Discussion Paper Series

RIEB

Kobe University

DP2021-17

Are Farmers "Efficient but Poor"? The Impact of Crop Choices on Agricultural Productivity and Poverty in Nigeria

Chisom UBABUKOH
Katsushi S. IMAI

Revised June 10, 2022

Research Institute for Economics and Business Administration
Kobe University
2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

Are Farmers "Efficient but Poor"? The Impact of Crop Choices on Technical Efficiency and Poverty in Nigeria

Chisom Ubabukoh†€ and Katsushi S. Imai †⁋

† The University of Manchester, UK

€ O. P. Jindal Global University, India

⁋ RIEB, Kobe University, Japan

This Draft: 15th September 2021

Revised on29th April 2022

Abstract

This paper aims to test the "efficient-but-poor" hypothesis" by estimating the determinants of smallholders' choice over cash or food crops and whether their crop choice affects technical efficiency and poverty using the national household panel data in Nigeria. As the crop choice is endogenous in the sense that the farmers' crop choice is also influenced by resulting revenue from the crop, we carry out stochastic frontier analyses with the Greene (2010) correction for sample selection about farmers' crop choice and find that smallholders are generally efficient in their resource allocations. A treatment effects model is employed to estimate farmers' crop choice in the first stage and the impact of their choices on technical efficiency and poverty outcomes in the second. The results show that farmers' access to free inputs, non-farm income and the use of seeds from the previous growing season are important determinants of crop choice. The adoption of cash crops by food-crop producing households will not generally reduce poverty, although it will improve technical efficiency marginally. However, if cash crops are commercialised, poverty tends to decline.

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

Keywords: Technical Efficiency, Poverty, Crop Choice, Stochastic Frontier Analysis, Treatment Effects Model, Nigeria

Corresponding Author: Katsushi S. Imai (Dr), Department of Economics, School of Social Sciences, The University of Manchester, Arthur Lewis Building, Oxford Road, Manchester M13 9PL, UK, Email: Katsushi.Imai@manchester.ac.uk.

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

Acknowledgments: We wish to thank Kunal Sen for his valuable comments and insights and thank William H. Greene for his advice on the use of the NLogit (software) and on the interpretation of some of the results presented in the paper. We would also like to thank the participants in the Royal Economic Society (RES) annual meeting, the Nordic Conference in Development Economics and several other seminars at Manchester for all their insightful comments and suggestions on the earlier version of the paper. Any mistakes are, however, ours.

Are Farmers "Efficient but Poor"? The Impact of Crop Choices on Technical Efficiency and Poverty in Nigeria

1. Introduction

The main purpose of this study is to test the "efficient-but-poor" hypothesis[1] by estimating the determinants of smallholders' crop choices and whether their 'endogenous' crop choices affect technical efficiency and consumption poverty. We focus on farmers' choice between cash crops and food crops where the former is defined based on the crop's exportability. Separately, we also analyse the effect of the extent to which cash or food crops are commercialised on technical efficiency and consumption poverty.

The challenge in estimating the effect of the crop choice on technical efficiency is that the former is endogenous in the sense that the farmers' crop choice is also influenced by the resulting revenue from the crop. To address this issue, we have carried out stochastic frontier analyses (SFA) by using the Greene (2010) correction for sample selection in estimating farmers' technical efficiency. The study is based on the household panel data constructed from two waves of Nigeria's General Household Survey-Panel, which is part of the World Bank's Living Standards Measurement Study. This is to our knowledge the first application of SFA with the Greene (2010) correction to Nigeria and one of the few applications to the agricultural productivity of households in developing countries.[2]

Producing cash crops was traditionally regarded as the forte of large-scale commercial farmers. However, there has been an argument in recent years that smallholder farmers could also take advantage of the large international market of their products while they attempt to raise overall productivity and improve their income from farming. We propose to examine this argument in greater detail by asking the research questions - "Have smallholder farmers who chose to grow a specific type of crops, such as cash crops with a higher degree of exportability

[1] "The poor but efficient hypothesis" - which is sometimes called "theory", "proposition", or "argument" by different scholars - was put forward by the Nobel laureate, Theodore Schultz (Schultz, 1964). It implies that farmers engaged in traditional agriculture are often poor with only a small area of land – either rented or owned – given the monopolistic and collusive land market in developing countries and they cannot easily get out of poverty due to both difficulties in undertaking the new investment as well as the low rate of returns to agricultural investment (e.g., Lundahl, 1987). However, Schultz "hypothesized" that smallholders in traditional agriculture are highly efficient in terms of their recourse allocations contrary to the previously-held view that they are constrained by tradition or culture (Abler and Sukhatme, 2006). Our finding that smallholders are more technically efficient than large farmers is consistent with this view as we will discuss later.

