# A Robust ELL Methodology for Poverty Mapping

Sumonkanti Das *
University of Wollongong, Wollongong, Australia – sd535@uowmail.edu.au
Ray Chambers
University of Wollongong, Wollongong, Australia – ray@uow.edu.au

**Abstract**

The ELL methodology for poverty mapping was developed by the World Bank and is widely used in developing countries. However, it has been criticized because of its assumption of negligible between area variability when used to calculate small area poverty estimates. In particular, the mean squared errors of these estimates are significantly underestimated when the between area variability in the income and expenditure data cannot be adequately explained by the explanatory and contextual variables in the model. In this paper a method of mean squared error estimation for ELL-type estimates is proposed which is robust to the presence of significant unexplained between area variability. Simulation results show that the proposed method performs better than the classic ELL methodology for poverty estimation.

**Keywords:** Mean Squared Error; Model Misspecification; Poverty Mapping; Small Area Estimation.

## 1. Introduction

The World Bank poverty mapping methodology (Elbers, Lanjouw & Lanjouw; 2002, 2003; henceforth ELL) has been applied in many developing countries. The basic idea of the ELL method is to combine household (HH) survey data with the HH records of a recent census, though any HH level linkage of these datasets is not required. The ELL method provides poverty estimates, with their estimated precision, at a specified local area level. Since the ELL method is based on an area homogeneity assumption, the main criticism leveled against it is its assumption of random effects at the survey cluster level rather than at the local area level. To reduce the impact of possible between area heterogeneity, a large number of explanatory and contextual variables are typically included in the regression model. Tarozzi & Deaton (2009) claim however that the resulting area homogeneity assumption may not be always true, leading to misleading inference when this assumption is violated. They also question the assumption of homoskedastic, independent and identically distributed cluster random effects since small areas within a large region are likely to be correlated. In such cases random cluster effects may be correlated if the model regressors fail to adequately capture this between cluster correlation. In response to these criticisms, Demombynes *et al.* (2007) show that they may be partially addressed if area-level contextual variables are included into the HH survey data. Furthermore, Elbers *et al.* (2008) describe the results of a validation study using the 2000 Census dataset of Minas Gerais state in Brazil, which is large enough to have heterogeneous local areas and so ELL homogeneity assumptions must fail. They check this assumption by developing a state-level model with municipality-level regressors and 853 municipality-level models with HH-level and enumeration-level regressors, concluding that the state level model is adequate for estimating welfare at area level as long as it captures local heterogeneity by including appropriate area-level regressors.

The estimated values of inter-cluster and intra-cluster correlation coefficients are important components of the simulation phase of the ELL method. In particular, the variance of the welfare predictions obtained following this simulation can be understated when inter–cluster and intra–cluster correlations are large and are not explicitly accounted for in the HH model (Tarozzi & Deaton, 2009). Since the estimated location effect cannot be separated into area-level and cluster-level effects in the ELL method, one has to assume that this effect is either entirely a cluster-level effect – an optimistic assumption that rules out any correlation at higher level – or entirely an area-level effect – a conservative assumption that will likely lead to an overstatement of mean squared error (MSE) estimates (Elbers *et al.*, 2002; Demombynes *et al.*, 2007). As Tarozzi & Deaton (2009) note, the "conservative" assumption should lead to imprecise and unusable estimates, while the "optimistic" assumption is obviously necessary for the validity of the ELL methodology, but leads to understated MSEs if incorrect. Elbers *et al.* (2008) applied both "optimistic" (also "standard") and "conservative" methods in the Minas Gerais study. When the "optimistic" method was applied, the confidence intervals were found to be narrower than those generated by the "conservative" approach. In particular, these authors found that 42% of municipalities could be statistically distinguished from one another using the "optimistic" method compared with 35% using the "conservative" method at a 95% level of confidence. They conclude that the ELL method performs well even under some violation of its assumptions. However, neither method provides a good solution for the problem of between area heterogeneity. Elbers *et al.* (2008) mention that the "conservative" method can be applied where some spatial correlation of errors remains after controlling for between cluster heterogeneity. However, as noted earlier, Tarozzi & Deaton (2009) argue that this "conservative" method may produce unstable and useless estimates. Consequently, the question of which approach a researcher should adopt becomes important.

