

# When Aggregation Misleads: Bias in Unit-Level Small Area Estimates of Poverty with Aggregate Data

Paul Andres Corral Rodas\*

## Abstract

This paper explores why small area poverty estimates from models at the household level that only use aggregate data as covariates, exhibit systematic bias. The analysis demonstrates that this bias stems from the model's inability to capture the complete between-household variation in welfare, as they rely solely on covariates aggregated at geographic levels. Through model-based simulations, the paper shows that the bias in these models is minimized when the empirical variability of simulated welfare based on the model is closest to the true empirical variance of welfare at the area level. This finding also has implications for bias in unit-level models.

**Key words:** Small area estimation; poverty mapping; satellite imagery; census; official statistics

**JEL classification:** C13; C55; C87; C15

---

\*The World Bank Group - Poverty and Equity Global Practice (pcorralrodas@worldbank.org). The author acknowledges financial support from the World Bank. Special thanks to Carlos Rodriguez-Castelan, Alexandru Cojocaru, Tara Vishwanath, and Isabel Molina for comments on an earlier draft. Full replication package for the results presented in this paper may be found in: [https://github.com/pcorralrodas/UC\\_source\\_of\\_bias](https://github.com/pcorralrodas/UC_source_of_bias)

# 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 has increased the reliance on indirect techniques that integrate supplementary data from censuses, registries, or larger-scale surveys. These methods are used to produce sufficiently precise statistics for granular populations. Small area estimation encompasses a broad range of statistical techniques designed to enhance the precision of estimates when household surveys lack the sample size required for the desired level of accuracy. Among these techniques, model-based approaches stand out by leveraging the concept of "borrowing strength" from larger datasets or auxiliary information. These methods use models that establish relationships across areas (e.g., regression techniques), enabling the creation of indirect estimators (Molina and Rao 2010).

Most model-based techniques fall into two main categories: unit-level models and area-level models. Unit-level models are generally applied when data on individual units (e.g., households) are available, while area-level models are used when only aggregate data for specific geographic areas (e.g., area means) are accessible, as described by Fay and Herriot (1979). In poverty estimation, unit-level models predict the welfare distribution first and then apply a threshold to determine the proportion of people below that threshold. In contrast, area-level models directly estimate the poverty rate for an area.

Unit-level models face limitations when survey and census data are from different years — a common issue in developing countries where censuses and surveys are conducted infrequently. Area-level models, however, provide a feasible alternative. These models rely on linear functional forms and perform estimation and prediction using only aggregate data for geographic entities of interest (Fay and Herriot (1979); Torabi and Rao (2014)). Another approach, unit-context models, employs an estimation stage where household-level measures are modeled as a linear function of area-level characteristics (Nguyen (2012); Lange et al. (2018); Masaki et al. (2020)). Unit-context models, like unit-level models predict the welfare distribution first and then apply a threshold to determine the proportion of people below that threshold, but do so only using area-level characteristics.

Although alluring, unit-context models, 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 compared to direct estimates are considerable (see Figures 7 to 9 in Edochie et al. (2024)). Moreover, it can be seen that the majority of the estimates fall above the 45 degree line, suggesting an

upward bias in the estimates. Using real world data, where samples are taken following methods implemented in developing countries, Corral Rodas et al. (2023) present evidence on how unit-context models produce biased estimates of poverty and that the noise and bias of unit-context models is considerably larger than that produced by area-level models.

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. Instead, the bias noted in this paper is related to the transformation bias noted by Würz et al. (2022). These authors attempt to address the same issue as Nguyen (2012), Lange et al. (2018) and Masaki et al. (2020) 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 as covariates 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 and is related to the bias noted by Würz et al. (2022). Because unit-context methods model transformed household-level welfare as a linear function of area-level covariates only, these are unable to adequately account for between-household variations in welfare. The simulations undertaken here show that areas where the variance of predicted welfare is aligned to the area’s true variance of welfare are those that exhibit the lowest bias across the entire welfare distribution. This illustrates that under unit-level and unit-context models, bias at the area level is a function of the dependent variable’s mean and the empirical variance of the linear fit. The dependent variable benefits from the use of EB methods, but the empirical variance is fully dependent on the auxiliary data at hand and how aligned it is to the population in each area.

