The Landscape of Farming: An Exploration of Spatial Bio-Economic Characterization Approaches

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

ALISON KAY BITTINGER JOGLEKAR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Adviser: Philip G. Pardey
Co-Advisers: Terrance M. Hurley and Jason M. Beddow

November 2015

Acknowledgements

I would like to extend my heartfelt appreciation to my advisor Dr. Phil Pardey for his mentorship and guidance throughout my graduate education. Dr. Pardey's passion about his work with InSTePP and the HarvestChoice project is infectious. It has been a joy to work with him, and I truly believe I am a better researcher because of his encouragement. I would also like to thank my co-advisers Dr. Terry Hurley and Dr. Jason Beddow. I was once told that while Dr. Hurley may be a quiet man, if he chose to speak, it is well worth your time to listen; this observation could not be truer. Throughout the development of my dissertation, Dr. Hurley helped flush out the technical details of my work and gave invaluable, constructive criticism. Dr. Beddow's input was also invaluable. His knowledge of spatial analysis, his coding prowess and his constant reminders to keep my research questions focused and accessible were vital to actually finishing this dissertation. I feel quite fortunate to have such a complementary team of advisers. Additionally, the input and critique from my other committee member, Dr. Jordan Chamberlin, was greatly appreciated and quite helpful in pulling together the final drafts of this dissertation.

I could not have completed this paper without the help and support of Ulrike Wood-Sichra and Zhe "Joe" Guo at IFPRI who were the lead developers for two of the datasets used in this dissertation. Thank you both for the long hours spent brainstorming, writing, editing and commiserating. I am especially grateful to Ulrike for the time she spent running the robustness tests used in Chapter 2 of this dissertation. Her efforts were truly instrumental in the formation of this particular chapter.

I would also like to thank Brooke Krause for her unwavering support and friendship throughout our entire PhD program, especially during the dissertation phase, and Michelle Hallaway, Xudong Rao, Senait Senay and the rest of my colleagues at InSTePP for their much appreciated roles as sounding boards for my work.

I would like to acknowledge the Bill and Melinda Gates Foundation and Dr. William Hueg who funded my doctorate education through the HarvestChoice project and the Hueg-Harrison Fellowship.

Finally, heartfelt thanks go to my friends and family, especially my husband, whose unconditional love, encouragement and support kept me sane throughout this entire process.

Dedication

For Nikhil and Surly.

Abstract

This dissertation consists of three inter-related but standalone papers focused on the theme of measuring the spatial, bio-economic attributes of production agriculture. Hitherto most of the agricultural development literature dealing with production agriculture has relied on data delineated in geopolitical (i.e., administrative district) boundaries of varying spatial resolutions, with some (increasingly of late) data reported for farm households. In some cases the household level data are geo-referenced, but in a majority of the studies the data are essentially aspatial. Many of the realities facing farmers however, including the agro-ecological (climate, soil, terrain and so on) attributes with which farmers have to work and their proximity to markets, are intrinsically spatial. Thus the location of farms and their physical and economic access to markets have a whole host of agricultural production and consumption implications that profoundly affect the economic circumstances of farm families. Spatially delineated data to facilitate analysis of the effects of location and its associated attributes on farm economies is still limited, but beginning to grow. This dissertation casts a critical eye over the nature and empirical plausibility of some key, spatially explicit datasets, including efforts to form spatially granular estimates of the location of crop production, area and yield worldwide; estimates of the proximity of African crop production to markets of varying sizes; and finally, the retail-level prices of key inputs (specifically fertilizer) faced by farmers throughout Tanzania.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

This dissertation consists of three inter-related but standalone papers focused on the theme of measuring the spatial, bio-economic attributes of production agriculture. Many of the realities facing farmers are intrinsically spatial. At the risk of oversimplification, agricultural producers face three major spatial environments: the agro-ecological conditions in which they operate, relative remoteness, and the socio-demographics of the surrounding community (Staal, Baltenweck, et al. 2002, Dixon, Gulliver and Gibbon 2001). Their agro-ecological circumstances in terms of climate, soil, terrain and related attributes dictate the available natural resources, affect farm productivity, and influence the state of communication and transportation infrastructure. The remoteness of the production system influences prices, market participation and the quantity and quality of agricultural extension services available (especially in developing countries), all of which ultimately influence the producer's choice of enterprise and input use. The demographics of the surrounding community dictate the relevant social constraints. Social institutions and personal circumstances condition preferences and influence how knowledge is transmitted, how common property is used, and the farmer's terms of trade. All three of these spatial environments have a direct effect on the producer's profit function, be it through prices, information attainment, or the functional form of the production function.

However, most of the agricultural development literature dealing with production agriculture has relied on data delineated in geopolitical (i.e., administrative district) boundaries of varying spatial resolutions, with some (increasingly of late) data reported for farm households. In some cases the household level data are geo-referenced, but in a majority of the studies the data are essentially aspatial. Understanding the spatial relationships that affect agricultural production is key to understanding the varied incentive structures that influence production. Ignoring the spatial heterogeneity of production could have serious implications on the applicability of programs and policies established to increase productivity and improve the welfare of agricultural producers, especially when these interventions are “one-size-fits-all” programs like many of the input subsidy programs presently in effect throughout sub-Saharan Africa (Jayne and Rashid 2013).

The use of spatially delineated data to facilitate analysis of the effects of location and its associated attributes on farm economies is still limited, but is beginning to grow. This dissertation casts a critical eye over the nature and empirical plausibility of some key, spatially explicit datasets. These include efforts to form spatially granular estimates of the location of crop production, area and yield worldwide, estimates of the proximity of African farms to markets of varying sizes, and finally, the retail prices of key inputs (specifically fertilizer) faced by farmers throughout Tanzania.

Chapter 2: The Impact of Methodological Choices on Estimates of Global Cropping Systems: An Analysis of the Spatial Production Allocation Model[1]

2.1 Introduction

By 2050, world population is estimated to reach 9.7 billion people (United Nations 2015). The increased food, fiber and fuel demand from this population will significantly affect land and resource use, climate change, the nature and prevalence of poverty, political agendas and technological development. Debate over how to best support the food demands of an increased population and temper the pressure on already scarce resources includes recommendations to close yield gaps, increase production capacities, reduce food wastage and change dietary preferences (Godfray, et al. 2010, Foley, et al. 2011, Beddow, Hurley and Pardey 2014). Regardless of the solutions sought, a clear and reliable understanding of the current spatial distribution of cropping systems is necessary to effectively implement, and study the prospective effects of, such recommendations.

Location is a particularly important factor in the bio-economic performance of agricultural production. Crop choice and yields are tied to the fertility of the soil and the unpredictable nature of weather patterns, as well as the inputs available (e.g., fertilizers, improved seed or mechanization) and the nature of output marketing opportunities. However, the diversity and spatial patterns that result from differences in agro-ecological conditions and market structures may not be adequately captured when only using national or sub-national statistics on agricultural production, such as those reported by the Food and Agricultural Organization (FAO) or household agricultural surveys. In response, pixelated representations of cropping systems, such as M3-Crops (Monfreda, Ramankutty and Foley 2008), the Monthly Irrigated and Rain-fed Cropping Areas (MIRCA) (Portmann, Siebert and Döll 2010), the Global Agro-Ecological Zones (GAEZ) (Fischer, et al. 2013) and the Spatial Production Allocation Model (SPAM) (You, Wood, et al. 2014, Wood-Sichra,

[1] This chapter is a joint effort between Ulrike Wood-Sichra and myself. While the text is my own, Ms. Wood-Sichra contributions were integral in drafting the concepts in the chapter and running the SPAM comparisons.

Joglekar and You forthcoming), are used to quantify the nature and impacts of agriculture on a more spatially granular level.

While all four global cropping system models mentioned above were created with similar objectives in mind, in a comparative study by Anderson, et al. (2015) found non-trivial differences in the year 2000 estimates of crop area and yields between M3-Crops, MIRCA, GAEZ and SPAM, despite the overlap in data inputs between the four models. Anderson, et al. (2015) primarily attributed these inconsistencies to differences in downscaling methodologies and the sub-national data used. The global 5 arc-minute (approximately 10 kilometers at the equator) M3-Crops data (Monfreda, Ramankutty and Foley 2008) contain estimates of harvested area and yield in the year 2000 for 175 crops, trees, forage and grassland. In M3-Crops, sub-national crop production statistics were rasterized using distributions proportional to the amount of cropland within a pixel relative to the total cropland in an administrative unit. MIRCA (Portmann, Siebert and Döll 2010) further disaggregated the M3-Crop estimates of harvested area for 26 crops and crop aggregates into estimates of monthly harvested area under irrigated and rain-fed conditions. GAEZ (Fischer, et al. 2013) combined the M3-Crops (Monfreda, Ramankutty and Foley 2008) sub-national data with other spatial information on population density, agro-ecological suitability and market access in an iterative minimization model to create gridded estimates of area, production and yield for 23 crops grown under irrigated and rain-fed conditions. For the SPAM year-2000 estimates, You et al. (2014) used their own collection of sub-national statistics, MIRCA's estimates of irrigated area, GAEZs estimates of agro-ecological suitability and an iterative allocation process similar to GAEZ to create gridded estimates of physical area, harvested area, production and yield for 21 crops and crop aggregates under four production systems defined by water source and input use (irrigated, rain-fed – high inputs, rain-fed – low inputs, rain-fed – subsistence).[2]

Pixelated data on crop production are needed to shape targeted and cost-effective interventions and programs aimed at improving the institutes and incentives that affect agricultural production. The higher spatial resolution of these data arguably provides

[2] The SPAM methodology has also been used to derive regional estimates for Brazil (You and Wood 2006) and sub-Saharan Africa (You, Wood and Wood-Sichra 2009).

researchers with more locally reliable information than is obtainable from data delineated by geopolitical boundaries. While Anderson, et al. (2015) compare the estimates of M3-Crops, MIRCA, GAEZ and SPAM, there has not yet been a comprehensive assessment of the robustness of the methodological choices made within these global cropping system models, the sensitivity of results to these choices or a validation of the pixelated estimates against measured views of on-the-ground realities.[3] A shortcoming of these types of pixelated data is that they require sizable amounts of underlying crop statistics to underpin calculations, and therefore leave a paucity of "out of sample" data on cropping practices for validation processes. Without validation and robustness procedures, it is difficult to classify inherent weaknesses in the data, which could lead to erroneous conclusions in future analyses based on these data.

The objective of this chapter is to examine and quantify the robustness of HarvestChoice's SPAM2005 (Wood-Sichra, Joglekar and You forthcoming) harvested area, production and yield estimates for 42 crops in light of changes to the methodological and data factors that underpin these estimates. Specifically, this chapter includes an examination of the impacts of five methodological choices in SPAM within nine countries:[4] the allocation method, biological suitability, economic suitability, the treatment of the "rest-of-crops" crop aggregate and the administrative level of underlying statistics used. The root mean squared error (RMSE) from a regression of each of the robustness scenario estimates on the baseline estimates is used to empirically quantify which of these choices had more influence on the results. The results of this investigation reveal significant differences in the sensitivity of estimates to the five methodological choices, which differed by country, crop and production statistic (i.e., area, production or yield). The largest RMSEs were from the level of sub-national statistics used and the choice of allocation method. To validate a portion of the SPAM2005 data, the estimates of harvested area are compared with high-resolution, remote sensed data on cropland in the United States.

---

[3] Efforts by You and Wood (2006) were made in their study on crop production in Brazil using SPAM to analyze alternative allocation methods.

[4] Brazil, China, Ethiopia, France, India, Indonesia, Nigeria, Turkey and the United States.

Some of the largest effects from differences in methodological choices highlighted by this robustness analysis were in low income countries for crops with multiple cropping seasons and low yields or crops with high yields but low suitable area for production. Each methodological choice resulted in a slightly different crop footprint. There were larger discrepancies between methodological choices in countries with relatively large (geographically) sub-national administrative units, which allowed for more shifting of the crop footprint. To the extent that underlying data in countries is less reliable and the footprint affects the results of an analysis, researchers should be aware of the potential for erroneous conclusions in their analysis.

Additionally, the developers of SPAM are currently working on the next release of SPAM for 2010 cropping systems. It is quite time consuming and computationally intensive to continually update estimates. This analysis provides critical feedback for future improvements to SPAM, as well as other efforts to estimate crop production statistics.

2.2 Data

The data examined in this chapter were sourced from HarvestChoice's Spatial Production Allocation Model (SPAM) 2005 Version 2.0 global estimates of physical area, harvested area, production quantity and yield in the year 2005 for 42 crops and crop aggregates.[5] SPAM2005 disaggregated its estimates by four production systems (i.e., irrigated, rain-fed – high inputs, rain-fed – low inputs and rain-fed - subsistence). This analysis focuses exclusively on the sum (or weighted average for yields) across these production systems. The estimates were calculated with a cross-entropy optimization model which used informed priors of the physical cropping area to estimate the physical area for each crop across a global 5 arc-minute grid subject to several constraints.[6] The informed priors of physical area by crop and production systems were developed using data on crop statistics, cropland, irrigated area, suitable area, population, crop prices and potential yields to create a pixelated allocation of physical area by crop. Allocated

[5] SPAM2000 and SPAM2005 are available for download at www.mapspam.info.

[6] The data and methods used to create the SPAM2005 cropping system estimates are fully documented in Wood-Sichra, Joglekar and You (forthcoming).

estimates of harvested area, production quantities and yield were derived from the modeled allocation of physical area. The analysis in this chapter focuses on nine countries that were chosen because they vary in agro-ecology, region, income level and geographical size (see Table 2-1).

[Table 2-1: Descriptive statistics on countries analyzed]

2.2.1 Data Inputs

Arguably the most important data collected for SPAM are national and sub-national statistics on harvested area and yield for each of the 42 crops and crop aggregates studied. To form the 2005 estimates, these data were collected from a variety of sources including AgroMaps (2012), CountrySTAT (2012), Eurostat (2012), national statistical offices, ministries of agriculture and households surveys. Coverage was complete (i.e., data on at least one crop) for all nine countries at the national level and sub-national administrative division level one (ADM1), but sub-national statistics at the administrative division level two (ADM2) for any crops were not available for France, Indonesia and Nigeria, and for the remaining six countries, coverage (of at least one crop) was between 66.8 and 91.2 percent. Statistics were averaged over the 2004–2006 period and scaled to the relevant 2004–2006 FAO (2015) national values.

SPAM2005 disaggregated the reported crop statistics by production system using the share of production in each of the four systems. Shares on irrigated production were derived by dividing the harvested area cultivated under irrigation, collected primarily from AQUASTAT (2015) and MIRCA (2010), by the total harvested area. The shares under the three rain-fed systems were largely based on generalized assumptions for individual countries and crops (e.g., fertilized areas were classified as either irrigated or rain-fed – high systems, so if known, non-fertilized areas could be split between rain-fed – low input and subsistence production systems). Production system shares varied significantly among the nine countries studied in this analysis. In China 41.3 percent of maize was irrigated, while 14.2 percent was irrigated in the United States and 0.9 percent in Nigeria and Brazil. The United States had no low input maize production, while the majority of production in Ethiopia, India, Nigeria and Turkey occurred under rain-fed – low input conditions.

Estimates of sub-national physical area under production were derived from the reported harvested area and indicators of seasonal production or multi-cropping. Hard data on cropping intensity existed for some countries and crops, but quite often cropping intensities were determined by grey literature and the informed judgment of the developers. Cropping intensity values were generally one in temperate and cool climates, and for crops which had long growing periods, such as sugar cane or coconuts. Cropping intensities larger than one were common for irrigated crops, like rice, and some beans (e.g., lentils and chickpeas). Of the nine countries examined in this chapter, Ethiopia and Nigeria had the highest cropping intensities, on average (e.g., 1.78 for maize in Ethiopia and 2.00 for vegetables in Nigeria).

The biophysical constraints on crop production were modeled by interacting three pixelated datasets on cropland, irrigated land and agro-ecological suitability. SPAM2005 used the global, 30 arc-second resolution (approximately one kilometer at the equator) cropland map developed by the International Institute for Applied Systems Analysis (IIASA) and the International Food Policy Research Institute (IFPRI) which represents the median and maximum estimates of cropland circa 2005 (Fritz, et al. 2015). These data were validated against high-resolution, remote sensed data collected through two independent crowdsourcing campaigns and determined to have an accuracy of 82.4 percent globally. Within the countries of interest in this chapter, most of the cropland in the United States, China, France and Ethiopia were marked with mid- to high-confidence levels; parts of India, Nigeria and Brazil were marked with high-confidence while the rest are low; and most of Indonesia and Turkey were marked with low confidence.

Pixelated estimates of irrigated areas were taken from the Global Map of Irrigation Areas (GMIA) version 5.0 which represents the amount of area equipped for irrigation circa 2005 at a 5 arc-minute resolution (Siebert, et al. 2007). The estimated area equipped for irrigation (as a percentage of arable land) in the focus countries was 6.5 percent in Brazil, 55.5 percent in China, 2.3 percent in Ethiopia, 11.4 percent in France, 38.9 percent in India, 29.2 percent in Indonesia, 0.8 percent in Nigeria, 22.4 percent in Turkey and 17.3 percent in the United States.

Agro-ecology varies throughout the world, and certain crops perform better under a particular set of thermal, moisture, and soil conditions than would others. IIASA and FAO (2012) developed the GAEZ methodology to provide a standardized framework for the assessment of biophysical limitations of cropland globally. Using these factors, GAEZv3.0 produced a crop- and production system-specific index of suitability at a 5 arc-minute resolution; these metrics were converted into a measure of suitable area by crop for integration into SPAM. Maps of the suitable area under irrigation, rain-fed – high inputs and rain-fed – low inputs are presented by Wood-Sichra, Joglekar and You (forthcoming). There is sizable variation between the focus countries in all three suitable area layers. Prior to estimation, cropland, irrigated area and suitable area were adjusted to satisfy the constraints on physical area set by the subnational statistics.

SPAM2005 included a measure of potential revenue in the model to take account of farmers' decisions to plant one crop over another. Potential revenue was modeled using crop-specific prices derived from the FAO's Gross Production Value (constant 2004–2006 International Dollars (I$)) (FAO 2012), a measure of accessibility based the Global Rural and Urban Mapping Project, Version 1 estimate of rural population density (CIESIN, IFPRI, the World Bank and CIAT 2011, Balk, et al. 2006) and potential yields as reported by GAEZv3.0 (2012). Potential revenue was calculated on a pixelated basis for all crops and production systems.

2.2.2 Processing

The estimates on potential revenue were combined with cropland and irrigated area to calculate a *prior* for physical area by crop and production system within each pixel. Using a cross-entropy allocation approach,[7] these priors were fed into a General Algebraic Modeling System (GAMS) based model to (iteratively) minimize the error between pre-allocated shares of physical area (i.e., the priors) and an allocated share of physical area in each pixel i by crop j and production system l, subject to several constraints. The model constraints specified the necessary relationships between the allocated physical area by pixel and cropland, suitable area, irrigated area and statistical physical area. If the model

[7] De Boer, et al. (2005) offer a good introduction to the cross-entropy method.

does not solve, a series of corrective measures were used including (i) adjusting the pre-processing parameters (i.e., cropland, irrigated land and suitable area), (ii) adjusting the entropy model constraints and (iii) adjusting the data harmonization rules (e.g., applying different production shares to crop aggregates). The resulting estimates of physical area were converted to estimates of harvested area, production and yield.

2.3 Robustness Scenarios

The estimates provided by SPAM allow for spatially disaggregated analysis on crop production in a variety of contexts (e.g., mapping at higher resolution than national or sub-national statistics allow or providing the basis for more heterogeneous policy interventions) but they are only as reliable as the methodology and data that underpin them. The following analysis examines the impact of five major methodological choices on the SPAM estimates.

2.3.1 Allocation Method

Spatial crop allocation models are heavily dependent on a reliable cropland layer. While this layer does not specify which crops are grown within a pixel, it does dictate precisely which pixels will be used within the allocation process. There are two main types of spatial allocation models: (1) simple models that only use the cropland layer to inform spatial distributions of crops (Monfreda, Ramankutty and Foley 2008, Portmann, Siebert and Döll 2010) and (2) complex, optimization models that attempt to mimic the localized, bio-economic decision environment of the farmer through the inclusion of several sources of geo-spatial information (Fischer, et al. 2013, You, Wood, et al. 2014, Wood-Sichra, Joglekar and You forthcoming). The first method has computational advantages, but the second arguably captures more nuanced factors of influence. You and Wood (2006) compared the effectiveness of alternative spatial allocation models in explaining the variance in municipality crop areas (from a secondary database) in Brazil and found that the cross-entropy approach in SPAM fared better than the simpler methods.

In the first robustness test, the baseline estimates from SPAM's cross-entropy optimization model are compared to those from a proportional allocation model similar to that used to derive the M3 estimates of global harvested area and, subsequently, yield by crop circa 2000. To create their M3 data, Monfreda, Ramankutty and Foley (2008) first collect statistics on harvested area and yield from the lowest national or sub-national administrative unit with data available (i.e., ADM0, ADM1 or ADM2). Statistics on harvested area are "downscaled" into pixels using the share of cropland within the respective administrative unit:

$$CropH_{ji} = CropLand_i \times \frac{CropH_{jk}}{CropLand_k}, \forall i \in k$$

where $CropH_{ji}$ is the estimated harvested area of crop j in pixel i, $CropLand_i$ is the total cropland area in each pixel i, $CropLand_k$ is the total cropland area calculated for each administrative unit k and $CropH_{jk}$ is the statistical harvested area of crop j in administrative unit k. Pixelated estimates of yield are subsequently calculated by the following:

$$\begin{array}{ll} CropY_{ji} = CropY_{jk} & if\ CropH_{ji} > 0 \\ CropY_{ji} = 0 & if\ CropH_{ji} = 0 \end{array}$$

where $CropY_{ji}$ is the estimated yield of crop j in pixel i and $CropY_{jk}$ is the statistical yield of crop j in administrative unit k. Details on the exact methodology used to calculate the estimates used in this robustness test are presented in Appendix A.

The alternative allocation method ties harvested area data to the lowest national or sub-national administrative unit available with data (most often the ADM2 level). While data is collected at the lowest administrative unit possible, SPAM standardizes statistics either at an ADM1 or ADM0 level depending on the size of the country. All nine countries studied in this analysis were modeled at the ADM1 level, which means that cropping statistics that represent a single ADM2 unit can actually be spread out among several ADM2 units within the greater ADM1 unit. Depending on the country and crop, this could account for a significant shift in the crop footprint between the two allocation methods.

2.3.2 Biological Suitability

Suitable area is one of the major variables included in SPAM that accounts for the biological constraints to agriculture. Its importance was tested by removing suitable area as a constraint in the allocation optimization model.[8] The specific constraint affected specifies that the allocated physical area by pixel, crop and production system must not exceed the relevant suitable area within the pixel. Removing this constraint may cause the model to distribute crops into unsuitable areas, which could overestimate the production and area of less suitable pixels.

2.3.3 Economic Suitability

Economic suitability, as represented by potential revenue, is a function of global crop prices, population density (a proxy for market access) and potential yields in SPAM. To measure the impact of this variable on SPAM, the model is re-run without variation in crop prices (i.e., all prices are set to I$/mt). Given the tradeoff between two crops, farmers will choose to plant the more profitable one, *ceteris paribus*, so removing this layer may alter the total harvested area under a particular crop within a pixel. However, a global price layer may not accurately reflect the localized profitability trade-offs for farmers around the world, so this methodological choice may not cause significant changes in the pixel-level distribution of crops.

2.3.4 "Rest of Crop" Allocation

One of the nine crop aggregates is a catch-all for the minor crops not covered by the other 40 crops and crop aggregates (e.g., spices, tree nuts, other sugar crops, mate and rubber), but are reported by FAO (2015). In SPAM2005, data were collected for this aggregate and modeled simultaneously with the other crops. However, in SPAM2000, the "rest-of-crops" aggregate was allocated after the model was run, and accounted for any unused cropland. This robustness run tests the difference between these two methods of "rest-of-crops" allocation. Actively allocating the rest-of-crops category may result in higher levels of displacement of other crops because it is now "competing" simultaneously for a location, rather than being allocated after the other crops have been optimally placed within the model.

[8] Suitable area was still used to adjust cropland and irrigated areas, as described in the Appendix B of Wood-Sichra, Joglekar and You (forthcoming).

2.3.5 Administrative Level of Statistics

The final robustness scenario focuses on the aggregation of crop statistics. In general, the accuracy of the SPAM model will increase if more spatially disaggregated crop statistics are fed into the model. However, it is quite costly to collect agricultural production data on a large enough scale to ensure the data are representative at these spatial levels. Often, agricultural household surveys are only representative at either an ADM0 or ADM1 level. Occasionally, censuses will contain large enough sample sizes to produce representative aggregates at an ADM2 level (e.g., Tanzania's 2007 Agricultural Sample Census (NBS 2011a)). Before running SPAM, statistics on crop production and productivity are compiled at an ADM1 level[9] and, if available, at an ADM2 level. If there is no information on a crop at that level then it is the ADM1 or ADM0 level statistic is used. The SPAM allocation model is better facilitated by more spatially disaggregated information, but what happens to the pixel-level estimates if information is only available at an ADM1 or ADM0 level? If there is less restriction on the boundary of crop allocation (i.e., ADM1 versus ADM2), then there is a higher likelihood of misallocating crops, and the crop footprints may shift dramatically; this is especially true if the relevant geopolitical units are large.

For this test, only Brazil and the United States were used because other areas would not solve with such little data. Within these two countries, four regions would not solve: Parana in Brazil and North Dakota, New Mexico and Idaho in the United States.

2.4 Results and Discussion

Descriptive statistics for the baseline and each of the robustness scenario estimates are presented in Table 2-2 for maize harvested area, Table 2-3 for maize production quantities and Table 2-4 for maize yields. Since SPAM estimates are scaled to match FAO national totals, the expected mean values across robustness scenarios should be the same as the baseline. While they are similar, differences occur due to the estimated number of pixels with positive crop allocation. The robustness tests on allocation method and level of

[9] For very few countries (e.g., geographically small or politically unstable countries) statistics are only available at an ADM0 level.

statistics resulted in the highest differences in pixels allocated and averages. Across all countries, the ‘allocation method’ assigned more pixels with positive maize production than the baseline scenario (the United States had the lowest difference – 8.7 percent more pixels and India had the highest – 64.2 percent more pixels). Across all of the robustness scenarios, Ethiopia had the most notable differences in allocated pixels and means from the baseline scenario. Thus, datasets derived under different allocation methods (e.g., M3-Crops and SPAM) may have significantly different crop footprints, especially between countries. For analyses where this footprint is a primary input, such as simulation models, (Franch, et al. 2015, Hutabarat, et al. 2012, Johnson, Takeshima and Gyimah-Brempong 2013) this could affect the results, especially for countries with larger administrative units connected to the underlying crop statistics (e.g., the average size of ADM2 units in Ethiopia is 14.7 square kilometers, as compared to 1.6 square kilometers in Brazil).[10]

[Table 2-2: Descriptive statistics on maize harvested area (ha) estimates]

[Table 2-3: Descriptive statistics on maize production (mt) estimates]

[Table 2-4: Descriptive statistics on maize yield (kg/ha) estimates]

For harvested area and production quantity, the percentage difference in number of pixels is generally reflected in the percentage difference in statistical averages across all countries (i.e., a 2.1 percent increase in the number of pixels results in a 2.1 percent decrease in the estimated average between the robustness scenario and the baseline). This is not the case for the yield robustness runs, which can be the product of crops with very low relative yields.

The objective of SPAM and other spatial crop distribution datasets is to capture the local heterogeneity in crop production. The root mean squared error (RMSE) from a regression of the robustness scenario estimates on the baseline estimates was used to empirically quantify and compare the degree of pixel-based differences from changes in methodological factors across countries using a single indicator. As the name of the statistic suggests, the RMSE finds the square root of the average differences between the baseline estimate (y_i) and the fitted values from the regression ($\hat{y}_i$), squared:

[10] Beddow and Pardey (2015) caution that using the incorrect spatial footprint of crop production will likely lead to erroneous conclusions.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

To examine the influence of these methodological factors across both countries and crops, two separate clusters of regressions were run. In the first, the robustness scenario estimates, crop type and an interaction between the two variables are regressed on the baseline estimates; these were run separately for each country, robustness scenario and production statistic type (i.e., harvested area, production or yield) for a total of 147 separate regressions. In the second cluster, the robustness scenario estimates are regressed on the baseline estimates separately for each country, robustness scenario, production statistic type and each of the 42 crops studied for a maximum of 6,174 regressions, depending on whether a crop is grown within a country. OLS regressions of the following forms were used:

Cluster 1: $baseline_{iks} = \beta_0 + \beta_1 scenario_{ikrs} + \boldsymbol{\beta_2 crop_j} + \boldsymbol{\beta_3}\left(scenario_{ikrs} \times \boldsymbol{crop_j}\right)$

Cluster 2: $baseline_{ijks} = \beta_0 + \beta_1 scenario_{ijkrs}$

where $baseline$ is the SPAM estimate under the baseline scenario, $scenario$ is the robustness scenario estimate and $crop$ is a vector of crop dummies. The subscripts represent pixel i, crop j, country k, robustness scenario r and crop production statistic s. Traditionally, RMSEs are used to assess the predictive accuracy of forecasting models. In this analysis, they are used to assess the degree of similarity between estimates from a robustness scenario and the baseline. Thus, higher RMSEs signal (relatively) higher levels of sensitivity within the SPAM model to the methodological factor being tested. The RMSE does not give a sense of the direction of differences between the baseline and robustness scenarios. The RMSEs from the first cluster regressions are presented in Figure 2-1.[11] The results from the second cluster regressions are presented in Appendix B.

[11] Since the allocation method robustness run does not include estimates of physical area, only regressions with respect to harvested area, production quantity and yields are included. The results for Brazil and the United States across all robustness scenarios were calculated with the four aforementioned ADM1 units removed to facilitate comparison. These results do not differ significantly from those that include the four ADM1 units. Finally, while it could also be a focus of analysis, for simplicity, the results delineated by production system (irrigated, rain-fed – high, rain-fed – low and subsistence) are not presented. Instead, the presentation only includes crop production and productivity estimates with respect to total production.

[Figure 2-1: Normalized RMSEs between robustness runs and baseline estimates]

2.4.1 Robustness Comparison across Countries and Crops

To facilitate comparison between three production statistics with different scales, RMSEs were normalized by dividing by the range in baseline estimates, specific to the production statistic. Figure 2-1 is a heatmap of the normalized RMSEs (NRMSE) by country and production statistic. Overall, SPAM was most sensitive to the allocation method choice and the level of underlying statistics used, especially with regard to yield estimates. It was least sensitive to the passive versus active "rest-of-crop" allocation, for all three production statistics. In Brazil and the United States, the NRMSEs were highest for the level of administrative units used, which reinforces the need to collect reliable data at higher levels of sub-national disaggregation. In these two countries the NRMSEs for 'level of statistics' were greater than those associated with the allocation method robustness run which may be due to the relatively small ADM2 units and strong data collection efforts within the countries.

Production quantity estimates were the least sensitive to methodological choices, while yield estimates were the most sensitive. This is due in part to how the proportional allocation method spatially distributes yield statistics. Yields are assigned to pixels if a positive harvested area has been assigned to the pixel. Thus, every pixel within the administrative unit represented by the yield statistic will be the same. In a country that has ADM2 units with relatively small areas, these localized differences (between the two allocation methods) may not be as apparent as in a country with larger ADM2 units.

There were noticeable differences in how the methodological choices studied affect SPAM from a county-level perspective. Overall, Ethiopia was the most sensitive to methodological choices, while France and the United States were the least. There could be a relationship between the income level of the country and the quality of underlying statistics collected. Other notable countries with relatively high NRMSEs included India for harvested area and production quantity estimates, and Indonesia and China for yield estimates. Figure 2-1 shows that the estimates from one country might be highly sensitive to a certain factor that has very little influence in another. For example, the estimates on harvested area in India were dependent on biological suitability constraint, but this was not the case in France, Turkey or the United States. While the three temperate countries studied, France, Turkey and the United States were all relatively unaffected by methodological changes with regard to the harvested area and production quantities estimated, there did not appear to be other noticeable patterns among the NRMSEs for other agro-ecological zones.

Heatmaps for the NRMSEs from the second cluster of regressions are presented in Appendix B-1 through Appendix B-9 for each of the nine countries studied. These plots show that, similar to the results from the first cluster of regressions, the relative impact of specific methodological factors on SPAM varies by crop as well as country and production statistic, and there are a few scenarios that stick out. In nearly every country, there were higher NRMSEs associated with sugar cane or sugar beets; the highest NRMSE was for sugarcane yields in Ethiopia (see Appendix B-3) with regard to the choice of allocation method. For most crops in the countries studied, the magnitude of the NRMSE was similar between production statistics and robustness scenarios, but there were cases where the NRMSE was high for one statistic and low for the others. For instance, there were relatively large differences in estimated harvested area of other cereals in Ethiopia (other cereals includes the Ethiopian staple crop teff) for all four robustness assumptions. Comparatively, there were low differences with respect to production quantities and yields.

The example of other cereals in Ethiopia can be used to demonstrate how the degree of change introduced by each robustness factor varies locally on a pixel-by-pixel basis. Figure 2-2, Figure 2-3 and Figure 2-4 map the differences between each of the robustness scenarios and the baseline scenario of harvested area, production and yield for other cereals in Oromia, Ethiopia. There are noticeable differences between each of the robustness scenarios and the baseline scenario. The pixels with maximum differences in harvested area indicate that there is a large, positive estimate of harvested area under other cereals under from the robustness run that is estimated as zero in the baseline (or vice versa) (Figure 2-2). The reason Ethiopia has such high RMSEs for other cereals under all four robustness scenarios is because the estimated cropping intensity on teff is high. Consequently, the yield estimate for the other cereals aggregate category is quite low (Figure 2-4); thus, its difference from zero would not result in a high RMSE.

SPAM estimates will be more susceptible to errors when a major crop in the country is included in an aggregate. Additionally, by nature of the alternative allocation method, if the crop footprint is quite different than under the baseline scenario, the corresponding yields could also be quite different, especially if they are relatively high (e.g., sugarcane in Ethiopia). Any pixel with positive harvested area within the relevant administrative unit (i.e., the administrative unit with the most disaggregated level of statistics available) will be assigned the yield associated with that relevant administrative unit. Any discrepancies introduced by a methodological choice will be bound by the borders of the relevant administrative unit, so countries with geographically larger administrative units are more susceptible to errors.

[Figure 2-2: Differences in other cereals harvested area (ha) estimates]

[Figure 2-3: Differences in other cereals production (mt) estimates]

[Figure 2-4: Differences in other cereals yields (kg/ha) estimates]

2.4.2 Comparison to Remote Sensed Data

Rasterized estimates of crop production statistics represent a "plausible" accounting of the spatial structure of crop performance within a country, conditioned on a host of source data and measurement factors. To validate the SPAM estimates, secondary data sets on crop production are needed, but finding statistics that have not already been used within the model is difficult, especially since these data have been shown to be important within SPAM. As an alternative, the present analysis utilized the high-resolution Cropland Data Layers provided by United States Department of Agriculture's (USDA) National Agriculture Statistics Service (NASS). These layers delineate the major crop or land cover categories (e.g., wetlands or forest) within each 30 meter pixel (USDA - NASS 2004, 2005, 2006).[12] The SPAM2005 estimates of harvested area in maize, soybeans, cotton, rice and wheat were compared to NASS estimates of physical area, averaged from 2004–2006 and aggregated to a 5 arc-minute grid resolution.[13] Only states with complete coverage in all

[12] While NASS does not publish accuracy assessment tables for 2004, 2005 or 2006, they do start doing so in 2008. At that time it was estimated that maize had a producer accuracy ranging from the high-80 percent to the mid-90 percent, depending on the state.

[13] The differences between estimates of physical area and harvested area in the United States was trivial during this time period, due to limited instances of double-cropping.

three years were used for the validation exercise: Illinois, Indiana, Iowa, Louisiana, Mississippi, Nebraska, North Dakota and Wisconsin. The Pearson correlation coefficients between physical area reported by NASS and harvested area reported by SPAM2005 are presented in Table 2-5 for the estimates from the baseline, allocation method, ADM0 only and ADM1 only scenarios.

[Table 2-5: Correlations between NASS and SPAM crop area estimates]

The baseline SPAM estimates are strongly correlated with the area estimates from NASS's remote sensed data, though the degree of correlation varies by state and crop. For example, the baseline SPAM estimates in Iowa are strongly correlated with the NASS estimates for maize ($\rho = 0.82$) and soybeans ($\rho = 0.78$), but weakly correlated for wheat ($\rho = 0.03$). While the primary crops in Iowa are maize and soybeans, there does not appear to be a persistent relationship between the degree of correlation and minor versus major crops across the eight states. Wheat represents four percent (on average) of the total harvested area in Illinois, Indiana, Louisiana and Wisconsin, but the correlation between the baseline SPAM estimates and NASS estimates for wheat in these four states is between 0.71 and 0.82. Among the eight states examined, estimates of area in Nebraska from SPAM most match those in the NASS data.

