THE FORMATION AND IMPACT OF BELIEFS ABOUT EXPECTED RETURNS TO AGRICULTURAL INPUT INVESTMENTS: AN ANALYSIS OF FERTILIZER USE AMONG TANZANIAN FARMERS

by

BRICE KARICKHOFF

(Under the Direction of Ellen McCullough)

The use of fertilizer among small-holder farmers is critical to agricultural and economic development in Sub-Saharan Africa. Farmers must decide if they will invest in fertilizer prior to receiving information about the ensuing returns to that particular investment. We, therefore, seek to understand how farmers form and use expectations about the returns to fertilizer use. We construct two distributions of fertilizer return outcomes: a simulated, “site-modeled” distribution, corresponding to typical conditions, and a “recent-year” distribution, corresponding to actual experienced returns in the past three years. We analyze how these two distributions each predict fertilizer use according to Tanzanian household survey data. We find that, with regard to both distributions, high levels of returns (associated with lower yields) and high variability in returns are both predictive of lower levels of fertilizer use.

INDEX WORDS: Fertilizer Adoption, Value-to-Cost Ratio, Agricultural Inputs

THE FORMATION AND IMPACT OF BELIEFS ABOUT EXPECTED RETURNS TO AGRICULTURAL INPUT INVESTMENTS: AN ANALYSIS OF FERTILIZER USE AMONG TANZANIAN FARMERS

By

BRICE KARICKHOFF

B.A., University of Georgia, 2018

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTERS OF SCIENCE

ATHENS, GEORGIA

2020

THE FORMATION AND IMPACT OF BELIEFS ABOUT EXPECTED RETURNS TO AGRICULTURAL INPUT INVESTMENTS: AN ANALYSIS OF FERTILIZER USE AMONG TANZANIAN FARMERS

by

BRICE KARICKHOFF

Major Professor: Ellen McCollough

Committee: Nicholas Magnan
Mateusz Filipksi

Electronic Version Approved:

Ron Walcott
Interim Dean of the Graduate School
The University of Georgia
August 2020

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION:

The development and introduction of new technologies has been pivotal to much of the improvements in yields, profits, and living standards among low-income rural farmers around the world. Among these technologies, fertilizer is one of the most impactful. Varying levels of fertilizer use across countries have been shown to have a great effect on yields (Evanson and Gollin 2003), and within many countries, experimental plots consistently testify to the positive and cost effective impact of fertilizer. The efficient use of fertilizer is surely pivotal to the agricultural and economic development of Sub-Saharan Africa.

Despite the general effectiveness of fertilizer, adoption rates remain low in many regions. For instance, Tanzanian and Kenyan farmers apply an average of 10 kg of fertilizer per hectare, compared to Brazilian and Indian farmers who apply 175 and 165 kilograms respectively (World Bank 2014). Theodore Schultz (1964) established that low-income farmers are indeed rational profit maximizers, yet fertilizer adoption rates in Sub-Saharan Africa remain below profit maximizing levels (Sheahan and Barrett 2017). Despite the evidence supporting the benefits of fertilizer use, real-world results are not always as compelling or straightforward (Yenggen et al. 1998). Consequently, we need more real world research, beyond agronomic trials, to better understand fertilizer adoption.

A large literature, which will be discussed in the next section, has emerged in an effort to measure the real-world effectiveness of fertilizer, and understand farmers' thinking about its use. While many of the findings have been interesting and helpful, if policy makers and economists hope to increase fertilizer use among Sub-Saharan African farmers, they must continue to pursue a deeper understanding of what motivates farmers to buy and use fertilizer.

This paper aims to contribute to the growing understanding of how farmers make decisions about fertilizer use. Generally speaking, this paper explores how farmers form expectations about the profitability of adopting fertilizer. Towards that end, we develop and then explore the explanatory power of several different frameworks describing how expectations about profitability might feature in a farmer's fertilizer adoption decisions. These frameworks are meant to compare the degree to which a farmer's fertilizer decisions are effected by the profitability of fertilizer use across two different distributions: a distribution containing fertilizer returns in the past three years, and a distribution representing expectations in a typical year.

We use data from the Living Standards Measurement Study (LSMS) in Tanzania to gather information on the fertilizer use decisions of individual maize-farming households. We then generate two distributions of data for each household using two measures of fertilizer returns: the Value-to-Cost Ratio (VCR - a measure of the profitability of fertilizer adoption) and yield difference (the difference between expected yields with and without fertilizer application). The two distributions are:

1. Recent Years - a set of the fertilizer returns during the three most recent years prior to the farmer's fertilizer decision

2. Site Modeled - a simulated set of 100 years of fertilizer return outcomes based on a locationally representative synthetic set of weather conditions developed by McCullough et al (2018)

The principle aim of this paper is to understand how these two distributions of fertilizer returns explain a household's fertilizer related decisions.

Several interesting patterns emerge in our results. We find that inorganic fertilizer use decreases with an increasing mean and standard deviation in inorganic fertilizer returns. This suggests two mechanisms. First, farmers view inorganic fertilizer adoption as a risky decision, and are less likely to adopt when faced with higher variability. Second, farmers are unlikely to adopt inorganic fertilizer when returns are historically high. We believe this may be explained by the low yields associated with non-adopters during years when returns are high. We find that, unlike inorganic fertilizer, organic fertilizer increases with higher fertilizer returns. Organic fertilizer use also increases with standard deviation, but only using the site modeled distribution. With few exceptions, the recent-year and site-modeled distributions yield similar results.

The remainder of this paper proceeds as follows. In the next section, Section 2, we overview the current literature on fertilizer adoption, and we discuss the McCullough et al (2018) paper that this paper largely relates to. In Section 3 we discuss our data sources. In Section 4 we discuss how we calculate and model returns to fertilizer use within the recent-year and site-modeled frameworks. Section 5 discusses our empirical strategy. Section 6 focuses on our results. Section 7 discusses our findings and the mechanisms that might explain our results. Section 8 concludes the paper.

2 LITERATURE REVIEW:

Adoption of fertilizer has been critical to the increase in yields experienced on a global level (Hopper 1993). However, Sub-Saharan African has lagged behind in terms of fertilizer use, leading economists across decades to call for an increase in fertilizer use in the region (Mellor, Delgado, and Blackie 1987; Quiñones, Borlaug, and Dowswell 1997; Otsuka and Kalirajan 2006; Holden 2018). It is well established that fertilizer adoption is a necessity in Sub-Saharan Africa at a macro-level, thus it is important that African governments, NGOs, and the international community find a way to address this incongruence between fertilizer profitability and adoption (African Union 2003). Highly effective solutions remain illusive.

Examples of policies meant to increase fertilizer use include fertilizer subsidies and productive safety nets (Holden 2018), but these policies have seen mixed results. Several African governments have opted to subsidize fertilizers, but these policies remain a contentious issue, and each government has handled them differently. Critics argue that subsidies should be scaled back due to fiscal constraints, corruption, poor administration, and the fact that they have not consistently proven to be an effective means of increasing adoption (Rashid et al. 2013). However, advocates accurately point out that adoption of new technologies, such as fertilizer, is an indispensable component of economic development on a large scale, and subsidies are among the best instruments to incentivize adoption (Dugger 2007).

African governments should certainly seek to improve technology use among rural farmers, but if programs aimed at improving adoption, such as fertilizer subsidies, are to be an effective means to that end, it is critical that they are targeted appropraitely among farmers whom they may actually incentivize to adopt. Those seeking to promote fertilizer face the reality that the determinants of fertilizer adoption are complex and not well understood. Economists must work together with policy makers to understand how a farmer decides whether or not they will purchase fertilizer.

Consequently, a large body of literature has emerged which addresses how farmers make decisions about adopting fertilizer. A lack of information is one factor in low adoption rates. Bold et al (2017) finds that farmers cannot be sure that the fertilizer they purchase has not been diluted, and Duflo (2007) finds that farmers are less likely to adopt if they are unsure of the exact quantity of fertilizer to apply. Two other constraints on the decision making of rural farmers are a lack of insurance and a lack of credit. These factors exacerbate the degree to which risk aversion and impatience might inhibit farmers from purchasing inputs such as fertilizer. Duflo, Kremer, and Robinson (2011) find that farmers procrastinate on purchasing fertilizer after a harvest, thus they struggle to purchase the input when they are lower on cash immediately preceding the next harvest.

A particularly relevant finding to this paper is the degree to which profitability varies from farm to farm. Research shows that the positive effects of adopting fertilizer are not as straightforward as field trials suggest, and this heterogeneity matters when farmer's make future input use decisions (Duflo, Kremer, and Robinson 2008; Jayne and Rashid 2013). McCullough et al (2018) explores this variability in profitability throughout Sub-Saharan Africa. The authors of this paper use an experimental crop trial metadataset, combined with geocoded weather and soil data, to predict the VCR as a fertilizer-response profitability measure across a broad geographical area. Using synthetic climate data simulated from historic climate data, they generate a site-specific distribution of VCRs for each location in their dataset based on predicted fertilizer-induced yield increases. I will build upon the McCullough et al. paper by applying their findings to maize farmers in Tanzania. Using this profitability analysis, I explore whether recent profitability realizations or a full site-level profitability distribution best explains fertilizer adoption.

