Agricultural Productivity and Anticipated Climate Change in Sub-Saharan Africa: A Spatial Sample Selection Model

Patrick S. Ward
Department of Agricultural Economics, Purdue University

Raymond J.G.M. Florax
Department of Agricultural Economics, Purdue University
Department of Spatial Economics, VU University Amsterdam

Alfonso Flores-Lagunes
Food and Resource Economics Department, University of Florida
Department of Economics, IZA, Bonn, Germany

May 3, 2010

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association 2010 AAEA, CAES, & WAEA Joint Annual Meeting, Denver, Colorado, July 25-27, 2010

Agricultural Productivity and Anticipated Climate Change in Sub-Saharan Africa: A Spatial Sample Selection Model*

Patrick S. Ward
Department of Agricultural Economics, Purdue University

Raymond J.G.M. Florax
Department of Agricultural Economics, Purdue University
Department of Spatial Economics, VU University Amsterdam

Alfonso Flores-Lagunes
Food and Resource Economics Department, University of Florida
Department of Economics, IZA, Bonn, Germany

May 3, 2010

Abstract

A cereal yield response function is estimated conditional upon environmental and topographical features to detect the effects of spatial heterogeneity and spatial dependence in explaining agricultural productivity across Sub-Saharan Africa. Controlling for direct and localized spillover effects, we then estimate the effect that projected changes in temperature and precipitation as a result of global climate change will have on agricultural production. We find that the estimated declines found in the climatological literature may overestimate actual declines, and factors such as spatial heterogeneity (i.e., country fixed effects) are profoundly more important to agricultural production.

Preliminary Draft. Please do not cite or circulate without permission of the author(s).

JEL codes: I3, Q18, C50

Keywords: Agricultural Production, Climate Change, Applied Spatial Econometrics, Sample Selection, Generalized Method of Moments Estimation

*Corresponding author: pward@purdue.edu. Helpful comments on an earlier draft were provided by Will Masters and are gratefully acknowledged. Special appreciation is extended to Jason Brown, Benoit Delbecq and participants in a Space, Health, and Population Economics workshop at Purdue University. All remaining errors are our own.

1 Introduction

Biophysical vulnerability,which is concerned with disruptions to biological and physical systems, is inherently a spatial phenomenon. Food security or sensitivity—specifically the sensitivity of food production systems—is a particularly precarious dimension of vulnerability, especially considering the impending threat of global climate change. But this biophysical vulnerability can also precipitate social vulnerability. Within the realm of global climate change shocks, which can be both idiosyncratic and covariate, it is by now widely acknowledged that the poorest of the world's citizens will find themselves particularly vulnerable, both because they often live in lower latitudes where temperatures are often already close to (or beyond) optimal, because they often live in areas prone to environmental disasters—disasters which can be enhanced as a result of climatic change—and because they lack the individual and societal wherewithal to adequately adapt to the shocks affecting their overall welfare (Diffenbaugh et al., 2007; Cline, 2007; IPCC, 2007). At present, there is no 'gold standard' definition of food security (Maxwell and Frankenberger, 1992), and thus there is no perfect measure. Even those measures that purport to gauge security according to the World Bank's definition[1] are subject to problematic measurements (De Araujo Marinho, 2008). Researchers are thus generally forced to use proxy measures that may be consequential to or symptomatic of food insecurity. Malnutrition—particularly juvenile malnutrition—is one such commonly cited measure. For example, Figure 1 shows the spatial distribution of malnutrition (percent of children under 5 suffering from malnourishment) in Sub-Saharan Africa. Only South Africa has less than 11% of childrn under 5 years of age suffering from malnutrition. In several countries, nearly 1 in 5 children under the age of 5 suffer from malnutritino. Dow and Downing (2007) report that, as of 2005, 20 of the 47 countries in Sub-Saharan Africa were already plagued with juvenile malnutrition (their measure of malnutrition is being moderately or severly underweight). The use of this proxy-and some others-seems to have its source in Malthusian philosophies, drawing upon the geometric growth of population expansion and the linear expansion in agricultural production as the main driving force behind hunger. While population expansion may be a significant threat to food security, the policy

[1] The World Bank defines food security as "access of all people at all times to enough food for an active, healthy life" (World Bank, 1986).

responses to this threat are quite scarce, and may be limited to coercive anti-reproductive policies (like those introduced China in 1979). Rather, recent food crises seem to support the hypotheses that future concerns for food security will need to pay close attention to both demand- and supply-side constraints. In this paper, we investigate and quantify the degree of biophysical vulnerability in Sub-Saharan Africa at the scale of geographic grid cells, where the biophysical vulnerability is specifically in regards to food production systems. Specifically, we examine the extent to which cereal productivity in Sub-Saharan Africa is a function of exogenous factors such as temperature, precipitation, irrigation, soil chemistry, elevation and distance to the seashore, taking into consideration both parametric spatial heterogeneity across countries and local spillovers producing spatial dependence. Because African food insecurity is expected to be exacerbated as a result of projected changes to the global environment, we aim to use simulated climate shocks to predict the extent to which yields in Sub-Saharan Africa are exposed to specific aspects of projected climate change, including increases in temperature and changes in precipitation patterns.

2 Conceptual Background

It is well documented that Sub-Saharan Africa is heavily dependent on agriculture. Indeed, next to the South Asian subcontinent, the subcontinent of Sub-Saharan Africa is the most heavily agriculture-dependent region in the world, with nearly 19% of the region's gross output being produced in the agricultural sector (World Bank, 2006). Additionally, many of the countries in Sub-Saharan Africa derive upwards of 30% of their GDP from agricultural production (See Figure 2). Despite this heavy dependence, agricultural productivity has been steadily declining in Sub-Saharan Africa over the last 50 years (See Figure 3 and Table 1). At first glance, the nexus of a heavy dependence on agriculture and a falling agricultural production would seem catastrophic. It is often quite tempting to think of food security as simply the risk of famine or mass starvation. Food production is generally the first variable that comes to mind when the word famine is pronounced (De Araujo Marinho, 2008). Many would suggest, however, that in an interconnected global community in which international trade can facilitate food shortcomings, it need not be so. However, there is much evidence to suggest that lower food production may be a precursor to lower food security, even when international food trade is allowed. Nafziger (2003) has shown that falling agricultural production may in fact have significant impacts on the availability of physical, social, and economic capital. Thus, while low food production may not be a necessary condition for a famine, it may be (under the right circumstances) a sufficient condition. "Declining rural productivity contributes not only to increased dog-eat-dog contention among severely impoverished rural populations, but also spurs rural-urban migration, increasing urban unemployment, underemployment, and political discontent, which contribute to humanitarian emergencies" (Nafziger, 2003, p. 300). For example, especially for countries heavily dependent on agriculture, the failure of food and agricultural development is a key element of overall economic stagnation, which has direct impacts on the provision of public goods and services. Such widespread economic stagnation also has a direct impact on wage income and overall employment, which limits the ability to purchase food to supplement production shortfalls (Sen, 1999). In economies heavily dependent upon agriculture, food scarcity could lead to sharp increases in food prices, which could lower the purchasing power of farm laborers, who constitute a significant proportion of the overall population. Similarly, it has been suggested that low agricultural production can result in rising social tensions, particularly when warning signs of low output lead to food panics and riots. In several African countries, agricultural stagnation has also been associated with overall slow technological and institutional modernization, unfavorable government policies and factor market distortions, and obsolete agrarian structures (Nafziger, 2003).

Particularly, cereals (wheat, rice, barley, maize, rye, oats, millet and sorghum) play an important role in the diets of people in Sub-Saharan Africa. Cereals constitute 47% of total caloric food consumption (Kcal/capita/day) for households in Sub-Saharan Africa (see Figure 4) and 50% of protein consumption (see Figure 5). Additionally, cereals provide calories more cheaply than do other sources of food. Clearly then, understanding and anticipating changes in cereal grains production are vitally important avenues for research. Recent food supply shocks caused by weather disruptions, livestock and crop disease, and export restrictions created serious crises in many of the poorest regions in the world. The challenges of addressing food security are even more pressing when we consider the impending climatic changes, if history is any indicator. Historically, temperature extremes in the peak growing seasons have been devastating to regional agricultural production, which can quickly produce chaos in international markets when policymakers intervene to restrict the limited domestic food supply from being exported.

There have been three primary strands of literature in studying the effects of climate change on agriculture: agronomic crop models, agro-ecological zone studies, and cross-sectional Ricardian models. Each of these models are unique and have their own strengths and weaknesses. The agronomic crop models utilize controlled experiments in which crops are grown in either field or laboratory settings under different climatic conditions and concentrations of atmospheric CO_2. Many of these test plots involve Free-Air Carbon dioxide Enrichment (FACE) methods, in which the plot is surrounded by a circular array of pipes that release CO_2 at a rate consistent with maintaining a specified, fixed atmospheric concentration, with sensors in the interior of the plot continually measuring the effective concentration. Based on the results of these experiments, scientists are able to estimate a yield response of specific crops to various conditions, from which they can extrapolate the projected losses (or gains) based on various climate change scenarios, generally utilizing global circulation models (GCMs). In a landmark study, Rosenzweig and Parry (1994) predicted that doubling of atmospheric carbon would have only a small negative effect on global crop production, but the effects would be more pronounced in developing countries. Recent climate simulations have predicted a similar pattern in the future. Compiling the results from various global circulation and Ricardian models, Cline (2007) predicts generally significant reductions in overall yields across Sub-Saharan Africa (See Table 7). Tebaldi and Lobell (2008) predicted rather remarkable declines in crop yields with a modest (1 degree) increase in average growing season temperature. Maize yields are predicted to decline by about $10\% - 15\%$ with a 1°C increase in growing season temperature, a temperature change lower than most published estimates (around 3°). Similarly, barley yields are predicted to decline by about 10% and wheat yields by about $5\% - 6\%$ with a 1°C increase in growing season temperature. Climate change researchers are essentially predicting a lateral shift in the distribution of mean seasonal temperatures. Battisti and Naylor (2009) used predictions from 23 GCMs and show that there is a greater than 90% probability that growing season temperatures in the tropics and subtropics will be higher than the observed extremes from 1900 through 2006.

They suggest that "the warmest summers during the past century will represent the norm by the end of this century".

The second primary strand of literature involves the analysis of changes in agro-ecological zone (AEZ) land usage. These models combine crop simulation models with land-use decision analysis and model changes in agronomic resources to assess changes in agricultural production, premised on lands shifting from one agro-ecological classification to another with changes in environmental conditions (Cline, 2007). A pioneering work in this literature is Darwin et al. (1995). Their model utilizes a multi-region, multi-sector computable general equilibrium (CGE) framework to assess changes in land use based on climatic changes, where the effects of climate change are derived from four global circulation models. These models explicitly recognize that land values may change as a result of climate change, and therefore the uses of land may change as well. A drawback of these models is that they are inherently dependent upon the underlying behavioral assumptions and the subsequent parameterization of the model. These models assume fluid prices and market closure, so may not capture structural rigidities that could realistically influence dynamic equilibria. Largely because of this assumption of fluid prices and market closure, Darwin et al. (1995) find that the adverse effect of climate change on yields will drive up food prices, ultimately resulting in increased land being devoted to food crop production, with little change in overall actual output.

The third strand of the literature involves cross-sectional reduced-form hedonic pricing models, which have come to be known as Ricardian[2] cross-section econometric models. Ricardian models use statistical methods to estimate the response of land values to climatic changes. Because land rents are assumed to reflect the value of the activity to which that land is allocated, these models are thought to embody adaptation, thus controlling for the "dumb farmer" scenario that were identified in traditional production function approaches. The first entry in this literature was Mendelsohn et al. (1994), which found that, within the context of the United States, higher temperatures generally resulted in lower land values, so they suggested that the effects of global warming on the agricultural sector might be lower than estimated. premised on the argument that land values account for the direct impact of climate on yields. The primary strength of these models is that

[2] These models are called 'Ricardian' after David Ricardo (1772-1823), who theorized that land is allocated according to whichever usage is able to pay the highest rents.

they explicitly allow for adaptation.

In a landmark study, Rosenzweig and Parry (1994) predicted that doubling of atmospheric carbon would have only a small negative effect on global crop production, but the effects would be more pronounced in developing countries. Recent climate simulations have predicted a similar pattern in the future. Compiling the results from various global circulation and Ricardian models, Cline (2007) predicts generally significant reductions in overall yields across Sub-Saharan Africa (See Table 7). Tebaldi and Lobell (2008) predicted rather remarkable declines in crop yields with a modest (1 degree) increase in average growing season temperature. Maize yields are predicted to decline by about $10\% - 15\%$ with a 1°C increase in growing season temperature, a temperature change lower than most published estimates (around 3°). Similarly, barley yields are predicted to decline by about 10% and wheat yields by about $5\% - 6\%$ with a 1°C increase in growing season temperature. Climate change researchers are essentially predicting a lateral shift in the distribution of mean seasonal temperatures. Battisti and Naylor (2009) used predictions from 23 GCMs and show that there is a greater than 90% probability that growing season temperatures in the tropics and subtropics will be higher than the observed extremes from 1900 through 2006. They suggest that "the warmest summers during the past century will represent the norm by the end of this century".

This study is somewhat of a hybrid, blending elements of the agronomic crop yield studies and the econometric approach of Mendelsohn et al. (1994). Whereas the agronomic studies use field experiments to estimate crop responses, this study uses statistical methods to isolate *ceteris paribus* effects of explanatory factors in explaining yield responses. Similar to Mendelsohn et al. (1994), we follow an econometric methodology, but with a different objective of analysis and an extension that considers the effects of spatial spillovers. This study does not explicitly use a Ricardian framework, as land values are not being explained by climatological and topographical factors, but rather we use these factors to explain yields, allowing for the isolation of spatial dependence[3].