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 paper are dependent on an assumed model. The nested error model used for small area estimation was originally proposed by Battese et al. (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:<sup>1</sup>

$$y_{ah} = x_{ah}\beta + \eta_a + e_{ah}, \quad h = 1, \dots, N_a, \quad a = 1, \dots, A, \quad (1)$$

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

$$\eta_a \stackrel{iid}{\sim} N(0, \sigma_\eta^2), \quad e_{ah} \stackrel{iid}{\sim} N(0, \sigma_e^2)$$

where the variances  $\sigma_\eta^2$  and  $\sigma_e^2$  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  $a = 1, \dots, A$ , and  $n_a$  is the sample size from area  $a$ . Finally,  $\beta$  is the  $K \times 1$  vector of regression coefficients.

One of the main assumptions is that errors are normally distributed. The assumption implies that, conditional on the observed characteristics, the model's errors are normally distributed. To obtain estimates, the first step is to fit the model from Eq. 1 to the observed sample data via any method providing consistent estimators. 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 specify a distribution. This yields the vector of parameter estimates:

$$\hat{\theta} = (\hat{\beta}, \hat{\sigma}_\eta^2, \hat{\sigma}_e^2).$$

The empirical best (EB) area effects for the model are estimated from:

$$\hat{\gamma}_a (\bar{y}_a - \bar{x}_a \hat{\beta}), \quad \hat{\gamma}_a = \frac{\hat{\sigma}_\eta^2}{\hat{\sigma}_\eta^2 + \hat{\sigma}_e^2 / n_a}$$

where  $\bar{x}_a$  and  $\bar{y}_a$  are the sample means in area  $a$  for  $x$  and  $y$ , respectively.

The variance for the area effects are given by (Molina and Rao, 2010):

$$\text{var}[\eta_a | \bar{y}_a] = \sigma_\eta^2 (1 - \gamma_a)$$

Making use of the parameters estimated from the model of Eq. 1, it is possible to produce a value of  $y_{ah}$  for every household in the census data:

$$y_{ah}^* = x_{ah}\hat{\beta} + \hat{\gamma}_a (\bar{y}_a - \bar{x}_a \hat{\beta}) + \varepsilon_{ah}^*$$

where  $\varepsilon$  is drawn from  $\varepsilon_{ah}^* \sim N(0, \sigma_\eta^2(1 - \gamma_a) + \sigma_e^2)$ . This Monte Carlo procedure is often repeated 100 times in order to derive indicators under each simulated population and then average across to

---

<sup>1</sup>For simplicity  $y_{ah}$  is considered the transformed welfare. The most common transformation applied is the natural logarithm.

derive the final EB estimate. Notice how in the simulated vectors  $x_{ah}\hat{\beta} + \hat{\gamma}_a(\bar{y}_a - \bar{x}_a\hat{\beta})$  does not vary across simulations, only  $\varepsilon_{ah}^*$  vary across the simulated vectors.

Alternatively, because normally distributed errors are assumed, the probability of being poor for any household  $h$  in area  $a$  is entirely dependent on its expected welfare,  $x_h\hat{\beta}$ , its idiosyncratic error,  $\hat{e}_h$ , and the predicted area effect  $\hat{\gamma}_a(\bar{y}_a - \bar{x}_a\hat{\beta})$  which are assumed to follow  $e_h \sim N(0, \sigma_e^2)$  and  $\eta_a \sim N(\gamma_a(\bar{y}_a - \bar{x}_a\beta), \sigma_\eta^2(1 - \gamma_a))$ , respectively. The empirical best estimated probability of a household being poor is given by:

$$Prob\ poor_{ha} = \Phi \left( \frac{z - \hat{y}_{ah}}{\sqrt{\hat{\sigma}_\eta^2(1 - \hat{\gamma}_a) + \hat{\sigma}_e^2}} \right) \quad (2)$$

where  $z$  is the transformed poverty threshold and  $\hat{y}_{ah} = x_{ah}\hat{\beta} + \hat{\gamma}_a(\bar{y}_a - \bar{x}_a\hat{\beta})$ . The poverty rate for the area is given by the average probability of being poor across households.<sup>2</sup> Within a given area, the only thing that varies across households is the covariates.

