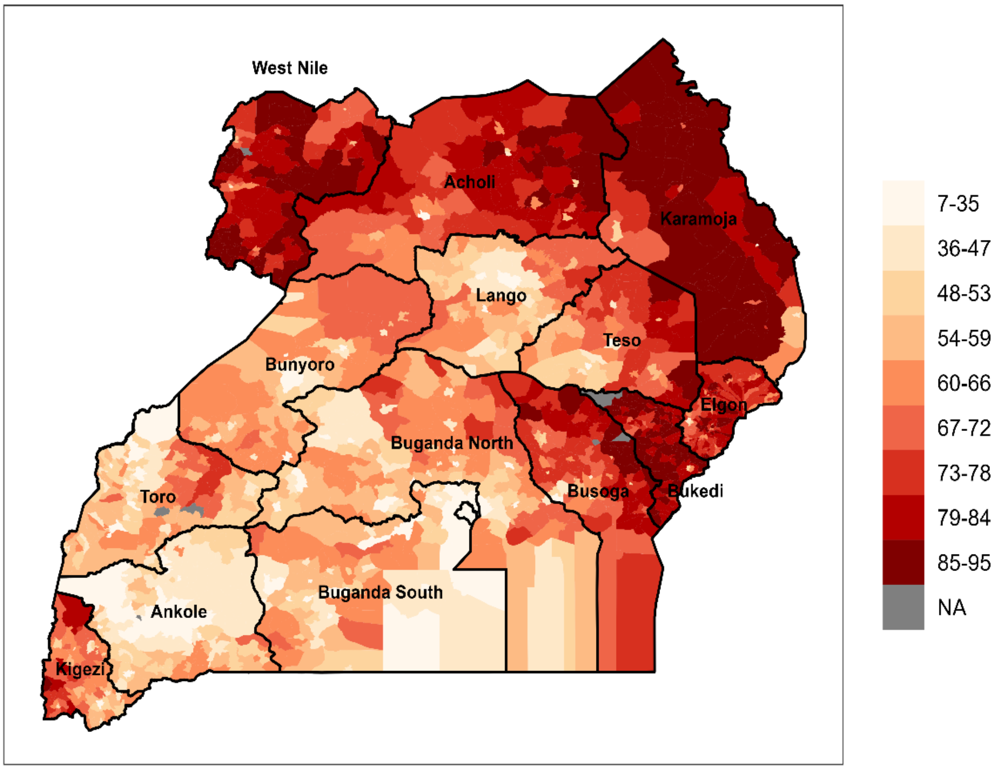
THE SPATIAL DISTRIBUTION OF MULTIDIMENSIONAL POVERTY IN UGANDA IN 2014

Small area estimates of multidimensional poverty in Uganda using the 2014 Census.

Marco Pomati and Héctor E. Nájera Catalán

April 2023

Estimated percentage of children in multidimensional poverty in 2014



ISBN 978-1-908469-35-9

EXECUTIVE SUMMARY

This report presents the first estimates of Consensual Approach multidimensional adult and child poverty in Uganda, using adult and child-specific measures of multidimensional poverty. It shows that adult and child poverty in Uganda have a clear geographical distribution and concentration, with high multidimensional poverty rates in the north.

However, the subcounty-level poverty estimates also suggest that poverty is very high in the north east, the south west and some areas in the north west. Moreover, although adults and children in Kampala experience lower rates, the distribution of poverty within Kampala is not homogeneous, and there are parishes with child poverty rates of up to three times higher than the average. Finally, we show that many of these small area spatial patterns match the distribution of public services, such as distance to public health facilities and primary schools.
ho
This report showcases best practice from the literature while keeping to a minimum technical language. It aims to be a comprehensive and accessible analysis of the spatial patterning and concentration of multidimensional poverty in Uganda and also provide a user-friendly and step-by-step framework for future SAE analyses in Uganda.

Acknowledgements

The analysis in this report was conducted by Dr Marco Pomati (Cardiff University) and Dr Héctor E. Nájera Catalán (Universidad Nacional Autónoma de México), under the guidance of and in close collaboration with Professor Dave Gordon (Bristol Poverty Institute), UNICEF Uganda and the Uganda Bureau of Statistics (UBOS).

The core task team in UNICEF consisted of Diego Angemi (Chief Social Policy and Advocacy) and Arthur Muteesasira (Information Management and GIS Officer) and Sarah Kabaija (Monitoring Specialist). The core task team at UBOS consisted of Mr Vincent Ssennono, Mrs Beatrice Winnie Nyemera and Mr Stephen Baryahirwa.

Suggested citation: Pomati, M. and Nájera Catalán, H.E. (2023), 'The spatial distribution of multidimensional poverty in Uganda in 2014. Small area estimates of multidimensional poverty in Uganda using the 2014 Census', Cardiff, Cardiff University.

ISBN 978-1-908469-35-9

Introduction

In 2016/17, the Uganda National Household Survey (UNHS) included a consensual deprivation question module designed to produce estimates of multidimensional adult and child poverty. Relevant data on material and social deprivation, as well as household expenditure and income, were collected. The design of effective and efficient anti-poverty policies can be aided by estimates of multidimensional poverty for small areas, which enable resources to be targeted at the areas with the greatest needs. However, although the UNHS is a robust and comprehensive survey its sample size of about 17,500 households means that it cannot be used to reliably measure poverty for areas smaller than sub-regions, such as districts, sub-counties and parishes.

To produce reliable small area estimates of the distribution of multidimensional poverty in Uganda it is necessary to combine the UNHS estimates with the national Census 2014 data. Small area poverty estimation (SAE) is a field in social statistics that provides a series of strategies and methods to estimate poverty rates for small areas by combining different data sources, particularly survey and census data (Rao and Molina, 2015). Drawing upon the SAE literature, this report includes the first small areas estimates of multidimensional adult and child poverty (UNICEF, 2019) for district and sub-district areas, obtained using the 2016/17 UNHS and 2014 Uganda Census data and recent advances in Small Area Estimation methodology (Rao and Molina, 2015; Pratesi, 2016).

The report is organized as follows. Section two explains the methodology used to produce the small area estimates. Section three presents the distribution of multidimensional poverty at the district and sub-county levels, as well as at the parish level for Kampala.

2. Methodology

Small Area Estimation (SAE) is a methodology used to tackle a common problem faced by policy-makers: targeted anti-poverty policies may require estimates of living standards for very small spatial areas. However, a typical Household Income and Expenditure Survey (HIES) does not have a sufficiently large enough sample size to produce statistically reliable estimates of poverty at small area level (Rao and Molina, 2015; Pratesi, 2016). By contrast, the national Census produces highly reliable estimates for small spatial areas, but the information collected in the Census is limited and so Census data on their own cannot often be used to produce statistically reliable estimates of poverty. By contrast, the UNHS contains a wide range of information from respondents, including information that can be used to create robust estimates of multidimensional poverty. However, the UNHS lacks the sample size to create reliable district and sub-district estimates of poverty, while the Census can be used to produce estimates for very small areas and collects some information that correlates with multidimensional poverty (e.g. occupational status, housing conditions, etc.), but not sufficient information to produce reliable and valid estimates of multidimensional poverty.

Contemporary Small Area Estimation methods are designed to combine the strengths of surveys and population censuses (Rao and Molina, 2015; Pratesi, 2016). They do this by exploiting the availability of common information in both national surveys and Census. This common information is used to create a statistical model to predict a variable of interest, such as multidimensional poverty in the national survey. After this statistical model has been tested and validated, it is then applied on the Census data to produce small area estimates of the variable of interest (e.g. multidimensional poverty).

Five main stages are involved in the Small Area Estimation presented in this report:

1. Assessment of the degree of similarity of the common variables in both UNHS and Population Census
2. Producing a good predictive model of multidimensional poverty using the UNHS data
3. Fitting the predictive model in step 2 using, in this case using a Hierarchical Bayesian estimator
4. Predicting and validating the multidimensional poverty estimate using the best model found in step 3 to the Census data
5. Producing small area estimates for different geographical areas

In the following, we describe these five stages.

2.1 Assessment of the common variables in both UNHS and Population Census

The UNHS contains information from the Uganda Consensual Approach Deprivation module to classify the population into multidimensional and not multidimensionally poor (UNICEF Uganda, 2019). Households, adults and children were classified as poor or not poor based on the poverty and social exclusion project (PSE) methodology (Gordon, 2006), after suitable modification to adapt it to the Ugandan context (Government of Uganda, UNICEF, CU, UBOS and BPI, 2019; Pomati et al. 2020)). This methodology draws upon Townsend (1979) theory of relative deprivation and upon Mack and Lansley's (1985) method to identify enforced lack of socially perceived necessities in the population. Based on the needs endorsed by the majority of the population, a valid and reliable deprivation score is produced to measure the extent of multiple material and social deprivation of each respondent. This methodology has been successfully applied in low, middle and high income countries and leads to highly reliable and valid multidimensional poverty indices (Guio et al., 2017; Lau et al., 2015; Pomati and Nandy, 2020). The multidimensional poverty measure derived from this methodology identifies poverty using both income/resources and age-speficic deprivation scores (sum of relevant deprivations experienced by adults and children, with age-specific cut-offs) (see and GoU et al, 2019 for a full overview and Table A2 in the Appendix for the questionnaire items).

A key task in SAE is to find an optimal set of variables that can be used to predict multidimensional poverty (i.e. dependent/response variable). These variables need to be measured in a similar way in both the survey and the Census so that it is possible to estimate poverty rates using the information in the Census. A random sample from the 2014 Population Census was provided by the Uganda Bureau of Statistics (UBOS) (in close collaboration with UNICEF). The sample contained socio-demographic information from 730,407 households, and it was large enough to produce estimates with confidence at the parish level. Ideally, when the Census and the Survey are undertaken at the same time (i.e. same year), the point estimates will be very similar in both sources, any differences are likely to be mainly due to sampling error. However, a range of factors can affect the comparability of the UNHS and the population Census:

a) There is a two to three-year gap between the 2014 Census and the 2016/17 UNHS survey.

b) The methods used to identify the number of household members were different. The UNHS provides information on usual residents who were able to respond, whereas the Census provided information about the usual residents and also about guests and household members who were not present during the interview.
c) Therefore, the identification of the household head may also have been different (i.e. the household head may have been away at the time of the UNHS survey interview), and therefore variables like the occupation of the household head may not be strictly comparable. Hence, some key variables like the household-head socio-demographic profile cannot be included in the model (the effect of this omission is explored below).
d) The Census data was a random sample of the whole Census, which includes all areas of Uganda, whereas the UNHS does not cover all areas of the country.

