



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

Borders and Barriers: Spatial Analysis of Agricultural Output Spillovers at the Grid Cell Level

Mesbah J. Motamed
Economic Research Service
U.S. Department of Agriculture
mmotamed@ers.usda.gov

Raymond J.G.M. Florax
Purdue University
VU University, Amsterdam
rflorax@purdue.edu

William A. Masters
Tufts University
william.masters@tufts.edu

May 21, 2012

Selected Paper prepared for presentation at the Agricultural & Applied Economics Associations 2012 Annual Meeting, Seattle, August 12-14. This paper has not been previously presented in any other meetings.

Borders and Barriers: Spatial Analysis of Agricultural Output Spillovers at the Grid Cell Level*

Mesbah J. Motamed, Economic Research Service-USDA[†]
Raymond Florax, Purdue University
William Masters, Tufts University

June 7, 2012

Abstract

Much of the world’s agricultural productivity growth over the past fifty years has been attributed to the international spillover of benefits from research and development (Alston, 2002). Because the impact of these spillovers often varies with distance, geography, and climate, attention has increasingly focused on measuring their spatial dimensions Evenson (1989). In this paper, spatial spillovers are examined in the context of country borders. Using high-resolution worldwide grid cell data covering yields for 11 major crops, we test whether spatial dependencies in agricultural yields fall in the presence of country borders, controlling for geography and climate as well as country-specific effects. Two different complementary analyses show that country borders significantly and sizeably diminish the transmission of spatial spillovers between locations. The results thus point to a clear “border effect” on an important determinant of agricultural productivity.

Keywords: agricultural productivity, spillovers, border effects

1 Introduction

Among the broad array of agricultural productivity’s determinants, one important contributor is the local sharing of knowledge, technology and information. Also known as spillovers, these neighbor-to-neighbor flows of indirect inputs have been observed to explain a large portion of worldwide productivity levels in agriculture (Alston, 2002). In this paper, we investigate the possibility that productive flows in Africa are being diminished or otherwise interrupted at international borders. Evidence of a border effect can potentially inform strategies for agricultural research, extension programs, and regional market and infrastructural designs.

*The views expressed are those of the authors and do not necessarily reflect the views of the Economic Research Service, United States Department of Agriculture.

[†]Corresponding author: mmotamed@ers.usda.gov

1.1 Border effects in the economic literature

Understanding how country borders matter to economic outcomes is an important objective for economic researchers and policy makers. Borders are significant inasmuch as they impede the international flow of goods and services and the transmission of prices, information, and technology. Measuring this so-called “border effect” hinges on the particular question being asked. For example, Engel and Rogers (1996) investigate the failure of the Law of One Price due to border effects. Examining the correlation of prices of consumer goods between the United States and Canada, they find that the variation in prices between two equidistant cities rises when a border lies between them. Crossing the border, they estimate, is akin to adding an additional 1,780 miles in distance between two same-country cities. Anderson and Wincoop (2003) estimate theoretically-consistent gravity models to obtain measurements of border effects in terms of an ad valorem tariff on traded goods. They compare trade flows between countries, controlling for distance and the size of each partner, but also examine border effects on domestically traded goods. Considered within a general equilibrium framework, the authors find that reducing border effects not only increases trade between countries, but also causes intra-country trade to rise. Approaching the concept from a different angle, Rietveld and Janssen (1990) examine the effects of borders on international telephone traffic inasmuch as telecommunications represents a complement to international exchanges of goods and services. In their analysis, the authors examine call volume both within and across countries, controlling for distance, long-distance charges, and of course borders. They estimate that borders reduce telephone traffic by approximately 30%.

These effects do not necessarily originate from the borders themselves. Rather, it is the geographic discontinuity in currencies, the origin and substitutability of goods, and markets and institutions, as summarized by a border, that typically underlie the outcome in question. Thus, another important vein of border effect research seeks to identify what factor (e.g., substitutability, exchange rate regimes, trade law, physical geography, or so-called psychic distance) is precisely responsible for the outcome. In a recent example, Evans (2003) decomposes the trade-volume impact of borders into several elements to argue that border effects are not as large as they seem and imply smaller welfare losses than are traditionally believed. The author demonstrates that, beyond simple distortionary barriers such as tariffs, cross-border trade is a function of the substitutability of the goods being

traded and the extra transaction and psychic costs associated with cross-border exchanges.

Meanwhile, a parallel stream of research approaches borders from another angle and investigates what benefits might accrue by virtue of sharing a border with a technology leader or a large market. For example, Conley and Ligon (2002) examine whether proximity to a trading partner can aid the growth process. To model the dependence of growth rates between countries, they conceive a spatial covariance matrix for which the off-diagonal elements are in part a function of the economic distance between countries. Data on commercial shipping and air travel serve as their proxy for economic distance. From their results, the authors observe not only considerable spatial dependence in the growth rates of countries, but also that, in certain instances, spillover effects explain a large portion of the spatial correlation between neighbors. In a different methodological approach, Rietveld and Wintershoven (1998) address whether spillovers in the supply of international transportation networks are smaller than their intranational equivalents. They construct a spatially explicit econometric specification with dual spatial weights matrices that identify within- and cross-country neighbors to measure the effects of neighbors who happen to lie on the other side of a border. Surprisingly, they find that cross-country neighbors exert a more positive influence on infrastructure development than within-country neighbors. Investigations into agricultural spillovers are also present in the literature. Alston (2002) surveys the benefits of agricultural research that spill over from one country to its neighbor and concludes that inefficiencies in the spatial distribution of agricultural research have led to suboptimal resource allocations.

The organization of this paper is as follows. Section 2 summarizes some of the pertinent literature covering agricultural productivity and spillovers. Section 3 introduces the data used for the analysis. Section 4 describes two methods that measure crossborder spillovers. Section 5 reports the results, and Section 6 concludes.

2 Background

2.1 Determinants of agricultural output

A vast literature covers the identification and measurement of direct and indirect contributors to agricultural output and productivity. Beginning at the smallest spatial scale, bio-

logical production functions relate output strictly to water, soil, fertilizer, and other inputs that combine with sunlight to produce a crop. These production functions are typically estimated at the plot level. Understanding why outputs might vary across plots, for instance from one farm to another, usually requires additional controls for labor and machinery inputs, cultivar, land gradient, and other factors. With these new variables, analyses can focus on how differences across farms (such as family size, animal or machine power, farm size, farmer education) might explain variations in yield. Zooming out to the largest spatial scale, country-level analyses permit an examination of additional determinants that operate over entire economies and regions, such as market structures, institutions, and technology.