Any time a farmer makes a decision to invest in an input, they do not know how it will effect their output and if it will be profitable (Chavas and Holt 1996). Therefore, farmers must make a decision based upon an internalized sense of the probability that the decision will be profitable. The literature reveals how complex this decision is – most papers discussed above offer some additional nuance to how a farmer might expect their fertilizer decision to pay off. My paper contributes to this literature by further enriching our understanding of input-use decision making among farmers in Sub-Saharan Africa.

3 DATA:

The key to our analysis rests in the generation of a cross-sectional dataset wherein each observation describes one household, including that household's decisions relating to fertilizer use. Additionally, each observation must contain information relating to how that household might develop expectations of its returns to fertilizer use. The primary measure of fertilizer returns we employ is the Value-to-Cost Ratio (VCR), which will be discussed at greater length in a later section. Each observation, in addition to containing fertilizer decisions and control variables, includes the VCRs of the three years preceding the fertilizer decisions and a set of 100 simulated VCRs that are meant to approximate the real probability distribution of fertilizer returns for that specific location.

In the next section, we will discuss how we obtained VCRs from our raw data in more detail. In this section we will focus more on the sources of our raw data. It is only important to note that climate and soil data is ultimately used to generate an expected yield with and without fertilizer use for each household. The difference between these yields is easily converted to a VCR value.

3.1 Fertilizer Use Household Data

We obtained household data using waves 1-3 of the Tanzanian Living Standards Measurement Survey (LSMS). Each wave of this dataset contained our outcome variables (fertilizer use decisions), and a set of household controls. The three waves were combined into one pooled cross-sectional dataset because we are ultimately concerned with each individual set of fertilizer use decisions. One household may appear in multiple waves of LSMS data, providing an opportunity for further analysis of how past fertilizer decisions effect future decisions, but in this paper, we treat multiple interviews with one household as multiple distinct observations.

In Sub-Saharan Africa, maize accounts for 27% of all cultivated cereal area, 34% of cereal production, and 31% of all calories consumed from cereals in the region (Smale, Byerlee, and Jayne 2013). Therefore, after combining the three waves of LSMS data, we focused on households that cultivated at least one maize plot in the year of their interview. This left us with a sample size of 3,112 households. Table 1 describes the fertilizer use decisions of those households which will ultimately be used as our dependent variables.

Table 1 - Fertilizer Use Among Sample of Maize Farming Households

	Inorganic Fertilizer	Organic Fertilizer	Either Variety
Use of fertilizer on any maize plot (Adoption)	22.3%	20.7%	37.1%
Share of maize plots receiving fertilizer (Share)	19.9%	16.9%	33.2%
Quanitity of fertilizer used on average in kg/ha (Volume)	39.188	84.561	x

3.2 Weather and Soil Data

In order to model the returns to fertilizer use experienced by each household in the three years preceding their fertilizer decisions, we used data on the weather and soil conditions for each household. This weather and soil data are used to predict the difference between the yield in a given location between a representative farmer that does use fertilizer and one that does not. This yield difference, along with a ratio between maize prices and fertilizer cost in that location, is then used to determine each VCR.

We obtained our weather data from The United States National Aeronautics and Space Administration's (NASA) AgMERRA-2 dataset. This dataset offers hourly observations on surface air temperatures and total precipitation at a 0.5 x 0.625 degree resolution. The crop growth period consists of the five months after planting (Banziger et al. 2006). Within this period, we calculated average temperatures and total precipitation in the 1-2, 3, and 4-5 month windows post planting. The third month was more specifically described because this period coincides with flowering and silking, a period that is considered especially sensitive to water and temperature stress.

We obtained soil data from the Africa Soil Information Service (AfSIS) dataset used by McCullough et al. (2018). This soil data contained information on several characteristics including pH, clay percentage, water retention capacity, etc.

Additionally, we matched Agro-ecological zone (AEZ) classifications from GAEZs to each household, and we determined household planting dates based on trial site data compiled by Lobell et al. (2011).

This compiled dataset could be used to predict site-specific yields with and without fertilizer use in the three years preceding the fertilizer decision contained within the LSMS data. The process used to determine these yield differences will be discussed in further detail in a later section. These yield differences enable us to calculate the VCRs in the three years prior to our observed fertilizer decisions. Figure 1 visualizes how the data were used to determine these three VCRs.

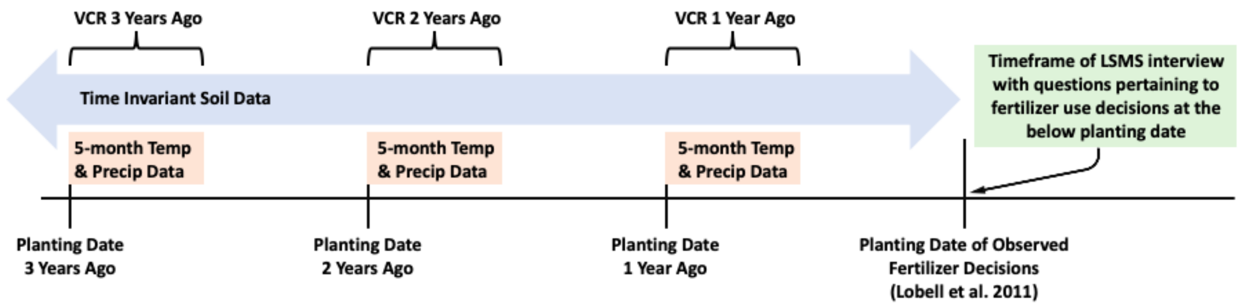

Figure 1: Timeframe of various data sources compiled to calculate VCRs

3.3 VCR Simulation Data

The last source used to compile our final dataset was the simulated site-specific yield differences and VCRs generated by McCullough et al. (2018). In their paper, the authors used the same soils data and a synthetic climate dataset to simulate site-specific, forward-looking predictions of fertilizer response across Sub-Saharan Africa. We geo-matched these simulated yields with and without fertilizer to each household in our sample, assigning 100 observations of simulated VCRs to each observation.

The details of yield difference and VCR estimation will be discussed at greater length in the next section, but one key aspect of our data is worth mentioning now. For continuity's sake, it was crucial that we developed our yield estimations (and consequently, VCRs) analogously to McCullough et al. (2018). One goal of this paper is to compare the impact of recent VCRs with the simulated VCR dataset. If there was any discrepancy in the generation of the two sets of

VCRs, we would ultimately compare the difference in methods for calculating the VCRs rather than the sets of VCRs themselves. To that end, we were sure to use analogous AgMERRA data, analogous soils data, and the same computational process to arrive at yield predictions from the weather and soils data in recent years that McCullough et al. used for their paper.

4 MODELING EXPECTED RETURNS TO FERTILIZER USE:

The key to understanding adoption of agricultural technologies among farmers lies in deciphering the web of factors they consider when making adoption decisions. These factors can be as simple as the distance to the nearest market, or as nuanced as time and risk preferences. Certainly, one of the key determinants of a farmer's decision is the profitability of the technology (Feder, Just, and Zilberman 1985). However, with respect to inputs such as fertilizer, farmers must decide whether or not to invest in a technology without yet knowing how yield increasing and profitable it might be. Therefore, we set out to understand how farmers internalize the probabilistic distribution of outcomes that will result from their use of fertilizer.

This section will consist of three subsections. First, we will discuss our two measures of fertilizer returns: yield difference and VCR. We will then investigate two specific frameworks through which farmers might form expectations about those returns: a site-modeled framework, and a recent-year framework. Finally, we will discuss how we plan to capture those expectations as explanatory mechanisms for a farmer's decisions.

4.1 Generating Measures of Returns to Fertilizer Use

The data discussed in the previous section allows us to, for any given location, estimate a farmer's expected yields with and without the use of fertilizer. Once we have made those estimations, we can calculate the *yield difference* outcome, and from the yield difference, by introducing a cost ratio, we can calculate a profitability measure, the *VCR*. Despite the fact that VCR is the fertilizer return measurement used for our primary analysis, we will discuss yield difference first as it is a central component of VCR.

Yield Difference

In order to capture the returns to fertilizer use experienced by a farmer, we must estimate their yields with and without the use of fertilizer. The difference in yields of crop y at location (*i*) and time (*t*) applying quantity (*q*) of fertilizer (*f*) can be represented by the following equation:

$$YIELD\ DIFFERENCE_{it} = y(f_q, M, X_i, w_{it}) - y(f_0, M, X_i, w_{it})$$

In this equation, the first term represents the estimated yield with fertilizer application, and the second term represents the estimated yield without fertilizer application. M refers to a vector of inputs and technologies used. Soil conditions (Sarkar and Wynjones 1982; Magdoff and Van Es 2000; Marenya and Barrett 2009) and climate (Uyovbisere and Lombim 1991; Haefele et al. 2006; You et al. 2010) also affect yields, therefore we use X_i to represent a vector of time-invariant locational soil characteristics and we use w_{it} to represent a vector of climate variables at location I and time t.

The key to calculating yield differences is to obtain estimates of $y(f_q)$ and $y(f_0)$ in each location. These estimates can only be obtained once we have estimates of the parameters of a representative farmer's production function. In order to estimate yields from our soils and climate data, we follow the procedure developed by McCullough et al. (2018). The procedure is discussed more clearly and at length in their paper, but a brief description follows.

