

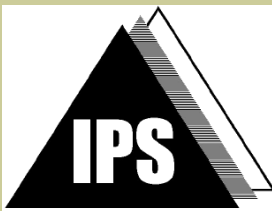


Politiques **E**conomiques et **P**auvreté
overty and **E**conomic **P**olicy

CRDI  IDRC

De La Salle University
AKI
Angelo King Institute
for Economic and Business Studies

 UNIVERSITÉ
LAVAL

 **IPS**

*Poverty Monitoring, Measurement and Analysis
(PMMA) Network*

*Incorporating Environment Factors
in Poverty Analysis Using
Small Area Estimation Techniques:
The Case Land Use Changes in Uganda*

*Paul Okwi
Uganda*

*A paper presented during the 4th PEP Research Network General Meeting,
June 13-17, 2005, Colombo, Sri Lanka.*

Incorporating Environment Factors in Poverty Analysis Using Small Area Estimation Techniques: The Case of Land Use Changes in Uganda

By

Patrick Birungi

Paul Okiira Okwi^{*}

Doreen Isoke¹

A Draft Report Prepared for PEP (April 2005)

^{*} The authors Patrick Birungi and Paul Okwi both lecturers at Makerere University, Faculty of Economics and Management, and ¹Doreen Isoke, working with the ministry of Finance, and Economic Planning, express their appreciation to the Poverty and Economic Policy (PEP) program, supported by IDRC for the financial support to the activities of this project. We also acknowledge and appreciate comments and advice from colleagues at Makerere University, and the unanimous reviewers from PEP. The views and any errors in this report are entirely the responsibility of the authors.

Abstract

This study combines census, survey and biophysical data to generate spatially disaggregated poverty/biomass information for rural Uganda. It makes a methodological contribution to small area welfare estimation by exploring how the inclusion of biophysical information improves small area welfare estimates. By combining the generated poverty estimates with national biophysical data, this study explores the contemporaneous correlation between poverty (welfare) and natural resource degradation at a level of geographic detail that has not been feasible previously. The resulting estimates of poverty measures have improved by the inclusion of environmental factors and the poverty estimates appear to be more robust, as the standard errors show a decline.

1.0. Introduction and motivation of the study

Environmental degradation can inflict serious damage on poor people, because their livelihoods often depend on natural resource use, and their living conditions offer little protection from the degraded environment. Environmental quality is a very important determinant of their health, earning capacity, security, energy supplies, and housing quality (Dasgupta et al., 2003). Studies have shown that the poor peoples' economic dependence on natural resources makes them particularly vulnerable to environmental degradation (Cavendish, 1999; Cavendish, 2000; Kepe, 1999).

If the above hypothesis is correct, then efforts to achieve poverty reduction in Uganda may be short lived because major environmental concerns need to be addressed. For example, despite the fact that over the last decade, poverty has reduced from 56% of the population below the poverty line in 1992; to 38 percent in 2002/03, (GOU, 2003; 2004) the country has experienced significant environmental degradation. This takes various forms that include, land degradation mainly due to soil erosion, deforestation and bush clearing, and over cultivation (NEMA, 2002). There have also been significant changes in landscape and land use patterns. For instance, the forest cover is said to be reducing at about 50,000ha (0.8% of the forestland) each year through deforestation most of which occurs in woodlands outside protected areas, which are mainly converted to agricultural land (NEMA, 2002). The Ugandan situation is unique because two decades ago, the country was faced with deteriorating economic, social and environmental conditions. However today, the social and economic trends have been greatly reversed, but it is not clear what the implications of these changes are for the natural resource base.

This raises concerns about the future supply of goods and services provided by the natural resources on which the poor depend for their survival and sustainability of agriculture in the country, that is a source of livelihood for the poor. Unless the environmental degradation is stopped or reduced, the current observed poverty reduction efforts might not provide long-run solutions.

Explaining theoretical links between poverty and environment dominates the literature on the subject. Existence of many theoretical papers without follow-up

empirical work is mainly due to lack of comprehensive data sets that cover comparable welfare and biophysical information. The few existing empirical works, are based on case studies, which are not representative, since environmental problems are spatial in nature. Using a combination of geo-referenced environmental information and household expenditure this study is able to explore the relation between poverty and the environment at a fine resolution that has not been possible before. This study therefore answers the following questions; What is the relationship between the location of the poor and the environment? How do changes in levels of poverty relate to changes in selected environmental indicators? Does incorporation of the biomass information help improve the precision of poverty estimates at a higher resolution?

This study draws upon earlier attempts, to improve poverty estimates using small area techniques in Uganda. Okwi et al., (2003) describe in detail how, using the integrated household survey (IHS) data and the population census, small area welfare estimators are derived for Uganda for 1992. Hoozeveen et al., (2004) on the otherhand shows how updated small area welfare estimators can be generated in the absence of a new census. Both studies understate the role of environmental/biomass variables on poverty analysis and propose to check their impact on welfare in their analysis. They however provide a good basis and framework for further analysis. Okwi et al., (2005), incorporate the role of environmental variables for a single point in time, using data for 1991. This study expands the analysis to capture how changes in poverty and environmental variables are correlated over time, using a panel of households interviewed in both periods. A key advantage of using panel data is that it allows us to control for unobserved time-invariant factors at the household and community levels.

A study of this nature is important for the country because it contributes to policy in a number of ways. First, documenting key relationships under a range of agro-ecological and demographic conditions provides a basis for deriving policy implications, within the ongoing framework of poverty reduction. This is very relevant, especially for Uganda where: 80% of population is engaged in agriculture; 90% relies on wood for fuel; 70% uses surface water for drinking and 39% is poor as of 2001/2002. Secondly, the study provides an opportunity for geographic targeting of

resources for investment in infrastructure and conservation of the environment. In Uganda where a program of decentralizing planning and fiscal responsibilities to lower local governments is being implemented, geographic targeting has much appeal. Thirdly, the paper demonstrates that despite structural changes to the Ugandan economy during the 1990s, it is possible to estimate a model for 1999/2000 per capita expenditure using household characteristics from 1992. The model is acceptable in part because of the accuracy of its coefficients and its R^2 . More importantly the welfare estimates derived from it are plausible in that they closely replicate stratum level estimates calculated directly from the household survey. The welfare estimates are satisfactorily precise as well. For instance 1999/2000 headcount rates of poverty for sub-counties (4th administrative level) have 95 percent confidence intervals of approximate the same width as those of stratum level estimates in the household survey.

Following this introduction, the remainder of this paper is structured as follows. Section 2 describes the data that form the basis of the research reported in this paper and provides an outline of the methodology. Section 3 sets out the empirical implementation of the model that underpins the analysis of the data, drawing extensively on the existing literature on small area estimation techniques. In section four, updated small area welfare estimators for rural Uganda are derived for 1992 and the panel of 1999/2000. The section also compares small area estimates for 1992 and 1999 derived with and without biomass information. Section five discusses the relationship between welfare and the environment using poverty maps. It presents a geographic profile of poverty and the environment for 1992 and 1999/2000 and how the changes in environmental factors are related to changes in poverty. Section 6 is the conclusion and discussion of policy implications.

2. Methodology and Data requirements

2.1. Data

The central element in this study is the availability of survey, census and biomass information. The poverty mapping portion of this project makes use of three household data sets: census data for 1991 and sample survey data from 1992 (IHS) and 1999/2000 (UNHS) to derive welfare estimates and maps. The IHS used a

stratified sample of 10,000 households in both rural and urban areas. The survey questionnaire collected information on household and demographic characteristics, education, assets, employment, income and expenditure (UBOS, 1992/93). This survey was based on four regions divided into rural and urban strata. In this study, we only use the 4 rural strata as for these strata we can also derive updated welfare estimates for 1999 (using a sample of 1263 households present in both the IHS and the UNHS).

The 1991 Population and Housing Census was conducted by the same institution (UBOS) and was meant to cover the entire population in both rural and urban areas. Two forms of questionnaires were used, a short and long form. The short form of the questionnaire covered mainly information on household members and education and was administered to all households in the country. The long form of the questionnaire covered housing characteristics and access to basic utilities and was administered to only 10% of rural areas (UBOS, 1991). The 10% is representative at district level. Although the census did not collect information on income and expenditure, it provides information on a number of characteristics likely to be correlates of poverty. The census and survey data have several common household variables such as household size composition, education, housing characteristics, access to utilities and location of residences.

The spatial analysis portion of this project used a variety of spatially referenced variables describing topography, land cover and land use, and roads. Geo-referenced information from the National Biomass Study of the Ministry of Water, Lands and the Environment is used. The project developed its own classification system based on a combination of land cover and land uses. This information covers changes in land cover such as broadleaved tree plantation or woodlots, coniferous plantations, tropical high forests (normal and depleted/encroached), woodland, grassland, wetlands, water resources and land use such as subsistence and commercial farmland, and changes in landscape among other aspects. In this project (NBS), the country was split into 9000 plots with 3 sample plots at each intersection. However, due to influences of population density and agro ecological zones on land cover and tree growth, some adjustments were made on the overall total sample plots. Topographic maps, land cover maps (1:50,000) and Global Positioning System (GPS) were used to locate the

field plots on the ground. There were four categories of data capture and processing i.e. mapping (spatial and its attributes), biomass survey (field plot measurements), monitoring of biomass and land cover change. This information details the woody biomass stock for each plot and it can be used to assess the relationship between tree cover and poverty. The data is extremely rich in bio-physical factors and also includes the distribution of infrastructure like markets, roads, schools and others. Besides, the GIS format of the data allows us to explore the possibilities of merging the data sets using GIS variables. Many of these variables required considerable cleaning, processing, and further transformation in order to generate the variables used in the spatial analysis

2.2. Using small area welfare methods to estimate the incidence of poverty

In Uganda, the availability of high-resolution data sets a strong foundation for us to produce and use poverty-biomass maps. Although several approaches have been developed to design poverty maps, there has been less effort to develop poverty/biomass maps. The approach we use to link these problems uses statistical estimation techniques (small area estimation) to overcome the typical limitations in the geographic coverage of household welfare that surveys provide and the lack of welfare indicators in the census data, and includes biomass information to assess these changes.

Our approach to the analysis of the links between poverty and land use changes using maps begins with the construction of a poverty map. We adopt the approach developed by Elbers, Lanjouw & Lanjouw - ELL (2003). The method is typically divided into three stages:

- Stage 0 involves identifying variables that describe household characteristics that may be related to income and poverty and that exist in both the household survey and in the census.
- Stage 1 estimates a measure of welfare, usually per capita expenditure, as a function of these household characteristics using regression analysis and the household survey data.

- Stage 2 applies this regression equation to the same household characteristics in the census data, generating predicted welfare for each household in the census. This information is then aggregated up to the desired administrative unit, such as a district or county, to estimate the incidence of poverty and the standard error of the poverty estimate.

The three sections below describe these methods in more detail and describe how they were applied in the current study.

Stage 0: Identifying household characteristics in the IHS, NHS and the Census

The first step was to compare the questionnaires of the 1992 Integrated Household Survey (IHS), 1999/00 UNHS and the 1991 Population and Housing Census to identify possible household characteristics found in both surveys that could be used as poverty indicators. The variables are derived from the comparable questions in the questionnaires. In addition to comparing the questionnaire, it is necessary to compare the values of the variables to ensure that they are in fact describing the same characteristics. A test is done to compare the means for the survey and census variables and the variables that pass the significance test are considered for the regression analysis.

Some household characteristics are categorical and, for regression analysis, must be represented by a number of dummy (binary) variables. For example, the main source of fuel used for cooking is a household characteristic, but for the regression analysis it must be represented by separate dummy variables for gas, electricity, fuelwood, kerosene, and so on. Based on this comparison, 162 household characteristics were selected for inclusion in the poverty mapping analysis.

Identifying identical variables between census and panel

Out of a total of 162 candidate variables 138, 148, 153 and 146 passed the means comparison test in respectively Central, East, North and West rural Uganda. 113 variables passed the test in all four rural strata. This is better than what was attained for the 1992 poverty map when respectively 143, 130, 128 and 130 variables passed

the means comparison test in Central, East, North and West rural Uganda and when 92 variables passed the test in all four rural strata.

Stage 1: Estimating per capita expenditure with a household survey

As mentioned above, Stage 1 of the poverty mapping method involves using the household survey data and regression analysis to estimate household welfare as a function of household characteristics. In this study, we use real per capita consumption expenditure from the 1992/3 and 1999 household surveys as the measure of household welfare. The explanatory variables are the household characteristics described above. Economic theory provides no guidance on the functional form, but generally a log-linear function is used:

$$\ln y_{ch} = \chi_{ch} \beta + \eta_c + \epsilon_{ch} \quad (1)$$

Where y_{ch} is the log of per capita consumption expenditure of household h residing in cluster c , X_{ch} are the observable characteristics of that household that are observable in both the survey and census data sets, and β is a coefficient vector. In our household survey, the clustering is done at regional (disaggregated into rural and urban) areas. The error term is composed of two parts. η_c applies to all households within the given cluster (location effect) while ϵ_{ch} is household specific component of the error term (heteroscedasticity). These two error components are assumed to be uncorrelated with one another and independent of the regressors. This specification of the error term allows for heteroscedasticity of the household specific error component. It also allows for the possibility of spatial autocorrelation. That is, location specific effects that are common to all households within a cluster. Because our main interest is predicting the value of $\ln(y)$ rather than assessing the impact of each explanatory variable, we are not concerned about the possible endogeneity of some of the explanatory variables. Elbers *et al* (2003) show that the probability that household i with characteristics X_i is poor can be expressed as:

$$E[P_i | X_i, \beta, \sigma^2] = \phi\left[\frac{\ln z - X_i \beta}{\sigma}\right] \quad (2)$$

where P_i is a variable taking a value of 1 if the household is poor and 0 otherwise, z is

the “overall poverty line” (see GSO, 2000, page 260), and ϕ is the cumulative standard normal function. If the predicted log per capita expenditure ($X_i\beta$) is equal to the log of the poverty line ($\ln(z)$), then the term in brackets is zero and the predicted probability that the household is poor is 50 percent. A lower predicted expenditure would imply a positive term in brackets and a higher probability that it is poor, while a higher predicted expenditure would imply a probability less than 50 percent.

To reduce the magnitude of the unexplained location specific component, we estimate a separate model to explain the cluster specific error terms. As regressors, cluster means of the household specific variables are obtained from the census and merged into the survey data set. This is a common procedure in poverty mapping. It amounts to explaining spatial autocorrelation between factors common to a household in a given Population Sampling Unit (PSU). To the extent that households attend the same school, make use of the same source of fuel wood or water and have similar access to markets, this procedure is likely to go a long way in explaining spatial autocorrelation. Yet, various rather obvious determinants of spatial autocorrelation cannot be obtained from the census. Population density, soil type and quality, access to infrastructure are examples of such information. By building an integrated data set with census and biomass information, we are able to include such bio-physical information in explaining spatial autocorrelation. We estimate equation 1 taking into consideration the location and heteroscedasticity component of the disturbance term. Survey weights are included in some of the regressions depending on the Hausman test (see Deaton 1997) results for whether the regressions should be weighted or unweighted.

Separate regressions were estimated for 1991 for each of the 4 rural strata of the survey data set. For 1999 only one model was estimated. We considered the set of variables that passed the test (zero stage) selection process and the final selection of variables was determined by a stepwise procedure.

Stage 2: Applying regression results to the census data

In Stage 2 of the standard poverty mapping method, the estimated regression

coefficients from the first step are combined with census data on the same household characteristics to predict the probability that each household in the Census is poor. This is accomplished by inserting the household characteristics for household i from the census, X_i^C , into equation 2. Thus, the expected probability that household i is poor can be calculated as follows:

$$E[P_i | X_i^C, \beta, \sigma^2] = \phi \left[\frac{\ln z - X_i^C \beta}{\sigma} \right] \quad (3)$$

Although this estimate is not very accurate for an individual household, it becomes more accurate when aggregated over many households. For a given area (such as a county or district), Elbers *et al* (2003) show that the proportion of the population living in households that are below the poverty line is estimated as the mean of the probabilities that individual households are poor:

$$E[P_i | X_i^C, \beta, \sigma^2] = \sum_{i=1}^N \frac{m_i}{M} \phi \left[\frac{\ln z - X_i^C \beta}{\sigma} \right] \quad (4)$$

where m_i is the size of household i , M is the total population of the area in question, N is the number of households, and X is an $N \times k$ matrix of household characteristics. The advantage of using the Census data, of course, is that the large number of households allows estimation of poverty headcounts for geographic units much smaller than would be possible with the household survey data.

Provided that a) the error term is homoskedastic, b) there is no spatial autocorrelation, and c) the full Census data are used, the variance of the estimated poverty headcount can be calculated as follows:

$$\text{var}(P^*) = \left(\frac{\partial P^*}{\partial \beta} \right)' \text{var}(\hat{\beta}) \frac{\partial P^*}{\partial \beta} + \left(\frac{\partial P^*}{\partial \hat{\sigma}^2} \right)^2 \frac{2\hat{\sigma}^4}{n-k-1} + \sum_{i=1}^N \frac{m_i^2 P_i^* (1-P_i^*)}{M^2} \quad (5)$$

where n is the sample size in the regression model. Thus, n , k , and σ^2 are from the regression analysis, while m_i , M , and N are obtained from the census data. The partial derivatives of P^* with respect to the estimated parameters can be calculated as follows:

$$\frac{\partial P^*}{\partial \hat{\beta}} = \sum_{i=1}^N \frac{m_i}{M} \left(\frac{-x_{ij}}{\sigma} \right) \phi \left(\frac{\ln z - X'_i \hat{\beta}}{\hat{\sigma}} \right) \quad (6)$$

$$\frac{\partial P^*}{\partial \hat{\sigma}} = -\frac{1}{2} \sum_{i=1}^N \frac{m_i}{M} \left(\frac{\ln z - X'_i \hat{\beta}}{\hat{\sigma}^3} \right) \phi \left(\frac{\ln z - X'_i \hat{\beta}}{\hat{\sigma}} \right) \quad (7)$$

The first two terms in equation 5 represent the “model error”, which comes from the fact that there is some uncertainty regarding the true value of β and σ in the regression analysis. This uncertainty is measured by the estimated covariance matrix of β and the estimated variance of σ^2 , as well the effect of this variation on P^* . The third term in equation 5 measures the “idiosyncratic error” which is related to the fact that, even if β and σ are measured exactly, household-specific factors will cause the actual expenditure to differ from predicted expenditure. These equations are described in more detail in Hentschel *et al.* (2000) and Elbers *et al* (2003).

Since we are using household level census data, the combination produces estimates of per capita expenditure for each household. Estimates of consumption for the census households must take into account the disturbance term, that is, the portion of the variation in consumption in the survey data that is not explained by variation in the regressors. If this is not done, the poverty estimates for the census data would be biased. We simulate the level of consumption for each household. The simulations draw the β coefficients from the multivariate normal distribution described by the point estimates and the variance-covariance matrix estimated in the first stage regression (equation 1). The set of simulated \hat{y}_{ch} values are then used to compute poverty estimates at different administrative levels. The poverty estimates are calculated at different levels (regional, district, county and sub county) for 1991 and

1999/2000. For each administrative unit or location, the means are the point estimates of the poverty rates, while the standard deviations are the standard errors of the estimates. The final step is to combine the generated welfare information with the GIS biomass data, generate overlays and carry out analysis.

2.3 Updating small area welfare estimators

At the core of the small area welfare estimation is an out-of-sample prediction of per capita expenditure using a set of representative household variables that is common to the survey and the census. A close correspondence between census and survey household characteristics is a pre-requisite to yield reliable welfare estimates. Much attention is therefore devoted to identifying common variables by assuring that variable definitions are identical between the census and the survey, that questions are phrased the same way, that coding and enumerator instructions are identical and that the survey and census are fielded contemporaneously. When the latter condition is not met -and this is more of a problem in rapidly changing economic environments, changes in the economic situation will be reflected in household characteristics. As a result, survey variables identified as common to the census, are actually not representative of the census and small area welfare estimates can not be derived.

The need for common, representative, regressors effectively closes the possibility to update poverty maps through the use of a household survey from a non-census year.¹ In the presence of panel survey data however, for which one of the waves has been collected at the time of the census, this problem can be avoided. The representativeness of the common survey variables with the census can be maintained by relying on household characteristics collected during the census year. Updated welfare estimates can then be based on expenditures obtained for the more recent period. More formally, and denoting time with subscript t , in the presence of panel data equation (1) can be re-written to:

$$\ln y_{ch, t+1} = E[\ln y_{ch, t+1} | X_{ch, t}] + \eta_{c, t+1} + \varepsilon_{ch, t+1} \quad (1^*)$$

¹ In reality survey and census are rarely administered at the same time, but the period between both is never long. And always much attention is devoted to assuring that household characteristics obtained from the survey are representative of those in the census.