Since this paper is concerned with country borders, the sizeable literature on cross-country differences in agricultural productivity is most relevant. A summary of this literature appears in Mundlak (2001). A prominent example from the early literature is Hayami and Ruttan (1970) who examined cross-country output differences by estimating per capita output's response to labor, land, livestock, fertilizer, machinery, and education. Differences in productivity levels between developed and less developed countries were attributed to levels of capital as well as farmers' education. As new data became available and significant events unfolded (e.g., the Green Revolution), researchers updated their estimates of input elasticities (Nguyen, 1979; Kawagoe et al., 1985). Additional research has examined other potential contributors, including price changes and the inequality of land distribution (Fulginiti and Perrin, 1993; Vollrath, 2007).

As part of this agricultural development literature, economists have also examined the contributions of a country's research and development (R&D) in agricultural science, technology, and extension to economic outcomes. Evenson (2001) conceptualizes these contributions in terms of three so-called gaps, specifically the gap between average observed farm yields and: (1) the yields achieved using best practices and technologies, (2) a farm's full productive potential, currently awaiting discovery via applied research activity, and (3) the underlying "science potential" of a location, achievable through fundamental scientific discoveries. Each gap is reduced through extension services and research in university and government trial plots and labs. To measure the effect of these activities on yields, productivity, and other economic outcomes, researchers have utilized variables describing the availability and intensity of extension services in a location, as well as the level of gov-

ernment expenditures dedicated to R&D. Estimated coefficients are then used to calculate the rate of return to these investments to society at large. In an early study, Griliches (1964) estimated an aggregate production function for the United States, and, in addition to the traditionally-observed inputs, tested a variable representing state-level expenditures on research and extension activities. More recently, Thirtle et al. (2003) used a system of equations to estimate the research elasticity of productivity across a variety of developing countries. From these estimates, the authors then calculated rates of return to agricultural research and their impact on poverty reduction worldwide.

Understanding the determinants of agricultural productivity at different scales, therefore, fixes the next section’s discussion of spillovers in the proper light. How do spillovers occur? What exactly gets shared? And how can they be measured?

2.2 Spillovers and barriers to spillovers

Within the literature on R&D’s effects on agricultural outcomes, one stream has focused on the occurrence and size of technology and knowledge spillovers from one location to another. Evenson (1989) presents a conceptual overview in which spillovers occur in four possible ways: (1) site-specific technological improvements that raise output in all locations with similar geography (typically neighbors), (2) knowledge spillovers from basic science research into technology development and invention, (3) spillovers that spread across multiple crops, and (4) spillovers attributable to the pricing of new technologies below their actual value in order to stimulate sales. Evenson (2001) calls these various beneficial spillovers “spill-ins” and proposes a measure of a location’s “research stock” to be a function of its neighbors’ research stock, weighted by those neighbors’ distances from the original location.

Spillovers are not just the result of centrally-located R&D and technologies diffusing across space. A farmer can observe her neighbors’ activities and knowledge and consequently react. For example, a neighbor’s use of a new seed variety may influence a farmer’s decision to adopt the same variety. Like extension services, neighbors’ activities may also confer indirect benefits. Pest management, hybrid variety conservation, and soil preservation techniques are examples of activities that spill over and potentially benefit neighboring farmers.

In his review of the state of empirical research on the topic, Alston (2002) argues that

spillovers explain a sizeable portion, often more than half, of agricultural productivity growth, and that the pace of agricultural development during the twentieth century owes much to spillovers of international technologies. Ultimately, an understanding of the value and the spatial extent of these spillovers can serve to more efficiently allocate resources dedicated to agricultural research and development. Notably, Alston (2002) lists a variety of studies that incorporate the geographic features of a region as “boundaries” within which spillovers might operate. Indeed, spillovers in agricultural productivity and knowledge, in contrast to other economic activities, are a function of the underlying geography of the neighboring locations.

However, in contrast to the general border effects literature highlighted in the introduction, comparatively little attention has focused on the degree to which R&D, technology, and knowledge spillovers are potentially interrupted and diminished by international borders. In his summary of “research spill-ins” Evenson (2001) also introduces a “spillover barrier index.” To understand this index, imagine two locations 1 and 2 (e.g., neighboring plots or neighboring regions) on which a common set of cultivars are tested. Denote the highest yielding cultivar in location 1 to be C_{i1} , while the highest yielding cultivar in location 2 is C_{j2} . One way to capture the barrier to spillovers is to transfer cultivar C_{j2} into location 1 and observe its yield, C_{j1} . In words, this represents the best cultivar from location 2’s yield performance in location 1. Equation 1 then captures the degree to which spillovers between the two locations are impeded:

$$SPB_{1,2} = 1 - C_{j1}/C_{i1} \tag{1}$$

This spillover barrier (*SPB*) equation can be generalized across all combinations of locations i and j and a variety of cultivars. The *SPB* measures the ease with which a particular cultivar can be grown in other locations. Evenson (2001) suggests that across locations, varying soil quality, climate, slope, or other agroecological characteristics are the primary barriers to beneficial spillovers. We extend this interpretation to include membership in different geopolitical institutions, the consequences of which include varying R&D objectives, strategies, and successes, market structures, and governance qualities. A review of the literature suggests this question has not been asked in agricultural settings, a surprising gap in the literature considering the importance of spillovers.

Empirically, detecting spillovers is accomplished indirectly. One common approach is to observe spatially clustered values on a map and infer that, among the neighbors within that cluster, some process is occurring by which an externality is transmitted. One popular measure of spatial clustering is the global Moran's I statistic which reports the degree to which a variable's values are clustered over an entire space and is interpreted along the lines of a correlation coefficient (Cliff and Ord, 1981). A local Moran's I statistic is an observation-specific measure that describes an observation's association with its neighbors (Anselin, 1995). As part of the Local Indicators of Spatial Association (LISA), a local Moran's I can identify so-called hot spots of like-valued clusters on a map as well as regions of instability, where high and low values both occur.

In this paper, we rely on a slightly different approach. It begins with the assumption that over the long run, spillovers will cause agricultural outputs in a given neighborhood to converge, holding direct inputs, geography, and climate factors constant. To represent neighborhood-level convergence, we calculate the standard deviation of its constituent grid cells' outputs. Low standard deviations imply relatively high convergence and fully operational spillovers, while higher standard deviations are interpreted to reflect little convergence and relatively inhibited spillovers.

3 Data

3.1 Description and sources

This paper uses newly available high-resolution grid cell data covering agricultural production across eleven commodities produced in sub-Saharan Africa. These data, taken from the Harvest Choice website, are generated using an entropy-based approach following You and Wood (2006) that assigns total production values to 5-minute grid cells, approximately 10 square kilometers in area, based on official production statistics at the lowest available administrative level, method of production, land cover data obtained from satellite images, crop-specific land suitability estimates, and population density. To reduce the number of observations to a manageable size, we aggregate the data to the 30-minute (half-degree) level by taking the average of all of its constituent 5-minute cells. Despite this aggregation, the basic patterns of sub-national variation in the data are preserved.

