Job Market Paper

"Why do Farmers Grow the Crops they Do?" The Impact of Crop Choice on Agricultural Productivity and Poverty

Ubabukoh Chisom*

*Department of Economics
University of Manchester

The aim of this paper is to examine the factors which determine the crop choices of small-holder farmers in Nigeria and how these choices affect productivity and welfare outcomes. Using the two-rounds of LSMS panel data from Nigeria in 2010/11 and 2012/13; I start by re-examining the old arguments surrounding whether small-holder farmers are indeed "efficient-but-poor". I find that smallholders are generally efficient in their allocation of resources (after estimating household crop productivity by stochastic frontier analysis), but are not necessarily rational in their crop choices because even when some crops are found to be more productive than others, the less productive crop is often chosen. To figure out why, a treatments effect model is employed to determine farmer selection into the choice of a type of crop in the first stage; and subsequently the impact of their choices on productivity and poverty. I find that access to free inputs, non-farm income and the use of seeds from the previous growing season are some of the important determinants of crop choice. In addition, I also find many similarities between the crop choice made and the idea of risk aversion.

JEL classification: D24, I32, N57, 013, 033

Keywords: Agricultural Productivity, Poverty, Crop Choice, Stochastic Frontier Analysis, Treatment Effects Model, Nigeria

1 Introduction

The main purpose of this research is to identify the nature of the relationship between the choice of smallholder farmers on what crops to grow on the productivity and welfare outcomes of the farm households. Their productivities are also analysed to determine the nature of the differences between the farmers who grow different crops. Additionally, the specific characteristics and properties of the subgroups of farmers within the sample are also examined to determine if any useful information can be obtained, and an attempt is made to identify the determinants of such choices. The data to be used comes from the two waves of Nigeria's General Household Survey-Panel (GHS-Panel), which is part of the World Bank's Living Standards Measurement Study – LSMS.

Furthermore, an attempt will be made to link this idea with the growing literature on what effects the risk attitudes of farmers have on their investment decisions (or on the crop they choose to grow). Dercon and Christiaensen (2011) showed that the less income a farmer has, the more risk averse he will be, and the more risk averse a farmer is, the less likely he will be to invest more in his farm operation or to adopt new technology.

Producing cash-crops have traditionally been looked upon as the forte of large-scale commercial farmers, but in more recent times, there have been arguments that perhaps smallholder farmers could also take advantage of the large international market these products have and synergise their efforts to raise overall productivity[1] and improve their incomes. Thus this research proposes to study these arguments in closer detail – do smallholder farmers who engage in the production of more exported crops experience significant productivity differences from those that don't and do they have better welfare outcomes?

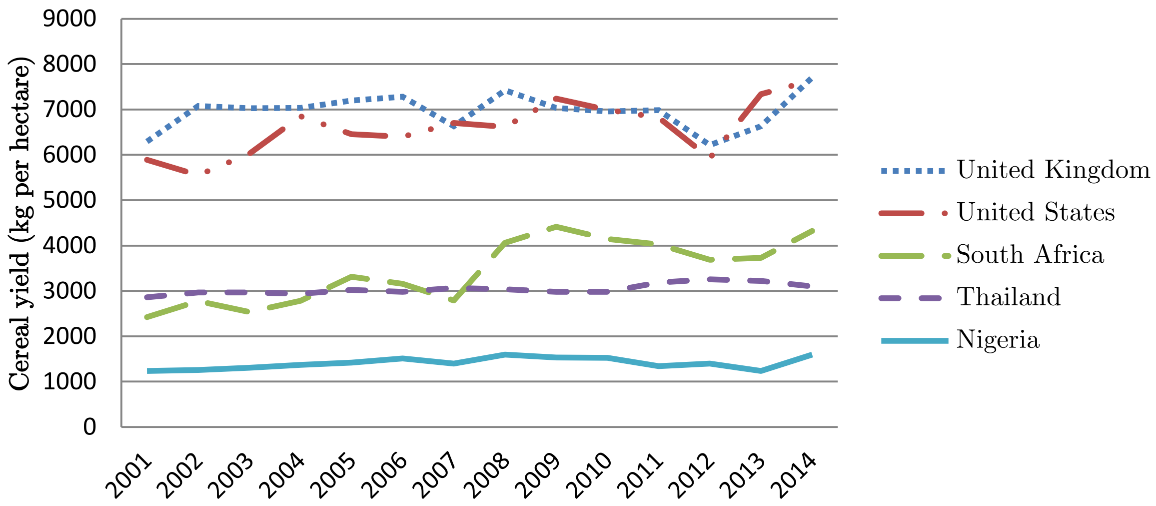
Nigeria is an appropriate country to use as a case study because it is the picture of a country with an agricultural sector trapped in a cycle of low productivity. Nigeria may be classified as a lower-middle-income country (by the World Bank definition) with a national GDP of $568.51 billion (as at 2014); and with an estimated population of 177.5 million people, has a gross national per capita GDP of $2,970 (World Bank, 2016). The average growth rate of GDP between 2007 and 2014 was 6.09%, which is comparatively higher than several other countries in Europe and even Africa who barely managed to average 5% within the same time frame.

However, about 82.20% of the population lives on less than $2 a day (World Bank, 2016), with a large proportion of this poor engaged in agriculture. Agriculture accounts for about 40% of the country's GDP and employs about 65% of the people (World Bank, 2016). Thus the agricultural sector is absolutely important in determining the quality of life and welfare of a large proportion of people in the country. Even though many people engage in agriculture, it has lagged behind other

[1] Productivity here is defined as total factor productivity or output after inputs have been accounted for.

sectors and the rest of the world in terms of productivity. To illustrate this, the graph below shows agricultural productivity of a few countries, proxied by cereal yields in kg per hectare and it is clear that in comparison, Nigeria is not doing as well as it can be.

Figure 1: Time Trend of Cereal Yield for Selected Countries



Source: Author's drawing from World Development Indicators (WDI), 2016 database

The low agricultural productivity in Nigeria could be due to a large number of factors ranging from poor soil quality due to erosion, pollution and leaching, to the scarcity and high cost of inputs. Others may be the continued use of crude implements, and traditional (non-modern) farming practices. However this paper intends to show that all other things being equal, the type of crop a farmer chooses to grow, even at the same levels of technology could be important for the outcomes of that household in terms of productivity (technical efficiency) and poverty.

To illustrate this further, below is a table of selected crops, the area of land planted with the crop, their prices, the average output in tonnes and their average revenues per hectare.

Table 1: Selected Crops with Outputs, Prices and Expected Revenues

Crop	Land Area ('000 ha)	Output ('000 metric tons)	Avg. Price per kg (Naira)	Avg. Revenue per ha ('000 Naira)
Yam	3236.16	37328.17	76.07	877.45
Cassava	3481.88	42533.17	65.31	797.79
Cocoyam	520.12	2957.09	80.00	454.83
Cotton	398.56	602.44	230.22	347.99
Melon	469.7	507.34	123.06	132.92
Rice	2432.64	4472.51	72.03	132.43
Maize	4149.33	7676.85	64.65	119.61
Guinea corn	4960.13	7140.96	73.08	105.21
Beans	2859.77	3368.24	83.03	97.79
Groundnut	2785.17	3799.15	69.02	94.15
Soyabeans	291.38	365.06	60.03	75.21
Millet	4364.16	5170.45	58.53	69.34

Source: Nigerian Bureau of Statistics (NBS), 2009

It is clear that some crops give more revenue than others (this, of course, does not include input and production costs, for which some crops would also have higher costs of production than others, thus narrowing the profit margins[2]), but this gives a general idea of the motivation. If there are crops that yield higher revenues, and importantly, a farmer is free to choose among all these crops equally, all other things being equal (like weather and soil variability), why would he choose to grow a crop that provides a smaller profit margin than the other crops? And by how much would choices of this kind impact on their productivity and household welfare? These are the questions we set out to answer in the paper.

This research is important for a couple of reasons. Firstly, from studying the productivity of farmers in Nigeria, it would be clearer where opportunities exist for improvements and the important factors which when increased (or reduced as the case may be), would result in the highest productivity increases. This is especially important as more countries and charity organisations channel a large part of their foreign aid investment into agriculture (Addison & Tarp, 2015). They, and the local authorities, would need to know where the greatest gains could be achieved. Should the focus be more on moving the technological frontier forward with innovation, or to try to raise efficiency on the current frontier by encouraging better use of inputs, or perhaps a combination of both?

Also, understanding the reasons why farmers choose to grow the crops they do could help policy makers know where to focus when trying to encourage the production of certain crops, for which they believe their country has a comparative advantage and where they feel the best national gains could be achieved, perhaps in terms of a reduction in foreign exchange expenditure or for food security. For

[2] This is however taken into account in our computation of productivity, and the analysis shows that there are indeed productivity differences between crops, even after accounting for all input and production costs.

example, each government that comes into power in Nigeria would often come with their own agenda for the agricultural sector, pushing for more production of certain "important" crops (Iwuchukwu & Igbokwe, 2012).

In addition, poverty and food security is a major concern for many sub-saharan African countries. It's a problem because if a country is not self-sufficient in the production of certain staple foods, they risk starvation when faced with adverse shocks in the production function. This thus becomes a matter of life and death for them. If the production of certain crops can improve the welfare outcomes of the farmers or reduce the food insecurity in bad seasons, this knowledge would be important. The angle of crop choice being a likely proxy for a measure of risk aversion could potentially also be important for researchers who face the daunting task of planning field experiments and using methods from behavioural economics or psychology to estimate the risk aversion of farmers.

It is only when the right links are known concerning what the drivers of productivity, or poverty perpetuation are, that progress could be made in determining the possible ways that intervention could be used in solving these problems. For example, according to Karlan et al. (2013), the unavailability of credit may not be biggest problem to the productivity question, in the sense that, even if all farmers had access to equal amounts of credit, not all farmers may decide to make use of the necessary amounts to raise their overall productivity. They argue that in this way, risk attitudes might potentially be even more important, therefore propping agricultural self or micro-insurance as a really vital piece of the puzzle. Because if farmers take out a large loan, but are not sure about what their output would be, it could present to them like just a really big gamble. Could it be that the type of crop a farmer decides to grow is a form of self-insurance? This illustrates why studying what types of crops farmers plant, why they choose those crops, and what effects such seemingly innocuous choices could have is important.

The results show that access to free inputs, non-farm income, the use of seeds from the previous growing season, household size, gender and the different regional differences are the main determinants of crop choice. Also, the choice influences the productivity and poverty of the households, although not in the ways that may be expected. In addition, commercialization was found to be important for poverty alleviation, but not for productivity improvements.

The rest of this paper is laid out thus: following this introduction, there will be a brief literature review highlighting some of the work that has been done on productivity of smallholders and the effects of decisions to grow a crop on productivity and welfare, both generally and in the specific Nigerian context; thereafter the economics behind the methodologies to be used are developed, starting with how the key crop choice variables are defined in this paper; and then the data section, analyses and results. Following this, there is a general discussion on risk aversion and how this could be related to crop choice.

2 Brief Literature Review

Productivity (Efficiency) and the Small Farmer

The starting point of any discussion on smallholder productivity would be the "poor-but-efficient" hypothesis, set out by Ted Schultz in his 1964 book "Transforming Traditional Agriculture". Therein, he posits that small farmers in traditional agricultural settings, though poor, are reasonably efficient in their allocation of resources. To quote from the book, he says:

> *"…there are comparatively few inefficiencies in the allocation of factors of production in traditional agriculture" (Schultz, 1964, pp-37)*

Over the last four decades or so since then, this has been an enduring theme in much of the rural development economics literature. Many have boldly challenged his claims (for e.g. Adams, 1986; Chakravorty, 1984; Myrdal, 1968; Shapiro, 1983, Ball & Pounder, 1996; Duflo, 2006; Ray, 2006) alluding to different reasons, tested with different methods and with different datasets, and some others have written in support (for e.g. Hayami and Ruttan, 1985; Stiglitz, 1989; Ruttan, 2003; Nerlove, 1999; Abler & Sukhatme, 2006). This paper attempts to bring further evidence to this debate using a new dataset from Nigeria. In addition, evidence concerning the inverse productivity relationship between land size and yields have also been around since Chayanov (1926), with papers like Eswaran & Kotwal (1986); Barret et al. (2009) trying to explain this seeming relationship by either differences in labour supervision costs or market failures. But the vast and growing literature on this subject implies that these arguments are still as relevant today as they were years ago.