Table 2-5 also includes correlation coefficients between three of the robustness scenario estimates and NASS: the allocation method, ADM0 only and ADM1 only. The correlation coefficients between the allocation method and NASS and the baseline and NASS are similar, but there are times when the allocation method is a better predictor of NASS and vice versa. In Mississippi, correlation coefficient associated with the allocation method is 25.5 percent higher for maize than the coefficient associated with the baseline. In general, the estimates from the ADM0 only robustness test are least correlated with the NASS estimates, especially for the southern states. For maize in Nebraska and soybeans in Iowa, the correlation coefficients are always high regardless of the methodological choice made.

A significant portion of the difference between SPAM and NASS could be the accuracy of the underlying cropland layer used in SPAM. Figure 2-5 contains pairwise comparisons between the cropland as reported by NASS and used within SPAM2005.

There are strong correlations between the two variables in all states, but Nebraska has the highest correlation ($\rho = 0.91$). However, from an aggregate perspective, there are significant differences between the two extents. In Louisiana, there is only a 0.3 percent difference in the total cropland (square kilometers) between SPAM and NASS. In Indiana and Mississippi there is a 25.9 percent difference NASS and SPAM. Since the cropland is a major input in all of the scenarios examined in this analysis, errors in this layer will exacerbate errors from other data, such as cropping intensities and suitability.

[Figure 2-5: Pairwise comparison between NASS and SPAM cropland estimates]

2.5 Conclusions

The pixel-level SPAM estimates are not necessarily meant to reflect micro-level on-the-ground realities. Rather, SPAM aims to provide a better-informed understanding of the heterogeneity of cropping systems than is can be gleaned from either national or sub-national statistics for geopolitical entities. However, the plausibility of these estimates is tied to the methodological decisions made in throughout the modeling process. While access to global, disaggregated data on localized cropping decisions is limited, it is possible to test the robustness of the methods and data that define the SPAM process. To that end, this paper presents the results of an examination of the relative influence of the overall allocation method, biological suitability layer, crop prices, crop choices, the method used to allocate the "rest-of-crops" variable and the level of sub-national statistics used.

The analysis presented here demonstrates that SPAM results are dependent on the underlying national and sub-national statistics on harvested area and yield, and especially on the level of disaggregation in these statistics within the nine countries studied. The results were also quite sensitive to using a simple model of distribution based on cropland proportions rather than a cross-entropy allocation method. The influence of methodological choices varied by country, crop and production statistic. While there were differences in the estimated crop harvested area, production and yields from each robustness run, there were also different crop footprints between the scenarios. It is important to recognize that taking (any) data as truth without understanding how it was generated could lead to erroneous conclusions. For instance, the RMSEs calculated in Ethiopia revealed that there were large discrepancies in the estimates of area harvested of other cereals (a crop aggregate that included the dominant crop teff) between the robustness runs. Misidentification of the cropping intensities or yields for major crops and especially crop aggregates, could introduce compounding errors through the allocation process.

Understanding the implications of the methodological choices within SPAM can also help direct future efforts to calibrate spatial models of cropping systems. A particularly important and unexpected result of robustness analysis was that removing crop prices from the model seemed to have little effect on the estimates. While subsistence farming is prevalent in many parts of the world, one would expect that most of the production decisions made within global agriculture are intended to improve profitability (and its variance). That prices do not seem to substantially influence the allocation could be an artifact of the way prices are incorporated into the model (i.e., within a revenue function rather than a profit function). However, it is more likely that using a global crop price fails to reflect the local profitability decisions made by farmers. The need for more spatially representative agricultural prices is further discussed in Chapter 4 of this dissertation.

Table 2-1: Descriptive statistics on countries analyzed

Country	Agro-Ecology	Income	Size (million km^2)	ADM1 Units (count)	ADM2 Units (count)	ADM1 Coverage (percent)	ADM2 Coverage (percent)
Brazil	Temperate/Tropics	Upper-Middle	8.59	31	5,360	87.10	98.75
China	Temperate/Tropics	Upper-Middle	9.45	32	2,373	100.00	97.81
Ethiopia	Tropics	Low Income	1.14	11	77	100.00	80.52
France	Temperate	High: OECD	0.57	22	96	100.00	-
India	Temperate/Tropics	Lower-Middle	3.04	34	570	100.00	98.07
Indonesia	Tropics	Lower-Middle	2.25	31	426	100.00	-
Nigeria	Tropics	Lower-Middle	0.92	37	528	100.00	-
Turkey	Temperate	Upper-Middle	0.81	12	26	100.00	100.00
USA	Temperate	High: OECD	9.58	51	3,106	98.04	79.97

Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and the 2015 World Bank country and lending groups definitions: http://data.worldbank.org/about/country-and-lending-groups

Note: ADM1 – administrative level one; ADM2 – administrative level two. Coverage statistics represent percentage of administrative units with statistics on at least one crop.

Table 2-2: Descriptive statistics on maize harvested area (ha) estimates

			Robustness Run					
Country	Statistic	Baseline	Allocation Method	Biological Suitability	Economic Suitability	"Rest of Crop" Allocation	ADM0 Only	ADM1 Only
Brazil	Mean	226.7	186.4	228.0	227.2	226.7	218.3	237.5
	Std. Dev.	490.4	592.3	580.9	490.0	490.2	367.2	474.5
	N	53,780	65,390	53,468	53,650	53,768	55,850	51,324
China	Mean	505.5	410.2	505.4	506.7	505.3		
	Std. Dev.	816.5	760.1	848.6	852.7	817.0		
	N	51,988	64,066	51,999	51,867	52,004		
Ethiopia	Mean	459.0	378.8	437.0	448.3	448.0		
	Std. Dev.	1,278.4	613.2	1,389.2	1,321.2	1,345.8		
	N	3,833	4,645	4,026	3,924	3,927		
France	Mean	244.2	198.4	244.1	244.2	244.2		
	Std. Dev.	276.7	268.3	277.5	276.9	276.7		
	N	6,750	8,309	6,753	6,749	6,750		
India	Mean	396.4	241.4	391.1	396.6	396.4		
	Std. Dev.	1,064.6	448.1	1,165.8	1,033.4	1,053.3		
	N	19,269	31,646	19,529	19,261	19,268		
Indonesia	Mean	170.2	138.1	173.9	170.3	170.1		
	Std. Dev.	771.0	517.1	828.8	756.3	770.3		
	N	20,207	24,914	19,776	20,204	20,220		
Nigeria	Mean	487.4	400.9	485.4	487.4	487.4		
	Std. Dev.	385.8	462.0	417.1	383.7	385.5		
	N	7,505	9,124	7,536	7,505	7,505		
Turkey	Mean	59.3	46.8	51.5	59.2	59.2		
	Std. Dev.	110.3	117.1	102.5	113.2	110.5		
	N	9,444	11,966	10,871	9,459	9,461		
USA	Mean	531.5	489.0	528.8	531.8	531.6	478.9	474.8
	Std. Dev.	815.7	1,019.2	814.4	815.8	815.7	476.9	719.7
	N	55,678	60,516	55,969	55,651	55,672	61,798	62,326

Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Table 2-3: Descriptive statistics on maize production (mt) estimates

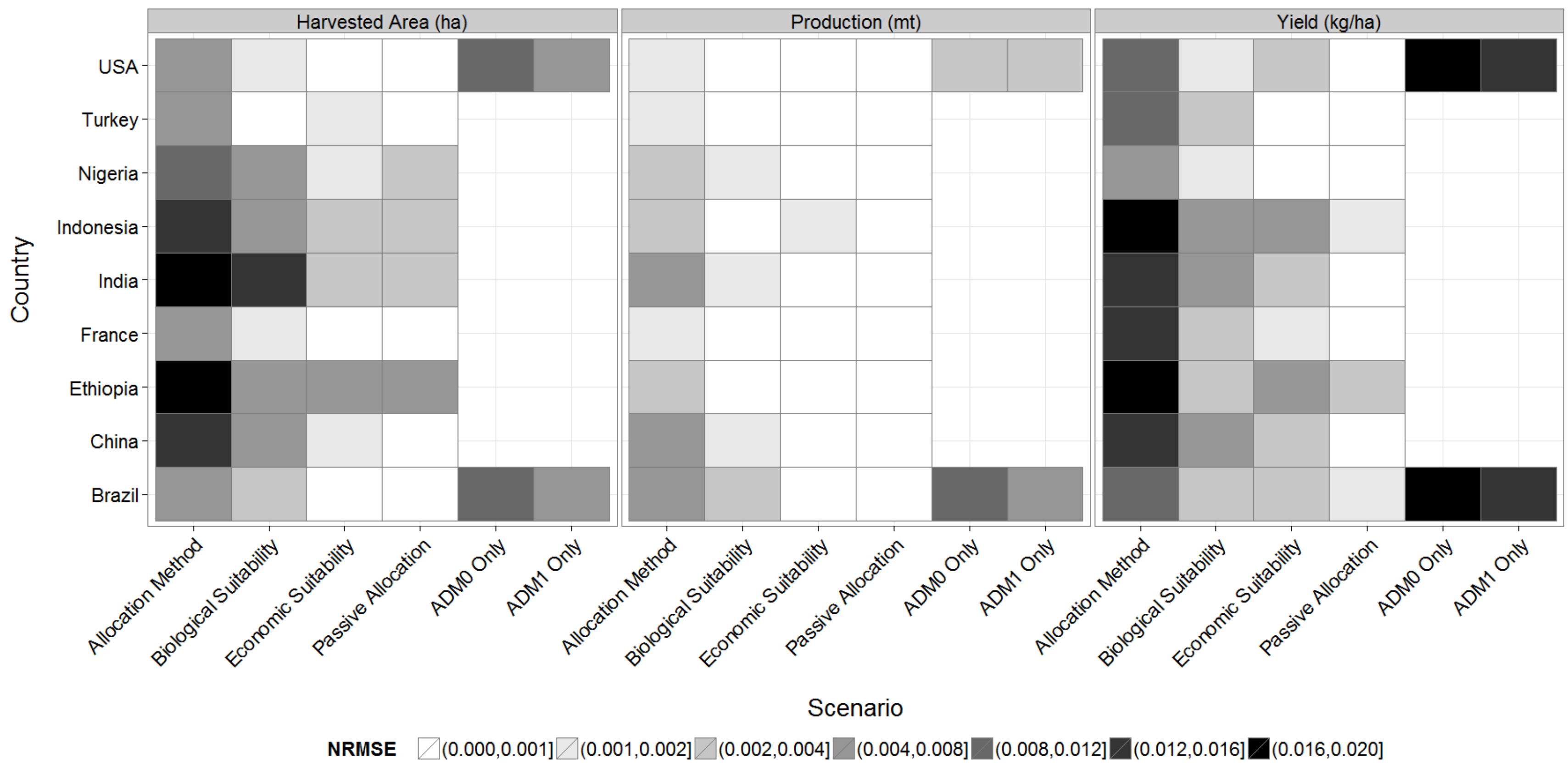
			Robustness Run					
Country	Statistic	Baseline	Allocation Method	Biological Suitability	Economic Suitability	"Rest of Crop" Allocation	ADM0 Only	ADM1 Only
Brazil	Mean	741.1	662.7	745.4	742.8	741.2	713.6	776.5
	Std. Dev.	1,935.1	2,301.4	2,279.3	1,943.0	1,933.6	1,509.9	1,935.7
	N	53,780	65,390	53,468	53,650	53,768	55,850	51,324
China	Mean	2,703.6	2,258.5	2,703.0	2,709.9	2,702.8		
	Std. Dev.	5,272.3	4,885.6	5,371.5	5,420.4	5,273.9		
	N	51,988	64,066	51,999	51,867	52,004		
Ethiopia	Mean	943.4	790.1	898.1	921.5	920.8		
	Std. Dev.	3,227.5	1,316.3	3,240.6	3,185.0	3,214.3		
	N	3,833	4,645	4,026	3,924	3,927		
France	Mean	2,115.3	1,746.4	2,114.4	2,115.6	2,115.3		
	Std. Dev.	2,763.5	2,450.4	2,771.7	2,767.8	2,763.6		
	N	6,750	8,309	6,753	6,749	6,750		
India	Mean	760.8	449.2	750.7	761.1	760.8		
	Std. Dev.	2,843.9	900.9	2,929.1	2,826.6	2,822.6		
	N	19,269	31,646	19,529	19,261	19,268		
Indonesia	Mean	583.3	485.5	596.0	583.4	582.9		
	Std. Dev.	3,255.3	1,908.4	3,385.1	3,167.3	3,253.3		
	N	20,207	24,914	19,776	20,204	20,220		
Nigeria	Mean	827.2	709.5	823.8	827.2	827.2		
	Std. Dev.	841.6	963.2	871.1	836.5	840.9		
	N	7,505	9,124	7,536	7,505	7,505		
Turkey	Mean	388.6	307.6	337.6	388.0	387.9		
	Std. Dev.	908.2	773.5	848.5	935.6	909.3		
	N	9,444	11,966	10,871	9,459	9,461		
USA	Mean	5,086.6	4,833.7	5,060.2	5,089.1	5,087.2	4,583.0	4,544.0
	Std. Dev.	8,613.2	10,803.4	8,597.7	8,613.7	8,613.5	5,195.7	7,544.4
	N	55,678	60,516	55,969	55,651	55,672	61,798	62,326

Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Table 2-4: Descriptive statistics on maize yield (kg/ha) estimates

			Robustness Run					
Country	Statistic	Baseline	Allocation Method	Biological Suitability	Economic Suitability	"Rest of Crop" Allocation	ADM0 Only	ADM1 Only
Brazil	Mean	2,447.4	2,352.5	2,452.0	2,451.3	2,453.2	2,979.7	2,911.4
	Std. Dev.	1,661.2	1,377.4	1,668.3	1,650.3	1,665.3	1,583.9	2,243.3
	N	53,780	65,390	53,468	53,650	53,768	55,850	51,324
China	Mean	4,254.4	4,542.1	4,337.9	4,264.9	4,253.7		
	Std. Dev.	2,478.3	1,839.6	2,497.6	2,477.0	2,478.3		
	N	51,988	64,066	51,999	51,867	52,004		
Ethiopia	Mean	1,494.7	1,882.9	1,512.6	1,464.0	1,495.9		
	Std. Dev.	897.9	545.6	902.5	887.8	897.1		
	N	3,833	4,645	4,026	3,924	3,927		
France	Mean	7,759.5	8,459.0	7,756.9	7,762.6	7,758.9		
	Std. Dev.	3,023.4	960.6	3,026.2	3,032.3	3,024.0		
	N	6,750	8,309	6,753	6,749	6,750		
India	Mean	1,712.4	1,946.7	1,719.9	1,701.9	1,712.4		
	Std. Dev.	1,424.8	1,423.3	1,418.3	1,418.5	1,425.3		
	N	19,269	31,646	19,529	19,261	19,268		
Indonesia	Mean	2,226.7	2,604.8	2,274.9	2,316.5	2,235.4		
	Std. Dev.	1,502.9	877.6	1,553.4	1,546.8	1,504.4		
	N	20,207	24,914	19,776	20,204	20,220		
Nigeria	Mean	1,571.1	1,571.1	1,562.3	1,558.4	1,558.2		
	Std. Dev.	505.1	505.1	593.4	592.6	593.6		
	N	9,124	9,124	7,536	7,505	7,505		
Turkey	Mean	5,562.1	6,663.2	5,712.1	5,522.9	5,552.1		
	Std. Dev.	2,663.7	564.1	2,629.6	2,628.8	2,651.6		
	N	9,444	11,966	10,871	9,459	9,461		
USA	Mean	8,068.4	8,032.7	8,085.4	8,071.1	8,069.6	9,553.3	8,385.0
	Std. Dev.	2,530.4	2,225.8	2,528.8	2,535.7	2,534.3	3,468.4	2,339.5
	N	55,678	60,516	55,969	55,651	55,672	61,798	62,326

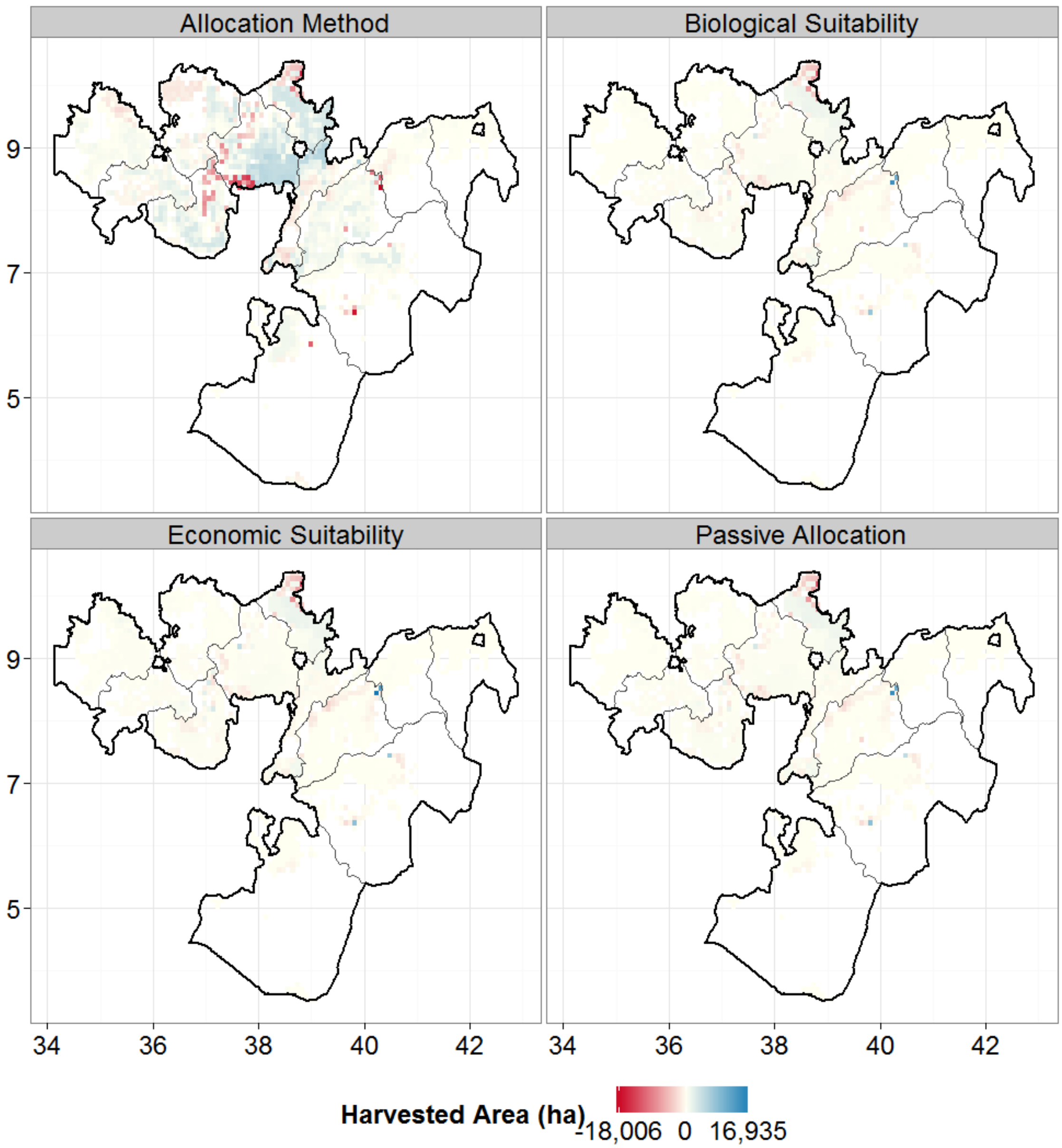
Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.



Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Note: To account for differences in scales between production statistics, root mean squared errors (RMSE) were normalized by dividing the RMSE by the range in baseline estimates from each production statistic.

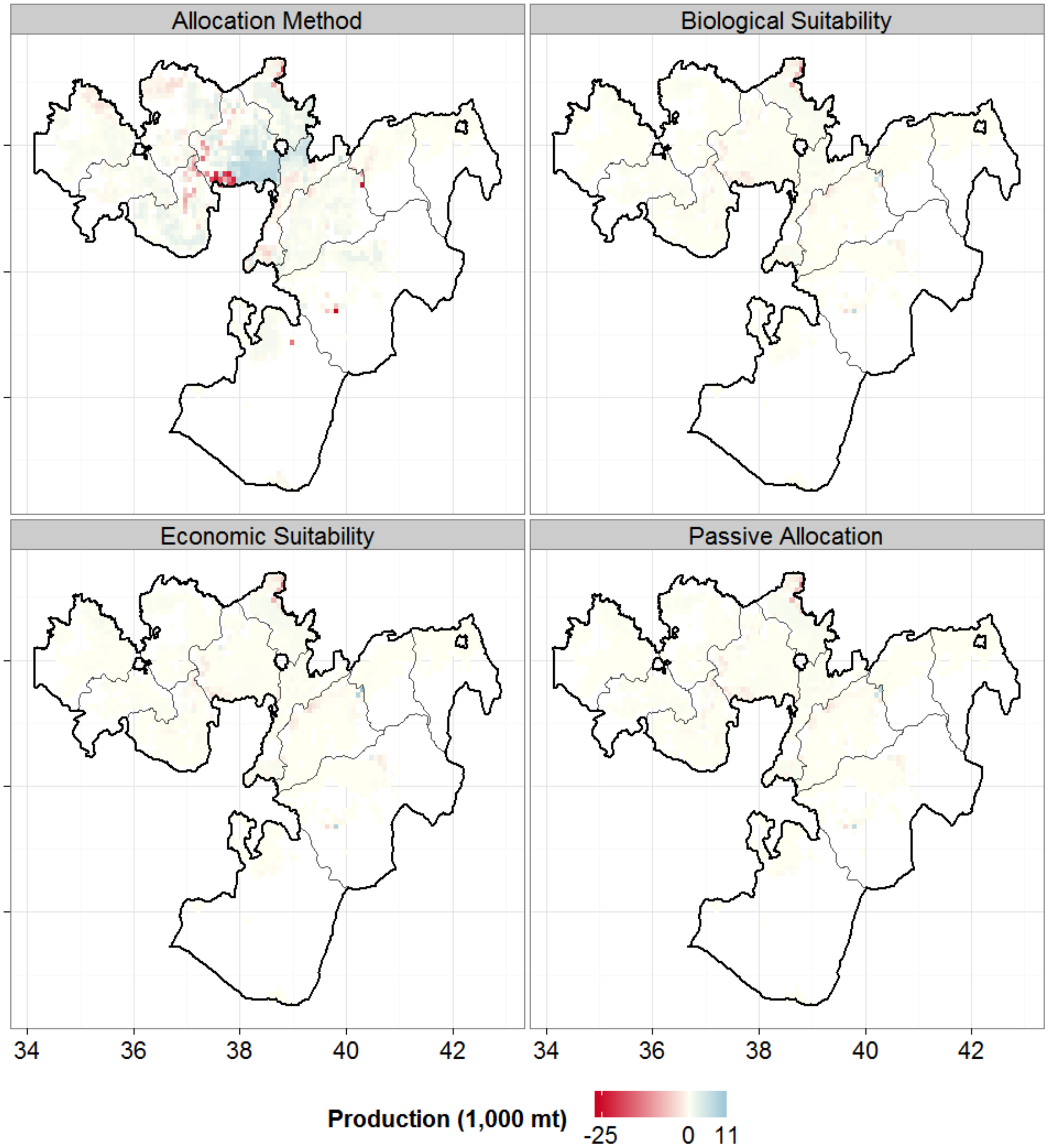
Figure 2-1: Normalized RMSEs between robustness runs and baseline estimates



Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Note: Map is of the Oromia region in Ethiopia; ADM2 unit borders are delineated in grey. Differences in estiamtes shown are between each of the four robustness runs and baseline.

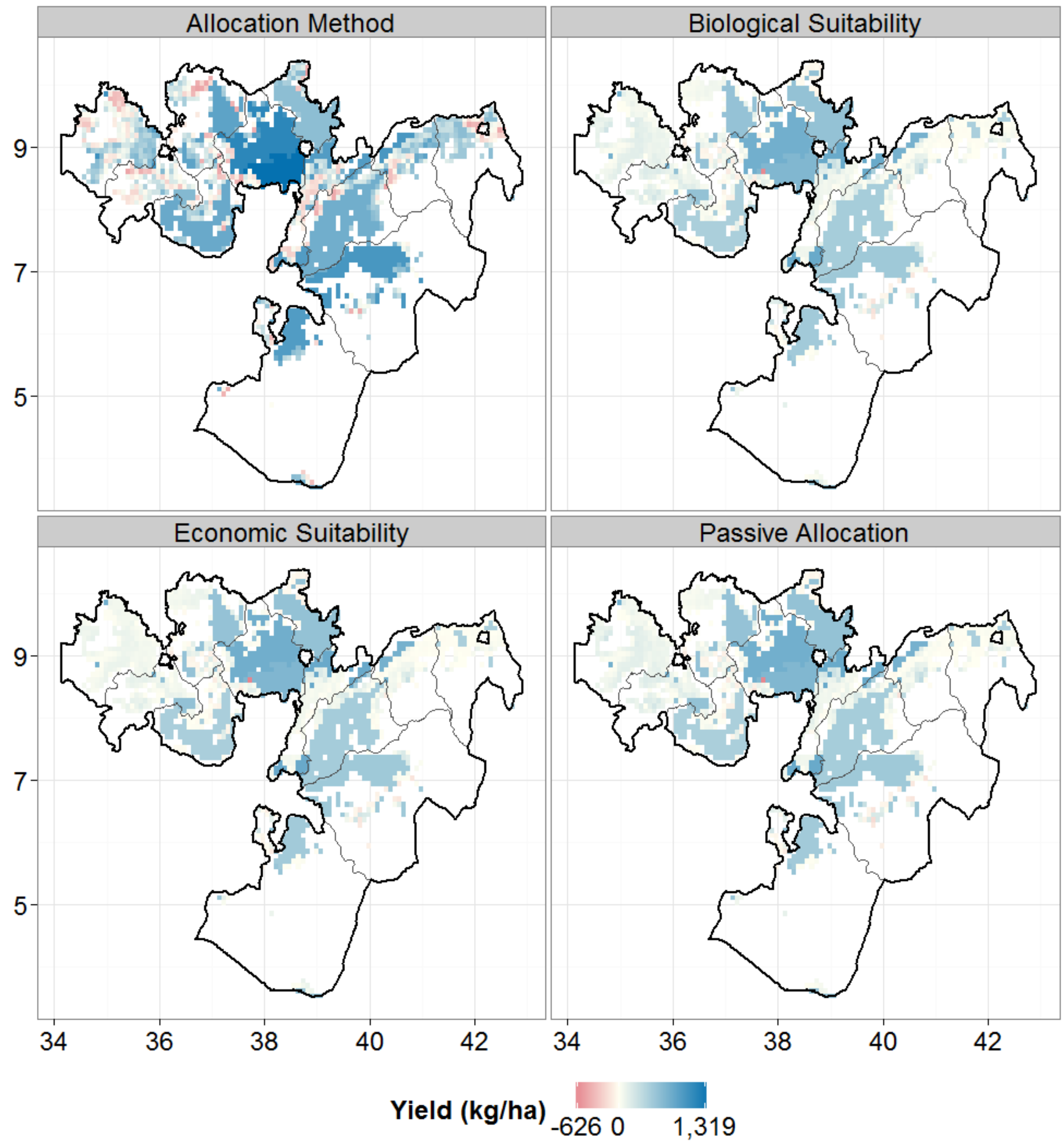
Figure 2-2: Differences in other cereals harvested area (ha) estimates



Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Note: Map is of the Oromia region in Ethiopia; ADM2 unit borders are delineated in grey. Differences in estiamtes shown are between each of the four robustness runs and baseline.

Figure 2-3: Differences in other cereals production (mt) estimates



Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

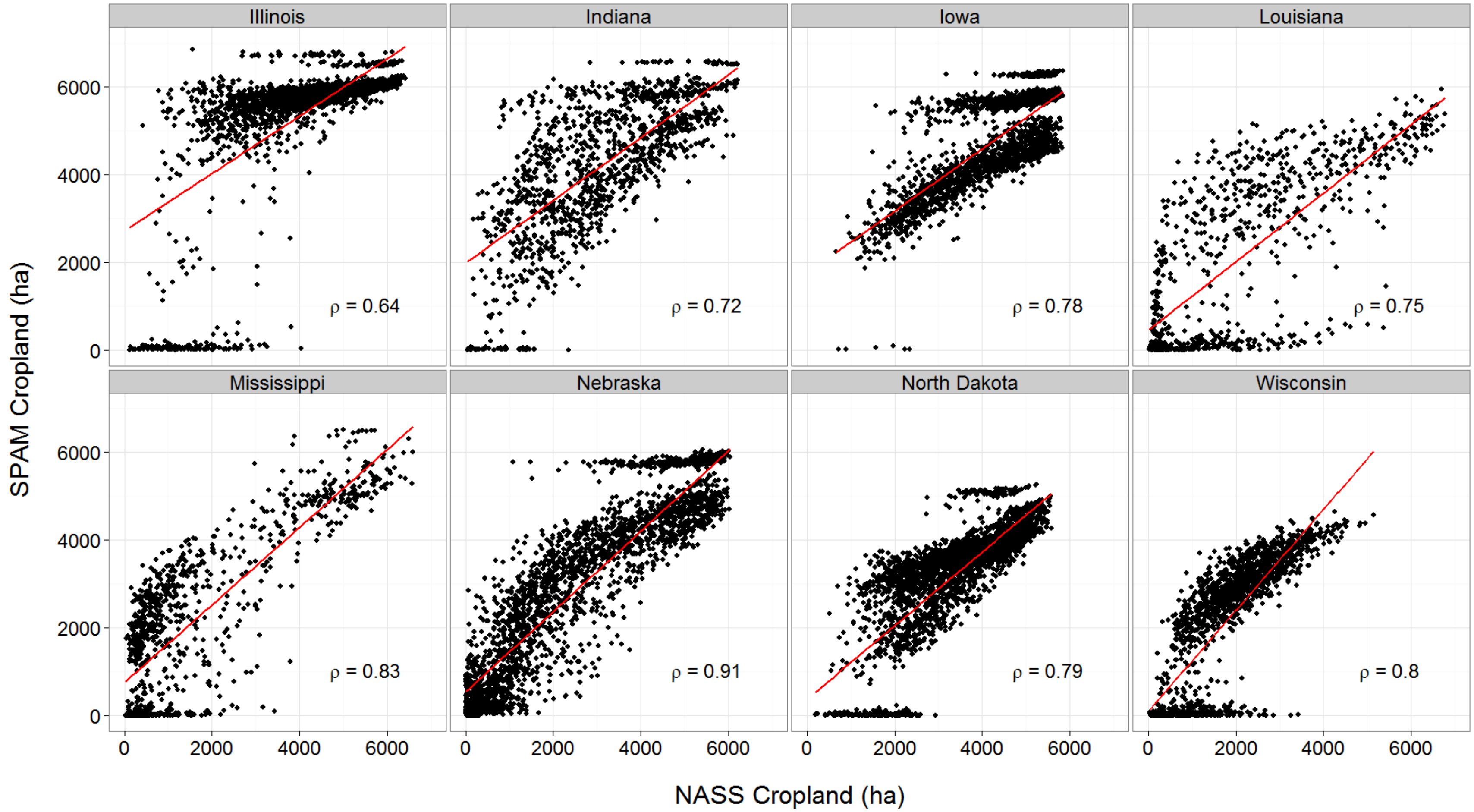
Note: Map is of the Oromia region in Ethiopia; ADM2 unit borders are delineated in grey. Differences in estiamtes shown are between each of the four robustness runs and baseline.

Figure 2-4: Differences in other cereals yields (kg/ha) estimates

Table 2-5: Correlations between NASS and SPAM crop area estimates

		Pearson Correlation Coefficients				
		Crop				
Scenario	State	Maize	Soybeans	Cotton	Rice	Wheat
Baseline	Illinois	0.78	0.71			0.82
	Indiana	0.82	0.77			0.71
	Iowa	0.82	0.78			0.03
	Louisiana	0.78	0.80	0.81	0.78	0.14
	Mississippi	0.55	0.90	0.76	0.78	0.63
	Nebraska	0.94	0.94			0.77
	North Dakota	0.83	0.79			0.55
	Wisconsin	0.69	0.72			0.81
Allocation Method	Illinois	0.82	0.75			0.83
	Indiana	0.78	0.77			0.67
	Iowa	0.76	0.71			0.01
	Louisiana	0.81	0.82	0.82	0.82	0.07
	Mississippi	0.69	0.89	0.74	0.81	0.67
	Nebraska	0.87	0.88			0.69
	North Dakota	0.81	0.92			0.59
	Wisconsin	0.81	0.81			0.80
ADM0 Only	Illinois	0.71	0.71	(0.04)	0.06	0.13
	Indiana	0.72	0.77			0.14
	Iowa	0.71	0.81			(0.04)
	Louisiana	0.42	0.49	0.50	0.55	0.05
	Mississippi	0.33	0.51	0.68	0.76	0.37
	Nebraska	0.82	0.67			0.23
	North Dakota	0.26	0.63			0.44
	Wisconsin	0.70	0.69			0.31
ADM1 Only	Illinois	0.56	0.61			0.21
	Indiana	0.74	0.78			0.15
	Iowa	0.73	0.81			(0.04)
	Louisiana	0.52	0.62	0.56	0.56	0.05
	Mississippi	0.34	0.76	0.68	0.76	0.30
	Nebraska	0.91	0.71			0.31
	North Dakota	0.83	0.79			0.55
	Wisconsin	0.73	0.71			0.33

Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015), the Cropland Landscape Layers (USDA - NASS 2004, 2005, 2006) and own calculations.



Source: Developed by author using data from the Cropland Landscape Layers (USDA - NASS 2004, 2005, 2006) and Fritz, et al. (2015).

Figure 2-5: Pairwise comparison between NASS and SPAM cropland estimates

Chapter 3: Proximity to Agricultural Markets in Sub-Saharan Africa

3.1 Introduction

Much of agricultural production in sub-Saharan Africa is characterized by highly fragmented, small-scale farming operations (most less than 5 hectares) with limited and uneven participation in off-farm market transactions (Fafchamps 2004, de Janvry, Fafchamps and Sadoulet 1991). Many of the factors affecting agricultural production decisions, not least the frequency, extent and nature of off-farm market participation, vary spatially. Thus, strategic policies, investments and development initiatives aimed at improving the performance of African agriculture can benefit from spatialized measures of these factors. This chapter focuses on exploring the nature of market isolation within sub-Saharan Africa, specifically for agricultural producers.

Reducing the isolation of farms can spur improvements in agricultural productivity and economic well-being in a host of ways, especially for small-scale famers in developing countries (Stifel and Minten 2008, Stifel, Minten and Dorosh 2003, Elbers, Lanjouw and Lanjouw 2003). Improved market accessibility for farmers makes it easier and more profitable to obtain yield-enhancing inputs, such as chemical fertilizer and improved seed, and promotes commodity market participation which can help mute commodity price volatility for consumers and producers alike. However, buying and selling in markets is costly. Farmers must search for buyers and sellers, negotiate acceptable terms, and physically transport goods to and from markets. These off-farm market participation activities are not free (explicitly or in terms of the opportunity cost of farm household members), which means that the effective cost of agricultural inputs to farmers is higher and the effective value of farm outputs lower than the corresponding (local) market prices. Measures of market accessibility help determine the nature of effective prices and the farm input and output decisions influenced by these prices.

Yoshida and Deichmann (2009, 3) defined accessibility as the "ability for interaction or contact with sites of economic or social opportunity." While this definition encompasses many facets of accessibility including spatial proximity, affordability, acceptability and availability, spatial proximity to agricultural markets is arguably the most important influential factor of market accessibility on farmers, and the focus of most studies of market access in developing countries (Apparicio, et al. 2008, Yoshida and Deichmann 2009). Commonly used measures of spatial proximity to markets are the distance- or time-based metrics from a starting location (e.g., a household, plot, or pixel) to some location that proxies for economic opportunity (e.g., road, urban center or agricultural cooperative outlet) (Chamberlin and Jayne 2013, Wood 2007). Within sub-Saharan Africa, many contemporary surveys of agriculture elicit measures of market proximity. For example, the 2008/09 National Panel Survey in Tanzania asked farmers to recall the distance traveled to sell their crops and the costs associated with that travel (NBS 2011b). This survey also collected each household's GPS coordinates and subsequently calculated the Euclidean distance (i.e., straight line distance) to key market centers.[14] These types of market proximity indicators are beneficial because they link household-specific decisions regarding agricultural production with the measures of spatial proximity relevant to that particular household. However, most survey statistics are only representative at a national or first sub-national administrative level, and cannot provide a more granular sense of the spatial heterogeneity of accessibility within a region.