Both 2016/17 UNHS and 2014 Census contain a common sub-set of variables such as sources of drinking water, types of fuel used for cooking, types of sanitation available to the household and types of wall and floors, and the ownership of certain household assets. However, several important differences were found between both samples[1]. These discrepancies are very likely to affect the validity of the models (i.e. its capacity to reproduce the findings in both the UNHS survey and the Census), and therefore, a form of correction is necessary to be consistent with the best practices in the international literature (Leulescu and Agafitei, 2013).

A common strategy involves reweighting one of the datasets to narrow the differences between both sources, i.e. to reduce random measurement error. In most cases, this means recalibrating the weights of the survey to match the Census data. Since it is highly likely that the Census has a smaller amount of error than the Survey, the weights of the UNHS were calibrated to match the Census distributions of the common set of variables used in the small area estimation model of multidimensional poverty (see below). A perfect match between the distribution of Census and survey variables is not a requirement for SAE, but calibration can help correct some of the larger differences between Census and survey data due to the larger sampling errors inherent in all survey data.

Table 1 shows the gap between the survey and the Census before and after the recalibration of the weights. This table only shows the variables that were good matches after reweighting the survey. Formally, we assessed the accuracy of the match using the Hellinger Distance statistic which is used to quantify the similarity of two distributions. This is zero when both distributions are perfectly matched, and a value below 5% difference is usually taken as the threshold of adequate similarity (Leulescu and Agafitei, 2013). In this case, before calibration, the distances are quite large when comparing Ugandan subregions. After calibration, most of the distances are considerably reduced. Although, some differences remain, they are sufficiently small to fit a good predictive model. SAE estimators (see next section) are designed to reduce the average differences in order to obtain reliable and valid predictions. For example, where the differences are large, and the estimated effect of the given variable is small, it would produce small deviations in the prediction, which will be further reduced by the shrinkage method inherent in Bayesian hierarchical models which further reduces sampling variation. We therefore assume that the SAE models used to produce the maps (see below) the differences between the Census and UNHS variables have relatively small effects on the results.

[1] Although only the variables with a reasonable good match were kept, a spatial analysis was included in section 3 to assess the relationship between multidimensional poverty and key predictors like: household head illiteracy and participation in paid work.

Table 1. Point estimates (%) of the values of different predictors from the Survey and the Census data

Names	Census	Uncalibrated Survey	Calibrated survey	Hellinger*
Urban area	28	28.2	28	0
Clothes deprivation	12	9.5	11	1
Shoes deprivation	31	35.7	34	2
Roof deprivation	26	24.1	25	1
Wall deprivation	53	46.4	52	1
Latrine	9	7.6	7	3
Covered pit latrine	21	30.7	27	5
Covered pit latrine with a slab	33	38.4	36	3
Covered pit latrine without a slab	7	3.1	5	3
Uncovered pit latrine with a slab	18	10.1	13	5
Uncovered pit latrine without a slab	1	0.2	1	1
No facility	8	7	8	0
Other	1	0.2	1	1
No Television	86	82.8	85	1
Improved wáter	73	79.1	74	1
Children in HH (0)	22	19	20	2
Children in HH (>0 & <4)	47	46	47	0
Overcrowding	40	35	39	1
No Bicycle	68	76	72	3

* A Hellinger distance below 5 the suggested threshold in the literature

2.2 Producing a good predictive model of multidimensional poverty using the UNHS data

The second step consists of finding a regression model capable of making good predictions of the poverty status (poor or not poor) of each respondent given a set of available variables - based on the list of variables in Table 1-. These predictors were used in the UNHS to predict multidimensional poverty, computed using the Consensual Approach (CA) method outlined in the 2019 UNICEF report (GoU et al, 2019).

In this case, the dependent variable is a binary variable which distinguishes poor and non-poor households (based on enforced lack of necessities as well as equivalised household expenditure cut-offs), so the model adopted is a logistic regression model. We first assessed the fit of different models, and we chose the simplest model that produced the best fit using a range of goodness of fit tests. Both the coefficients and the fit of the four models are shown in Table 2 below:

a) Model 1: Model without including Sub-region-level intercepts.
b) Model 2: Model 1 plus sub-region intercepts.
c) Model 3: Model 2 plus household head working in subsistence agricultural activities
d) Model 4: Model 3 plus household head literacy status

The coefficients in Table 2 indicate the differences in the log odds of of being identified as poor compared to the reference category for each variable, i.e. those suffering from shoes deprivation are more likely to be multidimensionally poor than those who do not suffer from shoes deprivation. The overall fit of the models was assessed using three criteria: Nagelkerke R-squared, the proportion of poor people correctly identified as poor by the model (sensitivity), and the proportion of the not poor correctly identified (specificity). These values are shown at the bottom of Table 2. This final model was formulated and finalised in collaboration with UBOS Uganda[2].

Table 2 Logistic regression models predicting individual-level poverty status

	Model 1[a]	Model 2
Urban	-0.15 *	-0.12
clothes deprivation	0.29 **	0.09
shoes deprivation	1.33 ***	1.23 ***
roof deprivation	0.29 ***	0.28 *
wall deprivation	0.44 ***	0.58 ***
Sanitation type (Flush toilet)		
Latrine	1.59 ***	1.68 ***
Covered pit latrine	2.55 ***	2.34 ***
Covered pit latrine with a slab	3.02 ***	2.92 ***
Covered pit latrine without a slab	2.60 ***	2.37 ***
Uncovered pit latrine with a slab	3.29 ***	3.15 ***
Uncovered pit latrine without a slab	2.13 *	2.01 *
No facility	3.39 ***	3.31 ***
Other	4.28 ***	4.47 ***
tv deprivation	1.59 ***	1.46 ***
Improved water	0.16 *	-0.04
Number of children	0.27 ***	0.31 ***
Overcrowding	0.52 ***	0.45 ***
bicycle deprivation	0.74 ***	0.56 ***
N	15646	15646
Nagelkerke R2	0.32	0.34
Specificity	0.75	0.78
Sensitivity	0.75	0.78

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. [a]Model 1 does not include Sub-region intercepts.

2.3 Fitting the selected model in step 2 but using a Hierarchical Bayesian estimator

Model 2 is a good predictive model but is not ideal for producing reliable small area estimates (Rao and Molina, 2015). Furthermore, given the differences between data sources, we would like to further improve the accuracy of the models. The models presented above assume that district area-level effects are zero. This is not a reasonable assumption, and it is rarely met in practice- as where you live can have a significant influence on your likelihood of being poor.

The literature on SAE proposes several ways (estimators) to make the most of the available data and its structure (Rao and Molina, 2015). The contemporary literature suggests implementing a hierarchical estimator, such as the HB, to allow for these contextual/area-based

[2] Other models 2a and 2b were fitted to assess the impact of not including key variables as employment status and literacy (See Table A1 in the Appendix). The conclusion, based on the overall fit of the models, is that there is only a marginal loss in the predictive power of the models. We also included information about the age and sex of the household-head in the model, but these independent variables did not improve the specificity and sensitivity. Furthermore, when literacy of the household-head was included in the models, we found that the prediction using the Census data diverged more from the Survey estimates.

effects. There are many options in the literature designed to achieve this desirable outcome (Rao and Molina, 2015).). In low-income countries, one of the most widely used is SAE models is the Elbers et al. (2003) method to produce small area estimates for income-based poverty measures (Haslett & Jones, 2010). However, empirical analyses have shown that it is far from the best method and also it is inadequate for use with categorical data (Guadarrama et al. 2014; Haslett and Jones 2010; Rao and Molina 2015). We have opted to rely on a more robust and flexible estimator, the Hierarchical Bayes estimator (HB), which is a much more powerful way to estimate prevalence rates for small areas and more adequate when working with real data (Guadarrama et al. 2014; Nájera et al. 2019).

The HB, like the EBLUP estimator[3], allows for the area-level variability by estimating the adjusted means of each district -random intercepts- and/or specific slopes for the explanatory variables for each district -random slopes-. This adds more information to the model and improves the accuracy of the coefficients of the household-level predictors. This, of course, adds to the complexity of the model but the HB has the great advantage of being feasible to calculate for complex models (unlike the Empirical Bayes estimator that relies on Maximum Likelihood and thus can be very computationally intensive, especially for categorical variables); thus some variants of model 2 were used to find an even better predictive model. Furthermore, for these estimates, we rely on the Hamiltonian Monte Carlo (HMC) approach, a significant breakthrough in Bayesian computation that makes estimation quicker and more efficient (Hoffman and Gelman, 2014). The HB model was thus fitted using the HMC, which was implemented on Rstan (Carpenter et al., 2016).

The variants were the following:

A) **Model 3**: Model 2 plus random intercepts at the district level. (See Table 4)
B) Model 4: Model 3 plus random slopes for different variables at district level.

The third step consisted of fitting the HB to the survey data (the next step is applying the model to the Census). Two criteria were used to evaluate the fit of this model:

A) Statistical fit: the WAIC (widely applicable information criterion) statistic of fit and Loo (leave-one-out cross-validation for fitted Bayesian models). The importance, relevance and robustness of Loo are discussed by Vehtari et al. (2017).
B) Capacity to reproduce the subregion point estimates, i.e. whether the model reproduces the observed data (design estimates of multidimensional poverty).

Based on the Loo index (Vehtari et al., 2017)., the best model was Model 4. However, the chosen model was Model 5 because of the predictive gains of Model 4 relative to Model 3 were minimal and Model 4 would take an unreasonably long time to compute on the Census data using currently available computer technology. The results below (Tables 3 and 4) summarise the accuracy of the prediction for the household-level model, i.e. by taking random samples and refitting the model, we could manage to predict accurately almost all households with the hierarchical model. This gives us confidence that when applying the model to a different data set, (i.e. the Census), the model will do a good job of predicting multidimensional poverty.

[3] The Empirical Best Unbiased Linear Predictor is for continuous variables but can be generalized for categorical data (Rao and Molina, 2015).