In this analysis, we train our attention on six crops in West Africa: cassava, cotton, maize, rice, millet and sorghum. The study area consists of 2,638 grid cells. Figure 1 presents a map of the study area. Note that not every grid cell is cultivated with each crop, and that some grid cells are not cultivated at all. Cells for which no production of a given crop is observed are excluded from that crop’s subsequent analysis. We define “no production” as output below 1 metric ton per 100 square kilometer area. Excluding “no production” cells is justifiable since it makes little sense to conduct a crop-specific spillover analysis in regions or grid cells where the crop is only grown in, at most, trivial quantities. Summary statistics of all the producing cells are presented in Table 1.



Figure 1: Geographic area of study, represented in white.

In this analysis, we ignore prices. It is possible to have multiplied crop-specific prices with outputs to more closely approximate agricultural productivity. Indeed, the Food and Agriculture Organization of the United Nations (FAO) reports prices for many crops produced in Africa. But unfortunately, they are reported only at the country-level and quite error-ridden.¹ Moreover, these price data, apart from their unreliable provenance, vary dramatically from one country to the next and consequently would introduce large and artificial productivity differences across borders. For these reasons, we rely only on Harvest Choice’s output estimates in the analysis. Measures of productivity, of course, typically account for price in order to make value comparisons across commodities. Since we lack price data, our

¹A description of FAO’s methods for reporting statistical data can be found at <http://faostat.fao.org/>

Table 1: Summary statistics of crop output and neighborhood variability

	Average output in metric tons per 5-minute cell				
	Count	Mean	Std. Dev.	Min.	Max.
Cassava	965	1,280.21	2,829.51	1.00	22,656.56
Cotton	885	75.19	148.89	1.00	2,031.08
Millet	1,050	235.20	435.83	1.00	2,652.90
Rice	1,121	212.15	370.63	1.00	4,814.88
Sorghum	1,151	231.73	354.47	1.00	2,960.97
	Grid cell-specific neighborhood standard deviation of output				
	Count	Mean	Std. Dev.	Min.	Max.
Cassava	965	896.71	1,440.37	0.00	8,240.93
Cotton	885	66.30	101.08	0.80	695.42
Millet	1,050	157.75	210.52	0.00	1,011.62
Rice	1,121	177.09	247.09	0.00	1,728.54
Sorghum	1,151	150.96	170.00	0.00	1,017.87

Note: All values derived from Harvest Choice data sets. Output values are reported in metric tons per 5-minute grid cell unit at the half-degree grid cell level.

estimates only permit same-crop comparisons. Of course, spillovers are not limited to crop-specific knowledge and technologies, for which reason our estimates more likely represent the lower bound of spillovers relative to the overall level.

In addition to these output data, we use grid cell level data on the Strahler order of rivers (Vörösmatry et al., 2000). The Strahler Index ranges from 1 to 6, with six representing the largest river into which all tributaries flow and one representing a small stream with no tributaries. Rivers frequently define the borders between two countries, and accounting for their presence permits effects attributable strictly to artificial borders, not rivers or potentially other features associated with rivers (e.g., varying weather conditions or soils), to be separately detected. To capture this, we arbitrarily indicate all rivers with Strahler Index greater than level three with the value of one. See Data Appendix for a discussion of this index.

Finally, to represent country borders, we calculate the kilometer distance of each grid cell to the nearest country border based on the Africa Equidistant Conical projection using ArcMap Version 9.2.² This does not include the distance to the nearest coast, and as such

²We select an Africa-centric projection to best approximate the arc-distances between points.

represents how close each grid cell is to a neighboring set of national institutions. For the entire map of Africa, the average distance between each grid cell and the nearest country border is approximately 149 kilometers, while the longest distance is 762 kilometers. It is worth mentioning that this variable is calculated with respect to the nearest country border along the *interior* of the study area. Thus, for example, the distances calculated in Chad are with respect to its western border. See the Appendix for more details.

3.2 Visualizing the data

A visual presentation of the data can reveal some potentially interesting features. Two sets of maps appear below. Figure 2 depicts the distribution and level of each crop’s production across the continent. Figure 3 shows the spatial distribution of the crop-specific neighborhood standard deviations across all the grid cells. The first obvious feature is the spatial clusters of crop production.

Judging from Figure 2, cassava production appears concentrated in the coastal stretch between Ghana and Nigeria. Cotton is grown more inland, with regional concentrations at the intersection of Ivory Coast, Burkina Faso, and Mali. Millet production is clustered in Senegal as well as northern Nigeria and its immediate neighbors. Sorghum appears primarily concentrated in Nigeria as well as Burkina Faso. Rice production, in contrast to the other crops, is distributed relatively evenly across the region.

Figure 3 presents a map of the crop-specific standard deviation of output centered on each grid cell and its 8 nearest neighbors. As the map reveals, large standard deviations occur where the levels of output fluctuate dramatically within a neighborhood. As described above, spillover-driven output convergence, as manifested in the observed neighborhood standard deviation, can occur through a variety of channels, common agroecological conditions being chief among them. Spillovers also occur via shared institutional features commonly contained within a country’s borders.

The maps do not generally reveal whether institutional differences, as reflected by a border’s presence, impede the overall transmission of spillovers, but Figure 4 may offer a clearer picture. The graphs present crop-specific local polynomial regressions of each grid cell’s neighborhood standard deviations on its distance to the nearest country border.³ The

³Local polynomial regressions are illustrative non-parametric estimates of the relationship between mul-

graphs exclude all grid cells for which no production is observed except for cells that belong to another producing cell's neighborhood. Millet, sorghum, and cotton exhibit a generally increasing standard deviation as the distance to the nearest border approaches zero. Cassava and rice behave less clearly. Of course, the absence of additional controls leaves considerable variability in the relationship unexplained, something which will be addressed in the later sections, but broadly speaking, the impact of a country's border on the similarity of outputs across neighboring grid cells appears negative, tentatively suggesting that borders may play a role in the inhibition of beneficial cross-country agricultural spillovers.

4 Model estimation

This section presents the two approaches we use to capture border effects on agricultural spillovers. The first approach relates to the convergence concept discussed earlier. The convergence approach tests whether the output variation of a grid cell's neighborhood is affected by its proximity to a country border. We define a grid cell's neighborhood to be the nearest 8 grid cells that surround it. For each grid cell in the data set, we calculate its unique neighborhood-specific standard deviation of output. This serves as the dependent variable in our analysis. Table 1 reports each grid cell's neighborhood standard deviation of output by crop. The second approach asks whether a location's output is explained more by within-country neighbors or cross-country neighbors using a higher-order spatial lag model.