Simulated log per capita expenditure is now derived as :

$$\ln \tilde{y}_{ch,t+1} = X_{ch,t}^T \tilde{\beta} + \tilde{\eta}_c + \tilde{\varepsilon}_{c,t+1}, \quad (8)$$

and welfare estimates are based on:

$$\tilde{\mu}_{t+1} = E[W_{t+1} | m_t, \tilde{y}_{h,t+1}] \quad (9)$$

This changes the original small area welfare estimation methodology in that instead of a contemporaneous association between per capita household expenditure and household characteristics, per capita household expenditure from a *different* time period is made conditional on household characteristics collected in the census year.

To implement the method three conditions have to be met: (i) the survey has to be reweighted, (ii) a set of common census-survey variables has to be identified and (iii) a sufficiently accurate expenditure model has to be estimated. Reweighting the survey is required because at the census based prediction stage only information on household size from the census year is available so that welfare estimates for year t+1 have to be based on information on household size from year t. To assure a close association between census and survey based welfare estimates for year t+1, it is needed to replicate the cross sectional per capita consumption distribution for year t+1 (based on $y_{h,t+1}$ and $m_{h,t+1}$) using $y_{h,t+1}$ and $m_{h,t}$. This implies reweighting the survey.

Reweighting the survey in one dimension (expenditure) may have consequences for its representativeness in other dimensions. Hence even if a set of representative variables has been identified between the survey and the census to make a poverty map for year t, it needs to be tested whether, with new weights, these common variables remain representative. After a set of common variables has been identified, a model for year t+1 per capita expenditure can be estimated with household characteristics from year t as regressors. Estimating a model of future expenditure on past household characteristics is unusual (though less so for permanent income adherents), but recall that the objective of equation (8) is to estimate the conditional

expectation of expenditure (from (1^{*})) and not a causal relation. The model is only usable if its coefficients are estimated accurately (to limit the variance attributable to model error) and if a reasonably high R² (to assure disaggregation for small target populations) is obtained. If these conditions are met, updating small area welfare estimates is feasible without the need for a new census.

2.4 Methods to estimate other measures of poverty

The methods described above allow one to estimate the incidence of poverty, defined as the proportion of people below the poverty line. We compute the welfare indicators measured by the conventional Foster-Greer-Thorbecke (1984) measures FGT (α). We report our estimates with p-values of 0, 1 and 2 reflecting respectively poverty incidence, poverty gap and the poverty gap squared.

These poverty measures can be expressed as follows:

$$P_{\alpha} = \frac{1}{N} \sum_{i=1}^M \left[\frac{(z - y_i)}{z} \right]^{\alpha} \quad (10)$$

Where z is the poverty line
 y_i is income (or expenditure) of person i in a poor household
 N is the number of people in the population,
 M is the number of people in poor households

Different values of α in equation 10 give different poverty measures. When $\alpha = 0$, this formula gives the incidence of poverty. This is because the term in brackets is always one, so the summation gives us the total number of people in poor households, which, when divided by N , gives us the proportion of people living in poor households. When $\alpha = 1$, it gives a measure called the depth of poverty (or the poverty gap). P_1 takes into account not just how many people are poor, but how poor they are on average. It is equal to the incidence of poverty (P_0) multiplied by the average percentage gap between the poverty line and the expenditure of the poor. When $\alpha = 2$, this equation gives a measure called the severity of poverty (or squared poverty gap).

P_2 takes into account not just how many people are poor and how poor they are, but also the degree of income inequality among poor households. It is equal to the incidence of poverty (P_0) multiplied by the average squared percentage gap between the poverty line and the income of the poor.

The poverty mapping method described in the above sections provide a method for estimating the proportion of people below a given poverty line, z , but do not provide any information on the distribution of income among the poor, which is necessary to calculate P_1 and P_2 . We can adapt the poverty mapping method to estimate P_1 and P_2 by noting that z does not have to be the poverty line. We can estimate the cumulative distribution of the population by level of per capita expenditure by running the poverty mapping calculations repeatedly for different values of z . More specifically, the following steps are used:

1. select 100 levels² of per capita expenditure, divided evenly along the range of per capita expenditure from the richest to the poorest household.
2. set z equal to the lowest of these 100 levels (call this z_1), run the poverty mapping calculations to calculate the proportion of the population with per capita expenditure below z_1
3. then repeat step 2 setting z equal to each of the other 99 expenditure levels (z_2 to z_{100}), storing the values of z_i and the proportion of the population below z_i in a file for further analysis.

As z_i rises from its lowest level to its highest level, the proportion of people with per capita expenditure below z_i rises from 0 to 100 percent. Thus, these results trace out the cumulative distribution of the population by per capita expenditure.

This information can be used to calculate the values of P_1 and P_2 . In the gap between each pair of z 's (z_i and z_{i+1}). we know the average per capita expenditure³ and the

² The use of 100 levels is arbitrary, the larger the number of levels, the more accurate the estimation of the cumulative distribution and hence, the more accurate the estimates of P_1 and P_2 . Increasing the number of levels, of course, also increases the computational burden and time to run the program.

³ strictly speaking, we only know the range of per capita expenditures in this group of households and we assume that the average is $(z_i + z_{i+1})/2$. But if we choose a large number of z 's, the difference

proportion of people with per capita expenditures in that range. Thus, each pair of z 's that are below the poverty line can be used to represent one value of y_i in equation 9, taking into account the number of households with per capita expenditure in that range.

3.0 Empirical Implementation

3.1 Zero Stage: Selection of Variables

In the “zero stage” we compared variables from the survey and census, and selected potential ones, which were later used in the regression models described in the methods above. Principally, the idea was to obtain variables from the household survey, which were comparable to those in the census. The initial step was to look at the questions in both the survey and census. This provides a clue as to whether the responses would provide similar information. However, it is not usually obvious that identical questions will yield similar responses for several reasons. For instance, the way the question was asked, the local language translation of the question, the ordering of the questions or even variations in interpretation of questions may provide major differences in the responses. To verify that the questions yielded similar answers, we conduct an assessment to determine whether the variables are statistically similarly distributed over the households in the survey and census. This assessment is done for each of the four strata and the comparison is done at regional level (four regions focusing only on rural strata).

After a comparison of wording, coding and instructions in the enumerator manual, we constructed a more disaggregated total of 162 potentially identical variables, which sometimes involved interactions among some variables. Then, using statistical criteria, we compare the stratum level means of the variables to assess the level of similarity. We do this by testing whether the survey mean for a particular variable lies within the 95 percent confidence interval around the census mean for the same variable. A third and final step is to do a comparison of the variables across the two categories of strata (rural and urban) to assess the level of uniformity in comparability. The selection of variables used in the first stage was based on criteria,

between z_i and z_{i+1} will be small, so the error in making this assumption will also be small.

which picked all continuous variables found to be comparable. For the dummy variables, we tested whether the census and survey means were identical (see Appendix A and B for list of variables and comparison, respectively).

3.2 Re-weighting

Despite being identified as potentially identical, household size did not pass the distribution comparison test. It differed consistently between the census and the survey in that small households are underrepresented in the survey. For instance, in Central rural the census mean for one-person households is 18.4 percent but the corresponding figure in the survey is 16.3 percent. As household size is crucial when deriving per capita welfare estimates, it was less of an option to drop it from the common set of variables. And fed by the suspicion that small households are underrepresented because of non-response and improper replacement (Hoogeveen, 2003) we decided to reweigh the survey.

The re-weighting strategy followed is known as post-stratification adjustment (Lessler and Kalsbeek, 1992). It ensures that the weighted relative frequency distribution among mutually exclusive and exhaustive categories in the survey corresponds precisely to the relative distribution among those same categories in the census. In total 13 different household size categories were distinguished, reflecting households of size 1-12 with category 13 reflecting households of size 13 and over, and re-weighting was done at the stratum level. A danger of re-weighting along one dimension, household size in this case, is that survey variables that were representative using the ‘old’ weights become non-representative once the weights have been adjusted to control for unrepresentativeness in other dimensions. On the other hand, if the adjustment corrects for a genuine sampling error, the comparability between the survey and the census should improve in all dimensions. As a check on the appropriateness of re-weighting, we compared the set of variables that were considered identical on the basis of wording, coding and enumerator instructions and how many passed the survey-census means comparison test before and after re-weighting. Re-weighting increased the number of variables that passed this test in all rural strata considerably from 23% to 43%, while improving the fit for household size related variables.

Having corrected for non-participation due to household size, another concern may be that survey participation varies with household wealth. Mistiaen and Ravallion (2003) demonstrate how such a wealth effect on survey compliance can be estimated using data on non-response across geographic areas. Using information on non-response rates per expenditure quintile at the district level (38 districts) we therefore also tested for wealth related non-compliance. Estimates for the linear model of non-compliance on per capita expenditure yielded insignificant results, whereas a quadratic specification turned out to be significant (at the 90% level). It shows an inverted-U shaped compliance-expenditure pattern with people in middle quintile groups more likely to comply than either the richest or the poorest. The difference in compliance rates is only marginal⁴, and we therefore only adjust for non-compliance related to household size.

3.3 First Stage

The first stage estimation is conducted using the household survey data, census and biomass data. Since we are analysing only rural data, the household survey is stratified into four sub-regions, and we estimate four different models. In this stage, we construct more interaction terms from the selected census, survey and biomass variables, then use a stepwise regression approach in SAS to select the variables which provide the best explanatory power to the log per capita expenditure. As is the case with other similar studies, we use a significance level criterion with no ceiling on the number of variables to be selected. The significance level used for selecting variables was 5 percent.

To capture differences between strata, stratum level models are usually estimated. This was the case for the 1992 poverty map. But with only 1071 rural panel households available, estimating separate models for each stratum could easily lead to over-fitting. In the North for instance as few as 160 panel households were interviewed. So for 1999/2000 one model is estimated with interaction terms for each region except Central which is subsumed in the constant term.

⁴ After correcting for wealth related non-compliance we estimate for the poorest quintile –which shows the largest divergence, that the true population proportion is 0.2097 (instead of 0.20); for the wealthiest quintile it is 0.1986.

Failing to account for spatial correlation in the disturbances would result in underestimated standard errors on poverty estimates. Sampling in the IHS and UNHS and household surveys is stratified into four regions (divided by rural and urban) and within each region primary sampling units (PSUs) are selected from the list of all census enumeration areas. Within the selected PSUs a number of households (typically 10) is randomly selected for inclusion in the survey. In the IHS, the PSU is therefore the level at which the cluster is defined and this is also the level at which the 1992 poverty map controls for location effects (Okiira Okwi et al. 2003). In the panel it often occurred that no, or only one panel household, was interviewed in a given PSU. So for the updated poverty map, the cluster is defined two administrative areas up from the PSU, at the county level.

To develop an accurate model of household consumption, we consider the model specified in equation 1. In this model, the error component is attributable to location and household specific effects. Presence of these errors makes our welfare estimates less precise. Since unexplained location effects reduce the precision of our poverty estimates, the first goal is to explain the variation in consumption due to location as far as possible with the choice and construction of explanatory variables. We attempt to reduce the magnitude of the location effect in four ways.

- i. We include in our specification district dummies and their interaction terms with key household level variables (household size, level of education, age of head of household).
- ii. We calculate means at the enumeration area in the census of household characteristics such as household size and composition, and the gender, age and average level of education of household heads. We then merge these EA means into the household survey and consider their interactions with household characteristics obtained from the survey for inclusion in the household regression specification.
- iii. For the information collected from the long form questionnaire, (for 10% of the rural households and representative at the district level) on housing characteristics, use of fuel, access to water sources etc. we calculate district means and interact these with household characteristics.

- iv. Finally, we include in our specification biomass variables and their interaction terms with key household level variables. The biomass variables include information on distance to roads, proportion of land under grassland, woodland, water, farmland and forests.

So far in the household model, cluster level means and biomass data interacted with household characteristics are included. To further select location variables we determine the common component in the household specific error terms and regress this on enumeration area and district means. We then select limited number (5 at most) variables that best explain the variation in the cluster fixed effects estimates. The number of explanatory variables is limited so as to avoid over-fitting. The selected location variables are included in the household regression model after which a combined model is estimated comprising of household specific and location variables.

A Hausman test described in Deaton (1997) is used to determine whether to estimate our final regression models for each stratum with household weights. We re-estimate the regressions in equation 1, but after adding weights to the selected explanatory variables. Then, using the Hausman test, we test the joint significance of the weighted explanatory variables, at 5 percent significance, and decide whether or not weighting is necessary for the regressions.

We model the idiosyncratic part of the disturbance by choosing variables from the set of potential variables selected from the census and survey, their squares and interactions. To select a subset of these variables, we use ε_{ch}^2 as the dependent variable in the stepwise regression and choose not more than 10 variables that best explain the variation in the household specific part of the residual.

Finally, we determine the distribution of η_c and ε_{ch} using the cluster residuals $\hat{\eta}_c$ and

standardised household residuals: $e_{ch}^* = \frac{e_{ch}}{\hat{\sigma}_{\varepsilon, ch}} - \left[\frac{1}{H} \sum_{ch} \frac{e_{ch}}{\hat{\sigma}_{\varepsilon, ch}} \right]$, respectively, where h

is the number of households in the survey. We use normal distributions for each of the error components. The consumption model is then re-estimated using the Generalised

Least Squares (GLS) method using the variance-covariance matrix resulting from the above equation.⁵

A major strengths of the poverty mapping method and inclusion of biomass data is that it calculates the standard errors, a measure of the accuracy of the estimate. Precisely, like any method of measuring poverty, the small area estimation method does not produce exact results. The household characteristics do not perfectly predict household expenditure in Stage 1. Even if they did, there may be differences between households in the IHS sample and those in the Census. Finally, our census data does not consist of the entire population for some housing characteristics, so there is some sampling error as well.

A number of factors affect the accuracy of the poverty estimate. First, if the Stage 1 regression equation is very good in predicting household expenditure based on the household characteristics, then the poverty estimates will be more accurate. Second, the accuracy of poverty estimates tends to be better for areas with poverty rates near 0 percent or near 100 percent. Third, the accuracy is better for areas with a large number of similar households than for areas with few and diverse households.

Standard errors help define the margin around the poverty estimates. There is a 95 percent chance that the “true” poverty estimate lies within two standard errors of the poverty estimate. For example, in the case of Central region, the estimated poverty rate is 54.3 and the standard error is 1.25. This means the 95 percent confidence interval of this poverty estimate is 54.3 percent \pm 2.5 percentage points ($1.25 * 2$). In other words, there is a 95 percent chance that the true poverty for Central rural is between 51.8 and 56.8 percent.

Table 1 below summarizes the results of the first-stage regression, and it shows that the adjusted R^2 s of the models for 1991 and the panel. The R^2 s for the 1991 model vary from 0.35 to 0.46⁶, (see also Tables C1 to C5 in the appendix C for examples of

⁵ For a description of different approaches to simulation see Elbers et al., (2001 ; 2003)

⁶ Note that the regressions are simply association models, therefore the parameter estimates should not be interpreted as causal effects. We do not claim to have tested for the presence of double causality in the model; in this study however, we are more interested in the associations and/or correlations between biomass variables and poverty indicators.

regressions results) and for the panel increased to 0.34 from 0.31. According to Table 1, inclusion of biomass information helped to raise the R^2 s by an average 2 percentage points for both the cross section and panel models compared to the models without them. The relatively low R^2 s in the rural areas may be attributed to at least two reasons. First, the number of variables in the census short forms is limited to mostly household composition, education and ethnic origin⁷. Secondly, household composition and education only change slowly over time. The returns to agriculture are variables much dependent on rainfall, illness of family labourers, incidence of pests and diseases and prices. Again some of this variation may be captured, for instance the age of the head of household and proneness to disease are correlated, but much of the cross sectional variation attributable to any of these sources will remain unexplained and gets subsumed in the error term.

Despite not being high, the explanatory levels are comparable to those attained elsewhere in Africa. For instance, in rural Madagascar the adjusted R^2 range from 0.239 to 0.460 (Mistiaen et al., 2002) and in Malawi it ranges from 0.248 to 0.448 (Machinjili and Benson, 2002). Considering that for Uganda, the long form of the questionnaire was available for only 10% of the rural households, the Ugandan R -squares seem to do relatively well.

Table 1: Summary Statistics of First Stage Regression Models (Rural Strata)

<i>Number of observations</i>	<i>Panel</i>	<i>IHS</i>			
	All rural strata	Central	East	North	West
Number of observations used in regressions	1071	1660	1640	1368	1637
Number of clusters ¹	117	163	165	144	163
Hausman test for weights	0.78	1.29	1.04	1.71	1.84
Regression weighted?	No	Yes	Yes	Yes	Yes
Adjusted R^2 without location means	0.30	0.27	0.32	0.39	0.31
Adjusted R^2 with location means no biomass	0.31	0.31	0.34	0.44	0.32
Adjusted R^2 with location means including biomass data	0.34	0.35	0.36	0.46	0.34

Note: In the IHS the cluster is defined by the census enumeration area; for the panel by the sub-county. In the panel, the predicted variance of the cluster effect is negative, and set to zero. Consequently in the predication stage cluster errors are not included for panel households. Information on the IHS is from Okiira Okwi et al. (2003).

⁷ Inclusion of all the variables from the short form and biomass data raised the R^2 but not to the urban strata levels implying we still needed to use more information such as housing and environmental characteristics to improve them.

A logical next step is to make the connection between welfare and environmental information. However, as already noted the regression analysis presents association and not causal models. The poverty-environment literature shows possible presence of the problem of double causality. Inadequate time series data on environmental as well as poverty variables renders it impossible to test empirically. In this study, we do not test for direction of causality. We are only interested in the associations and correlations between environmental and poverty related variables. There is need therefore for careful interpretation of the regression results. But it is important to note that obtaining information on biomass use for administrative units is not straightforward, because of confidentiality, different data formats, the intricacies of geo-analysis and because environmental conditions do not follow administrative boundaries. We consider a number of bio- physical factors, including proximity from parish centre to nearest main, tarmac and track roads separated into 1 to 5 kms, proportion of the parish land under woodlots, coniferous forests, tropical high forests, degraded forests, woodlands, grasslands, papyrus(wetland), subsistence and commercial farmland, water and impediments.

The regression results presented in Tables C1 and C5 in the appendix suggest some spatial correlation between poverty and some bio-physical variables. The ability of these variables to improve the explanatory power of the models is interesting but different variables were selected for the different strata. Once again, note that we are explaining spatial correlation and not causality. A few principal variables stand out to be clear correlates of poverty. Access to roads has much explanatory association to poverty in all the four rural strata. Despite the fact that the types of roads differ between the strata, the regression results indicate a close spatial correlation to poverty. In the rural central stratum, access to main and track roads was an important variable while in north rural, access to both main and tarmac roads was important. Likewise for east rural, access to track and tarmac roads was important and in the west rural, tarmac and track roads are important. The spatial correlation between poverty and access to roads is evident. Although our evidence is indirect, we conclude that access to various types of roads is potentially an important issue in Uganda. By implication, any policy focused on improving access to roads will yield disproportionate benefits for the poor.

Tables C1 and C5 in appendix C and Table E1 and E2 in Appendix E, summarize the available evidence of the association between poverty and other bio-physical information. Besides access to roads, the proportion of land under woodland, subsistence and commercial farms turned out to be the most important biomass variables associated with rural poverty in central rural. Meanwhile, in the east rural, proportion of land under commercial farms, woodland and the proportion of degraded forests were important spatial variables correlated with poverty. In the north, the proportion of land under water, subsistence farmland and subsistence farmland in the wetlands were the important spatial variables. The selection of water bodies and wet farmland is probably suggestive of the fact that northern region is generally dry and access to water or wetland could be important factors in explaining poverty, given that most of Uganda's rural population depends on agriculture. For west rural, the proportion of land under woodlots and subsistence farmland has spatial relations with poverty. In addition to the selected variables, biomass variables interacted with household characteristics also proved to be important in explaining the correlation between poverty and biomass. The results from the regression analysis clearly display regional variation in spatial correlation between bio-physical and poverty information. Although time series analysis would be useful, this evidence suggests that there is strong relationship between poverty and biomass variables. We conclude that access to subsistence and commercial farmland, wetland/water, woodlands, roads and grasslands are important spatial factors correlated with poverty in Uganda.