Applying modeling methods to rasterized data[15] enables researchers to measure spatial proximity in terms of physical-, time- or cost-distance to a market across relatively expansive areas and with finer scales of resolution within those areas rather than relying solely on the market access measures captured via the survey methods mentioned above. These raster-based measures are particularly helpful when seeking to answer questions regarding the geographical variation in market proximity, say, at a regional scale (e.g., Western versus Eastern Africa). This chapter uses rasterized data on travel time to markets

---

[14] Key market centers were identified by USAID FEWS NET, as demonstrated by the Production and Market Flow Map for Tanzania accessible from: http://www.fews.net/sites/default/files/documents/reports/tz_fullmap_maize_norm.pdf.

[15] Raster data consist of a matrix of regularly sized cells (or pixels) that are organized into rows and columns. Each pixel has an explicit spatial boundary, can be represented by its two-dimensional coordinate location, and contains a single value. In models used to calculate spatial proximity, this information represents the cost of travel through that pixel (e.g., slope, speed, monetary cost) and is referred to as a friction surface. Least-cost functions calculate multiple paths of travel between a starting point and a destination, and return the path with the lowest 'cost.'

of varying size to characterize the nature of market isolation for agricultural producers in sub-Saharan Africa.

There are few estimates of market proximity available for the entire sub-Saharan African region. Among the most widely used estimates are those created by Nelson (2008), who estimated travel time to cities with populations of at least of 50,000 people globally.[16] However, these estimates were formed using assumptions appropriate for a global context and are only relative to a single market size. Staple foods, such as maize and cassava, are often sold in smaller, local markets, while cash crops are more often sold internationally from larger markets. Thus a notion of market proximity to varying sized markets is useful for studying the variety of market participation decisions confronting farmers. Guo, Joglekar and Beddow's (forthcoming) estimates of travel time examined in this chapter were constructed using an approach similar to that used by Nelson (2008) but incorporate updated measures of road networks, land cover and city population. Additionally, Guo, Joglekar and Beddow's (forthcoming) estimates report travel time to markets throughout sub-Saharan Africa with populations of at least 20,000, 50,000, 100,000, 250,000 or 500,000 people.

The purpose of research related to agricultural production varies, for some the relevant attribute is the area in agriculture. For others, the market proximity of agricultural production, and in some instances the proximity of the agricultural labor force to population centers of varying sizes is of interest. I examine the estimates of market proximity with regard to each of these three variables and find they result in markedly different views of market isolation. Additionally, most small-scale famers reside in rural areas lacking extensive road networks. Thus, the nature of time spent traveling off-road is especially relevant for this group of farmers. I find there is substantial variation in the amount of time spent off-road exists, regardless of proximity to market. Thus, efforts to diminish the effects of market isolation for small-scale farmers may be well served by

[16] Verburg and Letourneau (2011) also make available pixelated estimates of market accessibility on a global scale, but their measure consists of a single index of travel time to national and international markets, and is more conducive to answering questions involving the influence of international trade than local markets.

focusing on efforts that reduce the time spent traversing on footpaths and tracks that lead to established road networks.

3.2 Accessing Agricultural Markets in Sub-Saharan Africa

The literature cites a host of inadequacies and imperfections that befall agricultural markets throughout sub-Saharan Africa, including excessively high transportation costs from underdeveloped infrastructure, missing or poorly functioning land and labor markets, weak bargaining power on the part of farmers and, relatedly, difficulties in enforcing contracts, high production risk, poor access to credit and asymmetric information (Dillon and Barrett 2014, de Janvry, Fafchamps and Sadoulet 1991). A recurring theme that contributes to many of these transactional problems is the remoteness of farms relative to markets (Key, Sadoulet and de Janvry 2000, de Janvry, Fafchamps and Sadoulet 1991, Omamo 1998a, 1998b). Mobile phones (Aker and Mbiti 2010, Aker 2010) have the potential to overcome some of the obstacles of market remoteness, but farmers are still confronted with the physical realities involved in hauling (often bulky) agricultural goods to and from markets.

The neoclassical and new institutional economics conception of transaction costs refers to the non-price costs of trade (e.g., time spent searching for a buyer or seller, negotiating and enforcing contracts), but in a broader sense they can also refer to the explicit and implicit costs associated with transportation, storage, processing, packaging, and so on. These costs influence the nature of market accessibility and the extent to which farm households throughout sub-Saharan Africa choose to participate in markets (if at all) by raising the effective cost of inputs faced by farmers and lowering the effective (or farm-gate) price they receive for their outputs (Renkow, Hallstrom and Karanja 2004, Staal, Delgado and Nicholson 1997, Alene, et al. 2008, Barrett 2008, de Janvry, Fafchamps and Sadoulet 1991).

The agricultural production literature is primarily focused on physical measures of access that deal with geographical closeness or quantity of services within a defined area. In a review of this literature, Wood (2007) groups these measures into three categories: distance, infrastructure and “other”. *Distance* measures define market proximity in terms of the “closeness” between one location (e.g., household, plot or pixel) and another (closest road, railway, urban center or market) measured in distance- or time-metrics. Jacoby (2000), for example, found a negative relationship between self-reported travel time to an agricultural cooperative or market center and farmland values in Nepal. *Infrastructure* measures account for the existence or density of services within a defined region. For example, Demeke, et al. (1998) found that Ethiopian farmers who lived in areas containing better road infrastructure tended to use more chemical fertilizer. *Other* measures included the use of categorical locations or regional dummies. Using such measures, Staal, Delgado and Nicholson (1997) found that proximity to Nairobi (as indicated by the relative proximity of the surveyed individual’s residing district) negatively impacted an individual’s decision to sell milk to the parastatal dairy cooperative. As geographic information systems (GIS) technologies become more commonplace, it becomes easier and cheaper to calculate geographically explicit measures of proximity.

3.3 Data

The time-to-market estimates developed by Guo, Joglekar and Beddow (forthcoming) are the focus of the assessments carried out in this chapter. They were calculated using a cost-distance function in ArcGIS to determine the minimum time cost of traveling from the centroid of each pixel to the nearest market or service location, where markets were defined as the centroid of a human settlement that meets one of five population thresholds (specifically, either 20K, 50K, 100K, 250K or 500K people or more).[17] The friction surface for this function was created using information on national road networks, land cover, and elevation layers. Time costs were determined using assumed vehicle travel speeds along different classes of roads and walking travel speeds across various classes of land cover. Elevation and, consequently, slope, were included as speed-reduction factors. Guo, Joglekar and Beddow’s (forthcoming) time-to-market layers were originally reported at an equal-area one kilometer squared resolution, which were then rescaled to a 1 arc-minute pixel (approximately two kilometers at the equator) for compatibility with other data used in this analysis.

[17] See Guo, Joglekar and Beddow (forthcoming) for more detailed documentation of the data set.

The baseline road network used in the model was a revised (though undocumented) version of the Vector Map Level 0 (VMap0) data obtained from the International Food Policy Research Institute (IFPRI). VMap0 is a widely used data set representative of the global road network circa 1992 and provided by the Environmental Systems Research Institute (ESRI) (NIMA 2000). The revised VMap0 network was first updated using unpublished road network data acquired through the World Bank from Michelin's regional map series. Additional efforts were also made to update road network data for eleven countries[18] using information and expert knowledge acquired from a variety of international collaborators (e.g., the World Bank and African Development Bank), centers within the Consultative Group on International Agricultural Research (CGIAR), universities and partners with on-the-ground knowledge of African roads. The updated road network data was validated against Google Earth and Bing Maps to address any inconsistencies.

The finalized road network layer identified three classes of roads throughout sub-Saharan Africa: primary (9.7 percent of total road length), secondary (22.7 percent) and tertiary (67.6 percent) roads.[19] These breakdowns differed slightly between regions:[20] Central Africa had the highest percentage of tertiary roads (70.9 percent), while Southern Africa contained the lowest (63.5 percent). However, it is difficult to know with certainty the extent to which a primary road in Chad is comparable with a primary road in South Africa. To incorporate the road network information into the cost-distance function calculations, the network must first be rasterized. If two roads of different classes existed

[18] Specifically, Mali, Kenya, Malawi, Senegal, Uganda, Nigeria, Tanzania, Burundi, Ethiopia, Ghana and Mozambique.

[19] Folding in these new data led to significant changes to the original VMap0 road network layer. Some new roads were added, and the placement of some roads was modified, but much of the change involved the reclassification of roads, and most often reclassification of a 'secondary' road into a 'tertiary' road. VMap0 originally classified 4.8 percent of its roads as primary (roughly half the corresponding road length compared with the updated data), 93.6 percent as secondary and 1.5 as other. While the exact definition of road segment classes varies by country, it is generally accepted that primary roads connect primary networks and are kept in good condition, secondary roads connect secondary networks and are kept in fair condition, and tertiary roads connect peripheral areas and may be in poor condition (Buys, Deichmann and Vheeler 2006, Gwilliam, Bofinger, et al. 2011).

[20] For this analysis, regions in sub-Saharan Africa are defined according to the UN classification system, with the exception that Sudan is grouped with Eastern Africa rather than Northern Africa http://esa.un.org/wpp/Excel-Data/country-classification.pdf .

within a single pixel, the pixel was classified according to the "highest quality" road class within that pixel (i.e., primary took precedence over secondary which took precedence over tertiary roads).

The land cover data used to construct the time-to-market estimates were taken from the GlobCover 2009 data developed by the European Space Agency (ESA) (Arino, et al. 2012). GlobCover classified satellite based-images according to 22 major land classes and four sub-classes to represent tree cover, shrub cover, herbaceous cover (including cultivated and managed areas), barren cover, urban areas, mosaic areas and water, and was reported at a 10 arc-second grid (approximately 300 meters at the equator). This data layer was chosen because it contained the most recent data with the highest resolution and complete coverage of sub-Saharan Africa.

The elevation data came from the U.S. Geological Survey (USGS) Shuttle Radar Topography Mission (SRTM) 90m Digital Elevation Database (DEM) (version 4.1) provided by the CGIAR's Consortium for Spatial Information (CGIAR-CSI) (CGIAR-CSI 2008, Jarvis, et al. 2008) and was reported at a 3 arc-second grid (approximately 90 meters at the equator). Slope was calculated from the elevation data using the Slope Analyst routine in ArcGIS. Elevation and slope were used to modify speeds with formulas similar to those deployed by Nelson (2008). Travel speeds were reduced for elevations above 2,000 meters, which accounted for just 1.3 percent of the 7.2 million (populated) pixels throughout sub-Saharan Africa. Sloped terrain affected travel speeds through an exponential downscaling of the (assumed) speed over flat terrain as the gradient of a pixel increased.

The friction surface used in the cost analysis was formed by first overlaying the road layer on the land cover layer. Each pixel was assigned a singular time cost based on the pixel's classification (e.g., primary road, mosaic cropland).[21] The main speeds used are

[21] Guo, Joglekar and Beddow's (forthcoming) time-to-market estimates do not account for travel by water or rail. While Nelson (2008) includes options for water and rail travel when forming his global estimates of travel time, other market proximity estimates (e.g., Pozzi and Robinson (2008)) also opted to set aside rail and water as a mode of transport. Both rail and shipping are quite costly in sub-Saharan Africa, and often, the time costs associated with loading and unloading goods are significant enough to make road travel a more desirable transportation option. Additionally, many of the region's rail networks have been poorly maintained, and do not offer a feasible transport option for (perishable) agricultural goods (Gwilliam, et al.

detailed in Table 3-1; speed assumptions were based on personal communication with Dr. Nelson regarding his global travel time estimates for the Joint Research Council (Nelson 2008). Given the friction surface, the cost-distance function calculates all possible paths between the centroid of the starting pixel and the centroid of the target market and returns the minimum time of travel, for every pixel in sub-Saharan Africa.

[Table 3-1: Speed and associated cost assumptions for friction surface]

Guo, Joglekar and Beddow's (forthcoming) time-to-market estimates were compiled for five market sizes: 20K, 50K, 100K, 250K and 500K people or more. To specify market geography, settlement locations were defined as the centroid of each urban extent reported by the Global Rural and Urban Mapping Project, Version 1 (GRUMPv1) (CIESIN, IFPRI, the World Bank and CIAT 2011, Balk, et al. 2006). Point data obtained from the World Gazetteer (a crowd-sourced source of population count data) were then used to update the GRUMPv1 settlements points to reflect estimates of urban population circa 2010 (World Gazetteer 2010).[22] Given these population data, it was estimated that there were 2,151 markets with populations of at least 20K people and 112 markets of at least 500K people throughout sub-Saharan Africa (see Table 3-2).

[Table 3-2: Number of populated pixels and markets in SSA]

The final estimates of time-to-market to pixels with a positive population (i.e., areas such as lakes are removed) are mapped in Figure 3-1[23] for three of Guo, Joglekar and Beddow's (forthcoming) layers (specifically, markets of 20K, 100K and 500K people or

2011). The time-to-market estimates also do not include a speed reduction factor for border crossings. Other studies have accounted for delays at international borders by assigning a specific, but often a seemingly arbitrary, time delay at the border (e.g., one hour to cross a one kilometer pixel) (Pozzi and Robinson 2008, Hartley, et al. 2007). Ignoring the time cost implications of border crossings may underestimate the travel time to particular markets.

[22] The updated layer of settlement points has not been formally validated, and some cities may be missing or misclassified. It is likely that any such measurement problems disproportionately affect the smaller settlements.

[23] Maps of travel time to markets of 50K and 250K people are presented in Guo, Joglekar and Beddow (forthcoming). To better highlight the market proximity of the continent's rural population, the travel times mapped in Figure 3-1 include only time-to-market estimates for pixels with a positive population. Estimates of the total number of people per grid square across Africa were taken from the WorldPop Project, Version 1.0, which is publically provided at a 30 arc-second resolution (WorldPop 2015). These data ostensibly represent the 2010 population, and so were recalibrated such that the sum of the population across all pixels within a country equaled the corresponding 2010 country total population reported by (United Nations 2012).

greater), and descriptive statistics for all five layers are provided in Table 3-3. As expected, travel times increase when seeking to access ever-larger markets. Across the continent, the average travel time to markets of at least 20K people is 14.0 hours; increasing to 16.6 hours for markets of 100K people or more and 20.2 hours for a 500K person market.[24] While there are some areas in the Sahara desert and the inner regions of the Democratic Republic of the Congo that have estimated travel times in excess of one week, these areas are minimally populated. In fact, 98.2 percent of Africa's population lives within one day of travel to a city of at least 20K people or more. There do appear to be significant regional differences in the landscape of travel time: Central Africa is the most remote overall, while Southern Africa is the least. However, in this analysis, I am concerned with understanding the nature of market isolation facing agricultural producers (specifically, crop producers) in sub-Saharan Africa, and a sizable amount of land in sub-Saharan Africa is not used for crop production.

[Figure 3-1: Time-to-market from all populated pixels]

[Table 3-3: Descriptive statistics for time-to-market estimates from all populated pixels]

Figure 3-2 maps the time-to-market estimates from pixels with positive crop production, as defined by the footprint of total harvested area from the SPAM2005 (Wood-Sichra, Joglekar and You forthcoming) data discussed in the previous chapter of this dissertation, to markets of 20K, 100K and 500K people or greater. Descriptive statistics for travel time to all five market cutoffs are provided in Table 3-4. On average, the travel time from cropland pixels to a city of at least 20K people is 8.4 hours (40.0 percent less than the time to markets of equivalent size for all populated pixels), 10.2 hours to a 100K person market (38.6 percent less than the time from all populated pixels) and 13.7 hours to a 500K person market (32.2 percent less than time from all populated pixels). Regional differences remain, but when only cropland is considered, Western Africa is the least remote region in sub-Saharan Africa (4.5 hours to a city of 20K or more on average), while Central Africa is still the most remote (11.6 hours to a city of 20K or more on average).

[Figure 3-2: Time-to-market from all cropped pixels]

[24] The average travel time to a city of 50K people as classified by HarvestChoice is 15.4 hours, this is 34 percent higher than the corresponding average reported by Nelson (2008).

[Table 3-4: Descriptive statistics for time-to-market estimates from all cropped pixels]

3.4 Evaluation

Modeled, pixelated measures of travel time represent a comparatively new way to assess (agricultural) market proximity. While they are able give a more granular, sense of proximity over large areas, they require substantial resources to calculate. This begs the questions: Do the particular estimates by Guo, Joglekar and Beddow's (forthcoming) generate plausible results, and how do such market proximity measures stack up against alternative (less resource demanding) estimates?

3.4.1 Assessment of Principal Assumptions

There are three principal assumptions underlying Guo, Joglekar and Beddow's (forthcoming) market proximity estimates, namely:

1. Travel times from a pixel to a population center provide a sensible and useful measure of market proximity.
2. Farmers (and rural populations more generally) engage in market based transactions at population centers of 20,000, 50,000, 100,000, 250,000 and 500,000 people or more.
3. Constraints on travel are (spatially and temporally) homogenous.

For highly localized transactions, farm-to-farm, or farm-to-local village transactions may incur transactions cost that are best represented by pixel to pixel (or even within pixel) market proximity measures. Be that as it may, the transactional economics of farmers sourcing inputs from non-farm, input suppliers are arguably better represented by the pixel to population proximity measures presented and discussed here. Similarly, selling outputs into off-farm markets also involves farm to population center value chains, and is likely best represented by the metrics discussed in this chapter.

The decision to access off-farm agricultural input or output markets is affected by a host of transaction costs including, the transport costs to the market, certainty of finding a buyer or seller, quantity of purchase or sale, expected price, terms of trade, ability to store unpurchased commodities and the nature of the commodity itself (Jagwe and Machethe 2011). Fixed transaction costs such as search, negotiation and enforcement costs are related to issues of information attainment, while variable transaction costs are affected by measures of distance and the mode of transport (Alene, et al. 2008). Generally, it is asserted that both these types of transaction costs increase with remoteness (Key, Sadoulet and de Janvry 2000, Omamo 1998b, 1998a), and are compounded to the extent that the quality and quantity of road and communication infrastructures decrease as physical distance from urban settlements increases (Gwilliam, et al. 2011).

The second principal assumption involves specification of the relevant market, and, in particular, the usefulness of the specific population centers used to form Guo, Joglekar and Beddow's (forthcoming) estimates. The nature of off-farm transactions likely varies among population centers of different sizes. Thus, by imputation, the frequency of these off-farm transactions is also likely to differ by market distance and size. For example, staple foods such as maize and cassava are more likely sold into nearby local villages, whereas export oriented crops are more likely sold into, or via market intermediaries, to larger, perhaps more distant, markets. Absent detailed data on the nature and number of off-farm transactions, it makes good analytical sense to present market proximity measures for a range of market sizes. While local market transactions occur within population centers of less than 20,000 people, the prospects of (accurately) identifying the location of these cities throughout sub-Saharan Africa with the presently available data are limited.

The final estimation assumption is, perhaps, the most questionable. For methodological tractability, pedestrian and vehicular travel speeds are fixed depending on the type of land cover or road class. One might imagine a number of factors that could affect the speed of on-road travel, including things like the actual mode of transport (e.g., a bicycle, car, bus or truck), the quality of the road, and the time-of-day or other factors that affect road congestion. However, the assumption of homogenous travel not only eases computational requirements but is a compromise that ensures the generalizability of the resulting proximity surfaces. The approach implemented by Guo, Joglekar and Beddow (forthcoming) is meant to reflect the relative proximity to various markets from a large-scale (e.g., country or region), time-invariant perspective. While there are certainly pixel-to-pixel or within pixel differences in wealth and, thus, the ability to access various modes of transportation, Guo, Joglekar and Beddow's (forthcoming) market proximity layers represent a baseline scenario. Depending on the analyses, wealth and seasonality may be controlled for with external data. Additionally, while congestion effects are not reflected explicitly in these estimates, road quality has been addressed by differentiating roads into three different classes (specifically, primary, secondary and tertiary roads) and assigning different road speeds accordingly.

Regardless of the mode of travel, the quality and density of road infrastructure available will likely reflect the major differences in regional proximity to larger agricultural markets. Thus, the road network used to form the friction surface is arguably the most important input to forming a proximity layer. There is no single authoritative source of digitized data for the African road network—different sources show different roads, with varying levels of accuracy across locations, and represent different points in time. So while much effort went into updating the road network layer used to form Guo, Joglekar and Beddow's (forthcoming) proximity layers, there are still likely to be omissions, especially in the tertiary and rural feeder roads. It is assumed that the majority of primary and secondary roads have been identified. If parts of the road network are missing and these roads are used for vehicular travel, then the estimates of travel time will be biased upward.

A sensitivity analysis of the road speed assumption shows that decreasing the assumed road speeds to 50 km/hr for primary roads (33.3 percent decrease), 35 km/hr for secondary roads (42.7 percent decrease) and 25 km/hr for tertiary roads (16.7 percent decrease) increased the average travel time from agricultural land to a market of 50K or more by 15.7 percent (1.5 hours). This result differed by region: the road speed decreases resulted in a 15.8 percent (2.1 hours) increase in the average travel time from cropland to a market of 50K in Central Africa, 14.0 percent (1.4 hours) in Eastern Africa, 22.5 percent (1.3 hours) in Southern Africa and 18.3 percent (0.9 hours) in Western Africa.

3.4.2 Alternative Metrics of Market Proximity

Guo, Joglekar and Beddow's (forthcoming) time-to-market metrics rely on a network-based algorithm, which is comparatively costly to calculate, both in terms of data collection as well as computational effort and complexity. For this reason, several studies use the "as the crow flies" measure of the Euclidean distance between two points as an indicator of proximity:

$$d_{ij} = \sqrt{(x_i - x_j)^2 - (y_i - y_j)^2}.$$

where d_{ij} is the distance between point i (x_i, y_i) and point j (x_j, y_j). It is argued that network-based proximity measures provide a better approximation of the true costs of travel over a Euclidean distance because they account for differences in the ease of travel and more likely approximate the path of travel. Given the minimal data requirements necessary to calculate a Euclidean distance, it is worth examining the nature (in terms of absolute and relative proximity to markets) of the differences in estimates of proximity generated by these different algorithms.

The Euclidean distance-to-market and time-to-market metrics for agricultural land in sub-Saharan Africa were only moderately correlated: the Pearson correlation coefficients ranged from $\rho = 0.38$ for travel to a 20K person market to $\rho = 0.50$ for travel to a 500K person market. These correlation coefficients appear to be substantially influenced by the pixels in Southern Africa (see Figure 3-3); for other regions within sub-Saharan Africa the concordance between these two metrics is much lower.[25] In this region, the majority of cropland was located in South Africa, an area that is comparatively flat with an extensive rural road network relative to the rest of sub-Saharan Africa. Both these factors make it more likely that the network and Euclidean travel paths to market are similar, resulting in higher correlations between the market proximity estimates derived using each of the metrics. The lowest correlations were in Central and Western Africa. In Central Africa, the Congolese forests act as a barrier to travel and in Western Africa the areas of poor correlation were both on the coastline of Cote d'Ivoire where there were several bays – water travel was set prohibitively high – and in pockets of the Saharan dessert where few roads existed. As suspected, geography plays a large role in determining how well a straightforward Euclidean distance metric approximates the analytically more demanding network metric of market proximity.

[25] For travel to a 20K person market, the Pearson correlation coefficient is $\rho = 0.53$ in Southern Africa, $\rho = 0.41$ in Eastern Africa, $\rho = 0.28$ in Central Africa and $\rho = 0.24$ in Western Africa. While the correlation coefficients increase across all regions when traveling to larger markets, the largest increase is in Southern Africa ($\rho = 0.86$ versus $\rho = 0.38$ in Central and Western Africa) (see Figure 3-3).

[Figure 3-3: Pairwise comparison between Euclidean distance- and time-to-market estimates]

3.5 Results and Discussion

3.5.1 The Nature of Market Proximity

The appropriateness of one market proximity measure over another is dependent on the question at hand. For some purposes the relevant attribute is the market proximity to agricultural areas or, alternatively, agricultural production (by quantity or value), while in other instances the proximity of rural populations (or the agricultural labor force) to settlement centers of varying sizes will be of interest. Answers to other questions may hinge on the (spatial) interrelationships between agricultural areas (or parts thereof, for example, the areas in corn or dairy production), agricultural output (or parts thereof), rural population and market proximity.

Table 3-5 presents the percentage of cropland pixels, persons and total value of crop production by category of travel time to a market of 20K people or more.[26] According to Guo, Joglekar and Beddow's (forthcoming) proximity layers, the majority of cropland is located within four to eight hours of a 20K person market and nearly 90 percent of cropland is within 18 hours of a 20K market. There is a markedly different view of market proximity when population or value of production are taken into account. It was estimated that 54.7 percent of sub-Saharan Africa's total population lives within two hours of a 20K market, while 90.0 percent of the population lives within eight hours.[27] In terms of where crops are produced, 61.6 percent of the total value of crop production is located within four hours of a 20K market and over 20 percent of the value of crop production is located more than eight hours from a market of the same size.

[26] In sub-Saharan Africa, pixel area ranges from 5.7 square kilometers to 8.6 square kilometers. Total population was calculated using the WorldPop (2015) estimates of population count within each pixel. Total value of production is calculated using the SPAM2005 (Wood-Sichra, Joglekar and You forthcoming) data described in Chapter 2. For each pixel, estimated crop production is multiplied by the relevant crop price and summed across all 42 crops and crop categories. SPAM2005 is reported at a 5 arc-minute resolution, so these total value of crop production estimates are disaggregated (by a factor of 5) to 1 arc-minute pixels to be matched with the travel time data.

[27] Reliable, gridded estimates of population broken down by rural and urban classifications were difficult to obtain, but the population living outside of two hours travel from a population center of 20K people are likely rural. Of this this population, nearly 90 percent reside within 12 hours of a 20K market.

[Table 3-5: Percentage of pixels, persons and value of crop production by remoteness]

Again, there are significant regional differences in the degree of isolation in sub-Saharan Africa, as viewed in relation to the location of cropland, people, and value of crop production. Central Africa is the most isolated region in sub-Saharan Africa with nearly 90 percent of its cropland pixels located within one day of travel to a 20K market. This proximity statistic falls to 18 hours with regard to the location of the value of crop production and 12 hours with regards to population. Regardless of the lens, Western Africa is the least isolated region where nearly 90 percent of the pixels, population and value of production are within eight hours of a 20K market.

A natural question arises in this context: specifically, is agriculture in sub-Saharan Africa substantially farther from the markets it serves than agriculture elsewhere in other regions of the world? The limited assessment done to date using proximity estimates compiled by Nelson (2008), suggests that, as expected, average travel times for crop production to markets in sub-Saharan Africa are at the higher end of the spectrum, but not inordinately so. For example, the average time to transport crop production (measured in value) to a market of 50K people is 5.4 hours in Brazil, 2.9 hours in India and 4.9 hours in China.[28] The corresponding average throughout sub-Saharan Africa is 5.1 hours, although for major producing countries such as South Africa and Nigeria the travel times are lower, averaging 3.6 and 4.3 hours respectively.

3.5.2 The Last Kilometer

Porter (2002) highlighted that much of the development literature has paid little attention to the accessibility concerns of marginalized, rural populations throughout sub-Saharan Africa, many of whom live “off-road” and for whom walking is the primary, if not only, means of transportation. The tracks and pathways that connect individual farms to local road network are often unrecorded, which means they will not influence the proximity measures in this analysis. The length of these unrecorded tracks and pathways in sub-Saharan Africa is estimated to be “one and a half to two times the local government road networks” (Gwilliam, Bofinger, et al. 2011, 22).

[28] These statistics were calculated using Nelson’s (2008) global estimates of travel time to a market of 50K people or more.

Guo, Joglekar and Beddow's (forthcoming) estimates reveal a sizable but varying time-to-market disconnect between where agriculture takes place and the small towns and growing cities where increasing amounts of agricultural consumption is projected to occur. For example, the United Nations (2014) estimates that 37.9 percent of the population in sub-Saharan Africa currently resides in urban areas, but that this will grow to 54.8 percent by 2050. With agricultural consumption projected to move off-farm at a rapid rate, getting produce from farm to settlement areas will become an increasingly pressing problem. Developing a more refined sense of the structure of the time-to-market impediments has obvious and increasingly important investment and policy value. While the market participation implications of undocumented (local) roads and pathways cannot be explicitly assessed, we can assess the implications of travel along primary versus secondary versus tertiary roads for the time to market. Which begs the question, is the biggest impediment to farmers in accessing a market the travel time along the more established roadways, or is it a last kilometer problem, that is, the time taken in simply getting from their farm to a road of any sort?

Guo, Joglekar and Beddow's (forthcoming) time-to-market metrics can be divided into a pedestrian (off-road) and vehicular (on-road) component by setting the road speed assumption for all three classes at extremely rapid rates (i.e., 500 km/hr) and re-running the model. The resulting time to market estimates thus approximate the time of off-road travel only, and are mapped (as a percentage of total travel time to a market of 20K people or more) in Figure 3-4. The largest fraction of time spent traveling off-road from pixels of cropland took place on the coast of Western Africa, in the Congolese forests, and through the eastern regions of Zambia and Mozambique and Madagascar. Off-road travel is the defining reality for the majority of agricultural pixels, population and value of production in sub-Saharan Africa. In fact, 60.5 percent of cropland pixels throughout sub-Saharan Africa are located such that more than 50 percent of the time spent traveling to a market of at least 20K people is spent off-road. Similarly, 53.8 percent of the population and 62.8 percent of crop production (measured in value) will spend over 50 percent of the time off-road on its way to a market of 20K or more. Table 3-6 shows that this is not just characteristic of the very remote areas. Even areas classified as relatively close to urban settlements can have production that will spend a substantial large share of the total travel time to market traversing off road terrains.

[Figure 3-4: Fraction of time spent traveling off-road]

[Table 3-6: Value of production (I$/mt) by remoteness and fraction of travel spent off-road]

Table 3-5 shows that one-third of sub-Saharan Africa's total value of crop production is located within two hours of a 20K market, meaning that two-thirds of the output value is located more than two hours away from even a relatively small township or city market. Table 3-6 shows that 43.5 percent of this agricultural output (measured in value) will be transported off-road for 40 percent of the time or more. Notably, as overall remoteness increases, so does the share of total travel time spent traveling off-road. For the output value located greater than one day travel from a 20K market, 91.2 percent of the travel is estimated to be off-road. Regardless of relative remoteness, 29.6 percent of agricultural output value in sub-Saharan Africa is transported off-road to a market of 20K people for 80 percent or more of the trip. There are regional differences with regard to the time spent traveling off-road, but they are less dramatic than the differences in total travel time between regions; specifically, 59.4 percent of the value of production in Central Africa, 58.5 percent in Eastern Africa and 59.8 percent in Southern Africa is estimated to spend over 50 percent of the time off-road getting to markets of 20K people or more. In Western Africa, 66.5 percent of the total value of production is located in areas where at least 50 percent of the time spent traveling to a market of 20K is off-road.

Over 75 percent of sub-Saharan Africa's population spends more than 40 percent of their time traversing off-road to get to a market of 20K people or more. This is a significant portion of the population that is often overlooked in the development literature, literally because they are less visible, a phenomenon dubbed "tarmac bias" (Chambers 1983, 13-16). Intermediate traders may be less willing to service this population, and those that do have higher bargaining power because households are less knowledgeable about current market prices, or may take a lower offer price for their produce if they are severely cash strapped. Beyond building or improving the quality of rural feeder roads, which may be costly, there are several other options that can help alleviate the burdens of rural populations living off-road. These may include increased telecommunications infrastructure, community owned transport or low-cost storage and processing technologies (Porter 2002). Other strategies may include the promotion of "intermediate means of transport" (IMT) such as wheelbarrows, bicycles, animal drawn carts and bike trailers and motorcycles (Porter 2002).

3.6 Conclusions

Increasing market participation is critical for economic growth and poverty reduction, especially for small-scale farmers (Barrett 2008). Since market participation is tied to the transaction costs of accessing these markets, the location of farms and their physical and economic proximity to markets have a whole host of agricultural production and consumption implications that profoundly affect the economic circumstances of farm families. Understanding the spatial nuances of these potential transaction costs will help tailor future strategies aimed at transforming agriculture.

The pixelated estimates of travel time throughout sub-Saharan Africa created by HarvestChoice are based on several simplifying assumptions that should be kept in mind when using these data. On balance these analytical simplifications are likely to result in travel times that are shorter than reality, but they should yield estimates that are sufficiently robust for assessing overall (relative) patterns of proximity within countries or regions.

At first glance, it seems as though most of sub-Saharan Africa is rather remote. Central Africa is the most remote region, while Southern and Western Africa are the least, depending on the market size referenced. However, the view of market proximity markedly changes when examined with respect to cropland, where the majority of the population resides and where crop value is located. The average time spent traveling to a market of at least 20,000 people is 8.4 hours on average when assessed in terms of the location of crop production, dropping to 4.6 hours when weighted by total value of crop production, and to just 3.4 hours when weighted by total population.

I find that a significant portion of sub-Saharan Africa's value of crop production spends the majority of its time traveling to market off-road. In sub-Saharan Africa, there is substantial variation in expenditures on road network maintenance, both by country and road class (Gwilliam, Bofinger, et al. 2011). On average, countries spend twice as much on maintaining main road networks than rural networks (Gwilliam, Bofinger, et al. 2011) which can help move agricultural output into urban areas, but may well by-pass the off-road population, at least in the near term. Thus, complementary strategies such as improved telecommunications infrastructure, community owned transport, low cost storage and processing technologies, or increased access to IMTs may be better ways of alleviating the burdens of rural isolation and be more influential in bringing African agriculture "closer" to the market.

Table 3-1: Speed and associated cost assumptions for friction surface

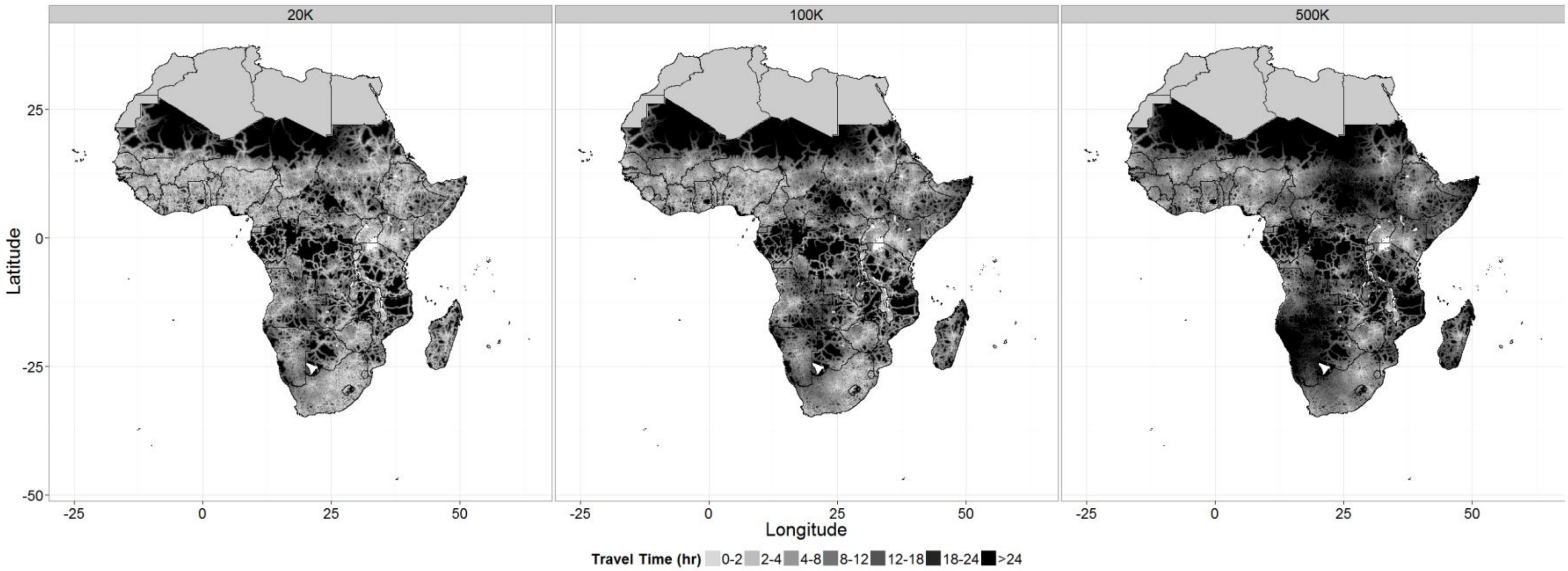
	Average speed (km/hr)	Average cost (min/km)
Road Class		
Primary roads	75.0	0.8
Secondary roads	60.0	1.0
Tertiary roads	30.0	2.0
Land Cover Type		
Tree cover	1.2	51.0
Shrub cover	1.7	36.0
Herbaceous cover (including cultivated and managed areas)	1.7	36.0
Barren cover	2.5	24.0
Urban areas	30.0	2.0
Mosaic areas	1.6	38.0
Water	0.3	180.0

Note: A complete list of land cover classifications and speeds can be found in Guo, Joglekar and Beddow (forthcoming).