Table 3. Model 3 Pareto k diagnostic values

	Household-level model
Good	15,636
OK	5
Bad	3
Very bad	1

Table 4. Model 3. (Model 2 plus random intercepts). Hierarchical Bayesian Model. Coefficients from the posterior distribution

	Model 3*	Rho**
(Intercept)	-5.2	1.0
Urban	-0.2	1.0
clothes deprivation	0.5	1.0
shoes deprivation	1.0	1.0
roof deprivation	0.2	1.0
wall deprivation	0.4	1.0
Electricity deprivation	0.6	1.0
No tv	1.5	1.0
Improved water		
Number of children	0.3	1.0
Overcrowding	0.3	1.0
Sanitation type (Ref: Flush toilet)		
Latrine	0.5	1.0
Covered pit latrine	1.0	1.0
Covered pit latrine with a slab	1.4	1.0
Covered pit latrine without a slab	1.6	1.0
Uncovered pit latrine with a slab	1.9	1.0
Uncovered pit latrine without a slab	2.2	1.0
No facility	1.8	1.0
Other	2.2	1.0
No radio	0.4	1.0
No bicycle	0.3	1.0
Standard Deviation (District Intercept)	0.7	1.0
N Households	**15645**	
Loo model	14582.8	
* Mean estimate (Bayesian model)		
** Values closer to 1 mean that the MCMC chains have good mixing		

2.4 Predicting poverty using the best model found in step 3 to the Census data

Having a good predictive hierarchical model is important, but one of the essential checks in SAE is the capacity of reproducing the UHNS (design) estimates of poverty using a different sample (Rao and Molina, 2015). This was one of the main criteria used by CONEVAL in Mexico to validate their small-area estimates of multidimensional poverty (CONEVAL 2011). The reproduction of the design estimate is a critical check as we know that the UNHS provides reliable direct estimates at the sub-regional level, and we would, therefore, expect our model to provide comparable estimates. Table 5 shows that Model 3 reproduces almost perfectly the design estimates from the UNHS data.

We then applied Model 3 (fitted on 2016/17 UNHS data) to the 2014 Uganda Census and checked the model prediction by comparing it to the direct estimates from the UNHS. As shown in the third column of Table 5, there are some minor differences between the direct estimates from the UNHS and the model estimates using the Census, but they are all within the margin of error of the UNHS estimates.. Many of these differences are likely due to small differences in the basic needs and household assets prevalence rates found in the 2016/17 UNHS and 2014 Census, which is a recurring problem in Small Area Estimation. However, all the differences are quite small. We therefore proceeded to apply the UNHS model to the Census to produce Small Area Estimates of multidimensional poverty at district, sub-county and, for Kampla, parish level.

Table 5 Comparison of UNHS direct and model estimates at the sub-regional level of percentage of people in poverty

SubRegion	Survey point estimate	Survey: Model prediction	Census: Model Prediction
Kampala	9 [6-11]	9	8
Central 1	27 [24-29]	27	27
Central 2	38 [34-41]	38	39
Busoga	59 [56-63]	60	59
Bukedi	74 [69-78]	74	74
Bugishu	63 [59-68]	63	64
Teso	53 [49-57]	53	52
Karamoja	75 [70-80]	75	75
Lango	42 [39-46]	42	42
Acholi	61 [57-65]	61	59
Westnile	70 [66-74]	70	72
Bunyoro	41 [37-45]	41	41
Tooro	41 [37-45]	41	39
Ankole	29 [25-34]	29	31
Kigezi	50 [44-55]	50	52

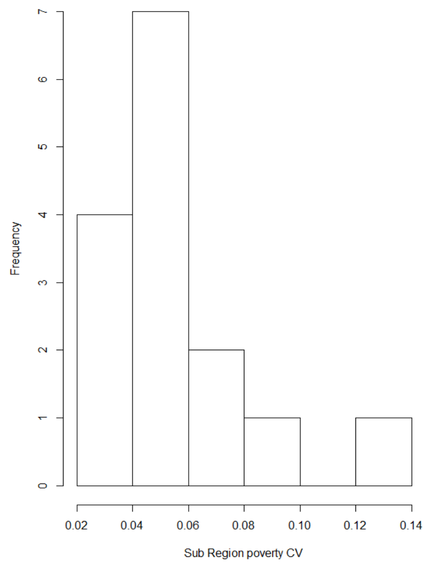
2.5 Producing small area estimates for different levels

Drawing upon the Rao and Molina (2015) SAE Bayesian estimator for categorical data, we estimated the prevalence of poverty for different geographies. The models were run at the household-level, and because all independent variables were at the household level, estimates for household-level poverty show the percentage of households predicted as multidimensionally poor in a given area. An essential validation of the predicted values is the low variation within the target areas (Molina and Rao, 2010). Figure 1 shows the coefficients of variation for four levels of disaggregation: SubRegions, Districts, Subcounty and Parishes.

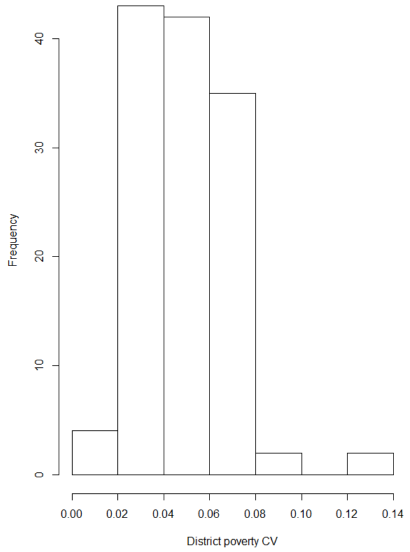
Figure 1 shows that the estimates are quite compact, and most are below 10%, as shown on the X-axis. This is a reasonable value for this standardised measure of the amount of variation (Molina and Rao, 2010), indicating that the model has managed to adequately shrink (ie reduce the effects of sample variation) the mean estimate of poverty for each spatial level.

Figure 1. Coefficient of variation HB estimates (Final Model) by different levels of disaggregation

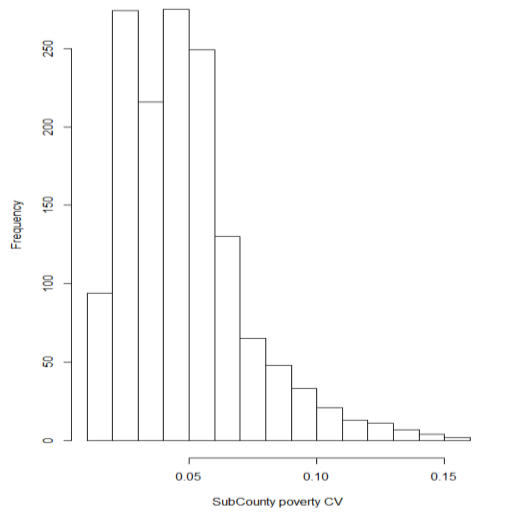
SubRegion Coefficient of variation



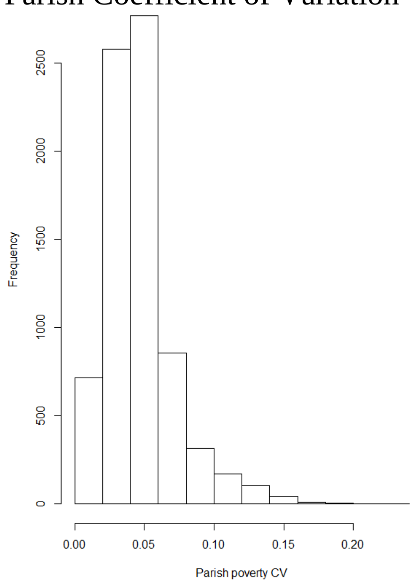
District Coefficient of Variation



Sub County Coefficient of Variation



Parish Coefficient of Variation



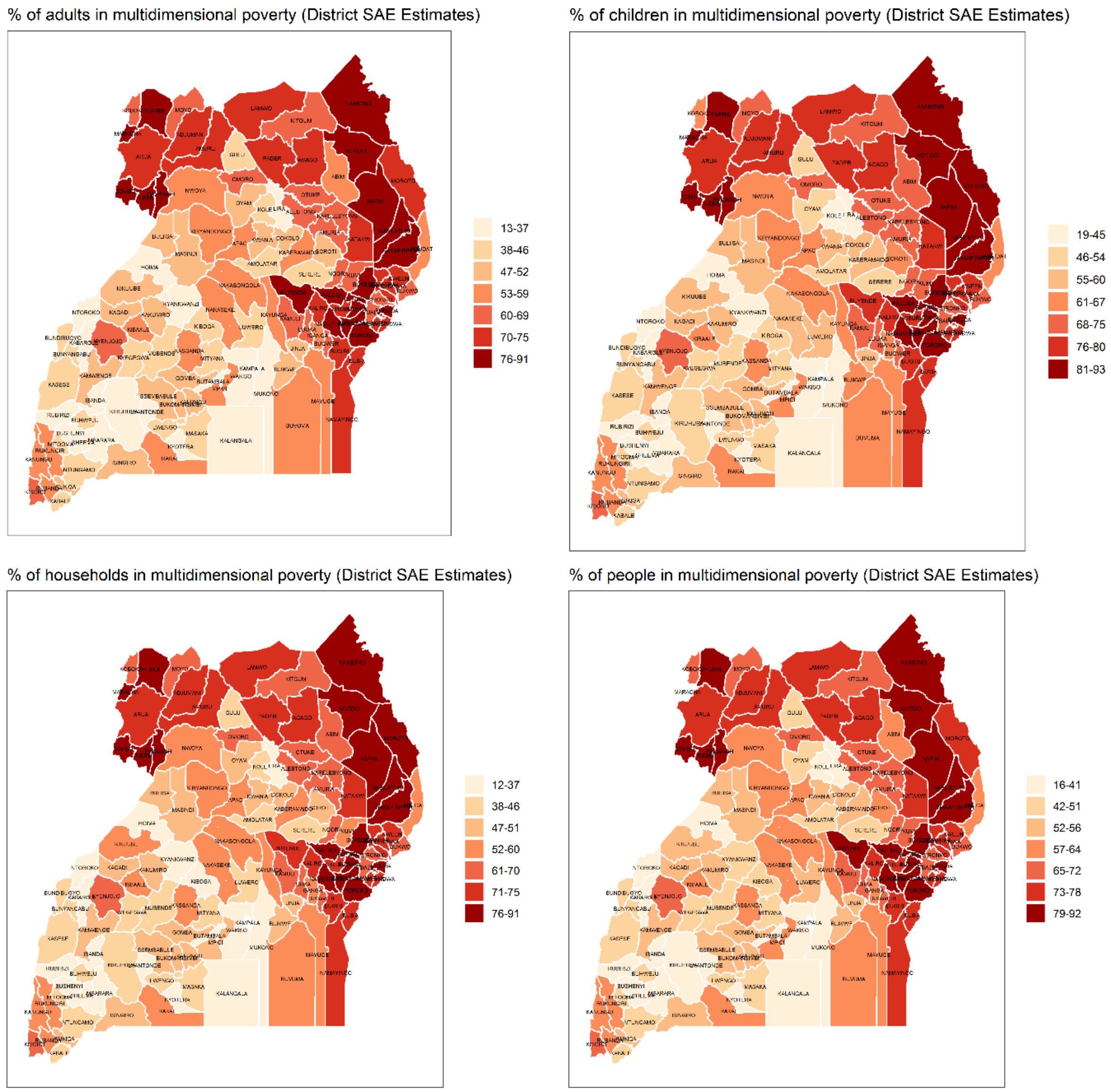
3. Results: The extent and distribution of household, adult and child poverty in Uganda

This section shows the SAE estimates for different populations using the following disaggregation levels: District, Subcounty and Parish. The prevalence of multidimensional poverty was mapped to facilitate the visualisation and interpretation of the results.