We propose a different approach to solving this problem of possible between area heterogeneity. We start by developing the analytical relationship between moment-based estimators of variance components when a level is ignored in a three level hierarchical model. From this relationship we show that the estimator of the level 2 variance component calculated under an incorrect two level model always underestimates the sum of the true level 2 and level 3 variance components. This relationship allows us to identify a robust variance estimator of the area mean that is unbiased under the three level model and is also approximately unbiased under the two level model. This robust variance estimator is then used to adjust the estimate of the level 2 variance component used in the ELL simulation procedure. We note in passing that this adjustment is aimed purely at improving the estimated MSEs of the poverty estimates produced by the ELL method. Standard ELL methods are still used to calculate confidence intervals for poverty measures, since under our approach only the cluster-level variance component is modified in the ELL bootstrap procedure. The paper is organized as follows: Section 2 summarizes the moment-based variance component estimation method both when a three level model is perfectly specified and also when it is misspecified by ignoring a level; Section 3 presents a robust variance estimator of the area mean under this type of model misspecification; Section 4 describes the standard ELL and the robustified ELL methodology; Section 5 illustrates the performance of these methodologies via a small scale simulation experiment; our conclusions are set out in Section 6.

## 2. Estimation of Variance Components

The variance components of a multilevel model are classically estimated using maximum likelihood (ML), restricted ML (REML), Henderson method III, expectation maximization (EM), and via moment-based least squares (Searle *et al.*, 1992). Here we focus on the moment-based estimators. In particular, let $y_{ijk}$ denote the value of Y for the $k^{th}$ HH (level-1) belonging to the $j^{th}$ cluster (level-2) of the $i^{th}$ area (level-3). A three level multilevel linear model for Y can then be written as

$$(1)\quad y_{ijk} = x_{ijk}\beta + \eta_i + u_{ij} + \varepsilon_{ijk};\quad i=1,2,...,D;\quad j=1,2,...,C_i;\quad k=1,2,....,N_{ij}\ .$$

Here $\eta_i \sim N\left(0,\sigma_\eta^2\right)$, $u_{ij} \sim N\left(0,\sigma_u^2\right)$ and $\varepsilon_{ijk} \sim N\left(0,\sigma_\varepsilon^2\right)$ are level-specific zero mean homogeneous random effects. The linear regression model implied by (1) can be fitted to the sample data using least squares, in which case the area-specific, cluster-specific, and HH-specific effects can be estimated as

$$\hat{\eta}_i = n_i^{-1}\sum_{j=1}^{C_i}\sum_{k=1}^{n_{ij}}\hat{e}_{ijk} = n_i^{-1}\sum_{j=1}^{C_i} n_{ij}\hat{u}_{ij},\quad \hat{u}_{ij} = n_{ij}^{-1}\sum_{k=1}^{n_{ij}}\hat{e}_{ijk},\quad \hat{\varepsilon}_{ijk} = \hat{e}_{ijk} - \hat{\eta}_i - \hat{u}_{ij},\ \text{where}\ \hat{e}_{ijk} = y_{ijk} - \hat{y}_{ijk}\ .$$

Under (1), the expectations of the sample residual variances at each level are

$$E\begin{bmatrix} s^{(1)} \\ s^{(2)} \\ s^{(3)} \end{bmatrix} = E\begin{bmatrix} (n-1)^{-1}\sum_{i=1}^{D}\sum_{j=1}^{C_i}\sum_{k=1}^{n_{ij}}\left(\hat{e}_{ijk} - \hat{\bar{e}}_{...}\right)^2 \\ (C-1)^{-1}\sum_{i=1}^{D}\sum_{j=1}^{C_i} n_{ij}\left(\hat{\bar{e}}_{ij.} - \hat{\bar{e}}_{...}\right)^2 \\ (D-1)^{-1}\sum_{i=1}^{D} n_i\left(\hat{\bar{e}}_{i..} - \hat{\bar{e}}_{...}\right)^2 \end{bmatrix} = \begin{bmatrix} 1 & (n-1)^{-1}\left(n-\bar{n}_0^{(2)}\right) & (n-1)^{-1}\left(n-\bar{n}_0^{(3)}\right) \\ 1 & (C-1)^{-1}\left(n-\bar{n}_0^{(2)}\right) & (C-1)^{-1}\left(n-\bar{n}_0^{(3)}\right) \\ 1 & (D-1)^{-1}\left(\sum_{i=1}^{D}\bar{n}_{0i}^{(2)} - \bar{n}_0^{(2)}\right) & (D-1)^{-1}\left(n-\bar{n}_0^{(3)}\right) \end{bmatrix}\begin{bmatrix} \sigma_\varepsilon^2 \\ \sigma_u^2 \\ \sigma_\eta^2 \end{bmatrix}$$

