The source of bias in unit-context models*

Paul Corral Rodas†

December, 2024

Abstract

This paper explores why poverty estimates from unit-context models exhibit systematic bias. The analysis demonstrates that this bias stems from these models' inability to fully capture between-household variation in welfare, as they rely solely on covariates aggregated at geographic levels. Through model-based simulations, we show that the bias in unit-context models is minimized when the share of total variance explained by the model matches the true variance of welfare at the area level.

Key words: Small area estimation; poverty mapping; satellite imagery

JEL classification: C13; C55; C87; C15

*The author acknowledges financial support from the World Bank.

†The World Bank Group - Poverty and Equity Global Practice (pcorralrodas@worldbank.org)

1 Introduction

Household surveys aimed at gauging a population's living standards often lack representativeness beyond broad regions or specific population demographics. Additionally, there is a risk that many pertinent locations or groups may be omitted from these surveys. However, detailed information on poverty is crucial for effectively targeting resources to alleviate it.

The demand for disaggregated statistics necessitates the use of indirect techniques, which integrate supplementary data from censuses, registries, or larger-scale surveys to generate sufficiently precise statistics for granular populations. Small area estimation encompasses a wide spectrum of statistical methods aimed at enhancing the precision of estimates in cases where household surveys lack sufficient sample size to achieve the desired level of accuracy. Among these methods, model-based approaches leverage the concept of "borrowing strength" from larger datasets or auxiliary information across different areas through models that establish connections between them (such as regression-type techniques), resulting in indirect estimators (Molina and Rao 2010). Most model-based techniques can be classified into two main groups: those based on unit-level models and those based on area-level models. The former are typically employed when data on individual units (e.g., households) are accessible, while the latter are utilized in situations where only aggregate data at the area level (e.g., area means) are available, as outlined by Fay and Herriot (1979).

Unit-level models are generally unsuitable when survey and census data correspond to different years — an issue common in developing countries where censuses and surveys are conducted infrequently. In contrast, area-level models offer a viable alternative. These models rely on linear functional forms but perform estimation and prediction using only aggregate data for the geographic entities of interest (Fay and Herriot (1979); Torabi and Rao (2014)). Another approach, unit-context models, features an estimation stage where household-level measures are modeled solely as a linear function of area-level characteristics (Nguyen (2012); Lange et al. (2018); Masaki et al. (2020)).

Unit-context models, however, have been discouraged due to the method resulting in biased estimates. Corral et al. (2022) show that the method is unable to fully replicate the welfare distribution which results in biased estimates. For a practical example of the resulting bias, see Edochie et al. (2024). In that work the authors present the aggregated estimates at the geographic level of representativeness and it is evident that across many areas the differences to direct estimates are considerable (see Figures 7 to 9). Moreover, it can be seen that the majority of the estimates fall above the 45 degree line, suggesting an upward bias in the estimates.

A key question is what is the source of the bias in unit-context models. In this note, the simulations implemented by Corral et al. (2022) are used to study the source of the method's bias. The bias studied here, goes beyond the potential bias of the method noted in Corral et al. (2021) which was related to a sampling issue. It is related to the transformation bias noted by Würz et al. (2022). These authors attempt to address the same issue as Nguyen (2012) of not having auxiliary microdata and relying on aggregate population-level auxiliary information. Würz et al. (2022)

note that relying on aggregate level data to model household-level welfare leads to first-order bias from back-transformation and second-order bias from using aggregate data. They note that using aggregate means instead of individual values introduces additional bias due to the convexity of the back-transformation function.

The bias studied here is related to the poor model explanatory power of unit-context models. Because unit-context methods model transformed household-level welfare as a linear function of area-level covariates these are unable to adequately account for between household variations in welfare. The simulations undertaken here show, that across the welfare distribution, areas where bias is lowest for these models are those where the ratio of total variance explained by the predicted welfare is closest to 1.

The note proceeds to present the assumed model for unit-level small area estimation and provides a discussion on why not accounting for variance leads to biased estimates of poverty. The method for creating simulated data is illustrated, followed by the results. Finally, conclusions are presented.