4. Updated small area welfare estimators for rural Uganda are derived for 1992 and the panel of 1999/2000

This section presents the welfare indicators derived from the out of sample predictions on the unit record census data. Mean per capita expenditure is presented along with measures of poverty. To this end the Foster-Greer-Thorbecke measures, $FGT(\alpha)$ are reported with α -values of 0, 1 and 2 reflecting respectively poverty incidence, the poverty gap and its square. As benchmark the official monthly per capita poverty lines (in 1989 prices) are used, i.e. Uganda shillings 15,947 for rural Central, shillings

15,446 for rural East, shillings 15,610 for rural North and shillings 15,189 for rural West.

Once the census and biophysical data sets are integrated, ELL welfare estimates can be improved (see for instance Mistiaen et al., 2002 for Madagascar). The preliminary poverty estimates for rural Uganda control for spatial autocorrelation solely by relying on PSU means calculated from the census. The second stage analyses sought to use the rural models highlight the importance of bio-physical factors in poverty estimation. First, the results of the second stage analysis are used to examine the extent to which the poverty estimates from the census and bio-physical data⁸ match the sample estimates at the level which the survey is representative (region). Secondly, we focus on the ultimate goal of the analysis, namely to produce disaggregated spatial profiles of poverty and biomass. Using poverty/biomass maps, we show how projecting poverty estimates and biomass information produces a quick and appealing way in which to convey a considerable amount of information on the spatial relationship between poverty and the natural environment to users. We use poverty and biomass overlays to show the spatial heterogeneity of poverty and land use.

4.1 Incidence of poverty: Cross sectional estimates including biomass data

Table 2 below summarizes the poverty and inequality estimates based on the predictions of the combined biomass and census at the regional level and the survey based estimates. The detailed estimates for the district level are presented in the appendices. To reduce clutter, the poverty estimates for the county and sub-county are presented in form of maps. In the map, the poorest areas are dark brown while the least poor areas are dark green.

Using the cross sectional data, the results confirm that poverty is most widespread in the North and Northeast, particularly in the sub regions of Karamoja and Acholi. At the stratum level, the results are reasonably close to those from the survey. Interestingly, most standard errors were lower than when no biomass data was

⁸ Some observations were missing in the census/biomass data therefore the populations represented may not be exactly the same as if it was census based data alone

included, in some cases by up to 40 percent. As shown in Table 2, the results show a consistent story with the survey and census-based estimates. Central rural emerges with the least level of poverty even when census/biomass data is used for prediction, while north rural remains the poorest of the four strata. When other measures of welfare such as the poverty gap (P-1) and the poverty gap squared (P-2) are used, the comparison among the rural strata still remains consistent with the survey rankings. The inclusion of the biophysical data improved the poverty estimates at the stratum level and lowered the census-biophysical based standard errors consistently. This is even when some parishes in the North and West did not have corresponding biophysical data (see Table 2 and Appendix D).

The inclusion of the biophysical information in the small-area estimation procedures can have two effects. First, the level of the poverty measures can change, and secondly, the standard errors of the estimates of poverty measures can change. Table 2 presents estimates of four poverty measures at the regional level in 1992. Poverty measures from three different sources compared. The survey-based estimates are directly calculated from the IHS database. The ‘Census predicted’ estimates are based on the ELL method without the use of biophysical information (see Okwi et al., 2003), and finally, the ‘Census/Biomass predicted’ estimates are from the present study. In this study we focus attention on the comparison of ‘Census’ and ‘Census/Biomass’ estimates.

Table 2 Poverty measures for four rural areas from different data sources, 1992.

Stratum			Central			East			North			West		
Poverty Measure			Estimate	Standard Error	CV	Estimate	Standard Error	CV	Estimate	Standard Error	CV	Estimate	Standard Error	CV
Poverty incidence FGT(0)	Survey		54.10	2.20	0,041	60.60	2.30	0,038	74.30	2.60	0,035	54.40	2.50	0,046
	Census*		54.10	1.69	0,031	63.80	1.57	0,025	74.50	1.84	0,025	55.50	1.69	0,030
	Census/Biomass		53.42	1.25	0,023	63.40	1.48	0,023	74.80	1.07	0,014	55.40	1.37	0,025
Poverty gap FGT(1)	Survey		18.60	1.20	0,065	23.00	1.30	0,057	28.30	1.90	0,067	19.80	1.40	0,071
	Census*		17.90	0.84	0,047	23.90	0.93	0,039	30.30	1.10	0,036	20.30	1.02	0,050
	Census/Biomass		17.85	0.71	0,040	23.90	0.93	0,039	32.00	0.70	0,022	20.10	0.77	0,038
Poverty gap squared FGT(2)	Survey		8.80	0.70	0,080	11.40	0.80	0,070	14.40	1.30	0,090	9.60	0.90	0,094
	Census*		8.10	0.73	0,090	11.70	0.60	0,051	15.60	0.72	0,046	10.00	0.91	0,091
	Census/Biomass		8.02	0.44	0,055	11.70	0.60	0,051	17.05	0.59	0,035	10.04	0.48	0,048
GINI	Survey		0.33	0.01	0,030	0.32	0.01	0,031	0.33	0.01	0,000	0.31	0.01	0,032
	Census*		0.31	1.81	5,839	0.32	0.84	2,625	0.31	0.89	2,871	0.34	1.72	5,059
	Census/Biomass		0.32	0.07	0,021	0.45	0.82	1,822	0.36	0.48	1,333	0.31	0.70	2,258
Mean Per Capita Expenditure	Survey		18131	629	0,035	15460	486	0,031	13899	636	0,046	16256	537	0,033
	Census*		17951	564	0,031	15049	382	0,025	12884	370	0,029	16954	509	0,030
	Census/Biomass		18202	345	0,019	19629	4073	0,207	13755	365	0,027	16210	314	0,019

* The 'Census' poverty measures are derived from Okwi et al., (2003). The 'Census' and 'Census/Biomass' estimates are predictions based on the ELL method, while the 'Survey' estimates are directly calculated from the IHS survey.

Table D1 Appendix D presents the poverty estimates at district level. These poverty estimates show some level of heterogeneity. All the standard errors fall below the stratum level survey based ones with the exception of Kalangala district in central region. The case for Kalangala district is an interesting and expected one. First, this is a small district with a total population of 14, 218 people which is significantly less than the population of any most sub-counties and even parishes in the region. For example, in Central region, the poverty estimates range from 25 percent to 63 percent at the district and 19.6 to 74 percent at the county. In Eastern, the poverty levels range from 39.5 to 82 percent at the district level. At the county level, the observed distribution is more interesting than at the district level. In the North, Arua is the least poor district (64 percent) while Kotido is the poorest with 91 percent poor. Similarly, Western region shows significant variation in poverty levels. Whereas Masindi has about 76 percent headcount ratio, Mbarara is the least poor with only 43 percent. Generally, there is wide variation in the poverty estimates in all the strata and we cannot categorically identify one region as being the poorest as there may be pockets of wealthy areas within the poorest region. The level distributions of poverty at various levels are shown in the graphs in Appendix F.

4.2 How well do (re-weighted) panel and survey estimates match at stratum level?

In Table 3 stratum-level welfare estimates for 1992 and 1999/2000 derived from respectively the IHS and from the re-weighted panel households (including biomass) are presented. For the IHS official estimates are presented. The table also presents, in the last column, census based predictions including biomass data for 1999/2000. Poverty maps relaying this information are presented in the appendices.

While the poverty maps in the appendices are useful in identifying the spatial patterns of poverty, Table 3 provides more detail, including the standard errors of the poverty estimates for each stratum. This table illustrates a number of points. First reweighting the IHS to adjust for household non-response does not affect the poverty estimates in a significant way. It should not be inferred from this that re-weighting is

superfluous.⁹ This depends on the research question. For instance, if the interest is in the fraction of non-poor living in small households then re-weighting makes a significant difference (at the 95% level of confidence) by increasing the fraction from 39.3 to 45.1%.

Secondly both for 1992 and for 1999/2000 we cannot reject at the 95% confidence level that the stratum-level poverty estimates derived for the panel households are the same as those derived for the complete surveys. This provides confidence in the post-stratification re-weighting procedure that was followed to assure the representativeness of the panel households. In combination with the large number of variables that passed the means comparison test, it provides a solid basis for deriving census based poverty estimates from the panel households. Unsurprisingly given the small number of panel observations, the strata-level standard errors based on panel data are considerably larger than those reported for either the IHS or the UNHS.

In addition, we analyze the extent to which the inclusion of spatial features can allow our poverty estimates to be robust. There are two major ways of determining the level of dis-aggregation at which the error becomes too big. They both yield similar conclusions in most cases. One way to approach this is to consider the absolute level of the standard error. The other method, which is used in this study, is to calculate the coefficient of variation (CV), which is the ratio of the standard error over the point estimate for each administrative unit and compare this with the survey-based ratios.

The inclusion of biomass variables has improved the standard errors (in some cases by upto 40 percent) of our estimators at the stratum level, except for the inequality¹⁰ index which are consistently lower than those obtained from previous analysis excluding biomass data (Okwi et al., 2003) and the household survey data alone.

⁹ The absence of any impact of reweighing on the poverty indicators can be traced to two aspects: (i), the fraction of poor one and two person households is small; and (ii) even after reweighing, members from small households make up only between 8% and 9% of the total population.

¹⁰ Similar results are obtained from other studies; see for example Mistiean *et al.* 2002 and Okwi *et al.*, 2003 and this is an expected result given the way inequality is measured.

Much higher inequality is observed in all the strata and overall, the standard errors for the inequality

Table 3: Poverty estimates for 1992 and 1999/2000 (Including environmental data)

		1992	1999/2000		
		IHS, official	UNHS official	Panel & census	Panel, Census and Environ. data
Poverty Incidence - FGT(0)	Central rural	54.3	25.7	24.5	23.50
		-2.2	-1.4	-1.5	-1.39
	East rural	60.6	38.4	34.3	31.09
		-2.3	-1.6	-2.4	-1.72
	North rural	73	67.7	66.5	69.35
		-2.9	-3.8	-2.2	-2.55
FGT(1)	West rural	54.3	29.5	31.7	33.60
		-2.4	-1.9	-1.5	-1.41
	Central rural	18.7	5.9	6.4	5.97
		-1.2	-0.4	-0.5	-0.47
	East rural	23	10.5	9.5	8.14
		-1.3	-0.6	-0.9	-0.61
Poverty Gap	North rural	29	26.4	27.2	29.27
		-2	-2.9	-1.5	-2.01
	West rural	19.2	7	10.8	9.61
		-1.3	-0.6	-1	-0.58
	Central rural	8.8	2	2.5	2.31
		-0.7	-0.2	-0.3	-0.22
FGT(2)	East rural	11.4	4.2	3.9	3.21
		-0.8	-0.3	-0.4	-0.29
	North rural	14.8	13.3	14.3	15.77
		-1.3	-2	-1.1	-1.48
	West rural	9.3	2.4	5.7	4.03
		-0.8	-0.2	-0.9	-0.30
Gini Coefficient	Central rural	0.329	0.313	0.31	0.30
		-0.01	-0.02	-0.01	-0.92
	East rural	0.321	0.31	0.295	0.28
		-0.01	-0.01	-0.01	-0.73
	North rural	0.33	0.314	0.39	0.40
		-0.02	-0.01	-0.03	-4.13
Per capita consumption	West rural	0.309	0.283	0.332	0.31
		-0.01	-0.01	-0.01	-0.77
	Central rural	18046	26423	26815	26588
		-638	-862	-784	-713
	East rural	15427	21219	21739	22307
		-480	-605	-701	-503
	North rural	13663	14095	15906	15314
		-632	-773	-972	-1661
	West rural	16368	22839	23249	22395
		-500	-528	-550	-488

Notes: The columns IHS (and UNHS) official present welfare estimates as released by UBOS. IHS reweighted adjusts the IHS sampling weights for household non-response. The column panel & census presents updated small area welfare estimates excluding environmental information. The columns UNHS panel and panel, census and biomass provide welfare estimates derived for the set of 1058 panel households. Standard errors are in parentheses and are corrected for survey design.

index have increased. However, technically we cannot explain this as a causal relationship but an association model.

It is useful to look at the changes in poverty using the updated census based, welfare estimates for 1999/2000. The last column of Table 3 shows that the stratum level sample survey estimates of poverty, the poverty gap and the poverty gap squared are closely replicated by the updated census based estimates after inclusion on biomass information¹¹. The size of the standard errors is generally smaller than the standard errors derived for IHS and is of similar magnitude as to those reported for the UNHS. The updated estimates including the biomass information not only replicate the poverty estimates well, the imputed per capita household expenditure is in all strata within the 95% confidence interval of the household survey. This reflects the fact that the distribution of explanatory variables is similar in the IHS panel and the 1991 census, and pays tribute to the care with which comparable variables have been identified.

Finally, this section offered insights about the inclusion of biophysical and other spatial features in poverty estimation. It demonstrated that relative improvements can be made in the estimation of welfare – with the inclusion of more explanatory spatial characteristics. That is, by controlling for biophysical characteristics at the estimation procedure, the efficiency of the derived poverty estimates may be improved, leading to more precise estimates and enhancing the level of spatial disaggregation that is attainable. Awareness of this association, combined with well designed policies are key factors that may support poverty reduction in these areas.

Comparing poverty in 1992 and 1999/2000

District level welfare estimates for Uganda are presented in the appendix. The 1992 estimates are copied from Okiira Okwi et al. (2003) and the 1999/2000 estimates are derived with the updating methodology.¹² From the discussion in the previous

¹¹ An exception holds for West rural where the poverty gap and its square differ significantly from the survey estimate.

¹² An issue requiring further investigation is that the standard errors for the 1992 and 1999 estimates are not independent as they are derived from the same census. Correlation may come through the

section it is clear that the 1999/2000 district estimates have to be interpreted with care. Though the 1999/2000 expenditure model performs better than the model estimated for 1992 in that the standard errors on the welfare estimates are smaller and that the stratum level estimates from the survey are more closely replicated, considerable divergence from the actual (but unknown) estimates is a real possibility. Keeping this caveat in mind but realizing that the results are correct on average, one could still consider changes and trends. This is possible because the expenditure aggregates that were calculated using the 1992 IHS and the 1999/2000 UNHS are comparable (Appleton, 2002). The results confirm the sample survey results mentioned in the introduction, that the drop in poverty incidence was highest in the Central region (where it dropped by 30 percentage points), and lowest in Northern Uganda where poverty dropped 8 percent points.

The census/biomass based estimates allow, unlike the survey, to consider changes in poverty at administrative levels below the stratum. So whereas the survey presents evidence that poverty declined in all regions, Maps in the Appendix F illustrate how the drop in poverty was widely distributed across districts: poverty dropped in almost all districts as well. There are three districts (Apac, Moyo and Kasese) in which poverty might have increased. The three districts have all been affected by insurgency in the 1992-1999 period so that it is plausible that poverty did not decline during the 1990s (the increases are not significantly different from zero). There is considerable within region variation in poverty incidence and reduction. For instance, in the Central region there are districts where the drop in poverty between 1992 and 1999/2000 was ‘only’ 15 percentage points, but there are also districts such as Rakai where the drop was close to 40 percent points.

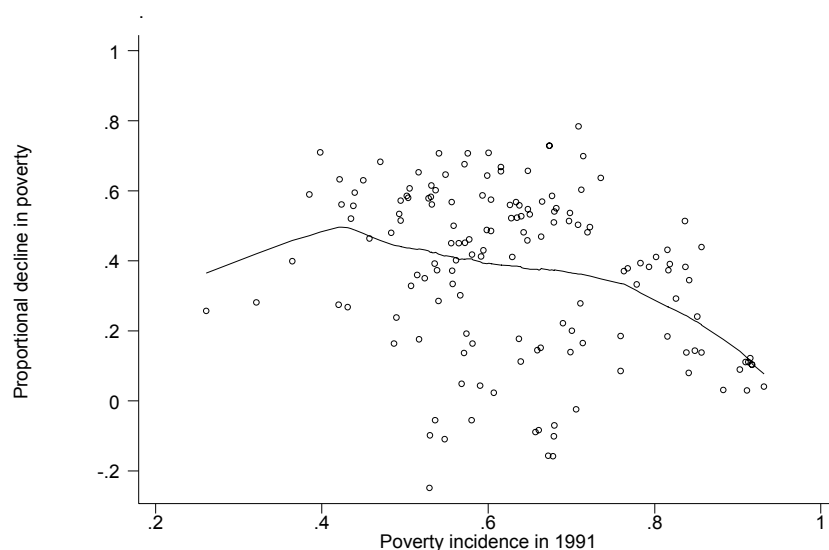
Figure 1 returns to the question raised in the introduction. Has poverty declined most in areas with lower initial levels of poverty? The figure presents a scatter plot with the proportional decline in poverty at the county level on the vertical axis and initial

$\hat{\beta}'s$, ε_{ch} 's and the cluster effects η_c 's. To control for this, simultaneous estimation of the 1992 and

1999 would be needed. The importance of correlation is likely to be limited, however, because the panel households are a small subset of the full IHS, because the various consumption models comprise different variables and because the cluster is defined at the PSU for 1992 and the county for 1999.

poverty on the horizontal axis. The scatter plot shows little in terms of a relation between the changes in poverty and initial poverty levels. The line however, which is a locally weighted smoothed function of the decline in poverty, suggests that areas with the highest levels of initial poverty did worst in terms of poverty decline. This is an alarming result as it means a growing divergence between Uganda's poorest and better off regions. The finding is, however, indicative at best. The negative slope may, for instance, have been brought about by measurement error. In the absence of any real correlation between poverty reduction and initial levels of poverty, a negative correlation would be found if the 1992 level of poverty was measured with error. Even if the negative slope is not brought about by measurement error, one still has to investigate whether the relation is statistically significant. Such an investigation is beyond the scope of this paper, as it requires taking into account that both right and left hand side variables are estimates with an standard error (but see Elbers, Lanjouw and Lanjouw 2003b on this issue).

Figure 5: Decline in poverty (as fraction of initial poverty) and initial poverty in 1992



5.0 Welfare and the environment in Uganda

5.1 Using maps to show the link between changes in poverty and the environment in Uganda

There have been attempts to link poverty to other socio-economic factors that do not follow administrative boundaries (e.g ILRI 2002), suggesting that combining poverty with other information (in this case on livestock) is key for a convincing integrated framework to address poverty issues for pastoralist populations. For Uganda, where most households are involved in agriculture, this finding motivates our attempt to combine poverty and environmental information. Further, to explain the link between certain bio-physical characteristics and poverty, we use overlays presented in Appendix F. The overlays are simply meant to provide a visual explanation of the relationship between poverty and land-use. The overlays generally have helped us to answer the following questions: Where are the poor? (Appendix F) Which poor (rich) areas have similar types of land-use features? Which areas provide which type/amount of ecosystem services? How do the land-use types overlap with poverty? How does the location of poverty compare to the distribution of ecosystem services? This information may help policymakers to design effective policies to improve the situation? For detailed maps, see the poverty and biomass maps for all strata in Appendix F.

a) Poverty and land use in 1992

Figures in the appendix F enable us to identify the poverty hotspots and correlate them with the type of land use in the area. According Appendix F, poverty incidence were higher in the North and Northeast. The type of land use in these areas is typically grassland and woodland. Economically, grasslands do not provide high returns to households and most of the households found in the grasslands are pastoralists. It is therefore not surprising that the areas of the north are generally poor. These areas are also characterized by poor climate and relatively less fertile soils compared to the Central region. The parts of the Northern region that show less poverty are those situated next to Lake Kyoga. These areas generally have low density of poverty and are generally wetlands. Typically, wetland farming (rice) is taking place in this area and this could explain the fact that households in this area are less poor. In Uganda, rice growing is becoming a major income source for households living near the wetlands.