Under the assumed model, variation in simulated welfare across the population ( $S_{y_a}^{2*}$ ) in a given area is the sum of two components: 1) the variability of the explained portion,  $S_{y_a}^{2*} = \frac{1}{N_a} \sum_h (\hat{y}_{ah} - \bar{y}_a^*)^2$ , and 2) variability in the simulated errors  $\frac{1}{N_a} \sum_h (\varepsilon_{ah}^*)^2$ :

$$\begin{aligned} S_{y_a}^{2*} &= \frac{1}{N_a} \sum_h (y_{ah}^* - \bar{y}_a^*)^2 \\ S_{y_a}^{2*} &= \frac{1}{N_a} \sum_h (\hat{y}_{ah} + \varepsilon_{ah}^* - \bar{y}_a^*)^2 \\ S_{y_a}^{2*} &= \frac{1}{N_a} \sum_h (\hat{y}_{ah}^2 + \varepsilon_{ah}^{*2} + \bar{y}_a^{*2} + 2\hat{y}_{ah}\varepsilon_{ah}^* - 2\hat{y}_{ah}\bar{y}_a^* - 2\bar{y}_a^*\varepsilon_{ah}^*) \\ S_{y_a}^{2*} &= \frac{1}{N_a} \sum_h (\hat{y}_{ah} - \bar{y}_a^*)^2 + \frac{1}{N_a} \sum_h (\varepsilon_{ah}^*)^2 \end{aligned}$$

The bias of poverty estimates at the area level will be determined by differences between the true  $S_{y_a}^2$  and the simulated  $S_{y_a}^{2*}$ . This is simplified if we assume that the dependent variable for a given area follows a log-normal distribution. Then the simulated poverty rate for a given area under a given threshold ( $\ln z$ ) under the assumption is:

$$FGT_{0a}^* = \Phi \left( \frac{\ln z - \bar{y}_a}{\sqrt{S_{y_a}^{2*}}} \right) \quad (3)$$

Consequently, the poverty rate for a given area is a function of  $\bar{y}_a$  and  $S_{y_a}^{2*}$ .

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 models are defined as models where household-level welfare is modeled using only area and sub-area level characteristics. Masaki et al. (2020) suggest

---

<sup>2</sup>Traditionally, this process is approximated via Monte Carlo simulations as noted above.

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:

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

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 that are nested in area  $a$ . 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  $\alpha$ ,  $\omega$  and  $\lambda$ . The random effects,  $\eta_{sa}$ , are specified in this model at the area level, the 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 ( $y$ ), with a coefficient of determination ( $R^2$ ) that is often quite lower than models that include household level covariates, and ranges between 0.15 and 0.25 in most instances. Define  $S_y$  as the standard deviation of the sample's dependent variable, the coefficient of determination,  $R^2$ , is given by:

$$R^2 = 1 - \frac{\hat{\sigma}_\eta^2 + \hat{\sigma}_e^2}{S_y^2}$$

Consequently, because unit-context models have lower explanatory power since they only rely on area-level covariates, if  $\hat{\sigma}_{\eta_{uc}}^2$  is the unit-context model estimate of  $\sigma_\eta^2$ , and  $\hat{\sigma}_{e_{uc}}^2$  is the unit-context model estimate of  $\sigma_e^2$ , then:

$$R_{uc}^2 = 1 - \frac{\hat{\sigma}_{\eta_{uc}}^2 + \hat{\sigma}_{e_{uc}}^2}{S_y^2} < R^2 = 1 - \frac{\hat{\sigma}_\eta^2 + \hat{\sigma}_e^2}{S_y^2}$$

$$\hat{\sigma}_{\eta_{uc}}^2 + \hat{\sigma}_{e_{uc}}^2 > \hat{\sigma}_\eta^2 + \hat{\sigma}_e^2. \quad (4)$$

The lower  $R^2$  of unit-context models implies that the explained portion of welfare is lower than the true portion of explained welfare.