The authors of McCullough et al. (2018) use a meta-experimental dataset from maize research trials, with each data point paired with its respective rainfall, locational, and management practice control variables. They use this dataset to estimate a production function by beginning with a yield model specification including soil and climate variables, fertilizer use, all possible interaction terms, and quadratics of the continuous variables. They select a subset of regressors based on Akaike Information Criterion (Lindsey and Sheather 2013). The authors then use a cross-validation procedure to select the yield model that best predicts yields out of sample.

We use the FGLS model generated by McCullough et al. (2018) to predict yield differences for farmers in our dataset. As discussed in last paragraph of the Data section, it is crucial that our yield difference generation in analogous to the McCullough et al. (2018) paper, as we used their simulated dataset. Both our raw data and our FGLS yield response function mirror those used in their simulation.

Value-to-Cost Ratio (VCR)

The VCR functions as a single profitability measure that approximates the benefits and costs that an individual farmer faces when they adopt fertilizer. The below equation describes how to calculate VCR of crop y using quantity q of fertilizer f:

$$VCR_y(f_q) = (delta(y) * p^y) / (delta(q) * p^f)$$

In this equation, p^y and p^f are the prices of the crop and the fertilizer respectively. Delta(y) is the yield difference and delta(q) is the application rate of fertilizer. For the purposes of our paper, the calculation of VCR is rather straightforward. We handle prices and fertilizer quantity using the same process as McCullough et al. (2018) for the sake of continuity. Fertilizer and maize prices (contained within the matched soils dataset) are multiplied by 260 USD/MT and 170 USD/MT respectively, according to world prices, and fertilizer treatment volume is assumed to be .1 MT/ha. Yield difference is determined through the FGLS estimation method discussed above. Therefore, the VCR equation at location I and time t could be represented as follows:

$$VCR_{yit}(f_q) = ([y(f_q, M, X_i, w_{it}) - y(f_0, M, X_i, w_{it})] * p^y) / (delta(q) * p^f)$$

It is also important to note for later analysis that VCR can also be decomposed to the following equation, which essentially is an interaction between yield difference and a cost ratio:

$$VCR_{yit}(f_q) = [y(f_q, M, X_i, w_{it}) - y(f_0, M, X_i, w_{it})] * [\, p^y /(delta(q) * p^f)]$$

A VCR over 1 indicates that the adoption of fertilizer is profitable. However, VCR has several limitations, including its exclusion of factors such as market access, agricultural favorability, and uncertainty faced by farmers. Consequently, researchers often look for a VCR over 2 to be confident that benefits outweigh the costs (Morris et al. 2007). However, we are more interested in the probabilistic distribution of a farmer's potential VCR, thus we do not measure the effect of VCR so bluntly. We will discuss our methods for describing a farmer's distribution of VCR's in section 4.3.

4.2 Frameworks for Internalization of Fertilizer Returns Probabilities

Recent-Year

This set of yield differences and VCRs consists of only three values. We obtained temperature and precipitation data for the three years prior to the season to which the LSMS fertilizer use questions refer. Using the processes laid out above, we used this weather data combined with time invariant data to calculate yield differences and VCRs for each recent year.

Conceptually, the recent year distribution of fertilizer return outcomes will affect a farmer's decision making if the farmer has a degree of recency bias. To the degree that weather is not correlated from year to year, this recency bias leads to a miscalculation on the farmer's part. It is worth noting that a farmer's fertilizer decisions may also be impacted by the prior year's VCR through lasting wealth effects. A rich or poor harvest, especially in the year immediately prior to a fertilizer decision, could determine how much cash a farmer has available to purchase fertilizer. We control for possible wealth effects by including the LSMS response to household expenditures per member as a control variable. However, even if this wealth effect does influence our results, the practical implications are the same. Whether a farmer makes decisions based on recent years because of recency bias, a wealth effect, or some other mechanism, it is still worth noting that this is how the decisions are made.

Site-Modeled

The site-modeled framework is meant to capture a full probabilistic distribution of returns to fertilizer use that a farmer at a given location will experience. Our site-modeled framework consists of a distribution of yields simulated with and without fertilizer for each location over multiple synthetic climate realizations. These simulated yields are based on geographically matched soil data and historic climate data that are held constant. The synthetic climate data are generated for each location and matched to the farmers with the farmer's geolocation. Each year of the simulated data contains two yield estimations – one with, and one without fertilizer. We take the difference of these yields to calculate a distribution of yield differences for each farmer, and multiply those yield differences by the cost ratio to calculate a distribution of VCRs for each farmer.

Conceptually, a farmer's fertilizer use decisions will be consistent with the site-modeled distribution if the farmer successfully internalizes typical weather outcomes, either through gained experience or disseminated information, *and* the farmer believes that weather outcomes are not serially correlated. Many farmers may not make decisions in accordance with this framework for lack of knowledge (an inefficient outcome, as an accurately internalized distribution of typical VCRs would incentivize most farmers to adopt fertilizer). Additionally, if a farmer believes that this year's weather outcomes are related to last year's weather outcomes, then a sort of recency bias develops, and the farmer will not determine fertilizer use based on the site distribution.

Comparison of Recent-Year and Site-Modeled Distributions

Once recent-year and site-modeled sets of VCRs have been assigned to every farmer, it is useful to illustrate what that data looks like across all farmers, and for each individual farmer.

Across all farmers, the recent-year and site-modeled data are generally similar. The general timeframe corresponding to the recent-year datasets turned out to be a time of mostly favorable fertilizer-use conditions. In particular, the three years preceding the wave 3 LSMS survey produced an average VCR of 11.614 across all farmers, which is statistically significantly above the average simulated VCR of 7.53. While this difference may seem large, a more detailed breakdown of VCRs illustrates how VCRs broke down across all farmers within each framework.

Figure 2 provides a comparison between the site-modeled and recent-year data using by using a density plot to describe each. The blue density plot, describing recent-year data, is slightly to the right of the pink density plot, describing site-modeled data, reaffirming that in the years preceding our sample, fertilizer returns were slightly higher than usual. Despite this slight difference, the plots are similar enough that they do nothing to invalidate our confidence that we estimated the recent-year data exactly analogously to the McCullough et al (2018) estimation of site-modeled VCRs.

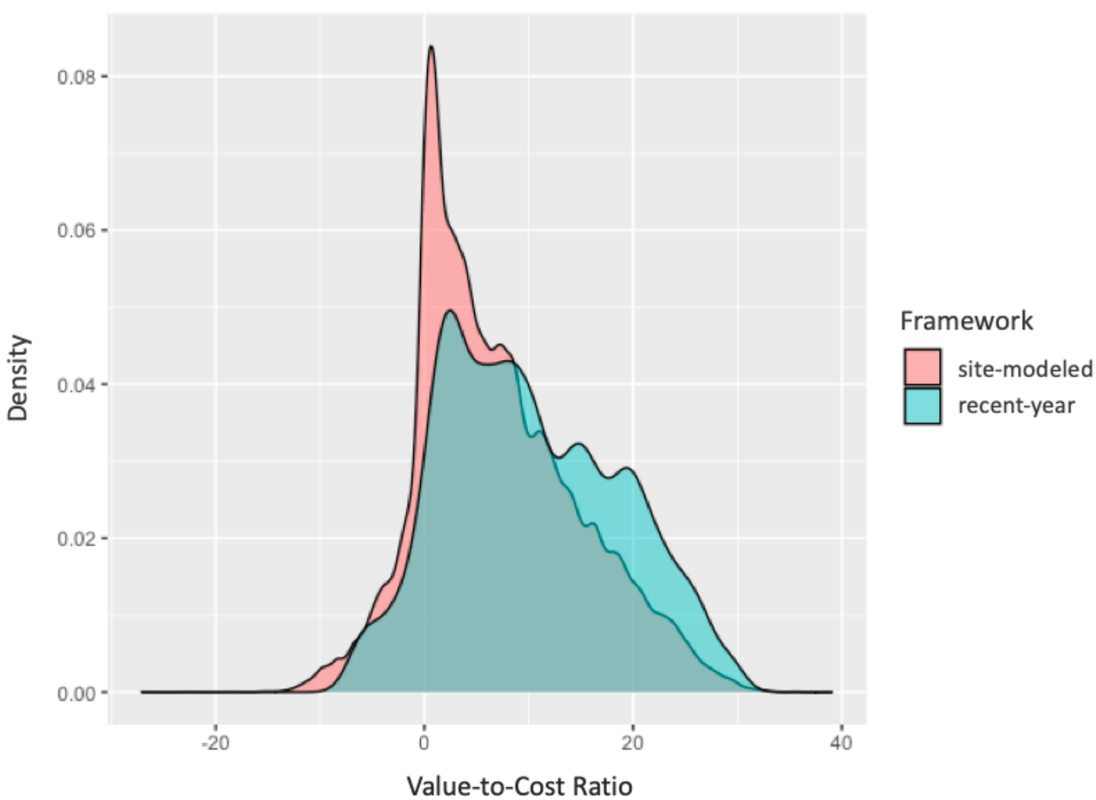

Figure 2: Density Plots of Aggregated Site-Modeled and Recent-Year Data

Before discussing how we will parameterize each farmers VCR distributions in the next section, it is useful to graphically illustrate the nature of the data for an individual farmer. Figure 3 displays the site-modeled and recent-year VCRs of a farmer selected randomly from the second wave of LSMS data.