Another picture that emerges from the north is that poverty is more pronounced in the parts which are typically wooded and grassland areas and less pronounced in the degraded lands of all the regions. The implication of the later result is that the poor are actually using the ecological resources to improve their welfare but in the process they degrade the natural environment as well. However, a contrasting picture emerges from the grassland areas in Western and Northern regions which portray less and more poverty respectively, see also according correlation coefficients with opposite signs in the Table C1-C5 in Appendix C. A question that emerges is why the difference? Possible explanations for the difference could be because the pastoral lands in Western Uganda have been modified by the people to produce high yielding varieties thus directly improving their welfare, while the pastoralists in the North are still held with the traditional norms of cattle rearing. In addition,

The Eastern and central region portray another interesting picture. The biomass map shows considerably more degradation in the areas surrounding Lake Victoria and the Mabira forest. The poverty map, however, shows that these areas relatively less poor (30-40 percent) compared to the areas in the same region. These maps reveal how land use (degradation) could be helping reduce poverty among rural households living along the Mabira forest and Lake Victoria. It should be noted that this explanation does not imply causality. Similarly, the land use map shows that areas that have typically high subsistence farming are generally poorer than the degraded areas.

b) Updated welfare and the environment Indicators

The poverty mapping method also generates estimates for changes in poverty and the environment. It is important to use these results with caution because the small number of panel households in some areas means that poverty estimates are not very reliable. In this section, we show how changes in poverty between 1991 and 1999 are related to changes in land use over the same period.

The spatial patterns in district and county poverty rates are shown in Appendix F. This maps provides considerably more detail than the regional poverty map. The

results from the analysis of poverty changes are encouraging, with large and widespread decreases in poverty seen countrywide. These trends should be viewed as indicative only, as cautious interpretation of the 1999 estimates is required due to the relatively small number of households surveyed in 1999. The 1999 maps will be updated to 2003 soon, making use of the new census data. The highest drops in poverty in rural areas between 1992 and 1999 can be seen in Central and parts of Western region in the districts of Kibaale, Luwero, Bushenyi, Rakai, Mpigi and Kisoro. Poverty was observed to have increased in Arua, Moyo and Apac in Northern region and Kasese district in Western. At the county level, the maps demonstrate how almost all rural areas in Uganda benefited from the growth that took place during the 1990's. Poverty worsened in 8 percent of Uganda's rural counties during this period. In terms of inequality, increasing inequality was observed in Northern region and some districts in Western region including Masindi, Kasese and Bundibugyo.

The maps showing how poverty has changed at the county level between 1991 and 1999 can be related to the changes in the environment. Appendix F typically shows which areas have had major changes in land use. With the exception of selected areas in the four regions, all the other districts and counties in Uganda have not experienced major changes in land use.

REGIONAL ROUND-UP

Central region: stood out as the least poor region in 1992 and 1999 for both rural and urban areas. However, the land use maps show increasingly more degraded areas. The region is mainly covered with subsistence farmlands which have increased in proportion compared to the total land area. Central region is the main coffee growing area in Uganda and has benefited from the rapid growth in coffee production during the 1990's. However, as can be seen from the maps, the areas that have experienced increases in degradation (forest) also have the least poor populations. Similarly, areas that are near the Lake, mainly wetlands, have experienced far more declines in poverty than the others. This relationship points to reclamation of wetlands and degradation of forests during the period 1991-1999. A relatively large population is involved in fishing in this area

Eastern region: With a rural population of 3.7 million people and 0.3 million found in urban areas, this region demonstrated the widest variability in poverty levels. Jinja district had the lowest poverty (38 percent) in 1992 while Kumi had the highest at 82 percent. County level variations were even higher. Like in the Central region, land use mainly changed in favour of subsistence farmlands and subsistence wetlands. Degradation was highest in the wealthy counties near Jinja. Poverty remained high in the grassland and wooded areas of Kumi, Katakwi and Soroti districts. However, areas near Mt Elgon experienced increased degradation and decreased poverty, an indication that the population in these areas are harnessing the forest resources from the Mt. Elgon reserve to improve their welfare.

Northern region: With over 75 percent of the population poor in 1992, this region remained the poorest region in Uganda in 1999. The poorest districts were Kotido and Kitgum with poverty incidences of 91 percent while Arua and Lira stood out as the least poor districts. There was significantly more variation in poverty in this region at both the district and county levels. This region, in contrast is generally wooded and grassland with a few pockets of wetlands. A few counties have poverty below 60 percent and the generally state of the environment has not changed much since 1991. The high incidence of poverty in this area is due to the fact that this is one of the most semi arid parts of Uganda, and the sandy soils make it difficult to practice intensive agriculture. This area is generally poorly served with roads and therefore access to markets is difficult. A relatively small population is involved in fishing in Lake Kyoga and River Nile. The fishing areas generally show improvements in welfare.

Western Region: This region ranked the second least poor in Uganda. More than half the rural population and one third of the urban population lived below the poverty line in 1992. Rural poverty was highest in Kisoro and lowest in Mbarara district. In 1999, there was a lot of variation in poverty incidence in this region. Masindi, Bundibugyo and Kasese had greater than 50 percent poverty incidence while relatively wealthy districts such as Mbarara and Bushenyi had poverty levels below 20 percent. This region showed the highest declines in poverty in the 1990's. The area generally has a mix of subsistence farming and cattle rearing. More areas have been reclaimed from grasslands into farms. However, there are pockets of high

degradation between 1991 and 1999 in the North-western parts of the region. These are areas close to the mountainous parts of Rwenzori with difficult access to roads and markets. Areas near the mid western have benefited from flat land and improved transportation (roads), all of which reduce poverty rates.

As mentioned earlier, the estimates of changes in poverty and land-use must be interpreted with caution. For the 1999 poverty rates, there were relatively a small number of households included in the panel, leading to relatively high margins of error in the poverty estimates. Similarly, the changes in land use are not bound by district and county boundaries and therefore subject to some measurement error. As indicated earlier, land use does not necessarily confine itself to administrative boundaries.

Finally, two notes of warning about putting small area welfare estimates on the map. This paper has placed considerable emphasis on the fact the census based poverty estimates are associated with a standard error. The maps do not reflect this, and in various instances counties that are classified differently on the map, have means for which a t-test cannot reject that they are identical. Next, poverty incidence is just one way to report poverty. Instead of reporting the fraction of poor, a geographic profile of welfare could also take into account land area and report poverty density –i.e. the number of poor per square kilometre. If one were to do so the geographic poverty profile becomes very different, with poverty being least an issue in the North and being most urgent near the Rwandan border in the South West and South of Mt Elgon in the East.

6. Conclusions and Implications for Policy

This study combines census, survey and bio-physical data to generate spatially disaggregated poverty/biomass information for rural Uganda. It makes a methodological contribution to small area welfare estimation by exploring the inclusion of bio-physical information. By combining the generated poverty estimates with national biophysical data, this study explores the contemporaneous correlation between poverty (welfare) and natural resource degradation at a level of geographic detail that has not been feasible previously. In this welfare estimation method,

association relationships are used to explain welfare rather than causal relationships are explored. However, the resulting estimates of poverty measures have improved by the inclusion of bio-physical information. In some cases the levels of poverty measures have changed. For North Uganda, the poverty gap and poverty gap squared increased compared to the estimates without biophysical information.

By providing comparable welfare and biophysical information for many data points, this study solves many problems faced by many previous studies. For instance, previous studies (see Atkinson and Brandolini, 1999) on poverty and the environment were based on case studies which are unrepresentative. This study presents results of a representative sample and population. Secondly, previous studies have also been cross-sectional thus raising data incomparability problems. By using data from one country and collected by the same institution, with comparable questions in the questionnaires and within a period of time less than 2 years, data incomparability problems are solved. Thirdly, this study has provided a practical analysis of the link between welfare and the environment. Other studies have only looked at the theoretical link between poverty and environmental degradation (Ambler 1999; Barbier, 2000; Roe, 1998; Chomitz, 1999; Ekbom and Bojo, 1999). This study has shown that accounting for spatial differences in welfare is key to high precision maps and explaining poverty environment relationships.

The poverty estimates appear to be more robust, as the standard errors show a decline in some cases by upto 40 percent. Moreover, the coefficient of variation, that is, the ratio of the standard error and the point estimate decline in general as well. Overall, we conclude that the estimates of the poverty measures are more robust when biophysical information is taken into account. Part of the output from this study are maps showing poverty and biomass overlays for Uganda. These maps can be used as a planning tool and for targeting purposes.

Updating requires panel data and estimation of an updated poverty map and will typically be done on a smaller survey data set than the one used to generate the poverty map for the census year. In the case of Uganda, the 1992 rural poverty map is based on a survey with 6,396 observations, whereas the updated map is based on

1,058 observations. This has implications. Updated welfare estimates for urban areas are not derived and the estimation procedure had to be adjusted. For instance one expenditure model with regional interaction terms was estimated instead of one for each of the four rural strata; district dummies could not be used because not all districts were represented in the panel and indicators of ethnicity obtained from the census were used instead. These deviations from the preferred poverty mapping methodology require careful scrutiny of the generated welfare estimates. Fortunately, in a typical case where a poverty map is updated, small area estimates already exist for the census year. The second important result from this exercise is that one should not only estimate an updated poverty map for the year of interest, but an ‘updated’ map for the census year should also be generated. The comparison of the updated census year map, with the actual poverty map for the census year, allows checking the accuracy of the method. Together with the R^2 of the updated expenditure model and the accuracy with which stratum level welfare estimates from the sample survey are replicated, it guides the decision on how to use updated small area results.

In terms of policy, by implication, any policy focused on improving access to roads is directly related to the welfare of the poor. Similarly, policies focused on conservation of wetlands and forests, improvement of grasslands (mainly pasture land), and access to water could be important policy issues to consider in understanding the relationship between poverty and the environment. Given that most of Uganda’s rural population depends on agriculture and the environment, and considering the spatial relationship between subsistence farming, degraded lands and poverty, the results suggest that focusing on improving production in the subsistence sector may prove important in reducing poverty and improving the biomass conditions. The results from the regression analysis clearly display regional upto county level variation in spatial correlation between bio-physical and poverty information and therefore imply region specific policy designs. Finally, in future research, with more information, the causal relationship will be analysed in more detail. Another conclusion that we reached is that without further verification the updated results should not be used as indicators for the welfare in specific sub-counties, counties or districts.

7.0 References

- Ambler J.(1999). Attacking Poverty while improving the environment. Towards Win- Win Policy Options. Background technical paper prepared for the September 1999 Forum of Ministers Meeting, under the UNDP-EC Poverty and Environment Initiative .
- Appleton, S., Emwanu, T., Kagugube, J. and Muwonge, J. (1999) "Changes in poverty in Uganda, 1992-1997". Centre for the Study of African Economies, Oxford Working Paper Series, WPS 99.22
- Appleton, Simon (2001), "Poverty in Uganda, 1999/2000: Preliminary estimates from the Uganda National Household Survey," University of Nottingham.
- Angelsen, A. and D. Kaimowitz. (1999), "Rethinking the Causes of Deforestation: Lessons from Economic Models," *The World Bank Research Observer*, Vol. 14, No.1: 73-98.
- Atkinson A.B. and A. Brandolini (1999). "Promises and pitfall in the use of "secondary" datasets: income inequality in OECD countries as case study". *Journal of Economic Literature* 39(3): 771-799.
- Barbier E. (2000). "The Economic Linkages between Rural Poverty and Land Degradation: Some Evidence from Africa". *Agriculture, Ecosystems and Environment Vol 82: 355-370*
- Besley T. and R. Kanbur, (1993). "The Principles of Targeting". In M.Lipton & J.van der Gaag, Including the Poor, proceedings of symposium organized by the World Bank and the International Food policy Research Institute (IFPRI). Regional and Sectoral Studies, Washington D.C: the World Bank
- Chomitz K. (1999). "Environment Poverty Connections in Tropical Deforestation". Discussion Notes prepared for the WDR workshop on Poverty and Development. Washington DC. July 6-8.
- Datt G. and Ravallion, M. (1998) "Why Have Some Indian States Done Better Than Others at Reducing Rural Poverty?", *Economica*, Vol. 65, No. 257, Feb., 1998, pp.17 -38.
- Deaton A. (1997). *The Analysis of Household Surveys: A Microeconomic Approach to development policy*. Baltimore, MD: Johns Hopkins University Press.
- Demombynes, G., Elbers, C., Lanjouw, J.O, Lanjouw, P., Mistiaen, J.A. and Ozler, B. (2002) "Producing an Improved Geographic Profile of Uganda:

Methodology and Evidence from Three Developing Countries”. Discussion paper 2002/39, WIDER, Helsinki, Finland.

Ekbom A. and J. Bojo (1999). “Poverty and Environment. Evidence of Links and Integration in the Country Assistance Strategy Process”. World Bank. Africa Region Discussion Paper no. 4 World Bank. Washington DC

Elbers C., Lanjouw, J.O and Lanjouw, P. (2002). “Welfare in Villages and Towns: Micro level estimation of Poverty and Inequality”. Policy Research Working paper, World Bank: Washington D.C.

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

Forest Department, 1988. National Biomass Project Document. Kampala, Uganda

Forest Department, 1992. National Biomass Study Phase I Technical Report. Kampala, Uganda.

Forest Department, 1992. National Biomass Project Review. Kampala, Uganda.

Forest Department, 1994. National Biomass Study Evaluation Mission Report. Kampala, Uganda.

Forest Department, 1995. National Biomass Review Mission Report. Kampala, Uganda.

Forest Department, 1996. National Biomass Study Phase III Project Document. Kampala, Uganda.

Forest Department, 2002. National Biomass Study Final Report. Kampala, Uganda.

Foster, J. Greer, J. and Thorbecke, E. (1984). “A Class of Decomposable Poverty Measures”, *Econometrica*, 52, pp. 761-66

Glewwe P. (1990). “Efficient allocation of transfers to the poor. The problem of unobserved household income”. Working paper No.70 Living Standards Measurement study. Washington D.C: the World Bank

Glewwe P. and J. van der Gaag, (1990). Identifying the poor in developing countries: Do different definitions matter? *World Development*, 18 (6).

Government of Uganda, (1991). “Uganda Population and Housing Census” Uganda Bureau of Statistics

Hentchel J., Lanjouw, J. O. Lanjouw p. and Poggi, J. (2000). “Combining Census and Survey Data to Trace Spatial Dimensions of Poverty: A Case Study of Ecuador”, *World Bank Economic Review* 14 (1) 147-65 Washington D.C: The World Bank

- Hoogeveen, J. G., T. Emwanu and P. Okiira Okwi 2004. Updating Small Area Welfare Indicators in the Absence of a New Census. Mimeo.
- International Livestock Research Institute (2002), Mapping Poverty and Livestock in East Africa. ILRI publications.
- Jones, D.W. and R.V. O'Neill (1995), "Development Policies, Urban Unemployment, and Deforestation: The Role of Infrastructure and Tax Policy in a Two-Sector Model," *Journal of Regional Science* 35:135-53.
- Kant, S. and A. Redantz (1997), "An Econometric Model of Tropical Deforestation," *Journal of Forestry Economics* 3: 51-86
- Machinjili, C. and Benson, T. (2002), "Malawi: An Atlas of Social Statistics" National Statistics office, Malawi
- Minot, N. (2000). "Generating Disaggregated Poverty Maps: An Application to Vietnam". *World Development* 28 (2).
- Mistiaen J. A., Ozler, B., Razafimanantena, T. and Razafindravonona, J. (2002). Putting "Welfare on the Map in Madagascar". The World Bank: African Region *Working Paper Series no. 34*.
- Moller, L. (2002). A practical guide to developing good poverty indicators. Based on Uganda's experience. *Mimeo*.
- NEMA (2002). National Environment Management Authority. *State of the Environment Report*, Uganda.
- Okwi, P.O., and Kaija, D. (2000). "The Distribution of Welfare in Uganda". Eastern Africa Social Science Research Review (Vol. XVI, No. 2, June 2000).
- Okwi, P.O., Emwanu, T and Hoogeveen, J.G. (2003). Poverty and Inequality in Uganda: Evidence from Small Area Estimation Techniques. Unpublished
- Ravallion, M. and Wodon, Q. (1997). "Poor areas, or only poor people?" *Policy Research Working Paper* No. 1798. Washington D.C: the World Bank
- Roe, E. (1998). Taking Complexity Seriously. Policy Analysis, Triangulation and Sustainable Development. Kluwer Academic Publishers: Boston USA
- Uganda Bureau of Statistics, (1991). "Uganda Population and Housing Census" Government of Uganda
- Uganda Bureau of Statistics (2001). *Statistical Abstract*. Government of Uganda.
- Uganda Bureau of Statistics (2002). *Statistical Abstract*. Government of Uganda

Uganda Bureau of Statistics (2003). *Statistical Abstract*. Government of Uganda

Wodon Q. (1997). “Targeting the poor using ROC curves”. *World Development*, 25 (12)

World Bank (2002), *World Development Report*. New York: Oxford University Press.

Appendices: Facts and Figures

This appendix is an additional report of Birungi *et al.* (2005). It contains a number of tables which accompany the results as presented in Birungi *et al.* (2005). Appendix A present the list of variables used in the analysis of preparing poverty estimates. The small area estimation approach relies data at different aggregation level, such as household surveys and census data. Therefore, the mean values of variables of both data sources are compared on a statistical basis, i.e. the zero stage comparison between the means of variables from the Integrated Household Survey 1992 in Uganda, and the 1992 census. Appendix B presents the results of the comparison tests while appendix C presents the first stage results for the cross section and panel analysis. Appendix D presents the poverty estimates at county level while appendix E presents the correlations and comparison of new and old estimates. Finally, Appendix F presents the poverty and environment overlays using 1992 and 1999 data.

A. List of variables

Table A.1 *List of variables from the Integrated Household Survey (HIS, 1991) for which Census information is available*

<i>Variables</i>	<i>Subcategories</i>
Expenditures	Total expenditures of the household
Household composition	
Household size	
Adult equivalents of household	
Relationship to head	Head of household Spouse Child
Sex	
Age	
Marital status	divorced separated single widowed married
Education	
Level	P1-P4 P5-P7 none O'level and higher below O'level primary school secondary school number of years
School attendance	
Literacy	
Education deficit	
Occupation	
Main occupation of household head	disabled employed household work clerical worker too old other self employed student unpaid family worker
Industry of main occupation	
Housing	
Type of housing unit	Housing type detached Housing type hut Housing type servant's quarters Housing type other

Table A.1 *List of variables from the Integrated Household Survey, 1991 (continued)*

<i>Variables</i>	<i>Subcategories</i>
Housing (continued)	
Number of rooms	
Livelihood subsistence farming	
Type of tenure of dwelling unit	Tenure own Tenure free Tenure other
Type of wall material	Wall burnt bricks Wall cement Wall mud Wall stone Wall unburnt bricks Wall wood Wall other
Roofing material	Roof asbestos Roof cement Roof other Roof thatched Roof tiles
Floor material	Floor burnt bricks Floor cement Mud floor Floor stone Floor wood Floor other
Type of kitchen	Kitchen inside and exclusive Kitchen outside and exclusive Kitchen shared Kitchen none
Type of fuel for cooking	Cooking with charcoal Cooking with electricity Cooking with gas Cooking with paraffin Cooking with wood Cooking with other fuel
Type of toilet	Toilet flush Toilet pit Toilet none Toilet other
Presence of bath	
Water availability	Water tap Water other
Water quality	Water safe Water unsafe
Type of fuel for lighting	Lighting electricity Lighting paraffin Lighting other

Table A.2 Land use covers in Uganda, 1991

Land use	Area in hectares	Proportion
Plantations Hardwoods – deciduous trees/broadleaves (hardwood)	18,682	0.1%
Plantations Softwoods- coniferous trees	16,384	0.1%
Tropical high forest (THF)- Normally stocked	650,150	2.7%
Tropical high forest (THF) – Degraded/depleted	274,058	1.1%
Woodlands – trees and shrubs (average height > 4m)	3,974,102	16.5%
Bush lands - bush, thickets, scrub (average height < 4m)	1,422,395	5.9%
Grasslands –rangelands, pastureland, open savannah including scattered shrubs and thickets	5,115,266	21.2%
Wetlands – wetland vegetation; swamp areas, papyrus and other sedges	484,037	2.0%
Subsistence Farmlands –mixed farmland, smallholdings in use or recently used, with or without trees	8,400,999	34.8%
Commercial Farmlands – mono cropped, non seasonal farmland usually without any trees for example tea and sugar estates	68,446	0.3%
Built up areas – urban or rural build up areas	36,571	0.2%
Water – Lakes, rivers and ponds	3,690,254	15.3%
Impediments – bare rocks and soils	3,713	0.0%
Total	24,155,058	100.0%

Source: National Biomass Study (Forest Department, 2002), Uganda.