When simulating vectors of the dependent variable at the national level, the empirical standard deviation of the dependent variable is approximated under unit-level and unit-context models. However, under unit-context models the empirical standard deviation at the area level is not properly approximated due to the model misspecification. Consequently, when creating simulated vectors of welfare under unit context models the variation in simulated welfare across the population in an area ( $S_{y_a}^{2*}$ ) will not match that of the true population. From Eq. 4, we know that:

$$\sum_h (\varepsilon_{ah_{uc}}^{*2}) > \sum_h (\varepsilon_{ah}^{*2}).$$

Additionally, given the poor model fit we also know that:

$$\sum_h (\hat{y}_{ah_{uc}} - \bar{y}_{a_{uc}}^*)^2 < \sum_h (\hat{y}_{ah} - \bar{y}_a^*)^2$$

Hence, under unit-context models some areas will have a value of  $S_{y_a}^{2*}$  that is larger than that of the true DGP and some will have a smaller or equal value of  $S_{y_a}^{2*}$ . This result, coupled with Eq. 3, implies that the bias of unit-context models will be larger in areas where  $S_{y_a}^{2*}$  is most different from the true model's  $S_{y_a}^2$ .

The focus here is the model's performance on estimating poverty. However, the poor model fit also affects predictions of welfare. Corral et al. (2022) present evidence that unit-context models will do a good job at predicting the model's dependent variable ( $y$ ), thanks to the use of empirical best predictors. A similar result is presented by Chen et al. (2024) who indicate that under model misspecification, unit-context models will perform just as well as area level models in predicting the model's dependent variable.<sup>3</sup> Nevertheless, this seems to only hold true when estimating the mean of the dependent variable of the model is the goal. It does not hold true when the goal is estimating the original untransformed equivalized income or expenditure. Under unit-context models, bias is introduced when the goal is a measure that may be distributionally sensitive – e.g., poverty or equivalized income or expenditure. Corral et al. (2022) present evidence from simulations on how unit-context models will produce biased predictions of the back transformed dependent variable, i.e. equivalized income or expenditure (see figure 5.3 in Corral et al. (2022)). Würz et al. (2022) also present evidence on how unit-context models will lead to biased estimates of welfare that arise from the back transformation of the dependent variable.<sup>4</sup>

Assuming log linearity, the simulated nested error model ( $y_{ah}^* = x_{ah}\hat{\beta}_0 + \hat{\gamma}_a(\bar{y}_a - \bar{x}_a\hat{\beta}_0) + \varepsilon_{ah}^*$ ) implies that  $\exp(y_{ah}^*) = \exp(\hat{y}_{ah})\exp(\varepsilon_{ah}^*)$ . Since  $\varepsilon_{ah}^* \sim N(0, \sigma_\eta^2(1 - \gamma_a) + \sigma_e^2)$ , then  $E[\exp(\varepsilon_{ah}^*)] = \exp\left(0.5\left(\sigma_\eta^2(1 - \gamma_a) + \sigma_e^2\right)\right)$ . Therefore, at the area level welfare will be equal to:

$$\left(\frac{1}{N_a} \sum_{h=1}^{N_a} \exp(\hat{y}_{ah})\right) \exp\left(0.5\left(\sigma_\eta^2(1 - \gamma_a) + \sigma_e^2\right)\right)$$

Consequently, just like for poverty, the predicted welfare under unit-context models will be more biased the greater the difference between  $S_{y_a}^{2*}$  and the true model's  $S_{y_a}^2$ .

<sup>3</sup>Under unit-level and unit-context models welfare is usually transformed to ensure that errors are normally distributed to conform to the model's assumptions.

<sup>4</sup>Under unit-level models welfare is usually transformed to ensure that errors are normally distributed to conform to the model's assumptions.

In the following section, I present a model-based simulation to illustrate how  $S_{y_a}^{2*}$  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.<sup>5</sup> The population size for the simulated data is  $N = 500,000$ , and the observations are allocated among  $A = 100$  areas ( $a = 1, \dots, A$ ). Within each area  $a$ , observations are uniformly allocated over  $c = 20$  clusters ( $c_a = 1, \dots, C_a$ ). 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 the 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  $\eta_{ac} \stackrel{iid}{\sim} N(0, 0.1)$ , area effects as  $\eta_a \stackrel{iid}{\sim} N(0, 0.15^2)$  and household specific residuals as  $e_{ach} \stackrel{iid}{\sim} N(0, 0.5^2)$ , where  $h = 1, \dots, N_{ac}$ ,  $c = 1, \dots, C_a$ ,  $a = 1, \dots, A$ .

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  $\lambda = 3(1 - 0.1\frac{a}{40})$ .
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  $\lambda = 3(\frac{c}{20} - \frac{a}{100} + u)$ , where  $u$  is a random uniform value between 0 and 1.

For the unit-context model variations implemented here the PSU level mean of each covariate is used as the eligible covariates to fit the model.

