Geographic Determinants of Poverty in Rural Kenya:
A National and Provincial Analysis.

Paul O. Okwi, Godfrey Ndeng'e*, Patti Kristjanson, Mike Arunga, An Notenbaert, Abisalom Omolo, Norbert Henninger**, Todd Benson***, Patrick Kariuki and John Owuor

A Report prepared for the Rockefeller Foundation,
Nairobi, Kenya.

April 2006

*Godfrey is with the Central Bureau of Statistics, Kenya.
**Norbert is with the World Resources Institute (WRI)
*** Todd is with the International Food Policy Research Institute (IFPRI)

Table of Contents

List of Tables

Table **Page**

Appendix

Acknowledgements

This report is a publication of a project jointly implemented by Central Bureau of Statistics (CBS) and International Livestock Research Institute and funded by the Rockefeller Foundation.

The authors would like to thank John Lynam for his initiative and contribution. The authors also express their appreciation to CBS, FAO and other organisations that made this data available. The contributions and comments from staff of CBS and Government of Kenya are acknowledged and appreciated. Special thanks are also due to John Owuor and Russ Kruska of the GIS Unit of ILRI for their help in GIS mapping and analysis, Luis Carlos Rodriguez and Maren Radeny for their comments, Douglas Ikong'o for logistical support and many others whom we have not mentioned here. The views and any errors in this report are entirely the responsibility of the authors.

Abstract
This paper investigates the link between poverty incidence and geographical conditions within rural Locations (administrative areas that usually contain several communities) in Kenya. Evidence from poverty maps for Kenya and other developing countries suggests that poverty and income distribution are not homogenous, with wide spatial variability. We use spatial regression techniques to explore the effects of geographic factors on poverty. The results show mixed effects of geographic variables at national versus provincial levels. Slope, soil type, distance/travel time to public resources, elevation, type of land use, demographic and income inequality variables prove to be significant in explaining spatial patterns of poverty. However, differential influence of these and other factors at the Location-level shows that Provinces in Kenya are highly heterogeneous; hence different spatial factors are important in explaining welfare levels in different areas within Provinces, suggested targeted pro-poor policies are needed. Policy simulations are conducted to explore the impact of various interventions on Location-level poverty levels. Investments in roads and improvements in soil fertility both are shown to potentially reduce poverty rates, with differential impacts in different regions of Kenya.

Key words: Poverty, spatial variables, rural Kenya, Provinces, Locations

1.0 Background

This study examines the determinants of poverty prevalence for small, spatially-defined populations in rural Locations[1] of Kenya. Evidence from poverty maps for East Africa (Kenya and Uganda) and other developing countries shows that poverty and income distribution are not homogenous and vary widely across space. There are significant differences in poverty and welfare levels between communities living in different geographical areas. Some of these differences are caused by differences in geographic and agro-climatic conditions (such as rainfall, soil fertility, altitude), infrastructural access to markets and public facilities (e.g. hospitals and schools), the presence or absence of natural resources such as forests or water bodies, as well as political and historical factors. Even though these factors have been identified as major contributors to differences in standards of living of populations in different areas, there has been little empirical work to ascertain the exact relationship between welfare levels and these factors. This type of analysis has been limited to date due largely to data deficiency and lack of appropriate analytical tools. Recent advances in spatial analytical software now allow such analyses to be undertaken.

Poverty, income inequality and natural resource degradation are severe problems in Kenya, especially in rural areas. Once among the most prosperous economies in Africa, today Kenya has among the highest rates of poverty and natural resource degradation in the developing world. National poverty prevalence is estimated to be around 53 percent (Central Bureau of Statistics (CBS), 2003) and natural resource degradation is reported to be increasing in various dimensions (NEMA, 2003). In the recent past, there have been several studies on poverty and income distribution in Kenya (see for example, Geda et al., (2001) and Mwabu et al. (2000)). Some of these studies have focused on the poverty profile, a descriptive tool that provides key information on welfare. The poverty profile is a bi-variate analysis which compares the poverty status of households or individuals to a range of selected characteristics of the households or individuals. The poverty profile is limited in its usefulness because it shows how poverty levels are correlated with one characteristic at a time, hence it tends to simplify complex relationships.

[1] Kenya's administrative units are Province, Division, District, Location and sub-Location.

Poverty maps are another tool that provides important information on the spatial distribution of poverty within a country. Kenya has in recent years developed poverty maps in an effort to improve resource allocation for poverty alleviation (CBS, 2003; CBS, 2005). However, like the poverty profile, the use of poverty maps does not furnish an estimate of the causal linkage between poverty and the variables that influence it; such maps furnish mainly a 'visual' picture of where the poor are located at a relatively high resolution. Below District-level welfare measures now allow a more insightful analysis of the empirical relationships between poverty, socio-economic and spatial indicators.

In Uganda, for example, using similar high-resolution poverty maps, high poverty rates have been found to be concentrated in environmentally fragile regions (Okwi et al., 2005). Demographic factors are also involved in complex ways and tend to exacerbate the problems of environmental degradation and poverty. In Kenya, like other developing countries, many households living within the same community have similar sources of income (e.g. from livestock in pastoral areas) and all households within a community are similarly affected by the geographic and agro-climatic conditions they face. They also deal with common circumstances such as good or poor access to roads, schools, hospitals and water. It is thus reasonable to assume that the environment in which people live has a large influence over their livelihood options, and in turn, their relative welfare levels.[2]

In this study, we attempt to explore this link between people and their local environments, that is, between empirical welfare information (poverty incidence at the Location-level) and Geographical Information System (GIS) based environmental data. An important aspect in developing this link is taking account of the fact that the dependent variable is of a different data type and of a different form of spatial aggregation than most of the independent spatial variables; while socioeconomic variables typically exist in a spatially discrete format based on administrative units, environmental data come in a spatially continuous nature. This poses methodological challenges. Data of different types and from different sources are used to generate the variables used in the analysis. We use global

[2] Which is not to underestimate the importance of household-level factors, such as education, on household-level welfare; there are many studies exploring household-level factors affecting poverty. Ours focuses on community, or meso-level influencing factors.

(spatial regression analysis) to examine the determinants[3] of the prevalence of poverty incidence in rural areas of Kenya. We also conduct policy simulations to understand the impact of some possible investments on poverty levels in different regions. Developing a better local-level understanding of poverty determinants, together with knowledge about how household-level factors and broader national policies affect household welfare, will assist policy makers and development practitioners in their efforts to enable rural Kenyans to improve their livelihoods and welfare levels.

This paper is structured as follows. The next section discusses the general conceptual issues guiding the analysis and specifically describes the research questions we explore and our approach to modeling the spatial determinants of poverty in Kenya. Section 3 describes in detail the data and summarizes the econometric methodology used. Section 4 presents and discusses the results of the spatial regression analysis and the policy simulations. Some concluding remarks and areas for future research are presented in the last section.

2. Conceptual issues in modeling the determinants of poverty

A number of conceptual issues emerge from this study of determinants of poverty at a meso-level (rather than an individual or household-level). First, we develop the key research questions and hypotheses. Second, we discuss the theoretical issues underpinning the analysis.

2.1 Research Questions

The key research questions in this study are as follows:

i. What spatial factors account for the spatial variation in Location-level poverty levels across rural areas of Kenya?

ii. Does the relationship between agro-climatic and other spatial variables and poverty differ significantly among provinces in rural areas of Kenya?

iii. What are the potential poverty impacts of investment/changes in some of the spatially-related factors found to influence poverty in different areas of Kenya?

[3] We follow the 'risk chain' theoretical approach taken by Benson that implies the spatial variables used as independent variables are largely exogenous to the outcome (a consumption-based indicator, such as the poverty indicator used here) and therefore can be interpreted as determinants, and not merely correlates of poverty. For further discussion of this distinction, see Benson, 2005 & Benson et al., 2004.

Generally, we investigate how well agro-climatic and market access variables account for the spatial variation seen in Location-level poverty rates across rural areas of Kenya. Overall, it is expected that the relationship between agro-climatic variables and poverty varies significantly from one Province of Kenya to another. The ability of agro-climatic variables to explain a large portion of the differences in poverty indicates that poverty in remote areas may be linked to agricultural potential, natural resource availability and lack of market access (see also Pender, et al., 1999; Pender, 2001). Better roads and access to markets are expected to favor production of high-value products and non-farm activities and should therefore contribute to better welfare or higher incomes (Pender et al., 2001). Areas with high agricultural potential (suitable climate and soil) may also have an absolute advantage in producing high value perishable vegetables and other crops. Other land use factors may have mixed results on welfare, depending upon the relative impacts on costs of productive factors (Angelsen, 1999). In terms of the demographic variables, population density is expected to influence labor intensity of agricultural production, including the choice of commodities as well as production technologies and land management practices, by affecting the land-labor ratio (Boserup, 1965; Pender, 2001). The effects of population density on welfare are mixed. Presence of social services such as hospitals, schools and markets may influence welfare in localities in Kenya by promoting better health, livelihood and other human capital variables. Such an understanding of the determinants of poverty can effectively guide governments' and others' effort to reduce poverty by adopting more location specific and precise policy options. It can also provide valuable policy lessons for other countries in the region.

2.2 Poverty approaches and Kenya studies

The determinants of welfare can be studied at many levels - regions, communities, households and individuals can all be considered, and many different approaches have been used. Most poverty studies concentrate on determinants of household welfare and at the individual and household-levels (see for example Glewwe 1991; Kyereme and Thorbecke, 1987; Datt et al., 2002; Datt and Jolliffe, 1999). In Kenya, analytical work on determinants of poverty is scanty and the few existing studies have focused on household-level poverty, or simply on descriptive and measurement issues. Mwabu et al. (2000) deal with measurement, profiles and determinants of poverty in Kenya by trying to explain household-level food and overall expenditures. However, a major weakness of their approach is that some factors, for example, income and assets, may increase consumption expenditure and reduce poverty, hence problems of endogeneity. Oyugi (2000) uses both discrete and continuous indicators of poverty as dependent variables and employs a much larger set of household characteristics as explanatory variables. This study analyses poverty at both micro (household) and meso (district level) using 1994 Welfare Monitoring Survey (WMS II). They find that being able to read and write, employment in off-farm activities, being engaged in agriculture, having a side business in the service sector, source of water and household size are important determinants of poverty. Geda et al., (2001) use household data collected in 1994 to explain probable determinants of poverty at the household level, using both binomial and polychotomous logit models. Their findings show that poverty is strongly and positively associated with level of education of household head, household size and engagement in agricultural (if occupation is in agriculture or otherwise) activity. However, most of these studies do not include spatial variables in their analyses and the results of such studies need to be treated with caution.

In this paper, we have sought to extend the analysis of poverty by modeling the determinants of poverty using spatial data for rural areas of Kenya and poverty estimates from the poverty maps obtained from Welfare Monitoring Survey III (1997) and Population and Housing Census (1999). An innovative aspect of this study is that we use *meso* (Location) level poverty rates using spatial regression analytical techniques. Our approach models the spatial determinants of Location-level poverty, or the factors that help explain spatial variation in the proportion of the rural population living below the poverty line across Kenya. We identify the important spatial determinants of the prevalence of poverty among relatively small communities, “Locations”, in rural areas of Kenya. We focus on rural areas because we expect the results to be more generalizeable across rural Kenya than if urban areas are included. Another reason we focus on rural areas is that agro-ecological conditions are important elements of livelihoods in rural areas, both as sources of risks and safety nets (Benson et al 2005). Also, urban areas often pursue a broader range of livelihood strategies which may not be adequately captured by spatial variables. The spatial regression analysis therefore carried out in this study involves estimating poverty incidence as a function of selected variables representing agro-climatic characteristics and market access to determine which of these factors are significant in explaining spatial patterns in this poverty measure. Similar approaches have been followed by Minot et al. (2003), Benson et al. (2004) and Benson (2005).

2.3 Estimation issues

The practical application of the determinants of poverty analysis presents a number of econometric and computational challenges, including issues relating to data, spatial autocorrelation, endogeneity and heteroscedasticity. Before discussing what sort of variables should be included among the set of explanatory variables, it is important to consider some of these issues relating to the model and analysis. First, we look at issues relating to potential heterogeneity of the model of living standards. While there can be different levels of heterogeneity, an argument can be made that, in particular, the rural areas of the Provinces of Kenya are sufficiently different that they warrant different models. For instance, one can argue that market access has different returns between different Provinces, hence has different implications on welfare in the Provinces. Also, we use separate models for the different Provinces because there are a number of geographic and community-related variables that are more relevant for certain Provinces, for example rainfall and soil type, that may not be relevant for other Provinces. Thus the analysis focuses on rural areas and separate models are estimated for each Province.