3.1 District-level multidimensional poverty estimates

The set of maps below display the distribution of multidimensional poverty at district level for different population groups: adults, children, households and total population. The maps are shaded in different colours to show areas with high and low levels of multidimensional poverty and the key is shown on the right-hand side of each map. For all maps, the darker the area, the higher the prevalence of multidimensional poverty and the lighter the lower the levels of poverty. The maps show that the highest rates of poverty are largely concentrated in the North East and North West of Uganda and that in general the lowest rates of poverty are in Kampala.

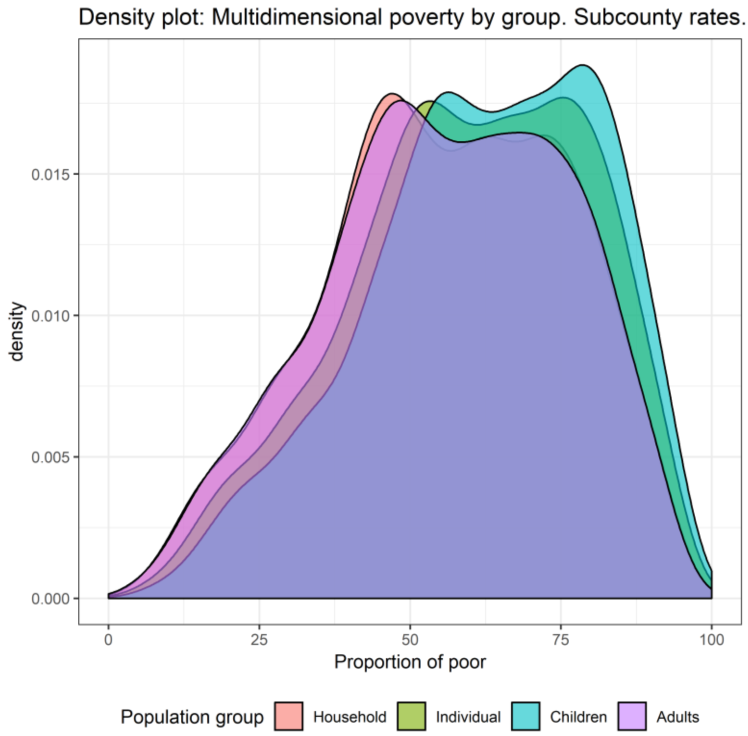
Figure 2. Multidimensional poverty maps. District level.



3.2 Subcounty-level multidimensional poverty estimates

Figure 3 shows the distribution of multidimensional poverty for each group. In all four cases, the shape of the densities shows that poverty varies a lot at the subcounty level but also that the poverty rates are quite high for the four population groups. The majority of subcounties have poverty rates above 25%. Among the four groups, the prevalence of child poverty is higher relative to the other groups (adults, households and total population)-the density is at the far right in the plot.

Figure 3. Prevalence of poverty at subcounty-level: adults, children, individual and household-level.



The four maps below (Figure 4) display the multidimensional poverty rates for the four populations groups (adults, children, households and total population) at subcounty level. In comparison with the district-level estimate, the pattern is more mixed pattern in both the north and the south west. The subcounties in the north tend to have higher poverty rates compared with the rest of Uganda. However, there are many subcounties in the central north with relatively low poverty rates – indicating that not all parts of the North of Uganda are poor. Similarly, the subcounties in the northwest constitute a large geographical area of very high poverty in Uganda.

Figure 4 (top pane). Multidimensional poverty maps. Subcounty level.

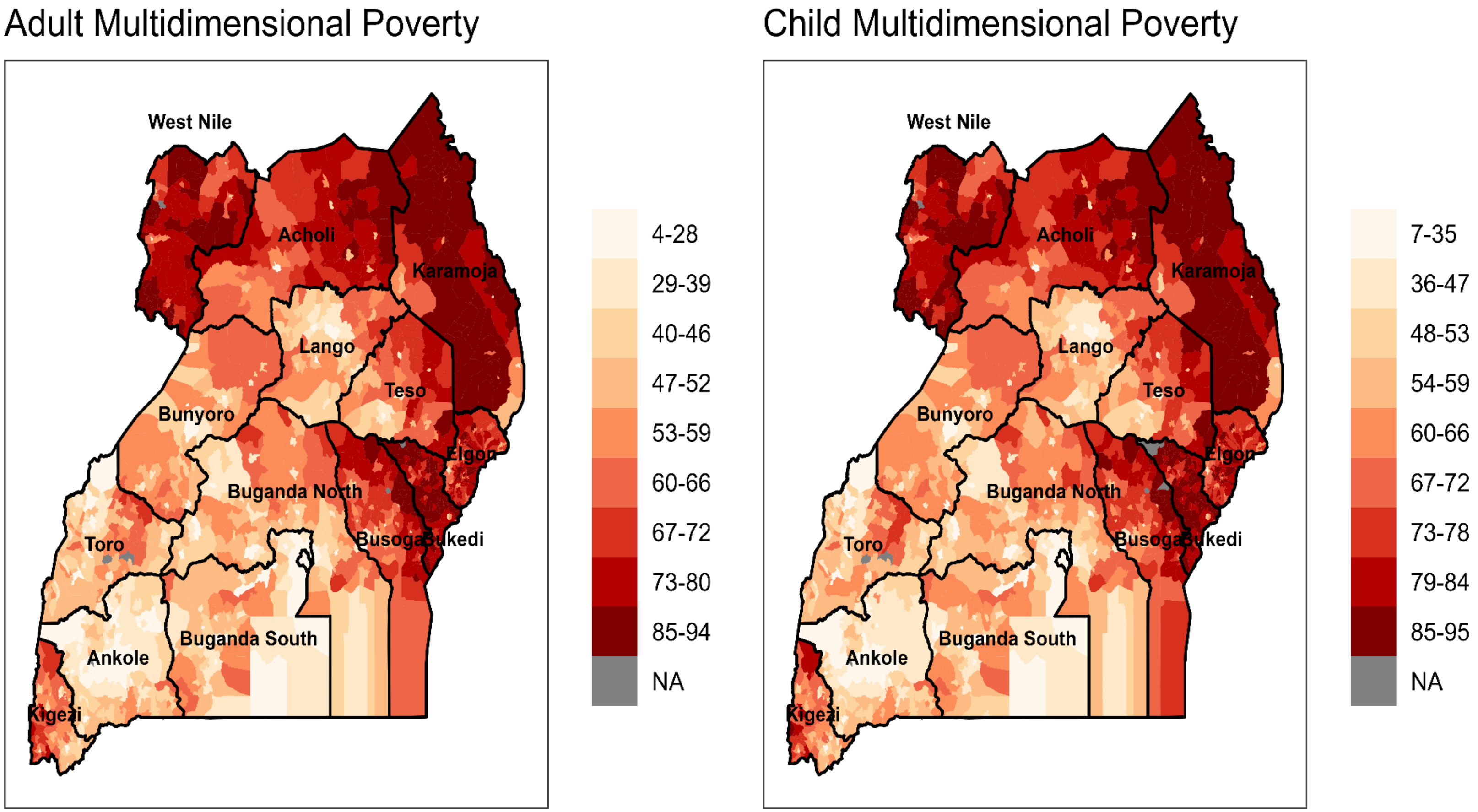
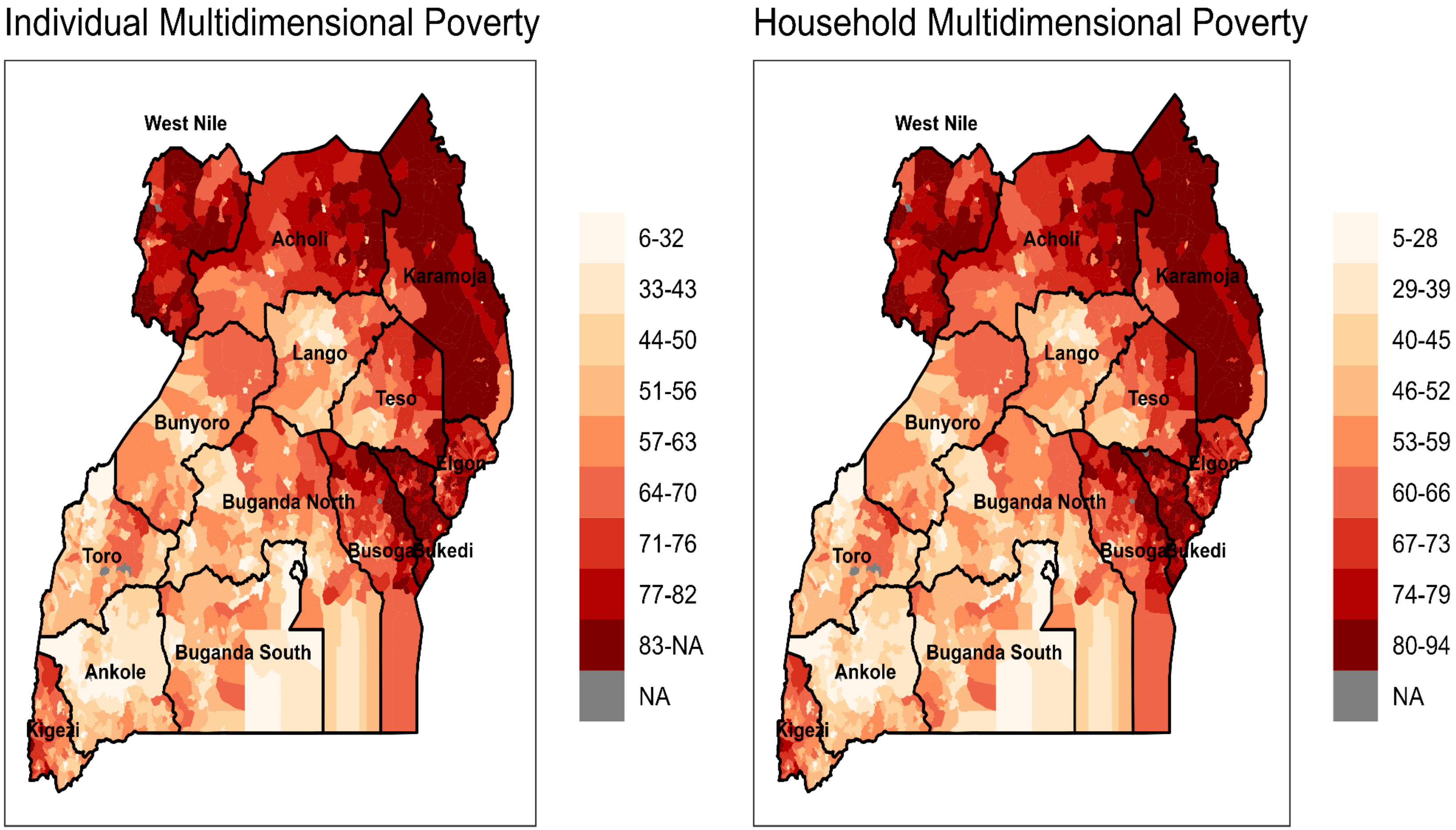