Table A.3 Distances to different kind of roads, 1991

Stratum
Proportion of a stratum within 5 km distance from a main road
Proportion of a stratum within 4 km distance from a main road
Proportion of a stratum within 3 km distance from a main road
Proportion of a stratum within 2 km distance from a main road
Proportion of a stratum within 1 km distance from a main road
Proportion of a stratum within 5 km distance from a tarmac road
Proportion of a stratum within 4 km distance from a tarmac road
Proportion of a stratum within 3 km distance from a tarmac road
Proportion of a stratum within 2 km distance from a tarmac road
Proportion of a stratum within 1 km distance from a tarmac road
Proportion of a stratum within 5 km distance from a track
Proportion of a stratum within 4 km distance from a track
Proportion of a stratum within 3 km distance from a track
Proportion of a stratum within 2 km distance from a track
Proportion of a stratum within 1 km distance from a track

Source: National Biomass Study (Forest Department, 2002), Uganda.

B. Survey and Census comparison

The variables in the household survey and the Census are first compared on definition and categorisation. If a variable of the household survey and the Census match on the basis of definition, the next step is to test whether the household survey means and the Census mean differ significantly. The test is set up as follows. Based on the household survey (SM), a 95% confidence interval is calculated with a lower bound (L95) and an upper bound (U95). If the Census mean is within this confidence interval, the variable is 'accepted', which means that the variable will be included in the first-stage regression of household expenditures.

Summary of the symbols in the tables of this chapter:

CM: Census Mean

SM: Survey mean

L95: Lower bound 95%

U95: Upper bound 95%

A: $A = 1$ if the Census mean of a variable lies within the 95% confidence interval of the Survey mean. Then, the variable is accepted to be included in the first-stage regression, otherwise it is rejected.

Table B 1 Zero stage comparison between census means and household survey means

Variable	Central rural					East rural					North rural					West rural				
	CM	SM	L95	U95	A	CM	SM	L95	U95	A	CM	SM	L95	U95	A	CM	SM	L95	U95	A
Number of males aged 0-5	0.50	0.54	0.49	0.58	1	0.55	0.53	0.49	0.58	1	0.54	0.56	0.52	0.61	1	0.61	0.62	0.58	0.67	1
Number of males aged 6-14	0.60	0.62	0.55	0.68	1	0.60	0.64	0.58	0.70	1	0.65	0.71	0.65	0.78	1	0.68	0.72	0.66	0.78	1
Number of males aged 30-49	0.33	0.30	0.28	0.33	1	0.37	0.36	0.33	0.38	1	0.38	0.36	0.33	0.40	1	0.37	0.38	0.36	0.41	1
Number of males aged 50 and older	0.25	0.24	0.22	0.27	1	0.27	0.26	0.24	0.28	1	0.20	0.19	0.16	0.21	1	0.22	0.24	0.22	0.27	1
Number of females aged 0-5	0.50	0.50	0.46	0.54	1	0.55	0.57	0.53	0.62	1	0.54	0.54	0.50	0.59	1	0.61	0.62	0.57	0.67	1
Number of females aged 6-14	0.57	0.57	0.51	0.62	1	0.57	0.61	0.56	0.66	1	0.63	0.67	0.60	0.73	1	0.67	0.64	0.59	0.70	1
Head male, divorced	0.15	0.14	0.12	0.16	1	0.07	0.06	0.05	0.08	1	0.05	0.05	0.03	0.06	1	0.06	0.05	0.04	0.07	1
Number of males education at least O' level	0.11	0.11	0.08	0.13	1	0.11	0.13	0.10	0.16	1	0.08	0.10	0.07	0.13	1	0.10	0.10	0.08	0.11	1
Number of males with at least secondary school	0.30	0.29	0.24	0.33	1	0.29	0.31	0.26	0.35	1	0.20	0.23	0.18	0.29	1	0.23	0.23	0.19	0.27	1
Adult equivalent size	3.48	3.43	3.28	3.59	1	3.89	3.85	3.68	4.02	1	4.00	3.94	3.75	4.14	1	4.07	4.05	3.88	4.21	1
Highest number of years of education in household	5.81	5.70	5.40	6.01	1	5.70	5.87	5.56	6.18	1	5.12	5.41	5.05	5.77	1	5.26	5.34	5.06	5.63	1
Household size	4.41	4.40	4.20	4.61	1	4.89	4.88	4.66	5.10	1	5.01	4.99	4.75	5.23	1	5.17	5.17	4.96	5.38	1
Household size squared	29.53	29.00	26.28	31.72	1	34.98	34.84	30.49	39.18	1	34.39	33.18	29.87	36.48	1	35.66	35.32	32.07	38.58	1
Household size = 1	0.18	0.18	0.16	0.21	1	0.12	0.12	0.10	0.14	1	0.08	0.08	0.06	0.09	1	0.08	0.08	0.06	0.10	1
Household size = 2	0.14	0.14	0.12	0.16	1	0.13	0.13	0.12	0.15	1	0.12	0.12	0.10	0.14	1	0.11	0.11	0.09	0.12	1
Household size = 3	0.14	0.14	0.12	0.16	1	0.14	0.14	0.12	0.16	1	0.14	0.14	0.12	0.16	1	0.14	0.14	0.12	0.15	1
Household size = 4	0.13	0.13	0.11	0.14	1	0.14	0.14	0.12	0.15	1	0.15	0.15	0.13	0.17	1	0.14	0.14	0.12	0.16	1
Household size = 5	0.11	0.11	0.09	0.12	1	0.12	0.12	0.10	0.14	1	0.14	0.14	0.12	0.16	1	0.13	0.13	0.11	0.15	1
Household size = 6	0.09	0.09	0.07	0.10	1	0.10	0.10	0.09	0.11	1	0.12	0.12	0.10	0.13	1	0.12	0.12	0.10	0.13	1
Household size = 7	0.07	0.07	0.05	0.08	1	0.08	0.08	0.06	0.09	1	0.09	0.09	0.07	0.10	1	0.09	0.09	0.08	0.11	1
Household size = 8	0.05	0.05	0.03	0.06	1	0.05	0.05	0.04	0.07	1	0.06	0.06	0.04	0.08	1	0.07	0.07	0.06	0.08	1
Household size = 9	0.03	0.03	0.02	0.04	1	0.04	0.04	0.03	0.05	1	0.04	0.04	0.03	0.05	1	0.05	0.05	0.04	0.06	1
Household size = 10	0.02	0.02	0.02	0.03	1	0.03	0.03	0.02	0.04	1	0.03	0.03	0.01	0.04	1	0.03	0.03	0.02	0.05	1
Household size = 11	0.01	0.01	0.01	0.02	1	0.01	0.01	0.01	0.02	1	0.01	0.01	0.00	0.02	1	0.02	0.02	0.01	0.02	1
Household size = 12	0.01	0.01	0.00	0.01	1	0.01	0.01	0.01	0.02	1	0.01	0.01	0.00	0.01	1	0.01	0.01	0.00	0.02	1

Table B.1 Zero stage comparison between census means and household survey means (continued)

Variable	Central rural					East rural					North rural					West rural				
	CM	SM	L95	U95	A	CM	SM	L 95	U95	A	CM	SM	L 95	U95	A	CM	SM	L 95	U95	A
Household size = 13	0.02	0.02	0.01	0.03	1	0.03	0.03	0.02	0.04	1	0.02	0.02	0.01	0.03	1	0.02	0.02	0.00	0.03	1
Log of household size	1.22	1.22	1.17	1.27	1	1.36	1.36	1.31	1.40	1	1.43	1.43	1.38	1.48	1	1.46	1.46	1.42	1.50	1
Proportion of males 0-5	0.09	0.09	0.09	0.10	1	0.09	0.09	0.08	0.10	1	0.10	0.10	0.09	0.11	1	0.11	0.11	0.10	0.12	1
Proportion of males at least 50 years old	0.09	0.10	0.08	0.11	1	0.09	0.08	0.07	0.09	1	0.05	0.06	0.05	0.07	1	0.06	0.06	0.06	0.07	1
Proportion of females 0-5	0.09	0.09	0.08	0.10	1	0.09	0.10	0.09	0.10	1	0.10	0.10	0.09	0.10	1	0.11	0.11	0.10	0.11	1
Pmedusec	0.07	0.07	0.06	0.08	1	0.06	0.07	0.06	0.08	1	0.04	0.05	0.04	0.06	1	0.05	0.05	0.04	0.06	1
maxed13	3.99	4.06	3.76	4.36	1	4.19	4.29	4.03	4.54	1	5.14	5.13	4.72	5.54	1	4.92	4.70	4.42	4.98	1
mned13	1.17	1.17	1.08	1.26	1	1.12	1.20	1.12	1.27	1	1.39	1.41	1.30	1.52	1	1.34	1.28	1.19	1.36	1
maxed132	35.91	37.44	34.56	40.31	1	38.09	39.40	36.76	42.04	1	49.82	49.68	45.34	54.03	1	46.48	45.61	42.68	48.54	1
Lnaesize	1.03	1.02	0.97	1.06	1	1.16	1.15	1.11	1.19	1	1.22	1.21	1.17	1.26	1	1.24	1.23	1.20	1.27	1
pm0_52	0.03	0.03	0.03	0.03	1	0.03	0.03	0.02	0.03	1	0.03	0.03	0.03	0.03	1	0.03	0.03	0.03	0.04	1
pm6_142	0.03	0.03	0.03	0.04	1	0.03	0.03	0.03	0.03	1	0.03	0.04	0.03	0.04	1	0.03	0.03	0.03	0.04	1
pm502	0.06	0.07	0.05	0.08	1	0.05	0.04	0.04	0.05	1	0.03	0.03	0.02	0.04	1	0.03	0.03	0.02	0.04	1
pf0_52	0.03	0.03	0.02	0.03	1	0.03	0.03	0.03	0.03	1	0.03	0.03	0.02	0.03	1	0.03	0.03	0.03	0.03	1
pf30_492	0.03	0.03	0.02	0.04	1	0.03	0.03	0.02	0.03	1	0.03	0.03	0.02	0.03	1	0.02	0.02	0.02	0.03	1
Pmeduola2	0.02	0.02	0.01	0.02	1	0.01	0.02	0.01	0.02	1	0.01	0.01	0.00	0.01	1	0.01	0.01	0.01	0.02	1
Pmedusec2	0.04	0.04	0.03	0.05	1	0.03	0.03	0.02	0.04	1	0.02	0.02	0.01	0.02	1	0.02	0.02	0.02	0.03	1
ped13	3.72	3.80	3.52	4.08	1	3.92	3.99	3.76	4.23	1	4.80	4.75	4.38	5.12	1	4.57	4.35	4.08	4.61	1
nmf0_5	1.00	1.04	0.97	1.10	1	1.09	1.11	1.04	1.18	1	1.08	1.11	1.04	1.17	1	1.22	1.24	1.18	1.31	1
nmf6_14	1.16	1.19	1.08	1.29	1	1.17	1.25	1.16	1.34	1	1.29	1.38	1.26	1.49	1	1.35	1.36	1.27	1.45	1
nm30plus	0.58	0.55	0.51	0.58	1	0.64	0.62	0.59	0.65	1	0.58	0.55	0.51	0.59	1	0.60	0.62	0.59	0.65	1
nm50plus	0.25	0.24	0.22	0.27	1	0.27	0.26	0.24	0.28	1	0.20	0.19	0.16	0.21	1	0.22	0.24	0.22	0.27	1
nmf30min	3.38	3.34	3.16	3.52	1	3.68	3.70	3.50	3.90	1	3.86	3.85	3.65	4.05	1	4.06	4.01	3.83	4.19	1
nmf25min	3.06	3.06	2.88	3.23	1	3.30	3.36	3.17	3.55	1	3.45	3.49	3.30	3.67	1	3.67	3.64	3.47	3.82	1
nmf20min	2.73	2.74	2.56	2.91	1	2.91	2.97	2.79	3.15	1	3.04	3.11	2.93	3.29	1	3.27	3.26	3.09	3.43	1

Source: Authors computations

Appendix C: First Stage results

Table C1. First stage regression: Central

Dependent Variable: log of per capita consumption expenditure

Number of observations: 1660

Number of Clusters: 163

Adjusted R-Square: 0.35

Variable	Parameter	Standard	t Value
	Estimate	Error	
Intercept	10.326	0.138	
Number of females aged 6 to 14 years	0.037	0.017	74.9
Household size squared	0.001	0.000	2.13
Log of Household size	-0.382	0.029	2.38
Prop. of males with secondary education	0.872	0.150	-13.36
Prop. of males with no education	-0.153	0.046	5.82
Prop. of males with A 'level education	0.426	0.136	-3.33
Heads ages squared	0.000	0.000	3.12
Mean years of education head squared	-0.005	0.001	-2.23
Number of females aged 45 plus	-0.056	0.025	-5.55
Buffer zone within 1km main road	0.341	0.078	-2.24
Buffer zone within 2km track road	-0.402	0.116	4.37
Buffer zone within 4km track road	-0.304	0.052	-3.47
Percentage of parish under woodland	0.380	0.144	-5.81
Log of heads age*Alur tribe	0.929	0.267	2.64
Log of heads age*Toro tribe	2.670	0.423	3.47
Log of heads age*Lugbara tribe	0.422	0.183	6.32
Log of heads age* males 30 plus	-0.213	0.036	2.3
Log of heads age* males 30 minimum	0.081	0.012	-5.86
Log of heads age*kitchen shared	0.703	0.198	6.74
Max. years of education*Ganda tribe	0.022	0.006	3.56
Log of heads age*proportion females 0-5 squared	-2.399	0.626	3.57
Log of heads age* Mubende district	-0.062	0.013	-3.83
Log of heads age* Percentage of parish under subsistence farms	0.083	0.020	-4.79
Log of heads age* Percentage of parish under commercial farms	0.183	0.058	4.2
Log of heads age* Percentage of parish under water	0.057	0.026	3.16
Mean household education 18 year* buffer within 5km tarmac	-0.026	0.009	2.21
Mean household education 18 year *Perc of parish comm. farms	-0.346	0.103	-3.08
Prop. Of males with A'level education*Kiboga district	-0.300	0.126	-3.35
Number of males education between P5-P7*prop. Grasslands	0.188	0.047	-2.39
head separated divorced*number of males 30 minimum age	-3.089	1.353	4.03
Heads education between P5-P7*prop. Under towns	2.999	0.766	-2.28
Heads education between P5-P7*prop.degraded forests	1.004	0.249	3.92
Number of males 30 minimum age* Mpigi district	-0.050	0.014	4.04
Japadhola tribe	-2.278	0.556	-3.67
Mugwere tribe	5.369	1.371	-4.09

Table C2. First stage regression: East

Dependent Variable: log of per capita consumption expenditure

Number of observations: 1640

Number of Clusters: 165

Adjusted R-Square: 0.36

Variable	Par. Est	St. error	t value
Intercept	9.379	0.142	66.15
Household size 10	-0.152	0.073	-2.08
Log of adult equivalent	-0.444	0.024	-18.19
Prop. of males with under secondary education squared	0.437	0.139	3.14
Number of Males aged 15 to 29	-0.061	0.018	-3.44
Heads age squared	0.000	0.000	-3.15
Prop. Household members with under O' and A'level education	0.457	0.115	3.96
Proportion of Males with education years 1 to 4 Squared	0.241	0.061	3.96
Buffer zone within 1km tarmac road	-0.255	0.105	-2.43
Percent of parish degraded forests	6.927	1.197	5.79
Percent of parish under commercial farms	4.100	0.706	5.81
Proportion of males with secondary education *Teso tribe	0.229	0.042	5.5
Number of males education between P5-P7*Ganda;	2.535	0.521	4.87
Maximum years of education 13* Rwanda tribe	-1.886	0.650	-2.9
Heads education between P5-P7*Ganda tribe	-2.824	1.261	-2.24
Log of heads age *Kamuli district	-0.069	0.016	-4.25
Log of heads age * Kapchorwa district	0.093	0.021	4.44
Log of heads age * Kumi district	-0.070	0.015	-4.53
I Log of heads age *Soroti district;	-0.070	0.014	-3.82
Maximum years of education*pit latrine	-0.070	0.005	6.35
Maximum years of education *Kamuli district	-0.070	0.008	2.11
Number of males education between P5-P7*Iganga district	0.062	0.021	2.89
Number of males education between P5-P7*_1km_track	0.049	0.021	2.32
Male head separated, divorced*Kamuli district	-0.348	0.131	-2.66
hnm30min*percent of parish under woodlot;	1.220	0.303	4.02
Number of males aged 30_49	0.584	0.184	3.18
Household size =1	1.722	0.281	6.13
Household size =8	1.587	0.546	2.91
Number of females ages 10 and below	-0.692	0.266	-2.61
Number of females aged 6 to 16 years	-1.449	0.235	-6.17
Number of females aged 15 and younger	1.112	0.288	3.86
Number of males with education P1-P4	0.444	0.105	4.22
Number of males with education P1-P4 squared	-1.904	0.738	-2.58

Table C3. First stage regression: North

Dependent Variable: log of per capita consumption expenditure

Number of observations: 1368

Number of Clusters: 144

Adjusted R-Square: 0.46

Variable	Parameter Est.	Std Error	t value
Intercept	10.225	0.093	110.38
Number of males with secondary education	0.061	0.029	2.09
Household size = 5	0.090	0.036	2.52
Household size =13	0.366	0.121	3.03
Max. education deficit of children aged 7-13 squared	-0.001	0.000	-2.82
Log of adult equivalent size	-0.681	0.052	-13.22
Proportion of females aged 30-49 squared	0.350	0.141	2.48
Number of males with education level P1-P4	-0.083	0.019	-4.33
Number of males with primary education	0.101	0.016	6.27
Proportion of males with education at O' level and higher	0.512	0.179	2.87
Number of females aged 30 or older	0.092	0.026	3.49
Proportion of parish within 1km from the main road	0.682	0.233	2.93
Proportion of parish within 1km from tarmac road	6.153	1.623	3.79
Proportion of parish within 3km from tarmac road	-8.865	1.692	-5.24
Proportion of parish within 4km from tarmac road	5.732	0.969	5.92
Percent of parish under subsistence farms	-0.130	0.054	-2.4
Percent of parish under subsistence farms in wetlands	-3.714	1.160	-3.2
Percent of parish under water	0.856	0.140	6.09
Interaction terms			
Age of household head * Lugbar tribe	0.007	0.002	3.24
Age of household head * Arua district	0.008	0.002	3.98
Head male and divorced* Head male and divorced	2.866	1.169	2.45
Highest number of years of education in household *Madi tribe	0.057	0.008	7.54
Number of males 30 and above* Arua district	-0.143	0.051	-2.82
Number of males 50 and above* Head male and divorced	-0.406	0.106	-3.81
Number of males 50 and above*Lugbar tribe	-0.569	0.114	-5
Number of females aged 15 or younger* Apac district	0.066	0.012	5.45
Age of household head * Prop. of parish 1km from main road	-0.020	0.004	-5.61
Log of adult equivalent size * Gulu district;	-0.346	0.087	-4
Log of adult equiv. size * Prop. of parish 1km from main road	0.253	0.115	2.2
Log of adult equivalent size* Prop. of parish 1km from track	0.105	0.047	2.27
Head male and divorced *Gulu district	0.564	0.233	2.42
Head male and divorced *Kitgum district	-0.445	0.176	-2.53
Head male and divorced *Nebbi	-3.059	1.076	-2.84
Highest number of years of educ. in household*Gulu district	0.059	0.011	5.27
Table C3. First stage regression: North Continued			
Highest number of years of educ. in household *Lira district	0.015	0.005	3.07
Highest number of years of educ. in household *Moroto district	0.106	0.041	2.62
Max. education deficit of children aged 7 - 18* Gulu district	0.025	0.012	2.11
Number of persons aged 30 or older*Moyo district	-0.177	0.063	-2.79
Number of persons aged 30 or older* prop of parish 1km from main	0.609	0.133	4.59
Proportion of females aged 0-5 squared	-4.514	1.140	-3.96
Number of females aged 45 or older	-0.599	0.134	-4.47