[3] We did not use a Ricardian framework for several reasons. First, it would be difficult—if not impossible—to derive land rental rates for sub-Saharan Africa, particularly at the level of geographic grid cells. Additionally, for many parts of sub-Saharan Africa, there are lax property rights, and such laxity severely reduces the validity of any land rental rates that could be identified. It could also be argued that rental rates in Africa may not truly reflect the economic opportunities of such land (this argument is, of course, correlated with the laxity of property rights) Finally, because many farmers in Africa grow crops for subsistence rather than for market, they are not likely to

A weakness of the existing yield response studies is that, while they may take location into consideration (i.e., the characteristics of a particular location), they often don't consider the effect of local technological spillovers or high-level institutions. Thus, while they may identify a statistical relationship between agricultural productivity and environmental variables (such as temperature and precipitation), they omit other relevant explanatory variables. We attempt to correct this oversight by modeling the process of agricultural productivity as a spatial process, taking into consideration spatially contemporaneous explanatory variables as well as the effect of spatial "neighborhoods" and parametric heterogeneity over space.

The proposed methodology is certainly not without flaws of its own. This approach does not explicitly allow for adaptation, which is one of the touted advantages of the Ricardian approach. Similarly, unlike the agronomic crop yield studies, this approach quantifies only statistical correlations, without necessarily identifying causality. Additionally, because this study focuses on a cross-sectional analysis, we forgo any dynamical elements that affect crop yields. If we assume trend stationarity, then it could plausibly be argued that cross-sectional differences in yields could capture some of the dynamical or temporal differences as well. Facing impending climate change, however, it is unlikely that current cross-sectional differences will be representative of dynamic differences over the ensuing decades (e.g., see Battisti and Naylor, 2009). Nevertheless, because this study allows for spillovers and heterogeneity, it is a novel contribution to the literature on yields and may provide insights into strategies for buffering the effects of climate change on agricultural production, which in turn may have impacts on food security.

3 Data Description

The areal data used in this analysis comes from several primary sources. To examine cereal yields (the dependent variable in this analysis), we utilize $5' \times 5'$ grid cell data on global cereal yields in the year 2000 from Monfreda et al. (2008). The spatial distribution of yields in Sub-Saharan Africa can be seen in Figure 6. Data on irrigation and soil carbon density (both at a $5' \times 5'$ grid cell level), and soil pH (at $0.5^\circ \times 0.5^\circ$ resolution) come from FAO-Aquastat, the International Geosphere-

adapt to climate change by switching their planting from cereals or other food crops to higher returning cash crops.

Biosphere Program (IGBP), and the International Soil Reference and Information Centre (ISRIC), respectively. The latter two data sets were acquired through the Oak Ridge National Laboratories Distributed Active Archive Center (ORNL-DAAC). Additionally, we utilize $1^\circ \times 1^\circ$ grid cell data on various other spatial, environmental, and economic factors obtained from William Nordhaus's G-Econ dataset (Nordhaus et al., 2006). The yield, irrigation, and soil chemistry data were aggregated to $1^\circ \times 1^\circ$ in order for it to be spatially joined to the G-Econ data applying standard GIS techniques. These data consist of 2,517 observations spanning all of sub-Saharan Africa[4]. The G-Econ dataset contains observations for gross cell product–hereafter GCP, the grid cell level equivalent of gross domestic product–which is a population-weighted estimate of grid cell income based on the gross domestic product for the country and the share of the population residing in a particular grid cell[5]. Summary statistics for these data can be found in Table 3. Generally, for the purposes of examining spatial effects, grid cell data are ideal in the sense that they can easily be thought of as a regular lattice, which forms a very simple and intuitive spatial system. From this regular lattice, it becomes a very simple procedure to construct neighborhood structures and weights matrices based on simple contiguity of either the rook or queen[6]. However, one drawback of the G-Econ data is that it cannot easily be converted from points to polygons—even simply Thiessen polygons. In an effort to be thorough in presenting the data, geographic coordinates—off of which the grid cells are based—are represented more than once if the area of the cell is shared by more than one country. For example, if the border between two countries passes through a grid cell, then the geographic coordinates corresponding to that grid cell will show up under both countries, though under both entries information is provided as to what portion of the total area of the grid cell belongs to which country (i.e., a "rate in grid"—or RIG—observation for each country) and the data for some of the variables is generally different for the countries in question.[7] These non-

[4]In the econometric analysis that follows, Zimbabwe and Somalia were removed from the sample due to missing data on some key explanatory variables. Most of the exploratory spatial data analysis centered on other variables proceeded with these countries included.

[5]For a detailed description of the construction of this variable, see Nordhaus (2006). While this is no doubt an imperfect measure for the actual level of income at the grid cell level, it is to date the best approximation at this level of spatial resolution

[6]These forms of contiguity take their name from the game of chess, in which rooks can only move in the vertical and horizontal direction, whereas the queen can move vertically and horizontally, as well as along any of the diagonals.

[7]Noteworthy exceptions to this general rule include precipitation, temperature, elevation and distance data, which are generally specific at the grid cell level and would therefore be expected to be homogeneous across countries in

unique entries make converting the data into polygons impossible since points in space would be represented multiple times. Thus, neighborhood structures based on strict contiguity are infeasible. Nevertheless, because the data points are the geographical centroids of the grid cells, one can *approximate* a neighborhood structure (and thus a row-standardized weights matrix) based on contiguity if one uses a distance-based neighborhood system in which the distance specified is the minimum distance required to ensure that each observation has at least one neighbor. This is possible over the entire spatial system because the grid cells form a regular lattice structure. When some cells are not represented in the spatial system because of missing observations for the dependent variable (as will be discussed below), such a distance-based neighborhood structure no longer approximates a contiguous neighborhood system, but does serve as a reasonable method for structuring the spatial system. Under both regimes, inverse-distance weights were applied to the neighborhood structures to account for rapid distance decay. Because the geographical scale involved in this analysis (i.e., the African continent), the size of the various grid cells no doubt is subject to the curvature of the Earth. In other words, because of the spherical shape of the Earth, there is "stretching" of the cells as one approaches upper latitudes (i.e., as one approaches either of the poles). Thus, the minimum distance to ensure that every cell has at least one neighbor will be based on the largest grid cell (in this case, one in South Africa).

The purpose of this analysis is to consider how grid cell level cereal yields are affected by exogenous environmental variables (such as temperature and precipitation) after accounting for spatial dependence and heterogeneity. In section 4, we proceed with performing exploratory spatial data analysis in order to identify unconditional spatial dependence and spatial clustering. In section 5 we introduce the basic model of interest and proceed with diagnostic tests to identify the spatial process by which to control for spatial dependence and heterogeneity. In section 6 we present the results of the spatial econometric estimation, including an interpretation of the marginal effects of the various conditioning factors on yields. Finally, we conclude in section 7.

the same grid cell.

4 Exploratory Spatial Data Analysis

Before proceeding with estimating the yield response model, including specifying the existent spatial process, consider the following general cross-sectional spatial model:

$$y = \mathbf{X}\beta + \lambda \mathbf{M}_j \mathrm{y} + \varepsilon, \quad |\lambda| < 1 \tag{1}$$

$$\varepsilon = \rho \mathbf{W}_h \varepsilon + u, \quad |\rho| < 1 \tag{2}$$

where y is the dependent variable to be explained, $\mathbf{X}$ is an $n \times k$ matrix of data, $\mathbf{M}_j$ is an $n \times n$ spatial weights matrix that defines the neighborhood structure for spatially correlated dependent variables, ε is a disturbance term, u is a normally distributed random error, $\mathbf{W}_h$ is an $n \times n$ spatial weights matrix that defines the neighborhood structure for spatially correlated errors, and λ and ρ are spatial correlation coefficients corresponding to the spatially autoregressive lags and errors respectively. Several key assumptions must be made concerning the general form of this model:[8]

Assumption 1: All of the diagonal elements of $\mathbf{M}_j$ and $\mathbf{W}_h$ must be zero. This assumption ensures that the neighborhood structures are specified such that a location cannot be considered its own neighbor.

Assumption 2: The matrices $(I - \lambda \mathbf{M}_j)$ and $(I - \rho \mathbf{W}_h)$[9] are non-singular for all $\lambda \in (-1/\tau_M, 1/\tau_M)$ and $\rho \in (-1/\tau_W, 1/\tau_W)$[10], where $\tau_i = max|\nu_{1,i}|, ..., |\nu_{n,i}|$ (where $|\cdot|$ represents the absolute value of a real number and the modulus of a complex number) and $\nu_{1,i}, ..., \nu_{n,i}$ are the eigenvalues of spatial weights matrix i.[11]

Assumption 3: The row and column sums of the matrices $\mathbf{M}_j$, $\mathbf{W}_k$, $(I - \lambda \mathbf{M}_j)^{-1}$ and $(I - \rho \mathbf{W}_k)^{-1}$ are bounded uniformly in absolute value.

Assumption 4: The design matrix $\mathbf{X}$ is of full column rank, and the elements of $\mathbf{X}$ are

[8] These assumptions are borrowed from (and where necessary, modified from) Kelejian and Prucha (1998, 2007)

[9] These matrices come from the reduced forms of equations 1 and 2, respectively

[10] In time series applications, the autocorrelation coefficients are usually restricted to be on the interval $(-1, 1)$. This assumption does not carry over to spatial applications, as there may arise instances in applications (particularly when the weights matrix is not row-standardized) where $(I - \lambda \mathbf{M}_j)$ is singular for some $\lambda \in (-1, 1)$.

[11] Kelejian and Prucha (2007) note that a closely related claim is that $\frac{1}{\omega_{min}} < \rho < \frac{1}{\omega_{max}}$ and $\frac{1}{\mu_{min}} < \rho < \frac{1}{\mu_{max}}$, where ω_{min} and ω_{max} represent the minimum and maximum eigenvalues of $\mathbf{W}_j$ and μ_{min} and μ_{max} represent the minimum and maximum eigenvalues of $\mathbf{M}_k$. This claim is only valid if (1) the weights matrix is row-standardized and (2) all of the eigenvalues are real. The condition stated here is general enough to be valid in all applications.

uniformly bounded in absolute value.

Assumption 5: The innovations u are independently and identically distributed with $E(u) = 0$, $E(u^2) = \sigma_u^2$, with $0 < \sigma_u^2 < \infty$.

These assumptions are critical for ensuring well-behaved spatial processes. For the analysis that follows, we assume that the underlying spatial process is a spatial error process, (i.e., one with spatially autoregressive errors). In the context of the above assumptions, therefore, we assume $\lambda = 0$. There are both theoretical and statistical grounds for making this assumption. Statistically, we used Lagrange multiplier tests on the OLS residuals (Anselin et al., 1996) to test for and specify the underlying spatial processes[12]. The robust LM tests indicate the presence of both spatially autoregressive lags (i.e., lagged dependent variables) and errors. Despite this relatively conclusive statistical evidence, we appeal to intuition and precedence in assuming that the only spatial process is in the error terms. Intuition would suggest there are no plausible reasons for yields to be spatially correlated except for correlation among unobserved factors (e.g., technological or knowledge spillovers), which would be captured more appropriately through a spatial error process. Additionally, the robust lag test developed in Anselin et al. (1996) and performed above has the highest power of all of the Lagrange multiplier tests included in their battery of specification tests, and thus the significance of this statistic may be due to an extremely significant spatial error coefficient. There is also precedence in modeling yield response functions as spatial error processes

Exploratory data analysis proceeded with detecting unconditional spatial dependence in cell-level cereal yields using Moran's I statistic, with random permutations used to verify the sensitivity of these results. The tests indicate an unconditional coefficient of spatial correlation of 0.7063, significant at the $\alpha = 0.001$ level based on 1,000 Monte Carlo replications. Thus, there is significant evidence of unconditional global spatial autocorrelation in agricultural productivity (See Figure 7). Clearly, there is significant spatial autocorrelation in yields across Sub-Saharan Africa. We suggest, however, that because the Moran's I statistic measures unconditional spatial autocorrelation, much of this is likely due to the spatial scale in question and the strong spatial similarities in

[12]While it is acknowledged that the necessary assumptions for OLS to produce unbiased and efficient estimates are violated (see below), for lack of better alternatives we proceed with the spatial model specification based on the the asymptotic properties of OLS

variables explaining yields, such as temperature and precipitation. If these truly exogenous factors are able to explain a great deal of the variations in yields, then it is likely that strong spatial autocorrelation in these explanatory factors will result in strong spatial autocorrelation in cereal yields, even when there is no plausible reason to suggest that the spatial autocorrelation has any causal interpretation. In addition to exploring the spatial correlation of yields, it is instructive to consider the possibility that spatial dependence could arise from unobservable disturbances. Thus, we again utilized Moran's I statistic to test for unconditional spatial dependence in the residuals of an OLS estimation of a simple non-spatial linear model (see the econometric model in equation 3 below). This test indicated unconditional spatial autocorrelation in the residuals of 0.2881, again significant at the $\alpha = 0.001$ level based on 1,000 Monte Carlo replications. So in addition to there being unconditional autocorrelation in the dependent variable (cereal yields), there is also significant autocorrelation in the OLS residuals (see Figure 8). Whether these significant spatial dependencies remain consistent after considering conditioning factors will be discussed in section 5.

To abstract from the global spatial system and focus on lower-level spatial autocorrelation, localized indicators of spatial autocorrelation (LISA) were used to identify patterns of localized spatial dependence. The results of the LISA analysis can be seen in Figure 9, which clearly indicates several clusters of highly productive areas neighboring other highly productive areas. Some of the most notable clustering occurs in southern Africa (including South Africa, Lesotho, and Swaziland), the island nation of Madagascar, near the Great Lakes region (including parts of Burundi, Kenya, Rwanda, Tanzania, and Uganda), in parts of Ethiopia, and in the tropical rainforests of western Africa, most notably in Ghana and Côte d'Ivoire. The LISA analysis also reveals several clusters of low productivity. This clustering is most pronounced in Sahelian regions of northwestern Africa, parts of the Republic of Congo (Brazzaville) and the Democratic Republic of Congo (Kinshasa), and semi-arid regions of Namibia and Botswana. It would be tempting to link these areas of low production with low precipitation, but the evidence is firmly against this broad generalization. In fact, rainforest regions of the Republic of Congo and the Democratic Republic of Congo are among the wettest areas in all of sub-Saharan Africa (see Figure 10), yet suffer from low agricultural productivity.