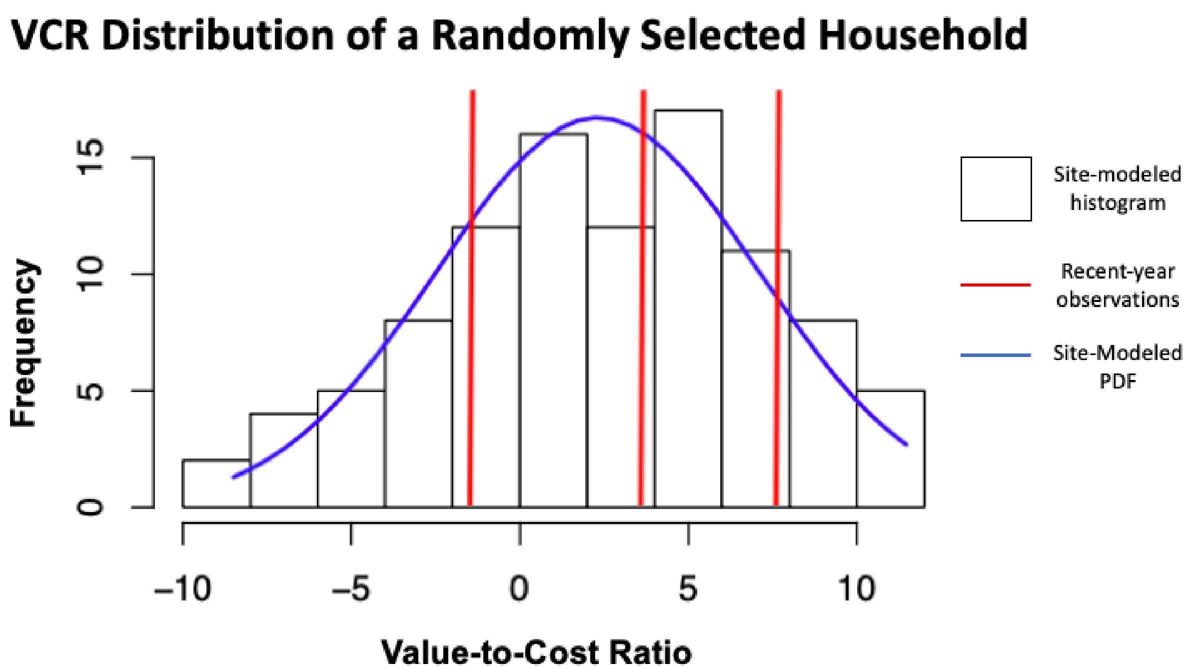

Figure 3: Comparison of recent-year and site-modeled VCR distributions for a randomly selected farmer

The histogram in Figure 3 illustrates the full set of 100 site-modeled VCRs for this farmer. The blue curve approximates a normal probability distribution based on the set of site-modeled VCRs. The spike-plot containing three red lines displays the three VCRs in the three years preceding this farmer's observed fertilizer use decision. The next section focuses on how we parameterize these two frameworks in such a way that they can be used as explanatory variables in our regressions.

4.3 Using Recent-Year and Site-Modeled Frameworks to Explain Decision Making

Our analysis focuses on the comparison of recent-year and site-modeled sets of VCRs to explain adoption, therefore, we must formulate methods to compare the two sets. Within the literature, there are many methods employed to analyze the effect of VCR values on adoption. When conducting *ex post* analysis of adoption, researchers simply use a farmers experienced VCR in a given season to analyze their fertilizer decision that season. This rather straightforward method assumes farmers have a deterministic, rather than probabilistic view, and isn't necessarily useful for explaining how a farmer is thinking when they make their fertilizer decision, *prior* to when that season's VCR is realized. Despite the fact that a farmer's thinking about fertilizer response is probabilistic (Anderson and Hardaker 2003), researchers may make a point estimate of VCR, requiring that they assume crop response and prices are certain (Spielman, Kelemwork, and Alemu 2011). Even in the presence of a probability distribution, researchers may be concerned with some critical value, and its position within the distribution. While all of these approaches possess pragmatic and statistical strengths, we seek to determine which distribution possesses the most explanatory power in explaining fertilizer use behaviors.

The first specification we use to compare the recent-year and site-modeled distributions simply involves their means and standard deviations. This method provides a straightforward path to explaining the farmers experience of expectations (mean), and risk (standard deviation).

Statistically, this method parametrizes the basic nature of the two distributions, and practically, this method offers a simple explanation of a farmer's decision-making.

The second specification we employ contains the standard deviation and includes one critical value. In our VCR analysis, this critical value refers to the proportion of VCR observations in the distribution over 1. In theory, a VCR over 1 indicates that fertilizer use was profitable, offering a conceptual justification for the selection of this critical value. This method expands on the first method by including a variable indicating the likelihood a farmer receives profitable returns. In this sense, it adds an additional practical angle to the exact shape of a farmer's distribution.

Our final specification is based entirely on critical values. More specifically, it includes variables indicating the proportion of a farmer's distribution of VCRs that is under 1, between 1 and 4, and over 4. These bins aren't selected entirely arbitrarily – a VCR of 1 indicates profitability in theory, and VCRs over 4 are almost certain to be profitable regardless of any exogenous circumstances. However, this model is interesting statistically in that it uniquely captures the shape of a farmer's distribution. In practical terms, some otherwise arbitrary value could have great significance to many farmers. Perhaps, for instance, a farmer is aware that if they were to experience any VCR under 4, they would risk entering into a poverty trap. Therefore, it is neither the mean, nor the standard deviation that determines their willingness to adopt fertilizer, but solely the likelihood that their returns are above or below this level.

5 EMPIRICAL STRATEGY:

Our empirical strategy can be broken into two sections. Our primary focus, the effect of site-modeled and recent-year VCR distributions on inorganic fertilizer use, is contained in the first section. In the second section, we isolate yield differences and analyze the effect of their distributions on both fertilizer varieties: inorganic *and* organic. Within each of the two sections, we use three outcome variables (adoption, volume, and plot share), and for each outcome variable, we run two regressions – one for the recent-year distribution, and one for the site-modeled distribution.

In this section, we will first provide a brief overview of the variables contained within our regressions. We will then provide our regression equations.

5.1 Variables

Outcome variables

We use three outcome variables in our regressions, concisely described as “adoption”, “volume”, and “plot share”. The adoption outcome is a dummy variable equal to one if the farmer used any of the given fertilizer variety on any of their plots in the year of the LSMS survey. The volume outcome represents the average application rate among adopting farmers of the given fertilizer variety across all cultivated maize plots farmed by a farmer in kg/hectare during the season of the survey. The share outcome variable represents the share of cultivated maize plots that received the given fertilizer variety in the season of the survey.

In the first section of results, these outcomes apply only to inorganic fertilizer use. The second section of results uses these three outcomes across three fertilizer varieties: *inorganic*, *organic*, and *either.* It is important to note that *either* pertains to the use of inorganic *or* organic fertilizer. Therefore, the adoption and share values, on average, will be between 1 and 2 times that of inorganic or organic fertilizer individually.

Fertilizer Response Explanatory Variables

These variables were largely discussed in section 4.3. It is only worth noting that our yield difference explanatory variables take the same form as our VCR explanatory variables (mean, standard deviation, and critical values).

Control Variables

Our vector of control variables within each regression include the following: a dummy equal to one if the household head is female, years of education of the household head, real expenditures per household member, an index for quarter of the survey, distance from the household to the nearest agricultural market in kilometers, the number of maize plots cultivated, the total acreage of the maize plots cultivated (as reported by the survey respondent), and a dummy equal to one if the household owns any of the maize plots they cultivated.

5.2 Regressions

Primary Regressions: Impact of VCR on Inorganic Fertilizer Use

The following generalized regressions will be used to evaluate the impact of the VCR distribution of framework *F* (recent-year or site-modeled) on farmer *i*'s fertilizer decisions according to the fertilizer use outcome, *P*:

$$Fertilizer\ Use_{P,i} = \beta_1(mean_{F,i}) + \beta_2(sd_{F,i}) + \delta(controls_{i}) + e_{i}$$

$$Fertilizer\ Use_{P,i} = \beta_1(pVCR > 1_{FW,i}) + \beta_2(sd_{F,i}) + \delta(controls_{i}) + e_{i}$$

$$Fertilizer\ Use_{P,i} = \beta_1(pVCR < 1_{FW,i}) + \beta_2(1 < pVCR < 4_{FW,i}) + \beta_3(pVCR > 4_{FW,i}) + \delta(controls_{i}) + e_{i}$$

The "pVCR" variables refer to the proportion of the distribution that is contained within the designated bin. When running the regressions, we omit the third bin, however, all three are displayed here to more clearly depict the goal of the third specification. It is important to understand how these regressions are generalized, and ultimately represent 18 different regressions. These three different regression forms are meant primarily to display the different sets of explanatory variables, or specifications, used to describe the VCR distributions. They are each run on both site-modeled and recent-year distributions, *and* the analysis is conducted across three outcome variables. The adoption outcome variable is analyzed through a probit regression, the volume is analyzed using a double hurdle model, and the plot share is analyzed using an OLS regression. Figure 4 illustrates each specific regression.