where $\bar{n}_{0i}^{(2)} = n_i^{-1}\sum_{j=1}^{C_i} n_{ij}^2$, $\bar{n}_0^{(2)} = n^{-1}\sum_{i=1}^{D}\sum_{j=1}^{C_i} n_{ij}^2$, $\bar{n}_0^{(3)} = n^{-1}\sum_{i=1}^{D} n_i^2$ . Note that the right hand side above can be written as $A\Lambda$ , where $\Lambda$ is the vector of variance components. If the coefficient matrix A is non-singular, unbiased estimators of the variance components can be easily obtained (Tranmer & Steel, 2001) as

$$(2) \qquad \hat{\Lambda} = \begin{bmatrix} \hat{\sigma}_\varepsilon^2 & \hat{\sigma}_u^2 & \hat{\sigma}_\eta^2 \end{bmatrix}^{/} = A^{-1}\begin{bmatrix} s^{(1)} & s^{(2)} & s^{(3)} \end{bmatrix}^{/} .$$

If the third level of (1) is ignored and a two level model is fitted instead to the sample data, the moment estimators can be expressed as

$$\begin{bmatrix} \hat{\sigma}_{\varepsilon(2)}^2 \\ \hat{\sigma}_{u(2)}^2 \end{bmatrix} = A_1^{-1}\begin{bmatrix} s^{(1)} \\ s^{(2)} \end{bmatrix} \quad \text{where, } A_1 = \begin{bmatrix} c_1^{(1)} & c_2^{(1)} \\ c_1^{(2)} & c_2^{(2)} \end{bmatrix} .$$

Here $\hat{\sigma}_{i(j)}^2$ is the $i^{th}$ variance component of the $j$-level model and $c_j^{(i)}$ is the value in the $(i,j)^{th}$ cell of A. The expectations of the estimated variance components under (1) are therefore

$$E_3\begin{bmatrix} \hat{\sigma}_{\varepsilon(2)}^2 \\ \hat{\sigma}_{u(2)}^2 \end{bmatrix} = \begin{bmatrix} \sigma_\varepsilon^2 \\ \sigma_u^2 + \left(n - \bar{n}_0^{(3)}\right) / \left(n - \bar{n}_0^{(2)}\right)\sigma_\eta^2 \end{bmatrix} .$$

Thus, when the third level is ignored, $\hat{\sigma}_{\varepsilon(2)}^2$ is unbiased but the expected value of $\hat{\sigma}_{u(2)}^2$ is not equal to the sum of the true second and third level variance components under the true three level model.

**3. Variance Estimation for an Area Mean under Model Misspecification**

If a level in the hierarchy of a model like (1) is ignored in inference, then the inferential model may be considered as misspecified. Under (1), the variance of an area mean and its plug-in estimator can be written as

$$\mathrm{Var}\left(\bar{Y}_i\right)\Big|_{(3)} = N_i^{-1}\sigma_{\varepsilon(3)}^2 + N_i^{-2}\sum_{j=1}^{C_i} N_{ij}{}^2\sigma_{u(3)}^2 + \sigma_{\eta(3)}^2 \quad \text{and} \quad \hat{\mathrm{V}}\left(\bar{Y}_i\right)\Big|_{(3)} = N_i^{-1}\hat{\sigma}_{\varepsilon(3)}^2 + N_i^{-2}\sum_{j=1}^{C_i} N_{ij}{}^2\hat{\sigma}_{u(3)}^2 + \hat{\sigma}_{\eta(3)}^2 .$$