As well as the productivity and land relationship, there is also a large literature on the importance of agricultural inputs in the agriculture of developing countries; including the relevance of access to finance for these inputs. Bravo-Ureta & Pinheiro (1997) also note the possibility of improving economic efficiency, not by increasing the use of inputs, but by increasing overall output, leaving the inputs constant.

Technical efficiency, in this context, is the ability of a farm household to obtain optimal output from a given amount of inputs (Farrell, 1957). The level of technical efficiency can be measured by how far away a particular farm household's production is from the maximum production frontier. Thus, a farm, whose productivity lies on the frontier can be said to be technically efficient.

Agricultural Productivity in Nigeria

To measure technical efficiency, two groups of methods can be employed: parametric and non-parametric methods. Among the parametric methods, stochastic frontier models are the most common. For Nigeria, these models have been used to compute farmer efficiency for a large variety of crops including rice, wheat and cassava, among others. Stochastic frontier methods are also applied in this paper, not for specific crops here, but for all the comparable farmers in the sample; in addition, we are able to make a more detailed analysis due to the availability of panel data. The difference between the parametric (like the SFA) and non-parametric methods is that whilst production functions are of a specified form for parametric analysis, there is no such restrictive functional form employed for the non-parametric method. An example of the non-parametric approach is the data envelopment group of models (Charnes, 1978). Models of this kind rely instead on the data rather than predetermined functional forms of the production functions (Ajibefun 1998). In addition to the above, some other studies have used some partial measures of productivity like yield per hectare in their analysis.

Adeyemo et al. (2010) compute an average technical efficiency score of 0.89 for cassava farmers in Ogun state, Ebong et al. (2009) do the same for food crop farmers in Akwa Ibom and recover an average TE of 0.81. In the South-East region, Onyenweaku & Ohajianya (2009) calculate an efficiency score of 0.65 for rice farmers in Ebonyi state. Finally, Amaza et al. (2005) do the same for food crop producers in Borno and calculate an average score of 0.68. Papers like these are an indication of the range of expectation for calculated efficiency scores, but this paper does a nationwide analysis using data from the nationally representative panel household survey of Nigeria. To the best of my knowledge, this is the first time both waves of this dataset have been combined to perform a SFA for efficiencies, thus this should add to the literature in an agreeable way.

Crop Choice, Productivity and Welfare

Using national household surveys from Mali, Delarue et al. (2009) studied the relationship between cotton production and household consumption, and they discovered that cotton producers consumed an average of 9 percent more food than non-cotton producing households. This aggregated small with large producers though, so that when they were disaggregated, it was found that the largest cotton producers consumed up to 22 percent more than the smallest cotton producers, however this paper intends to delineate causal relationships, rather than just correlation. Also, Loveridge et al. (2003) do something similar for Rwanda, but with coffee, and they discovered a positive but weak relationship between coffee production and consumption outcomes of the households. They speculated that this could have been as a result of the low prices for coffee in the world market as at the time of the survey, which was 2001. Murekezi and Loveridge (2009) use the same methodology to compare the 2001 season data of Rwanda to that of 2007, to assess the impact of policy reforms and they found that technology could be a factor in the effectiveness of cash-cropping among smallholders because those that used modern techniques spent 15 percent more on food and 17 percent more on all goods than the traditional producers. This research also takes differences in production technologies into account by separating crops that have vastly different methods of production from each other in one of the classifications for crop choice (i.e tubers and roots as against the other types of crops). Also, Maertens and Swinnen (2009) found that the welfare of rural households benefit strongly from participating in high-yield vegetable exports in Senegal and Cuong (2009) finds that commercial crops have positive poverty-reducing effects on rural households in Vietnam.

3 Defining Crop Choice

The initial idea that was intended to be examined here under crop choice was the cash-crop vs food-crop debate so as to make an attempt in answering the question of if one type of crop had quantitatively better production and welfare outcomes than the other. Normally, a cash crop is defined as an agricultural crop that is grown primarily for sale in order to make a profit. The term is often used to differentiate subsistence or food crops, grown for the family of the farmer from crops that are marketed. However, in most developing countries, the term 'cash-crop' has a connotation to export and the demand for that product from developed countries (especially for industrial purposes) and not necessarily just crops that are sold at the local level. According to the US Environmental Protection Agency, cash crops are typically purchased by organisations or commercial entities separate from the farm[3]. It is not common to use the term cash-crops for livestock or animal products, but only for actual crops grown from the ground. Given these definitions, if crops were to be divided by such a straight classification, it would be quite confusing and perhaps impossible to empirically test, especially when faced with the reality of data. This is also important as this paper intends to group similar crops together rather than study farmers who grow an isolated crop against all the others. The following are some of the reasons why this cash-crop/food-crop classification might be problematic.

[3] See: "Ag 101: Crop Glossary" (2009), US Environmental Protection Agency.

Firstly, when cash crops are mentioned, the first picture that comes to the mind of a listener is that of tree cash-crops such as cocoa, coffee, palm oil, rubber et.c. However, one of the objectives of this paper was to identify what determined the choice of crop planted and if tree crops are used in the analysis, this purpose would be defeated. This is because if we are trying to measure the effect of a planting choice on productivity and thence on welfare or poverty, it is necessary that the entire life cycle of the crop is captured within the year of interest. If tree cash crops are used, there would hardly be a basis for comparison with other farm households who do not produce these crops, mostly because a tree crop takes a relatively longer time to start producing output from when it is grown, and as such the inputs used for this would not be captured at all in our measure for productivity (which accounts for all the agricultural inputs and outputs within the production year under consideration). In addition to this, many of such trees could have even been planted by a previous generation, hence nullifying the premise that a choice has been made by the household to grow that tree crop. To be properly formal, tree cash-crops should be compared with tree food crops and annual cash crops with annual food crops. Therefore excluding all the farm households with livestock and tree crops listed as their primary output was the first thing that was done in creating the crop choice variable. Thus the focus will be restricted to annual crops (those crops that can complete a life cycle within a year).

The second reason why a cash-crop vs food-crop categorization might be impractical is that going by the formal definitions, it would be difficult to allocate one crop solely to one category, apart from a few strictly non-edible crops like cotton and rubber. For example, take a crop like cassava. This is one of Nigeria's largest agricultural exports, with an average of over 45,000,000 metric tons exported per year on average, making the country the largest exporter of the product in the world. However, it is also the raw material for a major local staple food – *'garri'*, which is consumed by most households in the country. Would this crop then be classified as a cash crop or a food crop?

For these reasons, this paper creates 3 different ways in which crops could be classified without too many of these same problems:

1. **C1** - by the most exported crops (most exported crops vs. others),
2. **C2** - by type (tuber and root crops vs. others),
3. **C3** - continuous variable for degree of crop commercialization.

It is important to mention that these are by no means an exhaustive list of ways in which crops could be classified. The point here is to simply illustrate that such divisions could be helpful to tell a story about the types of crops a farmer chooses to grow, depending on what the interest of the researcher is. For example, if a researcher is interested in the differences between farmers who choose to grow vegetables as against those who don't, or perhaps those who grow cereals as against those who don't; the sample could also be divided in a similar manner to investigate this.

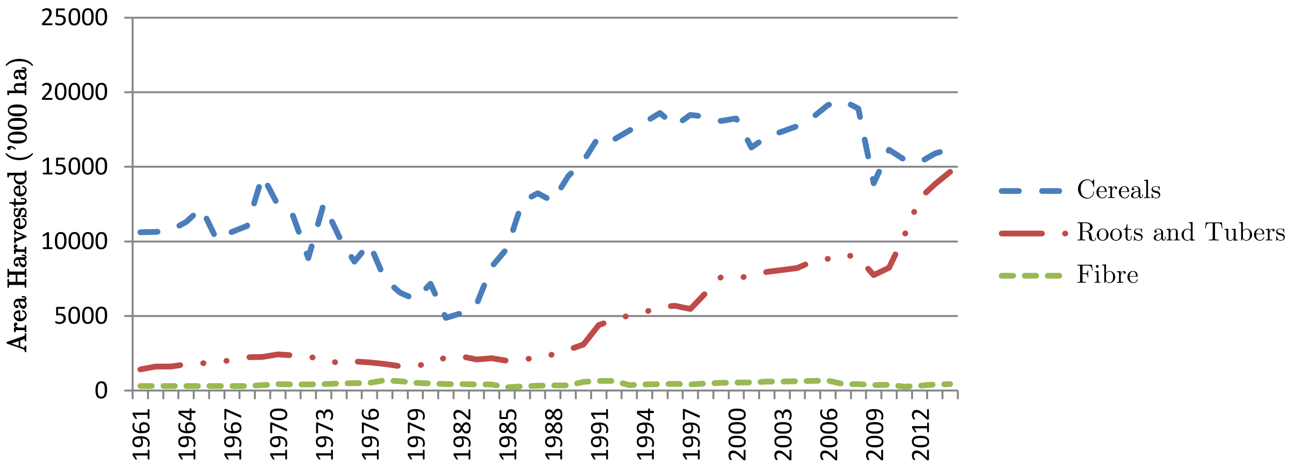
To create the variable for the first category by most exported crops, data from the FAO was examined to determine what crops were the most exported ones in Nigeria, and the farmers who grew the top 5 crops (and listed them as their primary product output) were classified as Crop-Choice 1 (C1) households. The purpose of this variable is to capture those farm households who grow those crops that are the most likely ones to be exported. As can be seen from table 2 below, 11.06% of the sample planted one of the five crops in the first wave and 7.14% planted these in the second wave. The crops used here are as follows:

Table 2: List of crops classified as C1 (by most exported)

Crops (C1)	Export ('000 metric tons)	% of sample (wave 1)	% of sample (wave 2)
Cassava	42,533.17	10.42	6.48
Sugarcane	1,429.57	0.04	0.04
Cotton	533.31	0.16	0.19
Ginger	167.29	0.08	0.08
Sesame seed (Beniseed)	127.60	0.36	0.35
Total	44790.94	11.06	7.14

For the second division, crops have been grouped by type, with tuber and root crops on the one hand against the others. This classification is important because root and tuber crops have long been recognised as particularly important to the agriculture and food security of many countries especially those in sub-saharan Africa. According to the Commission for Africa Report (2010), these types of crops are an important component of the diet for 2.2 billion people in developing countries, and in Nigeria, they were traditionally a store of wealth as you could tell how rich a person was by the size of his yam barn, for example. To illustrate this further, the first graph below (figure 2) shows that even though cereals like rice and maize have in the past been allocated more land for production by farmers than roots and tubers, this gap has been closing steadily as more and more land area is allocated to the latter. In fact there has been an upsurge in the production of tubers from around 2006.

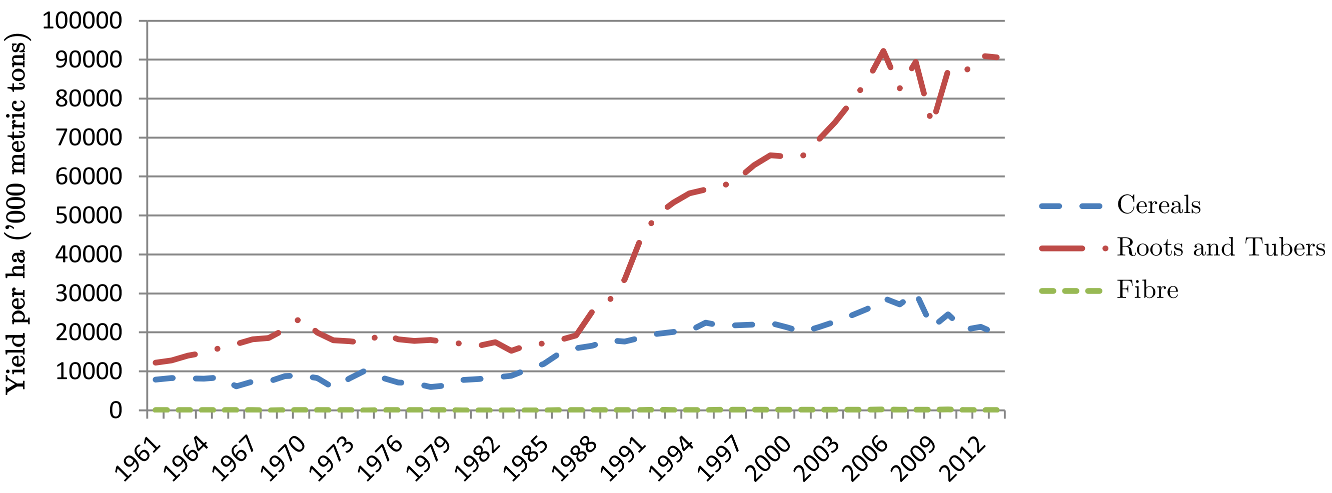
Figure 2: Time Trend of Area Harvested for Cereals, Roots and Tubers and Fiber in Nigeria



Source: Author's drawing from FAOSTAT database, 2016 database

The second graph below (figure 3) tells a similar story for yield, except that in this case, the data shows that roots and tubers have for long been a higher yielding crop type than cereals or fibers, and this productivity gap has increased dramatically over the last three decades. These diagrams just go to show how important crop divisions of this kind can potentially be.