A second issue that arises in this approach is the time horizon in question. Local communities invariably experience fluctuating living standards over time, associated in part with agro-ecological conditions. When using spatial data along with census and survey information as we are in this analysis, it is only possible to discuss the spatial determinants of welfare at a particular point in time, and not the driving forces behind longer-term welfare levels. Given this, a number of our chosen explanatory factors can be assumed to be exogenous in the short term, which allows us to ignore issues and approaches to deal with endogeneity issues.

A third issue relates to the degree to which these explanatory factors can be influenced by actions, interventions or policies. Obviously, many of them cannot. For example, we cannot influence rainfall levels, but market access can be improved through road investments, and soil fertility can be improved through added fertilizer and soil conservation measures.

We have adopted an underlying theory that attempts to understand how households cope (or fail to cope) with shocks, called the risk chain theory[4]. In this framework, household vulnerability depends upon the degree to which they are exposed to negative shocks to their welfare, and on the degree to which they can cope with such shocks when they occur (Benson et al, 2005). The outcome is whether or not the household is poor, which can be measured by a consumption-based welfare indicator (as we use in this study). At the community or Location-level, shocks such as droughts or floods are typically felt by all households, and their access to natural resource assets (soil, water, services, etc.) that help them cope with the shocks are also similar. Thus, Benson et al. (2005) argue that the independent variables used in this type of analysis are made up of an array of aggregate social and agro-ecological characteristics for the small local areas considered and, based on the underlying risk-chain theory, can be considered determinants of the local prevalence of poverty, and not simply correlates. That is, our results demonstrate more than just an association between levels of the independent variables and local levels of poverty, the chosen spatial independent variables found to be significant can be considered to be in part actually determining the observed poverty levels.

3. Data and Empirical Implementation

This section reviews the tools and methods used to estimate poverty as a function of variables representing agro-climatic characteristics and market access. We describe the data first and thereafter the methods employed.

3.1 Data

The Location-level poverty estimates of this project make use of data obtained from the 1997 Welfare Monitoring Survey (WMSIII) and the 1999 Population and Housing Census 1999. The survey questionnaire collected information on household and demographic characteristics, education, assets, employment, income and expenditure (CBS, 1998). The 1997 Population and Housing Census was conducted by the same institution (CBS) and

[4] See also Dercon (2001)

was meant to cover the entire population in both rural and urban areas. The census questionnaire included information on household members and was administered to all households in the country, with the exception of North Eastern Province. Although the census did not collect information on income and expenditures, it provides information on a number of characteristics that have been shown to be strong correlates of poverty (Elbers, 2000). The census and survey data collect the same information regarding household size composition, education, housing characteristics, access to utilities and location of residences. The small area estimation technique takes advantage of this, and uses this same set of similar variables to predict expenditure levels, and consequently, poverty measures, for all households, which are then aggregated up to the Location-level. It is these Location-level poverty estimates that serve as the dependent variable in our analysis (i.e. the proportion of the population falling below the rural poverty line, Kenya Shillings KShs 1,239/adult equivalent/month, referred to as the headcount poverty measure).

The spatial analysis portion of this project uses a range of spatially referenced variables describing topography, land cover and land use, climate, demography and market/town access, all derived from GIS data layers. Geo-referenced information from various government departments and institutions is used. The data on roads and other topographic data such as land cover, soils and climate data were obtained from Africover[5]. Data on demographic attributes was obtained from CBS, as described above. The project developed its own classification system based on a combination of land cover and land use. This information includes vegetative cover such as forests, grassland, wetlands, and water resources, and land use information, including proportion of each Location under subsistence farming, commercial farming, and buildings. The data also includes information on the distribution of road, market and town infrastructure. Many of these GIS variables required considerable cleaning, processing, and further transformation in order to generate the final set of variables used in the spatial analysis. We use a subset of these variables as our independent variables and the candidate independent variables are aggregated to the Location-level.

[5] Africover is a FAO environmental database for environmental resources. More info at http://www.africover.org/system/area.php?place=1

In selecting among potential determinants of welfare, a key consideration in this study has been selecting variables that are arguably exogenous to welfare or current consumption. Thus, for instance, we exclude several non-spatial characteristics of households such as type of welling or value of assets, because some of these items are already used in the derivation of welfare levels. Some of these excluded household characteristics may also be in part determined by living standards in the area, and would cause endogeneity concerns in the choice of modeling approach.

Potential determinants of the proportion of the rural population falling below Kenya's official rural poverty line (KShs 1,239/adult equivalent/month) are the independent variables in our analysis. Following Benson et al. (2005) and using the risk chain as a theoretical basis to guide our selection of independent variables, they were chosen as follows. A key consideration was the selection of variables that are arguably exogenous to household welfare. Thus, for instance, we tried not to include variables that may influence, or be influenced, by levels of community poverty. Our selection of potential determinants has also been guided by the results of the Kenya poverty profile and poverty mapping studies that suggest some significant correlates of poverty. The data was aggregated to the Location-level and covariance matrices were examined for all the variables. Where high levels of correlation were found between two variables, one was selected so as to limit problems of multi-collinearity.

Table 1 shows the key selected independent variables for the analysis and how they are hypothesized to affect poverty incidence. The variables are divided into two categories. Exogenous variables are those variables that are unlikely to be affected by the level of economic activity or poverty. An example of an exogenous variable is rainfall. This variable may influence poverty in an area but cannot be influenced by poverty. On the other hand, endogenous variables are those that may both influence poverty or be influenced by poverty. For example, the level of economic activity in an area may influence investments in transport infrastructure and market access, and similarly the density of markets and roads may influence the poverty rate or level of economic activity in the long run.

Table 1 Explanatory variables used in spatial regression analysis

Variables	Expected relationship to Poverty
Exogenous variables	
Rainfall	Negative (Low rainfall, higher poverty)
Rainfall variation	Positive (High variation, higher poverty)
Elevation	Positive (High elevation, higher poverty)
Slope	Positive (steeper slope. Higher poverty)
Type of land cover	Not known
Distance to towns/municipalities/cities	Positive (Greater distance, higher poverty)
Length of Growing period	Negative (Longer LGP, lower poverty)
Soil type	Negative (Good soil, lower poverty)
Possible Endogenous variables	
Population	Not known
Number and density of markets	Either direction
Transport time to towns, markets and cities	Either direction
Density of roads	Either direction

In the spatial regression analysis, we use data that were developed at several different scales. As pointed out in Benson et al., (2004), pooling data from different scales in such an analysis leads to the risk of drawing inferences about smaller analytical units from the aggregate characteristics of a group made up of several of those units. This is not a problem for us since the spatial factors identified in Table 1 are all collected at more local scales than the Location-level. More details on our independent variables are found in the appendices. Tables A1 and A2 provide detailed definitions and data sources and descriptive statistics, respectively. Appendix B shows the geographic distribution of the variables.

3.2 Estimation Strategy

To model the prevalence of poverty as a function of selected spatial variables, we carried out two different analyses: (1) a simple ordinary least squares regression, and (2) a global spatial regression. We also analyze poverty at two different levels, national and provincial[6]. We include a Provincial-level analysis because large differences in average poverty levels exist across Provinces and thus we expect that there will also be significant differences in the determinants of poverty across Provinces.

[6] In related studies, we also analyze the determinants of poverty among poor livestock keepers and across development domains in Kenya.

3.2.1 Generalized OLS regression model

Applied to this context, we estimate the OLS regression model as:

$$y_i = \beta X_i + \varepsilon_i \qquad (3)$$

where *Y* is a vector of observations on the dependent variable; *X* is a matrix of independent variables; *ß* is a vector of coefficients, and *e* is a vector of random errors. The explanatory variables, X, are specific variables influencing poverty rates. These vectors may include spatial and non-spatial factors. The interpretation of the coefficients is straightforward, as with any conventional regression analysis.

Despite the popularity of this approach, problems of spatial autocorrelation limit its application in analyzing spatial relationships. As indicated earlier, spatial autocorrelation occurs if variables in one area are affected by the value of that same variable in a neighboring area. Because poverty in one location may in fact be influenced by poverty in a neighboring location, it is important to consider the nature of the spatial dependence inherent in the data. An alternative way in which the problem of spatial autocorrelation manifests itself is through the correlation of error terms. Error terms may be correlated spatially, as evidenced by observations from locations near each other having model residuals of a similar magnitude. Therefore, unless we correct for spatial autocorrelation, the assumptions of OLS regression are violated, and the estimates derived from this method are likely to be biased. To assess spatial autocorrelation, the clustering of the residuals from the OLS model will be examined using the Moran's *I* statistic.

3.2.2 Global spatial regression model

To control for spatial autocorrelation in the model so that the estimates of the model are both unbiased and efficient, we modify the model by including a supplementary explanatory variable. This variable is meant to represent the spatial dependency of the dependent variable. This is commonly done using the spatial lag of the dependent variable. In this case, the spatial lag of the dependent variable is defined as the weighted mean of a variable for neighboring spatial units of the observation unit in question (Anselin 2002).

There are two major ways in which spatial autocorrelation can manifest itself, referred to as spatial lag dependence and spatial error dependence. Spatial lag dependence refers to a situation in which the dependent variable in one area is affected by the dependent variable in nearby areas. For example, in this study, the spatial dependence could be a result of the level of poverty (in this case our dependent variable) in a location affecting the level of poverty in the location in question through, for example, trade or investment linkages. Such a relationship is modeled as a spatial lag model and can be written as follows:

$$y_i = \delta \sum_{j \neq 1} w_{ij} y_j + \beta X_j + \varepsilon_j \qquad (2)$$

Where y_i is the dependent variable for area i
δ is the spatial autoregressive coefficient
w_{ij} is the spatial weight reflecting the proximity of i and j
y_j is the dependent variable for area j
β is a vector of coefficients
X_j is a matrix of explanatory variables, and
ε_j is the error term.

The spatial weights matrix, *w*, represents the degree of proximity between each pair of spatial observations. It is a binary variable if the two areas are contiguous, or else a continuous variable based on a function of the distance between the two areas or locations. Omitting this adjustment will result in the coefficients being biased and inconsistent.

A second type of spatial dependence can be attributed to the error term of the model. In this case, the error for the model in one area or Location is correlated with the error terms in its neighboring locations (Anselin, 1992). This kind of spatial dependence occurs if there are variables that are omitted from the regression model but in fact do have an effect on the dependent variable and they are spatially correlated. For example, the type of administration in an area may affect income and poverty levels, but is not easy to include in a regression model. Since the type of local administration is likely to be spatially correlated (all Locations in a given Province may be affected by poor administration), the error term in each location is likely to be correlated with those in nearby Locations. Such a relationship can be modeled as a *spatial error model*:

$$y_i = \beta X_j + \lambda \sum_{j \neq 1} w_{ij} y_j \varepsilon_j + \varepsilon_i \qquad (3)$$

Where y_i is the dependent variable for area i
λ is the spatial autoregressive coefficient
w_{ij} is the spatial weight reflecting the proximity of i and j
y_j is the dependent variable for area j
β is a vector of coefficients
X_j is a matrix of explanatory variables, and
ε_j is the error term.

Here, the error term is disaggregated into the spatial lag of the error term of neighboring locations and the residual error term for the spatial unit in question. When there is spatial error dependence, OLS coefficients will be unbiased but not efficient (the standard errors will be larger than if there were no omitted variables) making interpretation of the significance results difficult (Anselin, 1992).

Spatial autocorrelation can be detected using standard global and local statistics that have been developed, including Moran's Index, Geary's C, G statistics (Getis, 1992), LISA (Ansellin, 1995) and GLISA (Bao and Henry, 1996 Whenever there is either spatial error or spatial dependence, an appropriate model can be used to correct for the problem. For spatial dependence, the spatial lag model is used. In the case of spatial error, we use the spatial error model. In practice, there is usually very little difference between the two spatial models. However, in order to select which model to use, a Lagrange Multiplier test is used to assess the statistical significance of the coefficients in each model, respectively. Where spatial autocorrelation is likely, usually the result of the test on each will be significant. The preferred model in such a case is the one with the highest Lagrange multiplier test value (Anselin & Rey, 1991).