At first glance, these approaches might appear to be asking the same question. The difference, in fact, is subtle. In the convergence approach, we are testing whether a neighborhood's variation in output is at all influenced by its proximity to a country border. No assumption is made regarding the channel or origin of spillovers. In contrast, the spatial lag approach tests whether neighbors within the same country exert a greater influence on a location's outcome more than neighbors that happen to be located in the country next door. It imposes a structure on spillover transmissions by limiting them to contiguous grid cells.

tiple variables. The estimated curve derives from a local average taken around each observation. The precise bandwidth to calculate the local average is endogenously selected using a rule-of-thumb technique.

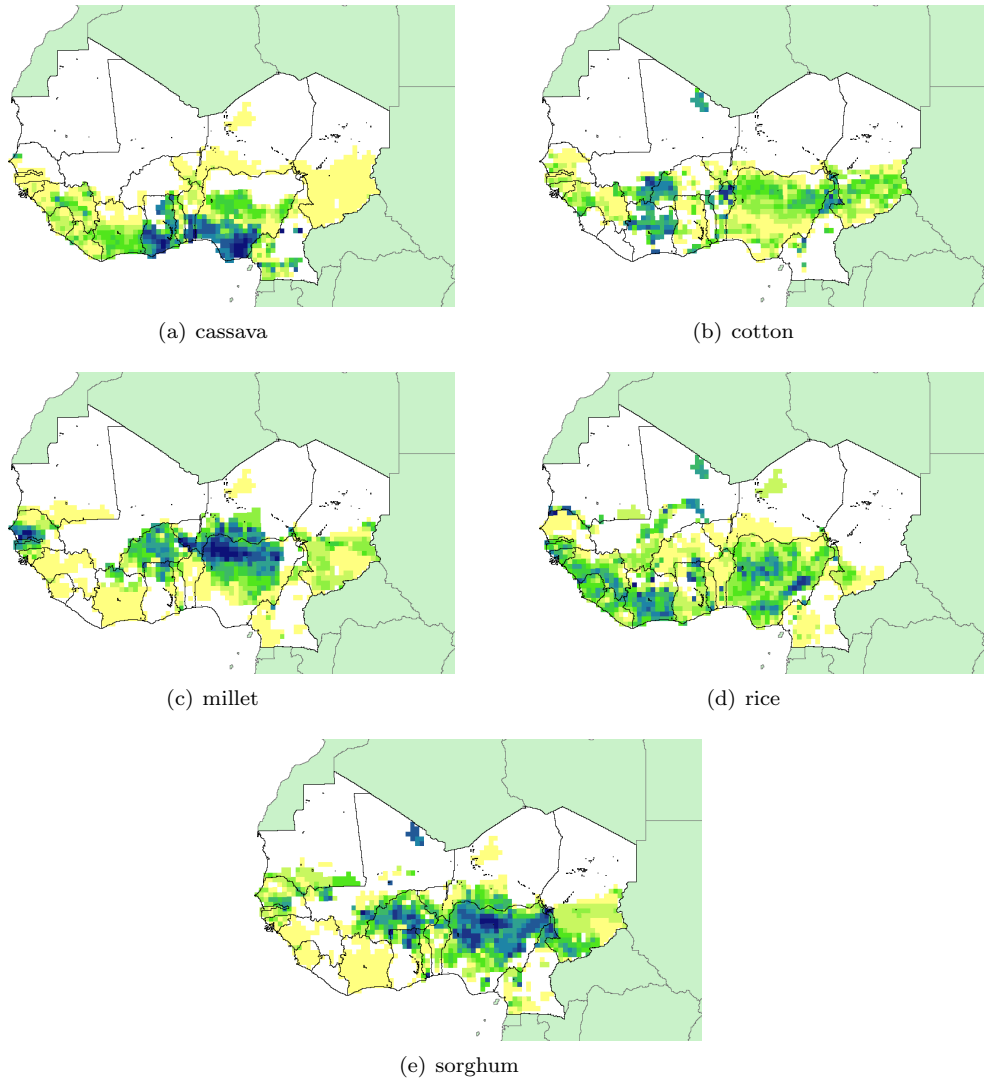


Figure 2: Distribution of crop outputs at the grid cell level. Note: Blue represents the highest output level, yellow the lowest, and white is no production.

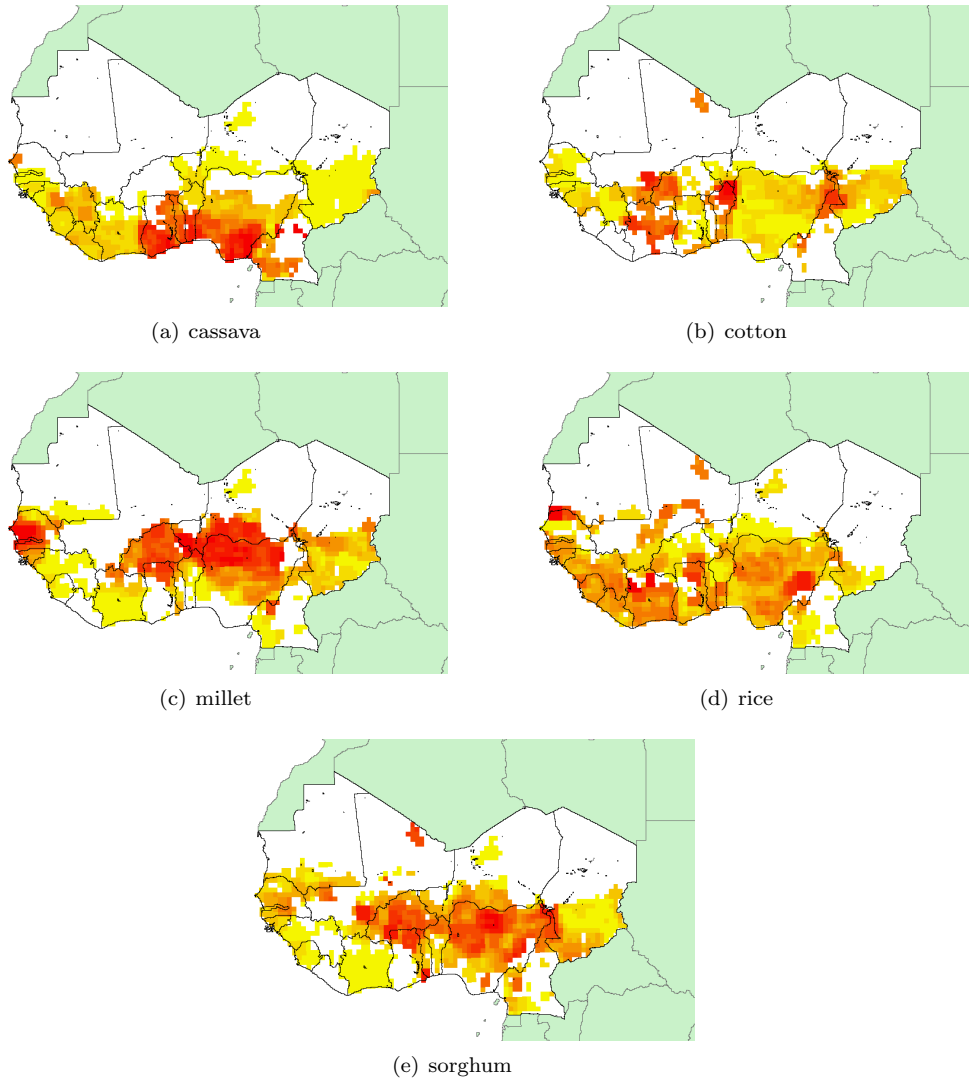
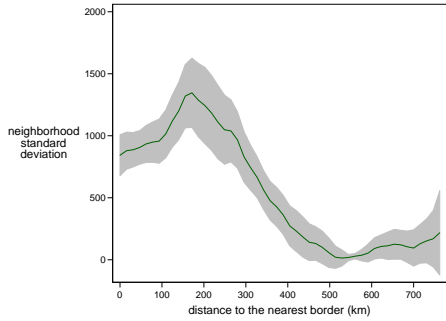
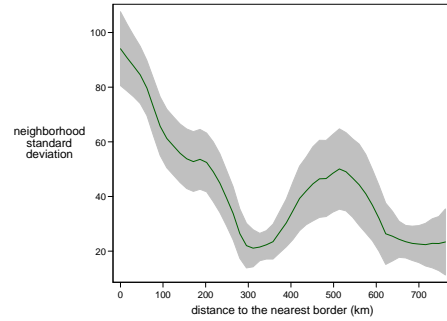