It is also of interest to consider spatial heterogeneity. In essence, spatial heterogeneity is an instability in the spatial system (i.e., relations between pairs of points in space is non-constant over the system). Anselin (1988) identified two primary sources of spatial heterogeneity: heteroskedastic errors and spatially varying parameters. Spatial process models, by their nature, induce heteroskedasticity. Incorporating a neighborhood structure in a spatial process model allows for the incorporation of spatial dependence, but often times, heteroskedasticity is due to spatial heterogeneity. Incorporating only spatial dependence implies parametric homogeneity over the spatial system, or at least it does not model spatial heterogeneity explicitly. While it is true that the spatial models artificially introduces heterogeneity, if the coefficients are likely to vary across space, it is generally preferable to explicitly incorporate heterogeneity in the model. Thus, it was also of interest to consider and control for discrete spatial heterogeneity, whether in the form of spatial shift operators or spatial regimes. Since the data—once aggregated—were at $1^\circ \times 1^\circ$ resolution, it seemed logical to identify spatial shift operators based on the country to which each grid cell belonged, since it is assumed there are country-level fixed effects which influence agricultural production, but which cannot be measured explicitly. Such factors could include (but are not limited to) governmental subsidies for fertilizers, extension programs to assist farmers' decision-making, a sound legal system which provides stringent property rights, a system of functioning input and output markets, etc. Since each of these factors would plausibly be the same for all grid cells within a country, but would vary across countries, even potentially across adjacent grid cells, it makes natural sense to include country-specific fixed effects in the model to account for this high-level heterogeneity. Such effects are captured through the use of country dummy variables. It is also relevant to allow for parametric heterogeneity through the identification of spatial regimes. In a critique of earlier Ricardian models, Schlenker et al. (2005) suggests that the effects of climate change on agriculture must be assessed differently in dryland and irrigated areas. Failure to account for irrigation, they argue, understates the water supply in irrigated regions. Thus, they suggest that irrigation be included in the set of explanatory variables in any such regression, but not simply as a linear term to shift the constant. Rather, they suggest that the effects of climatological variables should be different in irrigated areas than in rain-fed areas. Thus, interactions between irrigation and other explanatory variables allow for capturing an even greater degree of heterogeneity.

5 Model Specification

In specifying the model to be estimated, we initially specify a very general form, but follow the sequential (specific-to-general) spatial model specification procedure outlined in Anselin (2005), in which we sequentially impose a set of linear restrictions. We begin by considering the linear model:

$$\begin{aligned} y_i &= \mathbf{x}_i'\beta + \mathbf{h}_i'\pi + (\text{Irr}_i \cdot \tilde{\mathbf{x}}_i)'\,\xi + \varepsilon_i \\ \varepsilon_i &= \rho \sum_j \mathrm{w}_{ij}\varepsilon_j + u_i \end{aligned} \tag{3}$$

where y_i is cereal yield per grid cell (in tons per hectare), $\mathbf{x}_i'$ is a vector of explanatory variables for grid cell i containing observations on: average temperature (°C) and its square, the standard deviation of temperature, average precipitation (mm per month) and its square, the standard deviation of precipitation, average elevation (km), the roughness of the terrain (a measure of variation in elevation), the distance to the shore (m), the percentage of the cell that has irrigation, the average pH level for the soil in the cell, and the average carbon content in the cell's soil. The vector $\mathbf{h}_i'$ is a vector of country dummy variables used to control for country-specific fixed effects,[13] $\tilde{\mathbf{x}}_i \subset \mathbf{x}_i$ is a vector of explanatory variables excluding the cell-level irrigation proportion, ε_i is the disturbance term, w_{ij} is the (i, j) element of a spatial weights matrix,[14] and u_j is a disturbance term assumed to have a spherical, normal distribution. The vector β is a k-vector of parameters to be estimated, π is a vector of parameters to be estimated capturing country-specific effects that effect yields (i.e., intercept shifters), ξ is a $(k-2)$-vector of parameters to be estimated, and ρ is a scalar spatial correlation coefficient, also to be estimated. In the specific-to-general approach, we begin by assuming the absence of spatial autocorrelation (i.e., $\rho = 0$). In the absence of spatial correlation, the model

[13] To maintain a full-rank design matrix, responses for Angolan grid cells were omitted from the matrix $\mathbf{H}$.

[14] The weights matrix satisfies the requisite assumptions above.

reduces to:

$$y_i = \mathbf{x}_i\beta + \mathbf{h}_i'\pi + (\mathrm{Irr}_i \cdot \tilde{\mathbf{x}}_i)'\,\xi + u_i$$

which, under the assumptions of spherically distributed error terms, could be estimated consistently using OLS. Complications arise when testing the assumptions placed on the disturbance terms. Testing this model for heteroskedasticity using the studentized Breusch-Pagan (a.k.a. Koenker-Bassett) test results in test statistic of 205.21, indicating a clear rejection of homoskedastic errors. As such, estimation of equation (3) is inefficient; that is, it results in biased estimates for the variances, which invalidates any inferences that can be made based on simple OLS estimation of this model. A noted weakness of the Breusch-Pagan test is that, while it is very general and thus widely applicable, it makes no assumption about the source of the heteroskedasticity. The general approach taken in the case of unknown heteroskedasticity is to employ the White estimator for the variance of the least squares estimates. The heteroskedasticity-robust standard errors are shown beside the non-spatial OLS estimates in Table 4. In addition to heteroskedastic disturbances, the results of the Jarque-Bera test allow us to reject the normality the OLS residuals (with the distribution of residuals suffering from both positive skewness and positive kurtosis), thus violating one of the fundamental assumptions of the Gauss-Markov theorem.[15] As such, inferences based on normally distributed errors (e.g., t-tests, F-tests) would invalid.

Yields can take only non-negative values (that is, they are censored at 0), and, as such, the "true" yield y_i^* can be modeled as a latent variable. Because yields are only observed for a subset of the total grid cells in Sub-Saharan Africa, we potentially face a non-random sample. If the grid cells for which yields are observed have the same characteristics of the rest of the grid cells for which yields are not observed, then the subsample is truly random, and it is by mere coincidence that cereals were planted in the cells for which yields are observed and not planted in cells for which yields are not observed. If this were truly the underlying process, then we could proceed estimating the non-spatial yield response function using simple OLS or—to consider the censoring

[15] Rejection of the normality of the OLS residuals may result in the invalidation of the Lagrange Multiplier specification tests performed and outlined below, since they are based on the asymptotic properties of OLS, and thus assume spherical and normally distributed disturbances.

of the data—we could proceed using a censored regression (i.e., Tobit) estimator. If, however, the cereal planting is due to an endogenous decision (or selection) mechanism, then the cells for which we have observed yields no longer constitute a random sample, and the estimation suffers from sample selection bias. For the 382 grid cells without observed yields, it is presumed that the lack of yield is the result of no cereals being planted, not because conditioning factors were such that, though cereals were planted, the resultant yields were zero. Thus, if the probability of choosing to plant is conditional on factors outside those otherwise affecting yields (that is, factors not included in equation (3)), then the factors conditioning the choice to plant cereals are correlated with ε, and estimation of equation (3) is biased[16]. Consider the selection variable P^*, such that $P_i^* > 0$ if cereals were planted in grid cell i and $P_i^* = 0$ otherwise. The actual selection variable P_i^* is never explicitly observed, but rather a binary selection P_i is observed. The relationships between y_i^* and y_i and between P_i^* and P_i are as follows:

$$\begin{aligned} y_i &= y_i^* \quad \text{if } P_i^* > 0 \quad y_i = 0 \text{ otherwise} \\ P_i &= 1 \quad \text{if } P_i^* > 0 \quad P_i = 0 \text{ otherwise} \end{aligned}$$

If we allow for P_i^* to be conditioned by a set of exogenous variables, $\mathbf{z}_i$, then we can reformulate the selection model as:

$$P_i^* = \mathbf{z}_i'\alpha + \mu_i$$

with α being a j-vector of parameters. Under the assumption that $\mu \sim N(0, \sigma^2)$, the selection model is such that $\text{Prob}(P_i = 1|\mathbf{z}_i) = \Phi(\mathbf{z}_i'\alpha)$ and $\text{Prob}(P_i = 0|\mathbf{z}_i) = 1 - \Phi(\mathbf{z}_i'\alpha)$, where $\Phi(\cdot)$ is the normal cumulative distribution function. The regression model for the response is then:

$$y_i = \mathbf{x}_i\beta + \mathbf{h}_i'\pi + (\text{Irr}_i \cdot \tilde{\mathbf{x}}_i)'\,\xi + u_i$$

[16] For a detailed description of the properties of sample selection models and methods of correcting for sample selection bias, see standard textbook treatments in Greene (2003) or Davidson and MacKinnon (1993).

where y_i is observed only if $P_i = 1$. Thus:

$$E[y_i|P_i = 1, \mathbf{x}_i, \tilde{\mathbf{x}}_i, \mathbf{z}_i] = \mathbf{x}_i\beta + \mathbf{h}_i'\pi + (\text{Irr}_i \cdot \tilde{\mathbf{x}}_i)'\,\xi + \eta\sigma_u \left[\frac{\phi(\mathbf{z}_i'\alpha)}{\Phi(\mathbf{z}_i'\alpha)}\right] \tag{4}$$

where $\phi(\cdot)$ is the probability density function associated with the normal distribution and $\phi(\mathbf{z}_i'\alpha)/\Phi(\mathbf{z}_i'\alpha)$ is commonly referred to as the Inverse Mills Ratio (IMR). A two-step estimation procedure proposed by Heckman (1976, 1979) is widely used to estimate models of this nature. In the first stage of this procedure, the selection equation is estimated using a probit estimator and estimates of the IMR is constructed. In the second stage, the outcome (or response) equation is estimated via OLS, with the IMR included as an additional explanatory variable. Despite its popularity and wide usage, Davidson and MacKinnon (1993) point out that the consistency of this estimator depends crucially on the assumption of normally distributed errors. In specifying such a model, Wooldridge (2006) suggests that it is advantageous to impose the requirement that $\left[\mathbf{X}, \mathbf{H}, \tilde{\mathbf{X}}\right]' \subset \mathbf{Z}$ since (1) any variable relevant in explaining yields in equation (3) *should* also be considered as an explanatory variable in the selection equation (though the converse is not necessarily true) and (2) the requirement that $\left[\mathbf{X}, \mathbf{H}, \tilde{\mathbf{X}}\right]' \subset \mathbf{Z}$ automatically provides an exclusion restriction by which the model can be identified. In specifying the non-spatial model to be tested here, we allow $\mathbf{X} \subset \mathbf{Z}$, excluding the country dummy variables and the interaction terms from the selection equation. In addition to the variables in $\mathbf{X}$, the selection equation design matrix $\mathbf{Z}$ also includes observations on gross cell product and grid cell population, factors which plausibly affect the decision to plant but which should not have any power in explaining observed yields.[17]

Estimates of the model parameters computed via both OLS and Heckman's two-step estimation procedure are presented in Tables 4 and 5.

We now turn to modeling the cereal yield response function as a spatial process (i.e., we allow $\rho \neq 0$). For the time being, we abstract from the sample selection problem and consider the censoring of yields to be the result of an exogenous data generating process, and thus estimation

[17] Population is likely to affect the decision to plant cereals in two ways: If the population is too high (e.g., an urban area), it may be infeasible to plant cereals. Additionally, if the population is too low (for whatever reason), there may not be any cereals planted. To control for the first effect (i.e., urban areas), gross cell product is used as a conditioning variable in the selection equation. Grid cell population would therefore be considered to be the principal conditioning factor in sparsely populated grid cells.

of the parameters of the observed sample are consistent. Recall the general form of the model:

$$y_i = \mathbf{x}_i'\beta + \mathbf{h}_i'\pi + (\mathrm{Irr}_i \cdot \tilde{\mathbf{x}}_i)'\,\xi + \varepsilon_i$$
$$\varepsilon_i = \rho \sum_j \mathrm{w}_{ij}\varepsilon_j + u_i$$

which can be written in the reduced form

$$y_i = \mathbf{x}_i'\beta + \mathbf{h}_i'\pi + (\mathrm{Irr}_i \cdot \tilde{\mathbf{x}}_i)'\,\xi + \sum_j \left[(\mathbf{I}-\rho\mathbf{W})^{-1}\right]_{ij} u_j \tag{5}$$

where $\left[(\mathbf{I}-\rho\mathbf{W})^{-1}\right]_{ij}$ is the (i, j) element of the spatial multiplier matrix $(\mathbf{I} - \rho\mathbf{W})^{-1}$. This reduced form can be re-written in matrix notation

$$y = \mathbf{X}\beta + \mathbf{H}\pi + \left(\mathrm{Irr}' \cdot \tilde{\mathbf{X}}\right)\xi + (I - \rho\mathbf{W})^{-1}\,u \tag{6}$$

The parameter ρ is the coefficient of spatial correlation and the term $(I - \rho\mathbf{W})^{-1}$ represents the spatial multiplier, as so identified by Anselin (2003). Most econometric models involving spatial processes can be estimated using maximum likelihood approaches, but such approaches generally require the (perhaps strong) assumption of normally distributed disturbances, and the calculation of a likelihood function over a spatial system can be quite computationally demanding, particularly since they require the computation of the Jacobian of the $n \times n$ spatial weights matrix. Estimating equation (6) using a maximum likelihood approach was abandoned, primarily because diagnostic tests indicate the absence of normally and spherically distributed error terms, thus invalidating maximum likelihood estimates based on the assumption of normality.