Table 3-2: Number of populated pixels and markets in SSA

	Number of Populated Pixels (million)	Number of Markets (by market size)				
		20K	50K	100K	250K	500K
Sub-Saharan Africa	**7.2**	**2,151**	**1,015**	**555**	**219**	**112**
Central Africa	1.9	273	125	71	31	16
Eastern Africa	2.6	771	294	154	62	32
Southern Africa	0.7	350	197	106	35	17
Western Africa	1.8	757	399	224	91	47

Source: Developed by author using population data from the World Gazetteer (2010) and GRUMP (CIESIN, IFPRI, the World Bank and CIAT 2011).



Source: Developed by author using travel time data from Guo, Joglekar and Beddow (forthcoming) and pixelated 2010 population data from the WorldPop project (WorldPop 2015).

Note: Travel time to markets of 20K, 100K and 500K people or more are presented. Pixels with zero population have been masked (in white) from maps.

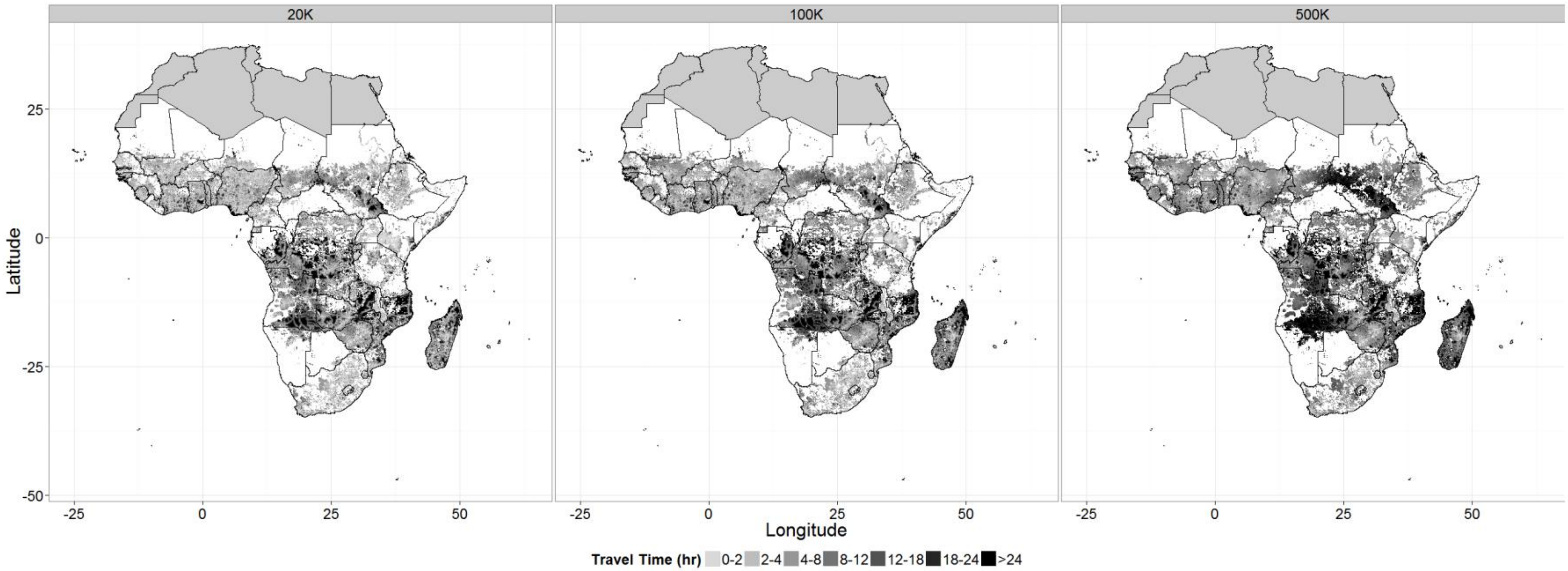
Figure 3-1: Time-to-market from all populated pixels

Table 3-3: Descriptive statistics for time-to-market estimates from all populated pixels

	Total Travel Time (hours)				
	Market Size				
	20K	50K	100K	250K	500K
Sub-Saharan Africa					
Range	170.60	183.14	185.44	186.60	190.54
Median	7.99	9.23	10.51	12.64	14.68
Mean	13.99	15.35	16.55	18.42	20.24
Standard Deviation	16.51	17.16	17.43	17.85	17.92
Central Africa					
Range	158.06	163.81	163.81	168.85	170.44
Median	10.51	12.05	13.22	15.88	17.02
Mean	18.17	19.67	20.75	23.11	24.11
Standard Deviation	20.73	21.14	21.30	21.47	21.60
Eastern Africa					
Range	170.60	183.14	185.44	186.60	190.54
Median	7.67	8.85	9.96	12.46	14.85
Mean	11.01	12.02	13.20	15.21	17.62
Standard Deviation	10.65	10.78	11.15	11.41	12.05
Southern Africa					
Range	69.58	74.20	75.62	75.70	78.01
Median	6.46	7.54	8.87	10.07	13.74
Mean	9.09	10.24	11.28	12.45	17.05
Standard Deviation	8.74	9.24	9.38	9.51	11.62
Western Africa					
Range	112.56	115.01	115.65	118.60	120.14
Median	6.86	7.87	9.51	10.62	11.30
Mean	16.10	17.89	19.34	20.83	21.34
Standard Deviation	19.42	20.75	21.12	21.95	21.73

Source: Developed by author using travel time data from Guo, Joglekar and Beddow (forthcoming) and pixelated 2010 population data from the WorldPop project (WorldPop 2015).

Note: Pixels with zero population were excluded from calculations.



Source: Developed by author using travel time data from Guo, Joglekar and Beddow (forthcoming) and crop harvested area data from SPAM2005 (You, Wood-Sichra, et al. 2015).

Note: Travel time to markets of 20K, 100K and 500K people or more are presented. Pixels with zero harvested area have been masked (in white) from maps.

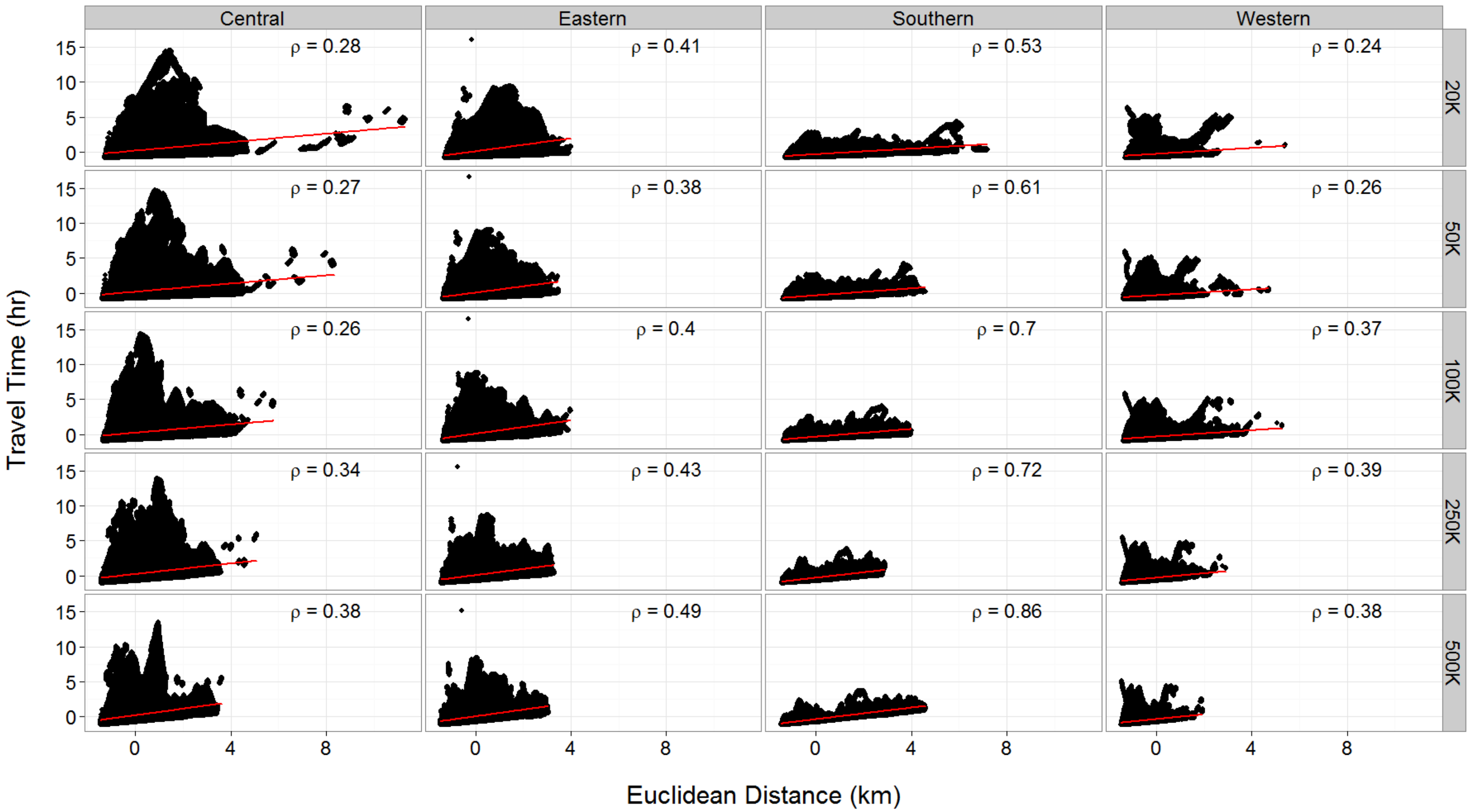
Figure 3-2: Time-to-market from all cropped pixels

Table 3-4: Descriptive statistics for time-to-market estimates from all cropped pixels

	Total Travel Time (hours)				
	Market Size				
	20K	50K	100K	250K	500K
Sub-Saharan Africa					
Range	170.60	183.14	185.44	186.60	190.54
Median	5.22	6.16	7.09	8.99	10.40
Mean	8.40	9.35	10.24	12.11	13.66
Standard Deviation	10.12	10.45	10.64	11.21	11.63
Central Africa					
Range	155.72	163.52	163.52	168.32	170.15
Median	7.61	9.06	10.18	12.70	13.77
Mean	11.55	12.88	13.88	16.32	17.49
Standard Deviation	12.96	13.38	13.46	14.00	14.24
Eastern Africa					
Range	170.60	183.14	185.44	186.59	190.54
Median	5.79	6.78	7.64	10.00	12.34
Mean	8.99	9.93	10.78	12.90	14.99
Standard Deviation	9.86	10.04	10.31	10.73	11.22
Southern Africa					
Range	52.10	53.60	53.75	55.00	57.01
Median	3.93	4.57	5.33	6.81	8.49
Mean	5.11	5.82	6.85	8.14	10.96
Standard Deviation	4.54	4.90	5.50	5.71	8.11
Western Africa					
Range	72.60	72.60	73.17	73.17	73.17
Median	3.29	3.93	4.66	5.67	6.64
Mean	4.48	5.06	5.82	6.72	7.52
Standard Deviation	4.60	4.73	4.95	5.01	5.00

Source: Developed by author using travel time data from Guo, Joglekar and Beddow (forthcoming) and crop harvested area data from SPAM2005 (You, Wood-Sichra, et al. 2015).

Note: Pixels with zero harvested area were excluded from calculations.



Source: Developed by author using Euclidean distance data from own-calculations, travel time data from Guo, Joglekar and Beddow (forthcoming) and crop harvested area data from SPAM2005 (You, Wood-Sichra, et al. 2015).

Note: Pixels with zero harvested area were excluded from plots and calculations.

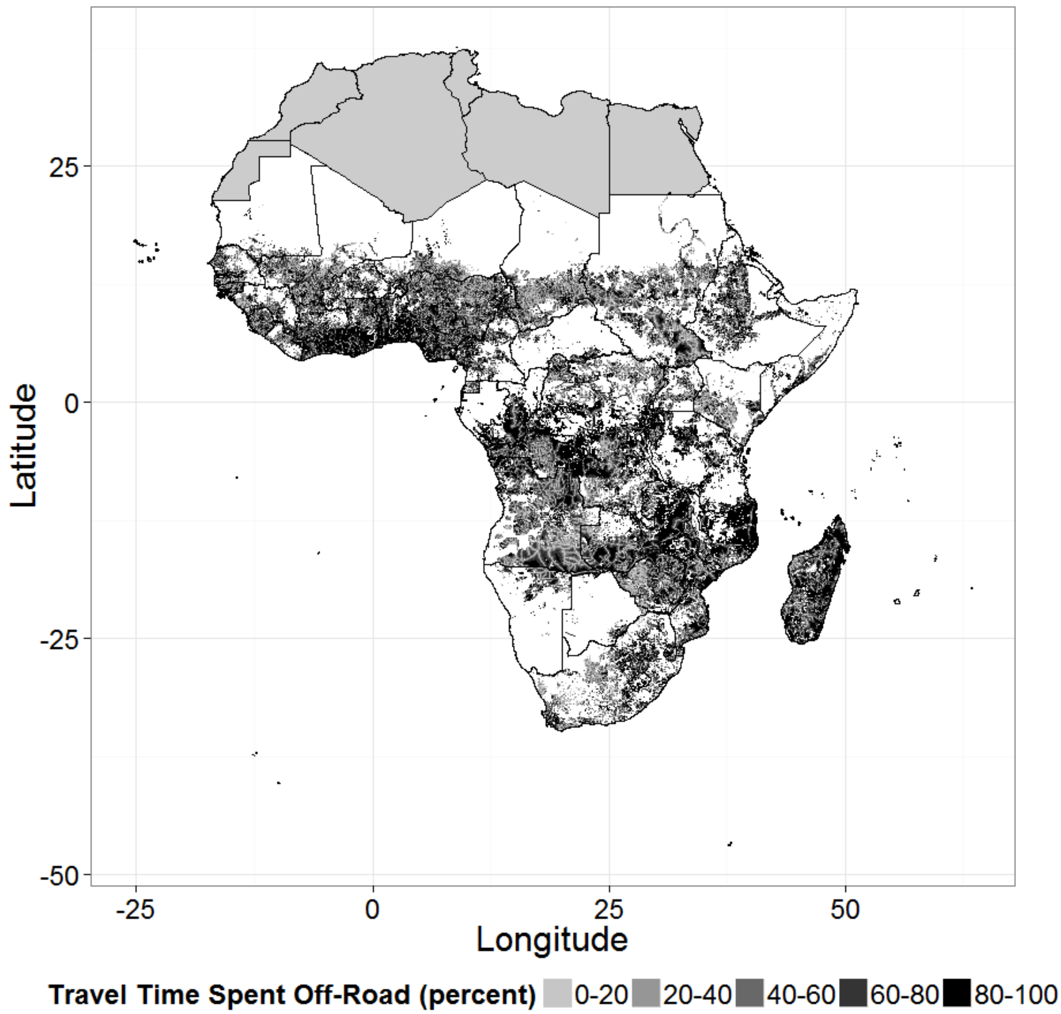
Figure 3-3: Pairwise comparison between Euclidean distance- and time-to-market estimates

Table 3-5: Percentage of pixels, persons and value of crop production by remoteness

	Percentage of Category						
	Travel Time						
	0-2 hours	2-4 hours	4-8 hours	8-12 hours	12-18 hours	18-24 hours	> 24 hours
Sub-Saharan Africa							
Pixels (count)	16.6	22.4	28.0	13.1	9.2	4.4	6.2
Persons (count)	54.7	19.9	15.4	5.0	2.6	1.0	1.3
Value of Production (I$/mt)	33.0	28.6	24.7	7.7	3.6	1.1	1.3
Central Africa							
Pixels (count)	8.0	16.1	28.1	17.0	13.4	6.9	10.5
Persons (count)	42.4	18.7	18.4	8.1	5.6	2.6	4.2
Value of Production (I$/mt)	20.0	25.5	28.6	10.7	6.9	3.2	5.1
Eastern Africa							
Pixels (count)	15.8	19.8	27.4	14.2	10.4	5.1	7.3
Persons (count)	53.1	19.7	16.1	5.7	3.0	1.1	1.2
Value of Production (I$/mt)	37.8	25.8	21.8	7.8	4.0	1.4	1.5
Southern Africa							
Pixels (count)	22.5	28.4	31.8	10.3	4.8	1.5	0.7
Persons (count)	61.3	19.0	14.1	4.0	1.1	0.3	0.2
Value of Production (I$/mt)	35.4	33.2	25.4	4.8	1.0	0.1	0.1
Western Africa							
Pixels (count)	27.0	32.9	27.8	7.4	3.2	0.8	0.8
Persons (count)	60.4	20.6	13.7	3.2	1.3	0.3	0.5
Value of Production (I$/mt)	32.4	30.3	25.6	7.4	3.1	0.6	0.6

Source: Developed by author using travel time data from Guo, Joglekar and Beddow (forthcoming), crop harvested area and value of production data from SPAM2005 (You, Wood-Sichra, et al. 2015), and pixelated 2010 population data from the WorldPop project (WorldPop 2015).

Note: Calculations were based on travel time to a market of 20K people or more. Pixels with zero harvested area were excluded from calculations.



Source: Developed by author using travel time data from Guo, Joglekar and Beddow (forthcoming), and crop harvested area data from SPAM2005 (You, Wood-Sichra, et al. 2015).

Note: Fraction of time spent traveling off-road to market of 20K people or more is presented. Pixels with zero cropping have been masked (in white) from map.

Figure 3-4: Fraction of time spent traveling off-road

Table 3-6: Value of production (I$/mt) by remoteness and fraction of travel spent off-road

	Percentage of Total Value of Crop Production						
	Total Travel Time						
	0-2 hours	2-4 hours	4-8 hours	8-12 hours	12-18 hours	18-24 hours	> 24 hours
Sub-Saharan Africa							
Off-Road: 0-20 percent	32.8	14.8	3.8	1.1	0.3	-	-
Off-Road: 20-40 percent	23.8	13.7	6.0	2.7	1.2	0.2	-
Off-Road: 40-60 percent	18.5	21.4	14.5	7.2	5.4	3.7	0.2
Off-Road: 60-80 percent	15.6	27.9	33.1	25.1	18.3	18.6	8.5
Off-Road: 80-100 percent	9.4	22.2	42.6	63.9	74.8	77.5	91.2
Central Africa							
Off-Road: 0-20 percent	35.0	26.0	11.8	4.7	1.1	-	-
Off-Road: 20-40 percent	23.9	18.8	13.3	9.1	3.7	0.6	-
Off-Road: 40-60 percent	17.2	21.2	21.9	18.9	13.1	7.2	0.4
Off-Road: 60-80 percent	14.1	20.6	31.1	35.4	31.9	25.6	11.7
Off-Road: 80-100 percent	9.8	13.4	21.9	31.8	50.1	66.6	87.9
Eastern Africa							
Off-Road: 0-20 percent	35.8	15.3	5.4	1.2	0.2	-	-
Off-Road: 20-40 percent	24.5	13.8	7.3	4.2	1.2	0.2	-
Off-Road: 40-60 percent	18.6	21.5	15.7	9.3	8.0	3.8	0.3
Off-Road: 60-80 percent	13.9	29.7	32.7	28.8	23.5	25.3	10.9
Off-Road: 80-100 percent	7.2	19.7	39.0	56.5	67.0	70.7	88.8
Southern Africa							
Off-Road: 0-20 percent	31.3	18.1	3.2	0.5	0.2	-	-
Off-Road: 20-40 percent	23.5	12.5	7.6	1.3	0.8	0.1	-
Off-Road: 40-60 percent	18.6	18.8	14.0	5.9	4.3	4.1	0.8
Off-Road: 60-80 percent	17.7	27.1	28.9	23.2	35.9	57.9	10.0
Off-Road: 80-100 percent	8.9	23.4	46.2	69.2	58.8	38.0	89.2
Western Africa							
Off-Road: 0-20 percent	30.6	12.2	1.3	-	-	-	-
Off-Road: 20-40 percent	23.3	12.9	3.5	-	-	-	-
Off-Road: 40-60 percent	18.6	21.7	12.3	2.6	0.1	-	-
Off-Road: 60-80 percent	16.7	28.4	34.3	20.0	7.8	1.1	-
Off-Road: 80-100 percent	10.9	24.8	48.5	77.4	92.2	98.9	100.0

Source: Developed by author using travel time data from Guo, Joglekar and Beddow (forthcoming), and crop harvested area and value of production data from SPAM2005 (You, Wood-Sichra, et al. 2015).

Note: Calculations were based on travel to a market of 20K people or more. Pixels with zero harvested area were excluded from calculations

Chapter 4: Getting (Fertilizer) Prices Right in Tanzania

4.1 Introduction

Contrary to most (if not all) economic sectors, agriculture has a distinctive, expansive footprint.[29] Farming takes place at disparate locations that can be quite distant from markets where off-farm inputs (like fertilizer) are sourced and where agricultural output is increasingly consumed as economies (and especially African economies) continue to urbanize. Moreover, agricultural inputs and outputs are often bulky, and thus entail significant (transaction) costs in moving goods to and from markets. It is the on-farm unit costs and returns of these agricultural goods and services that affect the bottom line for farm families, such that a better understanding of the spatial (market and farm-level) dispersion of agricultural prices is critical for a more nuanced understanding of the economics of farming, and the policies and intervention strategies built upon that understanding.

The uptake and regular use of inorganic fertilizers is woefully low in sub-Saharan Africa. In 2014, the region only accounted for approximately 1.6 percent of fertilizer nutrients consumed globally (FAO 2015). Many claim that farmers in sub-Saharan Africa use suboptimal levels of fertilizer from a technical standpoint (Sheahan, Black and Jayne 2013, Marenya and Barrett 2009, Matsumoto and Yamano 2011, Burke 2012), however, assuming farmers are rational optimizers, their decision to purchase fertilizer is likely motivated by objectives other than pure production profit-maximization. While African farmers may not purchase the amount of fertilizer necessary to obtain technical efficiency (i.e., the least amount of inputs necessary to produce a targeted output level), farmers arguably purchase the necessary level of inputs to obtain the highest net value from production (Beddow, Hurley and Pardey 2014). In other words, the demand for agricultural inputs matches the relative profitability needs of the farming household, such as concerns about food security or household expenditures on health and schooling, in addition to production profit maximization; needs that vary by household and location (Kelly 2006).

[29] Throughout sub-Saharan Africa, agriculture areas accounted for 43.2 percent of the total land mass in 2012 (FAO 2015).

Very few analyses use spatially-delineated prices, so little is known about the pattern of agricultural prices in sub-Saharan Africa, and even less so about the spatial distribution of input prices. Do they differ spatially? How are they measured? Do they make any economic sense? Getting answers to these basic measurement questions is an essential pre-condition to doing any meaningful analysis on the role of (relative) prices on farm decision making. To untangle the economic realities affecting fertilizer use decisions of African farmers requires going beyond a consideration of fertilizer prices at the national level. A working knowledge of the market retail and farm level prices faced by farmers is needed.

To start building that working knowledge, I examined the unit costs of inorganic fertilizer purchases in Tanzania with respect to the costs of transporting fertilizer from the port in Dar es Salaam to inland fertilizer retailers, proxied by the time-to-market metrics discussed in Chapter 3 of this dissertation. The unit cost of fertilizer purchases were calculated from the Agricultural Sample Census (ASC), which reported the quantity and cost of purchased fertilizer by season in the 2007/08 agricultural production year for 47,845 rural agricultural households in mainland Tanzania. To investigate the spatial patterns of fertilizer prices, it was necessary to know the location of production. The ASC survey did not report geographical coordinates, but did include the name of the villages surveyed. With that information I was able to approximate the geo-coordinates for 96.6 percent of the 3,507 villages sampled.

There are several factors that may influence the local supply and demand dimensions of inorganic fertilizers, and thus, the price of fertilizer, including transportation costs, government fees, regulations and policies, access to finance, knowledge regarding the product, agronomic response to fertilizer, and the use of complementary inputs (e.g., hybrid maize seed). However, the majority of variation in fertilizer prices is attributed to the high costs of transporting the bulky good and will likely vary spatially due to the size and (often fragmented) nature of fertilizer retail markets in Tanzania (Guo, Koo and Wood 2009, Chemonics International and IFDC 2007).

Despite the growing need for spatially delineated fertilizer prices, and the existence of data on the costs and quantities of fertilizer purchases in many contemporary household-level surveys (including the Living Standards Measurement Study (LSMS) surveys backstopped by the World Bank (LSMS-ISA 2015)), there is a woefully low use of these statistics in the literature. Large degrees of unexplained variation in the unit cost variables are likely the reason for the lack of use to this point in time. Household-level surveys require respondents to recall information from several weeks to months prior, and as a result, often suffer from sizeable amounts of measurement error, introduced by both the respondent and interviewer (Groves 1989, Deaton 1997).

As expected, there is a substantial amount of variation in the reported fertilizer unit costs from the ASC survey. I attempt to explain this variation with data on spatial location and transport costs, but the results are insignificant and the level of explained variation is low. Contemporary agricultural surveys suffer from issues of measurement error and omitted variables which renders some data (such as the unit costs of fertilizer examined in this chapter) useless. In terms of fertilizer purchases, while the use, quantity and cost of inputs is valuable information in itself, without knowledge of what exactly was purchased and from where, it is difficult to systematically understand the incentives facing farmers. This is invaluable information for policymakers and donors attempting to strategically and effectively increase the productivity capacity of agriculture in response to the ever increasing demands on our food systems.

4.2 Background and Motivation

Sub-Saharan Africa has experienced positive growth in agricultural productivity, but growth rates are relatively small and insufficient to meet food security and poverty reduction goals (Benin, et al. 2011). In response, policymakers and researchers recommend shifting from traditional cultivation techniques towards the integration of modern inputs, including inorganic fertilizers. Currently, sub-Saharan Africa has some of the lowest fertilizer usage rates in the world. In 2012, the average intensity of fertilizer used in the sub-Saharan Africa was 14.7 kilograms of fertilizer nutrients per hectare of arable land (kg/ha) (FAO 2015),[30] and these rates varied substantially among countries (see Figure

[30] Calculated as the ratio of the sum of the quantity of nitrogen, phosphate and potash nutrients consumed to the total amount of arable land reported by FAO (2015). Data on fertilizer consumption was not available for 18 SSA countries: Botswana, Cabo Verde, Central African Republic, Chad, Comoros, Djibouti, Equatorial

4-1). Comparatively, this rate was 123.2 kg/ha in Latin America and the Caribbean, 126.7 kg/ha in high income – OECD countries and 457.5 kg/ha in East Asia and the Pacific.

[Figure 4-1: Year 2012 fertilizer use (kg/ha) in sub-Saharan Africa]

There are a host of reasons cited for the relatively low levels of (inorganic) fertilizer use in sub-Saharan African agriculture. These include both demand side factors (e.g., poor quality fertilizers, lack of information, risk aversion and uncertainty regarding climatic variables, output market incentives, access to credit and complementary inputs) and supply side factors (e.g., bottlenecks in distribution, inadequate supply, high transport costs and inefficiencies at ports). In response to low uptake and usage rates there has been a recent resurgence in price support policies, namely input subsidy programs. According to a review by Jayne and Rashid (2013), ten sub-Saharan African countries spent approximately US$1.05 billion on these programs which was equal to 28.6 percent of their public expenditures on agriculture in 2011. In their review, Druilhe and Barreiro-Hurlé (2012) find that subsidies have been effective in raising fertilizer use, but given that rigorous evaluations of input subsidies are lacking, there is inconclusive evidence that these increases are economically efficient. One reason for this inefficiency may be that initial evaluations of the nature of adoption within a region do not include spatial measures of input and output prices faced by farmers (retail or farm-gate).

Ultimately, farmers will not choose to use fertilizer if it is not profitable (or relatively profitable given other household expenditures), so underpinning any program or policy aimed at increasing fertilizer use should be a comprehensive analysis of profitability. Researchers traditionally use three related metrics to quantify a farmer’s potential profitability: (i) the agronomic response: output/input quantity ratios, (ii) relative prices: input/output price ratios or (iii) net return: value/cost ratios. Regardless of the measure used, it is important to remember that the factors that influence yields and prices (e.g., agro-ecological conditions, the degree of infrastructure available and its quality and the nature of human capital) all vary spatially. Thus, a profitability assessment using aspatial prices will not reflect the varied economic realities affecting fertilizer use decisions on African

Guinea, Guinea-Bissau, Lesotho, Liberia, Mauritania, Mayotte, Réunion, São Tomé and Principe, Sierra Leone, Somalia and Swaziland.

farms (Guo, Koo and Wood 2009). A spatially explicit understanding of the current incentive structures faced by farmers, specifically fertilizer prices, has prospects for substantially improving the efficiency of and participation in fertilizer markets, and reducing the social inefficiencies incurred by initiatives set to increase fertilizer use.

Data on household-level fertilizer purchases are typically collected through agricultural censuses and integrated households surveys such as those conducted by national statistical agencies in collaboration with the Living Standards Measurement Study – Integrated Surveys on Agriculture (LSMS-ISA) (LSMS-ISA 2015) and various, sometimes complementary, national farm household surveys (Pardey, et al. Forthcoming). Traditionally, these national surveys simply reported the use of inorganic fertilizer as a binary variable, but more recently details on the amount purchased, cost, and type of fertilizer are included in the questionnaires. Some of these surveys include a question regarding location of purchases, but it usually is in the form of a categorical answer such as "government" or "private trader in local market." Beyond a vague measure of distance (e.g., the location is "within the village" or in "other region") there are few questionnaires that give a quantitative sense of the household's input purchase location. This is mainly to protect the confidentiality of the respondents, but all current LSMS surveys are released with household data tagged by village-level geographic coordinates which maintains confidentiality and still allows for the data to be spatially analyzed.[31]

There are trade-offs between the large scale census surveys and the small-scale integrated household surveys. While small-scale surveys, such as the LSMS-ISA, do collect a wealth of information on the nature of agricultural input use, they are typically not stratified below a sub-national level one administrative unit, which makes it difficult to get a reliable sense of retail-level variation in fertilizer prices. The problems with sample size of these small-scale surveys are further exacerbated by the woefully low use of fertilizer in sub-Saharan Africa. Large-scale surveys, such as an agricultural census will usually have a reasonable sample of respondents using fertilizer, but is less likely to have

[31] In some instances (e.g., Tanzania's National Panel Survey (NBS 2011b)) these integrated household surveys are complemented with a community questionnaire which collects market-level prices on selected agricultural outputs and inputs, including fertilizer. The Euclidean distance between households and the relevant market can be calculated using the geo-coordinates collected from the market location.

detailed information on fertilizer purchases. Additionally, any household-level survey based on recall data is likely wrought with measurement issues (Deaton 1988, Groves 1989, Deaton 1997).[32]

While not necessarily household specific, market-level fertilizer price data is also being collected from the seller directly (Omamo and Mose 2001, Benson, Kirama and Selejio 2012). The 2006 African Fertilizer Summit, called to address the role of fertilizers in reducing rural poverty in Africa, resolved to "set up a mechanism to monitor and evaluate the implementation" of the fertilizer-use resolutions tabled in the Summit declaration (IFDC 2007). This spurred the formation of AfricaFertilizer.org (AFO). AFO works to facilitate the exchange of information regarding fertilizer use in sub-Saharan Africa. To that end, they have partnered with the Regional Agricultural Input Market Information System (AMITSA) to provide monthly market-level input prices in 19 African countries since 2010. Their data is collected through other agencies and crowd-sourcing of local fertilizer prices (delineated by fertilizer type) via a web-based survey instrument.[33]

Getting a meaningful handle on the nature of (localized) fertilizer prices is complicated. While farmers' decisions are largely dependent on their relevant agro-ecological conditions, market isolation and the nature of output marketing opportunities, spatially-delineated data on fertilizer prices are only useful if the data is representative of reality. Thus, the remainder of this chapter focuses on assessing and analyzing the variation in the household-level fertilizer unit costs reported in Tanzania. But to do so, it is first necessary to have an understanding of the nature of fertilizer markets in Tanzania. Fertilizer markets are constrained by several factors that inhibit their performance. Such constraints include uncertain policy environments, weak regulatory systems, inadequate human capital, limited access to finance, lack of market information, the size of the market,

---

[32] Deaton (1988) emphasizes that if the quality (or type) of a good are not accounted for, then the unit costs are not direct substitutes for true market prices. If both quantity and cost of fertilizer are measured with error, not only will the unit cost be measured with error, but this could generate a spurious negative correlation between quantity and unit cost (Deaton 1988). For instance, households may report quantity by dividing cost by the price or vice versa.

[33] http://africafertilizer.org/prices_detailed.html

unnecessary product differentiation,[34] technical knowledge transmission and infrastructure (Gregory and Bumb 2006). All of which will impact the demand for and supply of fertilizer at the local market level.

4.3 Fertilizer Markets in Tanzania

4.3.1 Supply of Fertilizer

In Tanzania, fertilizer is primarily procured internationally and distributed to retailers (via wholesalers) through the private sector.[35] Retailers either sell the fertilizer at full-price (Scenario 1 in Figure 4-2) or accept a subsidized payment (Scenario 2 in Figure 4-2).[36] The number of fertilizer retailers throughout Tanzania is unclear but estimates put it between 2,500 and 3,000, equivalent to 1.3 retailers for every 10,000 farmers in the country (Benson, Kirama and Selejio 2012, Thapa 2012). Export-oriented companies (e.g., tea and coffee) also import their fertilizer through private companies and then distribute this fertilizer to their outgrower farmers (Scenario 3 in Figure 4-2).

[Figure 4-2: Fertilizer supply systems in Tanzania]

The main costs associated with the supply of fertilizer above the cost of the product itself are associated with international and domestic shipping and handling. According to

[34] Fertilizer markets are quite fragmented in Africa, but also sell many similar products that may be unnecessarily specialized (Morris, et al. 2007).

[35] The parastatal firm, Tanzania Fertilizer Company (TFC), has not imported fertilizer since 2008. In response to the sharp increase of food prices, the TFC was instructed by the government to purchase large amounts of phosphate fertilizers (both domestically and internationally). International prices for phosphate fell shortly after the purchase and TFC's competitors were able to offer DAP at lower prices, forcing TFC to sell its stock at a loss. The viability of the parastatal is questionable (Benson, Kirama and Selejio 2012).

[36] From 2003 to 2007, Tanzania had a subsidy program in place to compensate importers, wholesalers and retailers for their incurred transport costs (Chemonics International and IFDC 2007) which ultimately resulted in a standardized price for consumers. A year after this program ended, Tazania moved to a subsidy program that gave the purchasing power to the farmer, called the National Agriculutral Input Voucher Scheme (NAIVS). Until 2014, the program targeted 2.5 million full-time farmers with a maize- or rice cropping area of less than one hectare, and supplied them with vouchers for 50 percent subsidized inputs. These inputs included either one 50kg bag of DAP or two 50kg bags of MRP for a basal dressing, one 50kg bag of urea for top-dressing, and either 10kg of improved maize seed or 16kg of rice seed (Benson, Kirama and Selejio 2012). Fertilizer supplied through NAVIS was also distributed using the private importer-wholesaler-retailer supply chain, and retailers would redeem vouchers for reimbursement at the local National Microfinance Bank branch (a predominately privately-owned bank). Due to the bredth of the subsidy program, the government had strong influence on the quantity and allocation decisions of fertilizer procured and distributed in the country. The data used in this analysis represents an agricultural year when no subsidy was in place, namely 2007/08.

a 2006 study on fertilizer supply and costs in sub-Saharan Africa by Chemonics International Inc. and the International Fertilizer Development Center (IFDC) (2007), the average price of fertilizer at the retailer was US$419 per metric ton (mt) in Tanzania. Most (65.1 percent) of this price was attributable to FOB plus bagging and 22.4 percent was attributed to inland transport costs.[37] The remaining 12.5 percent of the retail price was derived from government fees, overhead, finance and margins. The study looked at three fertilizer types (Urea, Diammonium phosphate (DAP) and calcium ammonium nitrate (CAN)) and calculated cost chains for three observations of each (Table 4-1). Unit transport costs varied by fertilizer type and destination; Urea transported to Songea cost $0.07 per ton kilogram while transporting CAN to the same city cost $0.39 per ton kilogram.