Table C4: First stage regression West

Dependent Variable: log of per capita consumption expenditure

Number of observations: 1637

Number of Clusters: 163

Adjusted R-Square: 0.34

Variable	Parameter Estimate	Standard Error	t-value
Intercept	10.391	0.111	92.54
Number of females aged 6 to 14	0.047	0.017	2.77
Number of males with education at O' level and higher	0.079	0.037	2.17
Household Size = 2	0.004	0.001	6.14
Household size = 11	-0.343	0.101	-2.13
Log of Household size	-0.246	0.041	-3.29
Proportion of females aged 0 to 5 squared	0.934	0.235	-6.98
Proportion of females aged 30 to 49 squared	0.451	0.129	-2.98
Number of males without education	-0.077	0.013	3.99
Number of males with education P1 to P4	-0.076	0.016	2.62
Age of household head squared	0.000	0.000	-5.93
Proportion of parish within 1km from track	0.975	0.165	-4.69
Proportion of parish within 2km from track	-0.684	0.145	-2.63
Proportion of parish within 3km from tamarc	0.169	0.049	-2.39
Proportion of parish within 4km from track	0.226	0.066	5.58
Percent of parish under woodlot	-6.715	2.067	-4.65
Percent of parish under subsistence farms	-0.240	0.053	3.33
Percent of parish under subsistence farms in wetlands	1.096	0.300	3.33
Age of household head *Mukiga tribe	0.034	0.013	-3.24
Age of household head InHousehold heads age*Mukonjo tribe	0.206	0.028	-4.17
Age of household head InHousehold heads age*Munyankole tribe	0.107	0.013	2.88
Mean education deficit of children aged 7-18 *Alur tribe	0.216	0.082	7.13
Mean education deficit of children aged 7-18 * Munyankole tribe	-0.023	0.011	8.11
Mean education deficit of children aged 7-18*Munyoro tribe	-0.083	0.018	2.46
Household head without education *Alur tribe	-1.828	0.560	-2.39
Head male and divorced*Mukonjo tribe	0.574	0.250	-4.93
Highest number of years of education in household *Alur tribe	-0.230	0.062	-3.07
Highest number of years of education in household *Muganda tribe	0.231	0.077	2.6
Log of age of household head *Hoima district	0.071	0.018	2.52
Log of age of household head *Kasese district	-0.134	0.029	-3.74
Mean educ. deficit of children aged 7-18 *Perc. of parish under town	-1.815	0.881	2.85
Household head without education *Kabarole district	-0.157	0.052	3.99
Table C3. First stage regression: West Continued			
Proportion of males without education * Perc. of parish under town	-11.510	4.072	-4.77
Male household head male and divorced*Hoima district	0.478	0.198	-3.21
Household size=6*Kabarole district	0.348	0.098	-3.55
Household size=6*Kabale district	0.367	0.127	2.86
Number of males with education at O' level and higher	0.840	0.172	3.41
Number of males aged 35 or older	-0.419	0.096	2.83

Table C5: First stage regression: Panel

Dependent Variable: log of per capita consumption expenditure

Number of observations: 1058

Number of Clusters: 163

Adjusted R-Square: 0.34

Variable	Param. Est	Stand. Error	t-value
Intercept	10.07	0.06	180.84
Household size=4	0.10	0.04	2.61
Number of males with primary school education	0.02	0.01	2.43
Proportion of females aged 6-14 squared	0.59	0.21	2.76
Prop. of spouses with education at least secondary school	0.28	0.09	3.20
Proportion of females aged 30 to 49	0.33	0.11	2.92
Highest number of years of education in household*Muganda tribe	0.02	0.01	2.88
Interaction terms			
Log of household size*No bathroom	-0.29	0.05	-5.83
household size=1 + household size=2 + household size=3*Munyoro tribe	0.49	0.18	2.73
Household size=1 + household size=2 + household size=3*Mutoro tribe	-0.46	0.13	-3.43
Highest No of yrs of educ. in hh*prop of parish in 5km of tarmac	-0.01	0.00	-3.03
Number of males with at least secondary school*free house	1.72	0.47	3.63
Head female, divorced, separated or widowed *heads age squared* Mukonjo tribe	0.00	0.00	-1.66
Head female, divorced, separated or widowed *heads age squared* *prop. parish commercial farms	0.00	0.00	-3.05
Head female, divorced, separated or widowed *heads age squared* lighting electricity	0.00	0.00	-2.15
Iron roof*adult equivalence size*Muganda tribe*	1770.16	733.14	2.41
Iron roof* adult equivalence size*prop. Of parish _1km_track* dummy Northern	-1.79	0.57	-3.16
Log household size*Heads activity clerical work* dummy Northern	65.11	17.83	3.65
Log household size*heads activity other* dummy Northern	-0.54	0.08	-6.47
Household size=1 + Household size =2 + household size=3* household activity other * dummy Northern	-0.53	0.12	-4.40
Highest No of yrs of educ. in household *Karimojong tribe*dummy Northern	-0.33	0.13	-2.58
Mean number of years of education of adults * Karimojong tribe *dummy Northern	1.38	0.40	3.47
Mean number of years of education of adults *Madi tribe*dummy Northern	-0.36	0.10	-3.56
Number of males with at least secondary school*perc of parish under papyrus*dummy Northern	5.87	2.87	2.05
Household heads age2*perc. of parish under commercial farm*dummy Northern	0.00	0.00	1.84
Head female, divorced separated or widowed*Hh heads age square*Electricity lighting*dummy Northern	0.32	0.10	3.31
(household size=1+household size=2+household size=3)* Karimojong tribe *dummy Eastern	-90.53	36.36	-2.49
(household size=1+ household size=2+household size=3)*Madi tribe *dummy Eastern	66.02	27.56	2.40
log household size*Japadhola tribe*dummy Western	80.48	23.87	3.37
Mean education deficit of children aged 7-18*Mugisu tribe*dummy Western	-0.37	0.12	-2.97

Table C5: First stage regression: Panel continued

Number of males with at least secondary school *Acholi tribe*dummy Western	-109.76	34.34	-3.20
Household heads age2*Household heads age squared*dummy Western	0.01	0.00	2.23
Log household size*perc. of parish degraded*dummy Western	-0.62	0.27	-2.27
Highest No of yrs of educ. in household * stone wall*dummy Western	3.17	0.69	4.58
Household heads age squared*Lugbar tribe*dummy Western	0.00	0.00	-4.10
Household heads age squared*Heads activity clerical worker*dummy Western	0.02	0.01	3.51

Table C5: First stage regression: Panel Continued

Iron roof*adult equivalent size*Mukiga tribe*dummy Western	-0.17	0.03	-4.93
Female hh head, divorced, separated or widowed *Hh heads age squared*Lugbar tribe*dummy Western	-0.04	0.01	-3.80
Heads activity is student	-8.35	4.96	-1.68
Live in personal house	7.82	3.45	2.27
Tribe Rwanda	0.36	0.18	1.95

D. Poverty estimates at county level, 1992

Central		County	Population	Y		FGT0	Standard error
Code	District			Mean	Standard error		
11	Kalangala	Bujumba	7,265	24,583	1,488	0.292	0.039
		Kyamuswa	6,953	26,792	1,504	0.248	0.034
17	Kiboga	Kiboga	131,768	15,850	489	0.613	0.022
23	Luwero	Katikamu	117,255	18,753	692	0.524	0.023
		Nakaseke	89,962	19,364	651	0.475	0.025
		Wabusaana	102,685	16,276	542	0.601	0.024
24	Masaka	Bukomansimbi	125,322	18,599	518	0.495	0.022
		Bukoto	322,255	18,571	513	0.515	0.017
		Kalungu	139,088	19,229	556	0.482	0.020
		Lwemiyaga	19,109	21,955	2,036	0.426	0.041
		Mawogola	118,493	16,624	488	0.576	0.022
30	Mpigi	Busiro	235,573	20,448	674	0.461	0.022
		Butambala	69,740	17,546	762	0.540	0.036
		Gomba	113,453	15,091	463	0.655	0.023
		Kyaddondo	199,694	23,200	993	0.396	0.025
		Mawokota	132,525	17,306	591	0.566	0.022
31	Mubende	Busujju	64,586	14,042	891	0.703	0.039
		Buwekula	113,660	18,951	1,622	0.536	0.037
		Kassanda	140,479	16,249	832	0.629	0.027
		Mityana	130,064	15,356	757	0.649	0.031
32	Mukono	Bbaale	81,094	18,892	1,110	0.500	0.037
		Buikwe	188,946	20,005	796	0.474	0.023
		Buvuma	18,181	21,171	1,614	0.398	0.057
		Mukono	165,651	21,892	703	0.422	0.020
		Nakifuma	125,426	18,336	626	0.539	0.022
		Ntenjeru	126,135	20,296	1,084	0.480	0.029
35	Rakai	Kabula	46,454	14,123	625	0.692	0.030
		Kakuuto	65,435	15,272	720	0.641	0.033
		Kooki	129,641	16,794	504	0.567	0.023
		Kyotera	121,410	16,963	604	0.574	0.024

East

Code	District	County	Population	Y		FGT0	
				Mean	Standard error	Mean	Standard error
7	Iganga	Bugweri	76,229	15,805	713	0.602	0.032
		Bukooli	224,410	18,326	1,581	0.582	0.025
		Bunya	206,781	46,181	34,352	0.517	0.027
		Busiki	119,489	14,666	585	0.649	0.026
		Kigulu	129,678	15,383	732	0.617	0.032
		Luuka	128,811	16,095	622	0.589	0.027
8	Jinja	Butembe	82,936	125,034	133,230	0.374	0.038
		Kagoma	120,085	23,999	4,505	0.410	0.041
13	Kamuli	Budiope	124,776	12,091	793	0.772	0.033
		Bugabula/Buzaya	145,296	13,416	890	0.713	0.040
		Bulamogi	98,989	12,323	860	0.757	0.038
		Bugabula/Buzaya	91,621	13,255	950	0.715	0.044
14	Kapchorwa	Kongasis	22,935	20,405	2,077	0.428	0.066
		Kwen	35,901	18,690	1,713	0.463	0.069
		Tingey	43,183	18,652	1,708	0.470	0.066
21	Kumi	Bukedea	72,563	11,106	889	0.824	0.037
		Kumi	86,449	10,708	805	0.831	0.036
		Ngora	57,138	11,100	831	0.814	0.036
26	Mbale	Bubulo	176,144	15,464	481	0.610	0.023
		Budadiri	142,733	16,812	825	0.561	0.033
		Bulambuli	64,080	17,154	817	0.552	0.030
		Bungokho	179,705	16,563	783	0.585	0.027
		Manjiya	78,267	15,172	683	0.626	0.031
34	Pallisa	Budaka	98,826	15,257	707	0.621	0.031
		Butebo	62,398	15,462	567	0.609	0.026
		Kibuku	89,849	14,171	543	0.670	0.025
		Pallisa	96,863	14,885	564	0.639	0.025
37	Soroti	Amuria	44,620	12,845	1,695	0.744	0.071
		Kaberaido	36,629	11,494	780	0.797	0.035
		Kalaki	39,879	11,603	877	0.791	0.040
		Kapelebyong	16,805	10,492	840	0.835	0.036
		Kasilo	28,826	11,244	849	0.805	0.039
		Serere	54,197	12,038	804	0.772	0.037
		Soroti	69,230	11,386	728	0.805	0.033
		Usuk	68,266	11,874	742	0.779	0.033
38	Tororo	Bunyole	103,163	25,966	8,043	0.559	0.026
		Kisoko	156,173	13,877	587	0.684	0.029
		Samia-Bugwe	133,252	17,918	2,402	0.625	0.022
		Tororo	90,516	15,765	734	0.616	0.029

North							
Code	District	County	Population	Y		FGT0	
				Mean	Standard error	Mean	Standard error
1	Apac	Kole	113,620	15,873	893	0.645	0.030
		Kwania	83,067	15,180	771	0.652	0.033
		Maruzi	70,583	15,198	842	0.658	0.034
		Oyam	173,559	15,943	964	0.632	0.033
2	Arua	Aringa	98,081	13,867	835	0.729	0.036
		Ayivu	109,044	24,989	2,884	0.498	0.031
		Koboko	57,190	11,966	1,114	0.796	0.055
		Madi-Okello	69,239	17,500	1,655	0.563	0.061
		Maracha	106,073	14,873	725	0.674	0.035
		Terego	97,542	15,119	770	0.659	0.038
		Vurra	62,972	16,453	800	0.604	0.035
5	Gulu	Achiwa	62,851	12,773	749	0.769	0.024
		Kilak	84,215	10,208	505	0.855	0.017
		Nwoya	36,109	9,850	558	0.871	0.018
		Omoro	94,048	14,154	1,083	0.737	0.028
19	Kitgum	Agago	93,307	21,550	106,446	0.876	0.015
		Aruu	79,406	9,597	341	0.890	0.013
		Chua	86,781	9,436	321	0.901	0.013
		Lamwo	67,591	10,452	406	0.857	0.016
20	Kotido	Dodoth	45,539	8,517	466	0.919	0.015
		Jie	43,678	8,517	489	0.917	0.015
		Labwor	22,335	10,019	443	0.874	0.018
22	Lira	Dokolo	83,818	12,401	524	0.780	0.025
		Erute	150,126	14,182	788	0.706	0.029
		Kyoga	66,631	17,863	1,639	0.573	0.050
		Moroto	111,112	11,777	582	0.801	0.025
		Otuke	42,506	11,210	468	0.828	0.021
28	Moroto	Bokora	35,280	11,633	1,945	0.827	0.030
		Chekwii	30,998	11,391	1,822	0.835	0.036
		Matheniko	32,056	10,431	1,058	0.865	0.023
		Pian	17,393	12,609	2,450	0.811	0.038
		Upe	7,275	10,836	1,234	0.842	0.034
29	Moyo	East Moyo	62,091	13,948	719	0.703	0.030
		Obongi	21,558	12,791	900	0.753	0.036
		West Moyo	49,152	14,581	806	0.678	0.033
33	Nebbi	Jonam	62,868	10,754	495	0.850	0.021
		Okolo	117,348	9,570	339	0.895	0.013
		Padyere	106,136	10,080	389	0.879	0.016

West Code	District	County	Population	Expenditures		FGT0	
				Mean	Standard	Mean	Standard
					error		error
3	Bundibugyo	Bwamba	72,675	15,336	1,137	0.613	0.041
		Ntoroka	13,615	19,772	1,590	0.391	0.057
4	Bushenyi	Buhweju	48,718	16,219	922	0.532	0.045
		Bunyaruguru	15,339	21,419	1,681	0.357	0.050
		Igara	124,776	19,075	946	0.430	0.034
		Ruhinda	106,000	16,850	776	0.510	0.036
		Sheema	115,463	20,639	991	0.380	0.031
6	Hoima	Bugahya	115,625	17,274	1,457	0.532	0.058
		Buhaguzi	73,051	17,430	1,533	0.510	0.062
9	Kabale	Ndorwa	140,584	15,826	990	0.556	0.047
		Rubanda	145,865	15,236	876	0.582	0.043
		Rukiga	85,586	16,485	1,007	0.525	0.045
10	Kabarole	Bunyangabu	54,883	19,675	1,265	0.406	0.045
		Burahya	136,382	18,116	1,079	0.464	0.042
		Kibale	101,144	14,572	589	0.617	0.030
		Kitagwenda	53,191	15,526	754	0.577	0.036
		Kyaka	55,756	17,382	973	0.481	0.042
		Mwenge	157,259	16,624	889	0.516	0.039
		Bukonjo	129,236	15,597	1,323	0.568	0.066
15	Kasese	Busongora	79,688	16,555	1,504	0.532	0.065
		Bugangazi	46,086	13,618	865	0.676	0.042
16	Kibale	Buyaga	123,723	13,361	563	0.681	0.029
		Buyanja	37,104	12,761	840	0.714	0.043
		Bufumbira	174,013	12,930	749	0.703	0.041
25	Masindi	Bujenje	41,336	12,533	1,049	0.728	0.051
		Buliisa	43,153	10,666	1,163	0.824	0.046
		Buruli	68,910	12,378	1,001	0.737	0.045
		Buruli	68,910	12,378	1,001	0.737	0.045
		Kibanda	67,840	11,659	883	0.769	0.040
27	Mbalala	Bukanga	49,213	18,683	1,070	0.448	0.041
		Ibanda	117,569	18,154	801	0.462	0.032
		Isingiro	99,027	19,358	926	0.434	0.031
		Kashari	45,970	22,012	1,384	0.346	0.036
		Kazo	45,313	17,484	1,053	0.488	0.045
		Nyabushozi	50,564	19,513	1,205	0.405	0.045
		Bwampara	78,875	21,446	1,181	0.366	0.030
		Kinkizi	142,853	13,858	702	0.652	0.038
36	Rukungiri	Rubabo	90,086	14,699	808	0.610	0.040
		Rujumbura	108,483	13,148	717	0.692	0.039

Appendix E: Correlations between biomass and poverty

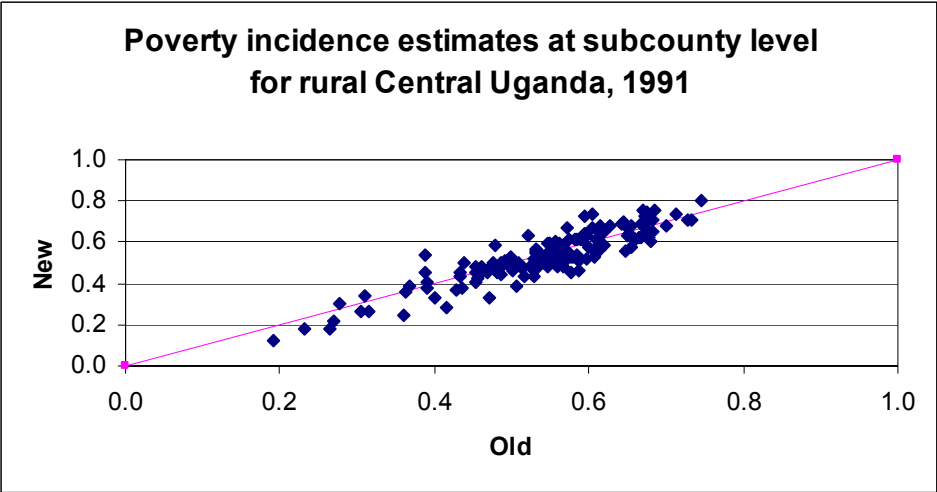
Table E1: Correlation between land use(biomass) and poverty incidence at county level