---

<sup>5</sup>Data is simulated following the same approach as Corral et al. (2022). The write-up here is also borrowed from Corral et al. (2022).



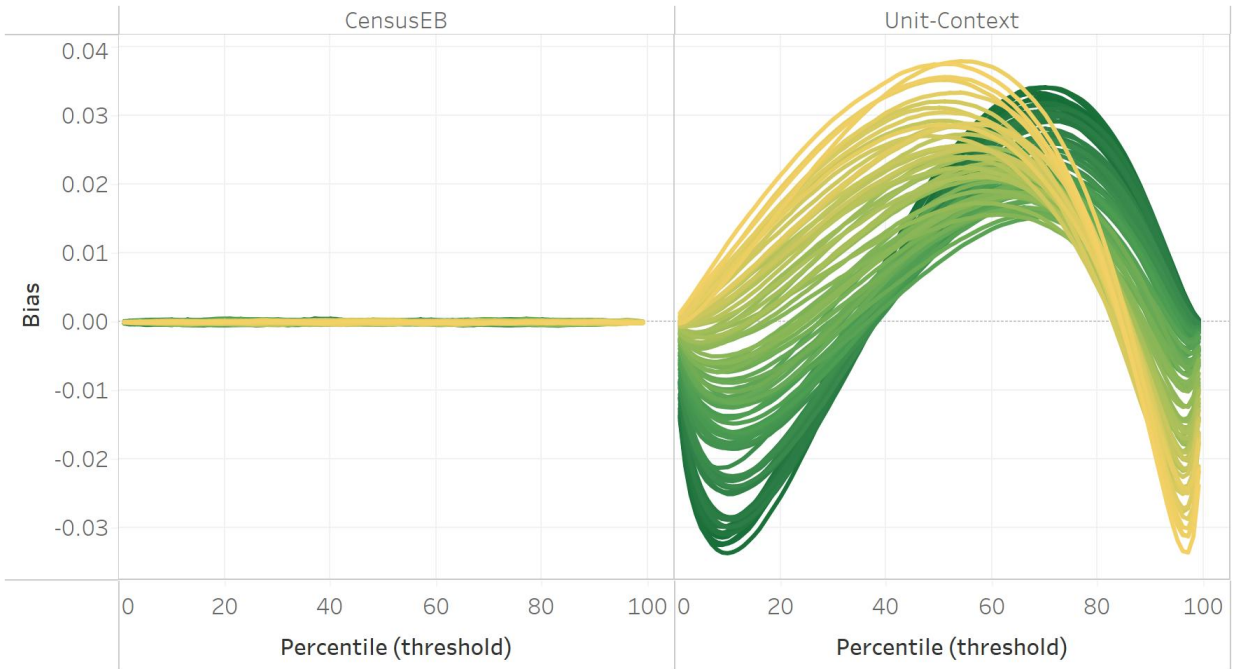
In this experiment, I take a grid of 99 poverty thresholds, corresponding to the 99 percentiles of the very first population generated. In total, 1,000 populations are generated. In each of the 1,000 populations, the following quantities are computed in every area for each of the 99 poverty lines:

1. True poverty indicators  $\tau_a$ , 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 represents the empirical bias for each area.

## 4 Results

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.

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 99 percentiles are considered since the goal

of unit-level small area estimation of poverty is to replicate the full welfare distribution and from it, estimate poverty. The simulations presented here, train the model and predict using the same data, thus sampling does not play a role. This is done in order to remove other potential sources of bias from unit-context models, such as the potential omitted variable bias in unit-context models 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.

Table 1: Statistics for the area level empirical variation ( $S_{\hat{y}}^2$ ) of the model’s linear fit

	True $S_{\hat{y}}^2$	UC $S_{\hat{y}}^2$	EB $S_{\hat{y}}^2$
Min	0.374	0.075	0.374
Max	0.861	0.106	0.861
Mean	0.568	0.088	0.568
p25	0.448	0.084	0.449
p50	0.549	0.088	0.549
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. The true level variation in the explained portion of the model ranges from 0.374 to 0.861, with the average across the 100 areas being 0.568.