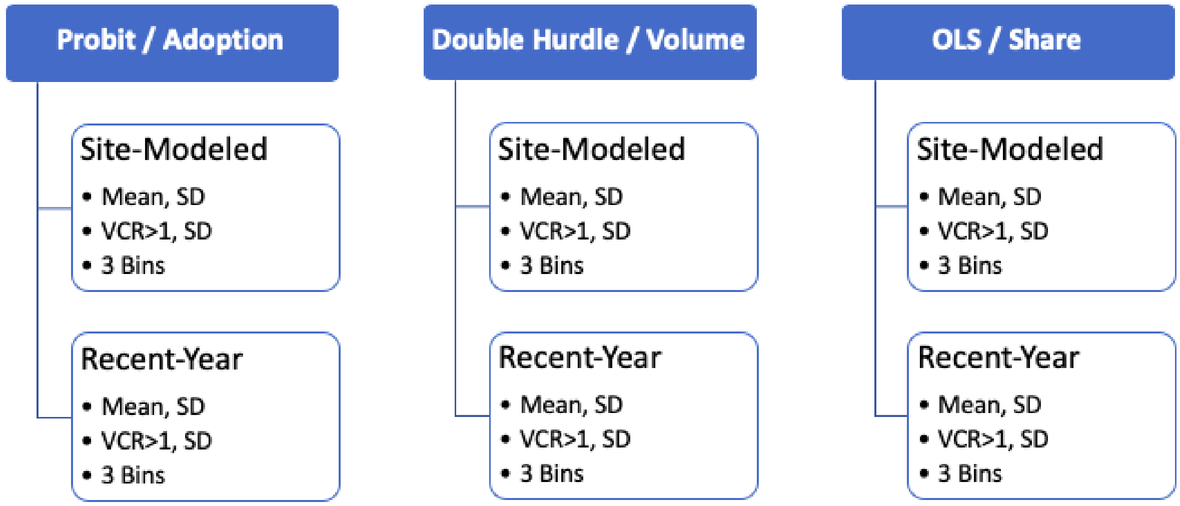

Figure 4: Breakdown of various regressions

Secondary Regressions: Decomposing the profitability components of fertilizer use

The VCR is a great tool for describing the returns to inorganic fertilizer use, the primary interest of this paper. However, because our VCR calculations include the price and quantity of inorganic fertilizer, it isn't a reasonable explanatory variable when we expand our analysis to include organic fertilizer. Therefore, we turn to the distribution of yield differences, backed out of the VCR values. Using these yield differences, we run the exact same regressions as above, extending our analysis to include outcomes for *inorganic*, *organic*, and *either* fertilizer variety.

The following generalized regressions describe our three sets of explanatory variables for farmer *i* using framework *F* to analyze adoption of fertilizer variety *V* measured by outcome *P*:

$$Fertilizer\ Use_{V,P,i} = \beta_1(mean_{F,i}) + \beta_2(sd_{F,i}) + \delta(controls_{i}) + e_{i}$$

$$Fertilizer\ Use_{V.P,i} = \beta_1(pYD > 1_{FW,i}) + \beta_2(sd_{F,i}) + \delta(controls_{i}) + e_{i}$$

$$Fertilizer\ Use_{V,P,i} = \beta_1(pYD < 1_{FW,i}) + \beta_2(1 < pYD < 3_{FW,i}) + \beta_3(pYD > 3_{FW,i}) + \delta(controls_{i}) + e_{i}$$

The regressions are similar to those listed in the VCR section with a few small exceptions. Data from the yield difference distributions is included, rather than the VCR distributions. The third bin in the third regression uses 3 rather than 4. This is simply because few yield differences exceed 4, therefore, the regression is more descriptive of the full distribution if we use a critical value of 3. Figure 4 can be helpful to understanding the nature of these regressions as well, with the caveat that this section of results focuses on three different fertilizer varieties.

6 RESULTS:

6.1 Primary Results

Our primary results pertaining to inorganic fertilizer use explained by VCR are contained in tables 2-4.

Table 2 contains several interesting findings. Using the first specification for describing the distribution, the effects of the recent-years and site-modeled distributions are similar. A higher mean VCR decreases the likelihood of adoption slightly, but the effect of the mean is less than that of the standard deviation. Both frameworks tell a similar story using the second specification. Both regressions of the second specification show decreasing adoption at higher standard deviations, and decreasing adoption when a greater share of the distribution is over 1. The final specifications are again similar for both frameworks. We omit the bin of VCRs over 4, so coefficients can be interpreted as the marginal effect relative to those observations of VCR values over 4. Generally speaking, both frameworks show that adoption increases with an increased share of VCRs below 1 and between 1 and 4 relative to an increased share of VCRs over 4. Within the recent-year framework, farmers with VCRs below 1 and between 1 and 4 have an increased probability of adopting fertilizer by .172 and .378 relative when the VCR is over 4. Within the site-modeled framework, the increase in probability of adoption below 1 is .320, and the increase in probability of adoption between 1 and 4 is .468 relative to VCRs over 4. Finally, we see slightly higher R^2 values within the site-modeled framework, suggesting that it has more explanatory power.

Table 2 - Probit on Adoption of Inorganic Fertilizer using VCRs

	Recent Years			Site Modeled		
Specification	DI1	DI2	DI3	DI1	DI2	DI3
mean	-.014*** (.002)			-.019*** (.002)		
sd	-.034*** (.010)	-.047*** (.012)		-.034** (.014)	-.064*** (.015)	
VCR>1		-.100** (.049)			-.232*** (.040)	
VCR<1			.172*** (.045)			.320*** (.035)
1<VCR<4			.378*** (.043)			.468*** (.065)
Obs	3023	3023	3026	3026	3026	3026
P. R-square	0.1683	.1129	.1700	0.1954	.1604	0.201

a. As the coefficient estimates show, DI1 refers to "distribution information 1", corresponding to a regression containing mean and standard error explanatory variables. Likewise, DI2 and DI3 refer to their respective sets of explanatory variables

b. * indicates p-value < .1

** indicates p-value < .05

*** indicates p-value < .005

c. Standard Deviations contained in parentheses

d. Marginal effects reported

The results of Table 3 are generally muddled and insignificant. Only a small handful of the coefficients reported in this table have a p-value under .5. Thus, for either statistical or real-world reasons, VCR distributions can tell us virtually nothing about the quantity of fertilizer farmers choose to use. A different method of reporting quantities, such as demeaning the values or reporting them in quintiles may be investigated in the future. In the next section, we will discuss why the volume results may be so insignificant.

The results in Table 4 largely substantiate and complement our findings in table 1. The means of the recent-year and site-modeled distributions have negative coefficients, but the standard deviations have stronger effects. According to the first specification, an increase of 1 in the standard deviation of the recent-year distribution decreases the share of plots that will

Table 3 - Double Hurdle on Volume of Inorganic Fertilizer using VCRs

Specification	Recent Years DI1	Recent Years DI2	Recent Years DI3	Site Modeled DI1	Site Modeled DI2	Site Modeled DI3
mean (use)	-13.851 (27.370)			-19.664 (351970)		
mean (qty)	-28.033 (63.668)			-53.368 (112.133)		
sd (use)	27.766 (72.089)	-39.901 (26.090)		-28.245 (93077)	-55.323 (75.916)	
sd (qty)	41.593 (336.622)	31.468 (328.521)		14.109 (605.189)	-29.469 (551.937)	
VCR>1 (use)		-114.873 (164.343)			-239.321 (707.637)	
VCR>1 (qty)		-548.231 (1278.401)			-792.418 (1472.587)	
VCR<1 (use)			177.905 (690290)			315.460 (535.530)
VCR<1 (qty)			596.673 (1288.363)			821.159 (1177.237)
1<VCR<4 (use)			335.689 (232752)			396.457 (404.339)
1<VCR<4 (qty)			201.155 (1147.600)			42.949 (2428.891)
Obs	3023	3023	3026	3026	3026	3026

Table 4 - OLS on Share of Fertilized Plots using VCRs

Specification	Recent Years DI1	Recent Years DI2	Recent Years DI3	Site Modeled DI1	Site Modeled DI2	Site Modeled DI3
mean	-.012*** (.001)			-.014*** (.001)		
sd	-.030*** (.004)	-.034*** (.008)		-.040*** (.007)	-.055*** (.013)	
VCR>1		-.121** (.058)			-.256*** (.051)	
VCR<1			.177*** (.055)			.318*** (.047)
1<VCR<4			.436*** (.060)			.481*** (.086)
Obs	3023	3023	3026	3026	3026	3026
P. R-square	0.1506	.1027	.1727	0.1668	.1497	.1870

receive fertilizer by 3%, and an increase of 1 in the standard deviation of the site-modeled distribution decreases the share of plots that will receive fertilizer by 4%. Table 4 also mirrors Table 2's results within the specifications that use bins. Following the findings in Table 2, we find that according to both specifications, a increase in the share of VCRs below 1 and between 1 and 4 increases adoption relative to an increase in the share of the VCR above 4. Lastly, we again see that the site-modeled distributions generate a slightly higher R^2 value within each specification.

6.2 Secondary Results

Our secondary results use focus on yield difference rather than VCR. This removes the ratio of maize price to the cost of inorganic fertilizer. It makes sense that these results would have slightly less explanatory power with regard to inorganic fertilizer, as the price effects are removed from the regression. Despite this fact, we still include inorganic fertilizer in this analysis because it offers some insight into how much inorganic fertilizer use is effected by yield difference alone, independent of associated price fluctuations. In the discussion section of the paper, we will divide our results by fertilizer variety. However, in this section we will use one table to describe each outcome variable in order to simplify the interpretation of the coefficients within each table.