Figure 3: Time Trend of Yield/Ha for Cereals, Roots and Tubers and Fiber in Nigeria



Source: Author's drawing from FAOSTAT database, 2016 database

However, as important as tuber and roots crops are, they have not been given as much attention as they deserve in policy making. One reason could be that in comparison to crops like wheat and rice, tuber crops are bulky, have higher water content and thus relatively shorter shelf lives (CIP Report, 2014). This constrains the development for innovations in their value chains, as well as the expansion of production and delivery at scale to processors and the markets.

Haven noted the above; four Consultative Group for International Agricultural Research (CGIAR)[4] organisations came together in 2011 to form a whole new research group devoted to the study and development of these often neglected crops – the Roots, Tubers and Bananas (RTB) research programme. The goal of the programme at launch was "to mobilize complimentary expertise and resources" in basically ensuring that sufficient research is devoted to improving the production outcomes and value-chain of these products.

In this paper, the crops classified under this category are as contained in table 3 below.

Table 3: List of crops classified as C2 (by being a tuber or root)

Crop	**% of sample (wave 1)**	**% of sample (wave 2)**
Yam	21.51	23.17
Cassava	10.42	6.48
Cocoyam	1.49	1.71
Groundnuts	1.79	1.45
Potatoes	0.58	0.64
Ginger	0.08	0.08
Total	35.87	33.53

C3 - Household Commercialization Index (HCI)

Finally, an index for the degree of commercialization of crop produce per household was used to capture the extent to which a farm household's crop production was oriented towards commercial agriculture. Following from Govereh et al. (1999) and Von Braun et al. (1995), which lay a standard in measuring commercialization; this can be calculated by taking the percentage of the value of the entire agricultural crop product in the year which is explained by the gross value of crops sold. This computation will result in a number between 0% and 100% in which a household with a HCI of 0% is one with none of its total crop product sold; while a household with an index of 100% will be one with all its crop output sold.

[4] The four organisations are: International Potato Center (CIP), which leads the research program, Bioversity International, the International Center for Tropical Agriculture (CIAT), the International Institute of Tropical Agriculture (IITA) - and the French Agricultural Research Centre for International Development (Cirad), which also represents INRA, IRD and Vitropi.

$$HCI = \left[\frac{gross\ value\ of\ crop\ sales}{gross\ value\ of\ all\ crop\ production}\right] x\ 100 \quad (1)$$

This is a neat way of transforming the binary crop choice variable into something that is continuous and which shows the range of possibilities between just the choices of which crops to produce. In addition, this variable allows for interactions to be made (with the other crop choice variables) to produce new parameters with interesting interpretations[5]. However, the limitations to the use and interpretation of this variable must be noted because it tends to give more weight to farms with smaller output who might sell a higher percentage of their output. To illustrate, if we consider a simple case where a farmer grows 5 stands of cassava, harvests and sells all 5; he would be classified as fully commercialized (100%), as opposed to a situation where a farmer grows 20 stands and sells 5. In the second case he would be measured as only semi-commercialized (with 25% on the index), even though they have both sold the same amount of crop. This notwithstanding, this measure is a useful one in describing agriculture in developing countries like Nigeria, because the smaller the farm, the more likely they would be selling less of their total output rather than consuming this at home (except for cases of higher value–added crops like cut flowers or vegetables) (Govereh et al., 1999).

4 Methodology

4.1 Stochastic frontier analysis

In this study, I estimate technical efficiency of crop production, for which I employed the aggregation of data at the household level; each observation representing a unique productive entity. Technical efficiency in this sense can be defined to be the ratio of the produced output of a farm household over the maximally possible output, given a set level of inputs. In order to achieve this, farm households had to be compared against some "ideal" farm as a benchmark; and this is often done in the literature through some form of frontier analysis. There are two major ways a production possibility frontier function may be estimated: the non-parametric Data Envelope Analysis (DEA), which was first proposed by Charnes et al. (1978) and the parametric Stochastic Frontier Analysis (SFA), independently proposed by Aigner et al. (1977) and Meeusen & van den Broeck (1977).

[5] For example, interacting commercialization with most likely to be exported crops would create a variable that represents how much of these crops are actually sold as opposed to consumed at home, which disaggregates the farm households growing this crop to some extent.

There are pros and cons to the use of either of the models and this mostly would depend on the research setting and what the researcher is trying to achieve. For the stochastic frontier model, its major disadvantage is that it requires assumptions to made on the functional form of the frontier function ahead of time i.e. either linear, quadratic, Cobb-Douglas and so on (Bogetoft and Otto, 2011). In this case, if a wrong functional form is assumed, the estimated parameters are biased. But for the data envelope model, which is deterministic, there is no stochastic error term representing the measurement errors of unobservable parameters; hence every deviation from the production frontier is explained by technical efficiency.

However, in agricultural research, there are many possible important stochastic shocks (for example, disease infestation, weather, motivation of the farmers or even luck), which could be experienced by the farm households and thus have to be accounted for (Coelli and Battese, 1996). The SFA model does this by partitioning the stochastic error term into two: systemic random/stochastic error to account for statistical noise and an inefficiency component (Battese and Coelli, 1992). Another advantage the SFA has over the DEA model is that it can be tested with conventional statistical tests due to its parametric nature (Singh et al., 2001). This is why the SFA is employed in most agricultural studies estimating efficiency by frontier and why it is also used here.

Aigner et al. (1977) and Meeusen & van den Broeck (1977) show how the error term in a stochastic frontier model can be split into: v_{it}, the stochastic error term and u_{it}, the inefficiency error term. The base model takes the form:

$$ln(Y_{it}) = ln(f(\boldsymbol{X}_{it})) + v_{it} - u_{it} \quad with\, u \geq 0 \tag{2}$$

v_{it} is either positive or negative and is assumed to be normally distributed with mean zero and constant variance, as it represents an unsystematic stochastic effect related with measurement errors and random influences (e.g luck, drought, flood, or other weather shocks, as earlier mentioned) while u_{it} is non-negative and either assumed to be half-normal or truncated normally distributed, measuring technical inefficiency, i.e the stochastic shortfall of output from the most efficient farm on the production frontier (Coelli and Battese, 1996).

Three conventional inputs are normally used in the computation of the agricultural production frontier function. These are *land* (total agricultural land area under cultivation), *labour* (total wage expenditures for labour including family labour[6]) and *inputs* (intermediate input costs like seed, fertilizer, pesticides, and cost of irrigation). In an ideal case, there would also be a variable for capital (depreciated

[6] Family labour is costed by multiplying number of hours supplied by family members with the going market wage rate per hour.

cost of machinery and buildings), but this is not included here due to data constraints. However, his should not be problem, because most smallholders in Nigeria usually own neither of these, apart from small implements like hoes and shovels and the farmers that want to mechanize would tend to rent the machines for the required period of time rather than buy them (these rental costs are included in the inputs variable already). These inputs are used to produce the output y_{it} defined as total revenue generated at the farm (including by-products). The Cobb-Douglas[7] model is employed here to fit the production frontier:

$$ln(Y_{it}) = \beta_0 + \beta_1 ln(land) + \beta_2 ln(labour) + \beta_3 ln(inputs) + v_{it} - u_{it} \quad (3)$$

Because of the non-symmetry of the conventional error term, ε_{it}, the expected value of which is defined here as, $E(\varepsilon_{it}) = -E(\varepsilon_{it}) \leq 0,\ \varepsilon_{it} = v_{it} - u_{it}$, the estimation by ordinary least squares (OLS) will provide consistent estimates of the parameters apart from the intercept. Moreover, the OLS estimation cannot extricate the technical efficiency component from its normal residual error. The maximum likelihood estimation (MLE) however can be used, as this selects values of the model parameters that produce the distribution most likely to have produced the observed data by maximizing the likelihood function; in addition, we would like the efficiency estimates to fall between 0 and 1. For this to work, we assume that the technical inefficiency error term (u_{it}) has a positive half-normal distribution and that u_{it} and v_{it} are independent. This is useful because the standard deviation of the distribution is able to concentrate efficiencies near zero or spread them out (with a zero cut off) (Aigner et al., 1977; Street, 2003).

Technical efficiency can then be derived for each farm household. It is the ratio of the output y_{it} over the stochastic frontier output when $u_{it} = 0$. The resulting technical efficiency would have a value between 0 and 1 and gives information about how far away the observation data points are from the production frontier:

$$TE_{it} = \frac{y_{it}}{exp(x_{it}\beta + v_{it})} = \frac{exp(x_{it}\beta + v_{it} - u_{it})}{exp(x_{it}\beta + v_{it})}) = exp(-u_{it}) \quad (4)$$

[7] Cobb-Douglas models without restriction and with restrictions (where the parameters are forced to be homogenous) were tried, but there was no significant difference. The time varying decay (TVD) estimation is also used as it most closely simulates a fixed effects regression, as against the time invariant (TI) version.

4.2 Treatment Effects Model

In this section, the intuition behind solving the problem of a potential selection bias in the creation of the key variables is discussed. Firstly, why should we suspect that the categorical variables we have created for crop choice (C1 and C2) might be biased by self-selection? This is because it is highly unlikely that farmers have chosen a particular crop to produce entirely at random, especially when they have chosen a crop with reduced productive capacity when they have been given an equal opportunity to grow one with higher productivity possibilities. It is far more likely that there are certain unobservable characteristics that influence their decision to produce these types of crops, and these would lead to the key variables being endogenous, as they become correlated with the error term of the main equation.

To try to mitigate these problems, we implement a treatment effects model, similar to the Heckit method (Heckman, 1979). It involves the use of a control function with an endogenous treatment variable which is the self-selection into the choice of crop a farm household has made. In addition, crop choice is likely to be an endogenous determinant of poverty and productivity. In this case, we are fortunate that we have panel data and thus are able to demean the data and control for time invariant characteristics of the sample.

The treatment effects model estimates the effect of an endogenous binary treatment, T_i (in this case crop choice), on a continuous, fully observed outcome variable, Y_i (in this case productivity and poverty in separate models); conditional on vectors of explanatory variables, $\boldsymbol{X}_i$ and $\boldsymbol{Z}_i$ (which would include exclusion restrictions). This can be modelled in the following way as our desired result:

$$Y_i = \beta T_i + \eta \boldsymbol{X}_i + \nu_i \tag{5}$$

In this case, β would be the parameter of interest as the average net effect of being treated on the outcomes. However, since T_i is endogenous, we would first need to model the selection into treatment. This could be written as:

$$T_i^* = \gamma \boldsymbol{Z}_i + \varepsilon_i \tag{6}$$

The selection into treatment T_i^* in this model is a function of ε_i, which is correlated with ν_i, the error term in the outcome equation of Y_i above. Thus, T_i^* is actually an unobserved latent variable (what is observed in the data is simply the choice, but not the underlying activity). The assumption is made that this is a linear function of the exogenous covariates $\boldsymbol{Z}_i$ and a random component ε_i. The relationship between the observed T_i and the latent T_i^* can be defined in this way:

$$T_i = \begin{cases} 1, & if \quad T_i^* < 0 \\ 0, & if \quad T_i^* \geq 0 \end{cases} \quad (7)$$

The problem here is that estimating equation 6 above directly by OLS would only be consistent if there is no correlation between ν_i and ε_i (notationally, this correlation is represented by ρ; so ideally, we want $\rho = 0$) (Green, 2008). But in this case, ρ is not zero, thus a different method would have to be used to estimate the coefficients consistently. More formally, if we assume that the treatment T_i is normally distributed, the expected conditional outcome of productivity and poverty (Y_i) could be written in this way:

$$\begin{aligned} E[Y_i|T_i, \boldsymbol{X}_i, \boldsymbol{Z}_i] &= \eta \boldsymbol{X}_i + \beta T_i + E[v_i|T_i, \boldsymbol{X}_i, \boldsymbol{Z}_i] \\ &= \eta \boldsymbol{X}_i + \beta T_i + \left[\rho_1 \sigma_{v_1} \{\phi(\gamma \boldsymbol{Z}_i)/\Phi(\gamma \boldsymbol{Z}_i)\} | T_i, \boldsymbol{X}_i, \boldsymbol{Z}_i\right] P(T_i = 1|\boldsymbol{X}_i) \\ &\quad + \left[\rho_0 \sigma_{v_0} \{-\phi(\gamma \boldsymbol{Z}_i)/1 - \Phi(\gamma \boldsymbol{Z}_i)\} | T_i, \boldsymbol{X}_i, \boldsymbol{Z}_i\right] [1 - P(T_i = 1|\boldsymbol{X}_i)] \end{aligned} \quad (8)$$