2 Small area estimation

The model based small area estimation methods described in this chapter, are dependent on an assumed model. The nested error model used for small area was originally proposed by Battese, Harter and Fuller (1988) to produce county-level corn and soybean crop area estimates for the American state of Iowa. For the estimation of poverty and welfare, Molina and Rao (2010) and Elbers et al. (2003) assume that the transformed welfare y_{ah} for each household h within each location a in the population is linearly related to a $1 \times K$ vector of characteristics (or correlates) x_{ah} for that household, according to the nested error model:

$$y_{ah} = x_{ah}\beta + \eta_a + e_{ah},\ h = 1, \ldots, N_c,\ a = 1, \ldots, A, \tag{1}$$

where η_a and e_{ah} are respectively location and household-specific idiosyncratic errors, assumed to be independent from each other, following:

$$\eta_a \overset{iid}{\sim} N\left(0, \sigma_\eta^2\right),\ e_{ah} \overset{iid}{\sim} N\left(0, \sigma_e^2\right)$$

where the variances σ_η^2 and σ_e^2 are unknown. Here, A is the number of locations in which the population is divided and N_a is the number of households in location a, for $a = 1, \ldots, A$, and n_a is the sample size from area a. Finally, β is the $K \times 1$ vector of regression coefficients.

One of the main assumptions is that errors are normally distributed. The assumption does not necessarily imply that welfare (the dependent variable) is normally distributed, but instead implies that, conditional on the observed characteristics, the residuals are normally distributed. Under the assumed model, variation in welfare across the population is determined by three components: the variation in household characteristics, the variation in household specific unobservables, and the variation in location specific effects. Within any given area, the welfare distribution is determined by the variation in household specific characteristics and household specific errors.

To obtain estimates, the first step is to fit the model 1 via any method providing consistent estimators. This yields the vector of parameter estimates:

$$\hat{\theta}_0 = \left(\hat{\beta}_0, \hat{\sigma}^2_{\eta 0}, \hat{\sigma}^2_{e0}\right).$$

Usual fitting methods under this approach are maximum likelihood (ML) or restricted maximum likelihood (REML), both based on the normal likelihood, and H3 method, which does not require to specify a distribution. The empirical best (EB) area effects for the model are estimated from:

$$\hat{\eta}_{a0} = \hat{\gamma}_a \left(\bar{y}_a - \bar{x}_a \hat{\beta}_0\right), \quad \hat{\gamma}_a = \frac{\hat{\sigma}^2_{\eta 0}}{\hat{\sigma}^2_{\eta 0} + \hat{\sigma}^2_{e0}/n_a}.$$

The variance for the area effects are given by:

$$\hat{\sigma}_{\eta_a} = \sigma^2_{\eta 0}(1 - \hat{\gamma}_a)$$

Because we assume normally distributed errors, the probability of being poor for any household h in area a is entirely dependent on its expected welfare, $x_h \hat{\beta}$, and its error, $\hat{e}_h$, and $\hat{\eta}_a$ which are assumed to follow $e_h \sim N(0, \sigma^2_e)$ and $\eta_a \sim N\left(\hat{\eta}_{a0}, \hat{\sigma}^2_{\eta 0}(1 - \hat{\gamma}_a)\right)$.

$$Prob\, poor_{ha} = \Phi \left(\frac{z - x_{ha}\hat{\beta} - \hat{\eta}_{a0}}{\sqrt{\left(\hat{\sigma}^2_{\eta_a} + \hat{\sigma}^2_e\right)}} \right) \tag{2}$$

where z is the transformed poverty threshold. The poverty rate for the area is given by the average probability of being poor across households. Within a given area, the only thing that varies across households are the covariates.

Unit-context models are an approximation to the assumed underlying data generating process from Eq. 1. Originally introduced by Nguyen (2012), and then re-introduced by Lange et al. (2018) and modified by Masaki et al. (2020). Unit-context versions (i.e. those with aggregated covariates only) are defined as models where household-level welfare is modeled using only area and sub-area level characteristics. Masaki et al. (2020) suggest that the model should include characteristics that explain variability at a geographic level below the one for which we aim to estimate poverty. A possible unit-context model follows:

$$y_{sach} = z_{sac}\alpha + t_{sc}\omega + g_s\lambda + \eta_{sc} + \varepsilon_{sach}$$

where s is used for an aggregation level that is over the target areas (a super-area) and c is used for subareas, e.g., clusters. Hence, z_{sac} contains subarea-level characteristics, t_{sa} includes area-level characteristics and g_s is composed of super-area-level characteristics (which may include super-area fixed effects). The regression coefficients across these levels are respectively denoted α, ω and λ. The random effects, η_{sa}, are specified in this model at the area level, same as in Eq. 1. Note that, among the set of covariates in this model, none is at the unit-level; covariates only vary at the subarea-level and above.