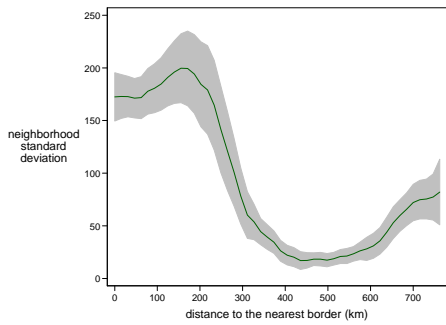
Figure 3: Crop-specific grid cell neighborhood standard deviation of output. Note: Red represents the highest standard deviation values, yellow represents zero.



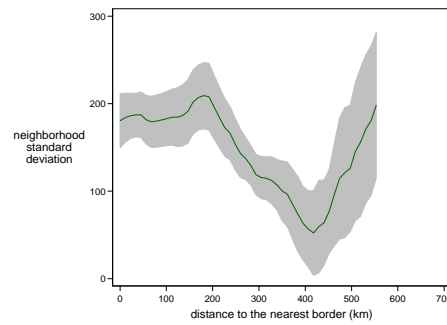
(a) cassava



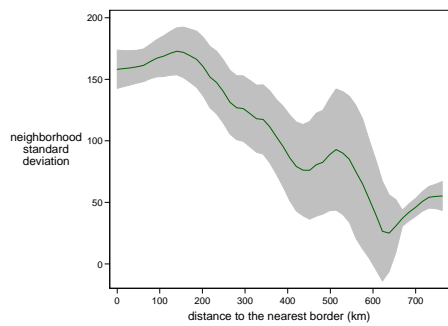
(b) cotton



(c) millet



(d) rice



(e) sorghum

Figure 4: Local polynomial regressions of neighborhood standard deviations of output on distance to nearest border. Note: Shaded area represents 95% confidence intervals. Bandwidths selected by rule-of-thumb using Stata's `lpolyci` command.

4.1 Testing for convergence as a function of borders

In addition to a grid cell’s proximity to a border, we include additional controls. These include a measure of the similarity of agroecological conditions in each grid cell’s neighborhood. Neighborhoods with dissimilar conditions are more likely to exhibit dissimilar outputs. We use two agroecological variables to control for these differences, the cultivation suitability index due to Ramankutty et al. (2002) and land elevation. Using these variables, we calculate standard deviations of cultivation suitability and elevation to represent the similarity of agroecological conditions in that cell’s neighborhood. In addition, we include a river variable, taken from Vörösmatry et al. (2000), to control for possible natural barriers to spillovers, particularly in the case where a country border coincides with a river.

The following specification operationalizes the hypothesis test:

$$SD_{ij} = \beta_0 + \beta_1 distance_{ij} + \beta_2 suitability_{ij} + \beta_3 elevation_{ij} + \beta_3 population_{ij} + \beta_4 river_{ij} + \delta_i + \epsilon_{ij}, \quad (2)$$

where SD is the crop-specific grid cell’s neighborhood standard deviation of output for a grid cell i in country j . The variable $distance$ is the distance in kilometers of the grid cell to the nearest country border. The grid-cell level standard deviations for the agroecological variables are $suitability$ for the cultivation suitability and $elevation$ for the land elevation in meters. The standard deviation of each grid cell neighborhood’s rural population is captured in $population$. The variable $river$ describes the order of a river that passes through a cell. Arguably, higher order rivers are more likely to represent boundaries of a particular agroclimatic zone, and thus, cause greater variation in crop outcomes within a cell’s neighborhood. Finally, since each country is believed to exert a unique influence on the similarity of outputs within its borders—owing to its institutions, R&D activities, potential data reporting biases, and other unobservable and unknown factors—the dummy variable δ_j is included to capture country fixed effects. The hypothesis of interest is that the coefficient on the variable $distance$ is positive and significant. That is, reducing the distance to the border raises the standard deviation of output for that grid cell’s neighborhood, implying that spatial spillovers are being impeded by country borders.

Since the standard deviations of neighborhood output are liable to be correlated over

space, we run the above specification using a spatial error model. We rely on the spatial error model since its interpretation fits more comfortably with the variable in question. That is, the spatial correlation of the variation in neighborhood outputs is more likely attributable to unobservable processes captured in the error term as opposed to the spatial lag term. The spatial lag model’s interpretation, that a grid cell’s neighborhood variation in output is explained by its neighbors’ neighborhood variation in output, is awkward and not intuitive given the dependent variable in question.