Thus, the expected outcomes for participants and non-participants have been disaggregated. The expected outcome for the treated would be:

$$E[Y_i|T_i, \boldsymbol{X}_i, \boldsymbol{Z}_i] = \eta \boldsymbol{X}_i + \beta T_i + \left[\rho_1 \sigma_{v_1} \{\phi(\gamma \boldsymbol{Z}_i)/\Phi(\gamma \boldsymbol{Z}_i)\} | T_i, \boldsymbol{X}_i, \boldsymbol{Z}_i\right] \quad (9)$$

And the expected outcome for the non-treated would be:

$$E[Y_i|T_i, \boldsymbol{X}_i, \boldsymbol{Z}_i] = \eta \boldsymbol{X}_i + \left[\rho_0 \sigma_{v_0} \{-\phi(\gamma \boldsymbol{Z}_i)/1 - \Phi(\gamma \boldsymbol{Z}_i)\} | T_i, \boldsymbol{X}_i, \boldsymbol{Z}_i\right] \quad (10)$$

Here, $\rho_1 \sigma_{v_1}$ represents the covariance between ν_i and ε_i for the treated, $\rho_0 \sigma_{v_0}$ represents the covariance between ν_i and ε_i for non-treated, $\phi(\gamma \boldsymbol{Z}_i)$ is the marginal probability of the standard normal distribution at $\gamma \boldsymbol{Z}_i$ and $\Phi(\gamma \boldsymbol{Z}_i)$ is the cumulative distribution function of the standard normal distribution at $\gamma \boldsymbol{Z}_i$. Equations 9 and 10 above include the "Inverse Mills Ratio" to control for the possible sample selection bias. The difference between the expected outcomes of the treated and non-treated becomes:

$$E[Y_i|T_i = 1, \boldsymbol{X}_i, \boldsymbol{Z}_i] - E[Y_i|T_i = 0, \boldsymbol{X}_i, \boldsymbol{Z}_i] = \beta + bias\ from\ selection \quad (11)$$

In this case, it is expected that there is a positive bias on the OLS estimates (that it overestimates the impact of crop choice on productivity and poverty), as ρ is positive. The coefficients are estimated by maximum log likelihood as this provides consistent estimates. The usual log likelihood equations are as follows:

$$lnL_i \begin{cases} ln\Phi\left\{\dfrac{\gamma \boldsymbol{Z}_i + (Y_i - \eta \boldsymbol{X}_i - \beta)\,\rho/\sigma}{\sqrt{1-\rho^2}}\right\} - \dfrac{1}{2}\left(\dfrac{Y_i - \eta \boldsymbol{X}_i - \beta}{\sigma}\right)^2 - \ln\left(\sqrt{2\pi\sigma}\right), & \boldsymbol{Z}_i = 1 \\ ln\Phi\left\{\dfrac{-\gamma \boldsymbol{Z}_i - (Y_i - \eta \boldsymbol{X}_i)\,\rho/\sigma}{\sqrt{1-\rho^2}}\right\} - \dfrac{1}{2}\left(\dfrac{Y_i - \eta \boldsymbol{X}_i}{\sigma}\right)^2 - \ln\left(\sqrt{2\pi\sigma}\right), & \boldsymbol{Z}_i = 0 \end{cases} \tag{12}$$

So in reduced form, there are two stages of regression; the first stage is the regression to estimate the probability of being treated, or for a farmer choosing to grow a type of crop, conditional on $\boldsymbol{Z}_i$; the inverse mills ratio was computed from the residuals and used in the second stage – an impact regression of the $\boldsymbol{X}_i$ and the IMR as an extra regressor to deflate the selection bias on productivity and poverty. The $\boldsymbol{Z}_i$ vector of variables used in the first stage would include selection restrictions, which are parameters that influence choice but do not "directly" influence productivity or poverty, and as such would not belong in the main impact equation of interest. Exclusion restrictions which have been used here are the amount of stored seed from the previous season used in planting the current season, and the amount of free seeds received by the farmer and used in planting. The distance of plot from the nearest extension provider was tried but proved a little problematic due to trying to find a proper distance proxy for this and multicollinearity issues[8]. Non-farm income has also been used as exclusion restriction for the productivity equation, but not from the poverty equation, as this is directly related to the mean per capita household expenditure.

For the C3 variable (the variable representing the Household Commercialization Index (HCI)) and its interactions, a different model is used in estimating its effects, mostly because this is a continuous variable (rather than a binary one), and hence, presents us with more opportunities to use a wider range of the data. A Fixed Effects (FE) model or a Correlated Random Effects (CRE) model can normally be used to address any endogeneity due to unobserved time invariant characteristics. The FE method addresses potential biases by using the variation in commercialization within a household over the two time periods to identify the causal effect of crop commercialization on productivity (Wooldridge, 2002).

[8] Defining what exactly the nearest extension provider is has been tricky. I tried proxying for this by using distance to the nearest major town with 20,000 plus residents, but this was highly correlated to the rural/urban area variable, as well as being directly correlated to poverty. Other proxies I tried include distance to nearest major road and distance of plot to town centre, but could not get very useful results. This could revisited in the future though.

However, a potential limitation of the use of the fixed effects model in this case is that we are unable to properly recover the coefficients on the time invariant observable characteristics such as regional dummies and when they are reported, must be interpreted with caution[9]. This can be an issue when important variables affecting productivity such as gender are time invariant. One other way suggested in correcting this is with the Correlated Random Effects (CRE) model. This model addresses endogeneity due to unobserved time invariant factors but still makes it possible to recover the coefficients on time invariant observed variables (Wooldridge, 2010; Sheahan et al., 2013).

5 Data

5.1 General description of data

For this analysis, I make use of the Nigerian General Household Survey-Panel (GHS-Panel) for 2010/2011 and 2012/2013, which is the most recent official comprehensive household survey for Nigeria and is part of the Living Standards Measurement Study-Integrated Surveys on Agriculture (LSMS-ISA) series from the World Bank. The panel version of the GHS was conducted by the Nigerian Bureau of Statistics (NBS) in collaboration with the Federal Ministry of Agriculture and Rural Development (FMA&RD), the National Food Reserve Agency (NFRA), the Bill and Melinda Gates Foundation (BMGF) and the World Bank. It covers all the 36 states of the country including the Federal Capital Territory (FCT), Abuja and within each state. They used a two-stage probabilistic sampling technique to select clusters (or neighbourhoods) at the first stage and households at the second stage. Clusters were selected from each of the 36 states that the country has and from the capital city. Sampling was carried out on both urban and rural Enumeration Areas (EAs) and it is thus nationally representative. According to the accompanying documentation, this panel component was created to focus on getting better information on the role of agriculture in households' economic well-being and it draws heavily on the Harmonized Nigerian Living Standard Survey (HNLSS) and the National Agricultural Sample Survey (NASS).

For the GHS-Panel, 5,000 households were surveyed out of 22,000 in the cross-sectional part. The survey for each wave was done in two stages: the post-planting period (lean season), once in 2010 and once in 2012 and the post-harvest period, once in 2011 and once in 2013. In addition, the post-planting survey includes the 22,000 cross-sectional households while the post-harvest survey includes just the 5,000

[9] This in fact renders some of the reported results uninterpretable.

households in the panel sample. The original objective was for the GHS-Panel to be repeated every two years while the normal Cross-Section component would continue to be carried out annually as it is currently done.

There are three detailed questionnaires contained in the survey, which cover a wide range of socioeconomic topics: the Household Questionnaire, the Agricultural Questionnaire and the Community Questionnaire. These questionnaires contain information on education, the observations' demographic characteristics, labour market, migration, credit and savings, household assets, non-farm enterprises, household food and non-food expenditures, food security and other non-labour income.

The household questionnaire in particular contains variables dealing with consumption, cash and non-cash income, savings, assets, food security, health, education, vulnerability and social protection. The agricultural questionnaire was only administered to the subset of the sample that was involved in non-aquatic agricultural activity, and included information on land size, agricultural inputs, access to extension services and production and marketing figures for main crops and livestock. The fishery questionnaire is not used because it does not provide the any information on crops planted, which is the focus of this research. Finally, the community questionnaire contains community or village-level data provided by several knowledgeable residents about community characteristics such as physical infrastructure, access to public services, economic activities and local retail prices of essential goods and services.

6 Results and Analysis

6.1 Agricultural Productivity in Nigeria

Table 4 below shows the results of the crop productivity estimation of farm households in Nigeria, using the methods previously outlined. The Cobb-Douglass specification applied here, does not force the coefficients to add up to one (this could be done by imposing constant returns to scale constraints on the maximum likelihood estimation of the production function, but there was no convergence in using this method and the estimates were not very different anyway). The result shows that all inputs are significantly important in the production function, but labour and land jointly contribute about 70% to output, with coefficient estimates 0.372 and 0.470 respectively. Other inputs like seeds, fertilizer, equipment etc. has a coefficient of 0.110. These results are indicative of the kind of agriculture Nigeria practices. The agricultural system is more labour intensive than capital intensive, which is fairly typical for traditional developing economies. This also shows that there might be potential for an overall frontier improvement by increasing capital intensity; whilst releasing the extra labour to other productive industries. This also ties in to the Lewis (1954) theory of surplus labour.

Table 4: Results of the Stochastic Frontier Analysis model

	Cobb-Douglas (Time Varying Decay-TVD)	
	Coefficient	SE
Constant	3.016	43.130
lnLand	0.372***	0.013
lnLabour	0.470***	0.004
lnInput	0.110***	0.005
$Sigma^2$	1.975	0.039
Gamma	0.163	0.023
$Sigma\_u^2$	0.322	0.048
$Sigma\_v^2$	1.652	0.052
$lnSigma^2$	0.680***	0.019
ilgtgamma	-1.633***	0.171
Mu	4.387	43.131
Statistics		
No. of obs.	5192	
No. of groups	3045	
Wald chi^2	1359.16***	

Note: *** represents significance at 1% alpha

Overall productivity of the farmers averaged about 68%. This is not very different from some of the other estimates that have been obtained by some other more crop specific studies (for eg. 89% by Adeyemo et al. (2010), 81% in Ebong et al. (2009), 65% in Onyenweaku & Ohajianya (2009), and 68% in Amaza et al. (2005)). This is also about the average obtained by studies designed to test the Schultz hypothesis of the efficient small farmer. Although, these productivity numbers are not too bad, there is a lot of room for improvement, even at the current levels of technology. Since non-labour variable inputs are a significant determinant of productivity, it is likely that the choice of crops grown itself is a source of inefficiency. Some evidence of this may be found in studying table 5, which shows the cross-tabulation of the crop choice variables and the average productivities of households. It will be noticed that there is, on average, higher productivities figures for households who grow either export oriented crops or tubers and roots. These differences range from 1.5% to about 5%. However, cross-tabulations are not really evidence, as they hide many possible explanatory variables for the differences. These differences are however tested in the following sections to see if they are significantly different from zero, using the distributional assumptions, utilizing the panel time framework, and controlling for other extenuating characteristics.

Table 5: Cross tabulation of crop choice variables and average technical efficiency

		C1		Difference between 1 & 0	C2		Difference between 1 & 0
		1	0		1	0	
TE	t = 1	0.660	0.640	0.020	0.666	0.651	0.015
	t = 2	0.644	0.611	0.033	0.670	0.620	0.04

By way of further analysis, table 6 shows the variation in productivity across the sample by gender, age and land size, just using the first wave alone (similar results are obtainable from the other wave also). The last row gives an overall productivity of each column division. Each of these variables provides useful information. Males in the sample are more productive than females with an average productivity of 66% as opposed to 62%. Following what we would expect, the most productive age range is between 20 and 60, and productivity appears to reduce as land size increases (as in Imai et al, 2015). Furthermore, in general most of the proportions of the sections fall within the 50-75% range of productivity.