A key feature of unit-context models is that the linear fit only explains a relatively small amount of the total variance of the dependent variable, with R^2 that range between 0.15 to 0.25 in most instances. The variance explained is, understandably, less that the household-level model. The variance that is not explained due to the exclusion of household-level covariates is captured in σ^2_η and σ^2_e. At the national level this leads to the variance of the linear fit, plus σ^2_η and σ^2_e will be approximately equal to $\mathrm{var}[y]$.This is true for unit-context and unit-level models.

At the area level, however, there will be differences and these will lead to bias. To see this, assume welfare for a given area follows a log-normal distribution. Also assume that the linear fit of Eq. 1 is normally distributed. Then the poverty rate for a given area under these assumptions is:

$$FGT_{0a} = \Phi\left(\frac{\ln z - \bar{x}\hat{\beta} - \hat{\eta}_a}{\sqrt{\mathrm{var}\left[x_h\hat{\beta}\right] + \hat{\sigma}^2_{\eta a} + \hat{\sigma}^2_e}}\right)$$

Because the use of EB ensures that for every sampled area a, $E\left[x\hat{\beta} + \hat{\eta}_a\right] = E\left[y_a\right]$,[1] then at the area level bias in poverty arises because:

$$\mathrm{var}\left[x_h\hat{\beta}\right] + \hat{\sigma}^2_{\eta a} + \hat{\sigma}^2_e \neq \mathrm{var}\left[y_a\right]. \tag{3}$$

And because the variance of the linear fit of unit-context models is smaller, unit-context models are more likely to yield biased estimates of poverty.

In the following section we present a model-based simulation to illustrate how Eq. 3 is related to biased estimates at the area level.

3 Simulation data

Data is generated for the simulations following the assumed model from Eq. 1.[2] The population size for the simulated data is $N = 500,000$, and the observations are allocated among $A = 100$

[1] Recall that y_{hc} is the transformed welfare.

[2] Data is simulated following the same approach as Corral et al. (2022). The write-up here is also borrowed from Corral et al. (2022).

areas ($a = 1, \ldots, A$). Within each area a, observations are uniformly allocated over $c = 20$ clusters ($c_a = 1, \ldots, C_a$). Each cluster c consists of $N_{ac} = 250$ observations. In this simulation experiment, a simple random sample of $n_{ac} = 10$ households per cluster is taken, and this sample is kept fixed across simulations. Using a sample, it is possible to compare with estimators based on FH model (see Corral et al. (2022); Molina et al. (2022); Rao and Molina (2015); Fay and Herriot (1979)). The model that generates the population data contains both cluster and area effects. Cluster effects are simulated as $\eta_{ac} \overset{iid}{\sim} N\left(0, 0.1\right)$, area effects as $\eta_a \overset{iid}{\sim} N\left(0, 0.15^2\right)$ and household specific residuals as $e_{ach} \overset{iid}{\sim} N\left(0, 0.5^2\right)$, where $h = 1, \ldots, N_{ac}$, $c = 1, \ldots, C_a$, $a = 1, \ldots, A$.

1. x_1 is a binary variable, taking value 1 when a random uniform number between 0 and 1, at the household-level, is less than or equal to $0.3 + 0.5\frac{a}{40} + 0.2\frac{c}{10}$.
2. x_2 is a binary variable, taking value 1 when a random uniform number between 0 and 1, at the household-level, is less than or equal to 0.2.
3. x_3 is a binary variable, taking value 1 when a random uniform number between 0 and 1, at the household-level, is less than or equal to $0.1 + 0.2\frac{a}{40}$.
4. x_4 is a binary variable, taking value 1 when a random uniform number between 0 and 1, at the household-level, is less than or equal to $0.5 + 0.3\frac{a}{40} + 0.1\frac{c}{10}$
5. x_5 is a discrete variable, simulated as the rounded integer value of the maximum between 1 and a random Poisson variable with mean $\lambda = 3\left(1 - 0.1\frac{a}{40}\right)$.
6. x_6 is a binary variable, taking value 1 when a random uniform value between 0 and 1 is less than or equal to 0.4. Note that the values of x_6 are not related to the area's label.
7. x_7 is generated from a random Poisson variable with mean $\lambda = 3\left(\frac{c}{20} - \frac{a}{100} + u\right)$, where u is a random uniform value between 0 and 1.