In order to assess spatial autocorrelation, we must develop a spatial weights matrix to define exactly the 'neighborhood' for each rural aggregated location. There are many ways to assign neighbor weights and the choice depends on the type of spatial application and on the research question. This specification requires a *priori* knowledge of the range and intensity of the spatial covariance between regions. Common methods include row standardization, length of common boundary and distance functions (see Lee & Wong, 2001, pp.136-145; Anselin, 2002, pp. 256-260). In this study, we used distance band weights. We conducted sensitivity analysis of the results obtained using different weighting schemes.

4. Results of the econometric analysis

All the models we estimated used the Location-level poverty rate (the proportion of individuals falling below the national rural poverty line of KShs 1239/adult equivalent/month) as the dependent variable. We undertook analyses at the national level first (for at total of 2232 rural Locations), followed by models at the provincial level (i.e. for each of Kenya's 7 rural based Provinces).

The results of the national OLS regression of poverty incidence on the set of independent variables are shown in Table A2. The adjusted R2 is 0.53, indicating that one half of what drives estimated poverty rates across all of Kenya's Locations is not explained by this model. This OLS model likely has biased estimates due to spatial autocorrelation in the model residuals. All the diagnostic tests show the presence of spatial dependence, as they are all highly significant (Table 2). Thus we chose to use a spatial regression model to control for this spatial autocorrelation.

4.1 Spatial regression (error) model

We constructed a spatial weighting matrix based on distance-band weights. Matrices based on the nearest neighbors were also derived and tested, but the distance-band weighted matrix was preferred[7]. To determine the type of spatial dependence model best to use (spatial lag or spatial error), we chose the spatial error model, based upon the largest value of the robust Lagrange Multiplier indicators, as suggested by Anselin, et al., 1996 and Benson et al., 2005 (Table 2).

[7] See also Anselin 1996, 2002 for details

Table 2 Diagnostics for spatial dependence

FOR WEIGHT MATRIX :(row-standardized weights)		
TEST	VALUE	Probability
Moran's I	2.97322	0.00295
Lagrange Multiplier (lag)	0.77618	0.37831
Robust LM (lag)	6.87594	0.00874
Lagrange Multiplier (error)	5.35721	0.02064
Robust LM (error)	11.45697	0.00071

In the spatial error model, a full set of variables hypothesized to have some spatial relationship with community level poverty is included. The model fit increases to 0.54, which is not a huge change from the OLS model, but by removing the nuisance caused by spatial autocorrelation, we can now have more confidence in our parameter estimates, and concentrate on the variables that are showing a strong spatial relationship to poverty prevalence at the Location-level.

The results of the spatial error model show 18 of the 23 explanatory variables are significant (Table 3). We discuss the results by each group of independent variables.

Slope

Given our expectations, and the findings of related studies (Minot et al., 2000) of a strong relationship between slope of land and poverty, it is not entirely surprising that two of the four estimated slope parameters are significant. Thus, we find that relative to the very flat areas (0-4 percent slope), Locations that have a high percentage of land made up of steep slopes have higher poverty levels. The coefficient is largest for Locations with more than a 30 percent slope area, a result that is consistent with theoretical explanations that point towards serious erosion, cultivation and irrigation-related problems associated with steep land.

Table 3. Results of the Spatial error Model

Dependent Variable	Poverty incidence		
Variable	Coefficient	Std.Error	Probability
CONSTANT	0.86067	0.03176	0.00000
Demographic and inequality			
Average GINI coefficient	-1.10110	0.07548	0.00000
population density	-0.00006	0.00001	0.00002
Provincial dummy variables			
reg2 (Central)	-0.14724	0.01140	0.00000
reg3 (Coast)	0.06534	0.01443	0.00001
reg4 (East)	0.10896	0.00883	0.00000
reg5 (North Eastern)	0.23308	0.01480	0.00000
reg6 (Nyanza)	0.13937	0.00917	0.00000
reg8 (Western)	0.09460	0.01181	0.00000
Distance and travel time			
Mean_distance to nearest town of 200,000 people	-0.00003	0.00000	0.00000
Average travel time to Type 1 or 2 Road (minutes)	0.00001	0.0000	0.0776
Land use			
Percent of location under grass	-0.00144	0.00026	0.00000
Percent of location under farmland	0.00008	0.00015	0.57360
Percent of location wooded	0.00036	0.00014	0.01020
Percent of location that is Built-up	-0.01330	0.00282	0.00000
Rangeland (Dummy)	0.01262	0.00651	0.05267
Natural Factors			
Average Elevation (meters above sea level)	-0.00174	0.00077	0.02279
Percent of location with 4 - 8% slope	0.00118	0.00020	0.00000
Percent of location with 8 - 15% slope	0.00002	0.00024	0.94537
Percent of location with 15 - 30% slope	-0.00019	0.00031	0.54445
Percent of location with over 30% slope	0.00213	0.00035	0.00000
Percent of location with LGP less than 60 days	0.00026	0.00013	0.05257
Percent of location with LGP 180 days	-0.00032	0.00010	0.00169
Good soil (dummy)	-0.01132	0.00516	0.02820
LAMBDA	0.19945	0.08684	0.02163
Observations	2232		
Adjusted R-squared	0.5320		
Akaike info criterion:	-3891.33		
Log likelihood	1968.665		

Land use

Land use variables emerge as strong determinants of poverty among rural locations in Kenya. The coefficients for the ‘percentage of the Location under particular land uses’ show mixed results. As expected, Locations that have large areas that are built-up (occupied by buildings) tend to have lower rates of poverty. This suggests that built-up areas represent tendencies towards urbanization and more urbanization is expected to result in lower poverty. In general, the poverty maps (CBS, ILRI 2003) show that urban areas are richer than rural areas in Kenya.

Our results suggest that Locations with large areas under grassland are likely to have lower poverty rates, a somewhat non-intuitive result (see also CBS and ILRI, 2003 and CBS, 2005). It may be that this result is reflecting the fact that there are very few people in grasslands areas, or it may indicate this variable is capturing something else.

With respect to the percentage of wooded area, another non-intuitive finding is that Locations with more wooded areas are associated with higher poverty rates in rural areas (given that woodlands often provide nuts, fruits and firewood for poor families).

Soil

To address the question of how sensitive poverty is to quality of soil, a dummy variable for soil quality was included (good versus poor soils, described in more detail in appendix C). We expect that Locations with good soils are likely to have high agricultural potential and thus have absolute advantage in producing high-value perishable vegetables and other crops. Indeed, we found that Locations with greater proportion of area with good soils are associated with less poverty. The magnitude of effect is not large - about 1 percent - i.e. improving soil fertility (from poor to good soil) would reduce poverty by up to 1 percentage point in rural areas of Kenya's Provinces. This strongly points to the policy of improving soil quality through the use of fertilizers and soil conservation techniques.

Elevation

Measured in meters above sea level (masl), elevation has a significant negative effect on Location-level welfare - communities at higher elevation are likely to be less poorer. This is expected, since obviously many communities living in the highlands are much better off than their counterparts in many parts of dry lowlands of Kenya (see CBS, 2003).

Agro-climatic variables

Several variables meant to capture agro-climatic conditions were tested in this model. Rainfall and its coefficient of variation, NDVI and length of growing period (LGP) were among those variables. As expected, the variation in poverty among rural communities in Kenya is strongly influenced by agro-climatic factors. The results show that Locations with longer growing periods are likely to have lower poverty rates relative to areas with shorter growing periods. The effect here is clear as most crops such as (maize, beans, millet, sorghum, peas) require more than 60 days to mature.

Livestock and rangelands

For the livestock-related variables, the analysis shows that communities living in rangelands are likely have higher poverty levels. Data on total livestock units[8] are not available for the entire country so we estimate a separate model for these areas (rangelands) later. Our results suggest that even though we do not have complete livestock data for the entire country, there appears a strong positive relationship between poverty and living in the rangelands. Recent studies have shown the rangelands have some of the highest poverty rates in Kenya (CBS, 2005). This is fairly intuitive, as they are also the areas with poorest access to roads, services (education and health) and general infrastructure in the country. We further explore the determinants of poverty in livestock keeping areas (rangelands) in a related study.

Demography and income inequality

Both the demographic variable (population density) and inequality variable (measured by the gini-coefficient) have significant negative effects on poverty in rural areas. The use of these variables as independent variables is justified even if they were used in derivation of poverty estimates (see Elbers 2004). Areas with high population densities are associated with lower poverty rates. Population density influences labor intensity of agricultural production, including the choice of commodities as well as production technologies and land management practices, by affecting the land-labor ratio. This result implies that

[8] It is worth noting that this variable may be somewhat endogenous, and the causality may run both ways: livestock ownership increases the communities' income and consumption through sale or consumption of animals and animal products, but better off communities may also purchase livestock as a form of investments or safety net.

people tend to settle in areas where they can enhance their incomes, for example, through farming, and such areas end up having relatively low poverty levels. Inequality also presents interesting results. Areas that have a high level of income inequality tend to have lower poverty levels. This result is supported by the findings from CBS (2003) and CBS (2005), which show that areas with lower poverty (particularly urban areas) also have higher inequality. A similar result was obtained in Uganda (UBOS and ILRI, 2005) and Malawi (Benson et. al., 2005).

Roads and Market access

Better roads and/or access to markets are expected to favor production of high-value products and non-farm activities that will contribute to higher incomes or lower poverty. The results of this study show that longer travel times to tarmac and murram roads significantly increase poverty levels. The standard explanation here is that the greater the travel time to a good road, the more difficult it is access markets, limiting livelihood options. Conversely, communities that have greater access to markets, good infrastructure (health, education) and public administration face lower transactions costs and more livelihood options, leading to lower poverty levels. The above results point towards the need for investment in improved rural roads if poverty is to be reduced in Kenya.

Provincial dummy variable

Finally, we investigate the evidence of regional heterogeneity regarding the effects of different spatial determinants on poverty. Thus, for all rural locations, we test for equality of parameter estimates for all Provinces except Nairobi and find that the homogeneity hypothesis is strongly rejected. Our reference Province is Rift Valley and the results show that with the exception of Central Province, all other Provinces are associated with higher poverty levels relative to Rift Valley Province. It is worth noting that the provincial dummies may be capturing a number of factors in the different regions (such as security, administration, infrastructure, culture) which are not captured in the other spatial variables. This heterogeneity strongly justifies the need for Province-specific estimations, the results and discussion of which are presented below.

We also explored the effects of spatial factors when we restricted the national-level rural poverty regression to include only variables that are likely to be exogenous to poverty, referred to as selective models. Restricting the model in this way helps us to explore the relative importance of the spatial explanatory variables. Variables representing distance and demographic characteristics were not included in the first selective model. Keeping the provincial dummy variables and excluding these variables (i.e. population density, distance to hospitals and major towns and income inequality) reduced the explanatory power by 5 percentage points, to 48 percent. When the dummy variables representing the seven Provinces were also excluded in the second selective model, the explanatory power of the model reduced to 36 percent. The exogenous spatial variables mainly land use and natural factors on their own are able to explain 36 percent of the variability in poverty rates that we see across Kenya.

The results of the national level analysis suggest there are concerns with the variations and significance of the variables. The Provincial related variables may be picking some omitted variables and yet they explain a high percentage of variation in rural poverty across locations, hence the need for Provincial-specific analysis and checking whether similar problems emerge with all or some Provinces.

4.2 Provincial determinants of poverty in Kenya

Separate models were run for each of the 7 Provinces in order to capture the differences in spatial poverty determinants across these very diverse Provinces. Table B3[9] shows the diagnostic tests for spatial dependence that identify the correct model to use and Tables B4-B10 present the model results for each Province, after correcting for spatial lag or spatial error problems, depending on the type of spatial dependence evident. Six of the seven Provinces showed significant presence of spatial dependence, mainly of the spatial lag type, except Central Province. North Eastern Province showed no presence of spatial autocorrelation and therefore we discuss their results based on the OLS estimates. We now turn to the discussion of regression results by Province. The variables that are significant for each of the Provinces, as well as at the national level are summarized in Table 4.