Table 6: Productivities of different segments of the population (from Wave 1)

	Male	Female	Age (<20)	Age (20-60)	Age (>60)	Land size (<1ha)	Land size (1-5ha)	Land size (5-10ha)	Land size (>10ha)
Productivity (<25%)	4%	15%	7%	2%	7%	9%	5%	11%	2%
Productivity (25-50%)	24%	35%	19%	9%	12%	19%	20%	19%	40%
Productivity (50-75%)	62%	48%	65%	70%	66%	69%	65%	65%	46%
Productivity (>75%)	10%	2%	9%	19%	15%	8%	10%	5%	12%
Overall Average Productivity	66%	62%	64%	70%	66%	69%	69%	64%	63%

6.2 Impact of crop choice on productivity and poverty

This section reports the results of the treatment effects model to estimate the determinants of crop choice and hence the impact of this choice on productivity and poverty, proxied for by mean per capita consumption expenditure (MPCE). In essence, following from the above analysis, this section tests whether the productivity and welfare differences between the two groups of farmers are significantly different from zero, after controlling for household characteristics and these are reported in Tables 7 and 8 respectively. This analysis is done using the two categorical crop choice variables as previously defined; columns 1 and 2 being results using C1, and columns 3 and 4 being results using C2. Columns 1 and 3 in both tables are the results of the first stage selection into treatment equation, determining the probability of being treated (growing C1 and C2). However, since these are drawn from probabilistic functions and not from linear probability modelling, the coefficients cannot be directly interpreted as probabilities, but with p-values indicating significance and direction of signage indicating direction of effect. Columns 2 and 4 are the results of the impact equation of the second stage, showing the average treatment effect on the treated.

The exclusion restrictions used for the productivity equation are the amount of free inputs used in production, the amount of non-farm income the household possesses and the amount of seeds used from the previous growing season, whilst for the poverty equation, only the free inputs and previous year's seeds are used because non-farm income is directly related to household expenditure, as previously explained in the methodology section. These variables were positive and significant in determining participation to growing export oriented crops and tubers or roots.

For the use of previous year's seeds variable, the data shows that the greater the amount of primary inputs like seeds that were saved from the previous year, the more likely it would be for that farm household to plant the same crop in the next growing season. The rationale behind the use of this variable was from informal discussions with local farmers and other people who had knowledge of the sector; and a constant theme that emerged as a major driver of the decision-making process of the farm household in choosing a crop to plant is the idea of tradition and culture. Farmers may not make a deep study of the different options available to them each growing season, if they already know enough about one crop from their years of experience working with and producing a particular crop. To help mitigate the perpetuation of potentially less productive traditional practices, extension services were introduced, and it would have been interesting to see how access to extension services would affect the crop choice of the farmers (but I have not included this variable because I could not find a good proxy that worked, as explained in the method section). Nonetheless, many other studies have examined this and found that extension played an important role in the productivity of farmers (eg. Imai et al, 2015).

The amount of free agricultural inputs received was significant at the 99% confidence level in all the regressions run. This indicates that at the point where farmers make a decision on the crop to produce, there is a large scope to influence their decisions by the amount of free agricultural inputs they are given. The coefficient is also positive indicating a positive relationship. What this implies is that the more inputs received, the more likely the households would be to choose to produce tuber or root crops and the more export oriented crops. It might not be immediately obvious why this is the case, but my suspicion is that this relationship exists because some types of crops require a greater initial investment to get going and that these free inputs act as a buffer to reduce the costs (or risks) of planting those crops which they believe could be more profitable.

Table 7: Treatment Effects Model Results for the Selection of Crop equation and the impact of Crop Choice on Productivity (Technical Efficiency)

	C1 – Chose a more exported crop		C2 – Chose a tuber/root crop	
	Selection (1)	Impact (2)	Selection (3)	Impact (4)
Crop Choice		0.0014 (0.005)		0.045*** (0.004)
Age	0.01 (0.35)	0.0010* (0.0006)	0.019 (0.45)	0.001 (0.001)
Age Square	-0.022 (0.22)	-0.0000 (0.0000)	-0.022 (0.22)	-0.000 (0.000)
Education	2.02e-05 (1.81e-05)	-0.0136*** (0.0028)	2.02e-05 (1.01e-05)	-0.012*** (0.003)
HH Size	0.128* (0.008)	-0.808*** (0.280)	0.129* (0.007)	0.003*** (0.001)
Sex	0.233*** (0.054)	0.766*** (0.316)	0.235*** (0.054)	0.028*** (0.007)
Rural	-0.22 (0.34)	0.005 (0.004)	-0.22 (0.34)	0.007 (0.004)
Female Dependants	-7.55e-05 (0.00)	-0.002 (0.001)	-7.05e-05 (0.00)	-0.002* (0.001)
Married	0.118* (0.063)	0.000 (0.001)	0.118* (0.063)	-0.000 (0.001)
Region1 (NW)	0.167 (0.209)	-0.008* (0.004)	-0.181 (0.150)	-0.006 (0.004)
Region2 (NC)	1.074*** (0.187)	0.036*** (0.004)	1.557*** (0.123)	0.019*** (0.004)
Region3 (SW)	1.737*** (0.212)	0.003 (0.007)	1.738*** (0.161)	-0.011 (0.007)
Region4 (SE)	1.031*** (0.192)	-0.020*** (0.005)	2.284*** (0.132)	-0.049*** (0.005)
Region5 (SS)	2.207*** (0.193)	0.003 (0.006)	2.885*** (0.157)	-0.031*** (0.006)
Free Inputs#	0.677*** (0.023)		0.334*** (0.033)	
Non-farm income#	0.118* (0.0638)		0.11** (0.062)	
Previous year's seeds#	0.420* (0.10)		0.484*** (0.064)	
Constant	-2.384*** (0.495)	0.588*** (0.018)	-2.538*** (0.419)	0.592*** (0.017)
N	2422	2422	2422	2422
Time Dummies	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; # Exclusion restriction

Table 8: Treatment Effects Model Results for the Selection of Crop equation and the impact of Crop Choice on Poverty (log MPCE)

	C1 – Chose a more exported crop		C2 – Chose a tuber/root crop	
	Selection	Impact	Selection	Impact
	(1)	(2)	(3)	(4)
Crop Choice		-0.183**		-0.161***
		(0.066)		(0.022)
Age	-0.007	0.008	0.019	-0.001
	(0.019)	(0.007)	(0.45)	(0.003)
Age Square	0.000	-0.000	-0.022	-0.000
	(0.000)	(0.000)	(0.22)	(0.000)
Education	-0.036	0.067	2.02e-05	0.090***
	(0.095)	(0.037)	(1.01e-05)	(0.017)
HH Size	0.128*	0.152***	0.129*	0.079***
	(0.008)	(0.008)	(0.007)	(0.004)
Sex	0.233***	-0.300**	0.235***	-0.004
	(0.054)	(0.096)	(0.054)	(0.044)
Rural	-0.22	0.011	-0.22	-0.142***
	(0.34)	(0.057)	(0.34)	(0.026)
Female Share	-7.55e-05	-0.079***	-7.05e-05	0.009
	(0.00)	(0.016)	(0.00)	(0.008)
Married	0.118*	-0.085***	0.118*	-0.056***
	(0.063)	(0.017)	(0.063)	(0.008)
Region1 (NW)	0.560*	-0.118*	-0.181	-0.267***
	(0.270)	(0.052)	(0.150)	(0.024)
Region2 (NC)	1.266***	-0.221***	1.557***	0.060*
	(0.257)	(0.056)	(0.123)	(0.027)
Region3 (SW)	1.276***	-0.038	1.738***	0.019
	(0.289)	(0.087)	(0.161)	(0.041)
Region4 (SE)	1.277***	-0.239***	2.284***	-0.159***
	(0.263)	(0.061)	(0.132)	(0.032)
Region5 (SS)	2.471***	-0.087	2.885***	0.140***
	(0.263)	(0.080)	(0.157)	(0.039)
Free Inputs#	0.677***		0.334***	
	(0.023)		(0.033)	
Previous year's seeds#	0.420*		0.484***	
	(0.10)		(0.064)	
Constant	-2.706***	11.084***	-2.538***	12.069***
	(0.619)	(0.235)	(0.419)	(0.109)
N	2422	2422	2422	2422
Time Dummies	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; # Exclusion restriction

The major other significant determinants of crop choice are the regions in which the household resides, the size of the household and the gender of the household head. The regions are obviously important because some crops grow better in some areas than others, and the simple imposition of topological or geographic constraints could influence the determination of crop produced. It is worth noting that within each crop choice division are crops that are capable to be grown profitably anywhere in the country, although with a distribution of productivities. The size of the household being significant and negative appears to indicate that the larger a household is, the less likely they are to plant tubers, roots or exportable crops. This is possibly due to the fact, as mentioned earlier, that different crops would require different capital outlays and the head of a larger household may be more reluctant to put up this sum. The connotation this has to risk aversion of the household, as well as the other variables is discussed briefly in the corollary at the end.

On the impact of the choice on productivity, there is a mixed result. Using C1 as measure of crop choice shows no statistically significant effect to productivity at all, but C2 is significant. This result, I think is to be expected given the trend described earlier in Figure 3, which showed roots and tubers having dominance over cereals and fibres in productivity. However, the difference between the productivities of the farm households who engage in the more export oriented crops is not that different from the rest.

From Table 8 however, both C1 and C2 have a significant effect and are important in explaining the differences in the poverty outcomes of the two groups of farmers, but in a strange direction. Their coefficients are negative implying that the farmers who have grown these types of crops have poorer mean household expenditures on the average. The only explanation I could think of for why this might be the case is that cassava which is a crop that features in both C1 and C2 divisions is the raw material for a major staple food in Nigeria, and as such, a lot of the produce is consumed within the household itself. If this is the case, such self-consumption would not be reflected in the household expenditure variable, therefore underestimating the real valuation of the welfare situations of the two groups of farm households. In the next section, the commercialization index is examined to clarify the dichotomy between home use and marketing of produce.

6.3 Impact of crop choice and commercialization on productivity and poverty

In this section, the results of the fixed effects and correlated random effects models to estimate the impact of commercialization and its interactions with the categorical choice variables on productivity and poverty are reported in Tables 9 and 10 respectively. Columns 1 and 2 are results for both models with commercialization only; columns 3 and 4 are for the impact of commercializing the export oriented crop grown; and columns 5 and 6 are for the impact of commercializing tuber and root crops.

The results show that the household index of commercialization is not a significant determinant of productivity, but of poverty. This is bit surprising because one might expect that the more commercialized a farm household is, the better its productivity should be due to the monetary incentives in producing the most output possible with the lowest amount of inputs. These results show however, that the incentives to the household head of increasing productivity to keep his family fed are greater than the incentives from doing so for the sake of the possible monetary value of his goods. This is an interesting result with potentially far reaching policy implications. It means that if the government is interested in increasing productivity, food security should be prioritised instead of commercialization. This ties in with the previous story of the efficient small farmer as well (if we are to take the more commercialized farms as farms with larger farm land holdings, even though we know, as was noted in the definitions, that the commercialization index is not correlated with land size). It is possible that some inefficiencies arise as costs increase when workers have to be hired and supervised.

On the other hand, commercialization is an important determinant of poverty (significant at the 99% confidence level). Thus if poverty alleviating policy is on the agenda, commercialization would be a policy to push forward and implement. It is not clear however, how these two relationships come together. From the coefficients of the interactions, it appears they simply echo and amplify the effects of the commercialization variable.