Similarly under a two level model, $\mathrm{Var}\left(\bar{Y}_i\right)$ and its plug-in estimator are

$$(3) \qquad \mathrm{Var}\left(\bar{Y}_i\right)\Big|_{(2)} = N_i^{-1}\sigma_{\varepsilon(2)}^2 + N_i^{-2}\sum_{j=1}^{C_i} N_{ij}{}^2\sigma_{u(2)}^2 \quad \text{and} \quad \hat{\mathrm{V}}\left(\bar{Y}_i\right)\Big|_{(2)} = N_i^{-1}\hat{\sigma}_{\varepsilon(2)}^2 + N_i^{-2}\sum_{j=1}^{C_i} N_{ij}{}^2\hat{\sigma}_{u(2)}^2 .$$

The expected values of these variance estimators under the true three level model are

$$(4) \qquad E_3\left[\hat{\mathrm{V}}\left(\bar{Y}_i\right)\Big|_{(3)}\right] = N_i^{-1}\sigma_{\varepsilon(3)}^2 + N_i^{-2}\sum_{j=1}^{C_i} N_{ij}{}^2\sigma_{u(3)}^2 + \sigma_{\eta(3)}^2 ,$$

$$(5) \qquad E_3\left[\hat{\mathrm{V}}\left(\bar{Y}_i\right)\Big|_{(2)}\right] = N_i^{-1}\sigma_{\varepsilon(3)}^2 + N_i^{-2}\sum_{j=1}^{C_i} N_{ij}{}^2\left\{\sigma_{u(3)}^2 + \left(n - \bar{n}_0^{(2)}\right)^{-1}\left(n - \bar{n}_0^{(3)}\right)\sigma_{\eta(3)}^2\right\} .$$

Note that $E_3\left[\hat{\mathrm{V}}\left(\bar{Y}_i\right)\Big|_{(2)}\right] < E_3\left[\hat{\mathrm{V}}\left(\bar{Y}_i\right)\Big|_{(3)}\right]$. We can adjust (or 'robustify') $\hat{\mathrm{V}}\left(\bar{Y}_i\right)\Big|_{(2)}$ to make it unbiased or approximately unbiased under (1). This leads to the adjusted estimator

$$(6) \qquad \hat{\mathrm{V}}^M\left(\bar{Y}_i\right)\Big|_{(2)} = N_i^{-1}\hat{\sigma}_{\varepsilon(2)}^2 + N_i^{-2}\sum_{j=1}^{C_i} N_{ij}{}^2\left\{\hat{\sigma}_{u(3)}^2 + \left(N_i^2 \Big/ \sum_{j=1}^{C_i} N_{ij}^2\right)\hat{\sigma}_{\eta(3)}^2\right\} .$$

The variance estimator (6) is robust under model misspecification, since under a true two level model, $\hat{s}_{h(3)}^2$ will be very small close to zero, and so the third term will be negligible. That is, estimator (6) is unbiased under a three level model and approximately unbiased under a two level model.

## 4. Standard and Robustified ELL Methodology

From the perspective of ELL methodology, the domains of interest (the small areas) are homogeneous and all between area variation is due to between cluster variation. Consequently, a two level regression model (HH as level one, cluster as level two) is used, ignoring domains. If there is a negligible domain area effect, the ELL method leads to unbiased estimators of poverty indicators, but underestimated MSEs. The basic steps of the standard ELL methodology can be explained briefly as follows:

**Step 1:** Fit a two level model to the sample data to obtain estimates of regression coefficients and variance components using a suitable estimation method. **Step 2:** Each cluster in the census data base is assigned a cluster-specific error $\hat{u}_{ij}$ drawn from $N\left(0,\hat{\sigma}^2_{u(2)}\right)$. **Step 3:** Each individual in the census data base is assigned an individual-specific error $\hat{\varepsilon}_{ijk}$ drawn from $N\left(0,\hat{\sigma}^2_{\varepsilon(2)}\right)$. **Step 4:** Simulated population values for $y_{ijk}$ are generated via $y_{ijk}=x^{/}_{ijk}\hat{\beta}+\hat{u}_{ij}+\hat{\varepsilon}_{ijk}$ and the corresponding values of area specific parameters such as mean, quantiles, poverty indicators are recorded. **Step 5:** Steps (2)-(4) are repeated a large number of times (say, 1000) and the mean and variance over these simulations of the area-specific parameters of interest are then used as their estimates and estimated MSEs, respectively.