[Table 4-1: Fertilizer cost chains from Dar es Salaam to multiple destinations]

In Tanzania, nearly all fertilizer is imported into the country through Dar es Salaam, with a negligible amount entering the country across the Kenyan border. (Thapa 2012, Benson, Kirama and Selejio 2012, Kamhabwa 2014).[38] In 2010/11, there were three firms actively engaging in importing fertilizer (Benson, Kirama and Selejio 2012). The costs associated with importing fertilizer include the c.i.f. price (FOB cost, insurance and frieght), port charges, duties, taxes and finance charges. Limited unloading berths that can only handle shipments of 20,000 mt or less, lack of warehouse space for storage while in port and the monopoly position of the Port Authority results in relatively high port charges in Dar es Salaam (Chemonics International and IFDC 2007).

Next to the cost of the product, transportation costs account for the largest component of the total cost of inorganic fertilizer (Chemonics International and IFDC 2007). Inland transport costs are high but competitive. The reported inland shipping cost of 62 US$/mt in Tanzania was found to be no higher per ton mile than the other countries (Chemonics International and IFDC 2007). Benson, Kirama and Selejio (2012) found similar transport

---

[37] Similarly, Thapa (2012) found that the retail price of fertilizer in Tanzania was equal to the c.i.f (cost, insurance, freight) price plus 41 percent of additional in-country costs. Benson, Kirama and Selejio (2012) found that the FOB (free-on-board) price of fertilizer was 64 percent of the retail price of urea and 66 percent of the retail price of DAP.

[38] There is production of Minjingu Rock Phosphate (MRP) in northern Tanzania. However, the agronomic response of the phosphate fertilizer is not immediately observable, and as such, demand for the product is low.

costs of 30 US$/mt to 50 US$/mt from Dar es Salaam to the main wholesale centers in the Southern Highlands. Road infrastructure in sub-Saharan Africa, especially rural feeder roads, is especially poor. Heavy rains and frequent security checks can also increase transportation costs. For these reasons, there are often less retail markets found in rural areas, which makes it more difficult for famers in these areas to access fertilizers.

The principle constraint for importers, wholesalers and retailers is financing. Since farmers tend to demand fertilizer at the same time (beginning of the rainy season) and storage is costly, fertilizer supply works on a compressed timeline. Depending on how long it takes to sell their product, importers usually need about 2 to 3 months of short-term finance. As long as retailers are able to access finance,[39] wholesalers can usually turn their product around in a month (Gregory and Bumb 2006).[40] Since fertilizer is imported in bulk, a significant amount of cash is needed upfront. Letters of credit (LC) often require high levels of collateral (e.g., 150 percent) and charges (e..g, upwards of 2.5 percent of the c.i.f. values) (Gregory and Bumb 2006). In response to the large costs associated with accessing bank finance, retailers are often forced to self-finance or split purchases into smaller increments which may not meet demand. Finance charges accounted for 3.9 percent of the average unit cost of fertilizer in Tanzania (Chemonics International and IFDC 2007).

4.3.2 Demand for Fertilizer

There are a myriad of fertilizers sold in Tanzania, but from 2002 to 2012, the majority of fertilizer nutrients consumed in Tanzania were nitrogen (81.9 percent in 2007/8), which was primarily in the form of urea (37.3 percent in 2007/8) (see Table 4-2). Fertilizer is mostly used on export crops (e.g., tobacco, tea, cotton, coffee), staple crops (e.g., rice, maize, millet, sorghum) and vegetables (FAO 2007).

[39] In an effort to strengthen human capital within the fertilizer retail network in Tanzania, the government offered a training program for retailers who were allowed to accept NAVIS vouchers. This program included training on business management, product knowledge, output marketing and corporate governance (Benson, Kirama and Selejio 2012). Those retailers who successfully completed the program were given the opportunity obtain loans from commercial lenders to build up inventory.

[40] Benson, Kirama and Selejio (2012) note that there are frequently delays associated with reimbursing retailers for NAVIS vouchers through the National Microfinance Bank. These delays prevent retailers from obtaining additional stock from wholesalers until they can afford to pay off their credit.

If agricultural input and output markets are fully functioning, farmers will choose to use inorganic fertilizer when it is profitable (i.e., the marginal cost of the last unit of fertilizer used is equal to the value of marginal return). Profitability is directly affected by the agronomic response to fertilizer use and input and output prices. However, Kelly (2006) emphasized that there is an important distinction between potential demand (the researcher's perception of profit incentives) and effective demand (the farmer's perception of profit incentives). Due to poor transportation and communications infrastructure, especially in rural areas, knowledge regarding the agronomic responses and market signals are not effectively communicated to potential fertilizer consumers. Thus, the effective demand for fertilizer is often lower than perceived potential demand.[41] There are nine major agro-ecological zones in Tanzania (Guo and Wood-Sichra, AEZ (16-class, 2009) 2015), but policies such as the fertilizer voucher program still have a "one-size-fits-all" recommendation on fertilizer use, so farmers are less likely to have access to appropriate recommendations for their particular situation. Additionally, the agronomic response to fertilizer use will likely change over time as climate patterns and consequently, agro-ecological zones shift. This could affect the location of crop production and ultimately the spatial patterns of prices (Beddow and Pardey, Moving Matters: The Effect of Location on Crop Production 2015).

[Table 4-2: Fertilizer consumption in Tanzania from 2002 – 2012]

Other factors that influence the demand incentives to use fertilizer include relative returns (i.e., fertilizer expenditures relative to other household expenditures) and profitability risks. HIV/AIDS is still a serious concern in Tanzania, so a significant amount of household resources may be allocated towards health.[42] Uncertainty is inherent in agricultural production and farmers have to account for both production risks (e.g., weather, input response) and price risks (e.g., output price fluctuation) in their decision to use fertilizer (Rowhani, et al. 2011, Minot 2010). Another issue related to uncertainty, is

[41] Mobile phones are making these types of signals easier to communicate. With smartphones, it is now possible to access weather updates and prices in real time (Aker and Mbiti 2010).
[42] The most recent HIV/AIDS and Malaria Indicator Survey found that HIV prevalence ranges from a low of less than 1 percent in the Pemba region to a high of 14.8 percent in the Njombe region (TACAIDS, ZAC, NBS, OCGS and ICF International 2013).

the farmer's perception on fertilizer quality. The impetus for the formation of the more stringent legislation on fertilizer regulation is based on anecdotal evidence that farmers are concerned with the high levels of adulterated fertilizer, but there is no objective evidence in Tanzania that this is the case (Benson, Kirama and Selejio 2012).[43] There are reported issues of caking, which is the result of storing fertilizers on the floor (Chemonics International and IFDC 2007).

A second, equally important, aspect of fertilizer demand is the capacity to acquire and use the product. There are few fertilizer retailers that service rural agricultural areas in sub-Saharan Africa, and those that do exist may struggle to obtain their product in a timely manner due to the poor quality of rural feeder roads and ability to access finance. Similarly, if a farmer determines that fertilizer use is profitable for their production, he might not be able to access the necessary credit to make the purchase. Without an influx of capital at the start of the agricultural season, it is difficult to afford small quantities of fertilizer, let alone the recommended amounts. Rural credit markets are thin and those that do exist may require large amounts of collateral that farmers do not have (Barrett, Reardon and Webb 2001). In response to own-credit constraints or low retail inventory levels, farmers often choose to buy fertilizer in smaller quantities than the traditional 50 kg bag. To service the demand for smaller purchases, retailers will re-bag fertilizer in-shop. Premiums and faulty scales could result in higher costs for farmers.

Even if farmers can effectively obtain fertilizer, there are still issues regarding the famer's capacity to use the product efficiently. If farmers do not know about available technologies or possess the skills to evaluate and adopt these technologies to their own production, they will not choose to use the technology (Kelly 2006). There are several types of fertilizers to choose from, but limited technical and extension support can result in farmers using fertilizers with the wrong nutrients for their soils.

As mentioned above, reported local market fertilizer price data, and especially farm or local market level data, are likely to vary for a variety of reasons, many of which are

[43] Tanzania's Fertilizer Act of 2009 replaced the Fertilizers and Animal Foodstuffs Act of 1962, in an effort to better regulate importation, distribution, storage and marketing of fertilizer within the country. At the time of their review, Benson, Kirama and Selejio (2012), mention that the regulatory authority meant to enforce the new legislation was still not in place, and efforts to monitor quality only took place at the port.

unobservable. Thus taking local fertilizer price data at face value can lead to spurious inferences about the nature of fertilizer prices and the fertilizer use behavior of farmers associated with these fertilizer prices. This analysis looks specifically at variation in self-reported unit fertilizer costs with regard to supply factors (e.g., transport time from a port to a retailer and institutional differences by region) and demand factors (e.g., production season, total harvested area, quantity of fertilizer purchased and household demographics) and market size.

4.4 Data

4.4.1 Household-Level Fertilizer Purchases

The main fertilizer purchase data used for this analysis are from the 2007/2008 Tanzanian Agriculture Sample Census Survey (ASC) (NBS 2011a). This household-level survey was conducted by the Tanzania National Bureau of Statistics in 2009 and solicited recall information from small-scale farmers in rural areas on their 2007/08 agricultural production activities.[44] While the ASC is spatially stratified at an administrative level two (ADM2) unit across Mainland Tanzania and Zanzibar, only mainland data are used for this analysis;[45] 47,845 households from 127 districts were sampled on the mainland.

The maintained hypothesis underpinning this analysis is that a significant portion of the reported variation in fertilizer unit costs paid by farm households at a local retailer depends on the location of that retail fertilizer outlet to a fertilizer port (Deaton 1988, Minten, Koru and Stifel 2013). I would expect the price of fertilizer to be higher in areas that are further from the Dar es Salaam port in response to increased transport costs. The 2007/08 ASC does not include geo-coordinate references for each of the farm households surveyed, so to assess these spatial relationships the names of sampled villages from the ASC were matched to spatial data from the International Livestock Research Institute (ILRI) (OpenMicroData 2010) to approximate the geographic coordinates of each

[44] Small-scale farms are defined as farms that (1) have between 25 meters squared and 20 hectares under production; and/or (2) between 1 and 50 head of cattle, and/or between 5 and 100 head of sheep, goats, or pigs; and/or (3) between 50 and 1000 chickens, turkeys, ducks, or rabbits (NBS 2011a).

[45] The metric used to calculate travel time does not account for water-based travel, so it would not be appropriate to use data from the island of Zanzibar.

household. A detailed description of this process is presented in Appendix A. It is worth noting that this procedure only yields the geographic coordinates of the village in which a given household is located, not the household itself. Thus, all households surveyed within a given village will be assigned the same geo-coordinate.

The ASC solicits information on the use of inorganic fertilizer by crop; specifically, the total quantity purchased and the total cost of the purchase for each crop grown in the long-rain or short-rain production season by household.[46] The survey does not distinguish between the types of fertilizer purchased (e.g., Urea, DAP, CAN) and purchases reflect the quantity of fertilizer purchased, not the amount of active nutrients. Additionally, the survey does not ask producers to specify where their fertilizer purchases are made. According to the survey, 12.5 percent of farms purchased inorganic fertilizer in 2007/08. However, there is substantial spatial variation in purchasing behavior of farmers. Figure 4-3a maps the percent of farm households within each ADM1 (administrative level 1) that purchased fertilizer in 2007/08; Figure 4-3b maps the average quantity of fertilizer purchased per farm household in each ADM1; Figure 4-3c maps the average cost of fertilizer purchased by farm households in each ADM1; and Figure 4-3d maps the unit cost of purchased fertilizer averaged across households in each ADM1.[47] The greatest percentage of household purchases occurred in the southwestern and Kilimanjaro regions, where around 40 percent of farming households made fertilizer purchases. The Kagera, Singida, Manyara, Dodoma, Pwani and Lindi regions had little to no fertilizer purchases. Households in the Rukwa region made the largest purchases for the largest sums of money, on average, but their usage rates were slightly below the national average. The highest average unit cost faced by households was in Manyara region – this was almost 5 times the national average. There did not appear to be any major discernable spatial patterns in the four maps.

---

[46] Specifically, the survey asks farmers to provide the following "for each crop planted during 2007/08 [Long/Short] rainy season:

- Area applied with fertilizer (used on less than ¼ of whole crop, ¼ of whole crop, ½ of whole crop, ¾ of whole crop or whole crop)
- Type of fertilizer used (organic versus inorganic)
- Quantity of fertilizer use (kilograms)
- Cost of fertilizer use (Tanzanian shillings)" (NBS, et al. 2012)

[47] There were 21 mainland ADM1 units in Tanzania in 2007/8, with a mean area of 4.5 million hectares per ADM1 ranging from 0.2 million hectares (Dar es Salaam) to 7.8 million hectares (Rukwa).

[Figure 4-3: Descriptive statistics on fertilizer purchases]

Tanzania has both unimodal and bimodal agricultural seasons: the unimodal rains (*Msimu* long-rain production) fall in the southern two-thirds of the country, while the bimodal rains (*Masika* long-rain production or *Vuli* short-run production) cover the north-northeastern regions. Fertilizer purchases are usually made during the planting phase of the long-rain season (Benson, Kirama and Selejio 2012, IFDC 2012). Planting runs from mid-September through October for the *Vuli* season, October through January for the *Msimu* season and mid-February through March for the *Masika* season (FEWS-NET 2013). Therefore, most farmers who reside in the unimodal regions would be expected to purchase fertilizer between October and January, and in the bimodal regions between February and March. The average price paid for fertilizer by households in Tanzania was 1.18 TSh/mt (approximately 1,000 US$/mt).[48] Compared to 419 US$/mt retail price found by Chemonics International and IFDC (2007), this seems rather high, but there was a dramatic increase in fertilizer prices in early 2008 which may have affected some Tanzanian farmers. The world price for Urea rose from 289.00 US$/mt in June 2007 to 628.38 US$/mt in June 2008 and the average world price for DAP rose from 493.00 US$/mt to 1175.00 US$/mt in the same time period (World Bank 2015).

4.4.2 Transportation Costs

According to Chemonics International and the IFDC (2007), 64.2 percent of the in-country costs associated with the retail fertilizer price are attributed to transportation. To proxy for these costs I used the measure of market accessibility developed by Guo, Joglekar and Beddow (forthcoming) described in Chapter 3 of this dissertation. The ASC did not specify where fertilizer purchases were made, but for this analysis, I assumed that fertilizer retailers were located in cities of at least 20,000 people – Guo, Joglekar and Beddow (forthcoming) determined that there were 161 cities in Tanzania that met this criteria at the time the data was collected.[49] The market accessibility layers were not dated, but the

[48] For this analysis, I used the average exchange rate from June 2007 – May 2008: 1,202.59 TSh = 1 US$.

[49] Kamhabwa's (2014) report for *AfricaFertilizer.org* lists 21 of the major fertilizer distribution centers in Tanzania; the minimum population in this list of cities is 40,000. I would assume that retailers are located in cities smaller than this. Additionally, according to Benson, Kirama and Selejio (2012), while no formal census has been done, estimates were given of 3,000 fertilizer retailers in Tanzania. Distribution of these

population centers used in the analysis were updated to reflect counts in 2010. At that time there was only one Tanzanian city greater than 500,000 people: Dar es Salaam. Conveniently, this is really the only port of concern for fertilizer imports in the country. Assuming that transportation from the port to the retailer (via the wholesaler) only occurs using a road network, and that fertilizer retailers are located within all cities greater than 20,000 people in Tanzania, I estimated the relevant travel time between the port and the retailers (500K minus 20K) reflected in each household's respective fertilizer price.[50] The relationship between port-to-retail travel time and fertilizer unit costs is expected to be positive.

4.4.3 Additional Variables

In addition to the port-to-retail travel time, I included regional dummy variables (relative to Dar es Salaam) to proxy for institutional factors that affect the supply of fertilizer. The main demand factors considered in the analysis were production season, total harvested area, quantity purchased and household demographics. There were substantially fewer producers using fertilizer in the short-rain season. This decrease in demand is likely associated with increased fertilizer unit costs in the short-rain season. Household's with larger areas under production may need larger amounts of fertilizer for their production and be able to secure quantity discounts on their unit cost. The (continuous) quantity purchased variable is categorized based on the number of 50 kilogram bags purchased – the common unit of import. If a farmer cannot afford or does not need a 50 kg bag, smaller units are often available for purchase, but there are likely penalties associated with re-bagging the fertilizer. Along the same line, purchases exceeding a 50 kg bag may experience bulk discounts. A collinearity between isolation and bulk purchases may exist (i.e., farmers who live further away attempt to save on transportation costs by purchasing more fertilizer at a time). To address this concern I interacted the quantity purchased with

---

retailers is inconsistent – most retailers are located in districts with high agricultural potential and remote areas may have little to no representation. Therefore, the assumption that fertilizer retailers are located in the 160 cities of 20,000 people or more in Tanzania seems reasonable.

[50] The data for Tanzania is not identical to the data used in the previous Chapter 3. The cost-distance function was run on Tanzania independent of other countries to remove border effects from nearby cities of 500,000, such as Nairobi, Kenya and Kigali, Rwanda.

a variable that measures the travel time from the fertilizer retailer to the household's village (i.e., the time-to-market metric to a market of 20K or more introduced in Chapter 3). Finally, households with better abilities to search and negotiate terms of trade may also be able to secure lower fertilizer unit costs. These transaction costs are proxied by head of household demographics including sex, age, education, literacy and a source of off-farm income.

Descriptive statistics of the variables used in this analysis are presented in Table 4-3. On average, 81 percent of the heads of households using inorganic fertilizer were male, 45 years old, and had six (of eight) years of primary schooling; 87 percent were literate. In terms of production, 88 percent of the households using inorganic fertilizer produced in the long-rain season and 18 percent in the short-rain season. There was a wide range in the quantity and costs of fertilizer purchases within the country, with highly unrealistic maximums, so the data has been cleaned by removing the top and bottom two percent. The majority of farmers purchased between 1 and 2 – 50kg bags of fertilizer, the average cost was 106,873.4 Tanzanian shillings (88.87 US$), and the average unit cost was 796.81 TSh/kg. (0.66 US$/kg).

[Table 4-3: Descriptive statistics on main analysis variables]

4.5 Analysis

Due to factors such as infrastructure quality, distance from the port of entry, weather patterns, institutional arrangements and an underdeveloped retailer's network, spatial variability in the retail prices paid by farmers would be expected. Figure 4-4 shows that while there is indeed spatial variability in the reported unit costs of fertilizer, there does not appear to be an identifiable relationship between distance from Dar es Salaam and the average unit costs. Additionally, the correlation coefficient between these two variables is very weak ($\rho = 0.06$). Given that transport is often cited as the main factor of the price markup after importation, it is interesting to note that quite often villages that reside the furthest from the importation port (Dar es Salaam) face relatively lower prices than some closer villages.

[Figure 4-4: Average household-level unit fertilizer costs by village]

The effect of port-to-retail time, and the other above-mentioned supply and demand factors, on average fertilizer unit costs reported by the household were calculated using multivariate OLS regressions. Before proceeding with any analysis, I checked to see if the data I was using matched the aggregates presented in the final report associated with the 2007/2008 ASC data (NBS, et al. 2012). Appendix B contains four tables that compare my own aggregates of the ASC data to those presented in the National Sample Census of Agricultural Small Holder Crop Sector – National Report, which is often referenced in the literature (NBS, et al. 2012). Tables B-1 and B-2 present the inorganic fertilizer use in the short and long rain season (respectively) by region, while Tables B-3 and B-4 present the same information aggregated by crop. My aggregates differ no more than 1 percent from the National Report.[51] Given these differences, I am confident that the data I am using is the same as that used for the national report.

4.5.1 Results

The regression results from the ASC survey are presented in Table 4-4. The main variable of interest is the travel time from a port to a fertilizer retailer (assumed to be at cities of 20,000 people or greater) which represents the transport costs of moving fertilizer from the port-of-entry in Dar es Salaam inland. A one hour increase in the travel time from a port-to-retailer significantly increased the household's unit fertilizer cost by an estimated 9.71 TSh (regression 1), but this significance disappeared when regional dummies were accounted for (regression 2). Given the literature on the markup of fertilizer prices with respect to transport costs, this is a surprising result. The small magnitude of the effect could indicate that I have not chosen the correct fertilizer market for each household, but a linear regression of fertilizer markets on unit cost indicate that of the 115 markets used, 106 are significant (and 99 of these are highly significant).[52] Thus, the unexpected magnitude is

[51] Any discrepancies between the two aggregates are believed to be due to the definition of the "Area Applied" variable for fertilizer. This variable is reported as a factor, with values equal to "Used for the whole crop", "Used on ¾ of whole crop", "Used on ½ of whole crop", "Used on ¼ of whole crop" and "Used on less than ¼ of whole crop." Since there was no directions as to what "Used on less than ¼ of whole crop" meant, I assumed that this was equal to 12.5 percent of the total area cropped under a particular crop. The aggregates in the National Report may have used a different definition of the category.

[52] The simple regression of fertilizer market on household unit cost of fertilizer was conducted without an intercept to be able to see the effects of each market individually (rather than in reference to a dropped dummy variable). The (significant) coefficients in this regression ranged from 478.6 TSh/kg to 2,777.8 TSh/kg.

likely attributed to measurement error in the travel time metric or unit cost. The roads connecting larger cities (of 20,000 or greater) in Tanzania are more likely to have been identified in the time-to-market model, but they could have been misclassified. More importantly, travel from Dar es Salaam is likely wrought with congestion, which is not accounted for in the time-to-market metrics. Thus, travel time from Dar es Salaam to a fertilizer retailer is likely underestimated.

The inclusion of the regional dummy variable in the regression model dramatically increased the coefficient of determination (R^2) by more than five times. A sizeable number of the villages that contain households which use fertilizer were clustered in certain locals, especially in the southwest regions (see Figure 4-4). The regional dummy may account for the effects of spatial clustering in fertilizer use. A dummy variable indicating whether the household produced (and bought fertilizer) in the short-rain season and the total harvested area were included in regression 3. Both variables were significant and of the expected sign. However, while the significance of short-rain season production persisted, the significant effect of harvested area disappeared when the quantity of fertilizer purchased was included (regression 4).

A sizable amount of variation in household unit fertilizer costs was attributable to the quantity purchased. Limited access to finance can constrain the size of fertilizer purchased, but fertilizer is usually bought by the retailer in 50 kg bags. Retailers may charge a higher price to break up and repackage fertilizer. Additionally, it is possible that they offer discounted prices for purchases made in bulk. Both of these notions are supported by the data. Relative to buying up to a quarter of a 50 kg bag of fertilizer, the effect of higher purchase quantities significantly decreased the unit cost of fertilizer, at an increasing rate.[53] This effect is more dramatic at lower quantities (i.e., moving from zero to a ¼ bag or ¼ to a ½ bag rather than moving from 1 to 2 bags or 2 to 4 bags). Somewhat surprising was that the majority of households sample bought 1 to 2 – 50 kg bags of fertilizer. While higher quantities of fertilizer purchases were associated with larger farm sizes, the average farm size for household that bought 1 to 2 – 50 kg bags of fertilizer was 0.72 hectares, which is

[53] Depending on the nature of measurement error in quantity and cost and how the respondent calculated the two variables, the negative relationship between quantity and unit costs could be spurious.

rather small. By all accounts in the literature, the 2003 transport subsidy had ended by 2007 and the 2008 voucher system had not yet started at the time of the survey. However, the higher quantities of purchase may have been a residual effect of the subsidy. Regression 5 incorporated both an interaction between each of the quantity variables and a time-to-market metric from the household's village to the fertilizer retailer, as well as farmer demographics (i.e., age, sex, education, literacy or income diversification). In a univariate regression, time village-to-retailer travel time was significant and positive, but this significance disappeared once the variable was interacted with quantity. None of the interaction terms were significant and only the sex (relative to a female head of household) demographic variable was significant (albeit positive and small in magnitude), however the inclusion of these terms did increase the coefficient of determination by six percent.

[Table 4-4: OLS regressions on household-level fertilizer unit costs, ASC survey]

4.5.2 Sensitivity Analysis

The fact that only region and quantity purchased appear to affect the household unit cost of fertilizer was surprising. This could be the result of measurement error in the variables used, especially the ASC variables on quantity and cost. The ASC survey was conducted in the summer of 2009, and asked farmers to recall the details of their production from the two seasons prior (2007/08). It would not be unreasonable to assume that farmers did not remember the exact details of their decisions, especially if they failed to keep written records. Another contributing factor to the small size of the effects could be attributed to omitted variables. From January 2007 to December 2008, the world price of DAP fluctuated between 1 to 2.5 times as high as the world price of Urea, and Table 4-1 shows different unit transport costs among fertilizers. The ASC does not contain any information on fertilizer type, which would likely help explain some of the differences in prices. However, the 2008/09 LSMS-ISA National Panel Survey (NPS) in Tanzania collects similar information as the ASC and does include a variable on fertilizer type. The NPS was also collected in 2009 to represent production in 2007/08, so it may suffer from the same recall issues as the ASC (depending on the questioning style of the enumerators).

Tanzania's LSMS survey is a nationally representative panel survey conducted over four waves on 3,265 original households. Of these households, 270 (8.3 percent) reported purchasing inorganic fertilizer during the 2007/08 agricultural year. The NPS included village-level geo-coordinates with its survey, and unit fertilizer costs were calculated in the same manner as the ASC; total cost of purchased fertilizer/total quantity of purchased fertilizer.[54] The results from the multivariate OLS regressions for NPS are presented in Table 4-5.

[Table 4-5: OLS regressions on plot-level fertilizer unit costs, NPS survey]

Using the NPS data, I found the opposite effects of travel time on fertilizer unit costs (regression 1). The effect of an additional hour in travel time between the port and retailer significantly decreased plot-level unit fertilizer costs by an estimated 30.69 TSh/kg (approximately US$ 0.03 per kilogram), and the sign and significance persisted throughout the rest of the regressions. Similarly, adding regional dummy variables (regression 2) substantially increased the coefficient of determination. Production in the short-rain season did not have a statistically significant effect on the fertilizer unit cost, but that was unsurprising since only 1.9 percent of households that used fertilizer did so in the short-rain season. In terms of quantity purchased, there did appear to be penalties for breaking up a 50 kg bag of fertilizer into smaller units, but not bulk discounts (either from harvested area or quantity purchases greater than one 50 kg bag of fertilizer) (regression 3). The significance on these small purchases did not persist once fertilizer type was controlled for (regression 4). There were eight types of fertilizer accounted for in the survey (the effects of different fertilizer type on unit costs are interpreted relative to DAP). Of the plots that were fertilized in the 2007/08 agricultural season, 58.9 percent were fertilized with Urea and DAP accounted for 14.0 percent of the fertilizer used. While the addition of fertilizer type increases the coefficient of determination by 7.7 percent, none of the fertilizer types significantly affected the unit cost, relative to DAP. The inclusion of interactions between village-to-retailer time and quantity and head of household demographics also increased the level of explained variation in the unit fertilizer cost, but did not have significant effects

[54] In the NPS survey, information is collected at the plot level (as opposed to the crop level in the ASC). To estimate the effects of fertilizer type on fertilizer unit costs, I had to use plot level unit costs rather than the average household unit costs used in the previous regressions on the ASC data. Regardless, regressions using the other independent variables in the NPS survey on average household unit costs produced similar results to those shown in Table 4-5.

on unit costs. The strange result on the negative effect of travel time on unit costs and the insignificance of quantity purchased could be due to the fact that the price data are too coarsely specified to be of much value in a plot- (or household-) level analysis.

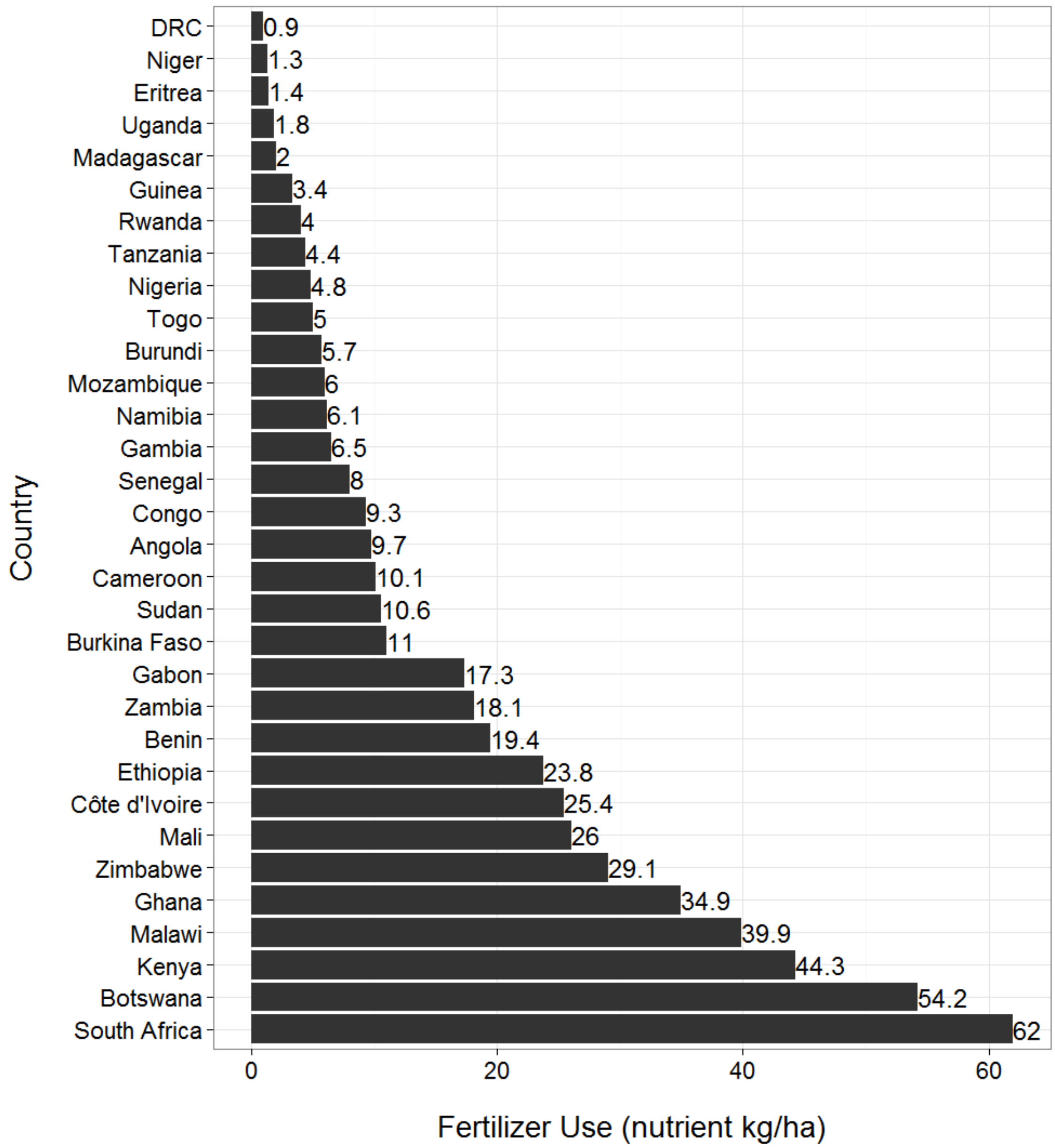
4.6 Conclusions

Input price data are important because they impact the production decisions made by agricultural households which directly affect the income and welfare of the household. Additionally, the relative levels of input and output prices influence the overall agricultural development of a country. As increasingly more people move into the cities, the burden on the rural population to feed their urban counterparts is much greater. Efforts to improve production capacities in these rural areas require increased input use, specifically fertilizer use. However, fertilizer usage rates in sub-Saharan Africa are some of the lowest in the world. While the development literature acknowledges the price of fertilizer as a major deterrent to increasing fertilizer use, most analyses use aspatial prices to investigate the profitability of the input. Agriculture is an inherently localized industry, thus, any assessment using aspatial prices will not reflect the varied economic realities affecting fertilizer use decisions on African farms (Guo, Koo and Wood 2009).

However, there are few reliable representations of spatially delineated input prices, including fertilizer prices. One option to address this issue explored in this analysis is to use the unit costs of fertilizer as reported by households in Tanzania's 2007/08 Agricultural Sample Census. There is quite a lot of variation in these prices, but the maintained hypothesis was that the majority of variation could be attributed to the high costs of transporting the bulky good and will likely vary spatially due to the size and (often fragmented) nature of fertilizer retail markets in Tanzania (Guo, Koo and Wood 2009, Chemonics International and IFDC 2007). This hypothesis was not supported by the data, and those few variables that did significantly affect unit fertilizer costs (i.e., quantity purchased and regional dummies) still did not explain much of the variation in the household unit fertilizer costs.

The variables on quantity and cost of fertilizer purchases appear to be wrought with measurement error. Additionally, there are likely several unobserved variables from this analysis that affect fertilizer prices, such as quality, a quantifiable notion of the household's knowledge about the product, the availability of fertilizers and when and where fertilizer was purchased. One variable that is not collected in the ASC survey is fertilizer type. The world prices of fertilizers varied substantially during the agricultural production year covered by the survey, so I would expect variation in fertilizer unit costs to be attributed to the type of fertilizer used. As they are, any use of these variables in analysis would likely lead to erroneous conclusions, which is probably why there has been a woefully low use of these types of variables in agricultural surveys.

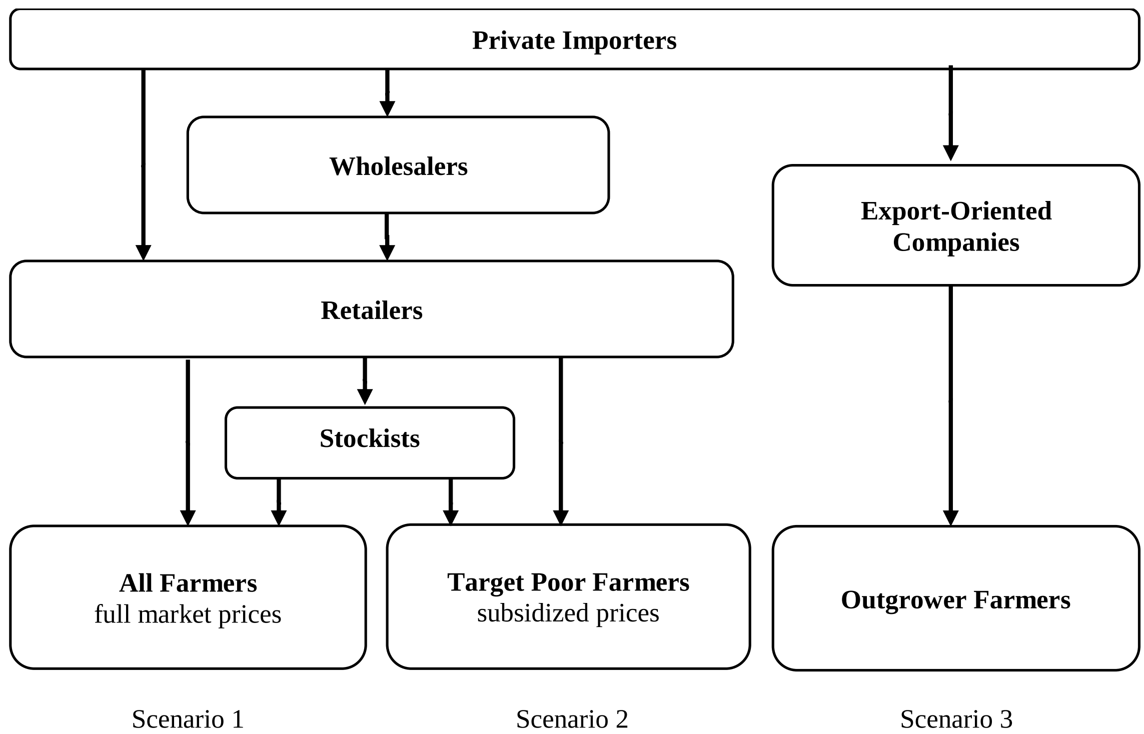
Spatially-delineated fertilizer prices can be used to target better intervention strategies and policies that improve agricultural production and household welfare. But for this data to be useful, better information needs to be collected to understand the more nuanced factors affecting fertilizer uptake by individual households. Future surveys, especially censuses, should include more variables such as the fertilizer type, cost, quantity, location and time of purchase as well as the price of outputs and the use of a voucher or subsidy program. Additionally, variables that reference the level of transaction costs incurred by households, such as cost of transport to and from a retailer will be helpful in creating a more detailed view of the household's incentive structure. Lastly, it is important for surveys to include disaggregated measures of spatial location (e.g., geo-coordinates or high-resolution shapefiles).



Source: Developed by author using data from the FAO (2015).

Note: Fertilizer use was calculated as the aggregated consumption of nitrogen (N), phosphate (P205) and potash (K20) fertilizer (kg of nutrients) per hectare of arable land. Arable land is defined by the FAO as land under temporary agricultural crops. In 2012, Seychelles' fertilizer intensity rate was 116.0 kg/ha and Mauritius' was 224.2 kg/ha. These are not included in Figure 4-1 for visualization purposes.