In this experiment, we take a grid of 99 poverty thresholds, corresponding to the 99 percentiles of the very first population generated. In total, 1,000 populations are generated. In each of the 1,000 populations, the following quantities are computed in every area for each of the 99 poverty lines:

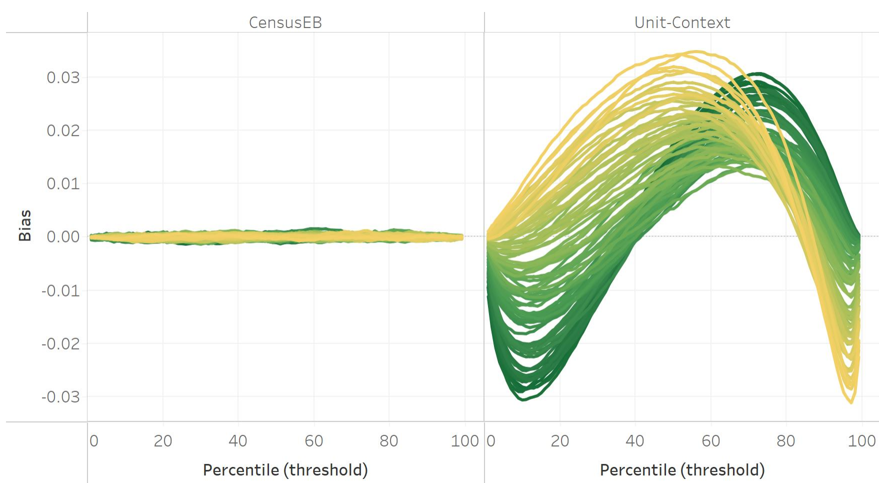
1. True poverty indicators τ_a, using the "census".
2. CensusEB estimators $\hat{\tau}_a^{CEB_a}$ presented in Corral et al. (2021), based on a nested-error model with only **area** random effects and including the unit-level values of the covariates. The R^2 for this model is roughly 0.60.
3. Unit-context CensusEB estimators $\hat{\tau}_a^{UC-CEB_a}$ based on a nested-error model with random effects at the **area-level**. This estimator follows the approach of Masaki et al. (2020). The R^2 of the resulting model hovers around 0.17.

The average difference between the true poverty indicator and the estimate across the 1,000 simulations represent the empirical bias for each area. The Stata script to replicate these simulations can be found in the appendix.

4 Results

Because poverty is predicted across the 99 thresholds noted in the previous section, it is possible to plot how bias across all lines and areas is present. The simulations presented here, train the model and predict using the same data. This is done in order to remove other potential sources of bias from unit-context models, such as the potential omitted variable bias noted by Corral et al. (2021). As can be seen in Figure 1, the bias of unit-context models is present across all lines, but as noted by Corral et al. (2022) the bias is lower for some areas and for some percentiles.

Figure 1: Empirical bias of poverty for CensusEB and Unit-Context small area estimation



Note: Simulation based on 1,000 populations generated as described in section 3. Each line corresponds to one of the 100 areas. The x-axis represents the percentile on which the poverty line falls on, and the y-axis is the empirical bias.

As argued in section 2, the bias arises because at the area level, the unit-context model does a poor job at capturing the entire distribution's variance. Table 1 presents the variance of the linear fit across areas. The true variance ranges from 0.37 to 0.86, a similar range to the CensusEB model. However, the range for the unit-context model is much smaller, from 0.07 to 0.1. Nevertheless, on average the total variance of the unit-context model matches that of the true welfare and that of the CensusEB estimates (Table 2). This is because the model is fit at the national level and σ_η^2 and σ_η^2 are estimated to be much larger since the covariates capture so little of the total variance. Consequently, the range for the total variance of the areas is minimal. Thus, in the unit-context models the larger estimates for σ_η^2 and σ_η^2 only works at the national level, across areas this leads to some areas where the total variance considerably larger than the true variance and some where the variance is considerably smaller than the real variance.