If domains have a small but significant effect, and are ignored in modeling, the ELL procedure needs to be modified to overcome the resulting MSE underestimation. Since the moment estimator of the level one variance component is unbiased even when the level three component of variance is ignored in the fitted model, an appropriately modified level two variance component must be used in the ELL simulation process. A number of such adjustments or robustifications are proposed below:

**Adjustment 1:** Replace $\hat{\sigma}^2_{u(2)}$ in the ELL procedure by $\hat{\hat{\sigma}}^2_u=k_1*\hat{\sigma}^2_{u(2)}$ where $k_1=\hat{\sigma}^{-2}_{u(2)}\left[\hat{\sigma}^2_{u(3)}+\hat{\sigma}^2_{\eta(3)}\sum_{i=1}^{D}\left(n_i^2\Big/\sum_{j=1}^{C_i}n_{ij}^{\ 2}\right)\right]$. This assumes $\left(n_i^2\Big/\sum_{j=1}^{C_i}n_{ij}^{\ 2}\right)\cong\left(N_i^2\Big/\sum_{j=1}^{C_i}N_{ij}^{\ 2}\right)$ for $i=1,2,...,D$; otherwise it will still underestimate. Also zero sample sizes in some areas mean that no adjustment occurs for them. **Adjustment 2:** As with adjustment 1, replace $\hat{\sigma}^2_{u(2)}$ by $\hat{\hat{\sigma}}^2_u=k_2*\hat{\sigma}^2_{u(2)}$ where $k_2=\hat{\sigma}^{-2}_{u(2)}\left[\hat{\sigma}^2_{u(3)}+\hat{\sigma}^2_{\eta(3)}\sum_{i=1}^{D}\left(N_i^2\Big/\sum_{j=1}^{C_i}N_{ij}^{\ 2}\right)\right]$. Theoretically, adjustment 2 will work better than adjustment 1 provided area-specific population sizes $\left(N_i\right)$ are known. **Adjustment 3:** The procedure used for Adjustment 2 will give higher weight $\left(k_2\right)$ to smaller areas and the lower weight to larger areas. To reduce this variation, clusters can be assigned to $P$ strata according to their population size (number of HHs here) and the second level variance component adjustment described in Adjustment 2 carried out separately within each stratum. That is, the single level two variance component used in simulation Step (2) of the standard ELL method is replaced by stratum-specific adjusted level two variance components of the form

$$\hat{\hat{\sigma}}^2_u\Big|_p=k_{3(p)}*\hat{\sigma}^2_{u(2)}\ ,\quad \text{where}\ \ k_{3(p)}=\hat{\sigma}^{-2}_{u(2)}\left[\hat{\sigma}^2_{u(3)}+\hat{\sigma}^2_{\eta(3)}\sum_{i=1}^{D_p}\left(N_i^2\Big/\sum_{j=1}^{C_i}N_{ij}^{\ 2}\right)\right]\ ;\ p=1,..,P\ .$$

Finally, the "optimistic" adjustment of Elbers *et al.* (2008) where the total location effect $\hat{\sigma}^{2(opt)}_{u(2)}=\hat{\sigma}^2_{u(3)}+\hat{\sigma}^2_{\eta(3)}$ is applied at cluster-level in the bootstrap procedure and the corresponding "pessimistic" adjustment of Tarozzi & Deaton (2009), where the location effect is defined at the area level rather than at the cluster level, are denoted "**Adjustment 4**" and "**Adjustment 5**", respectively.