Table 2: Statistics for the area level total empirical variance of the model simulated dependent variable

	True	Unit-context model		EB	
	$S_{y_a}^2$	$S_{y_a}^{2*}$	$S_{y_a}^{2*}/S_{y_a}^2$	$S_{y_a}^{2*}$	$S_{y_a}^{2*}/S_{y_a}^2$
Min	0.632	0.813	0.744	0.632	0.998
Max	1.119	0.844	1.295	1.119	1.003
Mean	0.826	0.826	1.026	0.826	1.000
p25	0.707	0.822	0.879	0.707	0.999
p50	0.808	0.826	1.024	0.807	1.000
p75	0.940	0.830	1.166	0.940	1.001

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

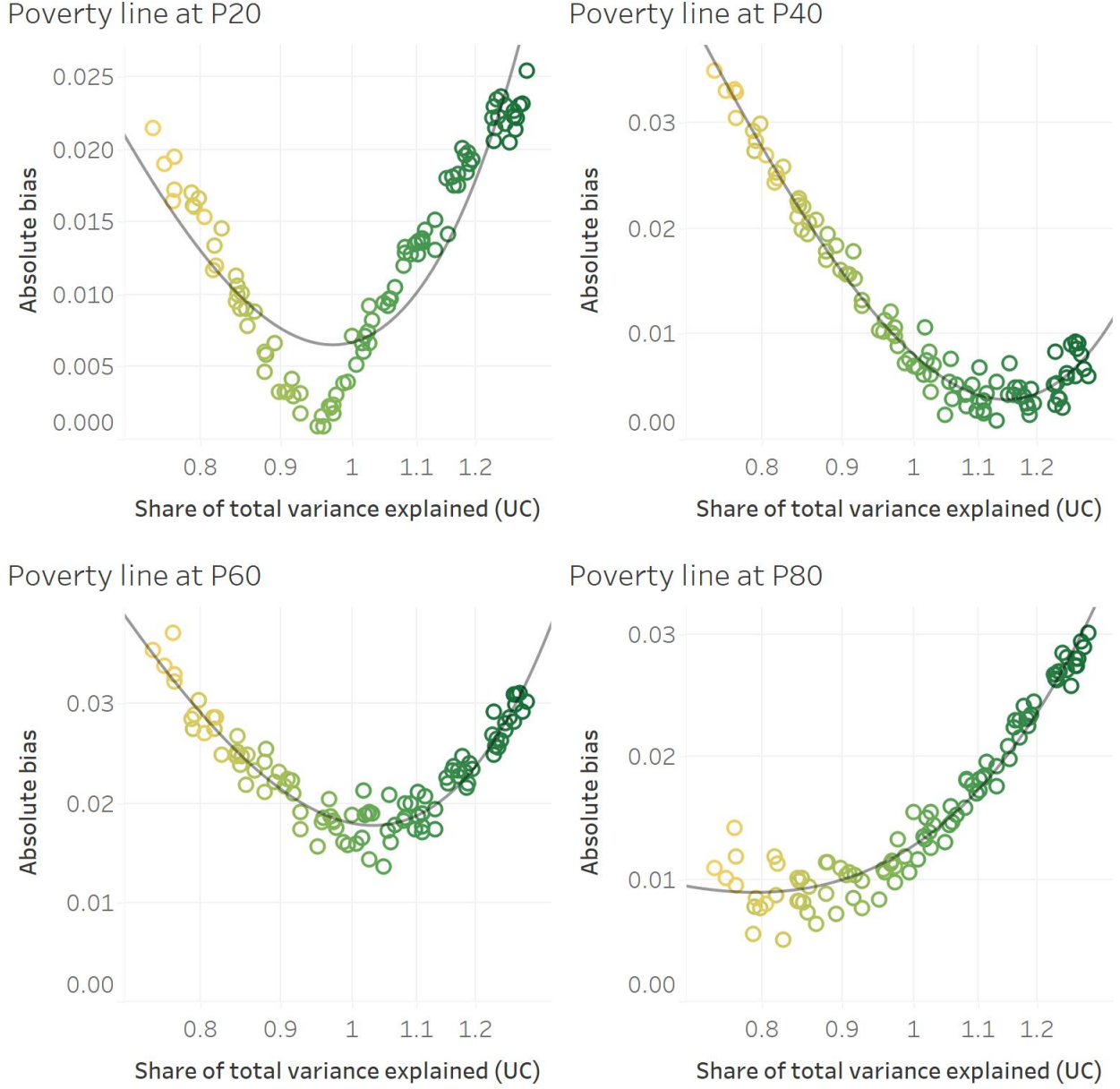
As argued in section 2, the bias arises because at the area level, the unit-context model does a poor job at capturing the empirical variation of the explained portion of the data,  $S_{\hat{y}}^2$ .<sup>6</sup> Table 1 presents the value for  $S_{\hat{y}}^2$  across areas. The true empirical variation of the explained portion,  $S_{\hat{y}}^2$ , ranges from 0.37 for the area with the lowest empirical variation to 0.86 for the area with the largest, a similar range to the EB model. However, the range for the unit-context model is much smaller, from 0.07 to 0.1. Nevertheless, on average, the total empirical variation of the dependent variable,  $S_y^2$ , is matched by unit-context models (Table 2; mean). This is because the model is fit at the national level and  $\sigma_\eta^2$  and  $\sigma_\eta^2$  are estimated to be much larger since the covariates capture so little of the total empirical variance of the dependent variable,  $S_y^2$ . Consequently, the range for the total

