi})] = V(\hat{\mathbf{p}}|\boldsymbol{\pi}) + (\hat{E}(\hat{\mathbf{p}}|\boldsymbol{\pi}) - \hat{\mathbf{P}})(\hat{E}(\hat{\mathbf{p}}|\boldsymbol{\pi}) - \hat{\mathbf{P}})'.$$

The first term on the right-hand side estimates the sample survey variance of $\hat{\mathbf{p}}$ where $\hat{\mathbf{p}}$ is the solution vector of the small area estimates based on SPREE given the census association structure $\boldsymbol{\pi}$, and the second term estimates the bias of the expected value of $\hat{\mathbf{p}}$ given $\boldsymbol{\pi}$ relative to some estimate $\hat{\mathbf{P}}$ of the true relative frequencies $\mathbf{P}$. Note that other parameters in the equation above are as defined in Sections 2.2 and 2.3.

Despite tradition, however, the true relative frequencies $\mathbf{P}$ are better considered as a superpopulation rather than a parameter fixed with respect to the survey design, since $\mathbf{P}$ is a parameter for a distribution that depends on the stochastic properties of $\boldsymbol{\pi}$ as well as $\hat{\mathbf{p}}$, rather than only on the survey design itself.

Letting $\boldsymbol{\xi}$ denote the expectation with respect to the ζ superpopulation distribution and E denote the usual expectation with respect to the survey design, with υ_ζ and V_d denoting the corresponding variances, we then have the following alternative interpretation of $[\widehat{\mathrm{MSE}}(\hat{\mathbf{p}}|\boldsymbol{\pi})]$ as an approximately unbiased estimate of the combined superpopulation and survey design variance of $\hat{\mathbf{p}}$.

Taking the joint expectation of equation (2.11) or the equation above first over the design (for which $\boldsymbol{\pi}$ is fixed) then over the superpopulation and assuming that the variance term is design unbiased yields

$$\begin{aligned}
\boldsymbol{\xi} E\{[\widehat{\mathrm{MSE}}(\hat{\mathbf{p}}|\boldsymbol{\pi})]\} &\approx \boldsymbol{\xi}[V_d(\hat{\mathbf{p}}|\boldsymbol{\pi})] + \boldsymbol{\xi} E[(E(\hat{\mathbf{p}}|\boldsymbol{\pi}) - \hat{\mathbf{P}})(E(\hat{\mathbf{p}}|\boldsymbol{\pi}) - \hat{\mathbf{P}})^T] \\
&\approx \boldsymbol{\xi}[V_d(\hat{\mathbf{p}}|\boldsymbol{\pi})] + \boldsymbol{\xi} E[(E(\hat{\mathbf{p}}|\boldsymbol{\pi}) - \mathbf{P})(E(\hat{\mathbf{p}}|\boldsymbol{\pi}) - \mathbf{P})^T] \\
&= \boldsymbol{\xi}[V_d(\hat{\mathbf{p}}|\boldsymbol{\pi})] + \boldsymbol{\xi}[(E(\hat{\mathbf{p}}|\boldsymbol{\pi}) - \mathbf{P})(E(\hat{\mathbf{p}}|\boldsymbol{\pi}) - \mathbf{P})^T] \\
&= \boldsymbol{\xi}[V_d(\hat{\mathbf{p}}|\boldsymbol{\pi})] + \upsilon_\zeta(E(\hat{\mathbf{p}}|\boldsymbol{\pi})) \\
&= E\boldsymbol{\xi}[(\hat{\mathbf{p}} - E\boldsymbol{\xi}(\hat{\mathbf{p}}))(\hat{\mathbf{p}} - E\boldsymbol{\xi}(\hat{\mathbf{p}}))^T],
\end{aligned}$$

which is the (unconditional) ζd variance of $\hat{\mathbf{p}}$.

Note that the first approximation replaces $\hat{E}(\hat{\mathbf{p}}|\boldsymbol{\pi})$ by $E(\hat{\mathbf{p}}|\boldsymbol{\pi})$ since the bias term being estimated in equation (2.11) is $E(\hat{\mathbf{p}}|\boldsymbol{\pi}) - \mathbf{P}$, the second approximation drops the terms involving $(\hat{\mathbf{P}} - \mathbf{P})$ as being rather smaller in magnitude than the quadratic term in $(E(\hat{\mathbf{p}}|\boldsymbol{\pi}) - \mathbf{P})$, and the last line follows from the usual formula for conditional variance.