## 5. Numerical Evaluation

A three level hierarchical population consisting of D=75 small areas, C=1650 clusters, and H=180450 HHs is used to evaluate the proposed adjustments using Monte Carlo simulation. The number of clusters per area varies in the range 15-29 and the cluster sizes vary in the range 96-120 HHs. The population is structured so that it can be partitioned into 5 sub-populations by size (number of HHs). A two stage sampling procedure is used, with 2-4 clusters randomly selected from each area and then 10 HHs randomly selected from these selected clusters. Using this procedure, a random two stage cluster sample of size n=2250 is drawn in each simulation. The following normal and lognormal population models were used to generate the population data for the simulation study:

$$(7) \quad y_{ijk} = 20 + \eta_i + u_{ij} + \varepsilon_{ijk};\; \varepsilon_{ijk} \sim N(0, 0.80);\; u_{ij} \sim N(0, 0.15);\; \eta_i \sim N(0, 0.05);$$

$$(8) \quad y_{ijk} = \exp\left(X_{ijk}\beta + \eta_i + u_{ij} + \varepsilon_{ijk}\right);\; \varepsilon_{ijk} \sim N(0, 0.20);\; u_{ij} \sim N(0, 0.035);\; \eta_i \sim N(0, 0.015).$$

The explanatory variables were drawn from a multivariate normal distribution as

$$X = \begin{bmatrix} X_0 & X_1 & X_2 \end{bmatrix};\; X_0 = 1_N;\; \begin{bmatrix} X_1 & X_2 \end{bmatrix}' \sim MVN\left[\begin{pmatrix} 0.50 \\ 0.75 \end{pmatrix}, \begin{pmatrix} 1.5 & 0.10 \\ 0.10 & 0.95 \end{pmatrix}\right];\; \beta = \begin{bmatrix} 6 & 0.5 & -0.55 \end{bmatrix}'.$$

In the simulation study, variances of area-specific means and the three main FGT poverty indicators (Foster, Greer, & Thorbecke, 1984), i.e. Head Count Rate (FGT0), Poverty Gap (FGT1) and Poverty Severity (FGT2), were estimated using the standard and the five different robustified ELL type approaches. In particular, variances of area means were investigated for the first population, and variances of FGT indicators at the $30^{th}$ percentile of Y were investigated for the second population. The MC simulation process was repeated 1000 times, with 1000 bootstrap simulations in each MC simulation. Performance measures corresponding to the 95% nominal coverage rate (CR) and the corresponding average confidence interval width (CIW) were calculated and then compared with true nominal 95% and true average CIWs respectively. The area-specific variance estimates of the parameters of interest were also compared with the corresponding simulation-based true variance estimates in terms of absolute relative bias (ARB) and relative root MSE (RRMSE).

The area averaged ARB values shown in Table 1 clearly show that the MSEs of the standard ELL estimates are underestimated, and that this underestimation is compensated for in all target parameters by the proposed adjustments. We also see that adjustment-3 does better than any of the other adjustments for both the mean and the FGT indicators in terms ARB, RRMSE, CR and CIW. Though the variance estimates for FGT0 still exhibit a slight downward bias, it is much smaller than the bias displayed by the stable but wrong standard ELL estimates (Figure 1). On the other hand, the "optimistic" approach underestimates the MSE while the "conservative" approach overestimates the MSE (Table 1 & Figure 1), as expected. Area-specific variance estimates (Figure 1) also confirms that Adjustment-3 is the best choice for the simulated population. The variation in the performance of poverty gap (PG) and poverty severity (PS) estimates (Figure 1) might be due to their complex non-linear form or back-transformation of the log-transformed response values to original scale when FGT indicators are estimated. However, estimates derived using the proposed robustified ELL approach still perform better than those obtained via the standard ELL approach.

## 6. Conclusions

We show that the basic ELL method will lead to incorrect inference if its area-homogeneity assumption fails. We then propose a robustified version of the ELL method that is based on the relationship between estimated variance components under an incorrect two level working model and corresponding estimates under a correct three level true model. The proposed method seemed to adequately correct the bias of MSE estimation of area-specific means and FGT indicators in our simulations of this case. However, although our simulation study indicates that modified ELL variance estimators perform well when a three level model is true, it is possible that this method may not be robust to non-normality of the different model errors. There is also the issue that the strata underpinning the modified bootstrap procedure of Adjustment-3 are based the fact that the level 2 variance component vary between these strata according to their population sizes. This raises the issue of optimal stratification for Adjustment-3. Moreover, the geographic locations of the areas of interest is neglected in the proposed adjustment methods. Extension of these ideas to where there is spatial correlation therefore seems worthy of further research.