[9] Detailed results for Provincial estimates are in the appendices

Table 4. Summary: Provincial Determinants of Poverty

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	■ ■ ■	■ ■ ■	■ ■ ■	■ ■ ■	■ ■ ■	[illegible]
[illegible]	■ ■	[illegible]	■ ■	■ ■	■ ■	[illegible]	[illegible]
[illegible]	■ ■ ■		■ ■ ■		■ ■ ■	[illegible]	[illegible]
[illegible]							
[illegible]	■ ■ ■	[illegible]	■ ■ ■		■ ■		■
[illegible]							
[illegible]		■			■ ■	[illegible]	
[illegible]				[illegible]		■ ■ ■	
[illegible]		[illegible]	■ ■		[illegible]	[illegible]	■ ■
[illegible]		[illegible]	■ ■		[illegible]	[illegible]	[illegible]
[illegible]			[illegible]	[illegible]			
[illegible]	■ ■	■ ■	■ ■		[illegible]	[illegible]	
[illegible]							
[illegible]			■ ■ ■				■ ■ ■
[illegible]	[illegible]						
[illegible]		■ ■ ■	■ ■	■	■	■	[illegible]
[illegible]		■ ■ ■	[illegible]		[illegible]	[illegible]	[illegible]
[illegible]			■ ■			[illegible]	■ ■
[illegible]			[illegible]			■ ■ ■	
[illegible]	■ ■ ■	■	■ ■ ■			[illegible]	[illegible]
[illegible]		■ ■					
[illegible]	■ ■ ■			■ ■			
[illegible]					[illegible]		
[illegible]		[illegible]	[illegible]			[illegible]	
[illegible]		■ ■ ■	■ ■			■ ■ ■	
[illegible]							
[illegible]	■ ■	[illegible]	[illegible]		■ ■	[illegible]	
[illegible]			[illegible]	■			

[illegible]							
[illegible]	..				[illegible]		[illegible]
[illegible]		[illegible]		..			.
[illegible]		[illegible]		[illegible]			
[illegible]	[illegible]	[illegible]				[illegible]	
[illegible]	[illegible]				[illegible]		[illegible]
[illegible]	...	...	...		...	...	...

[illegible] ... [illegible] ** Significant at 5 percent, * Significant at 10 percent, ns- Not sig

Central Province

In Central Province, 164 Locations are considered and the model fit is 0.50 (Table A5). The results show that limited access to roads is associated with higher poverty levels. The longer the travel time from the Location center to the nearest road (track, murram or tarmac), the poorer it is. Roads provide crucial access to markets and the result obtained here suggests that areas where it takes people a long time to reach a good road are typically poorer communities. Similarly, our findings show that Locations in Central Province that are mainly rangelands or are further away from public forests and on higher elevation are associated with higher poverty levels.

In contrast, higher population density and proportion of the wetland area of a Location are associated with lower poverty levels. The wetlands result suggests that people near wetlands may have enhanced livelihoods and it would be interesting to explore further what ecosystem goods and services they are benefiting from due to presence of these wetlands. Other factors were not significant in this Province.

Coast Province

A spatial lag model was estimated for this Province and the results (Table A6) show higher income inequality levels, percentage of the Location under wetlands, percentage of the Location with an 8-15% slope, probability of flooding and average length of growing period of 180 days or greater are associated with lower poverty rates. As expected, Locations with longer growing periods and thus much higher cropping potential are likely to be less poor.

Also among the significant variables, we see that the greater the percentage of the Location that is under water (water logged), with a slope of 4-8 percent, and travel time to the nearest road (feeder or murram) the lower the poverty. This reinforces the pattern displayed in the national and Central Provincial poverty results. The greater the distance from a Location center to the nearest tarmac or murram road, the higher the poverty. This reflects the importance of access to decent roads to community welfare levels.

Eastern Province

The Eastern Province model was also a spatial lag model that was able to explain 52% of the variation in poverty across Locations in this Province (Table A7). Locations that are relatively further from the nearest public forest, have 4-8% and 15-30% slopes, have more area under protected area and more farmland are poorer. Being far from a public forest has a highly significant influence on living standards. This suggests many people rely on forest resources such as firewood, fruits, nuts, charcoal and herbs.

Similar to the findings for Coast Province, variables that were significant and associated with lower poverty rates included income inequality, population density, elevation, proportion of the Location under wetlands, grasslands and Locations with an average growing period of 180 days or greater. These results portray the importance of agricultural potential and land use in poverty reduction.

North Eastern Province

We note that data from North Eastern Province should be treated with caution (Table A8). The poverty estimates used for North Eastern are derived estimates from the model for Coast Province, since the Household Budget Survey for 1997, which was used to estimate Location-level poverty levels for all the other Provinces, was not implemented in this Province due to security-related reasons. The model estimated is an OLS because there was no evidence of spatial autocorrelation. Perhaps not surprisingly in this arid Province, The coefficient of variation of rainfall stands out as a major determinant of poverty. In this region, Locations with higher rainfall variability also tend to be relatively poorer. Two distance-based variables are also key determinants of poverty in this Province. Distance to the nearest health centre and distance to the nearest town with 10,000 people are both positive and significant, i.e. the further the distance to the nearest health centre or town, the higher the poverty incidence. Health-related issues have been found to be a major factor influencing household descents into poverty in Kenya (Kristjanson et al., 2003), and this analysis suggests that accessibility to health services is important. Likewise, the distance to the nearest town is very important, and likely reflects the fact that basic services, including education, health, trade, and security are found in these towns.

Nyanza Province

For Nyanza Province, a spatial lag model was estimated (Table A9). Higher elevation, longer distance to the nearest public forest, a higher percentage of the Location under water, and longer distances to the nearest health facility are associated with higher poverty rates in Nyanza Province. Conversely, and as found elsewhere, the higher the income inequality and population density, the lower the poverty. Distance to the nearest city of 200,000 people is significantly and negatively related to poverty, i.e the further from a city, the lower the poverty, a non-intuitive result that suggests the benefits and services derived from large towns are equally available in small towns in Nyanza, hence nearness to large towns is not as important as it is in other less densely populated areas of Kenya.

Rift Valley Province

A spatial lag model for Rift Valley's 785 Locations was run (Table A10). The results from this Province provide very insightful information. The model fit is 0.60 and a number of variables have the expected sign. Variables representing rainfall and associated indicators, namely length of growing period, are significantly related to poverty. Once again, we found Locations with longer growing periods are associated with lower poverty rates. The higher the percentage of a Location that is built-up, the lower the poverty incidence. Inequality is consistently negative and significant, as found in other Provinces. In terms of variables associated with higher poverty levels, the significant indicators include flooding potential and slope. Areas with high potential for flooding are more vulnerable, other things being equal, and not surprisingly this has a negative impact on living standards.

Larger areas of sloping land may not be as conducive for settlement and farming, thus areas with larger slopes tend to have higher poverty rates.

Western Province

With the exception of a few variables, this Province has some interesting results which were not significant in the other regions (Table A11). Locations that are further away from small towns or have long distances to travel to reach public forests and protected areas, have higher poverty levels. This is consistent with the earlier notion that distance to facilities and resources is an important determinant of poverty. Locations with majority good soil have lower poverty. Conversely, the findings for Western Province suggest that the larger the percentage of grassland, and the more area with slopes of 15-30%, the lower the poverty, both results that are not very intuitive.

These findings suggest that the relationships between poverty and geographic factors do vary significantly and spatially in their effects across rural Kenya. It is clear that some spatial variables are important in influencing poverty in certain Provinces and not in others. Such a finding is critical for the formulation and targeting of anti-poverty programs. These results can be used to guide local actions aimed at reducing poverty.

5. Simple Poverty Simulations

5.1 The Methodology

Having estimated the poverty determinants, we can now generate simulations to predict reductions or increases in general poverty levels that result from changes in selected spatial characteristics. The purpose of these simulations is to illustrate how changes in levels of the determinants will alter aggregate poverty levels. These changes are such as those, which may result from the implementation of specific government policy aimed at reducing poverty. We briefly describe the methodology.

Using the estimated parameters of the model(s), we generate predictions of new poverty rates for every Location when the level of a particular independent variable x_j is changed. Of course, not all of our independent variables are amenable to policy changes – e.g. rainfall or slope – thus we target those that can be influenced by investments, such as roads and soils. The changes in explanatory variables result in changes in the predicted probabilities, and these are taken to be the effect of the policy. We do not consider higher order changes in this study. Because the results of the simulations assume that the considered changes in the determinant variables do not affect the model parameters or other exogenous variables, these results need to be interpreted as indicative only. While this is a plausible assumption for incremental changes, it warrants a more cautious interpretation for simulations that involve 'large' policy changes.

We simulate interventions aimed at reducing the proportion of poor people in a Location. When interpreting the simulation results, it is important to note that changes in poverty for each simulation will depend essentially on:

i) The magnitude and sign of the coefficients from the regression
ii) The proportion of the population affected by the simulation
iii) The size of the change considered in the determinants variable.

It is also important that we consider the resultant effects of the simulations as instantaneous because we estimate them from static models. In reality, the effects on community poverty realized from a change in an agricultural variable (say fertilizers for soil improvement) will only be observed in the next growing season, and the benefits from road construction will only be realized when the road is complete and market forces informed, perhaps two years later.

5.2 The simulations

Access to services and infrastructure simulation

Our simulations involve changing the variables at provincial level since the national results may be able to derive accurate inference. We choose to change variables that are significant and amenable to change in three of the 7 Provinces, namely: Central, Eastern and Western Province.

First, we consider the potential impact of a reduction in the travel time it takes to reach the nearest tarmac (All Weather Bound) or murram (All Weather Loose) or track road from the Location centre in Central Province. We reduce travel time to roads to one hour for all

Locations that have travel times of more than one hour (which is the mean travel time to the nearest road in this Province). In this simulation, we are trying to capture improvements in national road infrastructure as a means of improving accessibility of rural communities to markets and general infrastructure. Table A12 in the Appendices shows the possible impact of changes in certain selected variables. The analysis of this simulation highlights an important result: other factors constant, a general reduction in poverty in Central Province. The results show that a reduction in travel time, on average, from greater than 1 hour, to less than one hour to the nearest track, murram or tarmac roads within all Locations in Central Province could potentially lower average Location-level poverty rates by 0.8% (or the average Province-level poverty rate from 31.3 percent to 30.5 percent, which would imply 21,649 poor people escaping poverty. The result for Eastern Province is equally small (0.8 %).

Perhaps the disappointing aspect of this simulation is that the expected reduction in poverty is very small. This result holds true in terms of poverty reduction when we look at the sign of the coefficient. However, it should be noted that the small coefficients are a result of a change in only one variable. Roads alone may not be the panacea to the poverty problems in Central. There is need to consider other factors in this simulation. For example, easy access to good roads combined with high agricultural potential (better soils and reliable rainfall) may lead to larger reductions than roads alone. As already mentioned, we do not consider higher order changes in this study since the results of the simulations assume that the considered changes in the determinant variables do not affect the model parameters or other exogenous variables, that is these results are indicative only.

5.2.2 Soil

We simulate the potential direct impact of a change in soil fertility on poverty incidence in Western Province. Though soil did not show as a significant variable in the spatial models, we simulate its impact based on socioeconomic evidence about Western Province. This Province generally has among the highest poverty rates in Kenya, high HIV prevalence, small farm sizes, and high population density (Kristjanson et al., 2004). Soil was classified for each Location as being, on average, either good[10] for agricultural production or poor. Here we are not capturing the fact that soil fertility will be affected by other factors, particularly rainfall and slope. We assumed increases in soil fertility in all Locations classified with poor soils, and with more than 60 percent poverty rate.

The results suggest the poverty rate for Western Province could be lowered by 9.4 percentage points with investments leading to a change in average soil fertility from poor to good across all Locations that have poor soils and poverty levels above the mean for the Province (59%). Soil fertility improvements can be achieved through higher rates of fertilizer application and/or better soil management techniques, and thus the costs of such investments are difficult to estimate and would require further research. However, this result is indicative of the potentially substantial impact on poverty from improvements to soil fertility levels in western Kenya (a finding supported by numerous other studies). This approach, linked with further research, has potential for being able to quantify the potential costs, benefits, and impacts on poverty of investments aimed at enhancing soil fertility.

6. Conclusions and implications for policy

In this paper, we have sought to improve our general understanding of how (and which) spatial factors are related to poverty and how this varies across Kenya's diverse landscapes, how much of the variation in poverty incidence across Kenya can be explained by environmental/spatial factors, and how this approach can be used to evaluate the potential impact on poverty levels of investments in factors found to have a significant influence on poverty incidence. Our approach to modeling the geographic determinants of poverty is to use spatial regression techniques to investigate the impact of specific environmental variables on poverty in rural Kenya. The dependent variable is the location level poverty rate (P_0), and the explanatory variables include a wide range of spatial (GIS) variables such as; land use, slope, climate, elevation, distance and travel time to cities and public resources.