Table 9: Results of Impact of Crop Commercialization with Crop Choice on Productivity

	FE	CRE	FE	CRE	FE	CRE
	(1)	(2)	(3)	(4)	(5)	(6)
C3 – Commercialization	-0.011 (0.057)	0.067 (0.037)				
C3*C1 – by export and commercialization			0.014 (0.05)	0.00844 (0.34)		
C3*C2 – by tuber/root crop and commercialization					0.035*** (0.004)	0.055*** (0.004)
Age	0.096*** (0.027)	0.008 (0.007)	0.096*** (0.027)	0.008 (0.007)	0.096*** (0.027)	0.008 (0.007)
Age Square	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Sex	-1.740 (1.025)	-0.300** (0.096)	-1.740 (1.025)	-0.300** (0.096)	-1.740 (1.025)	-0.300** (0.096)
Education	0.096 (0.095)	0.067 (0.037)	0.096 (0.095)	0.067 (0.037)	0.096 (0.095)	0.067 (0.037)
HH Size	0.747*** (0.045)	0.152*** (0.008)	0.747*** (0.045)	0.152*** (0.008)	0.747*** (0.045)	0.152*** (0.008)
Rural	0.01 (0.35)	0.019*** (0.01)	0.01 (0.35)	0.019*** (0.01)	0.01 (0.35)	0.019*** (0.01)
Female Share	-0.022 (0.22)	-0.050 (0.041)	-0.022 (0.22)	-0.050 (0.041)	-0.022 (0.22)	-0.050 (0.041)
Married	2.02e-05 (1.81e-05)	0.358 (0.041)	2.02e-05 (1.81e-05)	0.358 (0.041)	2.02e-05 (1.81e-05)	0.358 (0.041)
Region1 (NW)		-0.808*** (0.280)		-0.808*** (0.280)		-0.808*** (0.280)
Region2 (NC)		0.766*** (0.316)		0.766*** (0.316)		0.766*** (0.316)
Region3 (SW)		-0.001 (0.00)		-0.001 (0.00)		-0.001 (0.00)
Region4 (SE)		-0.299 (0.270)		-0.299 (0.270)		-0.299 (0.270)
Region5 (SS)		-0.087 (0.080)		-0.087 (0.080)		-0.087 (0.080)
Constant	10.23*** (0.326)	11.095*** (0.229)	10.23*** (0.326)	11.095*** (0.229)	10.23*** (0.326)	11.095*** (0.229)
N	2422	4844	2422	4844	2422	4844
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses and *** p<0.01, ** p<0.05, * p<0.1

Table 10: Results of the Impact of Crop Commercialization with Crop Choice on Poverty

	FE	CRE	FE	CRE	FE	CRE
	(1)	(2)	(3)	(4)	(5)	(6)
C3 – Commercialization	-0.142***	-0.056***				
	(0.026)	(0.008)				
C3*C1 – by export and commercialization			0.019*	0.0178***		
			(0.00766)	(0.006)		
C3*C2 – by tuber/root crop and commercialization					-0.095*	-0.161***
					(-0.021)	(0.022)
Age	0.096***	-0.001	0.096***	-0.001	0.096***	-0.001
	(0.027)	(0.003)	(0.027)	(0.003)	(0.027)	(0.003)
Age Square	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Sex	-1.740	0.090***	-1.740	0.090***	-1.740	0.090***
	(1.025)	(0.017)	(1.025)	(0.017)	(1.025)	(0.017)
Education	0.096	0.079***	0.096	0.079***	0.096	0.079***
	(0.095)	(0.004)	(0.095)	(0.004)	(0.095)	(0.004)
HH Size	0.747***	-0.004	0.747***	-0.004	0.747***	-0.004
	(0.045)	(0.044)	(0.045)	(0.044)	(0.045)	(0.044)
Rural	0.01	-0.142***	0.01	-0.142***	0.01	-0.142***
	(0.35)	(0.026)	(0.35)	(0.026)	(0.35)	(0.026)
Female Share	-0.022	0.009	-0.022	0.009	-0.022	0.009
	(0.22)	(0.008)	(0.22)	(0.008)	(0.22)	(0.008)
Married	2.02e-05	-0.056***	2.02e-05	-0.056***	2.02e-05	-0.056***
	(1.81e-05)	(0.008)	(1.81e-05)	(0.008)	(1.81e-05)	(0.008)
Region1 (NW)		-0.267***		-0.267***		-0.267***
		(0.024)		(0.024)		(0.024)
Region2 (NC)		0.060*		0.060*		0.060*
		(0.027)		(0.027)		(0.027)
Region3 (SW)		0.019		0.019		0.019
		(0.041)		(0.041)		(0.041)
Region4 (SE)		-0.159***		-0.159***		-0.159***
		(0.032)		(0.032)		(0.032)
Region5 (SS)		0.140***		0.140***		0.140***
		(0.039)		(0.039)		(0.039)
Constant	5.198***	11.095***	5.198***	11.095***	5.198***	11.095***
	(1.233)	(0.229)	(1.233)	(0.229)	(1.233)	(0.229)
N	2422	4844	2422	4844	2422	4844
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors in parentheses and *** p<0.01, ** p<0.05, * p<0.1

7 Extension: Distributional Effects of Choice by Quantile Regressions

7.1 Model

As an extension to the previous analysis, quantile regressions are also done for a pooled sample of both waves of data to see the effects of crop choice on different quantiles of the poverty (MPCE) distribution, conditional on the control variables. The reason why this analysis is relevant is that it allows for a much richer characterization and description of what is actually going on in the data and can show if there are different effects of crop choice across the spectrum, and what nature these effects are. In addition, there is some flexibility here for modelling the data with heterogeneous conditional distributions; this would therefore produce a median regression (50^{th} quantile) that is often more robust to outliers.

The quantile regressions are described by the following equation:

$$y_i = \boldsymbol{x}_i'\beta_q + \varepsilon_i \tag{13}$$

Where β_q is the vector of unknown parameters (coefficients) associated with the q^{th} quantile, y_i is the mean per capita household expenditure (poverty variable), $\boldsymbol{x}_i$ are the explanatory variables including the crop choice variables and ε_i is the stochastic error term.

The quantile regression minimizes $\sum_i q|\varepsilon_i| + \sum_i (1-q)|\varepsilon_i|$ [10], a sum that gives the asymmetric penalties $q|\varepsilon_i|$ for underprediction and $(1-q)|\varepsilon_i|$ for overprediction.

The q^{th} quantile regression estimator, $\widehat{\beta_q}$ minimizes over β_q the objective function:

$$Q(\beta_q) = \sum_{i:y_i \geq \boldsymbol{x}_i'\beta_q}^{n} q|y_i - \boldsymbol{x}_i'\beta_q| + \sum_{i:y_i < \boldsymbol{x}_i'\beta_q}^{n} (1-q)|y_i - \boldsymbol{x}_i'\beta_q| \tag{14}$$

where $0 < q < 1$

The standard conditional quantile is assumed to be linear and for the j^{th} regressor, the marginal effect is the coefficient for the q^{th} quantile:

$$\frac{\partial Q_q(y|x)}{\partial x_j} = \beta_{qj} \tag{15}$$

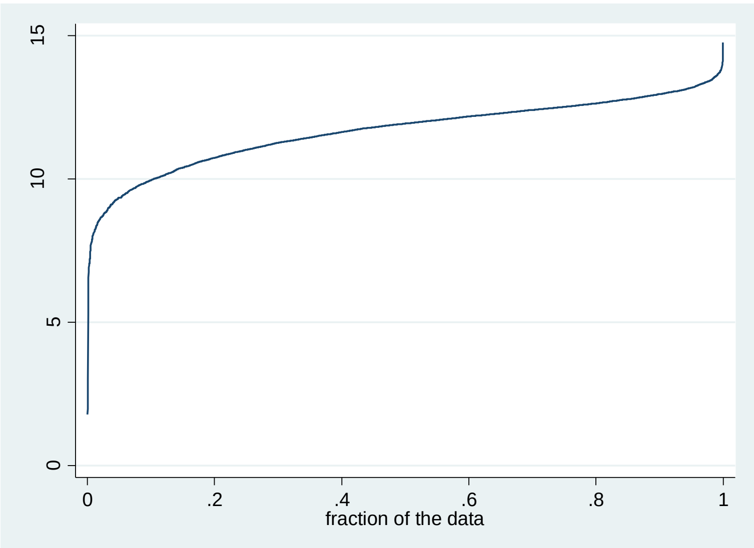
[10] As opposed to OLS, which minimizes: $\sum_i {\varepsilon_i}^2$ (sum of squares of model prediction).

In this way, we can interpret the coefficient or quantile regression parameter β_{qj} estimates as the change in a specific quantile q of the dependent variable y produced by a one unit change in the independent variable x_j. As is normal in quantile regressions, there are two kinds of significance of importance. The first being if the effect on each quantile value is significantly different from zero and the second, to check if they are different from a normal OLS, which would indicate that there are differences in the effect to different segments of the population. The three quantiles reported in the table are the 20th, the 50th and the 75th.

7.2 Results

Figure 4 is a plot showing the MPCE across different fractions of the sample. It can be seen that the bottom 20% have the worst poverty outcomes, with most of the rest around the average, while the top 10% have slightly better outcomes. This picture motivates the need to study the peculiarities of those households which fall within different quantiles of the mean expenditure distribution.

Figure 4: Quantile plot of MPCE



From the four regressions run on the pooled sample with both the C1 and C2 crop choice variables, what can be seen (from table 11) below, is that there is not a significant difference between the coefficients from the quantile regressions and the OLS estimates for crop choice. This result can also be inferred from Figure 8, a plot showing the estimated marginal effects of the different variables on the conditional quantiles of MPCE. It can be seen that the 95% confidence intervals band around the quantile function overlaps mostly with the 95% confidence interval band for the OLS regression, except for a little bit from around the 80th quantile.

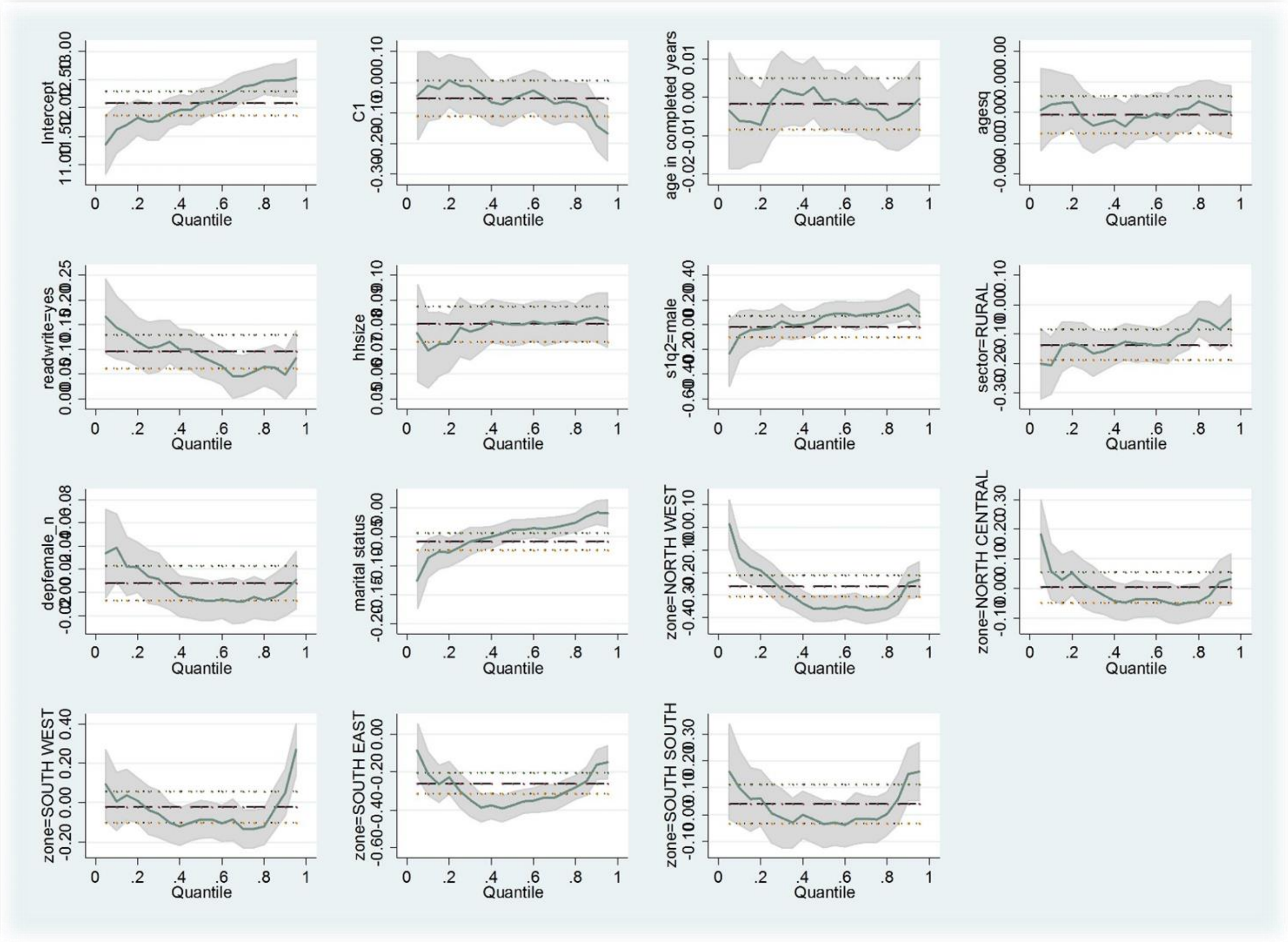
Table 11: Cross-sectional quantile estimation results for pooled cross-section