[2] They include Rahman (2011) and Martey et al. (2019).

improved their technical efficiency and reduced poverty?" and "How did commercialisation of each type of crops - cash or food crops - influence technical efficiency and poverty". In answering these questions, we also explore the underlying reasons for choosing to grow specific types of crops as well as the mechanisms for achieving, or not achieving, better technical efficiency or reducing, or not reducing, household poverty.

Nigeria has been selected because it is a country where the agricultural sector is trapped in a cycle of low productivity. Nigeria is classified as a lower-middle-income country with a national GDP of US$449.1 billion as of 2019 (which is about half a per cent of the global economy), an estimated population of 201.0 million people, and a gross national per capita GDP of US$2,230 (World Bank, 2021). The average growth rate of Nigeria's GDP between 2007 and 2014 was 6.49%, which is higher than the average of Sub Saharan African countries (4.84%, excluding high-income countries) and European Union Countries whose growth rate was only 0.59% in the same period. However, there has been a sharp decline in the GDP growth rate of Nigeria since then to an average of 0.61% between 2015 and 2017 due to a period of severe recession in 2016, after which it remained at around 2% in 2018 and 2019 (World Bank, 2021).

Despite the long period of high economic growth of Nigeria, about 23.2% (42.2%) of the population lived on less than US$1.90 (US$3.20) a day in 2009 (at 2011PPP) (World Bank, 2021). In 2017 Nigeria overtook India as the country with the largest amount of absolute poverty in the world; with a large proportion of the poor engaged in agriculture. Agriculture accounts for about 40% of the country's GDP and employs about 65% of the people (World Bank, 2021). Thus, the agricultural sector is important in determining the quality of life and welfare of a large proportion of people in the country. However, it has lagged behind other sectors and the rest of the world in terms of productivity.

The low agricultural productivity in Nigeria could be caused by many factors ranging from poor soil qualities due to erosion, pollution and leaching, to scarcity and high cost of inputs. Others may be the continued use of crude implements and traditional or non-modern farming practices. However, this paper will examine whether the type of crop a farmer chooses to grow influences household outcomes in terms of productivity or poverty, even at the same level of underlying agricultural technologies or other factors.

To illustrate this point briefly, Table 1 summarises for 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 shows that outputs or revenues per hectare vary considerably across different crops. The cross-crop variations in the use of inputs, e.g., land, to achieve a certain level of revenues would justify our focus on differences in technical efficiency across different crops.

[Table 1 to be inserted]

This research is important for several reasons. Firstly, our study provides policymakers with insights into how the improvement in technical efficiency or poverty reduction is achieved by re-allocating crops given the current set of available inputs and agricultural technology. Whenever a new government came into power in Nigeria, it would often seek to come up with an overarching agricultural agenda for the agricultural sector, for instance, encouraging the production of certain crops which it deems more "important" (Iwuchukwu and Igbokwe, 2012). Drawing upon the large-scale national household survey dataset, this paper aims to provide policy implications for the government on the agricultural policy regarding the promotion of particular crops. It should also be noted that poverty and food security remain a major concern for many sub-Saharan African countries, including Nigeria. In these countries, the cropping decision could have far-reaching implications for national food security. If the production of certain crops is found to improve the welfare outcomes of farmers, such as poverty or food security, our results would provide an important policy lesson.

Our results show that farmers' 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 their crop choice. Also, we find that the adoption of cash crops will not reduce poverty, although it will improve technical efficiency marginally. In addition, the commercialisation of cash crops is found to be important for poverty alleviation, but not for improvements in technical efficiency. However, if cash crops are commercialised, poverty tends to decline.