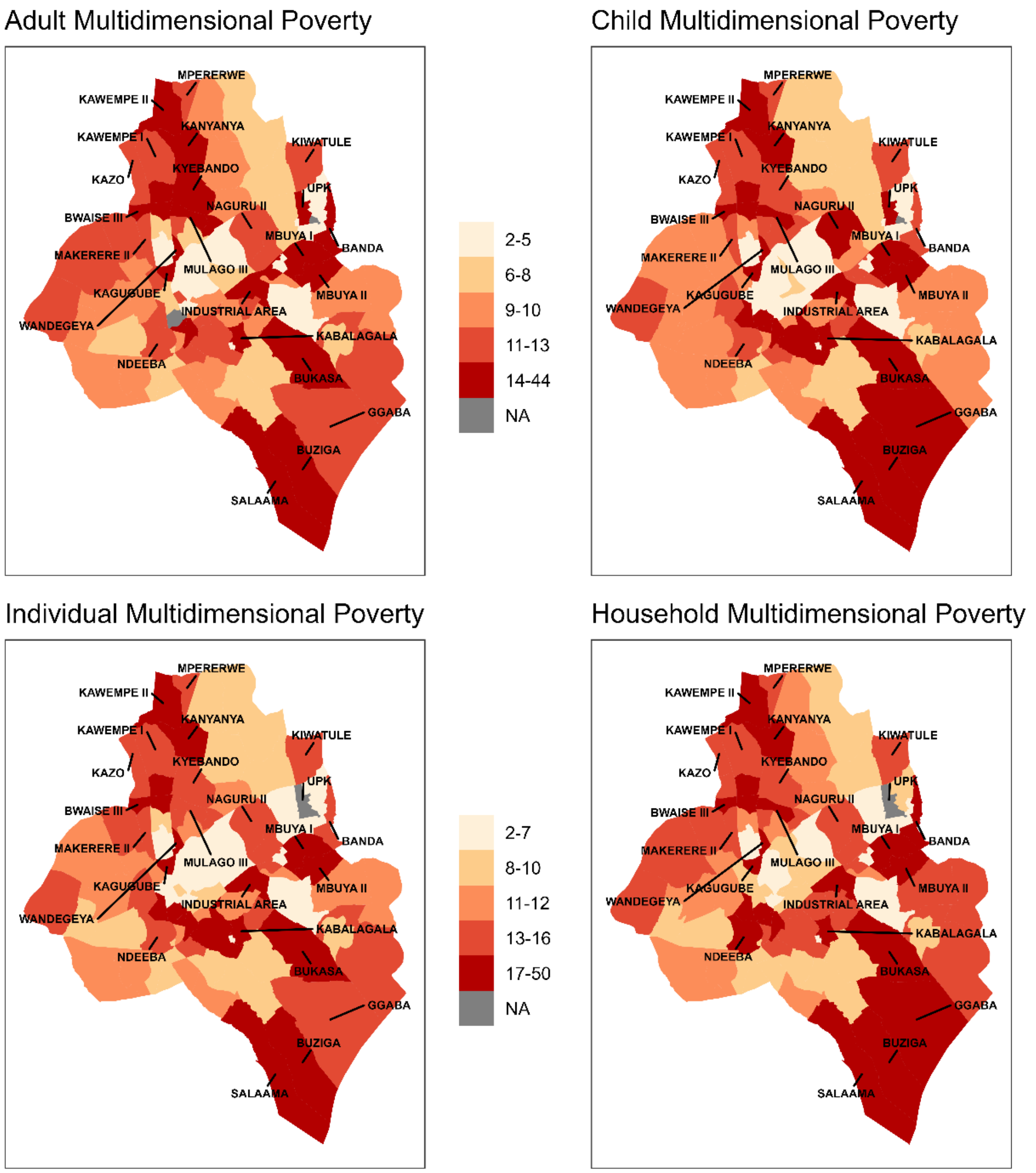
Figure 4 (bottom pane). Multidimensional poverty maps. Subcounty level.



3.3 Kampala parish-level multidimensional poverty estimates

The four maps below (Figure 5) show the prevalence of multidimensional poverty at parish-level in Kampala for adults, children, households and total population. The central area of Kampala shows very low poverty rates for the four population groups. In the north and in the south east of Kampala poverty are about twice as high as in the central area. Whereas on average Kampala has the lowest poverty rates in Uganda, it is important to underline that the non-central parishes have high poverty rates ranging between 20-60% for multidimensional child poverty.

Figure 5. Multidimensional poverty maps. Kampala. Parish-level.



3.4 Spatial analysis of multidimensional poverty

3.4.1 Spatial concentration of multidimensional poverty

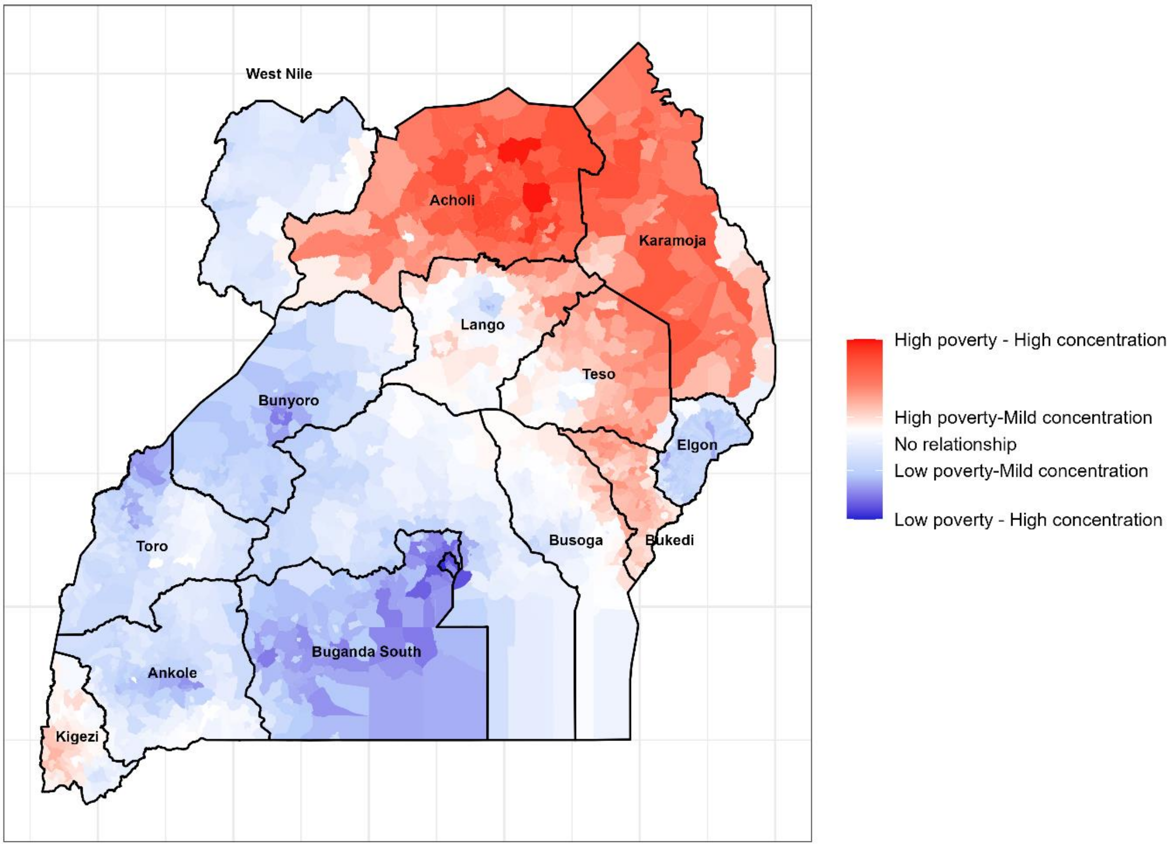
The subcounty-level maps suggest that multidimensional poverty is clustered in Uganda, i.e. high and low poverty rates tend to concentrate in certain areas. Table 6 provides a formal assessment of the geographical pattern. Global Moran's I statistics are a measure of spatial concentration i.e. how alike are neighbouring areas.. The more areas next to each have similar poverty ratesm, the closer the Global Moran's I will be close to 1. The values in Table 6 indicate that poverty, for all four groups, has a clustered spatial pattern.

Table 6. Global Moran's I statistic. Subcounty multidimensional poverty rates

Group	Moran's I (p-value)
Adult poverty	.40 (p<.01)
Child poverty	.39 (p<.01)
Household poverty	.39 (p<.01)
Total population poverty	.39 (p<.01)

Global Moran's I is a simple aggregate measure of geographical concentration. However, it provides only limited information as it does not show where exactly high or low rates of poverty are concentrated. One way to assess the specific clusters or hot spots of high or low poverty rates is by using the Local G statistics (Anselin, 1995). Figure 6 plots the significance tests of the Local G statistics, i.e. the areas where high or low concentrations of poverty are grouped into statistically significant clusters of geographic areas. The map shows that high poverty rates (shown in red on the map) are concentrated across the subcounties in the east of Uganda. There is another cluster of high poverty located in the south west. In the central area of the country, there is a cluster of subcountries, including Kampala and parts of Buganda South, with low multidimensional child poverty rates (shown in Blue on the map).

Figure 6. Local G statistic. Multidimensional Child poverty. Subcounty level estimates[4].