In the context of the theoretical expectations, some variables like distance to towns and

[10] Good soils in agriculture are those soils rich in nutrients with good drainage and readily workable. They have good texture lacking in high clay contents but instead having a balanced texture. They should also have a well developed profile thus lacking in stoniness at shallow depths.

land use did not have expected correct signs, hence the need for careful interpretation of their coefficients.

The results of the regression models demonstrate the statistical significance of certain spatial variables. At the national level, the set of important variables is diverse including regional dummies, land use, elevation, soil conditions/quality, and length of growing period, travel time to roads and towns (market access) and demographic conditions. This suggests the presence of a poverty-environment relationship and hence the impact of environmental factors on the welfare of the poor and on poverty reduction efforts. However, the strength of the provincial dummy variables shows that Provinces in Kenya are not homogenous. For this reason, different spatial indicators could be important in different Provinces, hence the need for a provincial level analysis. These region specific and not national level variables could be important for designing and evaluating Provincial specific poverty reduction strategies.

Our simulation results for three Provinces suggest that increasing access to roads and improving soil conditions would result in decline in the number of poor people in these Provinces. In Western Province, improving soil conditions in Locations with poor soil and high poverty rates (above 59 percent) would result in a 9 percent reduction in poverty levels across Locations in Western Province. We find the beneficial impact effect of improved soil quality is robust to whether we consider Locations with high or lower rates and proportion of the land that is arable.

While the beneficial effects of improvements in soil conditions in Western Kenya are significantly larger than the effects from other policy changes (such as road improvement in Central and Eastern Province), we do find positive effects from improvements in travel time to all the three different types of roads (track, murram and tarmac). Other variables, for example rainfall could be considered in these simulations but they are not amenable to change. The observed effects are, nevertheless, important even if particular pathways are difficult to identify. Among the key messages learned from the policy simulations is that it is important to improve access to roads (Central and Eastern) and increase agricultural productivity of soils (Western) through encouraging investment in soil conservation structures.

Since the results suggest that different spatial factors are important in different Provinces, the design and implementation of any poverty reduction strategies can be Province specific. However, in interpreting the importance of the results for poverty reduction, one should not assume these effects are instantaneous, even though we estimated them from static models. Road investments, for instance, have inherently long gestation while soil improvements can have immediate effects during the next planting season. Our results indicate that these variables can have powerful effects in terms of long-term reduction in poverty.

A few challenges were encountered in the implementation of this study. The data used was collected by a number of institutions with different interests. Combining this data was no easy task since the location identifiers were not unique. As such, some data was lost in the process of merging due to such inconsistencies. This exercise requires significant time and GIS skills. Also, there was a considerable degree of measurement error for a number of spatial variables on which spatial data were collected, including for instance rainfall, distance to facilities, and proportions of land use under different systems. While a considerable amount of effort was spent in cleaning the data (including computations and estimations using GIS techniques), the existence of measurement errors influenced the specification choices that were made in the analytical work. These limitations suggest the need for improvement in future spatial data collection, and the need for careful interpretation of the results in this study. It is therefore more judicious to focus on broad regularities than on exact numbers.

Finally, it should be reiterated that while this analysis has helped explain the geographic determinants of poverty, there is need to refine and extend this analysis, including more disaggregate analysis following development domains in Kenya, as well incorporating supplementary information from other data sources such as the livestock and agricultural census.

References

Aiwang, J.; P.B. Siegel; & S.L. Jørgensen. 2001. Vulnerability: A view from different disciplines. Social Protection Discussion Paper Series no. 0115. Social Protection Unit, Human Development Network, The World Bank, Washington, D.C. Photocopy.

Anselin, L. 2003. GeoDa 0.9 User's Guide. Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics, University of Illinois and Center for Spatially Integrated Social Science. June 15, 2003 revision. Photocopy.

Anselin, L., 2002. Under the hood: Issues in the specification and interpretation of spatial regression models. Agricultural Economics. 27: 247-267.

-----. 1992. Spatial data analysis with GIS: An introduction to application in the social sciences. Technical report 92-10, National Center for Geographic Information and Analysis, University of California - Santa Barbara. Photocopy.

Anselin, L & S. Rey. 1991. The performance of tests for spatial dependence in linear regression. Technical report 91-13, National Center for Geographic Information and Analysis, University of California - Santa Barbara. Photocopy.

Anselin, L., 1988. Spatial Econometrics: Methods and Models. Dordrecht: Kiuwer.

Benson, T.; with J. Kaphuka; S. Kanyanda; & R. Chinula. 2002. Malawi - An atlas of social statistics. Washington, DC: International Food Policy Research Institute & Zomba, Malawi: National Statistical Office.

Benson, T., Chamberlin, J., and Rhinehart, I. 2005. Why the Poor in Rural Malawi Are Where They Are: An Analysis of the Spatial Determinants of the Local Prevalence of Poverty. IFPRI Discussion papers 198.

Brooks, N. 2003. Vulnerability, risk and adaptation: A conceptual framework. Tyndall Centre

Baker, J. and Grosh, M., 1994. "Poverty reduction through geographic targeting: How well does it work?" World Development, Vol. 22, No. 7: 983-995

Baulch, B., 2002, "Poverty monitoring and targeting using ROC curve: Examples from Vietnam", IDS Working Paper 161, Brighton: Institute of Development Studies

Baulch, B., Truong, C, Haughton, D. and Haughton, J, 2003. "Ethnic minority development in Vietnam: a socio-economic perspective."

Bigman, D. and Fofack, H, 2000. Geographic Targeting for Poverty Alleviation: Methodology and Applications, Washington DC: World Bank Regional and Sectoral Studies

Brunsdon C., Fotheringham A. S. and Charlton M. E., 1996. "Geographically Weighted Regression: A method for exploring spatial nonstationarity", Geographic Analysis Vol. 28, No. 4: 281 – 298

Central Bureau of Statistics (CBS). 2003. Geographic Dimensions of Well-Being in Kenya. Vol. 1: Where are the Poor? From Districts to Locations. Government of Kenya, Ministry of Planning and National Development, Central Bureau of Statistics (CBS) in collaboration with International Livestock Research Institute (ILRI). CBS and ILRI, Nairobi, Kenya. 164 pp.

Central Bureau of Statistics (CBS). 2005. Geographic Dimensions of Well-Being in Kenya. Vol. II: Who and Where are the Poor? A Constituency Level Profile Government of Kenya, Ministry of Planning and National Development, Central Bureau of Statistics (CBS). Nairobi, Kenya. 164 pp.

Christiaensen, L.J. and K. Subbarao. 2001. Towards an understanding of vulnerability in rural Kenya. Paper presented at IFPRI-World Bank Conference on Risk and Vulnerability:Estimation and Policy Implications. September 23-24, 2002. Washington, D.C., USA. Photocopy.

Conway, T., 2001, "Using government data to target activities to poor communes and monitor poverty reduction: a review of options for the Cao Bang-Bac Kan Rural Development Project", Hanoi: Commission of the European Communities

Datt G., and Jolliffe, D**.** 1999.Determinants of Poverty in Egypt: 1997 . IFPRI FCND Discussion paper no. 75

Demombynes, G; C. Elbers; J. Lanjouw; P. Lanjouw; J. Mistiaen; and B. Ozler. 2002. Producing an improved geographic profile ofpoverty: Methodology and evidence from three developing countries. United Nations University (TJNU)/World Institute for Development Economics Research (WIDER) Discussion Paper no. 2002/39. UNU/WIDER, Helsinki, Finland. Photocopy.

Dercon, S. 2001. Assessing vulnerability to poverty. Paper prepared for the Department for International Development. Department of Economics, Oxford University, UK. Photocopy.

DflD (Department for International Development). 1999. Sustainable livelihood guidance sheets. http://www.1ivelihoods.org/info/info_gUidaflceSheetS.html. Accessed July 2004.

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

-----. 2000. Welfare in villages and towns: Micro-measurement of poverty and inequality. Tinbergen Institute Discussion Papers 2000-029/2. Tinbergen Institute, Amsterdam and Rotterdam, The Netherlands. Photocopy.

Epprecht, M. and Muller D., 2003. Linking People and the Landscape: GIS and Spatial Analytical Techniques for Agricultural Economists, presentation at the International Conference of International Association of Agricultural Economists, Durban, 2003

Foster, J.; Greer, J and Thorbecke, E. 1984. A class of decomposable poverty measures. Econometnca. 52(3): 761-765.

Fotheringham, A.S., Brunsdon C. and M. Charlton. 2002. Geographically Weighted Regression - The Analysis of Spatially Varying Relationships. Chichester, UK: John Wiley & Sons.

Government of Kenya, 2001, *Poverty Reduction Strategy Paper (PRSP) for the Period 2000-2003*. Nairobi: The Government Printer

Hoddinott, J. and Quisumbing A.R. 2003. Data sources for microeconometric risk and vulnerability assessments. Social Protection Discussion Paper Series no. 0324. Social Protection Unit, Human Development Network, The World Bank, Washington, D.C. Photocopy.

Henninger, N. and Snel, M. 2002. Where are the poor? Experiences with the development and use of poverty maps. World Resources Institute, Washington, D.C. and UNEP- GRID/Arendal, Arendal, Norway.

Hentschel, J., Lanjouw, J., Lanjouw, P. and Poggi, J., 2000. "Combining census and survey data to trace the spatial dimensions of poverty: a case study of Ecuador", World Bank Economic Review, Vol. 14, No. 1: 147-65

Kyereme and Thorbecke, E. 1987 Food Poverty Profile and Decomposition Applied to Ghana, World Development, 15(9): 1189-1199.

Lee, J. & D.W.S. Wong. 2001. Spatial Analysis with ArcView GIS. New York: John Wiley & Sons.

Minot, N.; B. Baulch, & M. Epprecht. 2003. Poverty and inequality in Vietnam: Spatial patterns and geographical determinants. A report produced in collaboration with the Inter-Ministerial Poverty Mapping Task Force of the government of Vietnam. International Food Policy Research Institute, Washington, DC, USA & Institute of Development Studies, University of Sussex, Brighton, UK. Photocopy.

Minot, N., 2000. "Generating disaggregated poverty maps: An application to Vietnam", World Development, Vol. 28, No. 2: 3 19-331

Minot, N. and B. Baulch. 2002a. "The spatial distribution of poverty in Vietnam and the potential for targeting." Discussion Paper No. 42. Markets and Structural Studies Division, International Food Policy Research Institute, Washington, D.C.

Minot, N. and B. Baulch. 2002b. "Poverty mapping with aggregate census data: What is the loss in precision." Discussion Paper No. 49. Markets and Structural Studies Division, International Food Policy Research Institute, Washington, D.C.

Mwabu, Germano, Wafula Masai, Rachel Gesami, Jane Kirimi, Godfrey Ndeng'e, Tabitha Kiriti, Francis Munene, Margaret Chemngich and Jane Mariara, 2000, "Poverty in Kenya: Profile and Determinants", Nairobi: University of Nairobi and Ministry of Finance and Planning.

Okwi, P., Hoogeveen, H, Emwanu, T, Linderhof V., and Beguman J.. 2005 Welfare and Environment in Rural Uganda: Results from a Small-Area Estimation Approach PREM Working Paper 05-04 - (download: http://www.prem-online.org/index.php?p=publications&a=show&id=51)

Oyugi, Lineth Nyaboke, 2000, "The Determinants of Poverty in Kenya" (Unpublished MA Thesis, Department of Economics, University of Nairobi). 14

Ravallion, M. 1992. "Poverty comparisons: A guide to concepts and methods", Living Standard Measurement Working Paper No. 88, Washington DC: World Bank

Ravallion, M and Datt, G. 2002. Why Has Economic Growth Been More Pro-Poor in Some States of India than Others? - Journal of Development Economics.

StataCorp, 2001. Stata Reference Manual, Release 7. College Station, Texas: Stata Press.

Tobler W., 1990. Frame Independent Spatial Analysis, pp. 115-122 of M. Goodchild, ed., Accuracy of Spatial Data Bases, Taylor and Francis, London.

Van de Walle, D. 2002."Choosing rural road investments to help reduce poverty." World Development 30 (4): 575-589.