The rest of this paper is laid out as follows. The next section highlights recent empirical studies on the productivity or the technical efficiency of smallholders and the effects of decisions to grow a crop on productivity and welfare. Section 3 discusses the methodology, starting with how the key crop choice variables are defined in this study, and then presents our main econometric models, namely, SFA and the treatment effects model. Section 4 explains the data and Section 5 presents the main results. The final section offers concluding observations.

2. Literature Review

Agricultural Productivity and Technical Efficiency in Nigeria

Technical efficiency is defined as the farmer's 'ability to produce maximum output given a set of inputs and technology' (Bravo-Ureta et al., 2007, p. 58), which is empirically measured by ‘the ratio of the produced output of an agricultural household over the maximally possible output, given a set level of inputs’. It takes the value between 0 and 1 where the higher value stands for higher efficient use of inputs in producing a unit of output given the agricultural technology. To measure agricultural farmers’ technical efficiency, two groups of methods can be employed: parametric and non-parametric methods. Among the parametric methods, stochastic frontier models have been most commonly used in the literature. For Nigeria, these models have been used to compute farmers’ technical efficiency for a large variety of crops including rice, wheat and cassava, among others (Adeyemo et al., 2010; Amaza et al., 2005; Ebong et al., 2009; Onyenweaku and Ohajianya, 2009). We also apply the stochastic frontier method, not for specific crops, but for a group of crops with the same characteristics as discussed later. In addition, our analysis draws upon the panel data and takes account of unobservable household characteristics. The difference between the parametric (like SFA) and non-parametric methods is that, while production functions are of a specified form for parametric analysis, there are no restrictive functional forms employed for the non-parametric method. An example of the non-parametric approach is the data envelopment group of models (Charnes, 1978). Other studies have used some partial measures of productivity such as yield per hectare in their analysis.

For example, Adeyemo et al. (2010) compute an average technical efficiency (TE) score of 0.89 for cassava farmers in Ogun state, while 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 and 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. Studies such as these are an indication of the range of calculated efficiency scores in particular regions, but this paper carries out a nationwide analysis using the nationally representative household panel data of Nigeria. To the best of our knowledge, this is the first study in which the nationwide panel dataset is used to perform the SFA to estimate technical efficiencies. This is expected to make a valuable contribution to the empirical literature.

Crop Choice, Productivity and Welfare in Developing Countries

In the papers reviewed below, household welfare is measured by domestic household per capita consumption. Using national household surveys from Mali, Delarue et al. (2009) studied the relationship between cotton production and household consumption and discovered that cotton producers consumed 9 per cent more food on average than non-cotton producing households where food consumption is a proxy for total consumption. When the authors disaggregated the results by the farm size, they found that the largest cotton producers consume up to 22 per cent more than the smallest producers, though these results imply correlations rather than causation. Loveridge et al. (2002) carried out a similar analysis of coffee for Rwanda and found a weak positive relationship between coffee production and the consumption outcomes of households. They speculated that this relationship could be explained by the low prices for coffee in the world market at the time of the survey, 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 found that technology could be a factor in the efficiency of cash-cropping among smallholders because those that used modern techniques spent 15 per cent more on food and 17 per cent more on all goods than the traditional producers. However, in addition to the methodology of Murekezi and Loveridge (2009), this study also takes into account differences in production technologies by distinguishing crops that are produced by vastly different methods of production from each other depending on the type of crops (i.e., cash and food crops). Similarly, Maertens and Swinnen (2009) found that the welfare of rural households vastly improved through their participation in high-yield vegetable exports in Senegal.

3. Methodology

Defining Crop Choice

This study proposes to address the research questions: 'Does choosing to grow a particular type of crop result in a higher level of technical efficiency and better household welfare outcomes or a lower level of household poverty?' This is closely related to "the cash-crop vs food-crop debate". As the name suggests, a cash crop is broadly defined as a crop that is grown primarily for sale to make a profit. Food crops are, on the other hand, grown primarily for the family of the farmer. However, in the literature of development economics, the term, 'cash crop', specifically denotes crops for exports and not necessarily the crops which are sold in the domestic market. According to the US Environmental Protection Agency, cash crops are typically purchased by organisations or commercial entities separate from the farm[3]. 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 using the real data. This is also important as our study intends to group similar crops rather than study farmers who grow an isolated crop against all the others. Therefore, we classify the crops with some modifications in our study.