Figure 4-1: Year 2012 fertilizer use (kg/ha) in sub-Saharan Africa



Source: Recreated from IFDC (2012).

Figure 4-2: Fertilizer supply systems in Tanzania

Table 4-1: Fertilizer cost chains from Dar es Salaam to multiple destinations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Fertilizer Type	Urea	Urea	Urea	DAP	DAP	DAP	CAN	CAN	CAN
Destination	Songea	Mbeya	Iringa	Sumbawanga	Morogoro	Arusha	Kigoma	Songea	Iringa
Travel Distance (km)	1,064	819	494	1,523	184	624	1,369	1,064	494
Travel Time (hr)	14.5	11.6	7.4	21.1	3.3	9.3	19.0	14.5	7.4
Product Cost (FOB + Bagging) (US$/ton)	287	286	286	307	307	308	225	225	225
Transport Cost (US$/ton)	79	83	70	126	73	88	135	103	84
Taxes and Levies (US$/ton)	1	1	1	2	2	2	2	3	3
Finance Costs (US$/ton)	21	20	20	15	15	15	15	14	14
Total Overheads (US$/ton)	10	10	10	7	7	7	7	7	7
Total Margins (US$/ton)	42	43	37	9	9	11	25	27	28
Retail Price (US$/ton)	**439**	**444**	**425**	**466**	**412**	**430**	**409**	**379**	**361**

Source: Chemonics International and IFDC (2007).

Note: FOB – Free on board; DAP – Diammonium phosphate; CAN – Calcium ammonium nitrate. Travel distance and travel time were estimated using Google Maps.

Table 4-2: Fertilizer consumption in Tanzania from 2002 – 2012

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	Share of 2007/08 Average
Consumption quantity in nutrients	**31,818**	**38,050**	**50,247**	**55,818**	**52,357**	**50,716**	**52,966**	**86,533**	**76,255**	**101,346**	**63,783**	**100.0**
Nitrogen Fertilizers (N total nutrients)	22,192	26,590	34,469	33,530	39,222	41,448	43,426	66,942	58,341	56,887	49,486	81.9
Phosphate Fertilizers (P205 total nutrients)	5,281	6,390	9,725	16,825	11,552	8,992	9,264	14,951	8,055	30,451	6,617	17.6
Potash Fertilizers (K20 total nutrients)	4,345	5,070	6,053	5,463	1,583	276	276	4,640	9,859	14,008	7,680	0.5
Consumption quantity by fertilizer	**64,846**	**68,894**	**85,002**	**99,794**	**155,140**	**179,554**	**201,682**	**142,260**	**190,985**	**233,820**	**124,040**	**100.0**
Urea (mt)	30,334	36,150	54,674	46,570	56,822	69,133	73,200	107,167	77,899	49,001	62,410	37.3
NPK complex >10kg (mt)	-	-	-	-	27,680	34,205	54,282	-	43,393	93,443	-	23.2
PK compounds (mt)	-	-	-	2,570	3,847	23,150	23,150	-	5,525	12	-	12.1
Diammonium phosphate (DAP) (mt)	6,515	3,897	10,551	26,588	21,438	19,408	20,000	-	1,405	34,106	-	10.3
Calcium ammonium nitrate (CAN) (mt)	12,577	21,494	12,680	15,460	25,589	12,079	12,079	12,936	22,063	19,934	9,978	6.3
NPK complex (mt)	-	-	-	3,704	1,239	6,031	6,031	-	-	-	-	3.2
Ammonium sulfate (mt)	12,600	4,099	2,593	1,554	4,877	4,620	4,620	3,276	11,147	7,559	18,503	2.4
NPK blends (mt)	-	-	-	-	5,199	3,347	3,347	-	-	-	-	1.8
Potassium nitrate (mt)	-	-	-	-	126	5,560	3,000	-	-	140	-	2.2
Other (mt)	2820	3254	4504	3348	8323	2,021	1,973	18,881	29553	29,625	33,149	1.0

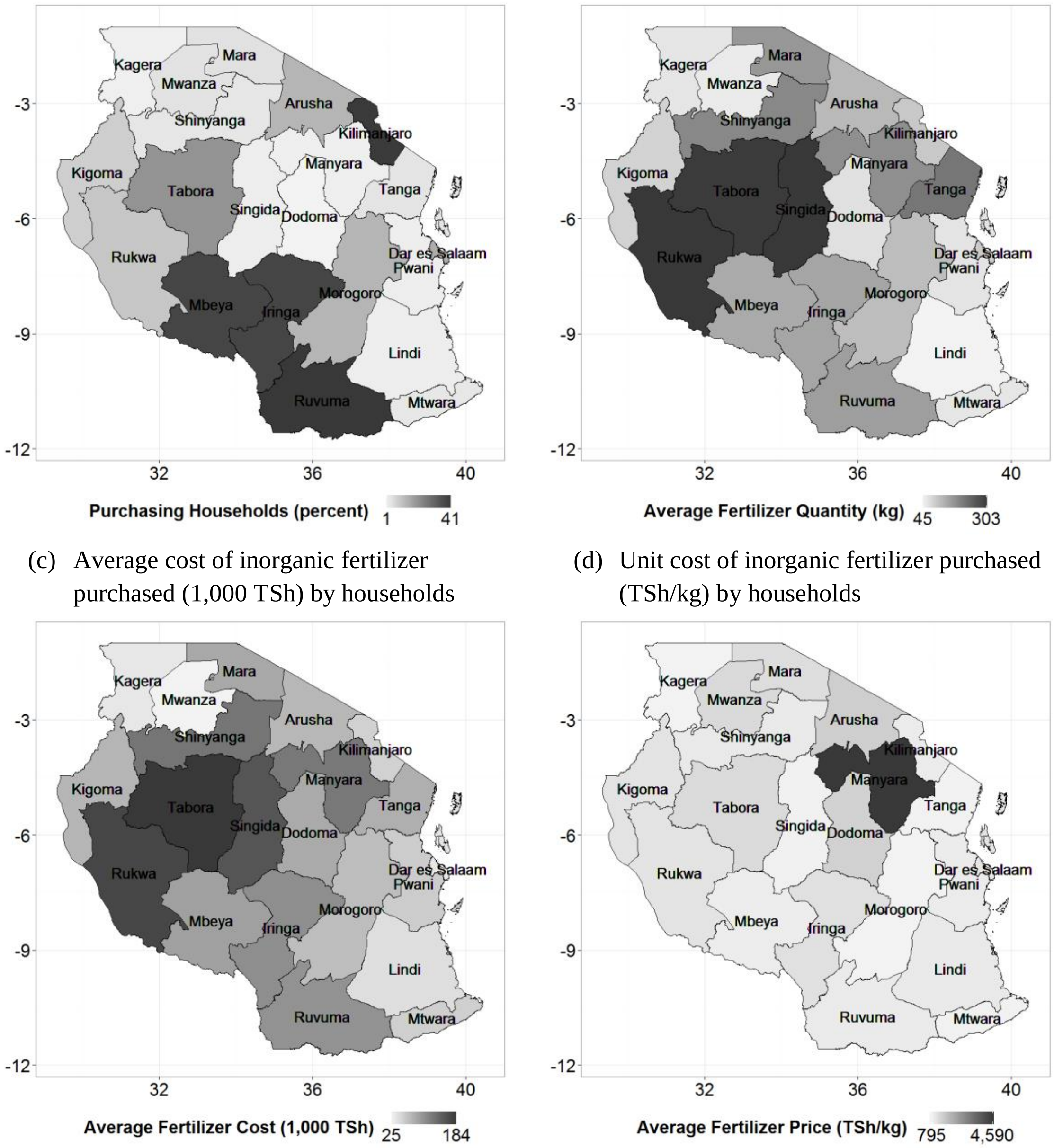
Source: FAOSTAT (2015).

(a) Households purchasing inorganic fertilizer (percent)

(b) Average quantity of inorganic fertilizer purchased (kg) by households

(c) Average cost of inorganic fertilizer purchased (1,000 TSh) by households

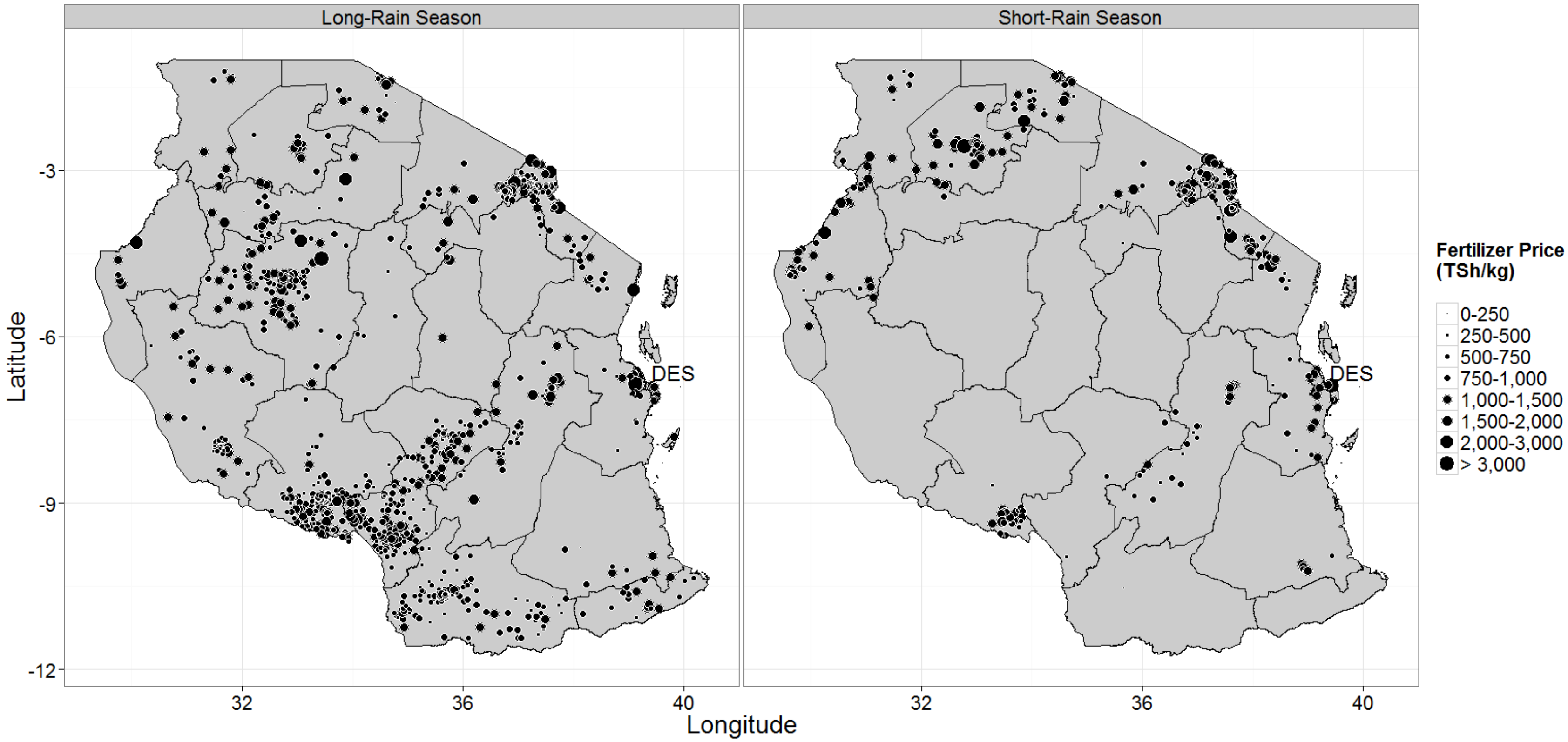
(d) Unit cost of inorganic fertilizer purchased (TSh/kg) by households



Source: Developed using data from Tanzania's ASC (NBS 2011a).

Note: Maps display statistics by region (ADM1 units). Unit cost of fertilizer purchased (panel (d) above) was calculated as the weighted average of household unit costs. Regional unit costs could have also be calculated as the sum of purchased cost (weighted) by the sum of purchased quantity (weighted) by region, which results in regional unit costs that are ten times lower, on average.

Figure 4-3: Descriptive statistics on fertilizer purchases



Source: Developed using data from Tanzania's ASC (NBS 2011a) and own calculations.

Note: Each village plotted has upwards of 15 households surveyed. Average unit costs are mapped by production season.

Figure 4-4: Average household-level unit fertilizer costs by village

Table 4-3: Descriptive statistics on main analysis variables

	Range	Median	Mean	Std. Deviation	Obs.
Port-to-Retail Travel Time (hr)	28.22	11.18	10.95	4.34	5,918
Village-to-Retail Travel Time (hr)	45.77	2.60	4.66	5.70	5,918
Household Head Demographics					
Sex (male = 1)	1.00	1.00	0.81	0.39	5,974
Age (yr)	81.00	42.00	44.64	14.35	5,974
Education (yr)	17.00	7.00	6.09	3.19	5,974
Literate: (yes = 1)	1.00	1.00	0.87	0.34	5,974
Off-Farm Income (yes = 1)	1.00	-	0.42	0.49	5,974
Production Season:					
Long-Rain Season	1.00	1.00	0.88	0.33	5,974
Short-Rain Season	1.00	-	0.18	0.38	5,974
Total Harvested area	19.42	0.73	0.91	0.98	5,974
Quantity of Fertilizer Purchased (kg)					
0 - 0.25 (50-kg bags)	1.00	-	0.07	0.26	5,753
0.25 - 0.50 (50-kg bags)	1.00	-	0.08	0.27	5,753
0.50 - 1.00 (50-kg bags)	1.00	-	0.03	0.17	5,753
1.00 - 2.00 (50-kg bags)	1.00	-	0.24	0.43	5,753
2.00 - 4.00 (50-kg bags)	1.00	-	0.20	0.40	5,753
> 4.00 (50-kg bags)	1.00	-	0.20	0.40	5,753
Cost of Fertilizer Purchase (1,000 TSh)	761.50	60.00	106.87	129.11	5,747
Unit Fertilizer Cost (TSh/kg)	2,941.15	742.86	796.81	415.19	5,356

Source: Developed using data from Tanzania’s ASC (NBS 2011a) and own calculations.

Table 4-4: OLS regressions on household-level fertilizer unit costs, ASC survey

	Fertilizer Unit Cost (TSh/kg)				
	(1)	(2)	(3)	(4)	(5)
Port Time	9.71***	-5.01	-5.43	-5.22	-3.93
	(1.30)	(3.99)	(3.97)	(3.86)	(3.86)
Season (SRS)			99.37***	100.65***	94.52***
			(17.61)	(17.13)	(17.17)
Area			-37.06***	2.8	2.27
			(6.50)	(7.21)	(7.21)
Quantity (0.25-0.5 Bags)				-149.57***	-158.63***
				(28.49)	(37.46)
Quantity (0.5-1 Bags)				-308.64***	-321.94***
				(24.81)	(33.32)
Quantity (1-2 Bags)				-344.05***	-322.56***
				(25.00)	(33.39)
Quantity (2-4 Bags)				-362.74***	-389.35***
				(26.05)	(34.57)
Quantity (> 4 Bags)				-423.02***	-442.83***
				(27.75)	(36.08)
Village Time					4.2
					(4.91)
Constant	696.52***	772.38***	765.89***	948.59***	915.16***
	(15.49)	(80.94)	(80.80)	(80.20)	(83.45)
Region		X	X	X	X
Village Time * Quantity					X
Farmer Demographics					X
Observations	5,321	5,321	5,321	5,321	5,321
R^2	0.010	0.053	0.064	0.116	0.123
Note:				*p<0.1; **p<0.05; ***p<0.01	

Source: Developed using data from Tanzania’s ASC (NBS 2011a).

Note: SRS – Short-Rain Season. The effects of quantity of inorganic fertilizer purchased are with respect to 0.0 – 0.25 50 kilogram bags.

Table 4-5: OLS regressions on plot-level fertilizer unit costs, NPS survey

	Fertilizer Unit Cost (TSh/kg)				
	(1)	(2)	(3)	(4)	(5)
Port Time	-30.39***	-31.10**	-31.35**	-30.80**	-32.35**
	(5.35)	(12.39)	(12.26)	(12.69)	(12.82)
Season (SRS)		4.21	-25.47	-17.41	-14.43
		(163.17)	(160.30)	(160.83)	(159.14)
Area		-12.31	-15.9	-15.28	-7.56
		(12.13)	(17.08)	(17.23)	(17.42)
Quantity (0.25-0.5 Bags)			-109.69*	-121.34**	-96.06
			(58.35)	(59.93)	(77.65)
Quantity (0.5-1 Bags)			-123.19**	-116.81**	-35.85
			(54.58)	(54.99)	(70.07)
Quantity (1-2 Bags)			-101.13	-95.46	30.09
			(64.78)	(65.03)	(89.90)
Quantity (2-4 Bags)			-108.55	-111.08	-141.01
			(77.68)	(79.22)	(109.46)
Quantity (> 4 Bags)			33.68	33.34	-129.16
			(102.78)	(116.15)	(169.27)
Fertilizer (Urea)				-69.66	-52.76
				(60.30)	(59.90)
Fertilizer (TSP)				90.98	123.51
				(185.36)	(184.02)
Fertilizer (CAN)				-33.71	-29.88
				(78.04)	(77.90)
Fertilizer (SA)				65.83	48.37
				(108.21)	(107.84)
Fertilizer (NPK)				-64.34	-43.47
				(110.83)	(110.89)
Fertilizer (MRP)				-261.03	-277.73
				(217.39)	(215.04)
Village Time					18.38
					(23.24)
Constant	1,199.45***	1,370.32***	1,445.52***	1,501.53***	1,418.73***
	(57.34)	(334.14)	(328.13)	(335.93)	(334.98)
Region		X	X	X	X
Village Time * Quantity					X
Demographics					X
Observations	418	418	409	409	409
R^2	0.072	0.248	0.267	0.277	0.304
Note:				*p<0.1; **p<0.05; ***p<0.01	

Source: Developed using data from the Tanzania's NPS (World Bank 2009).

Note: SRS – Short-Rain Season; TSP – Triple Super Phosphate; CAN – Calcium Ammonium Nitrate; SA – Sulphate of Ammonium; NPK – Nitrogen Phosphate Potassium; MRP – Minkingu Rock Phosphate. The effects of quantity of inorganic fertilizer purchased are with respect to 0.0 – 0.25 50 kilogram bags. The effects of fertilizer type are with respect to Diammonium Phosphate (DAP).

References

Aker, J.C. 2010. "Information from Markets Near and Far: Mobile Phones and Agricultural Markets in Niger." *American Economic Journal: Applied Economics* 2 (3): 36-59.

Aker, J.C., and I.M. Mbiti. 2010. "Mobile Phones and Economic Development in Africa." *Center for Global Development Working Paper No. 211*.

Alene, A.D., V.M. Manyong, G. Omanya, H.D. Mignouna, M. Bokanga, and G. Odhiambo. 2008. "Smallholder Market Particpation Under Transaction Costs: Maize Supply and Fertilizer Demand in Kenya." *Food Policy* 33: 318-328.

Anderson, W., L. You, S. Wood, U. Wood-Sichra, and W. Wu. 2015. "An Analysis of Methodological and Spatial Differences in Global Cropping Systems Models and Maps." *Global Ecology and Biogeography* 24: 180-191.

Apparicio, P., M. Abdelmajid, M. Riva, and R. Shermur. 2008. "Comparing Alternative Approaches to Measuring the Geographical Accessibility of Urban Health Services: Distance Types and Aggregation-Error Issues." *International Journal of Health Geographics* 7 (7).

Arino, O., J.J. Ramos Perez, V. Kalogirou, S. Bontemps, P. Defourny, and E. Van Bogaert. 2012. "Global Land Cover Map for 2009 (GlobCover 2009)." European Space Agency (ESA) & Université Catholique de Louvain (UCL).

Balk, D. L., U. Deichmann, G. Yetman, F. Pozzi, S. I. Hay, and A. Nelson. 2006. "Determining Global Population Distribution: Methods, Applications and Data." *Advances in Parasitology* 62: 119-156.

Barrett, C.B. 2008. "Smallholder Market Participation: Concepts and Evidence from Eastern and Southern Africa." *Food Policy* 33: 299-317.

Barrett, C.B., T. Reardon, and P. Webb. 2001. "Nonfarm Income Diversification and Household Livelihood Strategies in Rural Africa: Concepts, Dynamics, and Policy Implications." *Food Policy* 26 (4): 315-331.

Beddow, J.M., and P.G. Pardey. 2015. "Moving Matters: The Effect of Location on Crop Production." *Journal of Economic History* 75 (1): 219-249.

Beddow, J.M., T.M. Hurley, and P.G. Pardey. 2014. "Food Security: Yield Gap." *Encyclopedia of Agriculture and Food Systems* 3: 352-365.

Benin, S., A.N. Pratt, S. Wood, and Z. Guo. 2011. "Trends and Spatial Patterns in Agricultural Productivity in Africa, 1961 - 2010." *ReSAKSS Annual Trends and Outlook Report 2011.* International Food Policy Research Institute (IFPRI).

Benson, T., S.L. Kirama, and O. Selejio. 2012. "The Supply of Inorganic Fertilizers to Smallholder Farmers in Tanzania: Evidence for Fertilizer Policy Development." *IFPRI Discussion Paper 01230.* International Food Policy Research Institute.

Burke, W.J. 2012. "Maize Production in Zambia and Regiona Marketing: Input Productivity and Output Price Transmission." *Ph.D. Dissertation.* Lansing, MI: Michigan State University.

Buys, P., U. Deichmann, and D. Vheeler. 2006. "Road Network Upgrading and Overland Trade Expansion in Sub-Saharan Africa." World Bank Policy Research Working Paper 4097.

CGIAR-CSI. 2008. "SRTM Data Processing Methodology." Accessed June 2010. http://srtm.csi.cgiar.org/SRTMdataProcessingMethodology.asp.

Chamberlin, J., and T.S. Jayne. 2013. "Unpacking the Meaning of 'Market Access': Evidence from Rural Kenya." *World Development* 41: 245-264.

Chambers, R. 1983. *Rural Development: Putting the Last First.* London: Longman.

Chemonics International and IFDC. 2007. "Fertilizer Supply and Costs in Africa." *Report for the Bill and Melinda Gates Foundation.* Seattle, WA: Bill and Melinda Gates Foundation.

CIESIN, IFPRI, the World Bank and CIAT. 2011. "The Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Population Count Grid." Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC). Accessed January 2015. http://sedac.ciesin.columbia.edu/data/set/grump-v1-population-count.

—. 2011. "The Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Urban Extents Grid." Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). Accessed June 2010. http://dx.doi.org/10.7927/H4M906KR.

de Boer, P.T., D.P. Kroese, S. Mannor, and R.Y. Rubinstein. 2005. "A Tutorial on the Cross-Entropy Method." *Annals on Operations Research* 134: 19-67.

de Janvry, A., M. Fafchamps, and E. Sadoulet. 1991. "Peasant Household Behavior with Missing Markets: Some Paradoxes Explained." *The Economic Journal* 101: 1400-1417.

Deaton, A. 1988. "AssociationQuality, Quantity, and Spatial Variation of Price." *The American Economic Review* 78 (3): 418-430.

—. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy.* Washington, D.C.: World Bank Publications.

Demeke, M., V. Kelly, T.S. Jayne, A. Said, J.C. Le Vallée, and H. Chen. 1998. "Agricultural Market Performance and Determinants of Fertilizer Use in Ethiopia." *Working Paper 10, Grain Market Research Project.* Addis Ababa: Ministry of Economic Development and Cooperation.

Dillon, B., and C.B. Barrett. 2014. "Agricultural Factor Markets in Sub-Saharan Africa: An Updated View with Formal Tests for Market Failure." *Policy Research Working Paper 7117.* International Bank for Reconstruction and Development/World Bank.

Dixon, J., A. Gulliver, and D. Gibbon. 2001. *Farming Systems and Poverty: Improving Farmers' Livelihoods in a Changing World.* Rome: Food and Agriculture Organization and Washington, D.C.: World Bank.

Druilhe, A., and J. Barreiro-Hurlé. 2012. "Fertilizer Subsidies in Sub-Saharan Africa." *ESA Working Paper No. 12-04.* Food and Agricultural Organization of the United Nations (FAO).

Elbers, C., J.O. Lanjouw, and P. Lanjouw. 2003. "Micro-Level Estimation of Poverty and Inequality." *Econometrica* 71: 355-364.

European Commison. 2012. "Crops Products." *Eurostat.* Eurostat. Accessed May 2012. http://ec.europa.eu/eurostat/data/database.

Fafchamps, M. 2004. *Market Institutions in sub-Saharan Africa.* Boston: Massachusetts Institute of Technology.

FAO. 2012. *AGROMAPS.* Rome: Food and Agriculture Organization of the United Nations. Accessed June 2012. http://data.fao.org/ref/7c7aaeb7-7de5-45b2-9024-fbb4315dc771.html?version=1.0.

—. 2015. *AQUASTAT.* Rome: Food and Agriculture Organization of the United Nations. Accessed June 2012. http://www.fao.org/nr/water/aquastat/data/query/index.html?lang=en.

—. 2012. *CountrySTAT.* Rome: Food and Agriculture Organization of the United Nations. Accessed June 2012. http://www.countrystat.org/.

FAO. 2015. *Current World Fertilizer Trends and Outlook to 2018.* Rome: Food and Agriculture Organization of the United Nations.

—. 2015. "FAOSTAT Database." Food and Agricultural Organization of the United Nations.

—. 2007. "FertStat." Rome: Food and Agriculture Organization of the United Nations. Accessed August 2015. http://www.fao.org/ag/agp/fertistat/index_en.htm.

—. 2012. "Gross Production Value." *FAOSTAT.* Rome: Food and Agriculture Organization of the United Nations. Accessed May 2012. http://faostat3.fao.org/.

FEWS-NET. 2013. *East Africa Tanzania: Seasonal Calendar - Typical Year December 2013.* Famine Early Warning Systems Network (FEWS-NET). Accessed January 15, 2015. http://www.fews.net/east-africa/tanzania/seasonal-calendar/december-2013.

Fischer, G., F.O. Nachtergaele, S. Prieler, E. Teixeira, G. Tóth, H. van Velthuizen, L. Verelst, and D. Wiberg. 2013. "Global Agro-Ecological Zones (GAEZ) Version 3.0." Accessed October 2014. http://www.fao.org/nr/gaez/en/.

Foley, J.A., N. Ramankutty, K.A. Brauman, E.S. Cassidy, J.S. Gerber, M. Johnston, N.D. Mueller, et al. 2011. "Solutions for a Cultivated Planet." *Nature* 478: 337-342.

Franch, B., E.F. Vermote, I. Becker-Reshef, M. Claverie, J. Huang, J. Zhang, C. Justice, and J.A. Sobrino. 2015. "Improving the Timeliness of Winter Wheat Production Forecast in the United States of America, Ukraine and China using MODIS data and NCAR Growing Degree Day Information." *Remote Sensing of Environment* 161: 131-148.

Fritz, S., L. See, I. McCallum, L. You, A. Bun, E. Moltchanova, and M. Duerauer. 2015. "Mapping Global Cropland and Field Size." *Global Change Biology* 21 (5): 1980-1992.

Godfray, H.C.J., J.R. Beddington, I.R. Crute, L. Haddad, D. Lawrence, J.F. Muir, J. Pretty, S. Robinson, S.M. Thomas, and C. Toulmin. 2010. "Food Security: The Challenge of Feeding 9 Billion People." *Science* 327: 812-818.

Gregory, D.I., and B.L. Bumb. 2006. "Factors Affecting Supply of Fertilizer in Sub-Saharan Africa." *Agriculture and Rural Development Discussion Paper 24.* Washington, D.C.: The World Bank.

Groves, R.M. 1989. *Survey Errors and Survey Costs.* New York: John Wiley and Sons.

Guo, Z,, A.B. Joglekar, and J.M. Beddow. forthcoming. "Travel Time to Agricultural Markets in Sub-Saharan Africa: Technical Documentation." *HarvestChoice*

Working Paper. Washington, D.C.: International Food Policy Research Institute (IFPRI) and St. Paul: International Science and Technology Practice and Policy (InSTePP) Center, University of Minnesota.

Guo, Z., and U. Wood-Sichra. 2015. "AEZ (16-class, 2009)." *HarvestChoice Data Product.* Washington, D.C.: International Food Policy Research Institute (IFPRI) and St. Paul: International Science and Technology Practice and Policy (InSTePP) Center, University of Minnesota. Accessed August 2015. http://harvestchoice.org/data/aez16_clas.

Guo, Z., J. Koo, and S. Wood. 2009. "Fertilizer Profitability in East Africa: A Spatially Explicit Policy Analysis." *Contributed Paper prepared for presentation at the International Association of Agricultural Economists Conference.* Beijing, August 16-22.

Gwilliam, K., H. Bofinger, R. Bullock, R. Carruthers, A. Kumar, M. Mundy, A. Nogales, and K. Sethi. 2011. "Roads: The Burden of Maintenance." In *Africa's Transport Infrastructure: Mainstreaming Maintenance and Management*, edited by V. Foster and C. Briceño-Garmendia. Washington, D.C.: The International Bank for Reconstruction and Development and the World Bank.

Hartley, A.J., A. Nelson, P. Mayaux, and J-M. Gregoire. 2007. "The Assessment of African Protected Areas: A Characterization of Biodiversity Value Ecosystems and Threats, to Inform the Effective Allocation of Conservation Funding." *JRC Scientific and Technical Reports.* Institute for Environment and Sustainability, European Commission Joint Research Centre.

Hutabarat, B., A. Setiyanto, R. Kustiari, and T.B. Sulser. 2012. "Conjecturing Production, Imports and Consumption of Horticulture in Indonesia in 2050: A GAMS Simulation through Changes in Yields Induced by Climate Change." *Agro Economic Journal* 30: 23 pp.

IFDC. 2007. "Africa Fertilizer Summit Proceedings: June 9-13, 2006, Abuja, Nigeria." Muscle Shoals, Alabama: International Center for Soil Fertility and Agricultural Development (IFDC).

—. 2012. "Tanzania Fertilizer Assessment." *In Support of the African Fertiliser and Agribusiness Partnership.* Muscle Shoals, Alabama: International Fertiliser Development Center.

IIASA and FAO. 2012. "Model Documentation." *Global Agro-ecological Zones (GAEZ v3.0).* Laxenburg, Austria and Rome, Italy: International Institute for Applied Systems Analysis and the Food and Agriculture Organization of the United Nations.

Jacoby, H.G. 2000. "Access to Markets and the Benefits of Rural Roads." *The Economic Journal* 110: 713-737.

Jagwe, J.N., and C. Machethe. 2011. "Effects of Transaction Costs on Choice of Selling Point: A Case of Smallholder Banana Growers in the Great Lakes Region of Central Africa." *Agricultural Economics Research, Policy and Practice in Southern Africa* 50 (3): 109-123.

Jarvis, A., H.I. Reuter, A. Nelson, and E. Guevara. 2008. "Hole-filled SRTM for the Globe Version 4 Available from the CGIAR-CSI SRTM 90m Database." Accessed June 2010. http://srtm.csi.cgiar.org.

Jayne, T.S., and S. Rashid. 2013. "Input Subsidy Programs in Sub-Saharan Africa: A Synthesis of Recent Evidence." *Agricultural Economics* 44 (6): 547-562.

Johnson, M., H. Takeshima, and K. Gyimah-Brempong. 2013. "Assessing the Potential and Policy Alternatives for Achieving Rice Competitiveness and Growth in Nigeria." *IFPRI Discussion Paper 01301*. Washington, D.C.: International Food Policy Research Institute (IFPRI).

Kamhabwa, F. 2014. "Consumption of Fertilizers and Fertilizer Use by Crop in Tanzania." AfricaFertilizer.org/International Fertilizer Development Center.

Kelly, V.A. 2006. "Factors Affecting Demand for Fertilizer in Sub-Saharan Africa." *Agriculture and Rural Development Discussion Paper 23.* Washington, D.C.: The World Bank.

Key, N., E. Sadoulet, and A. de Janvry. 2000. "Transactions Costs and Agricultural Household Supply Response." *American Journal of Agricultural Economics* 82: 245-259.

LSMS-ISA. 2015. *Living Standards Measurement Study - Integrated Surveys on Agriculture.* The World Bank. Accessed July 2015. http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTLSMS/0,,contentMDK:21610833~pagePK:64168427~piPK:64168435~theSitePK:3358997,00.html.

Marenya, P.P., and C.B. Barrett. 2009. "State-Conditional Fertilizer Yield Response on Western Kenya Farms." *American Journal of Agricultural Economics* 91: 991-1006.

Matsumoto, T., and T. Yamano. 2011. "Optimal Fertilizer Use on Maize Production in East Africa." In *Emerging Development of Agriculture in East Africa: Markets, Soil and Innovations*, edited by T. Yamano, K. Otsuka and F. Place, 117-132. New York: Springer.

Minot, N.W. 2010. "Stape Food Prices in Tanzania." *Prepared for the Comesa Policy Seminar on "Variation in Staple Food Prices: Causes, Consequence and Policy Options", Maputo, Mozambique, 25-26 January 2010 under the African Agricultural Marketing Project (AAMP).*

Minten, B., B. Koru, and D. Stifel. 2013. "The Last Mile(s) in Modern Input Distribution: Pricing, Profitability, and Adoption." *Agricultural Economics* 44 (6): 629-646.

Monfreda, C., N. Ramankutty, and J.A. Foley. 2008. "Geographic Distribution of Crop Areas, Yields, Physiological Types, and Net Primary Production in the Year 2000." *Global Biogeochemical Cycles* 22: pp19.

Morris, M., V.A. Kelly, R.J. Kopicki, and D. Byerlee. 2007. *Fertilizer Use in African Agriculture: Lessons Learned and Good Practice Guidelines.* Washington, D.C.: The World Bank.

NBS. 2011a. "Agricultural Sample Census Survey 2007/2008 version 1.0." National Bureau of Statistics. Accessed June 2014. www.nbs.go.tz.

NBS. 2011b. *National Sample Census of Agriculture 2007/2008.* Volume 1: Technical and Operation Report, National Bureau of Statistics, United Republic of Tanzania.

NBS, Food Security and Cooperatives Ministry of Agriculture, Ministry of Livestock Development and Fisheries, Ministry of Water and Irrigation, Livestock and Enivronment, Zanzibar Ministry of Agriculture, Regional Administration and Local Government Prime Minister's Office, Trade and Marketing Ministry of Industries, and Zanzibar Office of the Chief Government Statistian. 2012. "National Sample Census of Agriculture 2007/2008: Small Holder Agriculture. Vol. 2, Crop Sector--National Report." Dar es Salaam: The National Bureau of Statistics.

Nelson, A. 2008. "Estimated Travel Time to the Nearest City of 50,000 or More People in Year 2000." Ispra: Global Environment Monitoring Unit - Joint Research Centre of the European Commisssion. Accessed June 2014. http://bioval.jrc.ec.europa.eu/products/gam/.

NIMA. 2000. "Vector Map Level 0 (Digital Chart of the World)." Bethesda, MD: National Imagery and Mapping Agency. Accessed June 2010. http://www.mapability.com/index1.html?http&&&www.mapability.com/info/vmap0_index.html.

Omamo, S.W. 1998a. "Farm-to-Market Transaction Costs and Specialization in Small-Scale Agriculture: Explorations with a Non-Separable Household Model." *Journal of Development Studies* 35: 152-163.

Omamo, S.W. 1998b. "Transport Costs and Smallholder Cropping Choices: An Application to Siaya District, Kenya." *American Journal of Agricultural Economics* 80: 116-123.

Omamo, S.W., and L.O. Mose. 2001. "Fertilizer Trade Under Market Liberalization: Preliminary Evidence From Kenya." *Food Policy* 26 (1): 1-10.

OpenMicroData. 2010. *Tanzania: Shapefiles for EAs, Villages, Districts and Regions.* December 16. Accessed June 2014. https://openmicrodata.wordpress.com/2010/12/16/tanzania-shapefiles-for-eas-villages-districts-and-regions/.

Pardey, P.G., I. Luby, C. Chang-Kang, J.M. Beddow, F.G. Liebenberg, and C. Azzarri. Forthcoming. "African Farming Landscapes Database, Version 2.0 Documentation." *HarvestChoice Working Paper.* St. Paul: International Science and Technology Practice and Policy (InSTePP) Center, University of Minnesota and Washington, D.C.: International Food Policy Research Institute (IFPRI).

Porter, G. 2002. "Improving Mobility and Access for the Off-Road Rural Poor through Intermediate Means of Transport." *World Transport Policy and Practice* 4: 6-19.

Porter, G. 2002. "Living in a Walking World: Rural Mobility and Social Equity Issues in Sub-Saharan Africa." *World Development* 30 (2): 285-300.