Fertilizer Adoption

Table 5 describes the effect of yield difference distributions on adoption of inorganic, organic, or either fertilizer variety.

Table 5 - Probit on Adoption of Fertilizers using Yield Differences

Specification		Recent Years DI1	Recent Years DI2	Recent Years DI3	Site Modeled DI1	Site Modeled DI2	Site Modeled DI3
Inorganic	mean	-.055*** (.007)			-.077*** (.008)		
	sd	-.078** (.040)	-.056 (.041)		-.102 (.067)	-.091 (.063)	
	YD>1		-.209*** (.037)			-.363*** (.033)	
	YD<1			.280*** (.035)			.645*** (.063)
	1<YD<3			.191*** (.041)			.469*** (.076)
	Obs	3051	3051	3054	3044	3044	3054
	P R-Square	0.1325	.1139	.1346	0.1636	.1642	.2074
Organic	mean	.017*** (.005)			.008 (.006)		
	sd	.005 (.030)	-.004 (.030)		.124*** (.044)	.139*** (.045)	
	YD>1		.017 (.029)			.079** (.036	
	YD<1			-.058** (.027)			-.049 (.033)
	1<YD<3			-.147*** (.024)			.027 (.032)
	Obs	3051	3051	3054	3044	3044	3054
	P R-Square	0.0414	.0370	.0513	0.0405	0.043	.0393
Both	mean	-.022*** (.008)			-.044*** (.009)		
	sd	-.052 (.040)	-.044 (.040)		.054 (.065)	.058 (.065)	
	YD>1		-.122*** (.045)			-.209*** (.044)	
	YD<1			.123*** (.042)			.300*** (.043)
	1<YD<3			.008 (.040)			.196*** (.042)
	Obs	3051	3051	3054	3044	3044	3054
	P R-Square	0.0598	0.0607	0.0602	0.0726	.0711	0.0792

The results relating yield difference distributions to inorganic fertilizer adoption are mostly complementary of the results in section 6.1. An increasing mean has a negative effect on adoption, but not this negative effect is weaker than the effect of an increasing standard deviation. Farmers are more effected by the variance of the distribution when making inorganic fertilizer decisions than they are by the mean. Within the bin regressions, the adoption is higher when a greater proportion of the distribution falls into the bins of lower VCR values. This could just have to do with the distribution being flatter in general, affording the farmers less certainty. In this case, an inclusion of negative yield difference critical values might increase the coefficients on the more positive critical values. However, we will discuss other mechanisms which might explain these negative coefficients in the next section.

Organic fertilizer adoption results are mixed. Within the first specification of the recent-year distributions, the mean has a strong positive effect on adoption. Favorable conditions in recent-years drive up organic fertilizer adoption. The first two specifications of the site-modeled distributions show that a higher standard deviation of yield differences significantly increases the probability of organic fertilizer adoption. The results of the third specification with regard to the site-modeled distribution are insignificant, however, the third specification of the recent-year distribution confirms reiterates the results of the first specification of recent-year data – lower yield differences in recent years decrease adoption of organic fertilizer.

Results related to the adoption of any fertilizer are also somewhat mixed. A higher mean yield difference in recent years decreases adoption, and the standard deviation coefficient is insignificant. A higher mean yield difference in the site-modeled distribution decreases the probability of fertilizer adoption and again, the effect of the standard deviation is not significant. The bins regression using the site-modeled data shows that adoption of any fertilizer increases when a higher share of VCR is contained in the bins of smaller values.

Fertilizer Volume

Table 6 displays our results for regressions of yield differences on kg/ha of inorganic and organic fertilizer. We exclude the “either” category from this analysis because it makes less sense in practice. The recommended quantity of each variety is different and plots don’t typically receive both varieties within our sample, thus the two outcomes aren’t intuitively additive. In similar fashion to our VCR analysis, these results are insignificant and inconclusive. Not a single value in table 6 is significant at the p-values we consider. We will address why the Volume results might be so muddled in the next section.

Fertilizer Plot Share

Table 7 shows the results of our regressions of yield differences on the share of plots receiving each of the varieties of fertilizer. The results are generally similar to those of the fertilizer adoption section above.

Inorganic fertilizer plot share decreases by 5.1% with an increase in 1 of the yield difference mean in recent years, according to the first specification. Plot share decreases by 8.0% with an increase of 1 in standard deviation in the yield difference in recent-years according to the first specification. Similarly to the adoption section, the main takeaway from the bins regression is that as higher proportions of the yield differences are over 4, inorganic fertilizer adoption decreases.

Organic fertilizer plot share generally increases with an increase in the mean yield difference. The standard deviation of yield differences in recent years does not have a significant effect on plot share, but an increase in 1 of the standard deviation in the

Table 6 - Double Hurdle on Fertilizer Volume Using Yield Differences

Specification		Recent Years DI1	Recent Years DI2	Recent Years DI3	Site Modeled DI1	Site Modeled DI2	Site Modeled DI3
Inorganic	mean (use)	-.51.554 (45.906)			-71.393 (119.379)		
	mean (qty)	-45.722 (298.736)			-124.306 (301.315)		
	sd (use)	-23.612 (132.406)	-3.905 (8719.013)		-21.171 (1374.552)	-4.217 (281.002)	
	sd (qty)	859.814 (1300.904)	880.284 (1274.481)		1281.749 (2897.862)	1364.576 (2874.305)	
	YD>1 (use)		-196.743 (1923.463)			-330.075 (358.127)	
	YD>1 (qty)		-195.623 (1351.776)			-348.451 (1392.08)	
	YD<1 (use)			244.8634 (8.75e7)			568.413 (28869)
	YD<1 (qty)			-82.841 (1442.69)			273.085 (1952.634)
	1<YD<3 (use)			144.326 (5.04e8)			384.404 (40566.3)
	1<YD<3 (qty)			-477.614 (1160.494)			-381.818 (2321.703)
	Obs	3051	3051	3054	3044	3044	3054
Organic	mean (use)	14.491 (16.587)			5.816 (23102.15)		
	mean (qty)	57.421 (50.280)			15.141 (66.644)		
	sd (use)	-6.179 (34.698)	-13.884 (7.91e6)		70.478 (72405.22)	81.776 (197949)	
	sd (qty)	-115.533 (320.813)	-139.459 (319.772)		-47.462 (616.290)	-22.254 (607.483)	
	YD>1 (use)		34.598 (1.65e7)			59.166 (1.35e6)	
	YD>1 (qty)		291.097 (283.107)			151.811 (360.424)	
	YD<1 (use)			-60.559 (302762)			-38.238 (822216)
	YD<1 (qty)			-314.511 (294.126)			-123.471 (363.931)
	1<YD<3 (use)			-98.883 (116312)			21.106 (467924)
	1<YD<3 (qty)			-120.814 (357.715)			70.268 (359.217)
	Obs	3051	3051	3054	3044	3044	3054

Table 7 - OLS on Share of Plots Receiving Fertilizer Using Yield Differences

Specification		Recent Years DI1	Recent Years DI2	Recent Years DI3	Site Modeled DI1	Site Modeled DI2	Site Modeled DI3
Inorganic	mean	-.051*** (.004)			-.068*** (.004)		
	sd	-.080*** (.019)	-.060** (.032)		-.090** (.035)	-.103* (.056)	
	YD>1		-.222*** (.044)			-.373*** (.041)	
	YD<1			.267*** (.041)			.426*** (.039)
	1<YD<3			.155*** (.038)			.175*** (.035)
	Obs	3051	3051	3054	3044	3044	3054
	P R-Square	0.1214	.1090	.1233	0.1433	.1537	.1659
Organic	mean	.017*** (.004)			.009** (.004)		
	sd	0.013 (.018)	.005 (.027)		.091** (.033)	.105*** (.039)	
	YD>1		.030 (.023)			.083*** (.030)	
	YD<1			-.069*** (.023)			-.061** (.027)
	1<YD<3			-.126*** (.021)			.019 (.029)
	Obs	3051	3051	3054	3044	3044	3054
	P R-Square	0.035	.0298	.0433	0.0323	.0355	.0328
Both	mean	-.026*** (.005)			-.045*** (.005)		
	sd	-.055** (.023)	-.046 (.037)		0.027 (.042)	.026 (.059)	
	YD>1		-.141*** (.044)			-.229*** (.045)	
	YD<1			.145*** (.041)			.293*** (.042)
	1<YD<3			.019 (.038)			.157*** (.037)
	Obs	3051	3051	3054	3044	3044	3054
	P R-Square	0.0721	.0734	.0727	0.0877	.0874	.0949

site-modeled data increases organic fertilizer plot share by 9.1% according to the first specification. The results from the regression using bins in recent years shows confirms the findings from the first specification In recent-years – higher levels of yield differences increase fertilizer use.

The share of plots that receive any fertilizer decreases by 2.6% with an increase of 1 in yield difference in recent years according to the first specification. Also in recent years, an increase of one in standard deviation decreases the share of plots that receive fertilizer by 5.5% according to the first regression. The first specification relating to the site-modeled data shows an increase of 1 in mean decreases share of plots by 4.5%. In both of the site-modeled regressions containing standard deviation, the effect of the standard deviation is insignificant. Finally, in the site-modeled distribution using bins, we find that plot-share increases with the proportion of the yield differences falling into the two smaller bins, compared with the bin of yield differences over 3.