Kelejian and Prucha (1999) developed an estimator for ρ based on the generalized method of moments (GMM). This approach has two advantages over the maximum likelihood approach (Bell and Bockstael, 2000). First, the GMM estimator is consistent regardless of whether the errors are normally distributed. Additionally, the calculation of the estimator is relatively straightforward, and the computational intensity is substantially lower compared with other spatial estimators, including maximum likelihood estimators. This approach requires neither the computation of the determinant of the spatial weights matrix, nor its eigenvalues. One negative of the GMM estimator, however, is that it does not allow the calculation of standard errors for the ρ parameter. This is not a significant concern, however, as the ρ term is generally viewed as a nuisance parameter. Let ε be the disturbance term of a spatial error process, as defined in equation (3). Then let $\tilde{\varepsilon}$ be a predictor of ε. Then let $\bar{\varepsilon} = \mathbf{W}\varepsilon$, $\bar{\bar{\varepsilon}} = \mathbf{W}\mathbf{W}\varepsilon$, and correspondingly $\tilde{\bar{\varepsilon}} = \mathbf{W}\tilde{\varepsilon}$ and $\tilde{\bar{\bar{\varepsilon}}} = \mathbf{W}\mathbf{W}\tilde{\varepsilon}$. Additionally, let $\bar{u} = \mathbf{W}u$. Then the three moment conditions are therefore:

$$E\left[\frac{1}{n}u'u\right] = \sigma^2 \qquad E\left[\frac{1}{n}\bar{u}'\bar{u}\right] = \sigma^2 n^{-1} Tr(\mathbf{W}'\mathbf{W}) \qquad E\left[\frac{1}{n}\bar{u}'u\right] = 0 \tag{7}$$

These moment conditions can be written in terms of ε:

$$\begin{aligned} E\left[\tfrac{1}{n}\varepsilon'(\mathbf{I}-\rho\mathbf{W})'(\mathbf{I}-\rho\mathbf{W})\varepsilon\right] &= \sigma^2 \\ E\left[\tfrac{1}{n}\varepsilon'(\mathbf{I}-\rho\mathbf{W})'\mathbf{W}'\mathbf{W}\mathbf{W}(\mathbf{I}-\rho\mathbf{W})\varepsilon\right] &= \tfrac{\sigma^2}{n}Tr(\mathbf{W}'\mathbf{W}) \\ E\left[\tfrac{1}{n}\varepsilon'(\mathbf{I}-\rho\mathbf{W})'\mathbf{W}'(\mathbf{I}-\rho\mathbf{W})\varepsilon\right] &= 0 \end{aligned} \tag{8}$$

Multiplying through, rearranging, and using $\bar{\varepsilon} = W\varepsilon$ and $\bar{\bar{\varepsilon}} = WW\varepsilon$, we have the following system of moment equations:

$$\mathbf{\Gamma}_n \left[\rho, \rho^2, \sigma^2\right]' - \gamma_n = 0 \tag{9}$$

where

$$\mathbf{\Gamma}_n = \begin{bmatrix} \frac{2}{n}E(\varepsilon'\bar{\varepsilon}) & \frac{-1}{n}E(\bar{\varepsilon}'\bar{\varepsilon}) & 1 \\ \frac{2}{n}E(\bar{\bar{\varepsilon}}'\bar{\varepsilon}) & \frac{-1}{n}E(\bar{\bar{\varepsilon}}'\bar{\bar{\varepsilon}}) & \frac{1}{n}Tr(\mathbf{W}'\mathbf{W}) \\ \frac{1}{n}E(\varepsilon'\bar{\bar{\varepsilon}} + \bar{\varepsilon}'\bar{\varepsilon}) & \frac{-1}{n}E(\bar{\varepsilon}'\bar{\bar{\varepsilon}}) & 0 \end{bmatrix} \quad \gamma_n = \begin{bmatrix} \frac{1}{n}E(\varepsilon'\varepsilon) \\ \frac{1}{n}E(\bar{\varepsilon}'\bar{\varepsilon}) \\ \frac{1}{n}E(\varepsilon'\bar{\varepsilon}) \end{bmatrix} \tag{10}$$

Using the OLS residuals as predictors of ε, we have the sample moment analogue to (9):

$$\mathbf{G}_n \left[\rho, \rho^2, \sigma^2\right]' - g_n = \nu_n\left(\rho, \sigma^2\right) \tag{11}$$

and the corresponding sample matrices

$$\mathbf{G}_n = \begin{bmatrix} \frac{2}{n}E(\tilde{\varepsilon}'\tilde{\tilde{\varepsilon}}) & \frac{-1}{n}E(\tilde{\tilde{\varepsilon}}'\tilde{\tilde{\varepsilon}}) & 1 \\ \frac{2}{n}E(\tilde{\tilde{\tilde{\varepsilon}}}'\tilde{\tilde{\varepsilon}}) & \frac{-1}{n}E(\tilde{\tilde{\tilde{\varepsilon}}}'\tilde{\tilde{\tilde{\varepsilon}}}) & \frac{1}{n}Tr(\mathbf{W}'\mathbf{W}) \\ \frac{1}{n}E(\tilde{\varepsilon}'\tilde{\tilde{\tilde{\varepsilon}}} + \tilde{\tilde{\varepsilon}}'\tilde{\tilde{\varepsilon}}) & \frac{-1}{n}E(\tilde{\tilde{\varepsilon}}'\tilde{\tilde{\tilde{\varepsilon}}}) & 0 \end{bmatrix} \quad g_n = \begin{bmatrix} \frac{1}{n}E(\tilde{\varepsilon}'\tilde{\varepsilon}) \\ \frac{1}{n}E(\tilde{\tilde{\varepsilon}}'\tilde{\tilde{\varepsilon}}) \\ \frac{1}{n}E(\tilde{\varepsilon}'\tilde{\tilde{\varepsilon}}) \end{bmatrix} \tag{12}$$

The GMM estimator for ρ and σ^2 is the restricted nonlinear least squares estimator

$$\left(\hat{\rho}, \hat{\sigma}^2\right) = \text{argmin} \left\{ \nu_n \left(\rho, \sigma^2\right)' \nu_n \left(\rho, \sigma^2\right) \right\} \tag{13}$$

such that ρ is bounded according to Assumption 2 above (to ensure the invertability of the spatial multiplier matrix) and σ^2 is a strictly positive real number. Once estimates of ρ and σ^2 have been obtained, estimates for the model parameters can be obtained by feasible generalized least squares (FGLS). For data matrix $\mathbf{Q} = \left[\mathbf{X}, \mathbf{H}, \left(\text{Irr}' \cdot \tilde{\mathbf{X}}\right)\right]'$ and vector of parameters $\Theta = [\beta, \pi, \xi]$, Kelejian and Prucha (1999) give the FGLS estimator as:

$$\Theta^{FGLS} = \left[\mathbf{Q}'\boldsymbol{\Omega}(\hat{\rho})^{-1}\mathbf{Q}\right]^{-1} \mathbf{Q}'\boldsymbol{\Omega}(\hat{\rho})^{-1} y$$

where $\boldsymbol{\Omega}(\hat{\rho})$ is the estimated variance-covariance matrix of the disturbance vector ε, $\boldsymbol{\Omega}(\hat{\rho}) = \hat{\sigma}^2 \left(\mathbf{I} - \hat{\rho}\mathbf{W}\right)^{-1} \left(\mathbf{I} - \hat{\rho}\mathbf{W}'\right)^{-1}$. Substituting in for $\boldsymbol{\Omega}(\hat{\rho})$, the FGLS estimator reduces to

$$\Theta^{FGLS} = \left(\dot{\mathbf{Q}}'\dot{\mathbf{Q}}\right)^{-1} \dot{\mathbf{Q}}'\dot{y}$$

where $\dot{\mathbf{Q}} = (\mathbf{I} - \hat{\rho}\mathbf{W})^{-1}\mathbf{Q}$ and $\dot{y} = (\mathbf{I} - \hat{\rho}\mathbf{W})^{-1}y$, which is simply the OLS estimator of a model transformed by a Cochrane-Orcutt transformation. Estimates of the parameters obtained via this procedure are shown in Table 6.

While this model *does* control for spatial dependence and heterogeneity, it also assumes that the sample is random and that the process for censoring yields at zero is exogenous. To remedy this oversight, we employ an estimator for spatial error models that controls for endogenous sample selection, recently developed by Flores-Lagunes and Schnier (2010). The original attempt to model sample selection in the context of spatial precess models was McMillen (1995), though the estimator he proposes is infeasible since it requires *a priori* knowledge of the model parameters and is computationally intensive. The estimator that Flores-Lagunes and Schnier (2010) proposes has the intuition of Heckman's model, but is within the broader family of GMM estimators. Specifically, it uses a selection equation analogous to the spatial probit estimator of Pinkse and Slade (1998). The estimates from the spatial probit are then used to construct the IMR, which is then included in the outcome equation. The Pinkse and Slade estimator yields consistent estimates of the selection equation, which are themselves necessary to obtain estimates of the parameters in the response equation. Additionally, because it is within the class of GMM estimators, it is computationally simpler than maximum likelihood estimators and does not require the strong distributional assumptions. Flores-Lagunes and Schnier note that when the parameters in the selection equation are different from those in the outcome equation, the appropriate IMR is a function of the spatial correlation coefficient in the outcome equation. Thus, to increase the efficiency of the estimates, all of the model parameters are estimated simultaneously through solving a system of stacked moment equations. While this estimator is less efficient than a maximum likelihood estimator, it is consistent and asymptotically normally distributed with an estimable variance-covariance matrix.

Considering the problem of sample selection within the context of our spatial model, we allow yields to once again be represented as a latent variable y^*. We therefore have the familiar cases:

$$\begin{aligned} y_i = y_i^* \quad & \text{if } P_i^* > 0 \quad \psi_i = 0 \text{ otherwise} \\ P_i = 1 \quad & \text{if } P_i^* > 0 \quad P_i = 0 \text{ otherwise} \end{aligned}$$

Explicitly modeling the selection and response equations taking into consideration spatial dependence in the errors, we have:

$$P_i^* = \mathbf{z}_i'\alpha + \varepsilon_{1i}, \qquad \varepsilon_{1i} = \rho_1 \sum_{j \neq i} \mathrm{w}_{ij}^1 \varepsilon_{1j} + v_i \tag{14}$$

$$y^* = \mathbf{x}_i'\beta + \mathbf{h}_i'\pi + (\mathrm{Irr}_i \cdot \tilde{\mathbf{x}}_i)'\,\xi + \varepsilon_{2i}, \qquad \varepsilon_{2i} = \rho_2 \sum_{j \neq i} \mathrm{w}_{ij}^2 \varepsilon_{1j} + u_i \tag{15}$$

where w_{ij}^1 is the (i,j) element of the $(n+b)\times(n+b)$ spatial weights matrix corresponding to the selection equation, and w_{ij}^2 is the (i,j) element of the $n\times n$ spatial weights matrix corresponding to the outcome equation. Note that both equation (14) and equation (15) exhibit (or are general enough to allow) spatial dependence in the errors (denoted by coefficients ρ_1 and ρ_2, respectively). The innovations v and u are assumed *iid* and multivariate normal such that $(v_i,u_i)\sim N(0,\Sigma)$, where

$$\Sigma=\begin{bmatrix}\sigma_v^2 & \sigma_{vu}\\ \sigma_{vu} & \sigma_u^2\end{bmatrix}$$

From these equations, we can write the model in its reduced form:

$$P_i^*=\mathbf{z}_i'\alpha+\sum_j\left[\left(\mathbf{I}-\rho_1\mathbf{W}^1\right)^{-1}\right]_{ij}v_j \tag{16}$$

$$y^*=\mathbf{x}_i'\beta+\mathbf{h}_i'\pi+\left(\mathrm{Irr}_i\cdot\tilde{\mathbf{x}}_i\right)'\xi+\sum_j\left[\left(\mathbf{I}-\rho_2\mathbf{W}^2\right)^{-1}\right]_{ij}u_j \tag{17}$$

where, again, $\left[(\mathbf{I}-\rho\mathbf{W})^{-1}\right]_{ij}$ is the (i,j) element of the spatial multiplier matrix $(\mathbf{I}-\rho\mathbf{W})^{-1}$, where the subscripts and superscripts refer to the selection equation and the outcome equation, respectively. Flores-Lagunes and Schnier (2010) note that the probit model with spatially autoregressive errors introduces a fully non-spherical variance-covariance matrix that renders the regular probit estimator inconsistent. Thus estimation of equation (14) must proceed with the spatial probit model of Pinkse and Slade (1998). The outcome equation (15) is estimated using the FGLS estimator in Kelejian and Prucha (1999).