Firstly, when cash crops are mentioned, the first picture that may come to the mind of readers is that of tree cash crops such as cocoa, coffee, palm oil, rubber etc. However, one of

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

the objectives of this paper is to identify what determined the choice of a crop planted and, if tree crops are used for cash crops, this purpose would be defeated. This is because if we try to measure the effect of a planting choice on productivity and poverty, we would need to capture the entire life cycle of the crop within one crop year. The production cycle of tree crops may span several years, which would make it difficult to compare their productivities or technical efficiencies with those of non-tree food crops. Therefore, we have excluded all the agricultural households with livestock and tree crops listed as their primary output in creating the crop choice variable. This ensures that our comparisons will be restricted to annual crops (that is, the crops that can complete a life cycle within a crop year).

The second modification is that we have restricted ourselves to the crops for which the data on export are fully available and there is likely to be a conflict in choosing between a food crop or a cash crop. For example, cassava 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. Cassava is often used in industry to produce ethanol and other biofuels. Therefore, we have classified farmers who produce cassava as cash-crop producers. It is clear that the type of crop produced alone does not determine how much of the farm product is marketed, so we have also included an index of commercialization as a variable of interest so that we could identify how much produce is sold versus self-consumed as an interaction with the type of crop produced. Here are further details about the crop classifications. Given the difficulty, we have grouped "representative" crops as either cash or food crops as in Tables 2 and 3. It should be noted that this classification is *exclusive,* that is, all the crops in our analyse are defined as *either* Cash Crops (C1) *or* Food Crops (C_2).

[Tables 2 and 3 to be inserted]

Cash Crops (C_1) – defined by the most representative cash crops, in terms of the overall share of exports (Table 2)

Food Crops (C_2) – defined by the most representative food crops, such as tuber and root crops and cereals (Table 3)

Cash Crops (C_1): To create the variable for the first category by most exported crops, data from the FAO were examined to determine which 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 Cash-Crops producing households (**C_1**). We aim to capture those agricultural households that grow crops that are most likely to be exported. As can be seen from Table 2, 11.06% of the sample planted one of the five crops in the first wave and 7.14% planted these in the second wave.

Food Crops (C_2): The major class of crops in the second category is made up of tubers and roots, which have long been recognised as particularly important for the food security of households in developing countries, especially those in Sub-Saharan African countries. Maize and rice are the only cereals included here because they are the most commonly consumed ones within households; while the others have traditionally only been grown by large-scale farmers rather than smallholders due to a lack of irrigation facilities and lack of sufficient financial capital (Grote et al., 2021), thus the data are not readily available on them. According to the Commission for Africa Report (2010), tubers are an important component of the diet for 2.2 billion people in developing countries. In Nigeria, they serve traditionally as a store of wealth as one could tell how rich a person was by the size of his or her yam barn (Obidiegwu and Akpabio, 2017). To illustrate this point further, Figure 1 shows that, even though farmers on average kept allocating a larger land for the production of 'cereals' like rice and maize than 'roots and tubers', the gap has been narrowed quickly in favour of the latter after 2009. In fact, there has been an upsurge in the production of tubers from around 2006, which explains an increase in the land area for roots and tubers.

[Figure 1 to be inserted]

Figure 2 further compares 'cereals' and 'roots and tubers' in terms of 'yield per hectare' as a rough measure of productivity. Figure 2 shows that roots and tubers have for long been a higher-yielding crop type than cereals and that this productivity gap has increased dramatically over the last three decades.

[Figure 2 to be inserted]

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

Household Commercialization Index (HCI)

In our empirical analyses, an index for the degree of commercialization of crop production per household is used to capture the extent to which an agricultural household's crop production - regardless of whether being for cash crops or food crops - was oriented toward commercial agriculture. Following Govereh et al. (1999) and Von Braun et al. (1994), which laid a standard for measuring commercialization, we calculate this index by taking the percentage of the value of the entire agricultural crop production in the year which is proxied by the gross value of crops sold. This computation will result in the number between 0 (%) and 100 (%) in which a household with an HCI of 0 is the one with none of its total crop production sold, while a household with an index of 100 is the one with all its crop output sold.