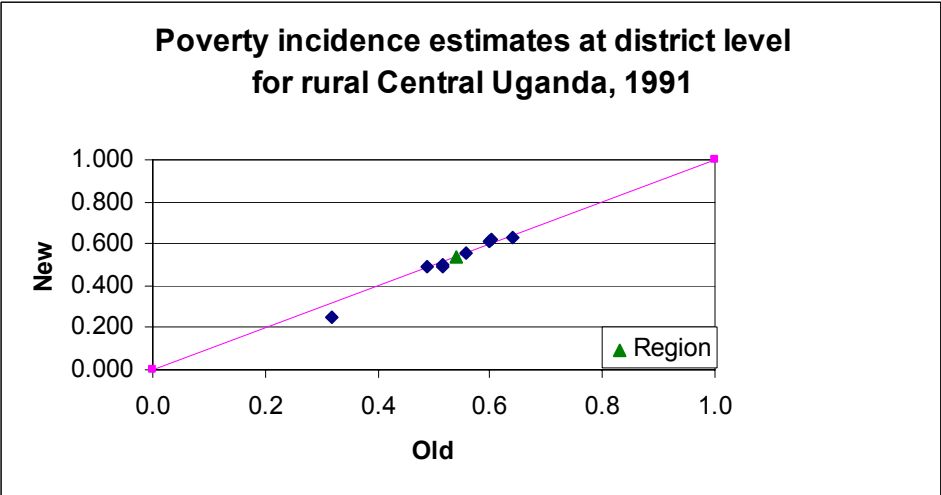
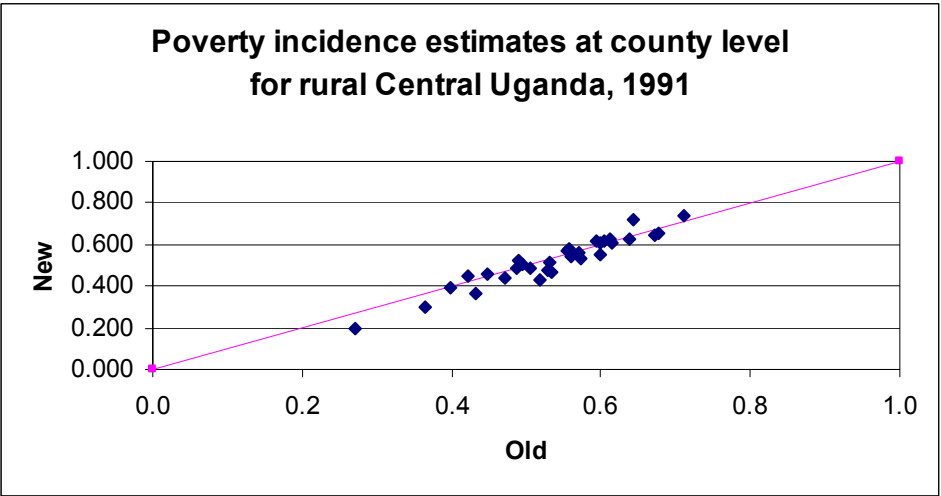
VARIABLE	Uganda (county level)	Central	East	North	West
WITHIN 1KM FROM TRACK	-0.014	0.100	-0.215	-0.456	-0.272
WITHIN 2KM FROM TRACK	-0.002	0.149	-0.239	-0.461	-0.267
WITHIN 3KM FROM TRACK	0.023	0.179	-0.235	-0.498	-0.254
WITHIN 4KM FROM TRACK	0.058	0.196	-0.208	-0.496	-0.245
WITHIN 5KM FROM TRACK	0.098	0.204	-0.176	-0.487	-0.238
WITHIN 1KM FROM MAIN	0.030	0.147	-0.520	-0.030	-0.513
WITHIN 2KM FROM MAIN	0.052	0.181	-0.526	-0.017	-0.508
WITHIN 3KM FROM MAIN	0.072	0.210	-0.528	-0.032	-0.510
WITHIN 4KM FROM MAIN	0.102	0.228	-0.511	-0.034	-0.510
WITHIN 5KM FROM MAIN	0.134	0.230	-0.478	-0.034	-0.501
WITHIN 1KM FROM TARMAC	-0.320	-0.001	-0.406	-0.504	-0.263
WITHIN 2KM FROM TARMARC	-0.309	0.038	-0.402	-0.501	-0.311
WITHIN 3KM FROM TAMARC	-0.302	0.071	-0.395	-0.504	-0.359
WITHIN 4KM FROM TARMAC	-0.296	0.096	-0.389	-0.503	-0.395
WITHIN 5KM FROM TARMAC	-0.292	0.115	-0.387	-0.501	-0.422
PERC_WOODLANDS	-0.221	-0.044	-0.350	-0.176	-0.429
PERC_CONIFEROUS	-0.058	0.017	-0.285	-0.110	0.205
PERC_TROPICAL HIGH FOREST	-0.302	-0.587	-0.181	0.195	0.011
PERC_DEGRADED FOREST	-0.219	-0.154	-0.284	0.203	-0.109
PERC_WOODED	0.370	0.072	0.487	0.341	0.254
PERC_GRASSLANDS	0.028	0.239	0.559	-0.499	0.268
PERC_PAPYRUS	-0.042	0.139	-0.029	-0.264	-0.352
PERC_SUBISTENCE FARM	-0.034	0.300	-0.483	0.032	-0.172
PERC_COMMERCIAL FARM	-0.176	-0.091	-0.381	0.184	-0.421
PERC_SUB_FRAM_WETLANDS	-0.037	-0.006	-0.354	0.183	0.033
PERC_TOWN	-0.252	-0.251	-0.553	0.105	-0.434
PERC_WATER	-0.278	-0.662	-0.084	0.062	-0.319

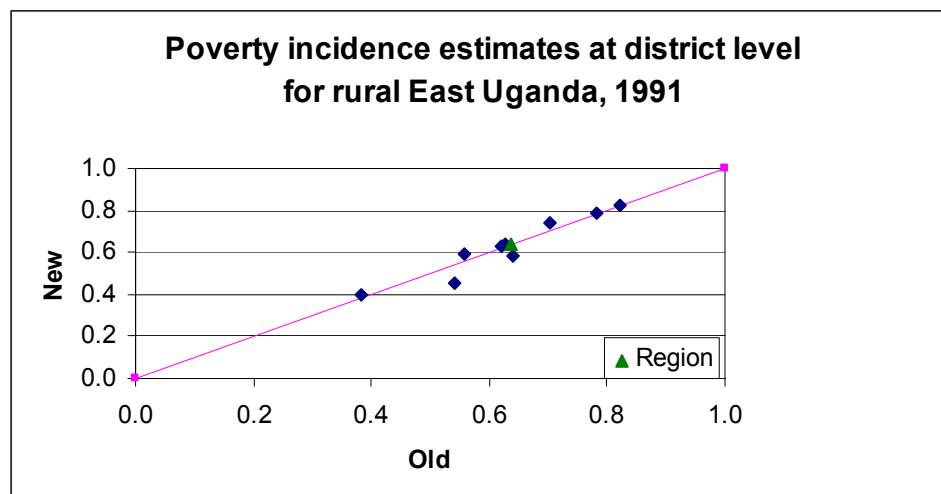
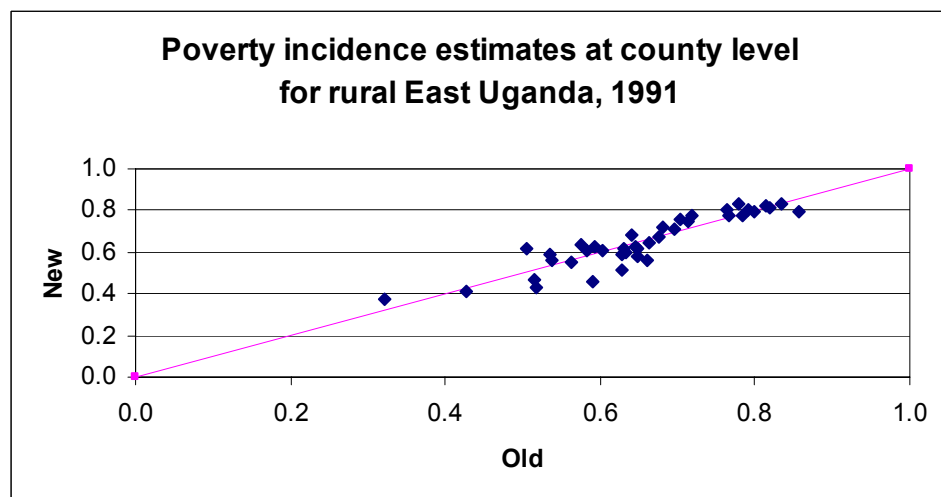
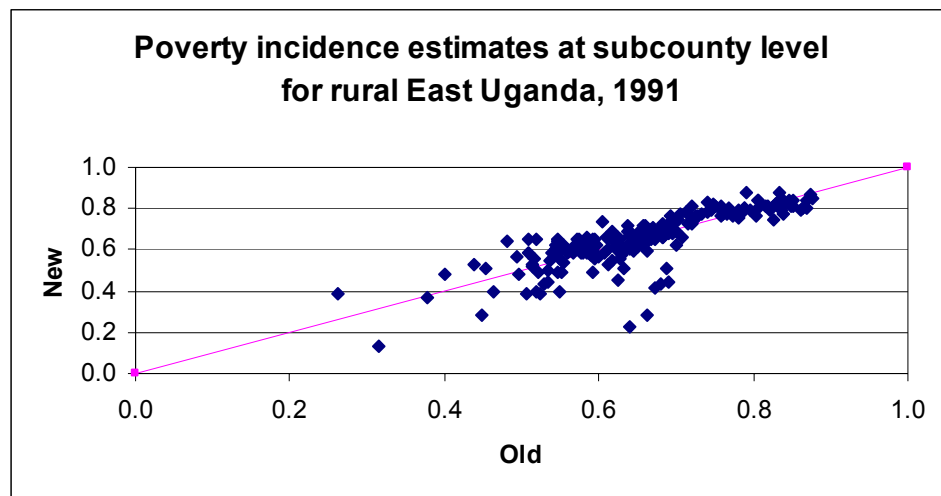
Table E2: Correlation between per capita expenditure: Survey, Predicted no biomass and predicted with biomass

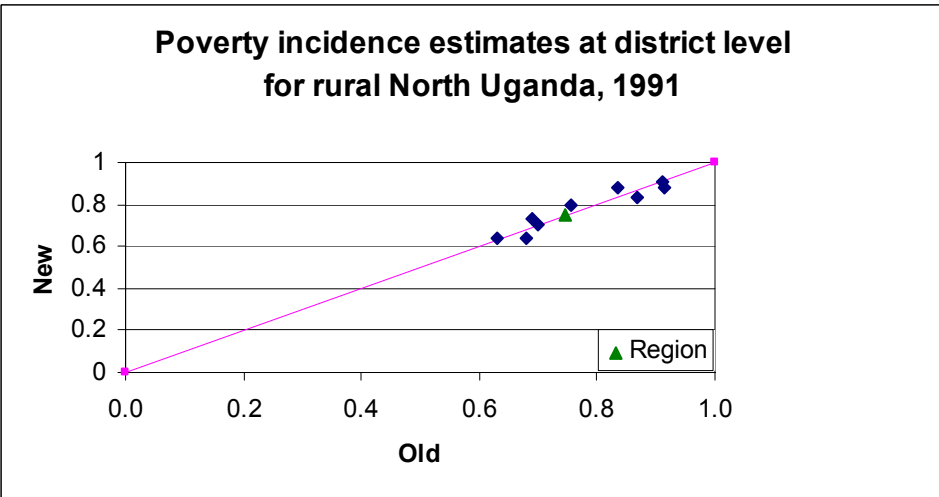
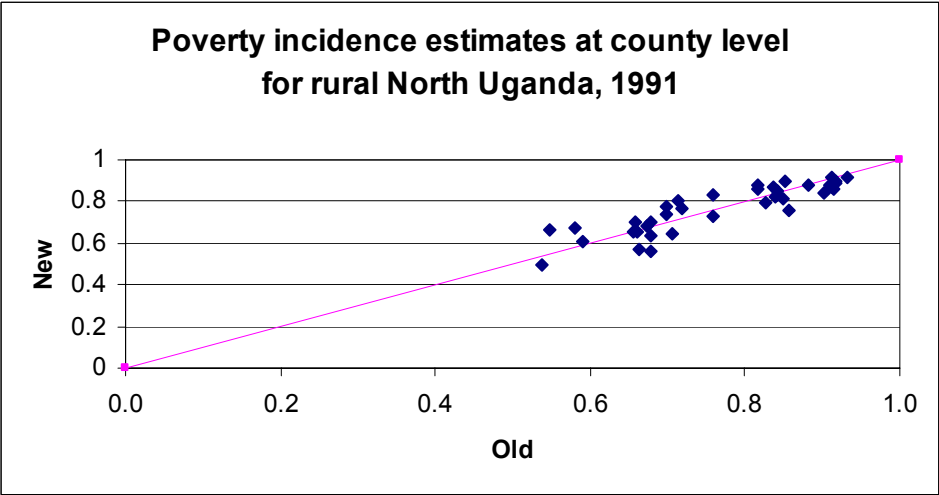
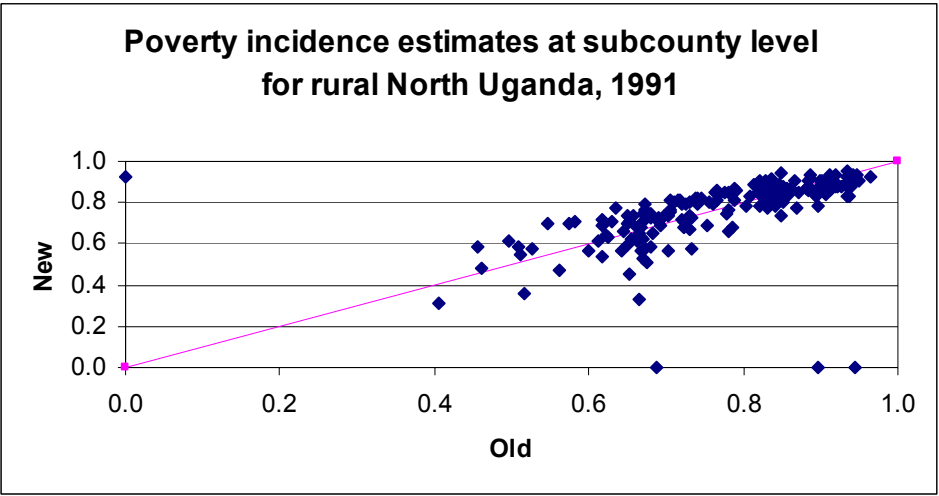
Panel 1992/99a		Survey	Predicted No biomass	Predicted Biomass
	Survey	1		
	Predicted No Biomass	0.9873	1	
	Predicted Biomass	0.9612	0.9898	1
Cross section 1992		Survey	Predicted No biomass	Predicted Biomass
	Survey	1		
	Predicted No Biomass	0.9646	1	
	Predicted Biomass	0.5752	0.5347	1

Appendix E Comparison of old and new poverty estimates









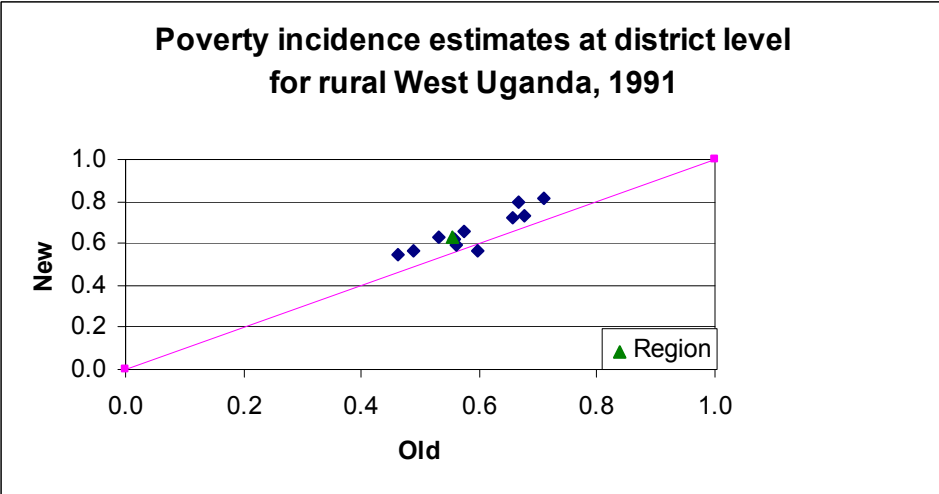
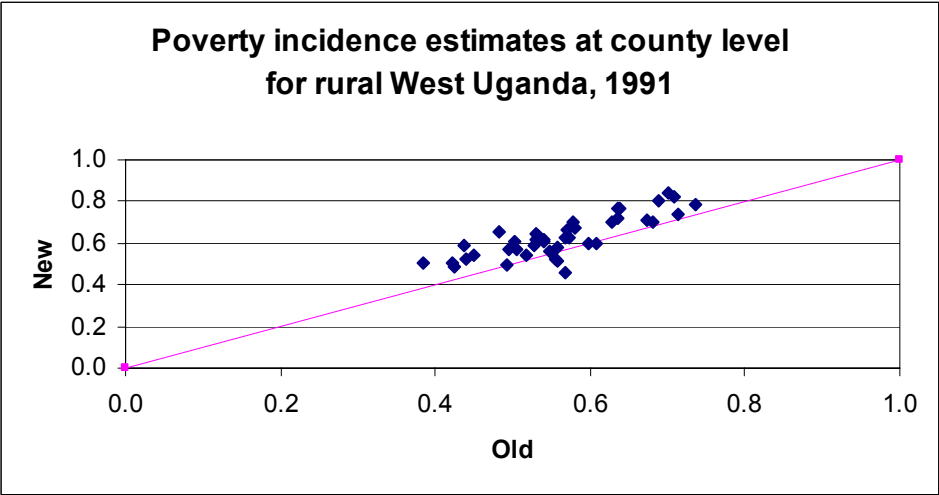
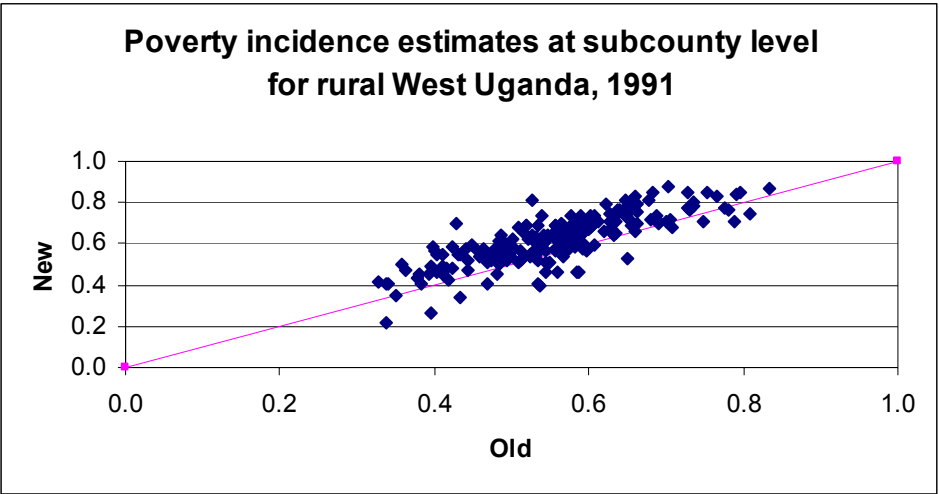


Figure F1: Map of poverty incidence in Uganda based on the poverty estimates with biomass, 1992.