From McMillen (1995), we have the following variance and covariance calculations:

$$\mathrm{Var}(\varepsilon_1)=\sigma_v^2\sum_j\left[\left(\mathbf{I}-\rho_1\mathbf{W}^1\right)^{-1}\right]_{ij}^2 \tag{18}$$

$$\text{Var}(\varepsilon_2) = \sigma_u^2 \sum_j \left[\left(\mathbf{I} - \rho_2 \mathbf{W}^2 \right)^{-1} \right]_{ij}^2 \tag{19}$$

$$E(\varepsilon_{1i}\varepsilon_{2i}) = \sigma_{vu} \sum_j \left[\left(\mathbf{I} - \rho_1 \mathbf{W}^1 \right)^{-1} \right]_{ij} \left[\left(\mathbf{I} - \rho_2 \mathbf{W}^2 \right)^{-1} \right]_{ij} \tag{20}$$

From $\text{Var}(\varepsilon_{1i})$, Pinkse and Slade (1998) construct "generalized" residuals with which to construct appropriate moment conditions for consistent estimation of the model parameters, taking into consideration the induced heteroskedasticity. Letting the vector of parameters in the selection equation be given as $\theta_1 = [\alpha', \rho_1]$ and letting $\delta_i(\theta_1) = \frac{\mathbf{z}_i'\alpha}{\sqrt{\text{Var}(\varepsilon_{1i})}}$ be the index function of a probit model weighted by the standard deviation of the residuals from the selection equation, the "generalized" residuals of the selection equation are:

$$\tilde{\varepsilon}_{1i}(\theta_1) = \sqrt{\sigma_v^2 \sum_j \left[\left(\mathbf{I} - \rho_1 \mathbf{W}^1 \right)^{-1} \right]_{ij}^2} \cdot \{P_i - \Phi[\delta_i(\theta_1)]\} \cdot \frac{\phi[\delta_i(\theta_1)]}{\Phi[\delta_i(\theta_1)]\{1 - \Phi[\delta_i(\theta_1)]\}} \tag{21}$$

The GMM estimator for θ_1 is given as

$$\theta_{1,GMM} = \text{argmin} \left\{ S(\theta_1)' \mathbf{M}_n S(\theta_1) \right\} \tag{22}$$

where $S(\theta_1) = \frac{1}{n}\mathbf{Z}'\tilde{\varepsilon}_1(\theta_1)$, where $\mathbf{Z}$ is the matrix of variables in the selection equation and $\tilde{\varepsilon_1}(\theta_1)$ is the vector of generalized residuals, and $\mathbf{M}_n$ is a conformable positive definite weighting matrix. Consistent estimates of θ_1 are then used to construct the "adjusted"-IMR (from McMillen, 1995) to be used in the outcome equation. The "adjusted"-IMR is given as:

$$\lambda_i \equiv \frac{\sum_j \left[\left(\mathbf{I} - \rho_1 \mathbf{W}^1 \right)^{-1} \right]_{ij} \left[\left(\mathbf{I} - \rho_2 \mathbf{W}^2 \right)^{-1} \right]_{ij}}{\sqrt{\sum_j \left[\left(\mathbf{I} - \rho_1 \mathbf{W}^1 \right)^{-1} \right]_{ij}^2}} \cdot \frac{\phi[-\delta_i(\theta_1)]}{1 - \Phi[-\delta_i(\theta_1)]} \tag{23}$$

This "adjusted"-IMR depends on the spatial correlation coefficient from the outcome equation (ρ_2), which is not estimated in this first step spatial probit. Likewise estimating the conditional outcome response requires the inclusion of this "adjusted"-IMR as an additional explanatory variable in the outcome equation. Thus, all of the parameters in both the selection and the outcome equations must be estimated simultaneously in order to increase the efficiency of the estimator and maintain an estimable variance-covariance matrix. To accomplish this, Flores-Lagunes and Schnier (2010) stack the moment conditions of the selection and outcome equations:

$$g(\mathbf{Z}, \mathbf{X}, \mathbf{H}, \tilde{\mathbf{X}}, \theta) = \left[s(\mathbf{Z}, \theta)', m(\mathbf{X}, \mathbf{H}, \tilde{\mathbf{X}}, \theta)'\right]' \tag{24}$$

where, now, $\theta = [\alpha', \rho_1, \beta', \pi', \xi', \eta, \rho_2]$. The components of this matrix are as follows:

$$s(\mathbf{Z}, \theta) = \mathbf{Z}'\tilde{\varepsilon_1}(\theta) \tag{25}$$

$$m(\mathbf{X}, \mathbf{H}, \tilde{\mathbf{X}}) = \left[P \cdot \left(\mathbf{X}, \mathbf{H}, \tilde{\mathbf{X}}, \hat{\lambda}\right)\right]' \tilde{\varepsilon}_2(\theta) \tag{26}$$

where $\tilde{\varepsilon}_2(\theta) = y - \mathbf{X}\beta - \mathbf{H}\pi - (\text{Irr} \cdot \tilde{\mathbf{X}}) - \eta\hat{\lambda}(\rho_1, \rho_2, \alpha)$. Stacking the generalized residuals from the spatial probit estimation with the residuals from the outcome equation, we get $\tilde{\varepsilon}(\theta) \equiv [\tilde{\varepsilon}_1'(\theta), \tilde{\varepsilon}_2'(\theta)]$. Then a consistent GMM estimator for all of the model parameters is:

$$\theta_{GMM} = \text{argmin}\left\{g_n(\theta)'\mathbf{M}_n g_n(\theta)\right\} \tag{27}$$

with $g_n = \dfrac{1}{n}\mathbf{Z}'\tilde{\varepsilon}(\theta)$ and, again, $\mathbf{M}_n$ is a conformable positive definite weighting matrix. In Flores-Lagunes and Schnier (2010), they propose two variants of this "spatial heckit" model, an equal weight version and an optimal weighting version. It has been shown that, at least in finite samples, the optimal GMM estimator may be biased, and thus an equally-weighted GMM estimator (i.e., one in which $\mathbf{M}_n$ is an identity matrix) may be preferable. The Monte Carlo simulations in Flores-Lagunes and Schnier (2010) suggest that, at least at larger sample sizes, the optimally-weighted GMM estimator performs slightly better than the equally-weighted GMM estimator in terms of estimating the model parameters (both in terms of bias and root mean squared error), but performs slightly worse than the equally-weighted GMM estimator in estimating the spatial correlation coefficients for both the selection and outcome equations.[18] For our purposes, we use the optimally-weighted GMM estimator, which estiamtes the model parameters in two steps, takes the estimates from a first-stage, equally-weighted GMM estimation and constructs the optimal-weights for use in a second-stage GMM estimation. The results from this estimation are in Table 7.

6 Discussion of Results

There is highly significant (and strong) spatial dependence in the unobserved error terms in both the selection equation and the response equation. In both equations, the correlation of spatial autocorrelation is roughly 0.83, which is likely so high because of the geographic scale involved and the nature of the spatial system. It is interesting, however, that the autocorrelation is so similar between the two equations, even though the different equations represent interactions among different spatial systems.[19]

While not reported in Table 7, it should be noted that all of the coefficients associated with the country dummy variables were significantly different from zero at the $\alpha = 1\%$ probability of Type I error. This suggests that, even after controlling for the effects of temperature, precipitation, and soil chemistry, country specific factors do significantly contribute to predicting cereal yields. In fact, all of the coefficients on these variables are positive, indicating that all countries should, other things equal, experience higher yields than the reference country, which in this case is Angola. The largest country fixed effect belongs to Swaziland, which should expect roughly 3,700 more pounds per hectare than Angola. The smallest country fixed effect belongs to Mozambique, which produces on average only about 488 more pounds per hectare than Angola.

The estimates in Table 7 don't provide the true marginal effects of small changes in the explanatory variable on the cereal yields, since the true marginal effects should take into consideration both the effect that the explanatory variable in question has on the outcome equation, but also on the

[18] For comparison purposes, we compared the simulated results for $\rho_1 = 0.75$, $\rho_2 = 0.75$, and a correlation between the errors in the two equations of 0.5, with sample selection of 25% and sample size of n=900.

[19] For concreteness, the selection equation represents all 2,505 grid cells, with some cells having zero (or unobserved) yields and some cells having positive (or observed) yields, while the outcome equation only considers those grid cells for which there are observed yields. Thus, the spatial system in the selection equation more closely approximates a regular lattice—and thus the spatial weights matrix approximates a weighted contiguity matrix—while the spatial system in the outcome equation is less regular in structure.

selection equation. From McMillen (1995), expected yields conditional upon the decision to plant cereals is given as:

$$E[y_i|P_i = 1, \mathbf{x}_i,\tilde{\mathbf{x}}_i,\mathbf{z}_i] = \mathbf{x}_i'\beta + \mathbf{h}_i'\pi + (\text{Irr}_i \cdot \tilde{\mathbf{x}}_i)'\,\xi + \frac{\sigma_{vu}}{\sigma_v}\hat{\lambda}_i \tag{28}$$

Of particular interest in this analysis is the effect of temperature and precipitation on agricultural production. Therefore, in order to estimate the effects of changes in temperature on cereal yields, the marginal effects can be computed as:

$$\frac{\partial E\left(y|P = 1, \mathbf{X}, \tilde{\mathbf{X}}, \mathbf{Z}\right)}{\partial \text{Temp}} = \beta_{\text{Temp}} + 2\beta_{\text{Temp}^2} \cdot \text{Temp} + \xi_{\text{Temp}} \cdot \text{Irr} + \frac{\sigma_{uv}}{\sigma_v}\frac{\partial \lambda}{\partial \text{Temp}} \tag{29}$$

Clearly, dividing by yields will result in the semi-elasticity of yields with respect to temperature; that is, the percentage change in yields resulting from a discrete change in temperature. Likewise, the effect of changes in yields brought about by changes in precipitation can be computed as:

$$\frac{\partial E\left(y|P = 1, \mathbf{X}, \tilde{\mathbf{X}}, \mathbf{Z}\right)}{\partial \text{Precip}} = \beta_{\text{Precip}} + 2\beta_{\text{Precip}^2} \cdot \text{Precip} + \xi_{\text{Precip}} \cdot \text{Irr} + \frac{\sigma_{uv}}{\sigma_v}\frac{\partial \lambda}{\partial \text{Precip}}$$

At present, it is not entirely clear whether a closed form analytical solution for these marginal effects can be derived, since it requires the partial derivative of the "adjusted" inverse Mills Ratio with respect to the various explanatory variables of interest. Additionally, because of the computational intensity of estimating the full set of model parameters simultaneously in this iterative GMM procedure, deriving any kind of numerical marginal effects also seems infeasible. If we assume that planting decisions are static, then changes in temperature, for example, *would not* affect the selection to plant cereals, but *would* affect the yield response. This implies, therefore, that $\frac{\partial \lambda}{\partial X_j} = 0$. This may be a strong assumption, since planting decisions are part of a dynamic process, and because changes in temperature and precipitation are expected to occur over a gradual period rather than instantly as such a marginal effect calculation would imply. Nevertheless, to make the analysis of these marginal effects tractable, such simplifying assumptions may be reasonable.

Based on this discussion, therefore, a 1°C increase in temperature would, on average, result in only a 6% decrease in yields, where this semi-elasticity is evaluated at the mean temperature, irrigation level, and current yield for those grid cells that are currently planting cereals. This is significantly lower than the 10%-15% reductions in maize yields, but roughly in line with the 5%-6% reduction in barley yields predicted by Tebaldi and Lobell (2008) with a similar temperature reduction. As would be expected, the average current yields for those grid cells without irrigation is significantly lower than the average irrigation for the entire sample (1,749 lbs. per hectare compared with 2,145 lbs. per hectare). Nevertheless, even for areas without irrigation, a 1°C increase in mean temperature would only be expected to reduce yields by only 8%. Because the coefficient on the quadratic temperature term is not significantly different from zero, and because the coefficients on the interaction terms (both with temperature and its square) are both not significantly different from zero, it may make more sense to assume that these terms have no effect on yields and thus only consider the linear effect of temperature on yields. Under this approach, we find that a 1°C increase in mean temperature would decrease yields by roughly 11.5% (still toward the low end of the Tebaldi and Lobell (2008) estimates) across the entire system, but by 14% for those areas without irrigation.

Interestingly, these results predict a small negative effect on yields with increased precipitation, though, strangely, an increase in the standard deviation of precipitation is expected to have a net positive effect on production. A 1% increase in precipitation is expected to result in a 7% reduction in cereal yields (or conversely, a 1% decrease in precipitation would—on average—result in a 7% increase in cereal yields), though it should be noted that only the quadratic precipitation coefficient is stastically different from zero at standard levels. Additionally, as yet there is not a significant consensus on the climate change-related effects on precipitation, as there are simply too many variables in flux to make realistic predictions, including the effects of aerosols in the atmosphere, evapotranspiration patterns, changing oceanic currents and wind patterns, etc. A generally accepted principle is that precipitation is expected to increase globally, but the distribution of precipitation and the magnitude of the increase are still uncertain. Thus even if the estimated marginal effects were significantly different from zero, it would be difficult to assess how changes in precipitation resulting from global climate change would be expected to effect cereal yields in

Sub-Saharan Africa.

The predictions reported in Table 7 and the marginal effects discussed here, of course, assume that all other factors are held constant and that there is no structural adaptation. This may not be a realistic assumption, as agriculture has often shown itself to be a progressive and innovative sector, and could very well exhibit strong adaptation to changes over the course of the next century. This analysis also ignores the effect of increased atmospheric carbon, which serves as a fertilizer and could realistically be expected to offset some of the negative effects of climate change on the agricultural sector. Additionally, we also assume that there are no other unforeseen effects that could affect the level of production, such as catastrophic climate extremes or other natural disasters that could dramatically affect the level of production. We leave it to future researchers to more effectively coordinate these more complex elements into a spatial econometric framework. However, because we incorporate spatial dependence among unobserved factors across the spatial system, we can control for the technicality that such unobserved factors would have spatial spillovers into neighboring grid cells. Controlling for this form of spatial dependence also controls for the effects of positive unobserved spillovers. Spatial correlation of this form, however, does not have any form of direct interpretation, and controlling for this through modeling spatial dependence only handles this statistical technicality, which, for example, Bell and Bockstael (2000) have referred to as a nuisance parameter.