<sup>6</sup>At the area level this is equal to:  $\frac{1}{N_a} \sum_h (\hat{y}_{ah} - \bar{y}_a^*)^2$

empirical variance of the areas is minimal, from 0.813 to 0.844, compared to the truth of 0.632 to 1.119. Thus, in the unit-context models the larger estimates for  $\sigma_\eta^2$  and  $\sigma_\eta^2$  only works at the national level, across areas this leads to some areas where  $S_{y_a}^{2*}$  is considerably larger than the true  $S_{y_a}^2$  and some where  $S_{y_a}^{2*}$  is considerably smaller than the true  $S_{y_a}^2$ .

Unit-context models are biased because these deviate from the truth by assuming welfare is not dependent on unit level characteristics, which leads to poor fitting models. Figure 2 illustrates how this is manifested across selected 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 2: Empirical bias of Unit-Context models across select lines

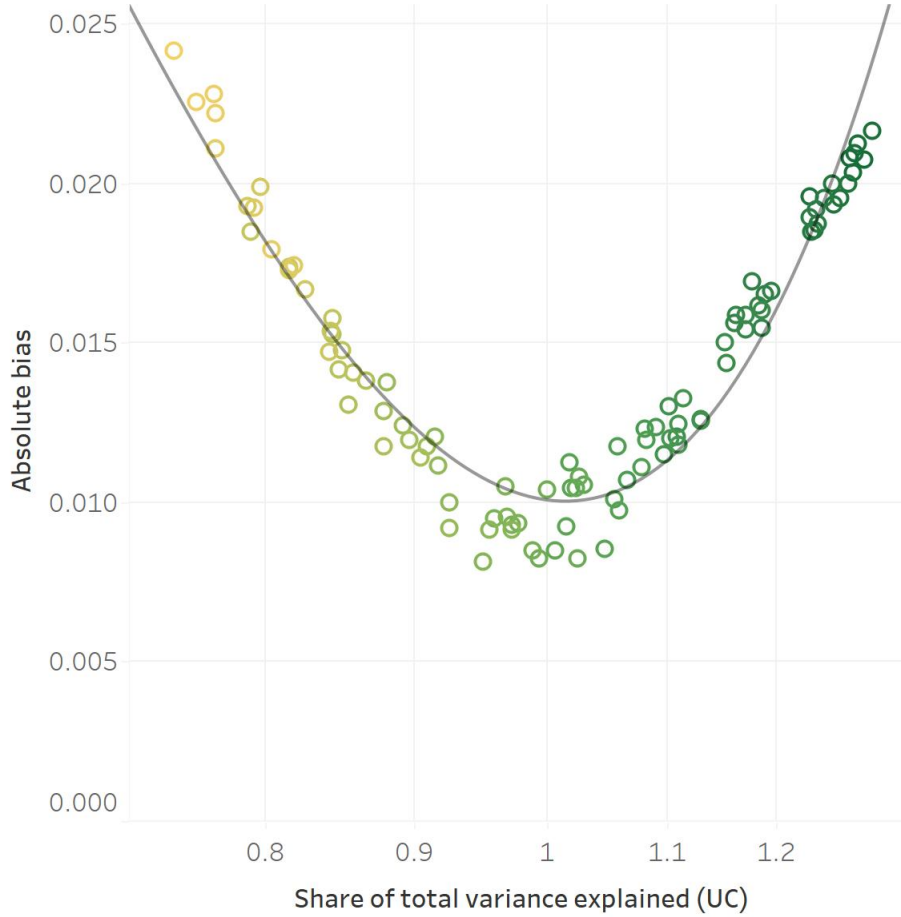


Note: Simulation based on 1,000 populations generated as described in section 3. Each dot corresponds to one of the 100 areas. The x-axis represents the  $S_{y_a}^{2*}/S_{y_a}^2$  for the UC model, and the y-axis is the absolute empirical bias. The gray line represents the quadratic fit plot.

Figure 3 averages across all 99 poverty lines the absolute bias from the model for each area.<sup>7</sup> This figure makes it more salient how bias is at its minimal point around the point where the unit-context model's  $S_{y_a}^{2*}/S_{y_a}^2$  is closer to 1 in that area. Hence, estimates are biased for some areas more than for others.

<sup>7</sup>For area  $a$ , the average absolute bias across all lines is given by:  $\frac{1}{99} \sum_p^{P=99} |\hat{\tau}_{pa} - \tau_{pa}|$ , where  $\tau$  is the head count poverty rate.

Figure 3: Empirical absolute bias of Unit-Context models across areas



Note: Simulation based on 1,000 populations generated as described in section 3. Each dot corresponds to one of the 100 areas. The x-axis represents  $S_{y_a}^{2*}/S_{y_a}^2$  for the unit-context model, and the y-axis is the absolute empirical bias across all 99 percentiles. The gray line represents the quadratic fit plot.

A similar issue is observed for welfare ( $\exp(y_{ah}^*)$ ) under unit-context models. The absolute bias decreases as  $S_{y_a}^{2*}/S_{y_a}^2$  is closer to 1 (Fig. 4, left). Considering that the true area values of untransformed welfare range from roughly 18 to 110, the bias for some areas is quite considerable. In accordance to what was noted in section 2, the method yields unbiased values of the model's dependent variable,  $\bar{y}_a$ , i.e. the transformed welfare (Fig. 4, right). This is a similar finding to that of Chen et al. (2024) who through simulations illustrate that, in the presence of substantial model misspecifications, the unit context model shows similar performance to that of area-level models with known variances. Nevertheless, Chen et al. (2024) do not provide results for the back-transformed variable. Considering that welfare usually requires transformation so that model assumptions are met, the author's results hold little value for international welfare and poverty monitoring.