Van de Walle, D. and Gunewardana, 2001. "Sources of ethnic inequality in Viet Nam", Journal of Development Economics, Vol 65: 177-207

World Bank, 2000. Panama Poverty Assessment: Priorities and Strategies for Poverty Reduction, Washington DC: World Bank Country Study.

Working Paper 38. Norwich, UK: Tyndall Centre for Climate Change Research & Centre for Social and Economic Research on the Global Environment, School of Environmental Sciences, University of East Anglia. Photocopy.

Appendices

Table A1. Description of Variables

Short description	Source	Explanation
Agroclimatological		
Annual Rainfall (mm)	The WorldClim interpolated global terrestrial climate surfaces. Version 1.3.	The average annual rainfall within the location boundaries, calculated as the sum of all the monthly rainfall figures derived from the original Worldclim1.3 dataset of monthly layers.
Rainfall coefficient of variation	The WorldClim interpolated global terrestrial climate surfaces. Version 1.3.	The average coefficient of variation (CV) of rainfall between the months within 1 year within the location boundaries. This variable was derived from the worldclim1.3 dataset of bio-climatic information, which describes the "rainfall seasonality".
Distance and Access to services		
Travel time to municipality	- Africover landcover multipurpose database (FAO) - NASA, Shuttle Radar Topography Mission (SRTM) - World Database on Protected Areas (WDPA - sea.unep-wcmc.org/wdbpa) - Roads - ASARECA - Settlements - CBS	This variable represents the average travel time from any place within the location to the nearest municipality (according to definitions of CBS). Travel time is a function of slope, road type and "impediments" (i.e. wetlands, water bodies and natural parks). The table below summarizes the travel times:
Travel time to town	Idem above	This variable represents the average travel time from any place within the location to the nearest town (according to definitions of CBS).
Travel time to trade centre	Idem above	This variable represents the average travel time from any place within the location to the nearest trade centre (according to definitions of CBS).
Travel time to market centre	Idem above	This variable represents the average travel time from any place within the location to the nearest market centre (according to definitions of CBS).
Travel time to type 1 road	Idem above	This variable represents the average travel time from any place within the location to the nearest road of type 1. *Type 1: Tarmac/All Weather Bound* *Type 2: Murram/All Weather Loose* *Type 3: Earth/Dry Weather*
Travel time to type 1 or 2 road	Idem above	This variable represents the average travel time from any place within the location to the nearest road of type 1 or 2.
Travel time to type 1, 2 or 3 road	Idem above	This variable represents the average travel time from any place within the location to the nearest road of type 1, 2 or 3.
Travel time to type 1, 2 or 3 road	Idem above	This variable represents the average travel time from any place within the location to the nearest road of type 1, 2 or 3.
Land use		
Percent of location under Protected Area	World Database on Protected Areas (WDPA - sea.unep-wcmc.org/wdbpa)	This variable represents the percent of location that is under the Protected Area.
Percent of location under Wetlands	Africover landcover multipurpose database (FAO)	The original land cover was produced from visual interpretation of digitally enhanced LANDSAT TM images (Bands 4,3,2) acquired mainly in 1999. Wetland areas are extracted on the basis of code1 of the original layer (considered to be wetland areas)

Percent of location Arable land (I.e. LGP > 60 days)	Jones P.G., 2004. Report on preparation of growing season days coverages for Hadley 3 scenarios A2 and B2, Consultant's report, ILRI	The variable describes the percentage of the location that is arable. Arable land was defined as land with a length of growing period of more than 60 days per year.
Arable land between 30 and 60 % (1=yes ; 0=no)	Jones P.G., 2004. Report on preparation of growing season days coverages for Hadley 3 scenarios A2 and B2, Consultant's report, ILRI.	This variable takes a value of 1 if the arable land is 30-60% of the location's area, and 0 otherwise. Arable land was defined as land with a length of growing period of more than 60 days per year.
Percent of location under water	Africover landcover multipurpose database (FAO)	The original land cover has been produced from visual interpretation of digitally enhanced LANDSAT TM images (Bands 4,3,2) acquired mainly in 1999. Water areas extracted on the basis of code1 of the original layer (considered to be water bodies: 7WP, 7WP-Y, 8WFP).
Percent of location that is Built-up	Africover landcover multipurpose database (FAO)	The original land cover has been produced from visual interpretation of digitally enhanced LANDSAT TM images (Bands 4,3,2) acquired mainly in 1999. Build-up areas extracted on the basis of code1 of the original layer (considered to be build-up areas: 5U, 5UC, 5UR, 5I, 5A).
Percent of location under forest	Africover landcover multipurpose database (FAO)	The original land cover has been produced from visual interpretation of digitally enhanced LANDSAT TM images (Bands 4,3,2) acquired mainly in 1999. Forest areas extracted on the basis of code1 of the original layer (considered to be forested areas). The resulting shapefile was converted to a raster with the following values: 100 = forest (covering about 100% of the area); 65 = mixed forest (covering approx. 65% of the area; 0 = non-forest
Percent of location under farmland	Africover landcover multipurpose database (FAO)	The variable contains the percentage of the location's area that is under agricultural land. The original land cover has been produced from visual interpretation of digitally enhanced LANDSAT TM images (Bands 4,3,2) acquired mainly in 1999. Farming areas were extracted on the basis of code1 of the original layer (considered to be agricultural areas). The resulting shapefile was converted to a raster with the following values: 100 = agriculture (covering about 100% of the area); 65 = mixed agriculture (covering approx. 65% of the area); 0 = non-agriculture
Percent of location under grass	Africover landcover multipurpose database (FAO)	The original land cover has been produced from visual interpretation of digitally enhanced LANDSAT TM images (Bands 4,3,2) acquired mainly in 1999. Grass areas extracted on the basis of code1 and code2 of the original layer (considered to be grassland areas)
Natural factors		
Arable land more than 60 % (1=yes ; 0=no)	Jones P.G., 2004. Report on preparation of growing season days coverages for Hadley 3 scenarios A2 and B2, Consultant's report, ILRI.	This variable takes a value of 1 if the arable land is more than 60% of the location's area, and 0 otherwise. Arable land was defined as land with a length of growing period of more than 60 days per year.
Percent of location with Arid or Semi-Arid land (i.e. LGP <= 180 days)	Jones P.G., 2004. Report on preparation of growing season days coverages for Hadley 3 scenarios A2 and B2, Consultant's report, ILRI.	This variable describes the percentage of the location that is arid or semi-arid (ASAL). ASAL was defined as land with a length of growing period of less than 180 days per year.
Elevation (masl)	NASA, Shuttle Radar Topography Mission (SRTM)	The average elevation in meters above sea level within the location.

Percent of location Steep land (I.e. > 10%)	NASA, Shuttle Radar Topography Mission (SRTM)	This variable represents the percentage of the location's area that is defined as steep. Steep land was defined as having a slope of more than 10%. The slope was calculated based on the elevation and can be expressed in degrees or percent.
Percent of location with 0 - 4% slope	NASA, Shuttle Radar Topography Mission (SRTM)	The percentage of the location's area with a slope between 0 and 4 %.
Percent of location with 4 - 8% slope	NASA, Shuttle Radar Topography Mission (SRTM)	The percentage of the location's area with a slope between 4 and 8 %.
Percent of location with 8 - 15% slope	NASA, Shuttle Radar Topography Mission (SRTM)	The percentage of the location's area with a slope between 8 and 15 %.
Percent of location with 15 - 30% slope	NASA, Shuttle Radar Topography Mission (SRTM)	The percentage of the location's area with a slope between 15 and 30 %
Percent of location with over 30% slope	NASA, Shuttle Radar Topography Mission (SRTM)	The percentage of the location's area with a slope of more than 30 %.

Table A2: Descriptive statistics

Variable	Label	Mean	Std. Dev.	Min	Max
avg_gini	Average gini	0.31	0.04	0.19	0.55
popden	Population density	205.64	226.05	0.12	2431.78
Elevation	Elevation	1345.19	659.38	2.82	3087.83
distance to forest	Distance to forest (m)	5658	8114	0.00	46756
Perc_water	Percent of location under water	0.44	2.63	0.00	60.20
Perc_built	Percent of location built	0.14	0.79	0.00	13.85
Perc_for	Percent of location under forest	4.42	11.13	0.00	84.66
Perc_farmland	Percent of location under farmland	28.28	27.26	0.00	97.95
Perc_grass	Percent of location under grassland	17.42	14.26	0.00	82.11
Perc_wooded	Percent of location under wooded	20.35	22.24	0.00	100.00
Perc_prota	Percent of location under protected area	1.60	9.39	0.00	100.00
Perc_wetlands	Percent of location under wetlands	1.60	6.13	0.00	97.43
Perc0_4slop	Percent of location with 0_4 slope	43.60	32.87	0.00	100.00
Perc4_8slop	Percent of location with 4_8 slope	24.08	16.18	0.00	73.76
Perc8_15slop	Percent of location with 8_15slop	17.14	14.79	0.00	59.27
Perc15_30slop	Percent of location with 15_30slope	10.79	13.26	0.00	62.13
Perc30_abovesl	Percent of location with 30_aboveslope	4.40	8.79	0.00	70.87
t_trav_munic	Travel time to municipality (minutes)	296.49	383.67	11.07	4417.17
t_trav_town	Travel time to town (minutes)	201.07	295.94	7.94	4323.98
t_trav_tcentre	Travel time to trading centre (minutes)	167.46	284.54	7.79	4323.98
t_trav_mrkt	Travel time to market (minutes)	128.81	254.57	7.53	3933.49
t_trav_road1	Travel time to road 1(tarmac) (minutes)	229.44	365.34	5.46	4308.73
t_trav_road12	Travel time to road 1 or 2 (tarmac or murram) (minutes)	175.17	297.74	5.46	4275.04
t_trav_raod123	Travel time to 1or 2 or 3 (tarmac, murram or dirt)(minutes)	116.49	251.88	3.93	3939.50
t_trav_hc	Travel time to health centre (minutes)	131.64	136.65	8.87	1302.03
Flood	Flood potential (Dummy)	0.40	0.49	0.00	1.00
Cvrain	Coefficient of variation (rainfall)	63.96	27.98	30.00	131.58
NDVI	Normalized Difference Vegetation Index	0.70	0.10	0.37	0.86
av_rainfall	Average rainfall (mm)	961.42	512.05	0.00	1987.00
lgparidsemi180	Length of growing period (LGP) 180 days	20.10	38.53	0.00	100.00
lgp60days	Length of growing period (less than 60days)	95.56	19.21	0.00	100.00
d_dist_disthosp	Distance to district hospital (meters)	24983.52	28494.55	1508.18	160466.00
d_dist_dispen	Distance to dispensary (meters)	7574.55	7987.97	1022.44	64203.69
Goodsoil	Good soil (dummy)	0.44	0.50	0.00	1.00
d_dist_10k2	Distance to nearest town of 10,000 pple (meters)	29853.15	32467.06	1407.87	234269.70
d_dist_50k2	Distance to nearest town of 50,000 pple (meters)	70740.03	108082.70	1756.27	547139.10
d_dist_100k2	Distance to nearest town of 100,000 pple (ms)	92253.60	121926.80	2366.59	638788.10
d_dist_200k2	Distance to nearest town of 200,000 pple(m)	152382.00	152727.90	4892.93	798293.00

Table A3: Ordinary Least Squares (OLS) Estimation

Dependent Variable: Poverty Rate		
Variable	Coefficient	t-statistic
Average GINI	-1.0266	(13.4131)**
Population density	-0.0001	(5.0228)**
Average Elevation (meters above sea level)	0.0000	(2.5578)*
reg2 (Central)	-0.1400	(14.0785)**
reg3 (Coast)	0.0622	(4.5167)**
reg4 (East)	0.0983	(12.1229)**
reg5 (North Eastern)	0.1817	(13.3528)**
reg6 (Nyanza)	0.1418	(15.8019)**
reg8 (Western)	0.0998	(9.4043)**
Percent of location under grass	-0.0016	(6.0789)**
Percent of location under farmland	0.0002	-1.1856
Percent of location wooded	0.0005	(3.3277)**
Percent of location that is Built-up	-0.0125	(4.3257)**
Percent of location with 4 - 8% slope	0.0013	(6.4337)**
Percent of location with 8 – 15% slope	0.0001	-0.4054
Percent of location with 15 - 30% slope	-0.0001	-0.3696
Percent of location with over 30% slope	0.0021	(5.7023)**
Percent of location with LGP less than 60 days	0.0005	(3.6689)**
Percent of location with LGP 180 days	-0.0006	(5.9913)**
Rangeland (Dummy)	0.0109	-1.6311
Good soil (dummy)	-0.0106	(2.0023)*
Average travel time to Type 1 or 2 Road (minutes)	ff0.0000	-1.6445
MEAN_distance to District Hospital	-0.0002	-1.3807
Constant	0.7901	(25.4316)**
Observations	2232	
Adjusted R-squared	0.5114	
Akaike info criterion : -3818.37	Log likelihood : 1933.19	