Portmann, F.T., S. Siebert, and P. Döll. 2010. "MIRCA2000-Global Monthly Irrigated and Rainfed Crop Areas Around the Year 2000: A New High Resolution Data Set for Agricultural and Hydrological Modeling." *Global Biogeochemical Cycles* 24.

Pozzi, F., and T. Robinson. 2008. "Accessibility Mapping in the Horn of Africa: Applications for Livestock Policy." *IGAD LPI Working Paper No. 11 - 08.* IGAD Livestock Policy Initiative.

—. 2008. "Accessibility Mapping in the Horn of Africa: Applications for Livestock Policy." *IGAD LPI Working Paper No. 11 - 08.* Intergovernmental Authority on Development Livestock Policy Initiative (IGAD LPI).

Renkow, M., D.G. Hallstrom, and D.D. Karanja. 2004. "Rural Infrastructure, Transaction Costs and Market Particpation in Kenya." *Journal of Development Economics* 73: 349-367.

Rowhani, P., D.B. Lobell, M. Linderman, and N. Ramankutty. 2011. "Climate Variability and Crop Production in Tanzania." *Agricultural and Forest Meteorology* 151: 449-460.

Sheahan, M., R. Black, and T.S. Jayne. 2013. "Are Kenyan Farmers Under-Utilizing Fertilizer? Implications for Input Intensification Strategies and Research." *Food Policy* 41: 39-52.

Siebert, S., P. Döll, S. Feick, K. Frenken, and J. Hoogeveen. 2007. "Global Map of Irrigated Areas Version 4.0.1." Frankfurt, Germany and Rome, Italy: University of Frankfurt (Main) and the Food and Agriculture Organization of the United Nations. Accessed July 2012. http://www.fao.org/nr/water/aquastat/irrigationmap/index10.stm.

Staal, S., C. Delgado, and C. Nicholson. 1997. "Smallholder Dairying Under Transaction Costs." *World Development* 25: 779-794.

Staal, S., I. Baltenweck, M. Waithaka, T. DeWolff, and L. Njoroge. 2002. "Location and Uptake: Integrated Household and GIS Analysis of Technology Adoption and Land Use, with Application to Smallholder Dairy Farms in Kenya." *Agricultural Economics* 27: 295-315.

Stifel, D., and B. Minten. 2008. "Isolation and Agricultural Productivity." *Agricultural Economics* 39: 1-15.

Stifel, D., B. Minten, and P. Dorosh. 2003. "Transaction Costs and Agricultural Productivity: Implications of Isolation for Rural Poverty in Madagascar." *MSSD Discussion Paper No. 56.* Washington, D.C.: IFPRI.

TACAIDS, ZAC, NBS, OCGS and ICF International. 2013. *Tanzania HIV/AIDS and Malaria Indicator Survey 2011-12.* Summary of Findings, Tanzania Commission for AIDS (TACAIDS), Zanzibar AIDS Commission (ZAC), National Bureau of Statistics (NBS), Office of Chief Government Statistician (OCGS) and ICF International.

Thapa, S. 2012. "Agribusiness Indicators: Tanzania." *Agriculture and Environmental Services Report: Number 74167.* Edited by G. Dixie. World Bank.

United Nations. 2012. "United Nations. World Urbanization Prospects—The 2012 Revision." New York: United Nations. http://esa.un.org/unpd/wup/.

—. 2015. "World Population Prospects: The 2015 Revision." New York: United Nations, Department of Economic and Social Affairs. http://esa.un.org/unpd/wpp/Publications/Files/Key_Findings_WPP_2015.pdf.

—. 2014. "World Urbanization Prospects: The 2014 Revision." *CD-ROM Edition.* United Nations, Department of Economic and Social Affairs, Population Division. Accessed November 2015.

USDA - NASS. 2004. "Cropland Data Layer: 2004." Washington, D.C.: United States Department of Agriculture - National Agricultural Statistics Service. Accessed September 2015. http://nassgeodata.gmu.edu/CropScape.

—. 2005. "Cropland Data Layer: 2005." Washington, D.C.: United States Department of Agriculture - National Agricultural Statistics Service. Accessed September 2015. http://nassgeodata.gmu.edu/CropScape.

—. 2006. "Cropland Data Layer: 2006." Washington, D.C.: United States Department of Agriculture - National Agricultural Statistics Service. Accessed September 2015. http://nassgeodata.gmu.edu/CropScape.

Verburg, P.H., E.C. Ellis, and A. Letourneau. 2011. "A Global Assessment of Market Accessibility and Market Influence for Global Environmental Change Studies." *Environmental Research Letters* 6: 12pp.

Wood, S. 2007. "Spatial Dimensions of the Regional Evaluation of Agricultural Livelihood Strategies: Insights from Uganda." *Ph.D. Dissertation.* Department of Economics, University of London.

Wood-Sichra, U., A. Joglekar, and L. You. forthcoming. "Spatial Production Allocation Model (SPAM) 2005: Technical Documentation." *HarvestChoice Working Paper.* Washington, D.C.: International Food Policy Research Institute (IFPRI) and St. Paul: International Science and Technology Practice and Policy (InSTePP) Center, University of Minnesota.

World Bank. 2015. "Commodity Price Data (Pink Sheet)." *Global Economic Prospects Data.* The World Bank. Accessed December 2014.

—. 2009. "Tanzania Living Standards Measurement Study - Integrated Survey on Agriculture 2008-2009." The World Bank.

World Gazetteer. 2010. "Population Figures for Cities, Towns, and Places of All Countries." GeoNames. Accessed June 2010. http://download.geonames.org/export/dump/.

WorldPop. 2015. "Version 1.0 2010 Estimates of Total Number of People Per Grid Square Across Africa, with National Totals Adjusted to Match UN Population Division Estimates, 2012 Revision (http://esa.un.org/wpp/)." Accessed June 2015. www.worldpop.org.uk.

Yoshida, N., and U. Deichmann. 2009. "Measurement of Accessibility and Its Applications." *Journal of Infastructure Development* 1 (1): 1-16.

You, L., and S. Wood. 2006. "An Entropy Approach to Spatial Disaggregation of Agricultural Production." *Agricultural Systems* 90: 329-347.

You, L., S. Wood, and U. Wood-Sichra. 2009. "Generating Plausible Crop Distribution Maps for Sub-Saharan Africa Using a Spatially Disaggregated Data Fusion and Optimization Approach." *Agricultural Systems* 99: 126-140.

You, L., S. Wood, U. Wood-Sichra, and W. Wu. 2014. "Generating Global Crop Distribution Maps: From Census to Grid." *Agricultural Systems* 127: 53-60.

You, L., U. Wood-Sichra, S. Fritz, Z. Guo, L. See, and J. Koo. 2015. "Spatial Production Allocation Model (SPAM) 2005 v2.0." Accessed June 2015. http://mapspam.info.

Appendices

Appendix A: Methodology Outline for Alternative Allocation Method

The alternative allocation method used in Chapter 2 follows the equal-probabilistic distribution model used to calculate the M3-Crops data (Monfreda, Ramankutty and Foley 2008). The specific variables and processes involved are outlined below:

Table A-1: Variables from initial data collection

Variables	Definition
$CropH_{jk}$	Total harvested area (H) by crop j and administrative unit k ($k \in k_0, k_1$), calculated from allocated SPAM estimates $$\sum_{i \in k} \sum_{l} AdjAllocH_{ijl}$$
$CropP_{jk}$	Total production (P) by crop j and administrative unit k ($k \in k_0, k_1$), calculated from allocated SPAM estimates $$\sum_{i \in k} \sum_{l} AdjAllocP_{ijl}$$
$CropY_{jk}$	Average harvested area (Y) by crop j and administrative unit k ($k \in k_0, k_1$), calculated from allocated SPAM estimates $$\frac{CropP_{jk}}{CropH_{jk}}$$
$CropLand_i$	Total cropland available by pixel i, equal to SPAM variable $AdjCropLand_i$
$IrrArea_i$	Total irrigated area by pixel i, equal to SPAM variable $AdjIrrArea_i$
$CropIntensity_{jIk}$	Irrigated cropping intensity by crop j, production system I and administrative unit k ($k \in k_0, k_1$), equal to SPAM variable $CropIntensity_{jIk}$
$CropIntensity_{jRk}$	Rain-fed cropping intensity by crop j, production system R and administrative unit k ($k \in k_0, k_1$), equal to weighted average of SPAM rain-fed cropping intensities $$CropIntensity_{jRk} = Percent_{jHk} \times CropIntensity_{jHk} + \left(1 - \left(Percent_{jIk} + Percent_{jHk}\right)\right) \times CropIntensity_{jLk}$$
$PixelArea_i$	Physical area in pixel i

Processing Steps:

(1) Calculate the total cropland area in each political unit

$$CropLand_k = \sum_{i \in k} CropLand_i$$

(2) Determine the ratio of the crop area to total cropland in each political unit

$$ShCropH_{jk} = \frac{CropH_{jk}}{CropLand_k}$$

(3) Calculate the crop harvested area in each grid cell

$$CropH_{ji} = CropLand_i \times ShCropH_{jk}\ \forall i \in k$$

(4) Calculate the crop yield in each grid cell

$$\begin{aligned} CropY_{ji} &= CropY_{jk} \quad && if\ CropH_{ji} > 0 \\ CropY_{ji} &= 0 && if\ CropH_{ji} = 0 \end{aligned}$$

(5) Scale area and yield grids so that country totals match

a. Area

$$SclCropH_{ji} = CropH_{ji} \times \frac{CropH_{jk_0}}{\sum_{i \in k_0} CropH_{ji}}$$

b. Yield

$$SclCropY_{ji} = \frac{\sum_{i \in k_0} \left(SclCropH_{ji} \times CropY_{ji} \right)}{\sum_{l} \left(CropH_{jlk_0} \times CropY_{ji} \right)}$$

Table A-2: New variables after processing

New Variables	Definition
$CropLand_k$	Total cropland in administrative unit k
$ShCropH_{jk}$	Ratio of harvested area to cropland by crop j in administrative unit k
$CropH_{ji}$	Total harvested area by crop j in pixel i
$CropY_{ji}$	Average yield by crop j in pixel i
$SclCropH_{ji}$	Total harvested area by crop j in pixel i, scaled by country total
$SclCropY_{ji}$	Average yield by crop j in pixel i, scaled by country total

Post-Processing Steps:

(1) Calculate multiple harvest ratio

$$MultHarvRatio_i = \frac{\sum_j SclCropH_{ji}}{CropLand_i}$$

(2) Calculate multiple harvesting potential in each grid cell

a. Calculate cropping intensity for each grid cell by irrigated and rain-fed production

i. Average over all crops

$$CropIntensity_{lk} = CropIntensity_{jlk} \times \frac{SclCropH_{ji}}{\sum_j SclCropH_{ji}}, \forall l = I, R$$

ii. Average over all administrative units and set for pixel

$$CropIntensity_{il} = CropIntensity_{lk_1}, \forall i \in k_0 \quad if\ CropIntensity_{lk_1} > 0$$
$$CropIntensity_{il} = CropIntensity_{lk_0}, \forall i \in k_0 \quad Otherwise$$

b. Calculate share of cropland under rain-fed and irrigated production

i. Calculate share of cropland under irrigated production

$$ShIrrArea_i = \frac{IrrArea_i}{PixelArea_i}$$

ii. Calculate share of cropland under rain-fed production

$$ShRainArea_i = 1 - ShIrrArea_i$$

c. Calculate cropping intensity for each pixel

$$CropIntensity_i = ShIrrArea_i \times CropIntensity_{iI} + ShRainArea_i \times CropIntensit_{iR}$$

(3) Adjust cells where estimated harvest ratio exceeds multiple cropping potential

$$AdjCropH_{ji} = SclCropH_{ji} \times \frac{CropIntensity_i}{MultHarvRatio_i} \quad if\ MultHarvRatio_i > CropIntensity_i$$
$$AdjCropH_{ji} = SclCropH_{ji} \quad Otherwise$$

(4) Scale adjusted area to match country total

$$Scl2CropH_{ji} = AdjCropH_{ji} \times \frac{CropH_{jk_0}}{\sum_{i \in k_0} AdjCropH_{ji}}$$

Table A-3: New variables after post-processing

New Variables	Definition
$MultHarvRatio_i$	Ratio of total harvested area in pixel i to total cropland in pixel i
$ShRainArea_i$	Rain-fed production as a share of cropland in pixel i
$ShIrrArea_i$	Irrigated production as a share of cropland in pixel i
$CropIntensity_i$	Area weighted average of cropping seasons in pixel i
$AdjCropH_{ji}$	Harvested area (H) by crop j in pixel i, adjusted by cropping seasons
$Scl2CropH_{ji}$	Harvested area (H) by crop j in pixel i, adjusted by cropping seasons and scaled to country total

Table A-4: Final variables of interest

Variable	Definition
$Scl2CropH_{ji}$	Total harvested area (H) by crop j in pixel i, adjusted by cropping seasons and scaled to country total
$SclCropY_{ji}$	Average yield (Y) by crop j in pixel i, scaled by country total
$SclCropP_{ji}$	Total production (P) by crop j in pixel i, adjusted by cropping seasons and scaled to country total $$SclCropP_{ji} = Scl2CropH_{ji} \times SclCropY_{ji}$$

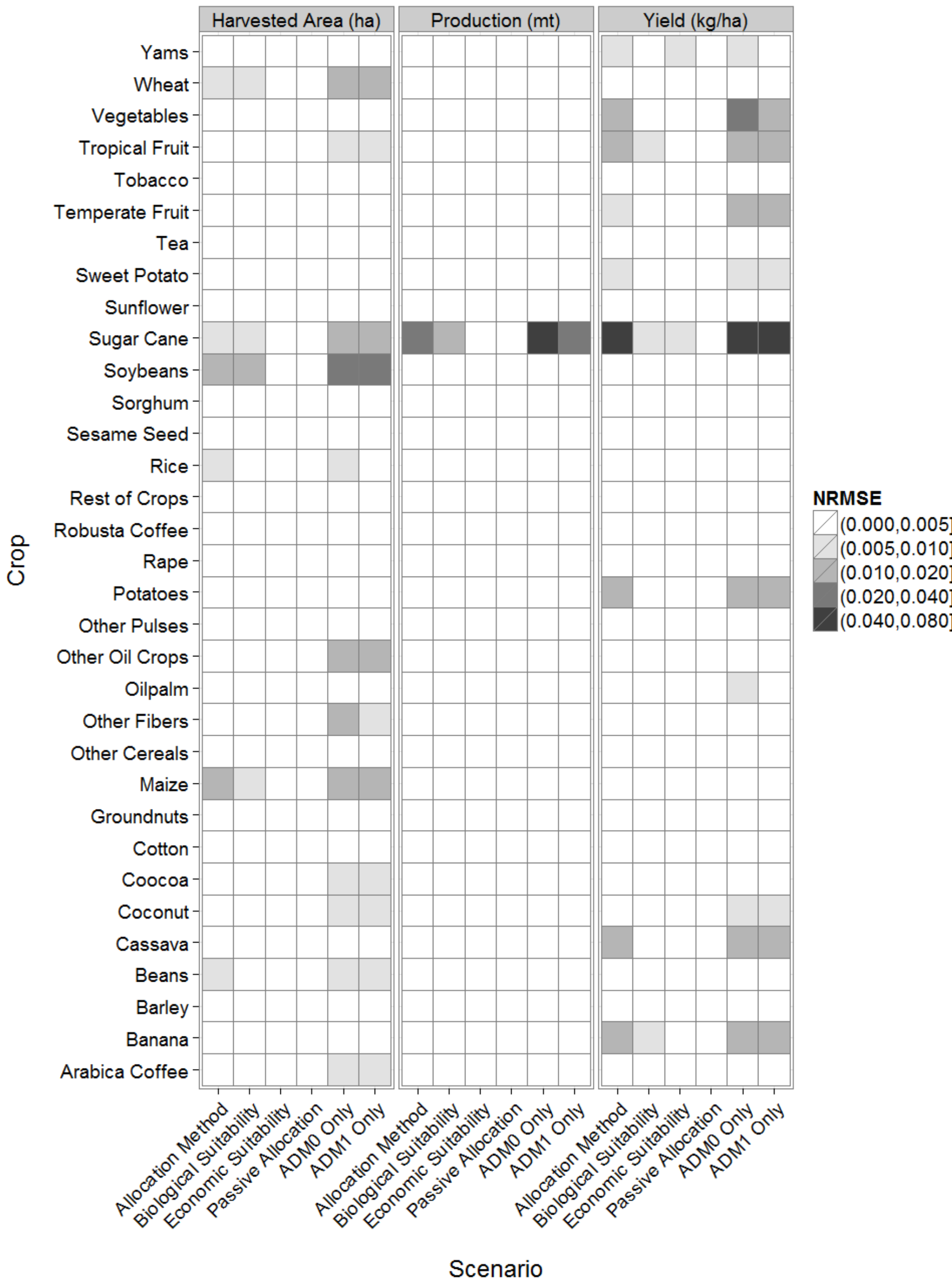
Appendix B: Root Mean Squared Errors by Country and Crops

There are substantial differences in the effects of methodological factors on the estimates from SPAM that vary by crop, as well as country and production statistic. The second cluster of regressions are simple linear regressions of the robustness scenario estimate on the baseline estimate. These are repeated for every combination of crop, country, robustness scenario and production statistic studied. For comparison, each RMSE is normalized using the range of the measured data:

$$NRMSE_{jksr} = \frac{RMSE_{jksr}}{\max_{\mathrm{r}}\left(baseline_{ijkr}\right) - \min_{\mathrm{r}}\left(baseline_{ijkr}\right)}$$

where $NRMSE$ is the normalized root mean squared error statistic for crop j, country k, robustness scenario s and production statistic r, $RMSE$ is the root mean squared error statistic for crop j, country k, robustness scenario s and production statistic r, $baseline$ is the original SPAM estimate for pixel i, crop j, country k and robustness scenario s. These normalized RMSEs are graphed in heatmaps, by country, in Appendix B-1 through B-9 to better visualize the large number of results.

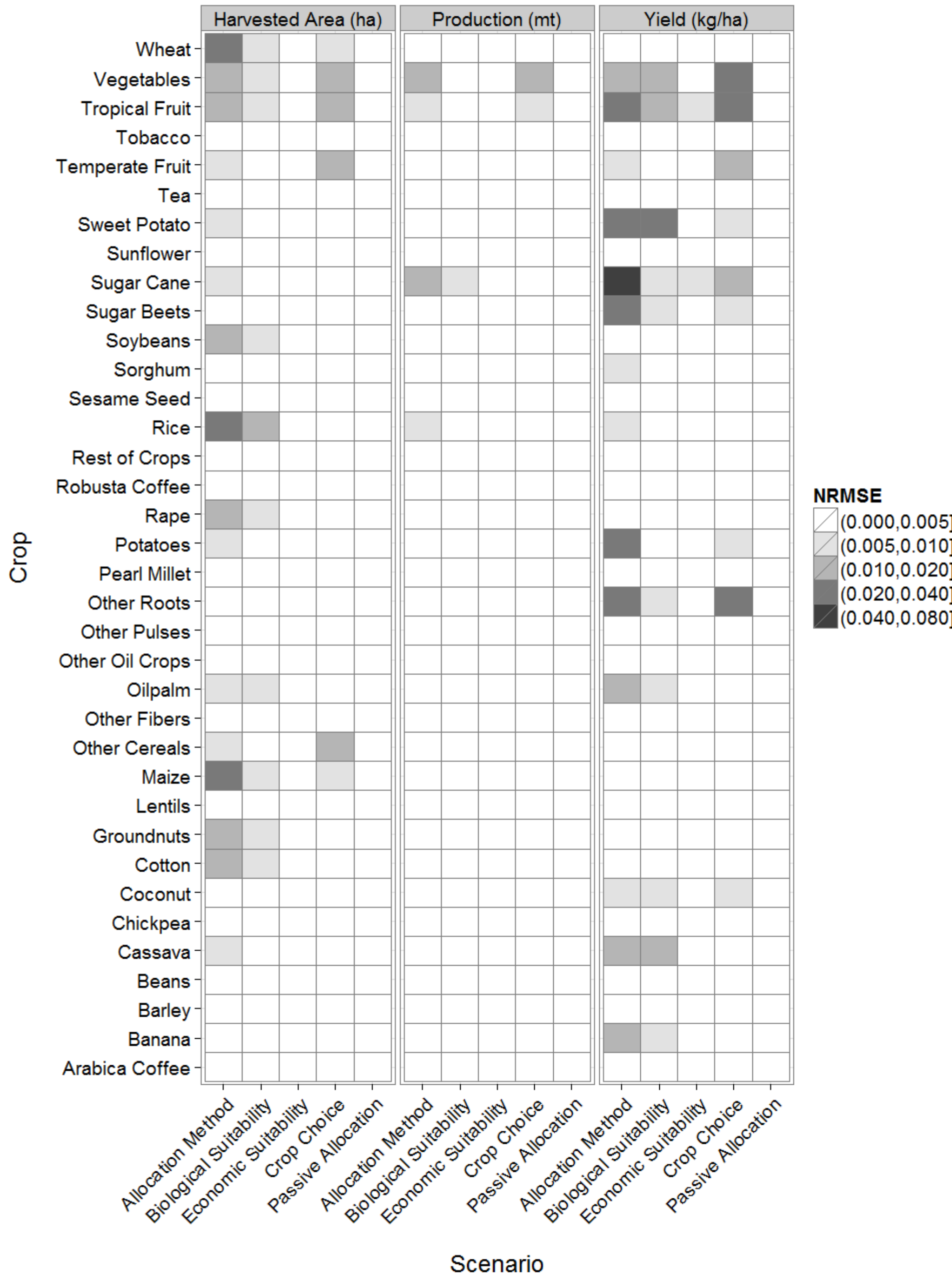
Appendix B-1: NRMSE between Robustness Runs and Baseline Estimates in Brazil



Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Note: To account for differences in scales between production statistics, RMSEs were normalized by dividing the RMSE by the range in baseline estimates.

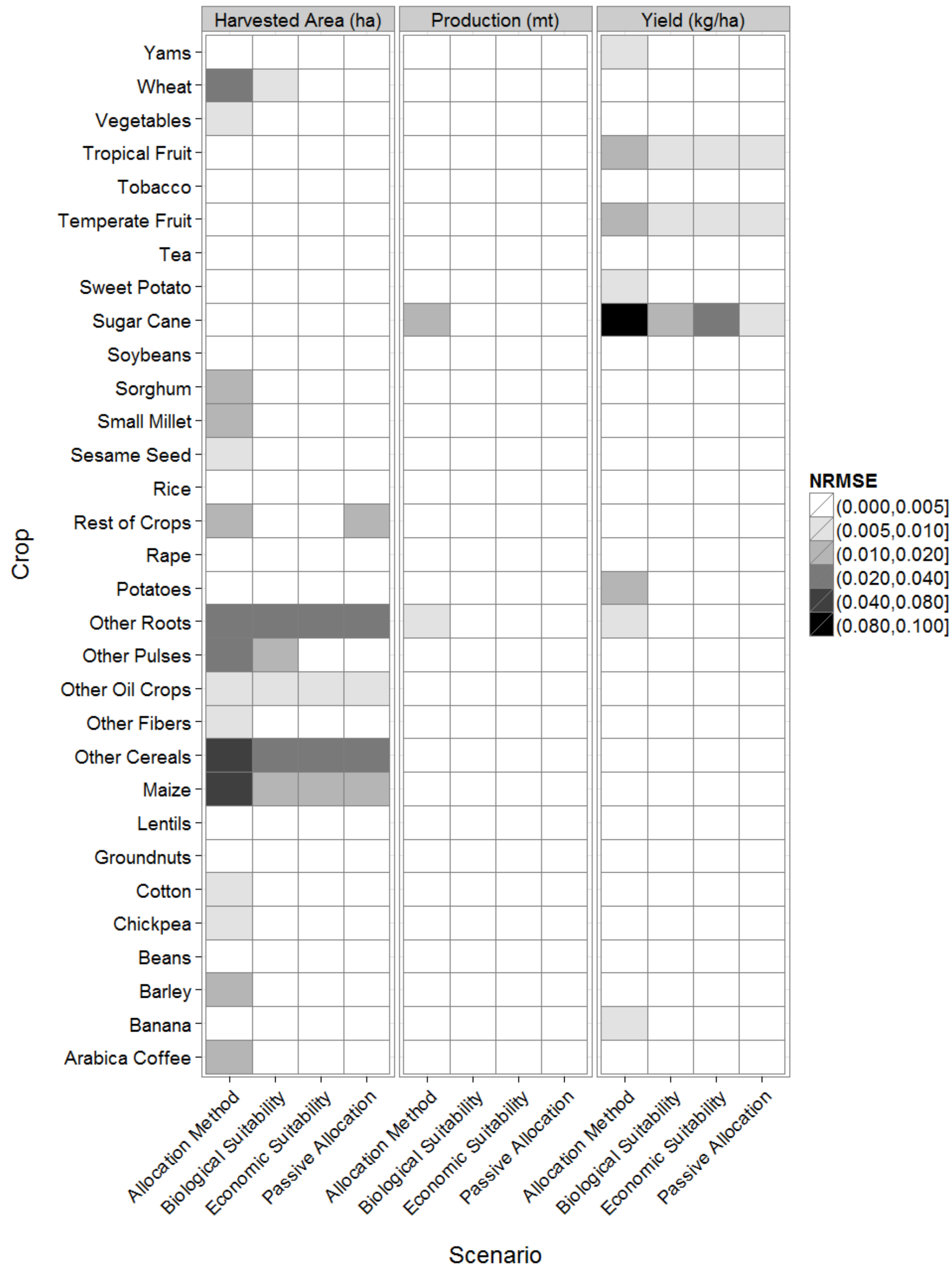
Appendix B-2: NRMSE between Robustness Runs and Baseline Estimates in China



Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Note: To account for differences in scales between production statistics, RMSEs were normalized by dividing the RMSE by the range in baseline estimates.

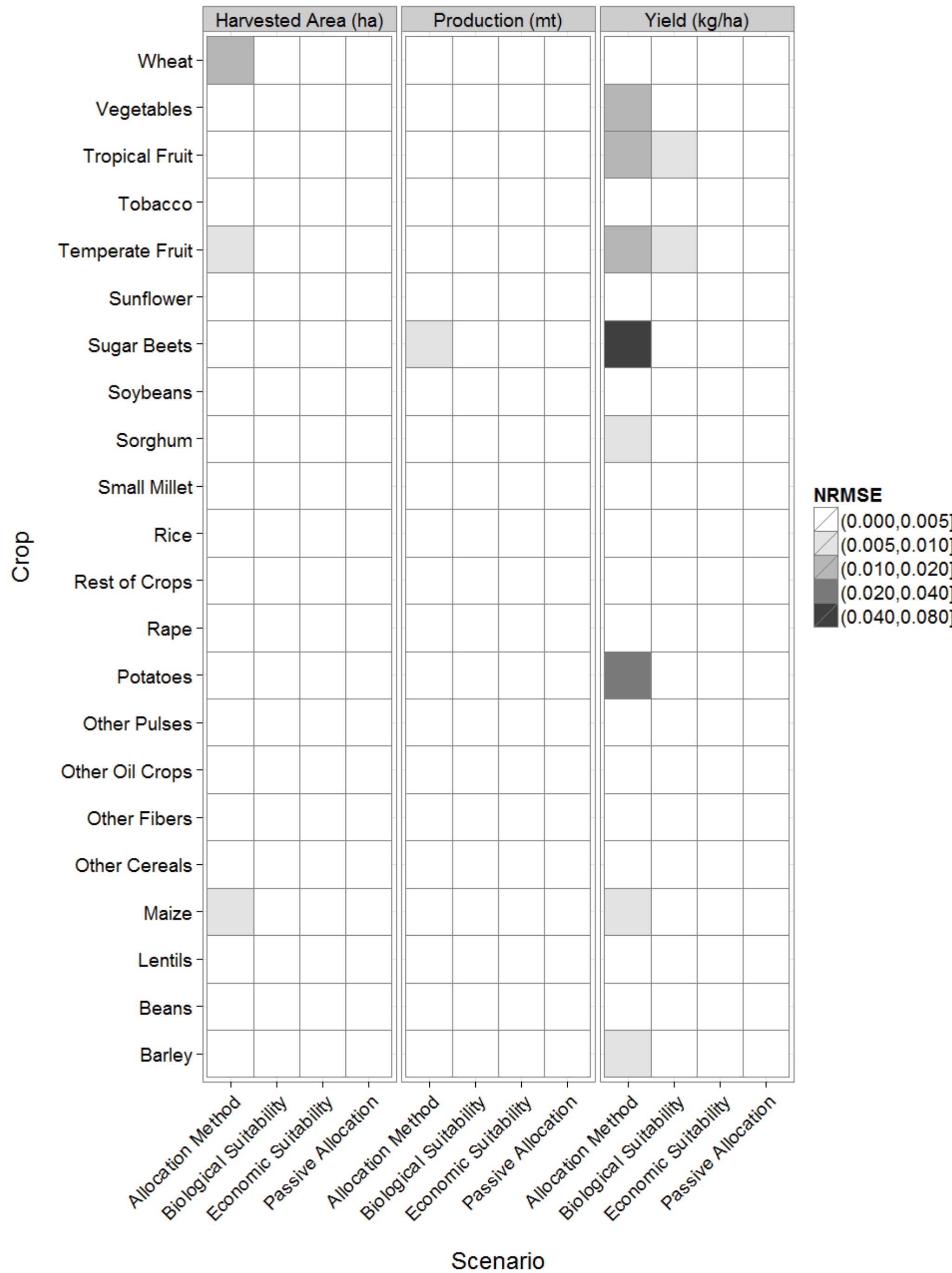
Appendix B-3: NRMSE between Robustness Runs and Baseline Estimates in Ethiopia



Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Note: To account for differences in scales between production statistics, RMSEs were normalized by dividing the RMSE by the range in baseline estimates.

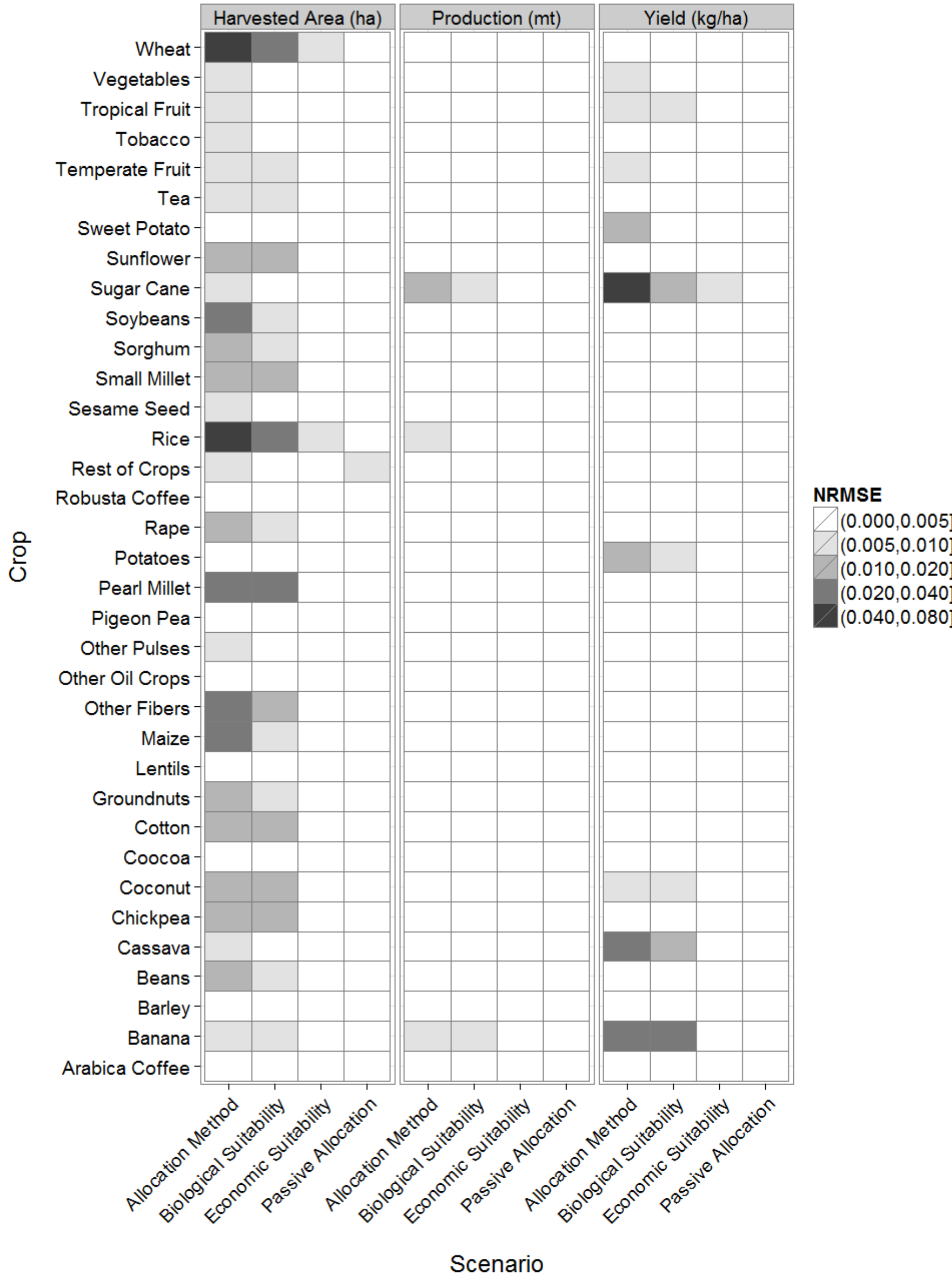
Appendix B-4: NRMSE between Robustness Runs and Baseline Estimates in France



Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Note: To account for differences in scales between production statistics, RMSEs were normalized by dividing the RMSE by the range in baseline estimates.

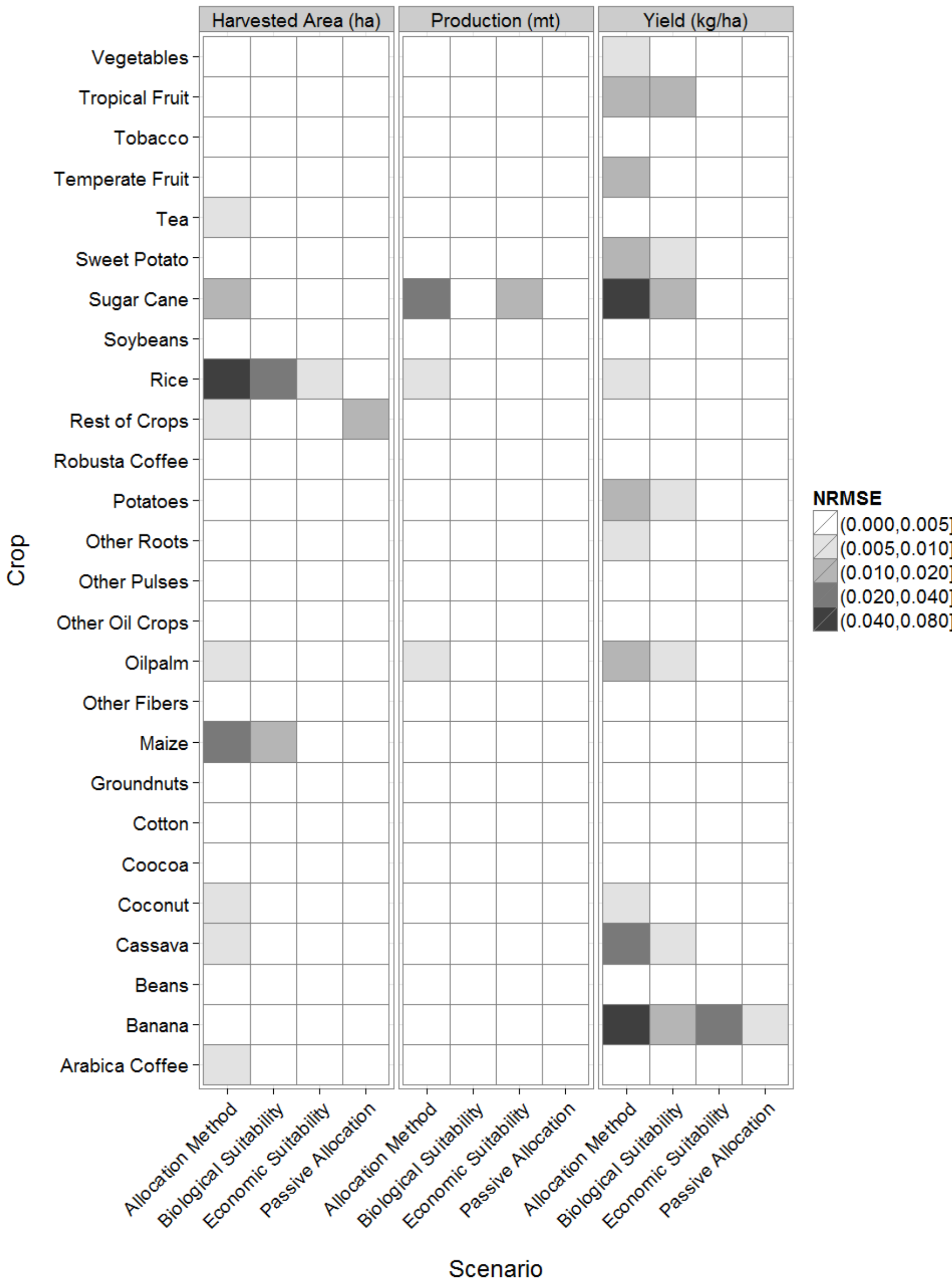
Appendix B-5: NRMSE between Robustness Runs and Baseline Estimates in India



Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Note: To account for differences in scales between production statistics, RMSEs were normalized by dividing the RMSE by the range in baseline estimates.