	C1 – Chose a more exported crop				C2 – Chose a tuber/root crop			
	Quantile Regressions				Quantile Regressions			
	0.20	0.5	0.90	OLS	0.20	0.5	0.90	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Crop Choice	0.005 (0.043)	-0.055 (0.036)	-0.061 (-0.034)	-0.052 (0.030)	-0.163*** (0.037)	-0.176*** (0.026)	-0.128*** (0.027)	-0.161*** (0.022)
Age	-0.007 (0.005)	-0.001 (0.004)	-0.003 (0.004)	-0.002 (0.003)	-0.007 (0.006)	0.000 (0.004)	-0.002 (0.004)	-0.001 (0.003)
Age Square	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)
Education	0.115*** (0.025)	0.085*** (0.020)	0.056** (0.019)	0.096*** (0.017)	0.097*** (0.028)	0.080*** (0.020)	0.057** (0.021)	0.090*** (0.017)
HH Size	0.072*** (0.005)	0.080*** (0.004)	0.081*** (0.004)	0.080*** (0.004)	0.074*** (0.006)	0.080*** (0.004)	0.081*** (0.004)	0.079*** (0.004)
Sex	-0.034 (0.063)	0.068 (0.052)	0.089 (0.050)	-0.015 (0.044)	-0.000 (0.072)	0.062 (0.050)	0.120* (0.053)	-0.004 (0.044)
Rural	-0.132*** (0.038)	-0.132*** (0.031)	-0.092** (0.029)	-0.136*** (0.026)	-0.161*** (0.043)	-0.125*** (0.030)	-0.098** (0.032)	-0.142*** (0.026)
Female Share	0.022* (0.011)	-0.007 (0.009)	-0.004 (0.008)	0.008 (0.008)	0.023 (0.012)	-0.005 (0.009)	-0.007 (0.009)	0.009 (0.008)
Married	-0.078*** (0.011)	-0.038*** (0.009)	-0.030*** (0.009)	-0.058*** (0.008)	-0.071*** (0.012)	-0.040*** (0.009)	-0.028** (0.009)	-0.056*** (0.008)
Region1 (NW)	-0.192*** (0.034)	-0.359*** (0.028)	-0.365*** (0.027)	-0.260*** (0.024)	-0.196*** (0.039)	-0.367*** (0.027)	-0.005 (0.011)	-0.267*** (0.024)
Region2 (NC)	0.052 (0.038)	-0.037 (0.031)	-0.049 (0.029)	0.004 (0.026)	0.117** (0.045)	0.009 (0.031)	-0.269*** (0.036)	0.060* (0.027)
Region3 (SW)	0.011 (0.058)	-0.085 (0.048)	-0.134** (0.046)	-0.022 (0.041)	0.084 (0.067)	-0.046 (0.047)	0.013 (0.041)	0.019 (0.041)
Region4 (SE)	-0.229*** (0.041)	-0.374*** (0.034)	-0.307*** (0.032)	-0.259*** (0.029)	-0.110* (0.052)	-0.268*** (0.036)	0.074 (0.061)	-0.159*** (0.032)
Region5 (SS)	0.061 (0.053)	-0.036 (0.044)	-0.017 (0.042)	0.040 (0.037)	0.183** (0.063)	0.083 (0.044)	0.154** (0.057)	0.140*** (0.039)
Constant	11.832*** (0.155)	12.081*** (0.128)	12.404*** (0.122)	12.083*** (0.109)	11.818*** (0.178)	12.073*** (0.124)	12.467*** (0.160)	12.069*** (0.109)
N	4,786	4,786	4,786	4,786	4,786	4,786	4,786	4,786

Note: Robust standard errors in parentheses and *** p<0.01, ** p<0.05, * p<0.1

What is more interesting however, are the conditional marginal effects of some of the other variables on the quantiles of MPCE. Education has a much higher significant effect on households in the bottom of the expenditure distribution than in the OLS case, and this can also be seen from the marginal plots diagram where the confidence interval bands for the lower quantiles is higher than the normal. This is an interesting result for proponents of education as a valid tool to raise the poorest households out of poverty the quickest. The actual quantile function line for household size also appears to be different from the OLS, but the confidence boundaries at that end of the distribution are higher, leading to substantial overlap. Finally, marital status is another variable of interest because the quantile regression shows that the head of household being married is more important to the welfare situation of that household, the poorer that household is. This could explain why in many poorer countries, marriages are contracted as a form of social security (Charsley and Liversage, 2013). Thus, finding a way to raise incomes might be a viable option to consider when thinking about tackling the problem of forced marriages and child brides.

Figure 8: Estimated marginal effects on the conditional quantiles for MPCE with C1



8 Corollary: Crop Choice and Risk Aversion

As a corollary to the discussion of results above, careful study of these results reveals a pattern that appears to emerge. This pattern is regarding the seeming similarity between the reasons why a farm household would choose to grow one crop instead of another and that household's vulnerability to risk. It has been shown extensively in the literature that vulnerability to risk is a dominant feature of the poor's livelihood, and this is particularly so for small farmers in sub-Saharan Africa

(Fafchamps, 2009). This is because when there are shocks to the production function, it could have a ripple effect on their incomes, assets, and hence the health and education of the next generation; especially as they have so little to begin with, and thus a lower margin for error.

So the point here that is hypothesised is that the household's desire to mitigate risks and protect themselves from adverse shocks could affect their production decisions. When risk aversion is to be computed for farmers (as in Binswanger, 1980 for his ICRISAT data project), it is normally done by applying psychological and game techniques via field experiments. This would involve the farmers making a choice between lotteries of sorts and this information would then be aggregated to produce an index of risk aversion. Similarly, the choice of which crop to produce could be likened to a range of lotteries each with their own distribution of expected returns. And the farmer would then choose the lottery that gives him the highest anticipated earnings. A similar idea is known and has been used in application for the adoption of new or modern agricultural technology, but to my knowledge, has not been expressly applied to the type of crop a farmer chooses to grow.

The channel through which this works is that farmers who are fearful of the uncertain return of growing a new (different) type of crop, might just be content to keep growing the crop they have always grown, simply because they know from experience what the output they would get at the end of the growing season is likely to be. And this information gets reinforced from year to year to such a point that it would take a great effort indeed to break the aversion to try something new. This is as opposed to the prospect of "shooting in the dark" and expecting the best, even though they may have heard on radio or been visited by extension agents who have tried to convince them that there is a crop they could grow which would be more productive.

Whilst this theory might seem logical, it would be more helpful if evidence for this can be provided. At the moment, I present mostly evidence from the literature and some anecdotal evidence, but intend to test the hypothesis with data at a later stage.

To start with, it is important to differentiate agricultural shocks from risk (even though some papers use them interchangeably). Following the book by Fafchamps (2003), he states that shocks could affect welfare and behaviour because they are often times unanticipated and as such, suitable precaution could not be taken against it. These could include severe weather disruptions like floods or droughts or unexpected influx of pests. To me, this is very similar to the idea of uncertainty; where the distribution of outcomes is unknown at the start of decision making. In contrast, although risk also involves unknown outcomes, the distribution of these outcomes can be predicted ahead of time. In other words, when people understand a shock is more likely to occur, any option that amplifies this likelihood becomes a more risky prospect. This difference is subtle but important because people can only adjust their behaviour ahead of time in response to risks and not shocks, for which they would have to respond after the fact.

There is surprisingly little research on the direct effects of actual risk attitudes on farm household behaviour. Most of the literature use shocks of different kinds as a proxy for risk because the effects are relatively easier to demonstrate econometrically than an index of risk aversion, which is not a concept that can be measured with completely assurance. For example, Portner (2008), Alderman et al. (2006) show the effects of weather shocks on agricultural yields and on nutrition and height of children respectively. Kurosaki and Fafchamps (2002) use a survey from Pakistani dairy farmers to show that the crops they plant are consistent with their desire to cover some of the feeding requirements of their cattle and hence reduce their risk exposure.

Studies from behavioural economics, that do use measures of risk aversion, often find that as males, who have more family burden and are less educated tend to have higher levels of risk aversion as opposed to the female, younger, single, and more educated individuals. It turns out that these characteristics, represented by gender, household size and education, are also significant determinants of the crops that farmers chose to grow. If it is the case that there is actually correlation between crop choice and risk aversion as we suspect, it might be possible to simply use the type of crop a farmer has grown at the start of the growing season as a proxy for his risk appetite instead of conducting expensive and often misleading field experiments to measure risk aversion directly. Also, this would be making use of a decision the farm household head had already made at the start of the season, rather than a spur of the moment answer to a set of questions in an interview or lottery game.

9 Conclusion

This intention of this research was to examine the arguments on whether or not smallholder farmers in Nigeria who produce certain types of crops (export oriented crops and roots and tubers) experience any significant productivity and welfare differences to those who do not and to examine the factors which determine the crop choices of these farmers. Using the two-rounds of LSMS panel data from Nigeria in 2010/11 and 2012/13; I started by re-examining the old arguments surrounding whether small-holder farmers are indeed "efficient-but-poor". I found that although smallholders were generally efficient in their allocation of resources (after estimating household crop productivity by stochastic frontier analysis), they were not necessarily rational in their crop choices because even when some crops are found to be more productive than the others, the less productive crop was often chosen. To figure out why, a treatments effect model was employed to determine farmer selection into the choice of a type of crop in the first stage; and subsequently the impact of their choices on productivity and poverty in the second stage. I found that access to free inputs, non-farm income, the use of seeds from the previous growing season, household size, gender and the different regional differences were the main determinants of crop choice. Also, the choice influenced the productivity and poverty of the households in different ways. While the choice of tuber and root crop improved productivity, they were found to reduce poverty outcomes, though this was likely due to constraints in how the poverty measure was computed. In addition, commercialization was found to be important for poverty alleviation, but not for productivity improvements.

There may not be enough evidence to suggest that all smallholders should switch from producing one type of crop to another, but the implications of crop choices by smallholder farmers having the effects reported could be that these different types of farm households may need to be targeted differently in terms of social welfare or aid. And depending on the poverty alleviation strategy of the government, it is possible that it could require far less effort to lift these groups of farmers out of poverty. Also, agricultural extension could be utilised to get more people within areas of comparative advantage to switch to these high production crops to improve their welfare outcomes.

References

Abler, D. G., and Sukhatme, V. A. (2006), "The "efficient but poor" hypothesis", *Rev. Agric. Econ.* 28, 338–343.

Adams, J. (1986), "Peasant rationality: individuals, groups, cultures." *World Development.* 14, 273–282.

Addison, T., and Tarp, F. (2015), "Lessons for Japanese foreign aid from research on aid's impact." *Japan's Development Assistance: Foreign Aid and the Post-2015 Agenda*, 295.

Adeyemo, R., Oke J. T. O., and Akinola A. A. (2010), "Economic Efficiency of Small Scale Farmers In Ogun State, Nigeria." *Tropicultura* 28(2).

Aigner, D., Lovell, C. A., and Schmidt, P. (1977). "Formulation and estimation of stochastic frontier production function models." *Journal of econometrics*, 6 (1), 21-37.

Ajibefun, I. A. (1998), "Efficiency of Small Scale Food Crop Farms In Ondo State, Nigeria. An Application of Parametric And Non-Parametric Frontier Production Functions."Department of Agricultural Economics and Extension, FUTA, Akure.

Amaza, P. S., Kamara A. Y. and Helsen J. (2005), "Determinants of food crop production and technical efficiency in Guinea Savannahs of Borno State, Nigeria". *African crop science proceedings* 7.

Amos, T. T., Chikwendu D. O. and Nmadu J. N. (2004), "Productivity, Technical efficiency and cropping patterns in the savanna zone of Nigeria." *J. Food Agric. Environ.* 2: 173-176.

Ayanwale, A. B., and Alimi, T. (2004), "The impact of the National Fadama facility in alleviating rural poverty and enhancing agricultural development in south-western Nigeria." *Journal of social sciences,* 9(3), 157-161.

Azariadis, C., and Stachurski, J. (2005). "Poverty Traps." Chap. 5 of *Handbook of Economic Growth*, Vol. 1, Part A, edited by Philippe Aghion and Steven N. Durlauf. Elsevier.

Banerjee, A. V., and Duflo E. (2011), "Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty." New York, NY: Public Affairs.

Banerjee, A., and Mullainathan, S. (2010), "The Shape of Temptation: Implications for the Economic Lives of the Poor." *NBER Working Paper* 15973.

Ball, R., and Pounder, L. (1996), "Efficient but poor revisited." *Economic Development Culture and Change* 44, 735–760.

Barrett, C.B., Bellemare, M.F., and Hou, J.Y., (2009). "Reconsidering conventional explanations of the inverse productivity-size relationship." *World Development* 38, 88–97.