7 DISCUSSION:

Our results suggest several straightforward mechanisms that might be at play when farmers make fertilizer use decisions, however, many of our results are either counterintuitive or inconclusive. Prior to a brief discussion our takeaways related to each fertilizer variety, it is important to consider a likely cause for our mixed results. Figure 5 illustrates the mean yields with and without fertilizer, and the mean yield differences, within different ranges of VCR.

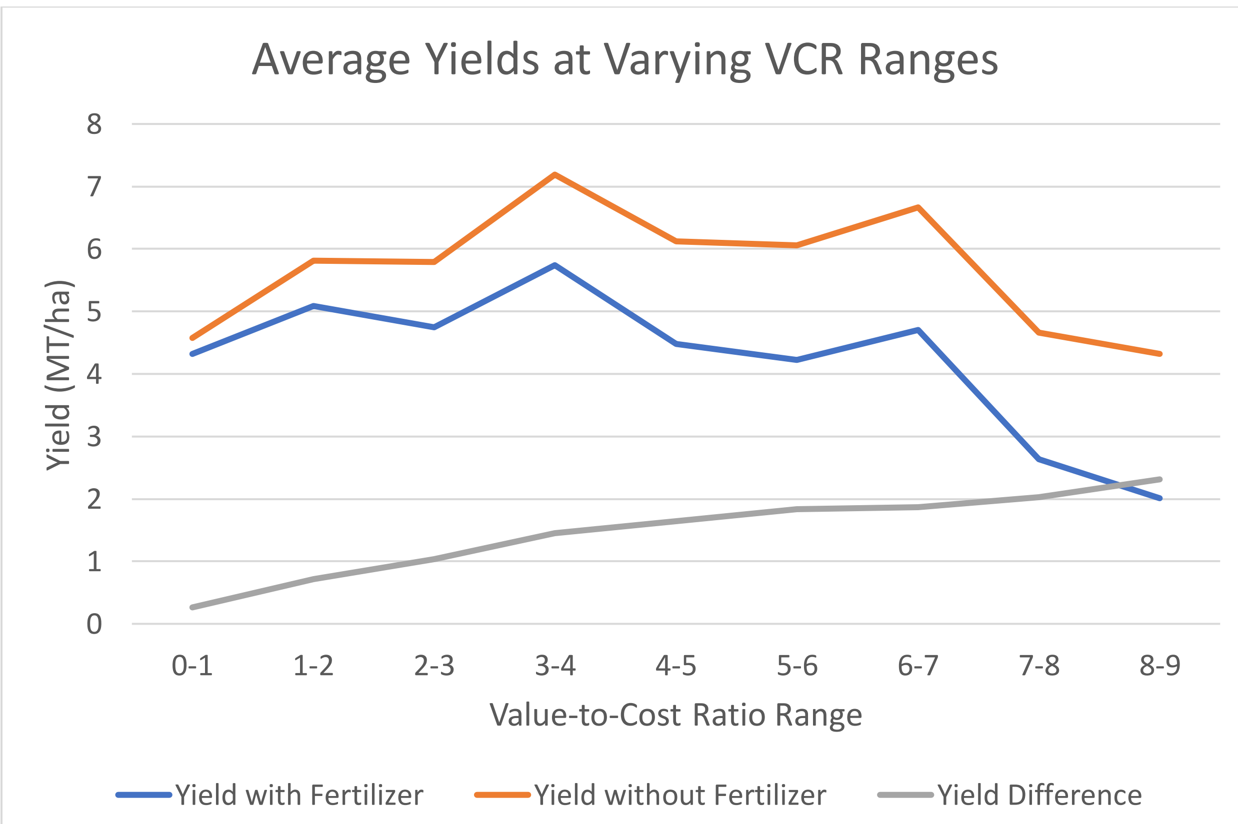

Figure 5: Line plot comparing yields and VCRs

Maximum yields occur around a VCR of 3-4, at which point, with the exception of yields at the 6-7 range, yields decrease as VCR and yield difference increase. This trend adds both clarity and complexity to our results. As we consider the effects of increased yield differences and VCRs, especially in recent years, it is crucial to realize that often times, farmer's simultaneously experience a good year in terms of returns to fertilizer use, and a bad year, generally speaking, in terms of yields. Within our dataset, farmers who experience extremely high or low returns to fertilizer use are usually also experiencing low yields in general, especially if they have failed to use fertilize in the past. With this in mind, we will discuss how fertilizer use returns effect the use of inorganic and organic fertilizer.

7.1 Inorganic Fertilizer Use

Several trends emerge relating inorganic fertilizer use to the yield difference and VCR distributions. Regardless of specification, distribution (recent-year or site-modeled), and explanatory variable (VCR or yield difference), we typically find that farmers are more likely to use inorganic fertilizer if the standard deviation of the returns is lower. This finding indicates a degree of risk aversion among farmers. It is crucial to a farmer that both in a typical year, and in the past three years specifically, the variability in the returns to fertilizer use is not too high. The lowering of perceived risk when adopting inorganic fertilizer may be even more important to the farmer than the expected mean of the return. Adoption decisions are significantly more nuanced than a simply assessment of expectations and risk, however it does seem evident that when targeting inorganic fertilizer programs, practitioners should consider the stability of the returns to fertilizer in the past.

The other trend that generally persisted across all related regressions was the decrease in use corresponding to high VCR and yield difference values. These results are largely confirmed by our specification using bins. In general, as a greater proportion of the returns fall into the bins of smaller returns, fertilizer use decreases. This counterintuitive finding illustrates the limits of our analysis, but affords us several possibilities to expand the analysis of our data. On the surface, it appears that as farmers experience higher returns to fertilizer use, even in recent years, they are less likely to adopt it in the future. However, this is likely an over-simplistic conclusion related to a more complex, but conceptually intuitive, mechanism.

The negative effect of higher returns in recent years could be due to lingering negative effects of recent harvests. As VCRs and yield differences increase, yields, particularly among the ~80% of farmers who didn't use fertilizer in the previous season, are low. We do control for household expenditures, which should absorb the effects of a lean year to some degree. However, inorganic fertilizer may be a luxury good with a rather elastic demand among farmers who have recently experienced high-VCR low-yield years. Consequently, despite the fact that inorganic fertilizer would have had an above average return in the previous years, farmer's cannot afford to invest. The similar effect in the site-modeled data could be less directly related to the same story. Farmers with higher VCRs in the site-modeled data might chronically suffer from low returns. This would perpetually lower adoption among the very farmers who would experience the highest returns. These mechanisms could be further explored in the future by converting our data to panel-data, or by analyzing heterogenous effects using additional variables available in the LSMS data.

One observation related to the comparison of recent-year and site-modeled data is that, in general, the R^2 values of the site-modeled regressions are slightly higher. The difference is not significant, and the regressions generally have similar results, saving practitioners the trouble of considering two opposing interpretations depending upon which distribution they favor. However, it does appear that the site-modeled distributions have slightly more explanatory power when predicting inorganic fertilizer use.

Finally, we observe that the regressions on fertilizer volume are entirely inconclusive. There are two potential reasons this might occur. First, we generated the variable by dividing self-reported fertilizer volume by self-reported field acreage. Reporting error is not only likely, but it could even be correlated with variables which are also correlated with VCR. Second, fertilizer volume decisions likely occur independently of the preceding adoption decisions, and past returns may have far more impact on the adoption decision. Within many of the volume regressions, the control variables for ownership and acreage were significant. Perhaps farmers consider past returns, decide whether or not to adopt, and then shift their thinking to the acreage and ownership status of their plot to answer the question of "how much?"

7.2 Organic Fertilizer Use

Our analysis of organic fertilizer use (and use of either fertilizer) reveals a few interesting trends. However, these regressions have significantly lower R^2 values compared to the inorganic fertilizer regressions. It is likely that organic fertilizer decisions largely occur independently of our yield difference values, therefore the mechanisms discussed in this section are less direct than those discussed in the previous section. Generally speaking, organic fertilizer adoption and plot share increases with yield difference mean and standard deviation. The regressions on volume are again inconclusive because of the same reasons discussed regarding inorganic volume.

Every regression containing the mean of yield differences, with the exception of the adoption regression using the site-modeled data, suggests that as the mean increases, organic fertilizer use increases. The regressions using bins largely substantiate this finding, as organic fertilizer use typically increases as a greater share of the returns fall into bins of higher VCRs and yield differemces. Contrary to inorganic fertilizer, these decisions indicate that the positive effect of higher returns outweighs the negative effect of lower yields. Organic fertilizer is often cheaper and more accessible than inorganic fertilizer, or it is available for free on the farm in the case of manure, making it the reasonable choice when a farmer is aware of high fertilizer returns, but has suffered from poor yields. This mechanism is further supported by the positive coefficients on high yield differences in recent years. Cash strapped farmers who have experienced low-yields and high fertilizer returns in recent years are likely to use organic fertilizer.