7 Conclusion

In this paper we have estimated a spatial econometric model to examine grid cell level agricultural productivity (specifically cereal grain productivity) across sub-Saharan Africa. As part of the econometric approach, we specified the model such that it allowed us to examine the dependence of agricultural production to a variety of exogenous factors, taking into consideration spatial dependence and heterogeneity, and explicitly control for the effect of an endogenous sample selection mechanism. Interestingly, we find that once we control for heterogeneity and spatial spillovers, changes in average temperature are expected to have a smaller negative effect on cereal production than recent climatological studies suggest. Also, contrary to previous studies that have attempted to predict the effect of anticipated precipitation changes on agricultural production, we find a small negative effect on grid cell level production, though the effects are not significantly different from zero. Rather, we find that one of the most important factors in determining a cell's level of agricultural productivity is the country to which it belongs. Dummy variables corresponding to the cell's country of origin capture country-specific effects even after controlling for the other explanatory variables in this model. All coefficients on these dummy variables are highly significant, suggesting a great deal of spatial heterogeneity across sub-Saharan Africa, independent of geographical factors like elevation or terrain or distance to the coastline, and also independent of environmental factors like precipitation and temperature. This heterogeneity could be due to many factors, including governmental subsidies for fertilizers, the general availability of fertilizers, extension programs to assist farmers' decision-making, a sound legal system which provides stringent property rights, a system of functioning input and output markets, supply chain management, price transmission, etc. These are vital components of a well-functioning agricultural sector, but in a continent like Africa in which it is nearly impossible to find a paradigm of good governance, the existence of these institutions and structures may provide an even greater marginal benefit than they do in other parts of the world.

References

Anselin, L. (1988). *Spatial econometrics: methods and models*. Springer.

Anselin, L. (2003). Spatial externatlities, spatial multipliers, and spatial econometrics. *International Regional Science Review 26*(2), 153–166.

Anselin, L. (2005). *Exploring Spatial Data with GeoDa : A Workbook*. Center for Spatially Integrated Social Science.

Anselin, L., A. K. Bera, R. Florax, and M. J. Yoon (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics 26*(1), 77–104.

Battisti, D. and R. Naylor (2009). Historical warnings of future food insecurity with unprecedented seasonal heat. *Science 323*(2), 240–244.

Bell, K. and N. Bockstael (2000). Applying the generalized-moments estimation approach to spatial problems involving micro-level data. *Review of Economics and Statistics 82*(1), 72–82.

Cline, W. (2007). *Global Warming and Agriculture*. Center for Global Development and Peterson Institute for International Economics.

Darwin, R., M. Tsigas, J. Lewandrowski, and A. Raneses (1995). World Agriculture and Climate Change. Agricultural Economics Report 703, United States Department of Agriculture, Economic Research Service.

Davidson, R. and J. MacKinnon (1993). *Estimation and inference in econometrics*. Oxford University Press.

De Araujo Marinho, E. C. (2008). *Economics of Hunger: Several Methods, levels and Scales*. Ph. D. thesis, Universite Catholique de Louvain, Belgium.

Diffenbaugh, N. S., F. Giorgi, L. Raymond, and X. Bi (2007). Indicators of 21st century socioclimatic exposure. *Proceedings of the National Academy of Sciences 104*(51), 20195–20198.

Dow, K. and T. E. Downing (2007). *The Atlas of Climate Change: Mapping the World's Greatest Challenge*. Berkely: University of California Press.

Flores-Lagunes, A. and K. Schnier (2010). Estimation of sample selection models with spatial dependence. *Journal of Applied Econometrics Forthcoming*.

Greene, W. (2003). *Econometric Analysis* (5 ed.). Prentice Hall Upper Saddle River, NJ.

Heckman, J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and social Measurement 5*(4), 475–492.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica 47*(1), 153–161.

IPCC, W. (2007). *Summary for Policy Makers*. Cambridge: Cambridge University Press.

Kelejian, H. and I. Prucha (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics 17*(1), 99–121.

Kelejian, H. and I. Prucha (1999). A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *International Economic Review 40*(2), 509–533.

Kelejian, H. and I. Prucha (2007). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. Working Paper 2448, CESifo.

Maxwell, S. and T. R. Frankenberger (1992). Household food security: Concepts, indicators, measurements: A technical review. Technical report, UNICEF, New York.

McMillen, D. P. (1995). Selection Bias in Spatial Econometric Models. *Journal of Regional Science 35*(3), 417–436.

Mendelsohn, R., W. Nordhaus, and D. Shaw (1994). The impact of global warming on agriculture: a Ricardian analysis. *The American Economic Review 84*(4), 753–771.

Monfreda, C., N. Ramankutty, and J. Foley (2008). Farming the Planet 2: Geographic Distribution of Crop Areas, Yields, Physiological Types, and Net Primary Production in the Year 2000. *Global Biogeochemical Cycles 22*, GB1022, doi:10.1029/2007GB002947.

Nafziger, E. W. (2003). Inequality and conflict. In E. Nnadozie (Ed.), *African Economic Development*, pp. 295–312. San Diego: Academic Press.

Nordhaus, W. (2006). Geography and macroeconomics: New data and new findings. *Proceedings of the National Academy of Sciences 103*(10), 3510–3517.

Nordhaus, W., Q. Azam, D. Corderi, K. Hood, N. M. Victor, M. Mohammed, A. Miltner, and J. Weiss (2006). The G-Econ Database on Gridded Output: Methods and Data. Technical report, G-Econ Project, Yale University, New Haven, CT, USA.

Pinkse, J. and M. Slade (1998). Contracting in space: An application of spatial statistics to discrete-choice models. *Journal of Econometrics 85*(1), 125–154.

Rosenzweig, C. and M. L. Parry (1994). Potential impact of climate change on world food supply. *Nature 367*(6459), 133–138.

Schlenker, W., W. Hanemann, and A. Fisher (2005). Will US agriculture really benefit from global warming? Accounting for irrigation in the hedonic approach. *American Economic Review 95*(1), 395–406.

Sen, A. K. (1999). *Development as Freedom*. New York: Anchor Books.

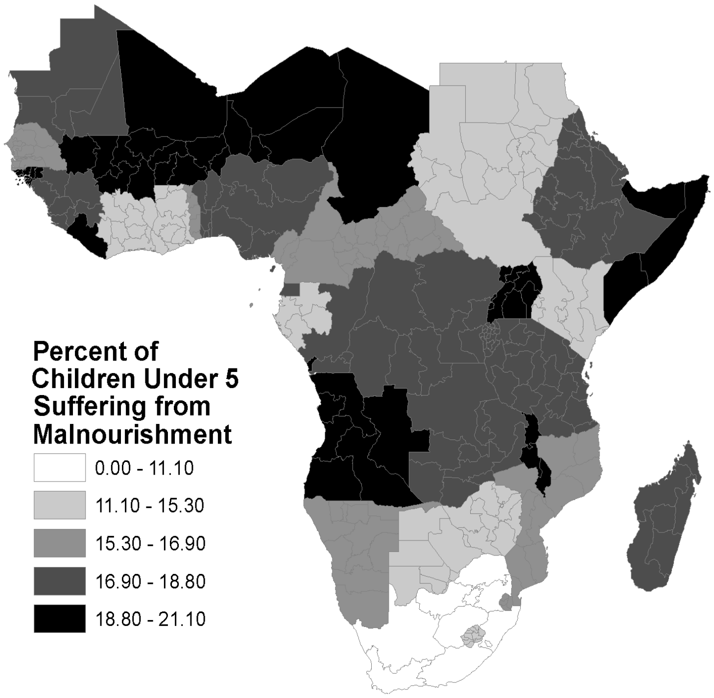
Tebaldi, C. and D. B. Lobell (2008). Towards probabilistic projections of climate change impacts on global crop yields. *Geophysical Research Letters 35*.

Wooldridge, J. (2006). *Introductory econometrics* (3 ed.). Thomson/South-Western: Mason, OH.

World Bank (1986). *Poverty and Hunger: Issues and Options for Food Security in Developing Countries*. Washington, D.C: World Bank.

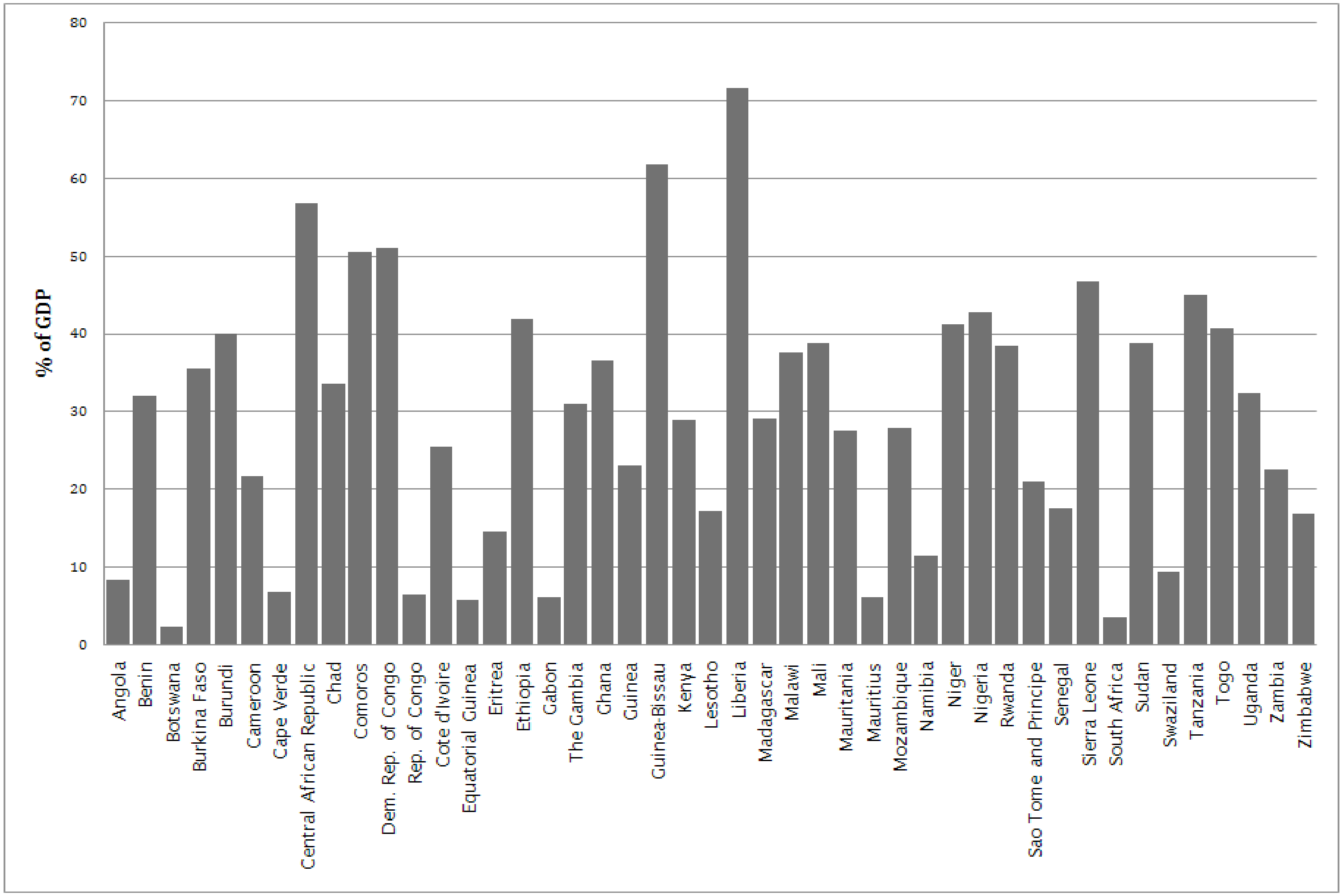
World Bank (2006). *World Development Indicators*. The World Bank Washington, DC.

Figure 1: Malnourishment in Sub-Saharan Africa



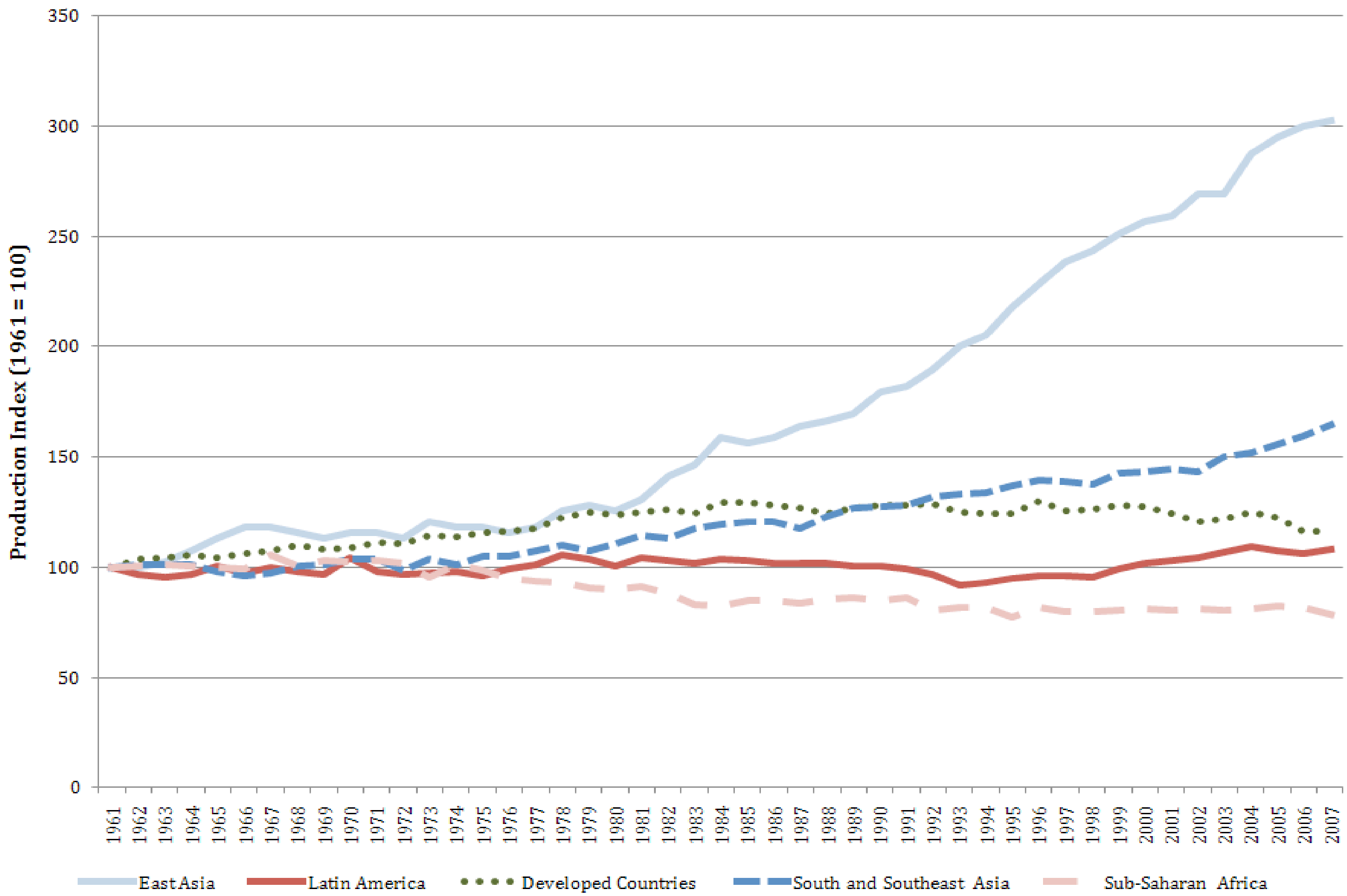
Source: Center for International Earth Science Information Network (CIESIN)