Figure 4: Empirical absolute bias of Unit-Context model’s predicted welfare,  $\overline{\exp(y_a)}$  and  $\bar{y}_a$  across areas



Note: Simulation based on 1,000 populations generated as described in section 3. Each dot corresponds to one of the 100 areas. The x-axis represents  $S_{y_a}^{2*}/S_{y_a}^2$  for the unit-context model, and the y-axis is the absolute empirical bias of the mean welfare for the area. The gray line represents the quadratic fit plot.

## 5 Conclusions

This paper provides evidence 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 (i.e. the model’s dependent variable) through empirical best prediction methods, they systematically fail to replicate the true welfare distribution within areas, leading to biased poverty estimates and welfare estimates.

The simulation results reveal several key insights:

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 may require having access to unit-level data and thus is not feasible when using satellite derived data.

Second, while unit-context models may be able to match at the national level the total empirical variation of welfare through the estimation of area and household error components ( $\sigma_\eta^2$  and  $\sigma_e^2$ ),

this compensation mechanism breaks down at the area level. Some areas end up with significantly over- or under-estimated variation of welfare, 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 simulated total empirical variation of welfare matches the true welfare’s variability in each area. Areas where this ratio approaches 1 show minimal bias, while areas with substantial mismatches exhibit larger biases.

These findings have important implications for practitioners. While unit-context models offer practical 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 simulated variability of welfare differs substantially from the observed welfare’s empirical variance in survey data. As noted by Corral Rodas et al. (2023) and Corral et al. (2022), using simulated and real world data, area-level models such as the well-known Fay Herriot model will outperform unit-context models.

Future research might explore methods to improve the empirical 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.
- Chen, Y., Lahiri, P., and Salvati, N. (2024). Effects of model misspecification on small area estimators. *arXiv preprint arXiv:2403.11276*.
- 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.