REFERENCES

AGRESTI, A. (2002). *Categorical Data Analysis*, 2nd ed. Wiley, New York. MR1914507

BELL, W., BASEL, W., CRUSE, C., DALZELL, L., MAPLES, J., O'HARA, B. and POWERS, D. (2007). Use of acs data to produce saipe model-based estimates of poverty for counties. Available at http://www.census.gov/did/www/saipe/publications/files/report.pdf.

BROWN, G., CHAMBERS, R., HEADY, P. and HEASMAN, D. (2001). Evaluation of small area estimation methods—an application to unemployment estimates from the UK LFS. In *Proceedings of Statistics Canada Symposium*. Hull, Quebec. Available at https://www.researchgate.net/publication/237775200_Evaluation_of_small_area_estimation_methods_-_An_application_to_unemployment_estimates_from_the_UK_LFS.

CHRISTIAENSEN, L., LANJOUW, P., LUOTO, J. and STIFEL, D. (2010). The reliability of small area estimation prediction methods to track poverty. Working paper No. 2010/99, United Nations University/World Institute for Development Economics Research (UNU/WIDER), 2010.

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

FOSTER, J., GREER, J. and THORBECK, E. (1984). A class of decomposable poverty measures. *Econometrica* **52** 761–766.

GHOSH, M. and RAO, J. N. K. (1994). Small area estimation: An appraisal. *Statist. Sci.* **9** 55–93. MR1278679

GONZALEZ, M. E. (1973). Use and evaluation of synthetic estimators. In *Proceedings of the Social Statistics Section, American Statistical Association, Joint Statistical Meetings* 33–36. American Statistical Association, New York.

GONZALEZ, M. E. and WAKSBERG, J. (1973). Estimation of the error of synthetic estimates. Presented at the First Meeting of the International Association of Survey Statisticians. Vienna, Austria; US Department of Commerce, Social and Economic Statistics, Bureau of Census, Washington, DC.

HASLETT, S. (2013). Small area estimation of poverty using the ELL/PovMap method, and its alternatives. In *Poverty and Social Exclusion: New Methods of Analysis* (G. Betti and A. Lemmi, eds.). Routledge, London.

HASLETT, S. (2016). Small area estimation using both survey and census unit record data: Links, alternatives, and the central roles of regression and contextual variables. In *Analysis of Poverty Data by Small Area Estimation* (M. Pratesi, eds.). Wiley, New York. MR3380626

HASLETT, S., GREEN, A. and ZINGEL, C. (1998). Small area estimation given regular updates of census auxiliary variables. In *Proceedings of the New Techniques and Technologies for Statistics Conference*. Sorrento, Italy.

HASLETT, S. J., ISIDRO, M. C. and JONES, G. (2010). Comparison of regression techniques in the context of small area estimation of poverty. *Survey Methodlogy* **36** 157–170.

HASLETT, S. J. and JONES, G. (2005). Local estimation of poverty in the Philippines. Philippine National Statistics Co-ordination Board/World Bank report. Available at http://siteresources.worldbank.org/INTPGI/Resources/342674-1092157888460/Local_Estimation_of_Poverty_Philippines.pdf.

HASLETT, S. J. and JONES, G. (2010). Small-area estimation of poverty: The aid industry standard and its alternatives. *Aust. N. Z. J. Stat.* **52** 341–362. MR2791524

HASLETT, S. J., JONES, G. and SEFTON, A. (2013). Small-area estimation of poverty and malnutrition in Cambodia. National Institute of Statistics, Ministry of Planning, Royal Government of Cambodia and the United Nations World Food Programme, Cambodia. Available at http://www.wfp.org/content/cambodia-small-area-estimation-poverty-and-malnutrition-april-2013.

HAWALA, S. and LAHIRI, P. (2012). Hierarchical bayes estimation of poverty rates. Available at http://www.census.gov/did/www/saipe/publications/conference.html.

HOOGEVEEN, J., EMWANU, T. and OKWI, P. (2006). Updating poverty maps with panel data. *World Dev.* **34** 2075–2088.

INGLESE, F., RUSSO, A. and RUSSO, M. (2008). Different approaches for evaluation precision small area model-based estimators. In *Proceedings of Q*2008 *European Conference on Quality in Official Statistics*. Rome. Available at http://q2008.istat.it/sessions/paper/37Inglese.pdf.

ISIDRO, M. C. (2009). Validation study for updated poverty estimates in the Philippines. A report submitted to the New Zealand Postgraduate Study Abroad Awards.