Organic fertilizer use also increase with the standard deviation of the site-modeled distribution specifically. This could be the result of two possible mechanisms. First, farmers might view organic fertilizer as a risk-mitigating choice. A deep analysis of the benefits and costs of inorganic and organic fertilizer at varying yield levels could shed some light on this notion. Perhaps, in a context of high yield and fertilizer return variability, inorganic fertilizer is too expensive to compensate for the degree to which it improves yields in years nearer to the tails of the harvest's distribution. In contrast, organic fertilizer might present itself as a more affordable means to mitigating the risk of a low yield in the presence of highly variable growing conditions. Second, organic fertilizer use might increase with site-modeled standard deviation because farmers with typically variable yields are likely to exist in a more agriculturally diverse context (Martin and Magne 2015). Organic fertilizer is derived from plant and animal sources that might be more abundant in a region with a higher standard deviation of site-modeled yield differences. In some cases, farmers may have even diversified to own livestock, and reported to the LSMS surveyor that they used organic fertilizer because they simply applied manure to their crops. Organic fertilizer might be more cheap or available to farmers who live in areas with more variable yields and yield differences.

8 Conclusion:

Our results, taken as a whole, suggest that a nuanced understanding is required of how farmers form expectations about fertilizer returns. Sweeping claims and policy suggestions remain illusive from our analysis, but there are a few available takeaways and room for further research.

Farmers are certainly less apt to use inorganic fertilizer in general as variability increases in both the recent-year and site-modeled data. This reaffirms the usefulness of viewing input adoption as a whole, and fertilizer adoption in specific, as a story of risk. Our data suggests that farmers may be more responsive to a risk mitigating policy, such as indexed insurance, as opposed to a policy that might increase expected returns, but does less to protect them from the variability of fertilizer returns. Similarly, in extension efforts there may be less use that one might assume in communicating to a farmer that fertilizer will increase their yields on average. A full view of our analysis suggests that the most compelling marketing to a farmer might truthfully suggest: "inorganic fertilizer will, on average, make your worst years less bad". Of course, for this to be true, the cost of inorganic fertilizer must be limited to some degree, thus policies such as subsidies are worth considering. However, we may benefit from framing subsidies through their effect on variability and risk. "Subsidies decrease fertilizer prices" may be an over simplistic story. An advocate of subsidies might suggest "subsidies decrease fertilizer prices, which decreases the variability of returns to fertilizer use according to most measures of returns."

There is ample opportunity to expand on our research, both through other papers, and through the continued improvement of this paper. We have only began to explore the potential lines of analysis afforded by a dataset combining simulated and recent-year fertilizer response distributions. Our dataset lends itself to two additional forms of learning that we will consider exploring: learning from neighbors and learning from past adoption decisions. Thus far, we haven't explored any village or enumerator id level analysis, and we have ignored the fact that some observations in our sample are actually the same household appearing over multiple years of LSMS data. Our analysis offers a useful first look into the formation of beliefs about fertilizer returns, but additional work is available to gain a more complete understanding of how farmers think about past and future fertilizer returns.

<u>REFERENCES:</u>

African Union. 2003. “Maputo Declation.” *Maputo: African Union*

Anderson, J.R., and J.B. Hardaker. 2003. “Risk aversion in economic decision making: Pragmatic guides for consistent choice by natural resource managers.” In J. Wesseler, H. P. Weikard, and R. D. Weaver, eds. *Risk and Uncertainty in Environmental Economics*. Gloucestershire, UK: Edward Elgar Publishing, pp. 171–188.

Banziger, M., P.S. Setimela, D. Hodson, and B. Vivek. 2006. “Breeding for improved abiotic stress tolerance in maize adapted to southern Africa.” *Agricultural Water Management* 80:212–224.

Bold, T., K. Kaizzi, J. Svensson, D. Yanagizawa-Drott. 2017. “Lemon Technologies and Adoption: Measurement, Theory, and Evidence from Agricultural Markets in Uganda” *The Quarterly Journal of Economics* 132(3): 1055-1100

Chavas, J.P., and M.T. Holt. 1996. “Economic Behavior Under Uncertainty: A Joint Analysis of Risk Preferences and Technology.” *Review of Economics and Statistics* 78:329–335.

Duflo, E., M. Kremer, and J. Robinson. 2007. “Why Don’t Farmers Use Fertilizer? Experimental Evidence from Kenya.” Unpublished.

Duflo, E., M. Kremer, and J. Robinson. 2008. “How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya” *American Economic Review* 98:2, 482-488

Duflo, E., M. Kremer, and J. Robinson. 2011. “Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya” *American Economic Review* 101(6): 2350-90

Dugger, C. 2007. “Ending Famine, Simply by Ignoring Experts” *New York Times*, December 2.

Evanson, R., D. Gollin. 2003. “Assessing the Impact of the Green Revolution, 1960 to 2000.” *Science,* 300(2): 758-62

Feder, G., R.E. Just, and D. Zilberman. 1985. “Adoption of agricultural innovations in developing countries: A survey.” *Economic development and cultural change* 33:255– 298.

Haefele, S.M., K. Naklang, D. Harnpichitvitaya, S. Jearakongman, E. Skulkhu, P. Romyen, S. Phasopa, S. Tabtim, D. Suriya-arunroj, S. Khunthasuvon, D. Kraisorakul, P. Youngsuk, S.T. Amarante, and L.J. Wade. 2006. “Factors affecting rice yield and fertilizer response in rainfed lowlands of northeast Thailand.” *Field Crops Research* 98:39–51.

Holden, S. 2018. “Fertilizer and sustainable intensification in Sub-Saharan Africa.” *Global food security* 18: 20-26

Hopper, W. 1993. “Indian agriculture and fertilizer: An outsider’s observations” *Keynote address to the FAI Seminar on Emerging Scenario in Fertilizer and Agriculture: Global Dimensions. New Delhi: FAI*

Jayne, T.S., and S. Rashid. 2013. “Input subsidy programs in sub-Saharan Africa: a synthesis of recent evidence.” *Agricultural Economics* 44:547–562.

Lindsey, C., and S. Sheather. 2013. “Variable selection in linear regression.” *Stata Journal* 13:65–76.

Lobell, D.B., M. Banziger, C. Magorokosho, and B. Vivek. 2011. “Nonlinear heat effects on African maize as evidenced by historical yield trials.” *Nature Clim. Change* 1:42–45.

Magdoff, F., and H. Van Es. 2000. *Building Soils for Better Crops.*, 2nd ed. Beltsville, MD: Sustainable Agriculture Network.

Marenya, P.P., and C.B. Barrett. 2009. "State-conditional fertilizer yield response on western Kenyan farms." *American Journal of Agricultural Economics* 91:991– 1006.

Martin, G., and M.A. Magne. 2015. "Agricultural diversity to increase adaptive capacity and reduce vulnerability of livestock systems against weather variability – A farm-scale simulation study" *Agriculture, Ecosystems, & Environment* 199: 301-311

McCullough, E., J. Quinn, A. Simons, L. Verteramo, J. Woodard. 2018. "Modeling climate smart soil health investments in Sub-Saharan Africa"

Mellor, J., C. Delgado, and M. Blackie. 1987. "Priorities for accelerating food production growth in Sub-Saharan Africa" *Indian Journal of Agricultural Economics* 45(3): 183

Morris, M., V.A. Kelly, R.J. Kopicki, and D. Byerlee. 2007. "Fertilizer use in African agriculture: Lessons learned and good practice guidelines." Washington, DC: World Bank.

Otsuka, K., K. Kalirajan. 2006. "Rice green revolution in Asia and its transferability to Africa: An introduction." *The Developing Economics* 44(2): 107-122

Quiñones, M., N. Borlaug, and C. Dowswell. 1997. "A fertilizer-based green revolution for Africa" *Replenishing soil fertility in Africa* 51: 81-95

Rashid, S., N. Tefara, N. Minot, and G. Ayele. 2013. "Can modern input use be promoted without subsidies? An analysis of fertilizer in Ethiopia" *Agricultural Economics* 44(6): 595-611

Sarkar, A.N., and R.G. Wynjones. 1982. "Effect of rhizosphere pH on the availability and uptake of Fe, Mn and Zn." *Plant and Soil* 66:361–372.

Schultz, T. 1964. "Transforming Traditional Agriculture" New Haven, CT: Yale University Press

Sheahan, M., and C. B. Barrett. 2017. “Ten striking facts about agricultural input use in Sub-Saharan Africa.” *Food Policy* 67:12-25

Smale, M., D. Byerlee, and T. Jayne. 2013. “Maize Revolutions in Sub-Saharan Africa.” In K. Otsuka and D. F. Larson, eds. *An African Green Revolution*. Dordrecht: Springer Netherlands, chap. 8, pp. 165–195.

Spielman, D.J., D. Kelemwork, and D. Alemu. 2011. “Seed, fertilizer, and agricultural extension in Ethiopia.”, pp.

Uyovbisere, E.O., and G. Lombim. 1991. “Efficient fertilizer use for increased crop production: The sub-humid Nigeria experience.” *Fertilizer research* 29:81–94.

World Bank. 2014. “Fertilizer consumption (kilograms per hectare of arable land)”

Yanggen, D., V. Kelly, T. Reardon, A. Naseen, C. Liedholm, and M.T. Weber. 1998. “Incentives for fertilizer use in sin-Saharan Africa: A review of empirical evidence on fertilizer response and profitability.”, pp.

You, L., C. Ringler, G. Nelson, U. Wood-Sichra, R. Robertson, S. Wood, Z. Guo, T. Zhu, and Y. Sun. 2010. “What is the irrigation potential for Africa?” Working paper.