Table 1: Statistics for the area level variance of the model's linear fit

	True var $[x\beta]$	UC var $[\hat{y}_a]$	CensusEB var $[\hat{y}_a]$
Min	0.374	0.074	0.374
Max	0.861	0.105	0.860
Mean	0.568	0.088	0.568
p25	0.448	0.084	0.448
p50	0.549	0.088	0.548
p75	0.682	0.091	0.682

Note: Simulations based on 1,000 populations generated as described in section 3 and averaged to the area level.

Table 2: Statistics for the area level total variance of the model

	True	Unit-context model		CensusEB	
	var $[y_a]$	var $[\hat{y}_a + \eta + e]$	Ratio to true	var $[\hat{y}_a + \eta + e]$	Ratio to true
Min	0.630	0.812	0.744	0.632	0.990
Max	1.119	0.844	1.298	1.118	1.008
Mean	0.826	0.826	1.026	0.826	1.000
p25	0.710	0.822	0.879	0.706	0.998
p50	0.804	0.826	1.025	0.807	1.000
p75	0.939	0.829	1.166	0.940	1.003

Note: Simulations based on 1,000 populations generated as described in section 3 and averaged to the area level.

A poor fitting model in unit-level small area estimation thus leads to biased estimates. Figure 2 illustrates how this is manifested across select poverty lines. The absolute bias for a given area decrease as the ratio of model explained variance to true variance for a given area gets close to 1. Therefore, even if the use of empirical best methods guarantees that the mean of the dependent variable of the model at the area level is unbiased, because poverty is distribution dependent the inability of unit-context models to approximate the full distribution at the area level leads to biased poverty estimates.

Figure 3 averages across all 99 poverty lines the bias from the model for each area. This figure makes it more salient how bias is at its minimal point around the point where the unit-context model's total variance is equal to the true dependent variable's variance in that area. Hence, estimates are biased for some areas more than for others.

5 Conclusions

This paper demonstrates that the bias in unit-context models for small area poverty estimation stems primarily from their inability to adequately capture the full variance of welfare at the area level. While these models may achieve unbiased estimates of mean transformed welfare through empirical best prediction methods, they systematically fail to replicate the true welfare distribution within areas, leading to biased poverty estimates.

The simulation results reveal several key insights:

Figure 2: Empirical bias of Unit-Context models across select lines

Note: Simulation based on 1,000 populations generated as described in section 3. Each line corresponds to one of the 100 areas. The x-axis represents the percentile on which the poverty line falls on, and the y-axis is the empirical bias.

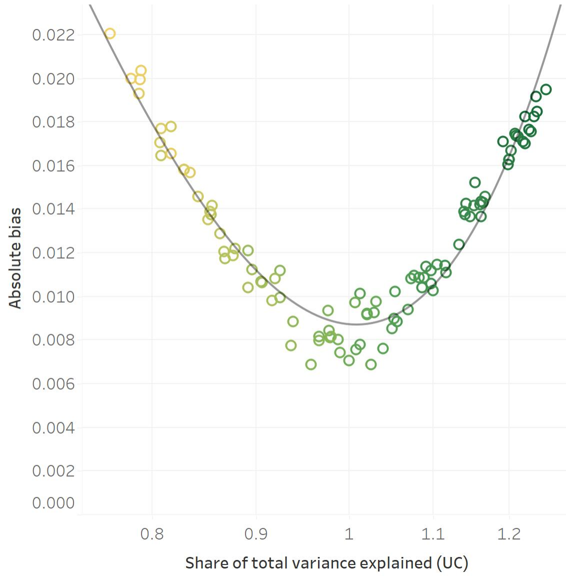
First, unit-context models typically explain only a small portion of the total variance in welfare (R^2 ranging from 0.13 to 0.25) compared to traditional unit-level models (R^2 around 0.50). This limited explanatory power arises because unit-context models rely solely on area-level covariates, omitting household-level variation. Adding more covariates to the unit-context model risks overfitting and could lead to further bias, thus the solution is not aligned to adding more data. An approach similar to the one presented by Würz et al. (2022) could work, but not when using satellite derived data.