4.2 Measuring spillovers in explicitly

Another way to capture spillovers is through a spatially explicit econometric specification in which a grid cell’s agricultural output is modeled as a function of its neighbors’ output. This is often achieved by adding an additional right hand side term, the spatial lag of the dependent variable.

$$Y_i = \beta_0 + \rho WY + X_i\beta + \epsilon_i. \tag{3}$$

Commonly labeled the “spatial lag” model, this specification tests whether the effects of a location’s neighbors exert a significant influence on the location’s outcome. The term W represents a spatial weights matrix, either distance-based or contiguity-based, that, upon multiplication with neighboring values of Y , essentially summarizes the effects of neighboring outcomes. A positive and significant coefficient ρ can be interpreted to suggest that neighboring observations exert a positive effect on the outcome of interest.

Rietveld and Wintershoven (1998), in an analysis of infrastructure spillovers in Europe, extend the spatial lag model to account for two types of neighbors, within-country neighbors and cross-country neighbors. The objective was to test whether spillovers within countries differed from spillovers across countries. Their proposed specification introduces two spatial weights matrices W_1 and W_2 corresponding to each type of neighbor. Each weights matrix thus has a unique effect on the dependent variable of interest. The likelihood function, the scores, and other supporting details of the methodology appear in Rietveld and Wintershoven (1998).⁴ For this analysis to work, of course, requires data at a sub-national level

⁴A similar approach appears in Lacombe (2004) who examines county-level determinants of food stamp payments including payments in neighboring counties, both within state and across states. The author

in which certain regions lie along international borders while other regions are entirely “domestic.” Using the subnational grid cell data described above, we follow the approach of Rietveld and Wintershoven (1998) to ask the same question: Are spillovers within countries different from spillovers that cross international borders, and by how much?

The following double-log specification models output as a function of the familiar geographic determinants, population, country dummies, and two additional regressors: the outputs of within-country neighbors and the outputs of cross-country neighbors.

$$\begin{aligned} \ln Y_{ij} = & \beta_0 + \beta_1 \ln cultiv_{ij} + \beta_2 \ln elev_{ij} + \beta_3 distance_{ij} + \beta_4 river_{ij} + \beta_5 frost_{ij} \\ & + \beta_6 \ln pop + \rho_1 W_1 \ln Y + \rho_2 W_2 \ln Y + \delta_j + \epsilon_{ij}, \end{aligned} \quad (4)$$

where Y is a grid cell i 's output in country j , $cultiv$ is a location's cultivation suitability index due to Ramankutty et al. (2002), $frost$ is the occurrence of seasonal winter frost following Masters and McMillan (2001), $elev$ is a measure of a grid cell's elevation, $distance$ is the cell's distance to the coast, a proxy for trade and market access, $river$ is the familiar Strahler Index value for rivers, and δ is a country dummy that absorbs all remaining effects relevant at the country-level, including but not limited to institutional quality, human capital levels, agricultural research, and other inputs. W_1 is a queen-criterion weights matrix composed of the neighbors of grid cells that lie strictly within a country's borders. W_2 , in contrast, is a weights matrix composed of the neighbors of grid cells that lie adjacent to or on top of a country's borders. The parameters of particular interest are ρ_1 and ρ_2 , the effects of a grid cell's neighbors on its own productivity.

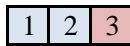


Figure 5: Within- and cross-country neighbors

What exactly are the elements of W_1 and W_2 ? Figure 5 represents a simple scenario of three grid cells divided across two countries. Cells 1 and 2 belong to the first country, and cell 3 is its own country. W_1 , the matrix that contains all domestic neighbors, is:

similarly employs two weights matrices to capture within state effects and cross state effects, but each weights matrix row sums to one because the distance criteria is not contiguity, but k -nearest neighbors, and as a consequence every observation has k -nearest in-state neighbors as well as k -nearest cross-state neighbors.

$$W_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Each (i, i) -element corresponds to grid cell i , and each (i, j) element represents a neighbor relation. The value 1 represents a contiguous connection, i.e. a shared border, and 0 represents no connection. Thus, the 1 in element positions (2,1) and (1,2) reflects the contiguity shared between grid cells 1 and 2. Note that the contiguity between cells 2 and 3 is coded with a 0 as well. This is because cells 2 and 3, while they are contiguous, are not domestic neighbors. To capture the international border contiguity, a 1 appears in matrix W_2 for element positions (3,2) and (2,3).

$$W_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

These spatial weights matrices, summed together, account for all contiguous neighbors, both domestic and international, in the data set. We follow Rietveld and Wintershoven (1998) and standardize the rows of $W_1 + W_2$ to sum to one. This permits the coefficients for each of the weighted neighbors to be directly comparable, since the number of border-sharing neighbors is smaller than the number domestic neighbors. Specifically, we test the hypothesis that neighbors' contribution to productivity among grid cells within a country is greater than the neighbors' contribution to productivity among grid cells that span country borders.

Rietveld and Wintershoven (1998) estimate this model using maximum likelihood. In our case, since the sample size is too large, we rely on an instrumental variables approach. A discussion of this method appears in Anselin (1988).

5 Results and discussion

Tables 2 present the results from the convergence approach. Distance from the nearest border is reported as a continuous variable, and the hypothesized sign on its coefficient is

negative. That is, as a cell's distance to a border decreases, its neighborhood's variation in output, as reflected in its standard deviation, is expected to rise. In all five crops examined, the sign on the variable *distance* is negative, four of which are significantly different from zero at the 10 percent level or higher. For example, in the case of cassava, the coefficient on border distance is -1.42 which implies that a 100 kilometer distance reduction from the border is associated with a 142 unit increase in the standard deviation. For cassava, this translates to a nearly 16% increase over the sample's average standard deviation of outputs. Table 3 summarizes the effect of a 100-kilometer reduction in border distance on each crop, relative to each crop's average level of neighborhood convergence.

Table 2: Output variability and distance to the nearest country border

	Cassava	Cotton	Millet	Rice	Sorghum
intercept	1383.26*** (365.03)	87.85*** (27.65)	135.02*** (55.55)	272.36*** (67.01)	154.78*** (39.63)
border	-1.42** (0.71)	-0.11* (0.06)	-0.15 (0.09)	-0.26* (0.15)	-0.15* (0.09)
$\sigma_{suitability}$	1052.08* (544.18)	40.33 (46.42)	89.35 (75.40)	1.03 (1.25)	-13.44 (68.84)
$\sigma_{elevation}$	-0.66 (0.59)	-0.02 (0.05)	-0.12 (0.09)	-1.35 (1.53)	-0.08 (0.08)
$\sigma_{ruralpop}$	1.69*** (0.49)	-0.03 (0.04)	0.13* (0.06)	-0.11 (0.11)	0.12* (0.06)
river	-7.54 (46.91)	3.51 (4.17)	0.95 (6.80)	-5.33 (10.45)	4.77 (6.12)
λ	0.95 (0.01)	0.92 (0.01)	0.95 (0.01)	0.93 (0.01)	0.92 (0.01)
<i>n</i>	965	885	1050	1121	1151

Note: Dependent variable is the standard deviation of output. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Country-level effects are included but not reported.