## References

Demombynes, G., Elbers, C., Lanjouw, J., & Lanjouw, P. (2007). How Good a Map? Putting Small Area Estimation to the Test. *World Bank Policy Research Working Paper, 4155*. Washington, DC: The World Bank.

Elbers, C., Lanjouw, J. O., & Lanjouw, P. (2002). *Micro-level Estimation of Welfare* (Vol. 2911). Washington D.C.: Development Research Group, the World Bank.

Elbers, C., Lanjouw, J. O., & Lanjouw, P. (2003). Micro-Level Estimation of Poverty and Inequality. *Econometrica, 71*(1), 355–364.

Elbers, C., Lanjouw, P., & Leite, P. G. (2008). Brazil within Brazil : Testing the Poverty Map Methodology in Minas Gerais. *World Bank Policy Research Working Paper, 4513*. Washington, DC: The World Bank.

Foster, J., Greer, J., & Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica: Journal of the Econometric Society*, 761-766.

Searle, S., Casella, G., & McCulloch, C. (1992). *Variance Components.* Wiley.

Tarozzi, A., & Deaton, A. (2009). Using census and survey data to estimate poverty and inequality in small areas. *The Review of Economics and Statistics, 91* (4), 773–792.

Tranmer, M., & Steel, D. G. (2001). Ignoring a level in a multilevel model: evidence from UK census data. *Environment and Planning A, 33*(5), 941-948.


**Table 1: Average of ARB, RRMSE, CR, and CIW of Estimates over Small Areas**

| Parameters | Performance Measures | Methods | | | | | |
|---|---|---|---|---|---|---|---|
| | | ELL | Adj. 1 | Adj. 2 | Adj. 3 | Adj. 4 | Adj. 5 |
| Mean: Population 1 | ARB | 0.824 | 0.740 | 0.162 | 0.048 | 0.823 | 2.496 |
| | RRMSE | 0.198 | 0.178 | 0.109 | 0.097 | 0.198 | 0.612 |
| | CR | 58.04 | 66.94 | 93.14 | 93.22 | 58.079 | 99.97 |
| | CIW (True: 0.942) | 0.393 | 0.476 | 0.944 | 0.933 | 0.394 | 1.755 |
| FGT0: Population 2 | ARB | 0.842 | 0.941 | 0.727 | 0.735 | 0.842 | 1.995 |
| | RRMSE | 0.041 | 0.037 | 0.016 | 0.015 | 0.041 | 0.100 |
| | CR | 55.16 | 64.66 | 91.10 | 91.19 | 55.24 | 99.91 |
| | CIW (True: 0.192) | 0.076 | 0.093 | 0.172 | 0.170 | 0.076 | 0.330 |
| FGT1: Population 2 | ARB | 0.844 | 0.755 | 0.190 | 0.108 | 0.844 | 2.037 |
| | RRMSE | 0.020 | 0.018 | 0.011 | 0.010 | 0.020 | 0.050 |
| | CR | 55.01 | 65.73 | 94.54 | 94.62 | 55.1 | 99.83 |
| | CIW (True: 0.094) | 0.037 | 0.046 | 0.098 | 0.097 | 0.037 | 0.163 |
| FGT2: Population 2 | ARB | 0.840 | 0.743 | 0.472 | 0.428 | 0.840 | 2.058 |
| | RRMSE*100 | 1.216 | 1.078 | 1.160 | 1.075 | 1.215 | 3.063 |
| | CR | 55.79 | 67.17 | 96.21 | 96.26 | 55.91 | 99.77 |
| | CIW (True: 0.057) | 0.023 | 0.028 | 0.067 | 0.066 | 0.023 | 0.099 |

**Figure 1: Variance of Area Specific Mean (Population 1), FGT0, FGT1, FGT2 (Population 2)**