Second, while the total variance at the national level matches the true welfare variance through the estimation of area and household error components (σ_η^2 and σ_e^2), this compensation mechanism breaks down at the area level. Some areas end up with significantly over- or under-estimated total variance, leading to systematic bias in poverty estimates.

Third, our analysis reveals that the magnitude of bias in poverty estimates is directly related to how well the model's total variance matches the true welfare variance in each area. Areas where this ratio approaches 1 show minimal bias, while areas with substantial variance mismatches exhibit larger biases.

These findings have important implications for practitioners. While unit-context models offer prac-

Figure 3: Empirical bias of Unit-Context models across select lines



Note: Simulation based on 1,000 populations generated as described in section 3. Each line corresponds to one of the 100 areas. The x-axis represents the percentile on which the poverty line falls on, and the y-axis is the empirical bias.

tical advantages in situations where household-level census data is unavailable or outdated, their inherent limitations in capturing welfare distributions should be carefully considered. Users should be particularly cautious when interpreting poverty estimates for areas where the model's variance differs substantially from the observed welfare variance in survey data. As noted by Corral Rodas et al. (2023), area-level models such as the well known Fay Herriot model will outperform unit-context models, when both use the same data.

Future research might explore methods to improve the variance approximation in unit-context models or develop diagnostic tools to identify areas where these models are most likely to produce reliable estimates. Additionally, investigating alternative approaches that better capture within-area welfare distributions while maintaining the practical advantages of unit-context models could prove valuable.

References

Battese, G. E., Harter, R. M., and Fuller, W. A. (1988). An error-components model for prediction of county crop areas using survey and satellite data. *Journal of the American Statistical Association*, 83(401):28–36.

Corral, P., Molina, I., Cojocaru, A., and Segovia, S. (2022). *Guidelines to small area estimation for poverty mapping*. The World Bank, Washington, DC.

Corral, P., Molina, I., and Nguyen, M. (2021). Pull your small area estimates up by the bootstraps. *Journal of Statistical Computation and Simulation*, 91(16):3304–3357.

Corral Rodas, P. A., Henderson, H. L., and Segovia Juarez, S. C. (2023). Poverty mapping in the age of machine learning. *World Bank Policy Research Working Paper*, (10429).

Edochie, I., Newhouse, D., Tzavidis, N., Schmid, T., Foster, E., Hernandez, A. L., Ouedraogo, A., Sanoh, A., and Savadogo, A. (2024). Small area estimation of poverty in four west african countries by integrating survey and geospatial data. *Journal of Official Statistics*, page 0282423X241284890.

Elbers, C., Lanjouw, J. O., and Lanjouw, P. (2003). Micro-level estimation of poverty and inequality. *Econometrica*, 71(1):355–364.

Fay, R. E. and Herriot, R. A. (1979). Estimates of income for small places: An application of James-Stein procedures to census data. *Journal of the American Statistical Association*, 74(366a):269–277.

Lange, S., Pape, U. J., and Pütz, P. (2018). Small area estimation of poverty under structural change. World Bank Policy Research Working Paper No. 8472.

Masaki, T., Newhouse, D., Silwal, A. R., Bedada, A., and Engstrom, R. (2020). Small area estimation of non-monetary poverty with geospatial data. World Bank Policy Research Working Paper No. 9383.

Molina, I., Corral, P., and Nguyen, M. (2022). Estimation of poverty and inequality in small areas: Review and discussion. *TEST*, pages 1–24.

Molina, I. and Rao, J. (2010). Small area estimation of poverty indicators. *Canadian Journal of Statistics*, 38(3):369–385.

Nguyen, V. C. (2012). A method to update poverty maps. *The Journal of Development Studies*, 48(12):1844–1863.

Rao, J. and Molina, I. (2015). *Small area estimation*. John Wiley & Sons, Hoboken, NJ, 2nd edition.

Torabi, M. and Rao, J. (2014). On small area estimation under a sub-area level model. *Journal of Multivariate Analysis*, 127:36–55.

Würz, N., Schmid, T., and Tzavidis, N. (2022). Estimating regional income indicators under transformations and access to limited population auxiliary information. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 185(4):1679–1706.