Table 3: Impact of 100 kilometer reduction of border distance

	Cassava	Cotton	Millet	Rice	Sorghum
distance effect magnitude	1.42	0.11	0.15	0.26	0.15
average convergence	896.71	66.30	157.75	177.09	150.96
reduction in convergence	15.8%	16.6%	9.5%	14.7%	9.9%

Note: Average convergence values are the average neighborhood standard deviations reported in Table 1. The percent reduction in convergence is calculated with respect to the average level of convergence.

In words, closer proximity to a border raises the variation of a crop’s output within a grid cell’s 3x3 neighborhood. This implies that any unobserved processes that drive the outputs of a neighborhood’s producers to converge are interrupted when a border intersects that neighborhood, controlling for variation in cultivation suitability and elevation, as well as the host of unobservable institutional and technological characteristics operating at the country-level. In this context, it is worth noting that this process does not explicitly capture cell-to-cell interactions, e.g. farmers in one cell sharing knowledge with farmers in the neighboring cell. Rather, the spillovers are understood to operate in the neighborhood without respect to their origin, direction, or destination. This might include regional extension services, common market institutions, or physical infrastructure.

However, cell-to-cell interactions are captured in the direct approach, the results of which are presented in Table 4. The specification tests the response of a grid cell to the outcomes of two types of neighbors: (1) within-country and (2) cross-country. Table 5 reports the test results of the hypothesis that domestic spillovers are greater than cross-country spillovers, or $\rho_1 - \rho_2 > 0$.⁵ None of the t -statistics reported in Table 5 are large enough to reject the null hypothesis at the 90% confidence level. From this, we tentatively conclude that having neighbors that fall in another country has no significantly different effect on a grid cell’s own output levels than neighbors that fall within the same country. In other words, the results suggest that borders do not impede the neighbor-to-neighbor transmission of output-specific spillovers. Again this result controls for differences in rural population (which proxies for agricultural labor), and several significant agroecological features, as well as country-level effects.

Despite these results, a direct comparison of ρ_1 and ρ_2 may not in fact be appropriate for two reasons. First, spatial lag models imply that the marginal effects of a shock vary with the precise location and size of that shock. Thus, it is possible that in certain locations on the map, a statistically significant difference between ρ_1 and ρ_2 indeed exists. Second, the spatial multiplier for each coefficient’s effect contains interactions between the two coefficients.

⁵Specifically we test the null hypothesis that the difference between the two coefficients equals zero, that is $\beta_1 - \beta_2 = 0$ (Greene, 2008). First calculate the standard error (se) of $\hat{q} = \beta_1 + \beta_2$.

$$se(\hat{q}) = (se(\beta_1)^2 + se(\beta_2)^2 + 2cov(\beta_1\beta_2))^{1/2}$$

$$t = \frac{\beta_1 - \beta_2}{se(\hat{q})}$$

A one-way t -test will reveal whether support for the alternate hypothesis $\rho_1 - \rho_2 > 0$ exists.

Table 4: Output and within- and cross-country spillovers

	Cassava	Cotton	Millet	Rice	Sorghum
W_1y	0.229** (0.107)	0.868*** (0.115)	0.719*** (0.083)	0.418*** (0.094)	0.659*** (0.079)
W_2y	0.268*** (0.099)	0.858*** (0.108)	0.714*** (0.083)	0.437*** (0.089)	0.648*** (0.082)
Intercept	5.416*** (0.758)	0.496 (0.467)	-0.018 (0.296)	1.456 (0.370)	0.664** (0.347)
suitability	-0.054 (0.231)	0.430** (0.188)	0.097 (0.164)	0.457** (0.208)	0.406** (0.179)
elevation	0.060** (0.027)	-0.031 (0.023)	-0.005 (0.022)	0.068*** (0.024)	0.006 (0.023)
distance	-0.270*** (0.036)	0.039** (0.020)	0.049** (0.021)	-0.016 (0.017)	0.047*** (0.018)
river	0.027 (0.131)	-0.109 (0.102)	0.366*** (0.099)	0.489*** (0.105)	0.120 (0.095)
frost	-0.011 (0.182)	0.055 (0.149)	-0.198 (0.137)	0.224 (0.144)	-0.154 (0.127)
population (10,000s)	0.015*** (0.004)	0.006** (0.003)	0.007** (0.003)	0.006** (0.003)	0.008** (0.003)
n	965	885	1050	1121	1151

Note: Dependent variable is the log of output. W_1y is the within-country spatial lag of the dependent variable, and W_2y is the cross-country spatial lag of the dependent variable. Country-level effects are included but not reported. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5: Differences between domestic and cross-country spillovers

	Cassava	Cotton	Millet	Rice	Sorghum
ρ_1	0.229** (0.107)	0.868*** (0.115)	0.719*** (0.083)	0.418*** (0.094)	0.659*** (0.079)
ρ_2	0.268*** (0.099)	0.858*** (0.108)	0.714*** (0.083)	0.437*** (0.089)	0.648*** (0.082)
$cov(\rho_1\rho_2)$	0.0103	0.0122	0.0067	0.0081	0.0063
t -statistic	-0.190	0.045	0.030	-0.105	0.069

Note: Coefficient values and standard errors taken from Table 4. The one-way t -statistic value at the 90% confidence level is 1.65.

Specifically, the spatial multiplier,

$$(I - \rho_1 W_1 - \rho_2 W_2)^{-1} \tag{5}$$

can be expanded to a second-order approximation to obtain the following:

$$I + \rho_1 W_1 + \rho_2 W_2 + \rho_1^2 W_1^2 + 2\rho_1 W_1 \rho_2 W_2 + \rho_2^2 W_2^2. \tag{6}$$

The fifth term in Equation 6 shows that the two individual effects interact to produce an additional effect which is not accounted for in the statistical tests reported in Table 5. This interaction effect can be observed if we (1) shock a variable in a domestic grid cell and observe effects on cross-country cells, and (2) shock a variable in a cross-country cell and observe its effects on domestic cells.

To illustrate this effect, we provide an example showing a variable's effect across weights matrices under two scenarios. In the first scenario, we shock rural population's value by 10% in only the domestic cells. In the second scenario, we shock rural population's value by 10% in only the cross-country cells. Table 6 compares how the direct, indirect, and total impacts of each shock are detectable across both categories of grid cells. Thus, under the "Domestic shock" column, we note that the impacts on the cross-country cells differ from the impacts under the "Baseline" regression, implying that the effects of a domestic shock are carried through onto the cross-country cells. For this reason, therefore, any conclusion derived from the statistical tests above must be treated with caution.