Figure 2: Sub-Saharan Africa's dependence on Agriculture (2003)



Source: World Bank (2007)

Figure 3: Growth in food production per capita, 1961-2007 (1961 = 100)



Source: Food and Agricultural Organization (FAO)

Table 1: Indices of Food Production per Head by Regions

Regions	1963-1967	1968-1973	1973-1977	1978-1982	1983-1987	1988-1992	1993-1997	1998-2002	2003-2007
World	82.2	83.4	85.5	88.0	90.6	92.6	94.6	99.2	104.8
Africa	106.4	108.4	102.5	95.2	90.8	95.2	96.2	99.4	102.8
Asia	58.0	59.2	61.0	65.2	73.4	80.4	91.0	99.2	108.4
China	38.8	39.6	40.5	46.2	57.2	66.0	83.8	99.6	114.8
India	72.6	75.0	76.3	78.2	84.2	91.0	95.6	97.4	101.0
Central America	89.2	89.6	90.8	94.8	93.0	90.2	94.2	99.0	106.4
South America	70.8	70.8	74.0	78.0	79.6	83.4	89.6	99.2	112.0
Europe	92.4	98.2	104.0	108.6	113.2	111.4	99.6	99.4	99.8
North America	78.4	80.0	87.75	93.6	89.8	90.4	95.8	98.4	99.8
United States	78.2	80.4	88.75	94.6	90.0	91.0	96.0	98.8	99.6

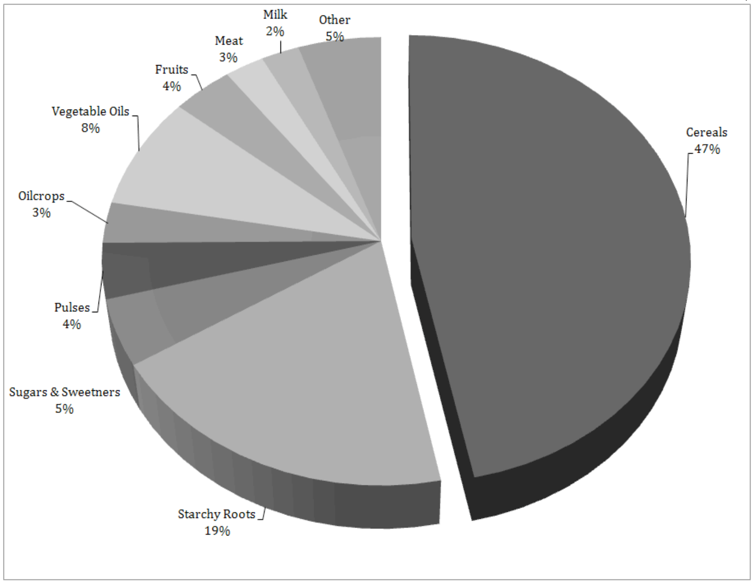
Source: FAO (2010) Note: the three-year average of 1999-2001 is used as the base for comparisons. Reported figures represent five-year averages in the Net per capita production index.

Table 2: Preferred estimates of impacts of baseline global warming by the 2080s on African Agriculture

			Impact without carbon fertilization		Preferred estimates		Changes in output (millions of 2003 dollars)	
Country	Farm Area (1,000 hectares)	Output per hectare (2003 dollars)	Ricardian Estimate (percent)	Crop models Estimate (percent)	Without carbon fertilization (percent)	With carbon fertilization (percent)	Without carbon fertilization	With carbon fertilization
Angola	3,300	360	-26.3	-25.3	-25.8	-14.7	-306	-174
Burkina Faso	6,830	190	-16.5	-32.1	-24.3	-13	-315	-168
Cameroon	7,160	768	-19.8	-20	-20	-8	-1,100	-441
DR Congo	7,800	422	-25.3	-14.7	-14.7	-1.9	-484	-64
Ethiopia	11,047	253	-31.4	-31.1	-31.3	-20.9	-873	-585
Ghana	6,331	434	-8.2	-19.8	-14	-1.1	-384	-30
Cote d'Ivoire	6,900	518	-8.8	-19.8	-14.3	-1.5	-511	-52
Kenya	5,162	446	15	-25.7	-5.4	8.8	-123	203
Madagascar	3,550	447	-20.3	-32.1	-26.2	-15.1	-416	-240
Malawi	2,440	267	-31.5	-31.1	-31.3	-21	-204	-137
Mali	4,700	350	-39	-32.1	-35.6	-25.9	-585	-426
Mozambique	4,435	253	-23.6	-19.8	-21.7	-10	-244	-112
Nigeria	33,000	460	-12.1	-24.9	-18.5	-6.3	-2,809	-953
Senegal	2,506	441	-84	-19.8	-51.9	-44.7	-573	-493
South Africa	15,712	407	-47	-19.8	-33.4	-23.4	-2,134	-1,495
Sudan	16,653	417	-81.1	-31.1	-56.1	-49.5	-3,892	-3,435
Tanzania	10,764	430	-16.3	-32.1	-24.2	-12.8	-1,122	-595
Uganda	7,200	280	-2.5	-31.1	-16.8	-4.3	-338	-86
Zambia	5,289	189	-47.1	-32.1	-39.6	-31	-395	-305
Zimbabwe	3,350	901	-44.7	-31.1	-37.9	-29	-1,144	-863
World	3,097,935	380	-10.1	-18.9	-15.9	-3.2	-186,510	-38,107
Median			-9.8	-19.8	-16.7	-4.2		

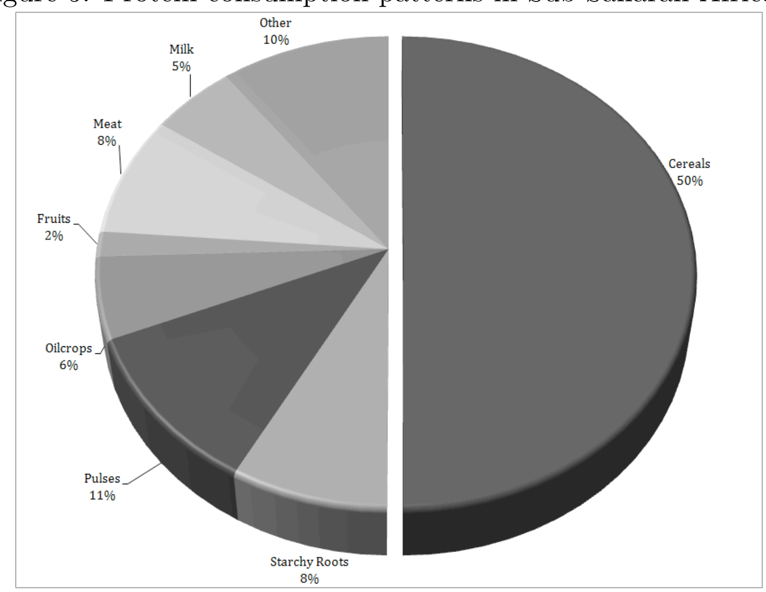
Source: Cline (2007)

Figure 4: Food consumption patterns in Sub-Saharan Africa (Kcal/capita/day)



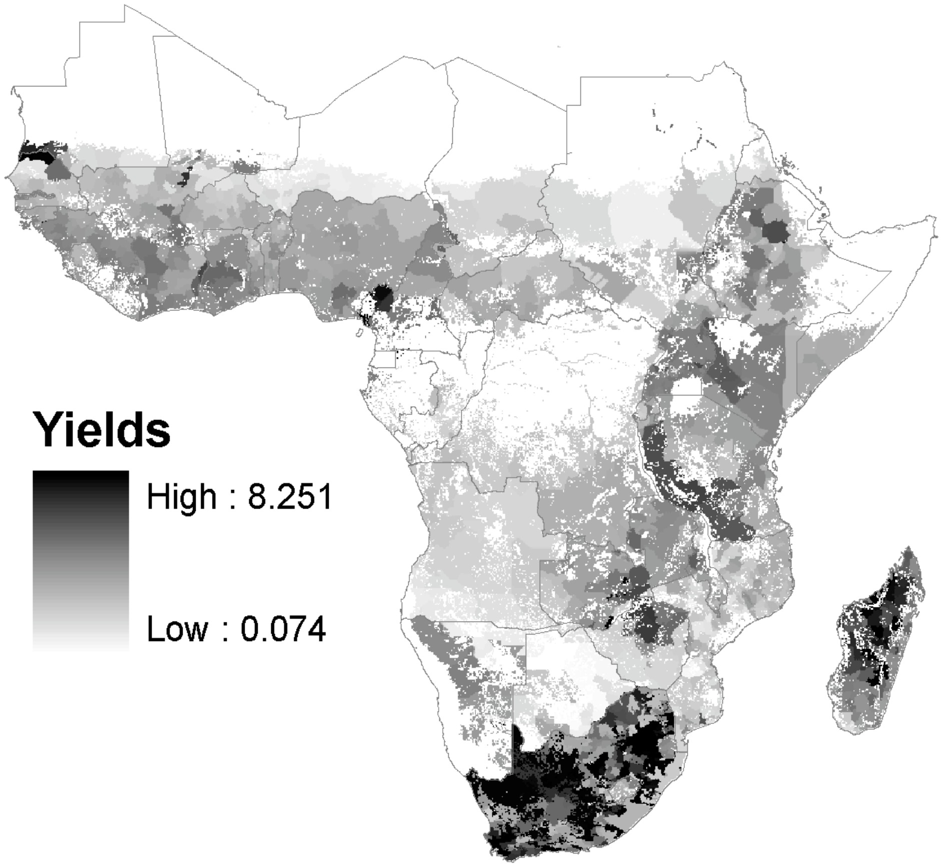
Source: Food and Agricultural Organization (FAO)

Figure 5: Protein consumption patterns in Sub-Saharan Africa (g/capita/day)



Source: Food and Agricultural Organization (FAO)

Figure 6: Cereal Yields in Sub-Saharan Africa



Source: Monfreda et al. (2008)

Figure 7: Moran scatterplot–Cereal Yields (2000)