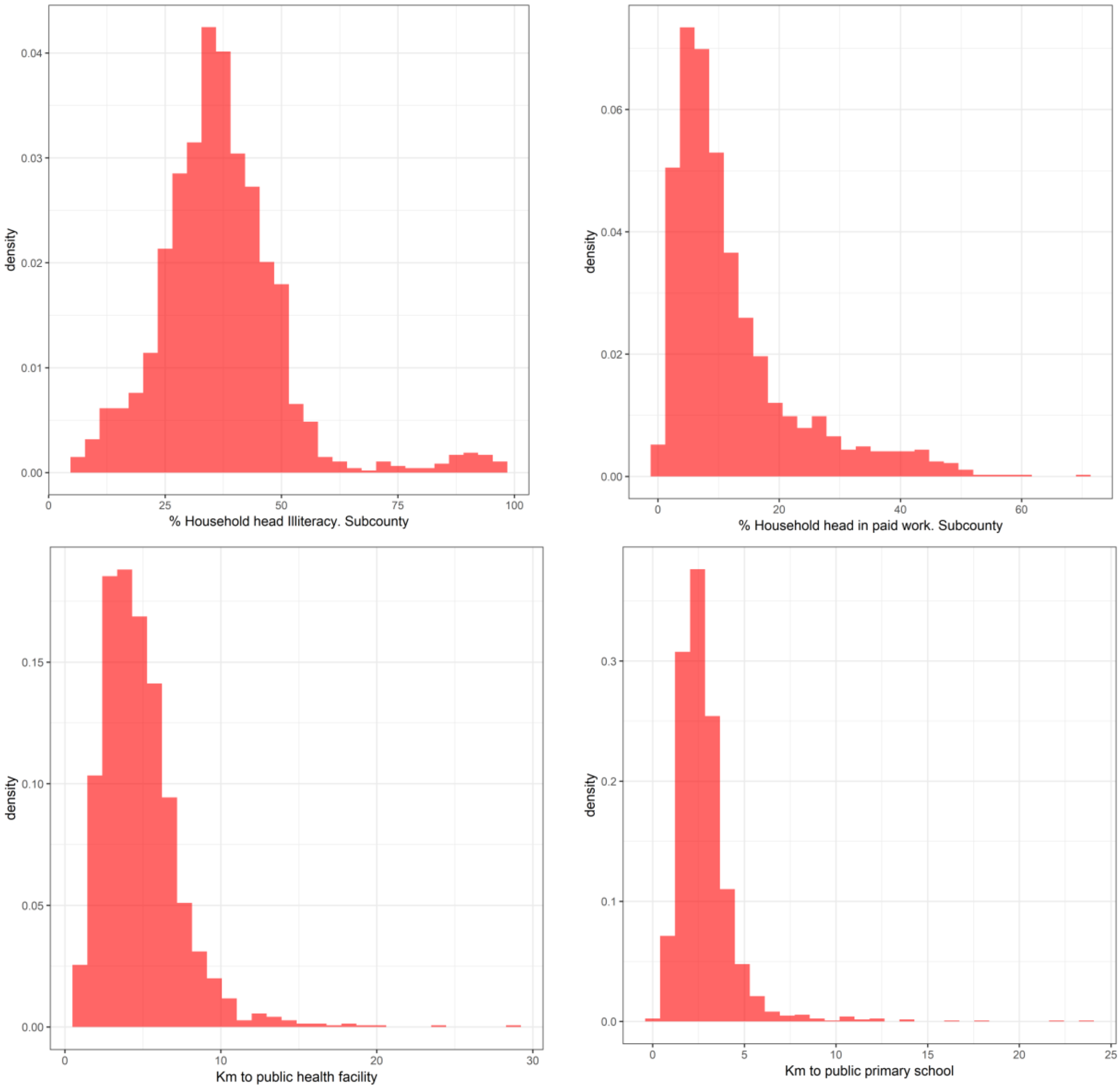
3.4.2 Local association between poverty and key socio-economic variables

The maps in previous sections show the geography of multidimensional poverty in Uganda. These spatial patterns of poverty in Uganda fit quite well with the results from other countries which also have high and low poverty areas clustered together (Davey et al., 2001; Dorling et al., 2007; CONEVAL, 2011; Nájera et al., 2019). This pattern is unlikely to be random, and it often mirrors policies which affect the geographical distribution of public services and the distribution of economic opportunities (Dorling et al. 2010; Venables, 2005). Therefore, it is important to describe the relationship between multidimensional poverty and key variables like household head illiteracy, household head participation in paid work, distance to public health facilities and to public primary schools (Figure 7).

Figure 7 plots the distribution, at subcounty-level, of the percentage household head illiteracy, the percentage of household head participation in paid work, the distance in km to public health facilities and to public primary schools. All of these variables vary considerably across subcounties. The question is how the spatial distribution of these important phenomena relates to distribution of multidimensional poverty.

[4] The maps for the other three groups (adults, total population and households) are not displayed as the patterns are quite similar.

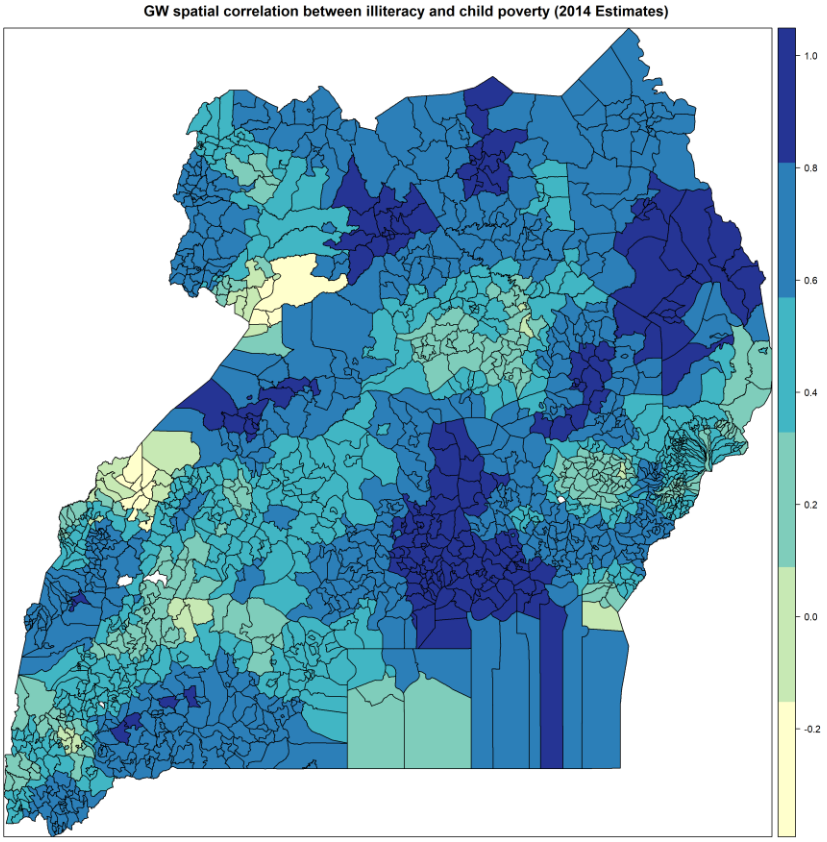
Figure 7. Distribution of key socio-economic and public provision variables. Subcountries. Uganda



In poverty research, most of the studies about the association between the chances of experiencing poverty and several socio-economic variables tend to focus on mean or aggregate relationships. In these types of analyses there is often an underlying assumption that the effect of increasing education in a population will have the same effect everywhere. From a geographical perspective, there are some features across regions and small-areas that might mediate and affect the effect of different policy variables upon poverty. Hence, it is important to have an idea of the varying relationship between different key variables and child poverty. To estimate such local or spatial relationship, we have to compute local correlations using Geographically Weighted correlations, which means allowing a correlation coefficient to vary across space (Brunsdon et al., 1996).

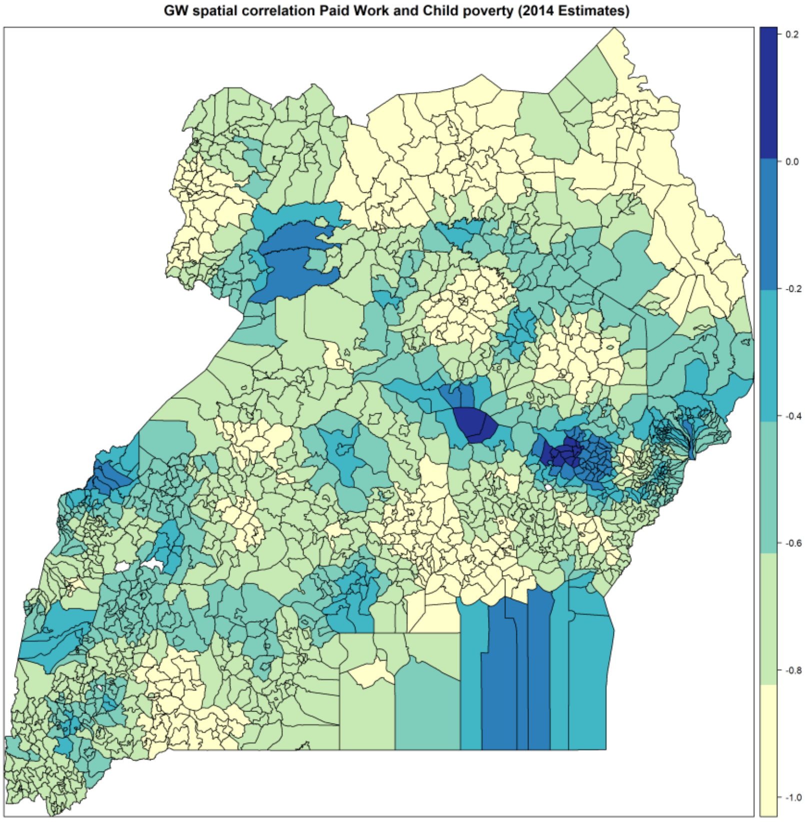
Figure 8 shows the spatial association, geographically weighted correlation, between the prevalence of child poverty and if the household head is illiterate or not. The legend on the right of the map shows the correlation coefficient for a given colour. Overall the map shows a strong association between illiteracy and child poverty, but it also shows that the strength of this relationship is not the same everywhere and varies across Uganda. The relationship tends to be stronger in the areas where child poverty is high than in areas where child poverty is lower. The association seems to be low in those areas with relatively low child poverty rates. This does not mean that there is no association between education and child poverty, but rather that the correlation between household head illiteracy and child poverty varies by the amount of poverty in an area.

Figure 8. Spatial correlation Household head illiteracy and multidimensional child poverty.