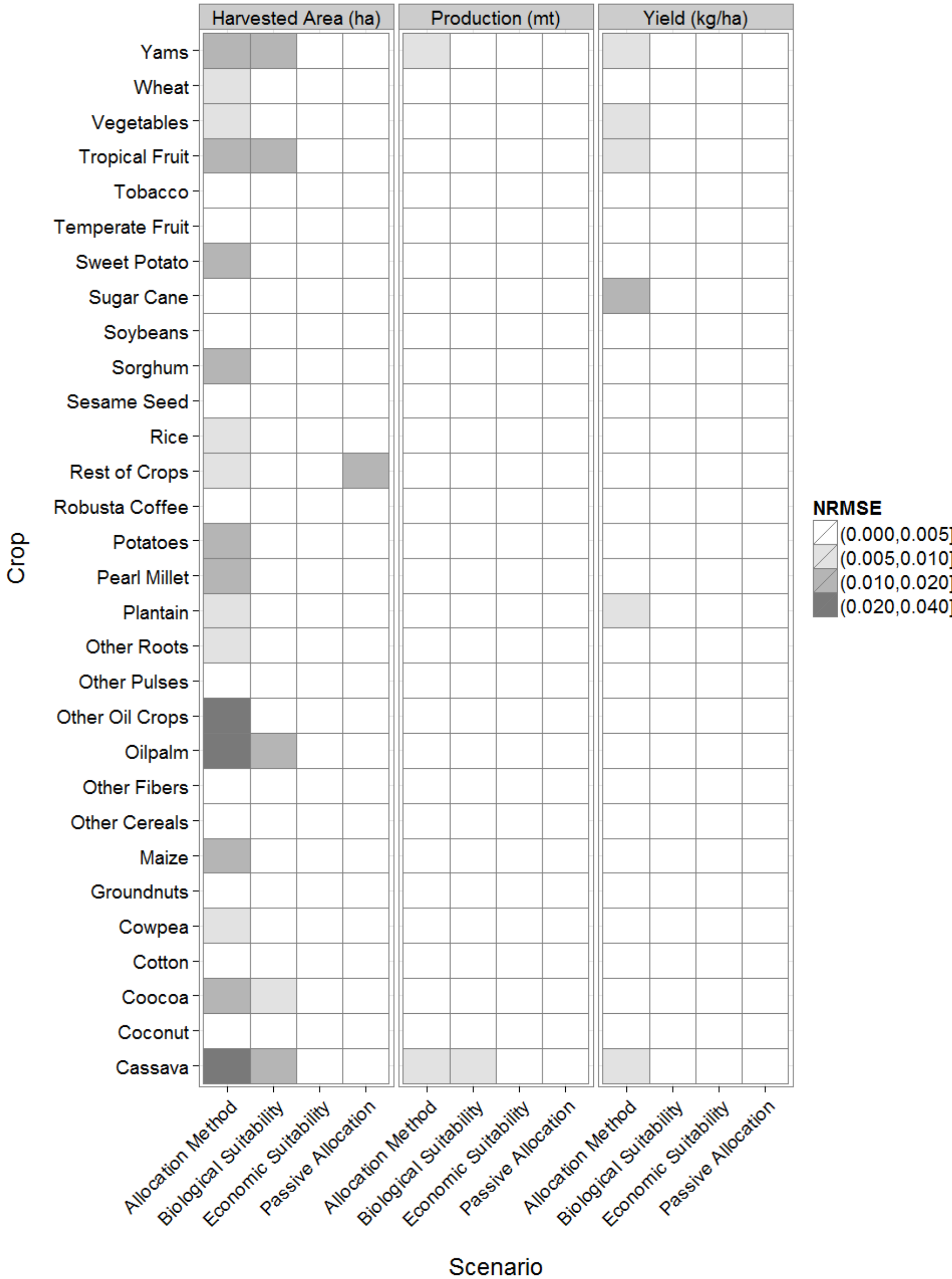
Appendix B-6: NRMSE between Robustness Runs and Baseline Estimates in Indonesia



Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Note: To account for differences in scales between production statistics, RMSEs were normalized by dividing the RMSE by the range in baseline estimates.

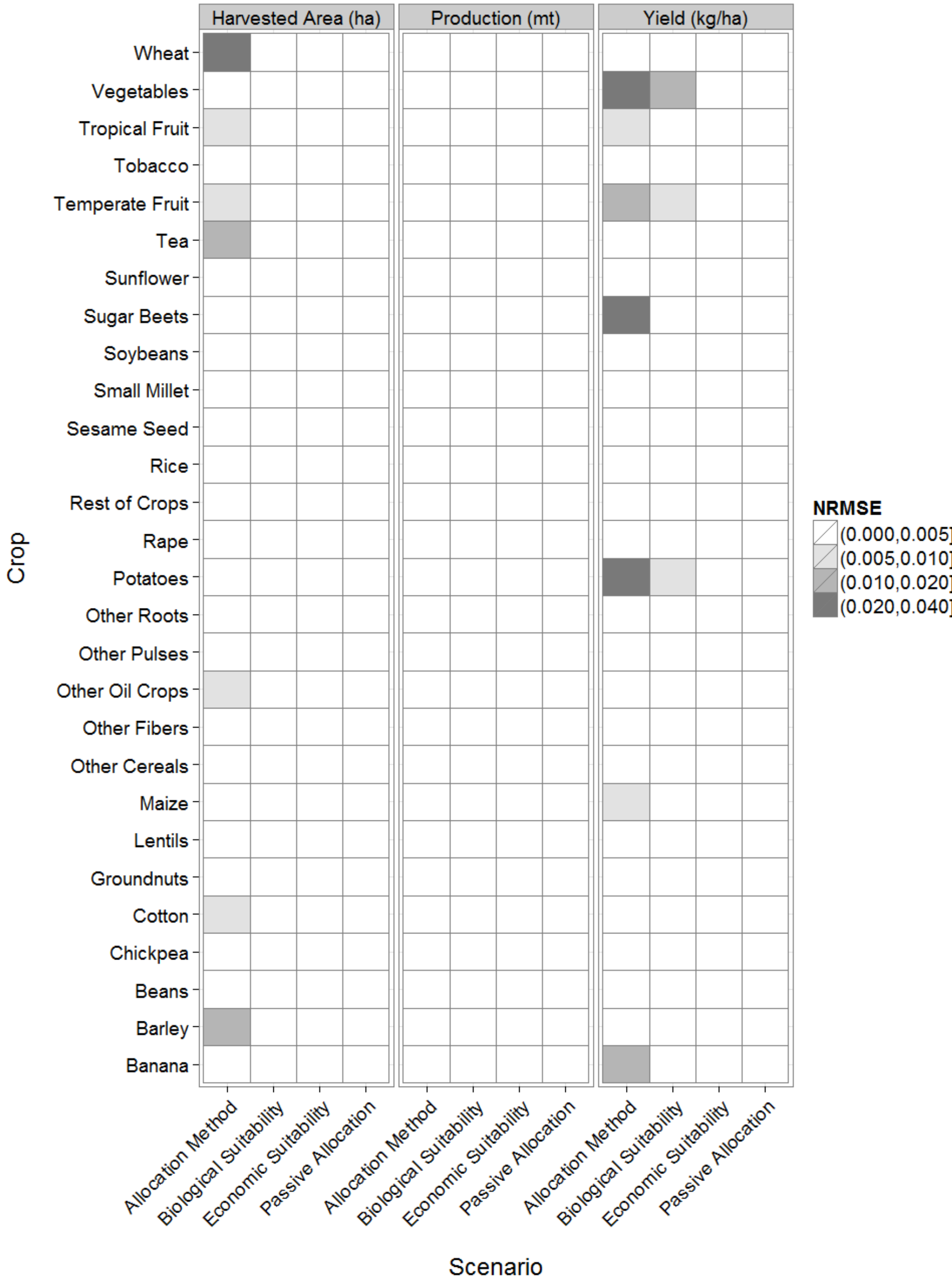
Appendix B-7: NRMSE between Robustness Runs and Baseline Estimates in Nigeria



Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Note: To account for differences in scales between production statistics, RMSEs were normalized by dividing the RMSE by the range in baseline estimates.

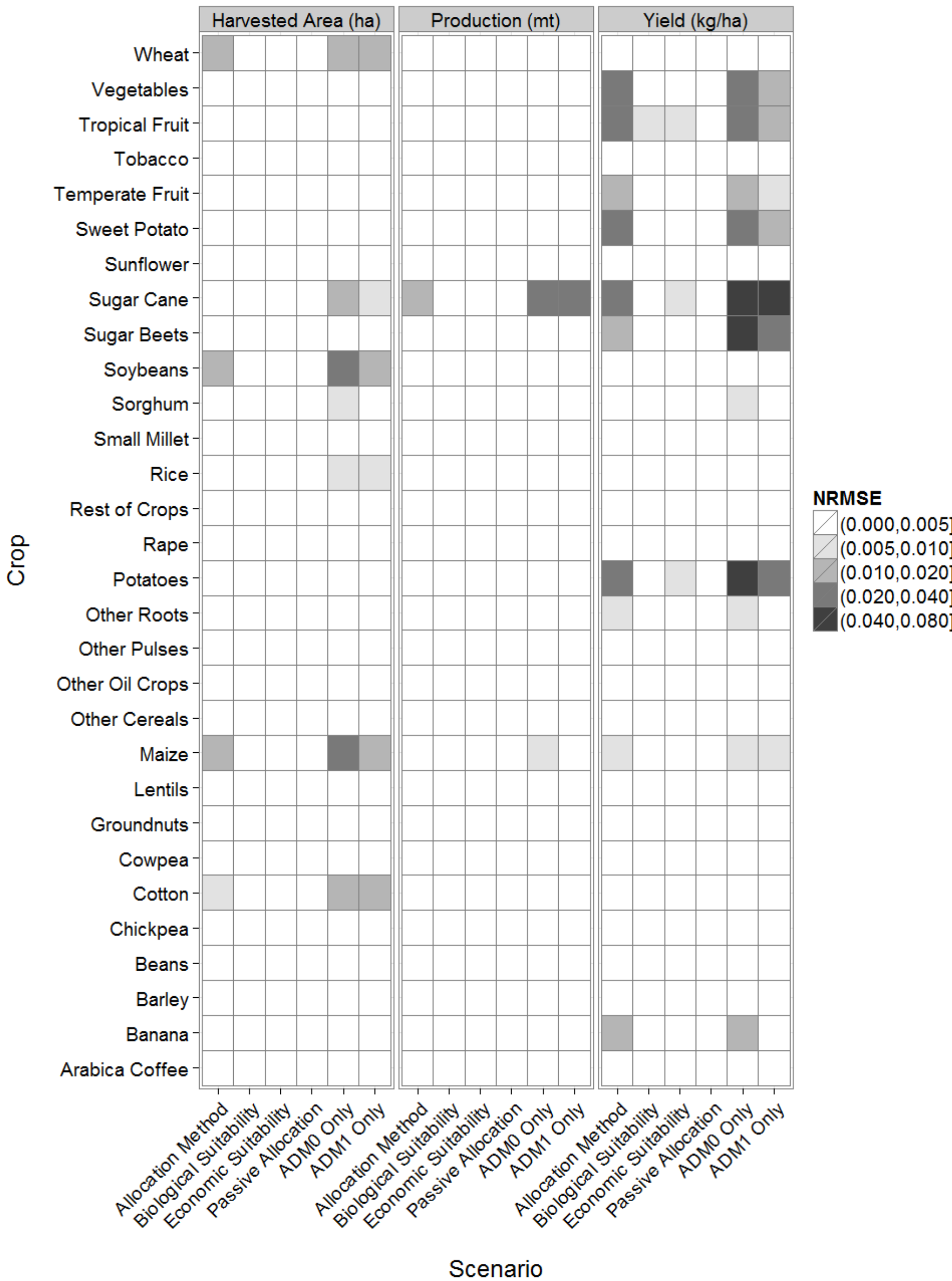
Appendix B-8: NRMSE between Robustness Runs and Baseline Estimates in Turkey



Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Note: To account for differences in scales between production statistics, RMSEs were normalized by dividing the RMSE by the range in baseline estimates.

Appendix B-9: NRMSE between Robustness Runs and Baseline Estimates in the United States

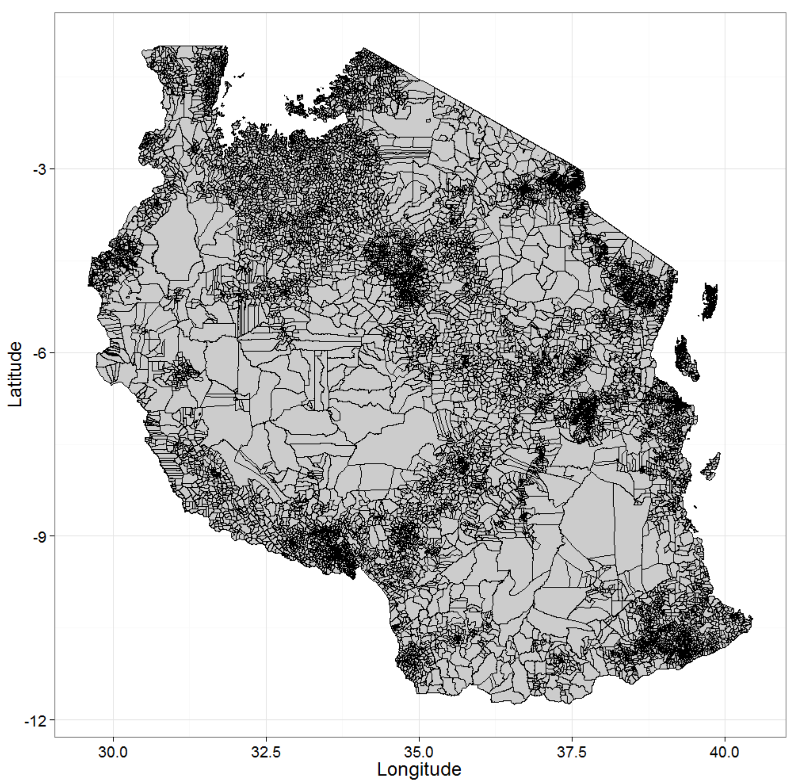


Source: Developed by author using data from SPAM2005 (You, Wood-Sichra, et al. 2015) and own calculations.

Note: To account for differences in scales between production statistics, RMSEs were normalized by dividing the RMSE by the range in baseline estimates.

Appendix C: Methodology to Calculate ASC Sampled Villages' Geo-Coordinates

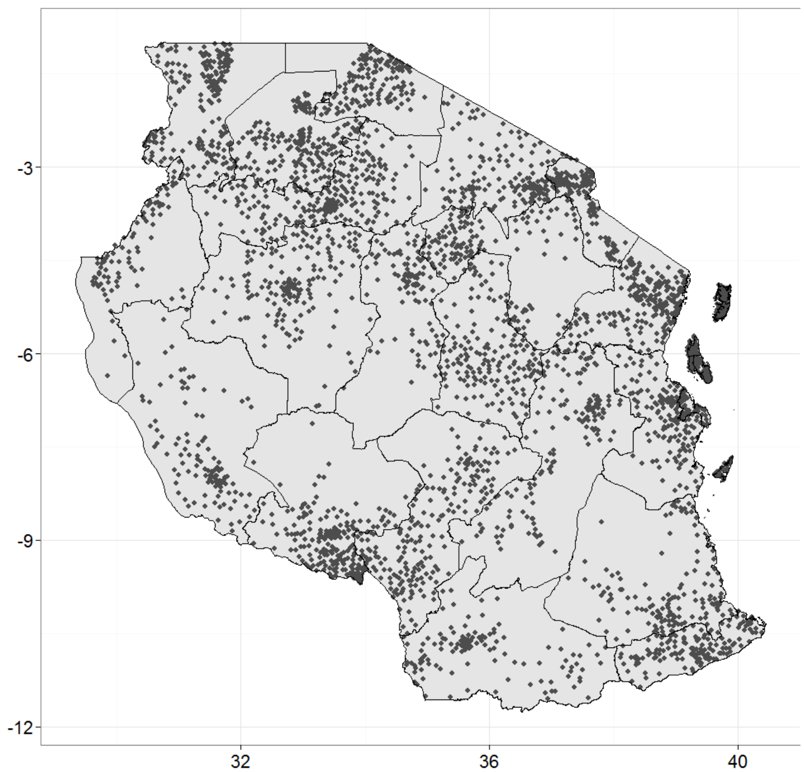
To approximate the location of households surveyed within Tanzania's 2007/08 Agricultural Sample Census, I first matched the names of surveyed villages to a village-level shapefile and then assigned the centroid coordinates of the polygon as the sampled village coordinates using ArcGIS. Each household within a village is assigned the same geo-coordinates. The names of the 3,434 surveyed villages are listed in Appendix 1 of the Technical and Operation Report associated with the ASC (NBS 2011b). The village-EA level shapefile was prepared by the National Bureau of Statistics (NBS) for the 2002 census and does not reflect boundary changes since that time (OpenMicroData 2010) and is mapped in Figure C-1.



Source: Developed using shapefile from OpenMicroData (2010).

Figure C-1: Tanzania village/enumeration area (EA) boundaries in 2002

I was able to merge 97 percent of the 2007 ASC sampled villages with the 2002 village boundaries based on the 2002 region code, district code, ward code and street/village codes. Of the unmatched villages, I was able to manually match all but ten based on region, district, ward and street/village names, and occasionally, reported 2002 population. This meant that I could not assign geo-coordinates to 3 percent of the surveyed households. The villages surveyed with available geo-coordinates are mapped in Figure C-2.



Source: Developed by author using data from the Tanzania ASC (NBS 2011a) and own-calculations.

Figure C-2: Sampled villages in the ASC data

Appendix D: Comparison of Own Aggregates to the National Sample Census of Agricultural Small Holder Crop Sector – National Report

Table D-1: Inorganic fertilizer use in short-rain season (*Vuli*) by region

	Planted Area with Fertilizer (ha)			Total Planted Area in Vuli (ha)			% of Planted Area using Fertilizer		
Region	Own- Aggregate	Reported Aggregate	Ratio	Own- Aggregate	Reported Aggregate	Ratio	Own- Aggregate	Reported Aggregate	Ratio
Arusha	5,156	5,158	1.00	33,982	33,997	1.00	15.2	15.2	1.00
Dar es Salaam	431	431	1.00	6,138	6,273	0.98	7.0	6.9	1.02
Dodoma	167	167	1.00	814	814	1.00	20.6	20.6	1.00
Iringa	288	288	1.00	1,115	1,116	1.00	25.8	25.8	1.00
Kagera	1,180	1,180	1.00	243,732	245,327	0.99	0.5	0.5	0.97
Kigoma	7,031	7,034	1.00	150,251	150,382	1.00	4.7	4.7	1.00
Kilimanjaro	14,701	14,707	1.00	78,748	78,824	1.00	18.7	18.7	1.00
Lindi	444	444	1.00	1,294	1,294	1.00	34.3	34.3	1.00
Manyara	500	501	1.00	10,856	11,175	0.97	4.6	4.5	1.02
Mara	3,391	3,392	1.00	170,010	170,217	1.00	2.0	2.0	1.00
Mbeya	9,702	9,707	1.00	52,550	52,573	1.00	18.5	18.5	1.00
Morogoro	17,475	17,482	1.00	237,052	237,174	1.00	7.4	7.4	1.00
Mtwara	-	-	-	666	666	1.00	-	-	-
Mwanza	4,298	4,300	1.00	511,814	512,358	1.00	0.8	0.8	1.05
Pwani	799	800	1.00	59,353	59,666	0.99	1.3	1.3	1.04
Rukwa	-	-	-	3,049	3,050	1.00	-	-	-
Ruvuma	25	25	0.99	500	501	1.00	4.9	4.9	1.01
Shinyanga	2,051	2,051	1.00	18,856	18,864	1.00	10.9	10.9	1.00
Singida	-	-	-	-	-	-	-	-	-
Tabora	-	-	-	601	601	1.00	-	-	-
Tanga	2,709	2,710	1.00	178,323	178,527	1.00	1.5	1.5	1.01
Mainland	**70,347**	**70,376**	**1.00**	**1,759,706**	**1,763,399**	**1.00**	**4.0**	**4.0**	**1.00**

Source: Developed using data from the ASC (NBS 2011a) and reported statistics from NBS, et al. (2012).

Table D-2: Inorganic Fertilizer Use in Long Rain Season (*Masika*) by Region

	Planted Area with Fertilizer (ha)			Total Planted Area in Vuli (ha)			% of Planted Area using Fertilizer		
Region	Own- Aggregate	Reported Aggregate	Ratio	Own- Aggregate	Reported Aggregate	Ratio	Own- Aggregate	Reported Aggregate	Ratio
Arusha	5,156	5,158	1.00	33,982	33,997	1.00	15.2	15.2	1.00
Dar es Salaam	431	431	1.00	6,138	6,273	0.98	7.0	6.9	1.02
Dodoma	167	167	1.00	814	814	1.00	20.6	20.6	1.00
Iringa	288	288	1.00	1,115	1,116	1.00	25.8	25.8	1.00
Kagera	1,180	1,180	1.00	243,732	245,327	0.99	0.5	0.5	0.97
Kigoma	7,031	7,034	1.00	150,251	150,382	1.00	4.7	4.7	1.00
Kilimanjaro	14,701	14,707	1.00	78,748	78,824	1.00	18.7	18.7	1.00
Lindi	444	444	1.00	1,294	1,294	1.00	34.3	34.3	1.00
Manyara	500	501	1.00	10,856	11,175	0.97	4.6	4.5	1.02
Mara	3,391	3,392	1.00	170,010	170,217	1.00	2.0	2.0	1.00
Mbeya	9,702	9,707	1.00	52,550	52,573	1.00	18.5	18.5	1.00
Morogoro	17,475	17,482	1.00	237,052	237,174	1.00	7.4	7.4	1.00
Mtwara	-	-	-	666	666	1.00	-	-	-
Mwanza	4,298	4,300	1.00	511,814	512,358	1.00	0.8	0.8	1.05
Pwani	799	800	1.00	59,353	59,666	0.99	1.3	1.3	1.04
Rukwa	-	-	-	3,049	3,050	1.00	-	-	-
Ruvuma	25	25	0.99	500	501	1.00	4.9	4.9	1.01
Shinyanga	2,051	2,051	1.00	18,856	18,864	1.00	10.9	10.9	1.00
Singida	-	-	-	-	-	-	-	-	-
Tabora	-	-	-	601	601	1.00	-	-	-
Tanga	2,709	2,710	1.00	178,323	178,527	1.00	1.5	1.5	1.01
Mainland	**70,347**	**70,376**	**1.00**	**1,759,706**	**1,763,399**	**1.00**	**4.0**	**4.0**	**1.00**

Source: Developed using data from the ASC (NBS 2011a) and reported statistics from NBS, et al. (2012).

Table D-3: Inorganic Fertilizer Use in Short Rain Season (*Vuli*) by Crop

	Planted Area with Fertilizer (ha)			Quantity of Fertilizer Purchased (kg)			Cost of Fertilizer Purchased (TZS)		
Crop	Own-Aggregate	Reported Aggregate	Ratio	Own- Aggregate	Reported Aggregate	Ratio	Own- Aggregate	Reported Aggregate	Ratio
Barley	-	-	-	-			-	-	-
Bulrush Millet	-	-	-	-			-	-	-
Finger Millet	116	116	1.00	3,288.24			2,693,717.47	2,693,717	1.00
Maize	34,836	34,858	1.00	64,633,707.24			3,434,601,261.38	3,434,601,261	1.00
Paddy	17,809	17,816	1.00	2,477,676.77			1,339,577,970.14	1,339,577,970	1.00
Sorghum	270	270	1.00	7,268.28			9,546,796.49	9,546,796	1.00
Wheat	433	433	1.00	924.55			2,990,337.74	2,990,338	1.00
Cereals	**53,464**	**53,494**	**1.00**	**67,122,865**			**4,789,410,083**	**4,789,410,083**	**1.00**
Cassava	44	44	1.00	86.80			347,196.14	347,196	1.00
Cocoyam	-	-	-	-			-	-	-
Irish Potatoes	1,369	1,369	1.00	11,364,900.98			184,080,334.26	184,080,334	1.00
Sweet Potatoes	392	392	1.00	34,531.14			31,063,606.12	31,063,606	1.00
Yams	-	-	-	-			-	-	-
Roots & Tubers	**1,805**	**1,806**	**1.00**	**11,399,519**			**215,491,137**	**215,491,137**	**1.00**
Bambara Nuts	-	-	-	-			-	-	-
Beans	3,053	3,055	1.00	1,656,126.44			280,144,681.42	280,144,681	1.00
Chick Peas	6	6	-	1,906,882.04			-	-	-
Cowpeas	151	151	1.00	15,153.36			19,293,534.65	19,293,535	1.00
Field Peas	77	77	0.99	67,377.41			36,370,824.91	36,370,825	1.00
Green Gram	-	-	-	-			-	-	-
Mung Beans	6	6	1.05	7,503.31			15,006.63	15,007	1.00
Pulses	**3,293**	**3,295**	**1.00**	**3,653,043**			**335,824,048**	**335,824,048**	**1.00**
Castor Seed	-	-	-	-			-	-	-
Groundnuts	132	132	1.00	3,825.42			3,727,790.19	3,727,790	1.00
Simsim	33	33	1.00	542.39			379,671.15	379,671	1.00
Soya Beans	28	28	1.01	3,499.99			3,499,987.50	3,499,987	1.00
Sunflower	135	135	1.00	3,860.09			12,040,795.63	12,040,796	1.00
Oil Seeds & Oil Nuts	**328**	**329**	**1.00**	**11,728**			**19,648,244**	**19,648,244**	**1.00**
Amaranths	213	213	1.00	335,605.78			29,484,591.22	29,484,591	1.00
Bitter Aubergine	338	338	1.00	43,914.32			48,765,903.15	48,765,903	1.00
Cabbage	696	696	1.00	1,676,510.69			105,094,128.93	105,094,129	1.00
Carrot	99	99	1.00	241,076.09			16,635,092.13	16,635,092	1.00

	Planted Area with Fertilizer (ha)			Quantity of Fertilizer Purchased (kg)			Cost of Fertilizer Purchased (TZS)		
Crop	Own-Aggregate	Reported Aggregate	Ratio	Own- Aggregate	Reported Aggregate	Ratio	Own- Aggregate	Reported Aggregate	Ratio
Chilies	536	536	1.00	420,720.96			118,028,966.55	118,028,967	1.00
Cucumber	231	231	1.00	63,826.98			45,055,915.79	45,055,916	1.00
Egg Plant	150	150	1.00	105,232.64			13,791,312.98	13,791,313	1.00
Okra	411	411	1.00	1,899,484.92			48,595,445.54	48,595,446	1.00
Onions	1,226	1,227	1.00	879,812.66			200,316,524.91	200,316,525	1.00
Pumpkins	26	26	1.00	414,662.74			4,364,465.71	4,364,466	1.00
Radish	94	94	1.00	11,943.88			4,777,551.29	4,777,551	1.00
Spinach	231	231	1.00	2,844,526.14			18,162,635.58	18,162,636	1.00
Tomatoes	3,607	3,609	1.00	6,018,518.46			544,983,583.79	544,983,584	1.00
Turmeric	-	-	-	-			-	-	-
Water Mellon	346	346	1.00	87,701.55			39,420,876.94	39,420,877	1.00
Fruits & Vegetables	**8,204**	**8,207**	**1.00**	**15,043,538**			**1,237,476,994**	**1,237,476,994**	**1.00**
Cotton	1,622	1,622	1.00	1,323,687.86			76,299,496.65	76,299,497	1.00
Jute	-	-	-	-			-	-	-
Pyrethrum	-	-	-				-	-	-
Seaweed	-	-	-				-	-	-
Tobacco	1,631	1,632	1.00	779,404.57			637,298,034.78	637,298,035	1.00
Cash Crops	**3,253**	**3,254**	**1.00**	**2,103,092**			**713,597,531**	**713,597,531**	**1.00**
Total	**70,347**	**70,384**	**1.00**	**99,333,785**			**7,311,448,038**	**7,311,448,038**	**1.00**

Source: Developed using data from the ASC (NBS 2011a) and reported statistics from NBS, et al. (2012).

Table D-4: Inorganic Fertilizer Use in Long Rain Season (*Masika*) by Crop

	Planted Area with Fertilizer (ha)			Quantity of Fertilizer Purchased (kg)			Cost of Fertilizer Purchased (TZS)		
Crop	Own-Aggregate	Reported Aggregate	Ratio	Own- Aggregate	Reported Aggregate	Ratio	Own- Aggregate	Reported Aggregate	Ratio
Barley	-	-	-	-	-	-	-	-	-
Bulrush Millet	89.99	90.00	1.00	13,897.86	13,898.00	1.00	7,182,329.40	7,182,329.00	1.00
Finger Millet	1,245.63	1,246.00	1.00	74,911.02	74,911.00	1.00	73,075,999.82	73,076,000.00	1.00
Maize	349,847.46	350,133.00	1.00	354,027,898.32	354,031,970.00	1.00	41,034,004,069.15	41,035,714,280.00	1.00
Paddy	46,393.26	46,373.00	1.00	10,554,868.12	10,554,868.00	1.00	3,638,796,664.59	3,638,796,665.00	1.00
Sorghum	379.36	380.00	1.00	87,586.63	87,587.00	1.00	34,235,620.23	34,235,620.00	1.00
Wheat	2,062.09	2,063.00	1.00	285,938.18	285,938.00	1.00	168,319,187.71	168,319,188.00	1.00
Cereals	**400,017.79**	**400,285.00**	**1.00**	**365,045,100.12**	**365,049,172.00**	**1.00**	**44,955,613,870.89**	**44,957,324,082.00**	**1.00**
Cassava	32.96	33.00	1.00	32,575.45	32,575.00	1.00	17,590,742.09	17,590,742.00	1.00
Cocoyam	96.57	97.00	1.00	11,213.74	11,214.00	1.00	21,238,648.55	21,238,649.00	1.00
Irish Potatoes	14,112.11	14,118.00	1.00	12,235,454.77	12,235,455.00	1.00	3,261,089,491.20	3,261,089,491.00	1.00
Sweet Potatoes	1,208.83	1,209.00	1.00	138,787.51	138,788.00	1.00	122,729,043.91	122,729,044.00	1.00
Yams	-	-	-	-	-	-	-	-	-
Roots & Tubers	**15,450.46**	**15,457.00**	**1.00**	**12,418,031.47**	**12,418,031.00**	**1.00**	**3,422,647,925.75**	**3,422,647,926.00**	**1.00**
Bambara Nuts	2.35	2.00	1.18	1,226.17	1,226.00	1.00	2,056,346.54	2,056,347.00	1.00
Beans	14,969.18	14,976.00	1.00	20,777,113.99	20,777,114.00	1.00	1,425,549,861.77	1,425,549,862.00	1.00
Chick Peas	-	-	-	55,555.93	-	-	-	-	-
Cowpeas	477.23	477.00	1.00	-	55,556.00	-	55,977,999.15	55,977,999.00	1.00
Field Peas	1,778.15	1,779.00	1.00	218,064.57	218,065.00	1.00	239,495,127.53	239,495,128.00	1.00
Green Gram	3.77	4.00	0.94	558.83	559.00	1.00	670,599.08	670,599.00	1.00
Mung Beans	134.29	134.00	1.00	38,101.38	38,101.00	1.00	20,270,282.65	20,270,283.00	1.00
Pulses	**17,364.97**	**17,372.00**	**1.00**	**21,090,620.87**	**21,090,621.00**	**1.00**	**1,744,020,216.73**	**1,744,020,217.00**	**1.00**
Castor Seed	-	-	-	-	-	-	-	-	-
Groundnuts	1,873.77	1,875.00	1.00	753,761.23	753,761.00	1.00	156,259,729.72	156,259,730.00	1.00
Simsim	60.76	61.00	1.00	3,127.47	3,127.00	1.00	1,827,147.80	1,827,148.00	1.00
Soya Beans	47.46	47.00	1.01	7,020.93	7,021.00	1.00	5,433,173.41	5,433,173.00	1.00
Sunflower	3,030.37	3,032.00	1.00	1,521,148.06	1,521,148.00	1.00	242,640,835.66	242,640,836.00	1.00
Oil Seeds & Oil Nuts	**5,012.35**	**5,014.00**	**1.00**	**2,285,057.70**	**2,285,058.00**	**1.00**	**406,160,886.59**	**406,160,887.00**	**1.00**

Crop	Planted Area with Fertilizer (ha) Own- Aggregate	Reported Aggregate	Ratio	Quantity of Fertilizer Purchased (kg) Own- Aggregate	Reported Aggregate	Ratio	Cost of Fertilizer Purchased (TZS) Own- Aggregate	Reported Aggregate	Ratio
Amaranths	729.31	730.00	1.00	285,364.78	285,365.00	1.00	63,032,350.70	63,032,351.00	1.00
Bitter Aubergine	504.26	504.00	1.00	103,257.04	103,257.00	1.00	65,551,051.25	65,551,051.00	1.00
Cabbage	1,254.13	1,255.00	1.00	320,371.36	320,371.00	1.00	259,411,435.36	259,411,435.00	1.00
Carrot	129.28	129.00	1.00	37,262.77	37,263.00	1.00	27,891,821.36	27,891,821.00	1.00
Chilies	722.49	723.00	1.00	252,710.95	252,711.00	1.00	134,639,593.57	134,639,594.00	1.00
Cucumber	344.06	344.00	1.00	69,124.59	69,125.00	1.00	32,078,431.73	32,078,432.00	1.00
Egg Plant	63.32	63.00	1.01	104,489.03	104,489.00	1.00	3,213,949.90	3,213,950.00	1.00
Okra	334.92	335.00	1.00	47,129.15	47,129.00	1.00	29,336,344.74	29,336,345.00	1.00
Onions	2,058.56	2,059.00	1.00	583,618.60	583,619.00	1.00	481,428,495.89	481,428,496.00	1.00
Pumpkins	137.59	138.00	1.00	1,362,560.21	1,362,560.00	1.00	7,497,519.92	7,497,520.00	1.00
Radish	12.08	12.00	1.01	597.19	597.00	1.00	716,632.69	716,633.00	1.00
Spinach	687.78	691.00	1.00	359,847.31	360,738.00	1.00	40,367,957.58	41,259,139.00	0.98
Tomatoes	8,309.59	8,307.00	1.00	26,527,048.78	26,527,049.00	1.00	1,508,625,346.80	1,508,625,347.00	1.00
Turmeric	8.06	8.00	1.01	995.41	995.00	1.00	1,791,736.34	1,791,736.00	1.00
Water Mellon	260.67	261.00	1.00	66,261.74	66,262.00	1.00	23,828,532.61	23,828,533.00	1.00
Fruits & Vegetables	**15,556.11**	**15,559.00**	**1.00**	**30,120,638.90**	**30,121,530.00**	**1.00**	**2,679,411,200.44**	**2,680,302,381.00**	**1.00**
Cotton	309.30	309.00	1.00	940,961.79	940,962.00	1.00	87,081,456.66	87,081,457.00	1.00
Jute	-	-	-	-	-	-	-	-	-
Pyrethrum	53.36	53.00	1.01	263.71	264.00	1.00	115,372.66	115,373.00	1.00
Seaweed	-	-	-	-	-	-	-	-	-
Tobacco	50,366.08	50,388.00	1.00	65,452,261.82	65,452,262.00	1.00	19,508,116,960.42	19,508,116,960.00	1.00
Cash Crops	**50,728.74**	**50,750.00**	**1.00**	**66,393,487.32**	**66,393,487.00**	**1.00**	**19,595,313,789.74**	**19,595,313,790.00**	**1.00**
Total	**504,130.42**	**504,438.00**	**1.00**	**497,352,936.38**	**497,357,899.00**	**1.00**	**72,803,167,890.14**	**72,805,769,282.00**	**1.00**

Source: Developed using data from the ASC (NBS 2011a) and reported statistics from NBS, et al. (2012).