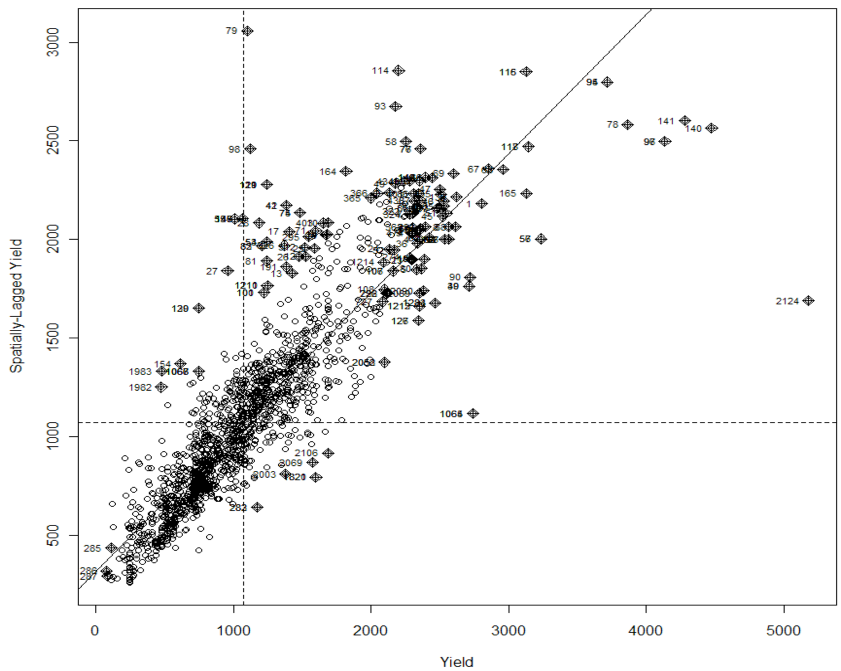


Figure 8: Moran scatterplot–OLS Residuals

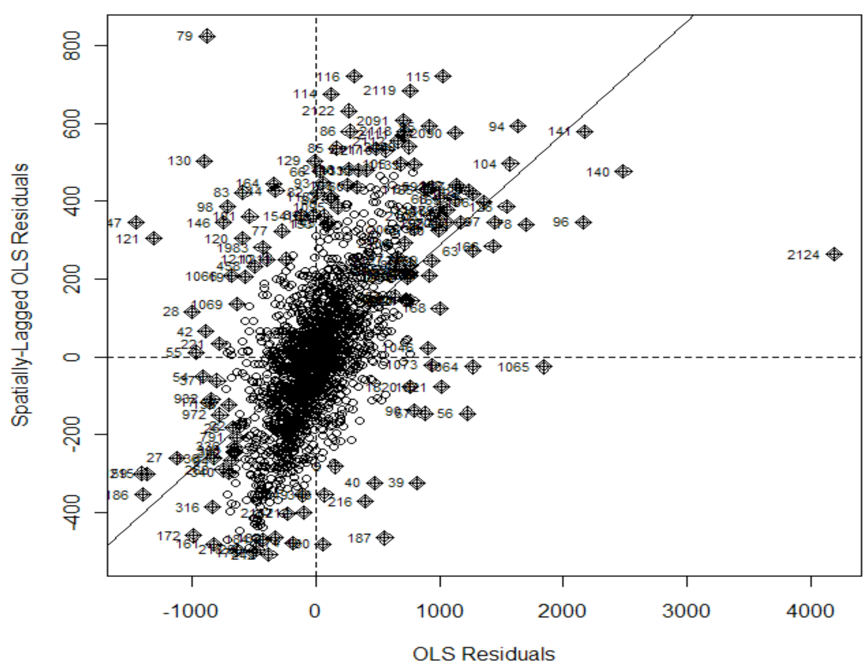


Figure 9: Local Indicators of Spatial Autocorrelation: Cereal Yields

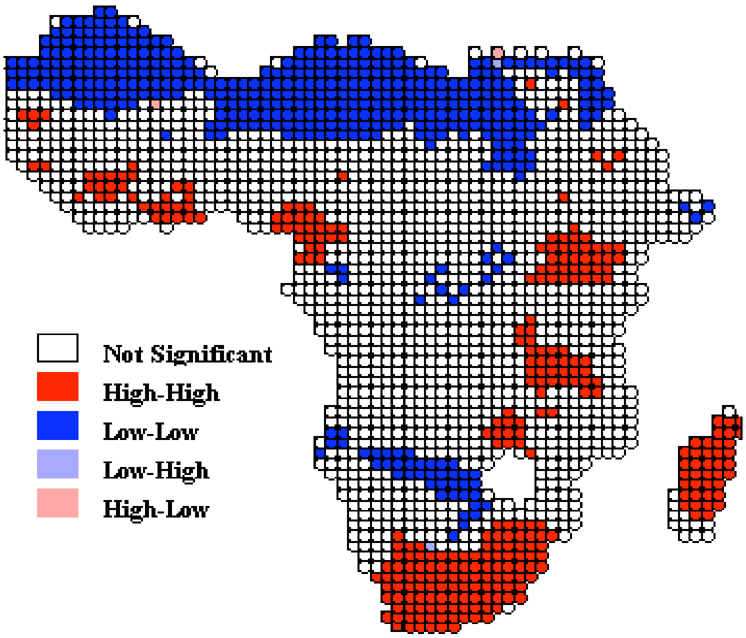


Figure 10: Distribution of Precipitation in Sub-Saharan Africa

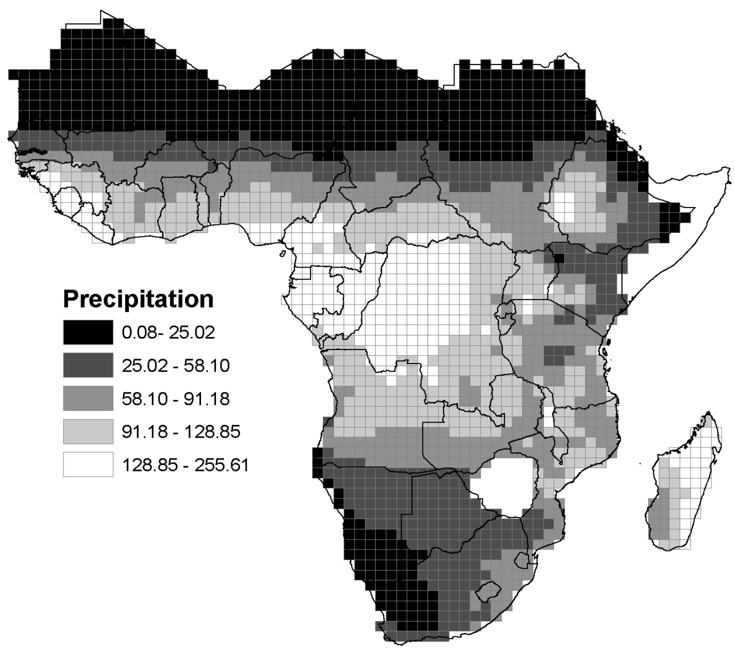


Table 3: Summary Statistics for Economic and Environmental/Geographic Data

Variable	Description	N	Mean	Std Dev	Min	Max
y	Cereal yields, 2000 (pounds per hectare)	2,124	2145.1227	1074.6486	157.1400	10362.0000
Gross Cell Prod	Gross Cell Product, 2000 ($US millions)	2,505	285.1281	1597.0102	0.0000	55089.9900
Pop	Population (,000s)	2,505	234.6821	510.1857	0.0010	6984.1200
Temp	Average annual temperature (degrees Celsius)	2,505	24.5655	3.4192	10.2100	30.3200
Std Dev–Temp	Standard deviation of temperature	2,505	2.6879	1.7725	0.2700	7.7200
Precip	Average precipitation (mm per month)	2,505	70.7741	48.8425	0.0800	255.6100
Std Dev–Precip	Standard deviation of precipitation	2,505	59.4667	36.1757	0.1300	246.5700
Elev	Elevation (m)	2,505	669.7979	428.2270	4.5943	2575.2210
Rough	Roughness of elevation	2,505	9.2080	10.1118	0.0000	60.0000
Dist	Distance to coast (km)	2,505	697.9023	434.9767	4.1000	1686.7000
Irr	Irrigation (% of grid cell with improved irrigation)	2,505	1.4293	3.6743	0.0000	52.8933
pH	Soil pH (Index ranging from 0 to 99)	2,505	32.6009	12.3194	15.0625	99.0000
Carbon	Soil carbon density (kg C/m^2)	2,505	8.3005	3.7966	0.0000	27.7155

Table 4: Selected Coefficients: OLS Estimation

	Estimate	Std Error		Robust Std Error	
Constant	4,560.9506	1,200.9436	***	1,643.2880	***
Temp	-152.6618	91.0974	*	132.7378	
Temp^2	0.0993	1.9152		2.9559	
Std Dev–Temp	15.6707	39.9436		73.6618	
Precip	-0.4124	3.1637		3.2216	
Precip^2	0.0016	0.0121		0.0118	
Std Dev–Precip	2.9202	1.2920	**	1.2184	**
Elev	-0.4516	0.1333	***	0.1980	**
Rougness	-0.4191	0.0754	***	2.6558	
Dist	4.2196	2.6347		0.0687	***
Irr	-187.5909	274.8689		381.2729	
pH	3.9913	3.0338		2.5573	
Carbon	20.5116	5.9612	***	4.6336	***
Irr · Temp	16.3065	20.4696		28.0050	
Irr · Temp^2	-0.2937	0.4245		0.5758	
Irr · Std Dev–Temp	34.8060	10.1579	***	13.1456	***
Irr · Precip	-0.9582	0.9053		1.0944	
Irr · Precip^2	0.0036	0.0033		0.0037	
Irr · Std Dev–Precip	1.2292	0.4501	***	0.4725	***
Irr · Elev	0.0541	0.0251	**	0.0260	**
Irr · Rough	-1.2134	0.8919		0.9040	
Irr · Dist	-0.0396	0.0212	*	0.0208	*
Irr · pH	-2.4196	0.9827	**	1.0745	**
Irr · Carbon	-6.4894	1.6474	***	1.8038	***
	$R^2 = 0.5606$				
	N = 2,124				

0 *** 0.01 ** 0.05 * 0.1

Table 5: Selected Coefficients: Non-Spatial Heckman Two-Stage Estimation

Selection Equation			
	Estimate	Std Error	
Constant	18.88829	3.680936	***
Gross Cell Prod	-0.00011	0.000053	**
Pop	0.59672	0.591912	
Temp	-1.78091	0.305817	***
Temp^2	0.03757	0.006292	***
Std Dev–Temp	0.04240	0.071687	
Precip	0.02450	0.008396	***
Precip^2	-0.00019	0.000037	***
Std Dev–Precip	0.05286	0.005735	***
Elev	0.00100	0.000385	***
Rougness	-0.00057	0.000172	***
Dist	-0.00157	0.009234	
Irr	0.15023	0.038662	***
pH	-0.01729	0.005856	***
Carbon	0.12620	0.030913	***

N = 2,505
Pseudo-R^2 = 0.7317
Percent "Successes" Correctly Predicted: 98.26%
Percent "Failures" Correctly Predicted: 86.88%
Total Percent Correctly Classified: 96.53%

Outcome Equation			
	Estimate	Std Error	
Constant	5,218.92154	1,171.55119	***
Temp	-246.70040	90.39324	***
Temp^2	2.33343	1.91052	
Std Dev–Temp	-69.89141	39.36085	*
Precip	2.45838	3.16929	
Precip^2	-0.02653	0.01271	**
Std Dev–Precip	7.41904	1.39617	***
Elev	-0.35576	0.13151	***
Rougness	-0.47011	0.07454	***
Dist	3.71275	2.66919	
Irr	-301.20845	275.20771	
pH	2.27741	2.91373	
Carbon	30.14600	6.20358	***
Irr · Temp	28.25790	20.72398	
Irr · Temp^2	-0.51905	0.43019	
Irr · Std Dev–Temp	36.82399	9.92721	***
Irr · Precip	-0.67608	0.92048	
Irr · Precip^2	0.00411	0.00337	
Irr · Std Dev–Precip	0.54944	0.46810	
Irr · Elev	0.05508	0.02567	**
Irr · Rough	-1.11866	0.90746	
Irr · Dist	-0.02738	0.02135	
Irr · pH	-2.81572	0.95641	***
Irr · Carbon	-7.89991	1.69728	***

N = 2,124
R^2 = 0.5754

0 *** 0.01 ** 0.05 * 0.1

Table 6: Selected Coefficients: Estimation of Spatial Autoregressive Error (SAE) Process

	Estimate	Std Error	
Constant	4,851.3060	1,502.5969	***
Temp	-114.5274	108.5669	
$Temp^2$	-0.6597	2.2445	
Std Dev–Temp	140.1086	63.7916	**
Precip	-0.6777	4.3599	
$Precip^2$	-0.0146	0.0156	
Std Dev–Precip	6.4747	2.1585	***
Elev	-0.5044	0.1910	***
Rougness	-0.4353	0.1706	**
Dist	2.2112	2.7554	
Irr	833.0684	261.2403	***
pH	0.5243	3.0612	
Carbon	8.9912	7.0127	
Irr · Temp	-59.6912	19.1703	***
Irr · $Temp^2$	1.2500	0.3926	***
Irr · Std Dev–Temp	-19.4439	9.1861	**
Irr · Precip	-1.6560	0.8610	*
Irr · $Precip^2$	0.0050	0.0032	
Irr · Std Dev–Precip	0.8783	0.4338	**
Irr · Elev	0.0215	0.0234	
Irr · Rough	-0.2270	0.7701	
Irr · Dist	0.0130	0.0189	
Irr · pH	-1.3549	0.8753	
Irr · Carbon	-3.5413	1.3921	**
ρ	0.8356		

0 *** 0.01 ** 0.05 * 0.1

Table 7: Selected Coefficients: Spatial Heckit Estimation

Selection Equation			
	Estimate	Std Error	
Constant	18.96727	6.64593	***
Gross Cell Prod	-0.00011	0.00003	***
Pop	0.00061	0.00033	*
Temp	-1.76596	0.52983	***
$Temp^2$	0.03779	0.01071	***
Std Dev–Temp	0.04215	0.12972	
Precip	0.02483	0.01412	*
$Precip^2$	-0.00019	0.00006	***
Std Dev–Precip	0.05376	0.00870	***
Elev	0.00102	0.00053	*
Rougness	-0.00058	0.00030	*
Dist	-0.00166	0.01260	
Irr	0.15022	0.09982	
pH	-0.01792	0.00719	**
Carbon	0.12567	0.03782	***
ρ_1	0.83167	0.02143	***

Selection Equation			
	Estimate	Std Error	
Constant	5,199.34287	1600.0305	***
Temp	-247.68161	129.1513029	*
$Temp^2$	2.32920	2.8120844	
Std Dev–Temp	-71.33907	67.1558629	
Precip	2.45716	3.2870627	
$Precip^2$	-0.02682	0.0139533	*
Std Dev–Precip	7.43186	1.945101	***
Elev	-0.35641	0.1718461	**
Rougness	-0.47020	0.0708304	***
Dist	3.69493	2.6055351	
Irr	-301.42896	359.7912169	
pH	2.34983	2.9159933	
Carbon	30.23793	5.4022809	***
Irr · Temp	28.23859	27.0950252	
Irr · $Temp^2$	-0.51679	0.5545297	
Irr · Std Dev–Temp	36.85476	12.0627811	***
Irr · Precip	-0.68368	0.9897805	
Irr · $Precip^2$	0.00410	0.0033838	
Irr · Std Dev–Precip	0.54866	0.4781438	
Irr · Elev	0.05496	0.0239988	**
Irr · Rough	-1.13132	0.8364094	
Irr · Dist	-0.02737	0.0191155	
Irr · pH	-2.82356	1.0006531	***
Irr · Carbon	-7.90629	1.8936299	***
IMR	864.93819	333.6108644	***
ρ_2	0.83040	0.0285239	***