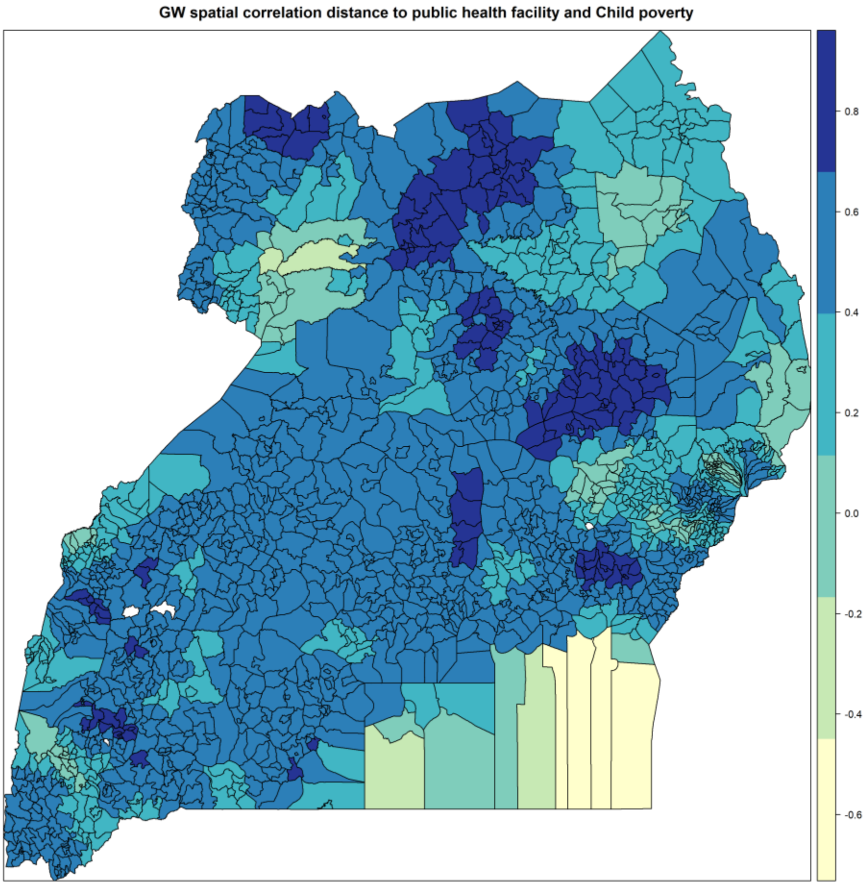
Using geographically weighted correlations, Figure 9 displays the spatial association between multidimensional child poverty and the participation of the household head in paid work. The association is negative across virtually all areas in Uganda. This means that the more people work in paid employment, the lower the rates of multidimensional child poverty. The association again is strong but not the same across the different subcounties in Uganda – it is much weaker in the far North East, North West and South West (the deep blue areas in Figure 9).

Figure 9. Spatial correlation Household head paid work participation and multidimensional child poverty.



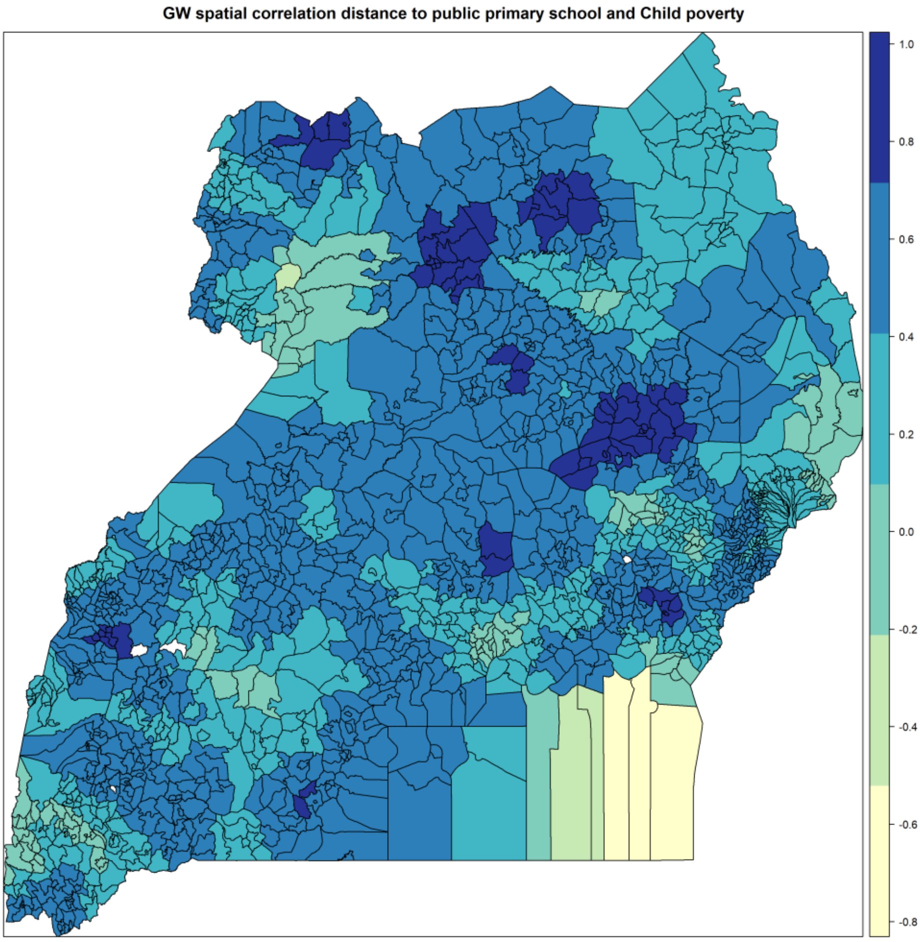
Both availability and proximity to health care are key challenges in developing countries. The average distance people have to travel to reach a public health facility in Uganda is around 4.8 kms. To assess the relationship between poverty and public health facilities, we computed the spatial correlation between these two variables. The geographical association between the mean distance to public health facilities and child poverty is displayed in Figure 10. The map suggests that across Uganda, the higher the child poverty rate, the further away the health facility is likely to be, from the household. The areas with higher poverty rates are also likely to be the areas with the greatest health needs but the worst health services – an example of the 'Inverse Care Law' (Tudor, 1971).

Figure 10. Spatial correlation. Mean distance to a public health facility and multidimensional child poverty.



Availability and proximity to primary schools are central to enhance children's rights and reduce poverty. In Uganda, the mean distance to a school is 2.8 km, and Figure 11 shows the geographical association between the mean distance to public primary schools and multidimensional child poverty. The association is positive across most subcounties in Uganda. That is, the areas with higher poverty rates also have schools that are farther away from the household's location.

Figure 11. Spatial correlation. Mean distance to a public primary school and multidimensional child poverty.



Conclusion

This report presents the first small areas estimates of multidimensional adult and child poverty for district and sub-district areas in Uganda, based on the 2016/17 Uganda National Survey (UNHS) and 2014 Uganda Census data. The estimates were produced following some of the best practices in the statistical matching of different sources and implementing one of the best computational procedures for small area estimation - the hierarchical Bayesian estimator.

The results show that adult and child poverty in Uganda have a clear geographical distribution and concentration. The areas in the north, particularly in the north east, tend to have very high multidimensional poverty rates (above 60%). This is in line with the overall picture presented by UNHS sub-regional poverty estimates. The subcounty-level multidimensional poverty estimates suggest that in 2014 there are clusters of high multidimensional poverty in the north east, the south west and some areas in the north west. The estimates for smaller areas, however, also show that there are pockets of high poverty in subcounties that do not appear to have very high poverty rates at sub-region level.

Kampala has very low multidimensional poverty rates relative to the rest of the country. The prevalence rate is on average 8%. However, the distribution of poverty within Kampala is not homogeneous, and there are parishes with child poverty rates of up to three times higher than the average.

The geographical analysis shows that multidimensional child poverty is highly correlated, at a spatial level, with high illiteracy rates and low participation in paid work but also that this association is not the same everywhere. The clusters of high concentration of high child poverty (in subregions such as Acholi, Karamoja and Teso) generally show strong association with illiteracy. The spatial analyses also show that there is a strong association between the areas with high rates of child poverty and the need to travel long distances to health care facilities and primary schools. The maps presented in this report help identify areas across Uganda that showed a particularly high correlation between multidimensional poverty and these important determinants of poverty.

Small area estimation involves making several assumptions about the quality of the data, the comparability between data sources and the plausibility of the model underlying the prediction. Therefore, there are many sources of error that affect the uncertainty around the estimates for a given small area. For future exercises, it is recommended to reduce the differences in the way key variables are measured in surveys and Census and, if possible, undertake the survey shortly after the census. These estimates can be helpful to inform policies and within a reasonable margin of error provide useful estimates of the subcounty and parish geography of poverty in Uganda

References

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, *27*(2), 93-115.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis*, *28*(4), 281-298.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Riddell, A. (2016). Stan: A probabilistic programming language. Journal of Statistical Software, 20, 1–37.

CONEVAL (2011) – Medición de la pobreza en los municipios de México.

Davey Smith, G., Dorling, D., & Shaw, M. (Eds.). (2001). *Poverty, inequality and health in Britain: 1800-2000: A reader*. Policy Press.

Dorling, D., Rigby, J., Wheeler, B., Ballas, D., Thomas, B., Fahmy, E., ... & Lupton, R. (2007). *Poverty, wealth and place in Britain, 1968 to 2005*. Bristol: Policy.

Dorling, D., & Pritchard, J. (2010). The geography of poverty, inequality and wealth in the UK and abroad: because enough is never enough. *Applied Spatial Analysis and Policy*, *3*(2-3), 81-106.

EURAREA. (2012). EURAREA small area estimation. European Comission.

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

Gordon, D. (2006). The Concept and measurement of poverty. In C. Pantazis, D. Gordon, & R. Levitas (Eds.), Poverty and Social Exclusion in Birtain: The Milenium Survey (pp. 29–69). Bristol Policy Press.

Government of Uganda, UNICEF, Cardiff University, Ugandan Bureau of Statistics & Bristol Poverty Institute (2019) Multidimensional Child Poverty and Deprivation in Uganda Report : Volume 1. Kampala: UNICEF

Guadarrama, M., Isabel, M., & Rao, J. N. K. (2014). A comparison of small area estimation methods for poverty mapping. *Statistics in Transition New Series*, *17*(2450–0291), 41–66. https://doi.org/10.21307/stattrans-2016-005.

Guio, A.-C., Gordon, D., Marlier, E., Najera, H., & Pomati, M. (2017). Towards an EU measure of child deprivation. Child Indicators Research. doi: 10.1007/s12187-017-9491-6

Haslett, S. J., & Jones, G. (2010). Small-area estimation of poverty: The aid industry standard and its alternatives. *Australian & New Zealand Journal of Statistics*, *52*(4), 341–362. https://doi.org/10.1111/j.1467-842X.2010.00588.x.

Hoffman, M. D., & Gelman, A. (2014). The No-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, *15*(1), 1593–1623.

Lau, M., Pantazis, C., Gordon, D., Lai, L., & Sutton, E. (2015). Poverty in Hong Kong. China Review, 15(2), 23–58.

Leulescu, A., & Agafitei, M. (2013). Statistical matching: a model based approach for data integration. *Eurostat-Methodologies and Working papers*.

Mack, J., & Lansley, S. (1985). Poor Britain (J. Mack & S. Lansley, Eds.). London, George Allen & Unwin.