Absolute value of t-statistics in parentheses

* significant at 5% level; ** significant at 1% level

Table A4. Diagnostics for Spatial Dependence by Province

Statistic Province		Moran's I	Spatial error: Lagrange multiplier	Robust Lagrange multiplier	Spatial lag: Lagrange multiplier	Robust Lagrange multiplier
Central	Statistic	1.319	5.616	5.198	0.469	0.050
	p-value	0.187	0.018	0.023	0.494	0.822
Coast	Statistic	5.066	10.372	2.710	36.645	28.984
	p-value	0.000	0.001	0.100	0.000	0.000
Eastern	Statistic	8.140	27.428	0.246	47.494	20.312
	p-value	0.000	0.000	0.620	0.000	0.000
North Eastern	Statistic	1.639	0.420	0.077	0.653	0.311
	p-value	0.101	0.517	0.781	0.419	0.577
Nyanza	Statistic	10.796	92.469	7.621	95.299	10.451
	p-value	0.000	0.000	0.006	0.000	0.001
Rift Valley	Statistic	25.126	540.038	50.392	516.655	27.010
	p-value	0.000	0.000	0.000	0.000	0.000
Western	Statistic	7.255	31.056	1.814	43.660	14.418
	p-value	0.000	0.000	0.178	0.000	0.000

Table A5: Results of the Spatial Corrected Models: Central Province

Variable name	Coefficient.	Std. Err.	P>z
Demographic/Inequality			
Average gini coefficient	-0.20696	0.18931	0.27400
Population density	-0.00006	0.00002	0.01000
Distance and travel time			
Distance to forest (km)	0.00279	0.00098	0.00400
Distance to district hospital (km)	0.00390	0.00172	0.02300
Distance to nearest town of 200,000 people (meters)	-0.00005	0.00005	0.32300
Travel time to road all road types (track, tarmac or murram) (minutes)	0.00029	0.00006	0.00000
Land use			
Percent of location under bush	0.00601	0.00487	0.21800
Percent of location under wetland	-0.00537	0.00232	0.02100
Natural factors			
perc4_8slop	0.00040	0.00040	0.30700
Mean rain coefficient of variation	0.00279	0.00071	0.00000
Elevation (km above sea level)	0.06258	0.02198	0.00400
Goodsoil (dummy)	0.01996	0.01586	0.20800
Rangeland (dummy)	0.04994	0.02104	0.01800
lambda	0.10698	0.01580	0.00000
Intercept	0.00124	0.00774	0.87300
Number of observations	164		
Adjusted R-squared	0.5071		
Log likelihood	216.75		

Table A6: Results of the Spatial lag Models: Coast Province

Variable	Coefficient	Std.Error	Probability
Demographic/Inequality			
avg_gini	-1.762	0.408	0.000
Popden	0.000	0.000	0.723
Distance and travel time			
d_dist_hc	0.000	0.000	0.808
t_trav_road12	0.000	0.000	0.070
d_dist_for2	0.000	0.001	0.678
d_dist_50k2	0.000	0.000	0.926
d_dist_200k2	0.000	0.000	0.907
Land use			
perc_grass	-0.002	0.001	0.115
perc_farmland	0.001	0.001	0.460
perc_wetland	-0.003	0.001	0.020
perc_water	0.010	0.006	0.098
Natural capital			
perc4_8slop	0.002	0.001	0.009
perc8_15slop	-0.004	0.001	0.002
lgp60days	0.001	0.002	0.722
Lgparids~180	-0.001	0.001	0.006
Flood	-0.047	0.024	0.049
_cons	0.839	0.301	0.005
Rho	0.474	0.096	0.000
Number of observations	167		
Adjusted R square	0.726		
Log likelihood	143.5168		

Table A7. Results of the Spatial lag Model: Eastern Province

Variable	Coefficient	Std.Error	Probability
Demographic/Inequality			
avg_gini	-0.8609	0.2058	0.0000
Popden	-0.0001	0.0000	0.0160
Distance and travel time			
d_dist_for2	0.0012	0.0002	0.0000
d_dist_dist hosp	0.0000	0.0000	0.4050
t_trav_road12	-0.0001	0.0000	0.0000
Land use			
perc_grass	-0.0016	0.0007	0.0180
perc_farmland	0.0007	0.0004	0.0350
perc_wooded	0.0001	0.0004	0.7830
perc_wetlands	-0.0033	0.0013	0.0100
perc_protected area	0.0024	0.0006	0.0000
Natural capital			
perc4_8slop	0.0009	0.0005	0.0500
perc8_15slop	0.0003	0.0006	0.6420
perc15_30slop	0.0021	0.0009	0.0140
perc30_aboslop	0.0018	0.0012	0.1510
lgp60days	-0.0003	0.0006	0.6030
Lgparids180	-0.0004	0.0002	0.0160
Elevation	-0.0001	0.0000	0.0000
_cons	0.6912	0.1030	0.0000
Rho	0.5320	0.0871	0.0000
Number of observations	416		
Adjusted R square	0.5220		
Log likelihood	415.9590		

Table A8. OLS Model: North Eastern Province

Variable name	Coefficient.	Std. Err.	t
Demographic/Inequality			
avg_gini	-0.2384	0.0655	-3.6400
Popden	-0.0002	0.0001	-2.3000
Distance and travel time			
d_dist_disthosp	0.0000	0.0000	1.7400
d_dist_10k2	0.0002	0.0001	2.2400
d_dist_50k2	0.0000	0.0000	0.3800
d_dist_200k2			
t_trav_road12	0.0000	0.0000	-0.6000
Land use			
perc_built	0.0079	0.0052	1.5300
perc wooded	0.0003	0.0002	1.6200
Natural capital			
perc0_4slop	-0.0005	0.0005	-0.9600
perc4_8slop	-0.0021	0.0011	-1.9600
meanraincv	0.0003	0.0002	2.0900
Intercept	0.6899	0.0540	12.7800
Adj R-squared =	0.1785		
Number of obs =	202		

Table A9. Results of the Spatial lag Models: Nyanza Province

Variable	Coefficient	Std.Error	Probability
Demographic/Inequality			
avg_gini	-1.6590	0.1604	0.0000
Popden	-0.0001	0.0000	0.0280
Distance and travel time			
d_dist_for2	0.0006	0.0003	0.0430
d_dist_disthospl	0.0000	0.0000	0.0300
d_dist_200k2	-0.0001	0.0000	0.0010
Land use			
Rangelandyes	-0.0168	0.0115	0.1460
perc_water	0.0021	0.0009	0.0270
perc_grass	0.0000	0.0006	0.9580
perc_farmland	-0.0002	0.0002	0.2760
perc_wetlands	-0.0011	0.0007	0.1100
Natural factors			
perc4_8slop	0.0006	0.0003	0.0540
perc8_15slop	-0.0001	0.0004	0.8690
av_rainfall	0.0000	0.0000	0.3140
Elevation	0.0001	0.0000	0.0030
Goodsoil	-0.0168	0.0110	0.1280
_cons	0.6864	0.0818	0.0000
Rho	0.4938	0.0470	0.0000
Number of observations	305		
Adjusted R square	0.6150		
Log likelihood	357.2417		

Table A10. Results of the Spatial lag Models: Rift Valley Province

Variable	Coefficient	Std. Error	Probability
Demographic/Inequality			
avg_gini	-1.2703	0.1125	0.0000
Popden	0.0000	0.0000	0.3965
Distance and travel time			
D_dist_forest	-0.0001	0.0001	0.4393
T_trav_road12	0.0000	0.0000	0.3813
D_dist_disthsop	0.0000	0.0000	0.9310
D_dist_201	0.0000	0.0000	0.5695
land use			
perc_water	-0.0014	0.0009	0.1345
perc_built	-0.0263	0.0034	0.0000
perc_grass	-0.0004	0.0003	0.2172
perc_farmland	-0.0003	0.0002	0.1163
perc_wetland	0.0003	0.0006	0.6496
natural factors			
perc4_8slope	0.0005	0.0003	0.0758
perc8_15slope	0.0000	0.0002	0.8448
perc15_30slope	-0.0002	0.0003	0.5015
perc30_abo	0.0011	0.0003	0.0001
Flood	0.0184	0.0068	0.0066
lgp60days	-0.0002	0.0002	0.2267
Lgparidsem	-0.0009	0.0001	0.0000
Constant	0.5458	0.0463	0.0000
Rho	0.7621	0.0382	0.0000
Number of observations	785		
R-squared	0.6016		
Log likelihood	995.284		

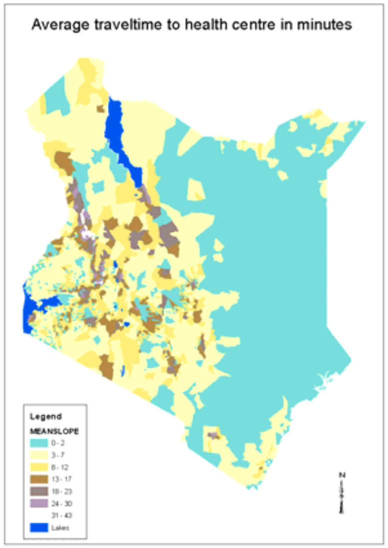
Table A11. Results of the Spatial lag Models: Western Province

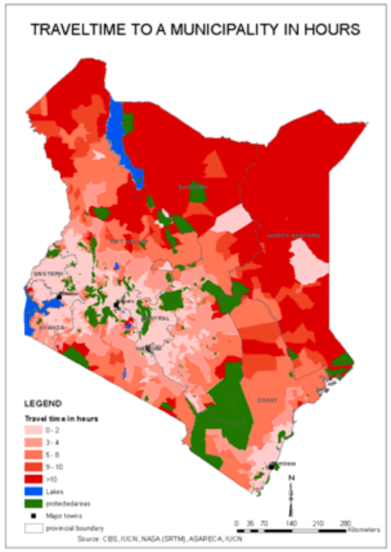
Variable	Coefficient	Std. Error	Probability
Demographic/Inequality			
avg_gini	-0.0163	0.0886	0.8540
Popden	0.0000	0.0000	0.2460
Distance and travel time			
t_trav_road12	0.0000	0.0000	0.2930
d_dist_10k	0.0000	0.0000	0.0900
d_dist_for2	0.0008	0.0004	0.0620
Land use			
perc_grassland	-0.0013	0.0006	0.0400
Rangelandyes	0.0200	0.0187	0.2850
perc_farmland	0.0002	0.0003	0.4810
perc_protected area	0.0055	0.0020	0.0070
Natural factors			
Elevation	0.0000	0.0000	0.6680
perc4_8slope	-0.0002	0.0003	0.4320
perc8_15slope	0.0004	0.0004	0.3360
perc15_30slope	-0.0018	0.0007	0.0110
Goodsoil	0.0133	0.0112	0.0440
_cons	0.1984	0.0671	0.0030
Rho	0.5860	0.0677	0.0000
Number of observations	193		
Adjusted R sq	0.6180		
Log likelihood	300.9238		

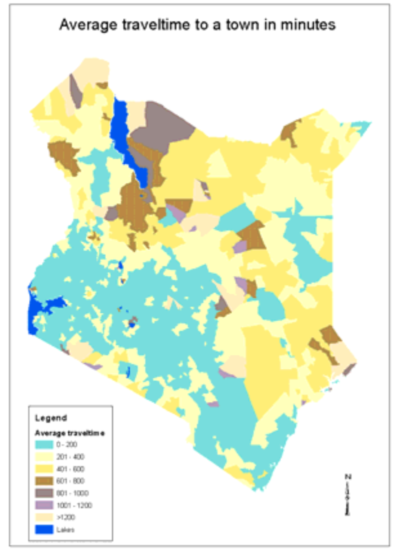
Table A12: Impact of changes in soils and travel time: An Illustrative Simulation