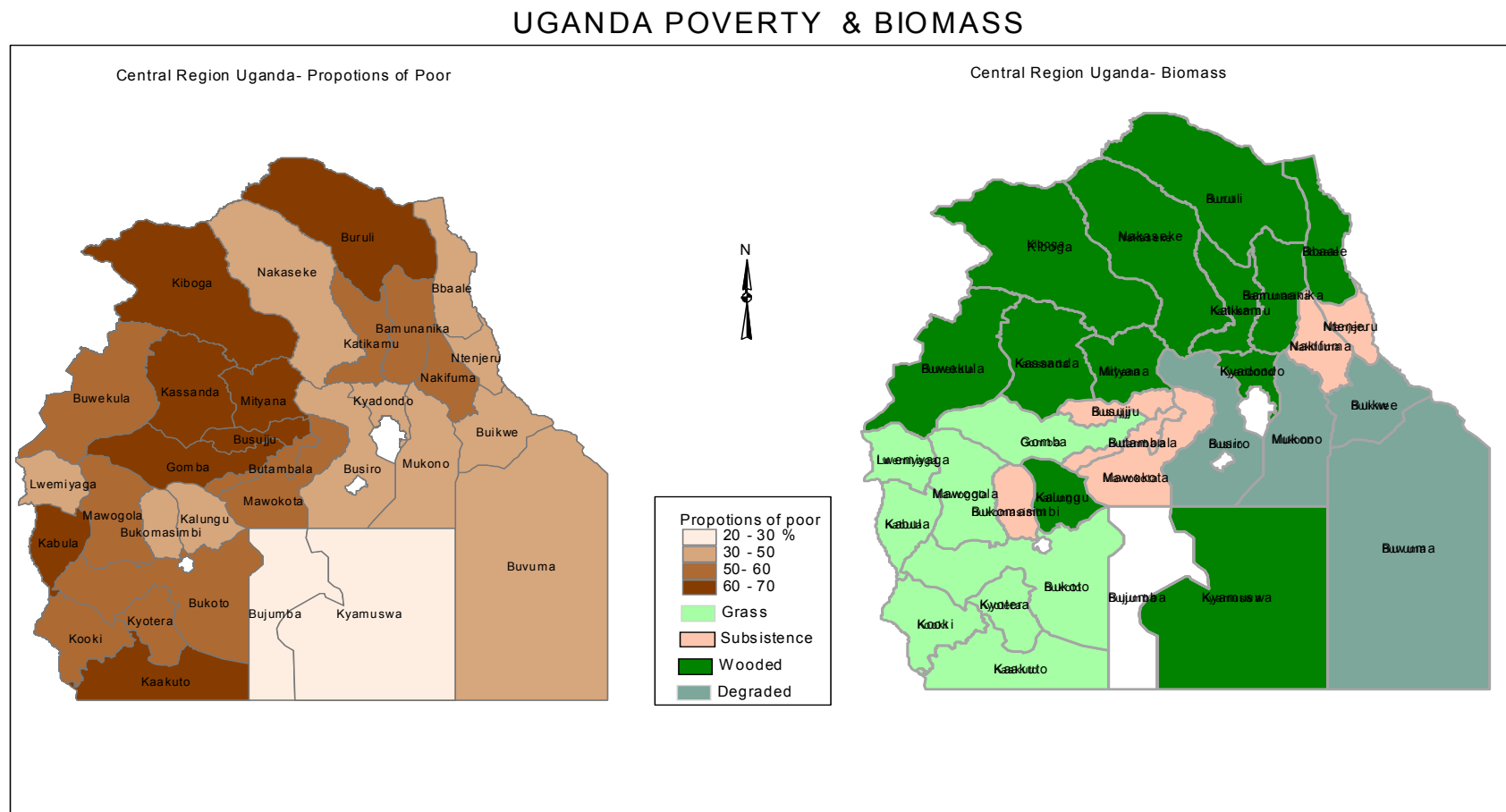


Figure F2: Poverty and biomass in Central region, Uganda, 1992

UGANDA POVERTY & BIOMASS

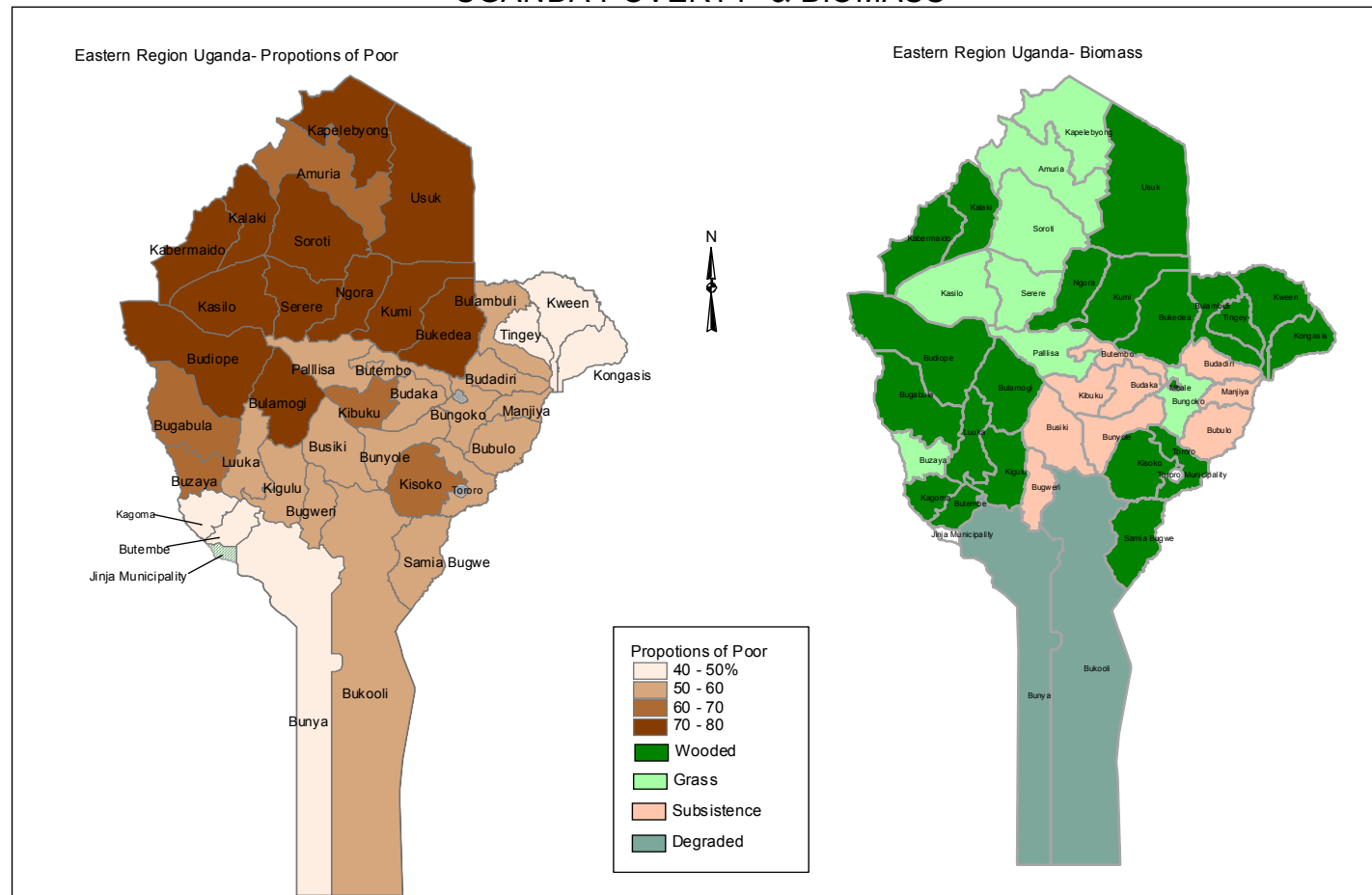


Figure E3: Poverty and biomass in Eastern region, Uganda, 1992

UGANDA POVERTY & BIOMASS

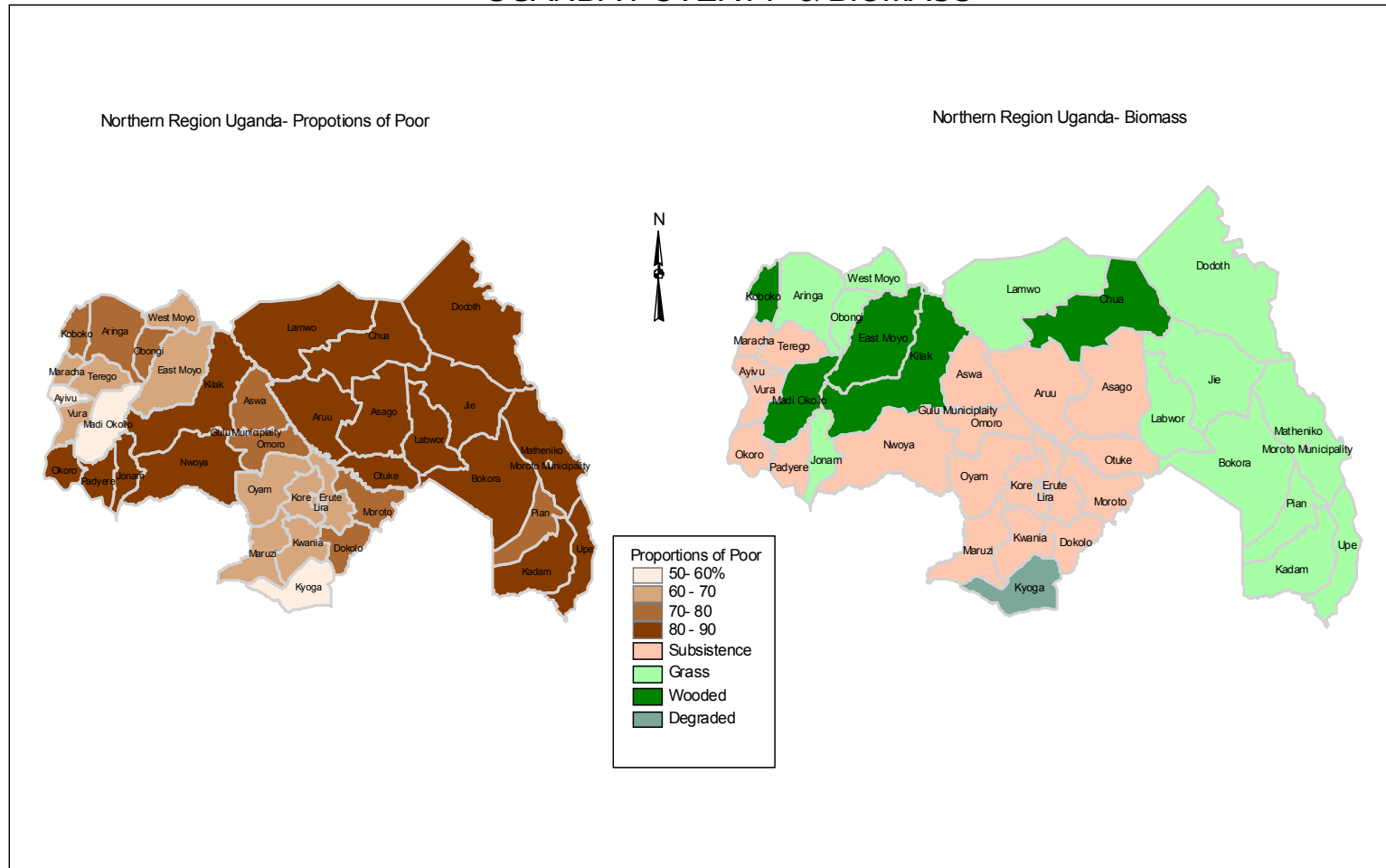


Figure F4: Poverty and biomass in Northern region, Uganda, 1992

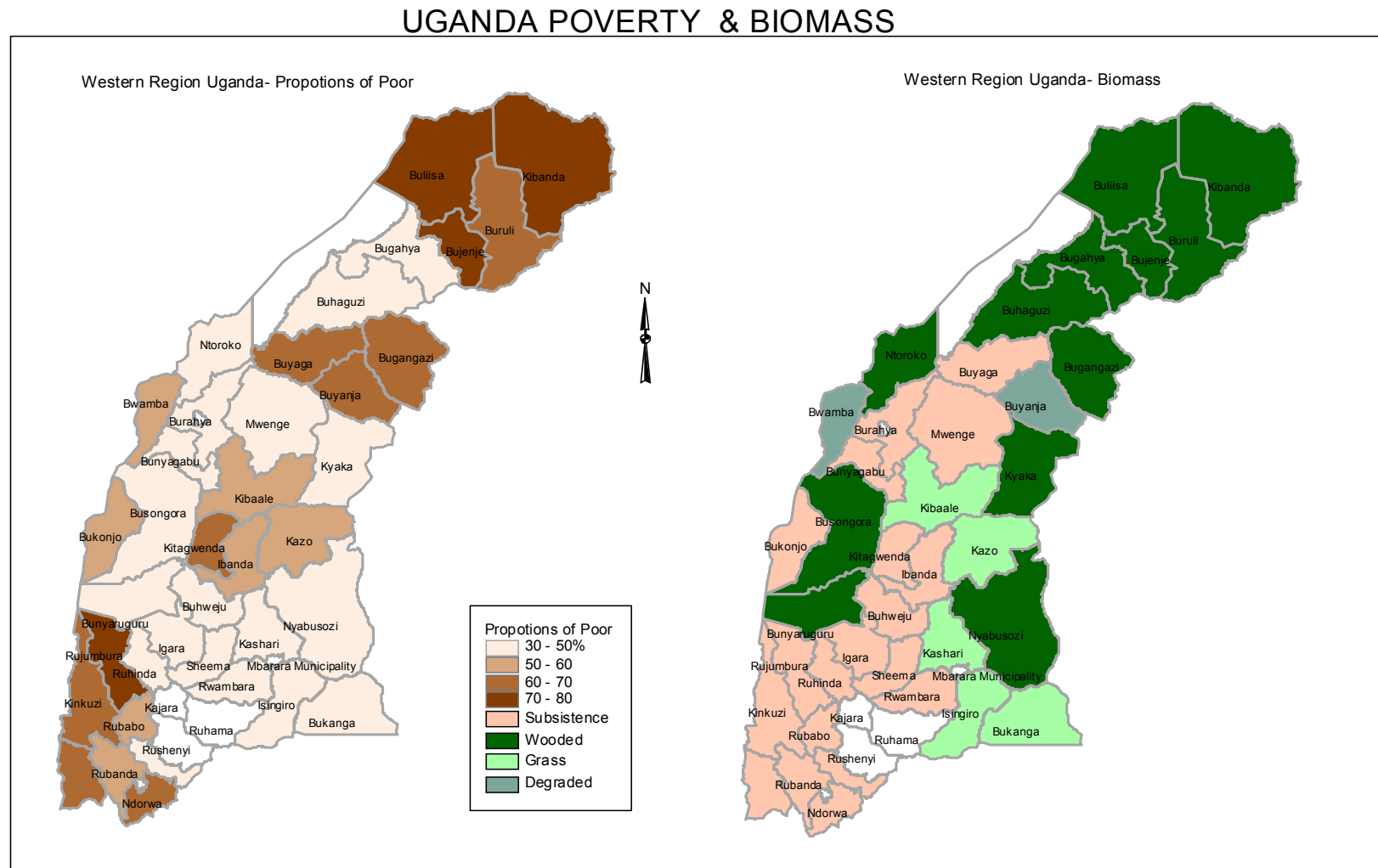
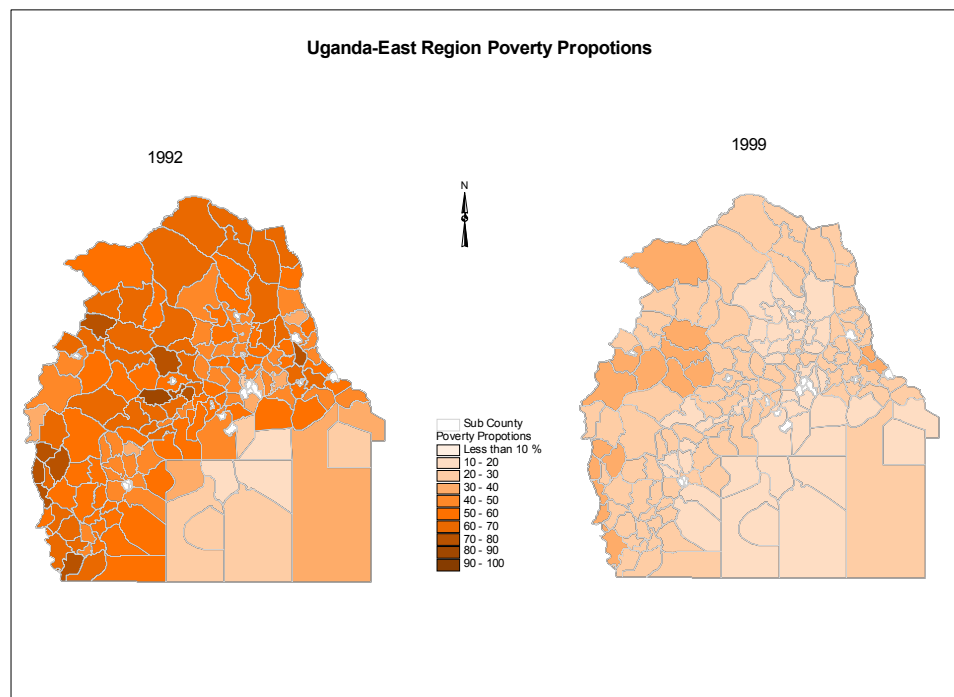
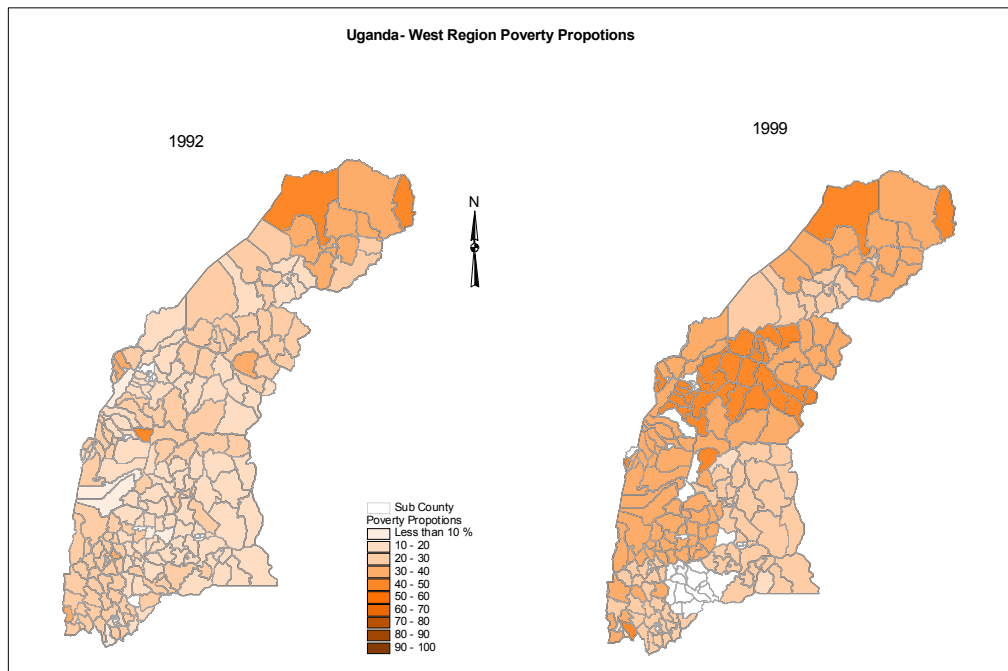


Figure F6. Uganda Poverty rates comparison between 1992 and 1999



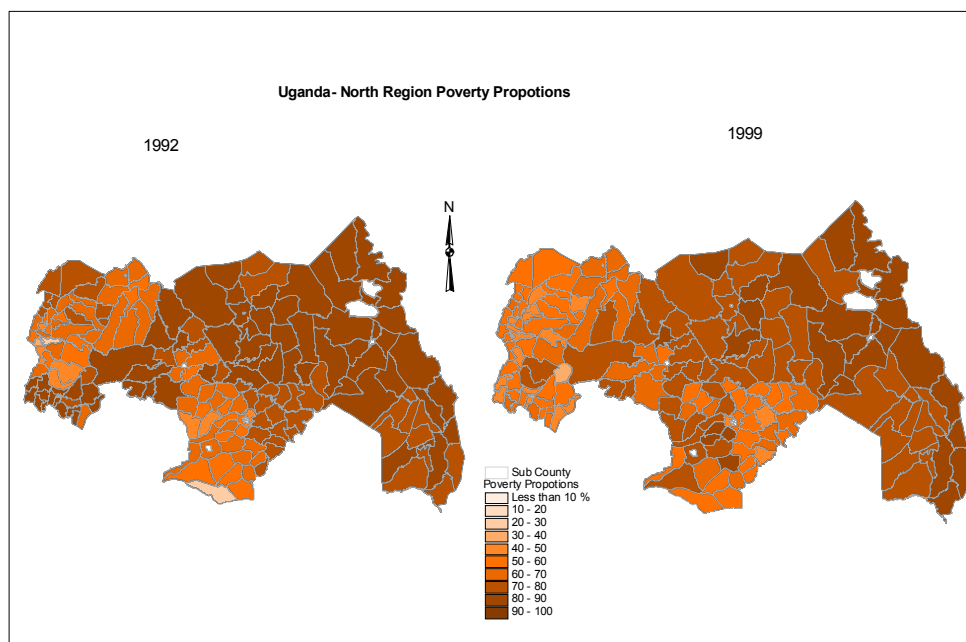
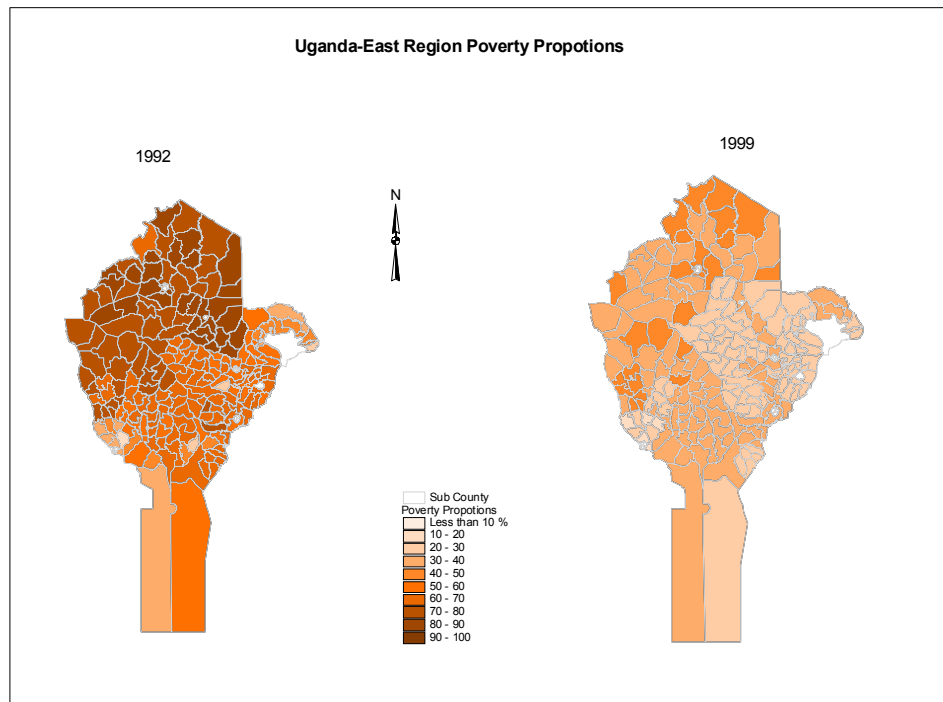


Figure F7. Uganda Land use changes between 1992 and 1999