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

Nájera, H., Fifita, V. & Faingaanuku, W. Small-Area Multidimensional Poverty Estimates for Tonga 2016: Drawn from a Hierarchical Bayesian Estimator (2019) *Appl. Spatial Analysis* doi:10.1007/s12061-019-09304-8

Pomati, M. and Nandy, S. (2020) Measuring multidimensional poverty according to national definitions - operationalising target 1.2 of the sustainable development goals. *Social Indicators Research,* 148, pp. 105-126

Pomati, M., Nandy, S., Angemi, D., Gordon, D., Depio, S., Ahaibwe, G., Kasirye, I., Muwonge, J., Ssennono, V., Baryahirwa, S., Twesigye, B., Eroglu-Hawksworth, S., Fahmy, E., Oloya, A. and Fifita, V. (2020) Multidimensional child poverty and the SDGs: from measurement to action. In: Lawson, David, Angemi, Diego and Kasirye, Ibrahim eds. *What Works for Africa's Poorest Children. From Measurement to Action*. Practical Action Publishing, pp. 19-42.

Pratesi, M. (2016). *Analysis of poverty data by small area estimation*. John Wiley & Sons.

Rao, J. N., & Molina, I. (2015). *Small-Area Estimation*. Wiley Online Library.

Townsend, P. (1979). Poverty in the United Kingdom: A Survey of Household Resources and Standards of Living. Penguin Books

Tudor Hart, J. (1971). The Inverse Care Law. The Lancet. 297: 405–412. https://doi.org/10.1016/S0140-6736(71)92410-X

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and computing*, *27*(5), 1413-1432.

Venables, A. J. (2005). Spatial disparities in developing countries: cities, regions, and international trade. *Journal of Economic Geography*, *5*(1), 3-21.

Appendix

Table A1.

Table 2 Logistic regression models predicting individual-level poverty status

	Model 1[a]	Model 2	Model 2a	Model 2b
Urban	-0.15 *	-0.12	-0.13	-0.12
clothes deprivation	0.29 **	0.09	0.09	0.08
shoes deprivation	1.33 ***	1.23 ***	1.23 ***	1.15 ***
roof deprivation	0.29 ***	0.28 *	0.29 *	0.28 *
wall deprivation	0.44 ***	0.58 ***	0.59 ***	0.57 ***
Sanitation type (Flush toilet)				
Latrine	1.59 ***	1.68 ***	1.69 ***	1.63 ***
Covered pit latrine	2.55 ***	2.34 ***	2.34 ***	2.29 ***
Covered pit latrine with a slab	3.02 ***	2.92 ***	2.92 ***	2.84 ***
Covered pit latrine without a slab	2.60 ***	2.37 ***	2.37 ***	2.31 ***
Uncovered pit latrine with a slab	3.29 ***	3.15 ***	3.16 ***	3.09 ***
Uncovered pit latrine without a slab	2.13 *	2.01 *	2.01 *	1.80 *
No facility	3.39 ***	3.31 ***	3.33 ***	3.21 ***
Other	4.28 ***	4.47 ***	4.47 ***	4.33 ***
tv deprivation	1.59 ***	1.46 ***	1.46 ***	1.41 ***
Improved water	0.16 *	-0.04	-0.03	-0.03
Number of children	0.27 ***	0.31 ***	0.30 ***	0.31 ***
Overcrowding	0.52 ***	0.45 ***	0.44 ***	0.47 ***
bicycle deprivation	0.74 ***	0.56 ***	0.56 ***	0.52 ***
Household head working in subsistence agriculture			-0.32	-0.40 *
Household head Illiterate				0.47 ***
N	15646	15646	15646	15645
Negelkerke R2	0.32	0.34	0.34	0.35
Specificity	0.75	0.78	0.78	0.78
Sensitivity	0.75	0.78	0.78	0.78

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. [a]Model 1 does not include Sub-region intercepts.

Figure A1. Comparison of the final HB model with a model where both literacy and paid-work status of the household head are included in the model.

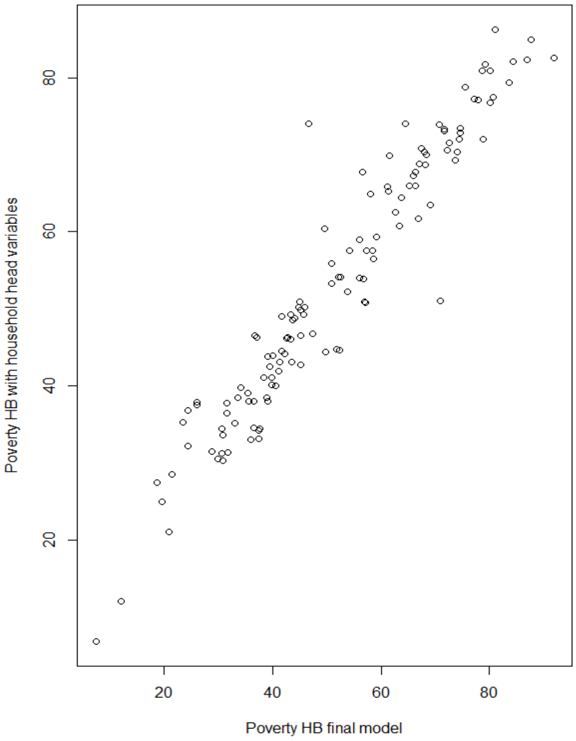


Table A2. 2016/17 Uganda National Household Survey. Consensual Approach questionnaire

Children's items (relevant to household members under 18 years of age)

Please say whether you think each of the following is essential for <u>all children (< 18 years) to be able to afford in order for them to enjoy an acceptable standard of living</u> in [COUNTRY] today. If you think it is essential, please say 'ESSENTIAL'. If you think it is desirable but not essential, please say 'DESIRABLE'. If you think it is not essential and not desirable, please say 'NEITHER'. So, the three possible answers are 'ESSENTIAL', 'DESIRABLE' or 'NEITHER'.

Item	Essential	Desirable, but not essential	Neither	DK	Have it	Don't have, can't afford	Don't have, don't want	Don't have, for another reason	DK/NA
QC1. Three meals a day	1	2	3	8	1	2	3	4	8
QC2. Two pairs of properly fitting shoes, including a pair of all-weather shoes	1	2	3	8	1	2	3	4	8
QC3. Toiletries to be able to wash every day (e.g., soap, hairbrush/comb)	1	2	3	8	1	2	3	4	8
QC4. Books at home suitable for their age (including reference and story books)	1	2	3	8	1	2	3	4	8
QC5. Some new clothes (not second hand or handed on/down)	1	2	3	8	1	2	3	4	8
QC6. Educational toys and games	1	2	3	8	1	2	3	4	8
QC7. A visit to a health facility when ill and all the medication prescribed to treat the illness	1	2	3	8	1	2	3	4	8
QC8. Own bed	1	2	3	8	1	2	3	4	8
QC9. Own blanket	1	2	3	8	1	2	3	4	8
QC10. Two sets of clothing	1	2	3	8	1	2	3	4	8
QC11. Presents for children once a year on special occasions, e.g., birthdays, Christmas, Eid	1	2	3	8	1	2	3	4	8
QC12. All fees, uniform of correct size and equipment required for school (e.g., books, school bag, lunch/lunch money, stationery)	1	2	3	8	1	2	3	4	8
QC13. To be able to participate in school trips or events that cost money	1	2	3	8	1	2	3	4	8
QC14. A desk and chair for homework for school-aged children	1	2	3	8	1	2	3	4	8
QC15. Bus/taxi fare or other transport (e.g., bicycle) to get to school	1	2	3	8	1	2	3	4	8
QC16. Own room for children over 10 years of different sexes	1	2	3	8	1	2	3	4	8

Household items (relevant to all household members)

Please say whether you think each of the following is essential for <u>everyone to be able to afford in order for them to enjoy an acceptable standard of living</u> in [COUNTRY] today. If you think it is essential, please say 'ESSENTIAL'. If you think it is desirable but not essential, please say 'DESIRABLE'. If you think it is not essential and not desirable, please say 'NEITHER'. So, the three possible answers are 'ESSENTIAL', 'DESIRABLE' or 'NEITHER'.

Item	Essential	Desirable, but not essential	Neither	DK	Have it	Don't have, can't afford	Don't have, don't want	Don't have, for another reason	DK/NA
QH1. Enough money to repair or replace any worn out furniture	1	2	3	8	1	2	3	4	8
QH2. Enough money to repair or replace broken electrical goods, e.g., a refrigerator	1	2	3	8	1	2	3	4	8
QH3. To be able to make regular savings for emergencies	1	2	3	8	1	2	3	4	8
QH4. To be able to replace broken pots and pans for cooking	1	2	3	8	1	2	3	4	8
QH5. Enough money to repair a leaking roof for the main living quarters	1	2	3	8	1	2	3	4	8
QH6. Have your own means of transportation (e.g., car, bike, motorcycle, etc.)	1	2	3	8	1	2	3	4	8

Adult items (relevant to household members over 18 years of age)

Please say whether you think each of the following is essential for <u>every adult (18+ years) to be able to afford in order for them to enjoy an acceptable standard of living</u> in [COUNTRY] today. If you think it is essential, please say 'ESSENTIAL'. If you think it is desirable but not essential, please say 'DESIRABLE'. If you think it is not essential and not desirable, please say 'NEITHER'. So, the three possible answers are 'ESSENTIAL', 'DESIRABLE' or 'NEITHER'.

Item	Essential	Desirable, but not essential	Neither	DK	Have it	Don't have, can't afford	Don't have, don't want	Don't have, for another reason	DK/NA
QA1. A visit to a health facility when ill and all the medication prescribed to treat the illness	1	2	3	8	1	2	3	4	8
QA2. Toiletries to be able to wash every day (e.g., soap, hairbrush/comb)	1	2	3	8	1	2	3	4	8
QA3. Two pairs of properly fitting shoes, including a pair of all-weather shoes	1	2	3	8	1	2	3	4	8
QA4. A small amount of money to spend each week on yourself	1	2	3	8	1	2	3	4	8
QA5. Replace worn-out clothes by some new (not second hand) ones	1	2	3	8	1	2	3	4	8
QA6. To get together with friends/family (relatives) for a drink/meal at least once a month	1	2	3	8	1	2	3	4	8
QA7. Celebrations on special occasions, such as Christmas, Eid.	1	2	3	8	1	2	3	4	8
QA8. Attend weddings, funerals and other such occasions	1	2	3	8	1	2	3	4	8
QA9. Able to access to safe, reliable public transport, such as buses and boats	1	2	3	8	1	2	3	4	8
QA10. Enough money to pay school fees for children	1	2	3	8	1	2	3	4	8
QA11. Enough money to take children to a medical facility when sick	1	2	3	8	1	2	3	4	8