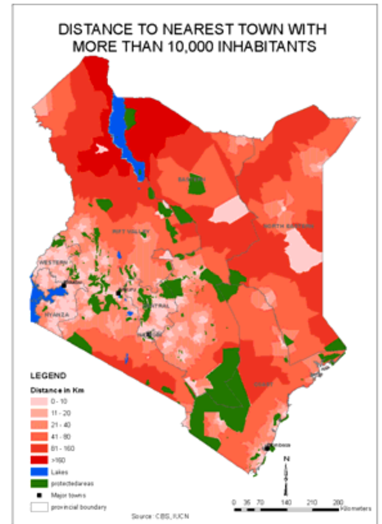
Travel time simulations			
Central	Obs	Poverty rate	Overall effect
Base poverty rate before road improvement	164	31.3	
Poverty rate after road improvement	164	30.5	Reduces poverty
Eastern			
Base poverty rate before road improvement	416	57.7	
Poverty rate after road improvement	416	56.9	Reduces poverty
Western			
Base poverty rate before road improvement	193	59.2	
Poverty rate after road improvement	193	58.9	Reduces poverty
Soil Improvement			
Western province			
Base Poverty Rate before soil improvement	193	59.2	
Poverty rate after soil improvement	193	49.8	Reduces poverty
Livestock systems			
Base Poverty Rate before soil improvement	1159	55.9	
Poverty rate after soil improvement	1159	50.4	Reduces poverty
Poverty rate after road improvement	1159	48.3	Reduces poverty

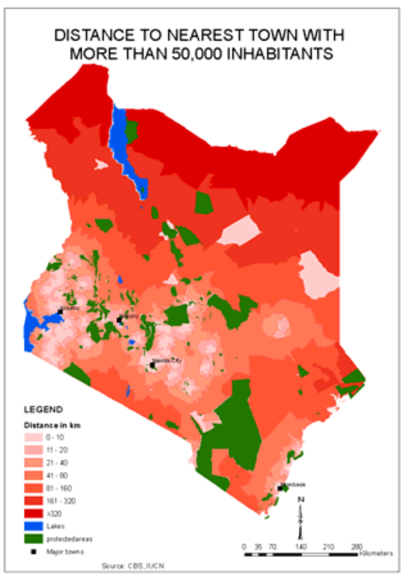
Appendix B: Maps of the spatial distribution of Variables

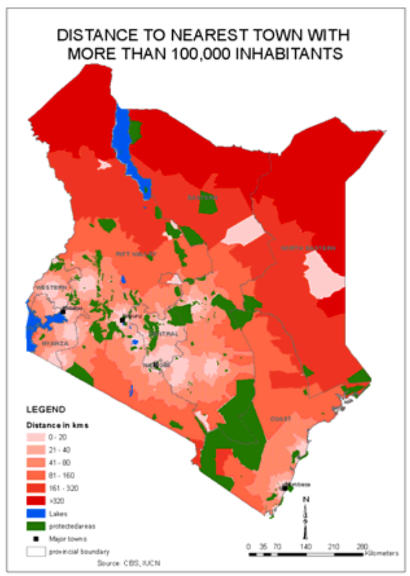


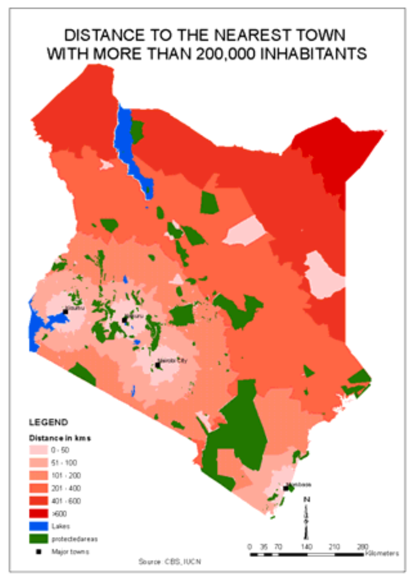


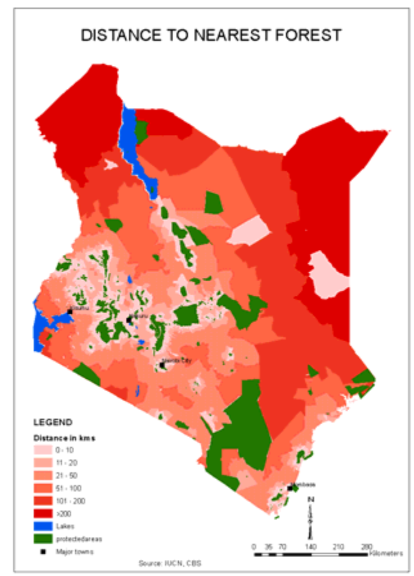




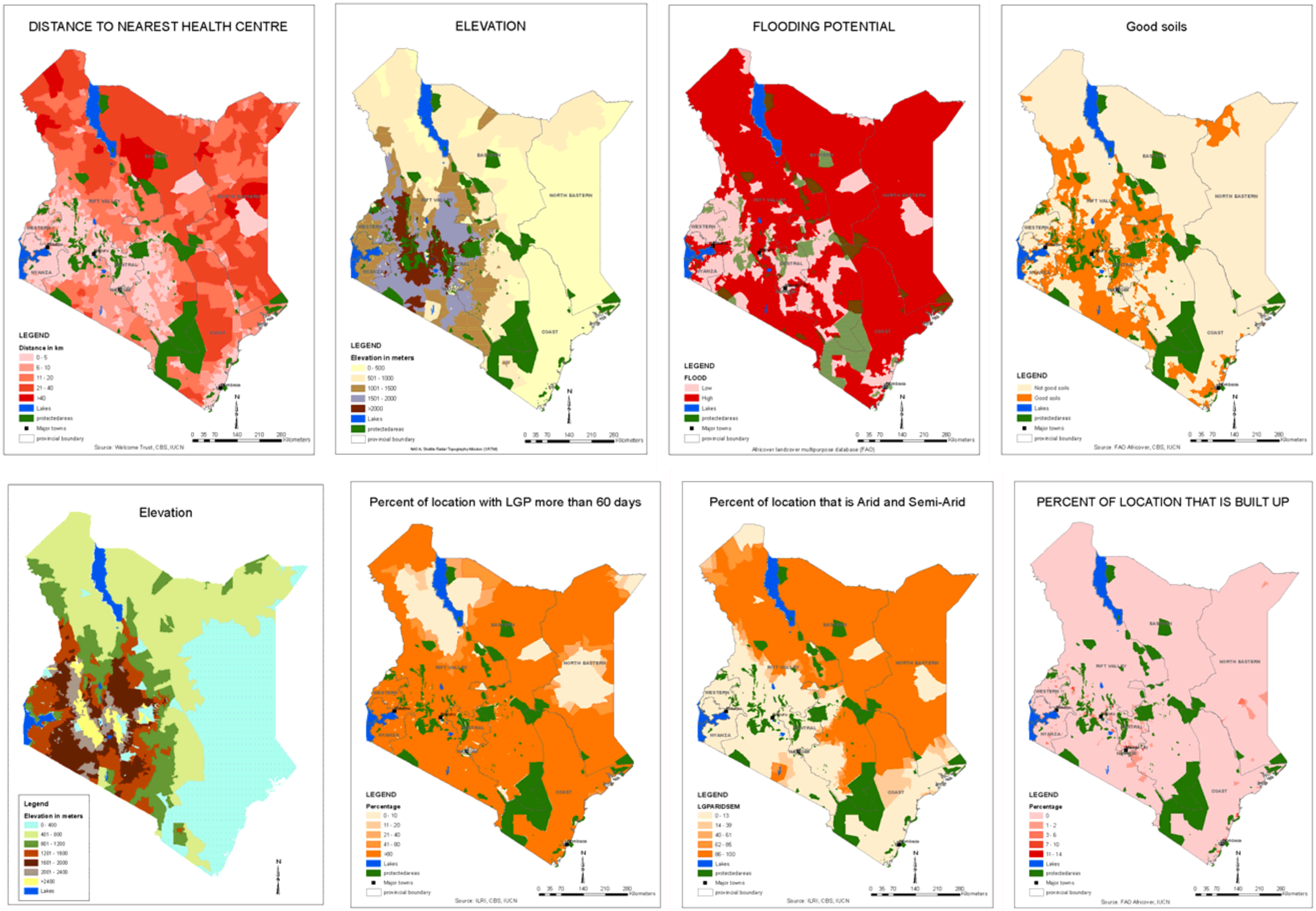




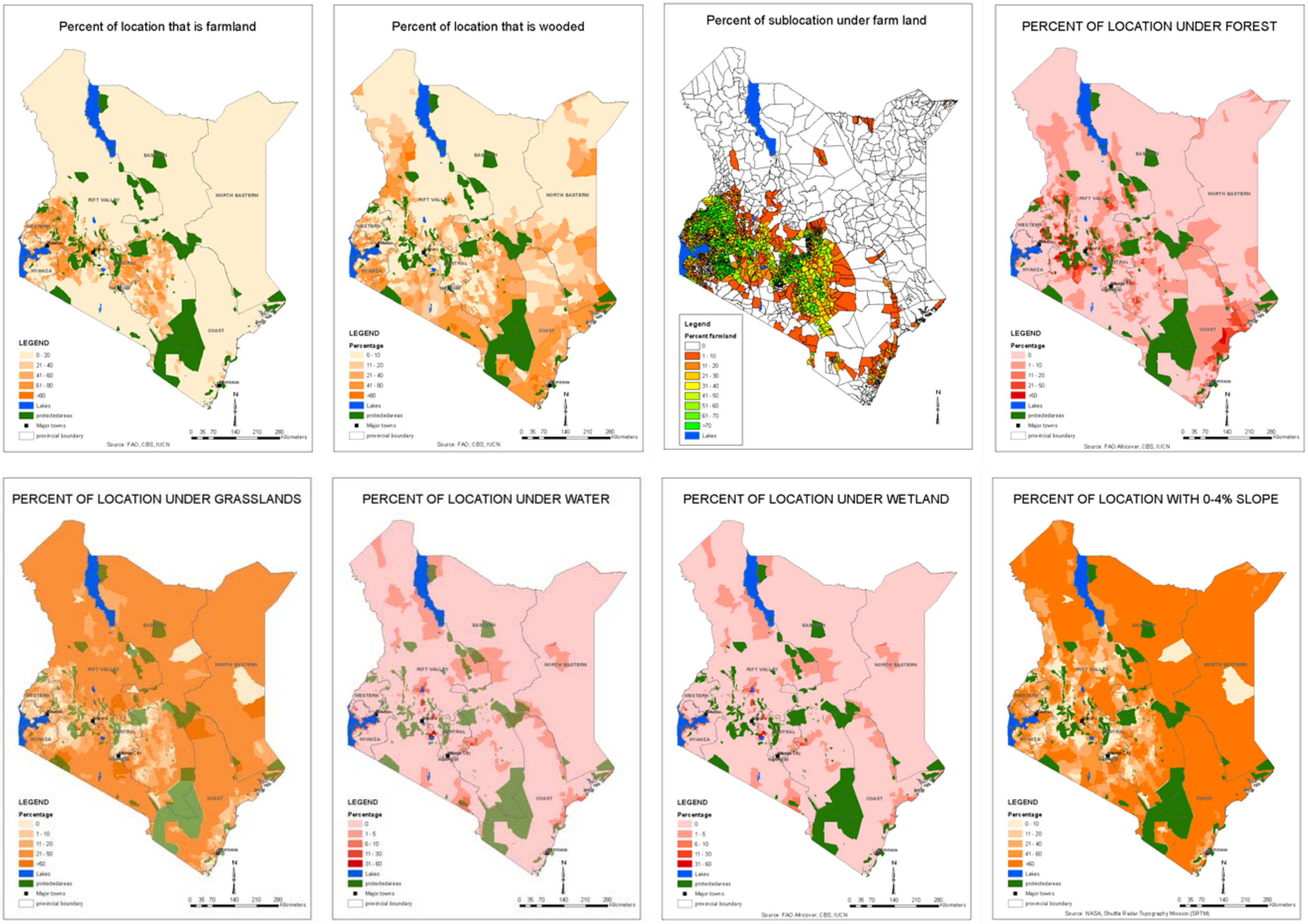




Appendix B. continued. Maps of spatial distribution of variables



Appendix B. continued. Maps of spatial distribution of variables



Appendix B continued. Maps of spatial distribution of variables