Table 6: Domestic and cross-country impacts of a shock

	Baseline		Domestic shock		Cross-country shock	
	domestic cells	cross-country cells	domestic cells	cross-country cells	domestic cells	cross-country cells
direct impacts	4.56	3.38	4.45	3.31	4.76	3.50
indirect impacts	17.70	12.85	17.36	12.78	18.38	13.09
total impact	22.27	16.23	21.81	16.09	23.14	16.59

Note: Shock was a 10% increase in the rural population in 2000.

6 Conclusion

In this essay, high-resolution agricultural output data from Sub-Saharan Africa are exploited to test the hypothesis that beneficial spillovers among and between neighboring producers are mitigated by the presence of a country border. Building on the assumption that unobservable spillovers in a region will push outputs to converge *ceteris paribus*, we test the hypothesis that country borders that intersect a region will raise that cell's neighborhood output variation. From the results, it appears country borders indeed inhibit such spillovers from occurring, even after controlling for variations in agroecology and country-level institutional factors. However, the hypothesis that borders interrupt the cell-to-cell transmission of spillovers is not substantiated by the results as they are currently interpreted. That is, the influence of one cell's output on its neighbors appears not to hinge on whether a border happens to cross between them.

Refinements in this analysis can be imagined. Currently, the definition of a grid cell's membership in a particular country depend on which country its majority area falls inside. Decomposing the areal units of each grid cell into their constituent country-specific parts would not only offer greater precision in the estimates, but also shine a closer light on the spatial scale of the spillovers themselves. Currently, the analysis is premised on the assumption that the relevant distance over which spillovers operate is approximately 55 kilometers, or the width of one grid cell. But spillovers, particularly neighbor-to-neighbor spillovers posited in the direct estimation approach, may occur at a smaller scale. Indeed, the original Harvest Choice are reported 5-minute grid cells. This implies a distance far more amenable to the transmission of knowledge spillovers between producers.

Nevertheless, the analysis reveals that country borders play an inarguably significant role in explaining the outcomes of agricultural producers via the channel of spillovers. Policy makers, increasingly aware of the relevance of their interventions' geographic scope, can interpret these results to suggest that country borders potentially diminish the impact of their interventions, and as such, additional measures might be needed to counteract their effect. Alternatively, efforts towards minimizing the differences in institutions, infrastructures, and markets on opposite sides of borders can also serve to more fully capture the benefits of spillovers that are currently lost.

References

- Alston, J. (2002). Spillovers. *The Australian Journal of Agricultural and Resource Economics* 46(3), 315–346.
- Anderson, J. E. and E. v. Wincoop (2003). Gravity with gravitas: A solution to the border puzzle. *The American Economic Review* 93(1), 170–192.
- Anselin, L. (1988). *Spatial econometrics: methods and models*. Springer.
- Anselin, L. (1995). Local indicators of spatial association-LISA. *Geographical analysis* 27(2), 93–115.
- Cliff, A. and J. Ord (1981). *Spatial processes: models & applications*. Pion Ltd.
- Conley, T. and E. Ligon (2002). Economic distance and cross-country spillovers. *Journal of Economic Growth* 7(2), 157–187.
- Engel, C. and J. H. Rogers (1996). How wide is the border? *The American Economic Review* 86(5), 1112–1125.
- Evans, C. L. (2003). The economic significance of national border effects. *The American Economic Review* 93(4), 1291–1312.
- Evenson, R. (2001). Economic impacts of agricultural research and extension. In B. L. Gardner and G. C. Rausser (Eds.), *Handbook of Agricultural Economics*, Volume 1 of *Handbook of Agricultural Economics*, Chapter 1, pp. 573–628. Elsevier.
- Evenson, R. E. (1989). Spillover benefits of agricultural research: Evidence from u.s. experience. *American Journal of Agricultural Economics* 71(2), 447–452.
- Fulginiti, L. E. and R. K. Perrin (1993). Prices and productivity in agriculture. *The Review of Economics and Statistics* 75(3), 471–482.
- Greene, W. (2008). *Econometric analysis (6th international edition)*. Pearson Education, New Jersey (US).
- Griliches, Z. (1964). Research expenditures, education, and the aggregate agricultural production function. *The American Economic Review* 54(6), 961–974.
- Hayami, Y. and V. W. Ruttan (1970). Agricultural productivity differences among countries. *The American Economic Review* 60(5), 895–911.
- Kawagoe, T., Y. Hayami, and V. W. Ruttan (1985). The intercountry agricultural production function and productivity differences among countries. *Journal of Development Economics* 19(1-2), 113 – 132.
- Lacombe, D. (2004). Does Econometric Methodology Matter? an Analysis of Public Policy Using Spatial Econometric Techniques. *Geographical Analysis* 36(2), 105–119.
- Masters, W. A. and M. S. McMillan (2001). Climate and scale in economic growth. *Journal of Economic Growth* 6(3), 167–186.
- Mundlak, Y. (2001). Production and supply. In B. L. Gardner and G. C. Rausser (Eds.), *Handbook of Agricultural Economics*, Volume 1 of *Handbook of Agricultural Economics*, Chapter 1, pp. 3–85. Elsevier.
- Nguyen, D. n. (1979). On agricultural productivity differences among countries. *American Journal of Agricultural Economics* 61(3), 565–570.

- Ramankutty, N., J. A. Foley, J. Norman, and K. McSweeney (2002). The global distribution of cultivable lands: current patterns and sensitivity to possible climate change. *Global Ecology and Biogeography* 11(5), 377–392.
- Rietveld, P. and L. Janssen (1990). Telephone calls and communication barriers. *The Annals of Regional Science* 24(4), 307–318.
- Rietveld, P. and P. Wintershoven (1998). Border Effects and Spatial Autocorrelation in the Supply of Network Infrastructure. *Papers in Regional Science* 77(3), 265–276.
- Thirtle, C., L. Lin, and J. Piesse (2003). The impact of research-led agricultural productivity growth on poverty reduction in Africa, Asia and Latin America. *World Development* 31(12), 1959–1975.
- Vollrath, D. (2007, 02). Land distribution and international agricultural productivity. *American Journal of Agricultural Economics* 89(1), 202–216.
- Vörösmatry, C. J., B. M. Fekete, M. Meybeck, and R. Lammers (2000). Global system of rivers: Its role in organizing continental land mass and defining land-to-ocean linkages. *Global Biogeochemical Cycles* 14, 599–622.
- You, L. and S. Wood (2006). An entropy approach to spatial disaggregation of agricultural production. *Agricultural Systems* 90(1-3), 329–347.