ISIDRO, M. C. (2010). Intercensal updating of small area estimates. Unpublished Ph.D. thesis, Massey Univ., New Zealand.

JITSUCHON, S. and LANJOUW, P. (2005). Updating poverty maps: Emerging experience with available methods and data from Thailand. Preprint.

LANJOUW, P. and VAN DER WIEDE, R. (2006). Determining changes in welfare distributions at the micro-level: Updating poverty maps. Powerpoint presentation at the NSCB Workshop for the NSCB/World Bank Intercensal Updating Project.

LUMLEY, T. (2014). Survey: Analysis of complex survey samples. R package Version 3.30.

LUMLEY, T. and SCOTT, A. (2015). AIC and BIC for modeling with complex survey data, Version 3.30, DOI:10.1093/jssam/smu021.

MARKER, D. (1999). Organization of small area estimators using a generalized linear regression framework. *J. Off. Stat.* **15** 1–24.

MOLINA, I., NANDRAM, B. and RAO, J. N. K. (2014). Small area estimation of general parameters with application to poverty indicators: A hierarchical Bayes approach. *Ann. Appl. Stat.* **8** 852–885. MR3262537

MOLINA, I. and RAO, J. N. K. (2010). Small area estimation of poverty indicators. *Canad. J. Statist.* **38** 369–385. MR2730115

NAMAZI-RAD, M.-R. and STEEL, D. (2015). What level of statistical model should we use in small area estimation? *Aust. N. Z. J. Stat.* **57** 275–298. MR3366374

NOBLE, A., HASLETT, S. J. and ARNOLD, G. (2002). Small area estimation via generalized linear models. *J. Off. Stat.* **18** 45–60.

NOBLE, A., HASLETT, S. J. and ARNOLD, G. (2006). Erratum to: Small area estimation via generalized linear models. *J. Off. Stat.* **22** 159–160; *J. Off. Stat.* **18** (2002) 45–60.

NSCB (2005). Estimation of local poverty in the Philippines. National Statistical Coordination Board, Philippines. Available at http://www.nscb.gov.ph/poverty/sae/NSCB_LocalPovertyPhilippines.pdf.

NSCB (2009). 2003 City and Municipal level poverty estimates. National Statistical Coordination Board, Philippines. Available at http://www.nscb.gov.ph/poverty/sae/2003%20SAE%20of%20poverty%20(Full/%20Report).pdf.

PURCELL, N. J. (1979). Efficient estimation for small domains: A categorical data analysis approach. Unpublished Ph.D. thesis, Univ. Michigan.

PURCELL, N. J. and KISH, L. (1980). Postcensal estimates for local areas. *Int. Stat. Rev.* **48** 3–18.

RAO, J. N. K. (2003). *Small Area Estimation*. Wiley, Hoboken, NJ. MR1953089

RDC-I (2008). Approving the common masterlist of priority program benefeciaries (CMBPP) for region I poverty reduction program. RDC Resolution No. 57 S.

STATACORP (2007). Stata Statistical Software: Release 10. StataCorp LP, College Station, TX.

TAROZZI, A. and DEATON, A. (2009). Using census and survey data to estimate poverty and inequality for small areas. *Rev. Econ. Stat.* **91** 773–792.

WOLTER, K. M. (1985). *Introduction to Variance Estimation*. Springer, New York. MR0799715

M. ISIDRO
G. JONES
INSTITUTE OF FUNDAMENTAL SCIENCES
MASSEY UNIVERSITY
PB 11222
PALMERSTON NORTH, 4442
NEW ZEALAND
E-MAIL: maris_isidro@yahoo.com
g.jones@massey.ac.nz
URL: www.massey.ac.nz
www.massey.ac.nz/massey/learning/colleges/college-of-sciences/about/fundamental-sciences/staff/?stref=482430

S. HASLETT
INSTITUTE OF FUNDAMENTAL SCIENCES
MASSEY UNIVERSITY
PB 11222
PALMERSTON NORTH, 4442
NEW ZEALAND
AND
STATISTICAL CONSULTING UNIT
AUSTRALIAN NATIONAL UNIVERSITY
ACTON
CANBERRA
ACT 2601
AUSTRALIA
E-MAIL: stephen.haslett@anu.edu.au
s.j.haslett@massey.ac.nz
URL: www.anu.edu.au
https://services.anu.edu.au/business-units/professor-stephen-haslett