Barrett, C. B., Marenya, P. P., McPeak, J., Minten, B., Murithi, F., Oluoch-Kosura, W., Place, F., Randrianarisoa, J. C., Rasambainarivo, J., and Wangila, J. (2006), "Welfare Dynamics in Rural Kenya and Madagascar." *Journal of Development Studies*, 42(2): 248-77.

Binswanger, H. P. (1980), "Attitudes toward risk: Experimental measurement in rural India." *American journal of agricultural economics*, *62*(3), 395-407.

Bogetoft, P. and Otto, L. (2011). Benchmarking with DEA, SFA, and R, volume 157. Springer.

Bravo-Ureta, B. and Pinheiro E. (1997), "Technical, Economic, And Allocative Efficiency In Peasant Farming: Evidence From The Dominican Republic." *The Developing Economies,* XXXV-1: 48–67.

Chakravorty, S. (1984), "Power structure and agricultural productivity." In: Desai, M., Rudolph, S., Rudra, A. (Eds.), Agrarian Power and Agricultural Productivity in South Asia. Oxford University Press, Delhi.

Charsley, K., and Liversage, A. (2013). "Transforming polygamy: migration, transnationalism and multiple marriages among Muslim minorities." *Global Networks*, *13*(1), 60-78.

Charnes, A., Cooper, W. W., and Rhodes, E. (1978). "Measuring the efficiency of decision making units." *European journal of operational research*, 2 (6), 429-444.

Chavas, J., and Aliber M. (1993), "An analysis of economic efficiency in Agriculture: A non-parametric approach." *Journal of agricultural and resource economics*, 8(1):1-16.

Chayanov, A.V. (1926), in: D. Thorner et al., eds., The theory of peasant economy (Irwin, Homewood, IL).

Ciaian, P., Falkowski J., and Kancs D., (2012), "Access to Credit, Factor Allocation and Farm Productivity: Evidence from the CEE Transition Economies." *Agricultural Finance Review* 72(1), 22-47.

Clemens, M. (2012), "New Documents Reveal the Cost of 'Ending Poverty' in a Millennium Village: At Least $12,000 per Household." *Center for Global Development.*

Coelli, T. J. and Battese, G. E. (1996). "Identification of factors which influence the technical inefficiency of indian farmers." *Australian Journal of Agricultural and Resource Economics,* 40 (2), 103-128.

Coelli, T. J., Rao, D. P., O'Donnell, C. J., and Battese, G. E. (2005). An introduction to efficiency and productivity analysis. Springer.

Cuong, N. V. (2009), "Measuring the Impact of Cash Crops on Household Expenditure and Poverty in Rural Vietnam", *Asia Pacific Trade and Investment Review*, United Nations Economic and Social Commission for Asia and the Pacific (ESCAP), 16(2), 87-112.

Dasgupta, P., and Ray D. (1986). "Inequality as a Determinant of Malnutrition and Unemployment: Theory." *Economic Journal* 96(384): 1011-34.

Delarue, J., Mesple-somps, S., Naudet, J., and Robilliard, A. (2009), "The Sikasso Paradoxe: Cotton and Poverty in Mali", Development Institutions and Analysis Working Paper. [Accessed at: www.dial.prd.fr/dial˙publications/PDF/Doc˙travail/2009-09GB.pdf]

Dercon, S. and L. Christiaensen, "Consumption risk, technology adoption and poverty traps: evidence from Ethiopia," Journal of Development Economics, 2011.

Dixon, J. A., Gibbon, D. P., & Gulliver, A. (2001), "Farming systems and poverty: improving farmers' livelihoods in a changing world." Food & Agriculture Org..

Duflo, E., (2006), "Poor but rational?" In: Banerjee, A., Benabou, R., Mookherjee, D. (Eds.), Understanding Poverty. Oxford University Press, Oxford, pp. 367–378.

Ebong, V. O., Okoro U. S. and Effiong E. O. (2009), "Determinants of Technical Efficiency Of Urban Farming In Uyo Metropolis Of Akwa Ibom State, Nigeria." *Journal of Agriculture and Social Science* 5: 89–92.

Eswaran, M., and Kotwal, A. (1986). "Access to capital and agrarian production organisation." *The Economic Journal*, *96*(382), 482-498.

Farrell, M. (1957), "The Measurement of Productive Efficiency." *Journal Of The Royal Statistical Society*, 120(3): 253–290.

Fasoranti, M. M. (2006), "A Stochastic Frontier Analysis of Effectiveness Of Cassava-Based Cropping Systems In Ondo State, Nigeria." Phd Thesis, Department of Agricultural Economics and Extension, FUTA, Akure.

Gaiha, R., Azam, M.S., Annim, S., and Imai, K. (2013), "Agriculture, Markets and Poverty - A Comparative Analysis of Laos and Cambodia," *Asian Journal of Agriculture and Development,* 9(1):97-111.

Govereh, J., Jayne, T. S., and Nyoro, J. (1999), "Smallholder commercialization, interlinked markets and food crop productivity: Cross-country evidence in eastern and southern Africa." *Michigan State University, Department of Agricultural Economics and Department of Economics*, *39*.

Greene, W. H. (2008), Econometric Analysis. 6th ed. Upper Saddle River, NJ: Prentice Hall.

Gulati, A., Minot N., C. Delgado C. and Bora S. (2006), "Growth in high-value agriculture in Asia and the emergence of vertical links with farmers". In Swinnen, J. (ed.) *Global supply chains, standards, and poor farmers*. London: CABI Press.

Hansen, L. P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, 50, 1029-1054.

Hayami, Y. and Ruttan, V. W., (1985), "Agricultural Development: an International Perspective." Johns Hopkins University Press, Baltimore, MD.

Heckman, J., R. LaLonde, and J. Smith (1999), "The Economics and Econometrics of Active Labor Market Programs," in *Handbook of Labor Economics Vol.III*, ed. by O. Ashenfelter, and D. Card, pp. 1865–2097. Elsevier, Amsterdam.

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

Henderson, D. (2003), *"The Measurement Of Technical Efficiency Using Panel Data."* Mimeo, State University Of New York At Binghamton, New York.

Imai, K. S., Hasan, M. F., and Porreca, E. (2015), "Do Agricultural Extension Programmes Reduce Poverty and Vulnerability? Farm size, Agricultural Productivity and Poverty in Uganda", *Discussion Paper Series*, Rieb Kobe University, DP2015-06.

Iwuchukwu, J. C., and Igbokwe, E. M. (2012), "Lessons from agricultural policies and programmes in Nigeria." *Journal of Law, Policy and Globalization*, *5*, 11-21.

Kaplinsky, R. and Morris M. (2002), "A Handbook for Value Chain Research." IDRC. [Accessed at: http://www.ids.ac.uk/ids/ global/pdfs/VchNov01.pdf]

Karlan, D., Osei, R., Osei-Akoto, I., and Chris Udry, (2013), "Agricultural Decisions after Relaxing Credit and Risk Constraints", *NBER working papers* No.18463, National Bureau of Economic Research, Inc, Dec.

Kazianga, H., and Udry, C. (2006), "Consumption Smoothing? Livestock, Insurance and Drought in Rural Burkina Faso." Journal of Development Economics, 79(2), 413–446. doi:10.1016/j.jdeveco.2006.01.011.

Kraay, A., and McKenzie, D. (2014) "Do Poverty Traps Exist?" *Policy Research Working Paper 6835*. Washington, DC: World Bank.

Loveridge, S., Nyarwaya, J. B., and Shingiro, E. (2003), "Decaffeinated? Situation, Trends and Prospects for Smallholder Coffee Production in Rwanda", Agricultural Policy Synthesis No. 5E. [Accessed at: http://www.aec.msu.edu/fs2/rwanda /ps˙5e.pdf]

Maertens, M. and Swinnen, J. (2009), "Trade, Standards, and Poverty: Evidence from Senegal", *World Development*, 37(1), 161-178.

Mazumdar, D. (1959), "The Marginal Productivity Theory of Wages and Disguised Unemployment." *Review of Economic Studies* 26(3): 190-97.

Meeusen, W. and van den Broeck, J. (1977). "Efficiency estimation from cobb-douglas production functions with composed error." *International economic review*, 18 (2), 435-444.

Murekezi, A., and Loveridge, S. (2009), "Have coffee reforms and coffee supply chains affected farmers' income? The case of coffee growers in Rwanda", in Agricultural and Applied Economics Association's 2009 AAEA and ACCI Joint Annual Meeting, Milwaukee, WI, July 26-28, 2009.

Murillo-Zamorano, L. (2004), "Economic Efficiency and Frontier Techniques." *Journal of Economic Surveys* 18(1): 33–77.

Nerlove, M. (1999), "Transforming economics: theodore W. Schultz, 1902–1998: In Memoriam." *Economic Journal*. 109, 26–48.

Otitoju, M. and Arene C. J. (2010), "Constraints and Determinants of Technical Efficiency in Medium-Scale Soybean Production in Benue State, Nigeria." *African Journal of Agricultural Research* 5(17): 2276–2280.

Ojo, M. A., Mohamed U. S., Ojo A. O., and Olaleye R. S. (2009), "Return To Scale And Determinants Of Farm Level Technical Inefficiency Among Small Scale Yam Based Farmers In Niger State Nigeria: Implications For Food Security." *International Journal of Agricultural Economics and Rural Development*.

Onyenweaku, C.E. and Ohajianya D.O. (2009), "Technical Efficiency of Rice Farmers In Ebonyi State, Nigeria." *Journal of Agriculture and Food Sciences* 3 (2): 175–184.

Ravallion, M. (1997), "Famines and Economics." *Journal of Economic Literature*, 35(3): 1205-42.

Ray, B., (2006), "Aspirations, poverty, and economic change." In: Banerjee, A., Benabou, R., Mookherjee, D. (Eds.), Understanding Poverty. Oxford University Press, Oxford, pp. 409–421.

Rivers, D. and Vuong, Q. (1988), "Limited information estimators and exogeneity tests for simultaneous probit models", *Journal of Econometrics*, 39, 347–366.

Roy, A. (1951), "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135–145.

Rubin, D. (1974), "Estimating Causal Effects to Treatments in Randomised and Nonrandomised Studies," *Journal of Educational Psychology*, 66, 688–701.

Ruttan, V. W. (2003), "Social Science Knowledge and Economic Development: an Institutional Design Perspective." University of Michigan Press, Ann Arbor, MI.

Schultz, T. W. (1964), Transforming traditional agriculture. *Transforming traditional agriculture.*

Shapiro, K. H. (1983), "Efficiency Differentials in Peasant Agriculture and Their Implication for Development Policies." *Journal of Development Studies.* 19, 179–190.

Singh, S., Coelli, T., and Fleming, E. (2001). "Performance of dairy plants in the cooperative and private sectors in India." *Annals of Public and Cooperative Economics,* 72 (4), 453-479.

Skees, J. R., Gober, S., Varangis, P., Lester, R., & Kalavakonda, V. (2001), "Developing rainfall-based index insurance in Morocco", *The World Bank*, Policy Research Working Paper 2577, April 2001.

Stiglitz, J. E. (1989), "Rational peasants, efficient institutions, and a theory of rural organization: methodological remarks for development economics." In: Bardhan, P. (Ed.), The Economic Theory of Agrarian Institutions. Oxford University Press, New York.

Street, A.. (2003). "How Much Confidence Should we Place in Efficiency Estimates?" *Health Economics*, 12(11): 895-907.

UNCTAD (2006), "Economic Development in Africa: Doubling Aid: Making the 'Big Push' Work." *United Nations Conference on Trade and Development*, Geneva.

Von Braun, J. (1995). "Agricultural commercialization: impacts on income and nutrition and implications for policy." *Food policy*, *20*(3), 187-202.

Von Braun, J., Bouis, H. and Kennedy, E., (1994). "Conceptual Framework", In: Agricultural Commercialization, Economic Development and Nutrition, Chapter 2, Von Braun and Kennedy (eds), The Johns Hopkins University Press, London.

Wang, H., Dong, X., Rozelle, S., Huang, J., and Reardon, T. (2009), "Producing and Procuring Horticultural Crops with Chinese Characteristics: The Case of Northern China", *World Development*, 37(11), 1791-1801.

Webber, C. and Labaste, P. (2010), "Building competitiveness in Africa's Agriculture: A guide to value chain concepts and applications", World Bank. Washington, D.C.

Wooldridge, J. M. (2012), "Quasi-Maximum Likelihood Estimation and Testing for Nonlinear Models with Endogenous Explanatory Variables", Department of Economics, Michigan State University, June 2012.

World Bank. (2016). Data retrieved February 10, 2016, from World Development Indicators Online (WDI) database.