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

We will use in our regression analyses the interaction term between HCI and a variable of either cash or food crop choice. For example, interacting HCI with the cash crop in the poverty equation creates a variable that represents how much of these crops are sold, rather than consumed at home. Although this approach is limited as it ignores the absolute value of crop sales, the measure is still useful for describing agriculture in developing countries like Nigeria, because the smaller the farm is, the more likely it would consume a larger proportion of their total output at home for subsistence reasons rather than selling them (except for cases of higher value-added crops like cut flowers or vegetables) (Govereh et al., 1999).

Stochastic Frontier Analysis (with the Greene (2010) Correction for Selection Bias)

To estimate the technical efficiency of crop production, we will aggregate the data at the household level where each observation represents a unique productive entity. 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_i, the stochastic error term and u_i, the inefficiency error term. To illustrate, the base model takes the form:

$$ln(Y_i) = ln(f(\boldsymbol{X}_i)) + v_i - u_i \quad with\ u \geq 0 \qquad (2)$$

where v_i is either positive or negative and is assumed to be normally distributed with a mean zero and constant variance, as v_i represents an unsystematic stochastic effect related to measurement errors and random influences (e.g. luck, drought, flood, or other weather shocks, as mentioned earlier). On the other hand, u_i 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). However, as discussed earlier, the variable on crop choice is likely to be an endogenous variable. We have thus followed Greene (2010) who demonstrated that selection bias could make a significant difference if ignored in the computation of a production frontier. We estimated Greene's selection model for the stochastic frontier analysis in a panel data framework (Pitt and Lee, 1981) to take into account the household unobservable heterogeneity.

Three conventional inputs are 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[4]) and *inputs* (intermediate input costs like seed, fertilizer, pesticides, cost of irrigation, and costs to rent farm equipment/machinery). To gain some perspective on the results of this analysis, it may be useful to examine the nature of land distribution in Nigeria, especially as it relates to agriculture.

In an ideal case, there would also be a variable for capital (the depreciated cost of machinery and buildings), but this is not included due to data constraints. However, this is not a problem in our study context 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 rather than purchase them. It should also be noted that these rental costs are included in the inputs variable already. These inputs are used to produce the output y_{it} defined as the total revenue generated at the farm level, including by-products. Both the Cobb-Douglas model[5] and the trans-log model have been estimated, but we have adopted the trans-log model as it is a more general specification and performs better in our data context.

$$\begin{aligned} ln(Y_{it}) = \ & \beta_0 + \beta_1 ln(land) + \beta_2 ln(labour) + \beta_3 ln(other\ inputs) + \beta_4 ln^2(land) \\ & + \beta_5 ln^2(labour) + \beta_6 ln^2(other\ inputs) + \beta_7 ln(land) ln(labour) \\ & + \beta_8 ln(land) ln(other\ inputs) + \beta_9 ln(labour) ln(other\ inputs) + v_{it} \\ & - u_{it} \qquad (3) \end{aligned}$$

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

[5] The Cobb-Douglas model is specified as: $ln(Y_{it}) = \beta_0 + \beta_1 ln(land) + \beta_2 ln(labour) + \beta_3 ln(inputs) + v_{it} - u_{it}$.

Because of the non-symmetry of the conventional error term, ε_{it}, the expected value is defined as $E(\varepsilon_{it}) = -E(\varepsilon_{it}) \leq 0,\ \varepsilon_{it} = v_{it} - u_{it}$. The estimation by OLS will provide inconsistent estimates of the parameters apart from the intercept and cannot extricate the technical efficiency component from its normal residual error. The maximum likelihood estimation (MLE) will be thus employed in our study. MLE selects values of the model parameters that produce the distribution most likely to have produced the observed data by maximizing the likelihood function. 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 so that the efficiency estimates will be in the range between 0 and 1. This is useful because the standard deviation of the distribution can concentrate the efficiencies near-zero or spread them out (with zero cut off) (Aigner et al., 1977; Street, 2003).

Technical efficiency can then be derived by Equation (3) for each agricultural 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:

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

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, the categorical variables we have created for crop choice (**C**$_i$) might be biased by self-selection because farmers are unlikely to choose a particular crop to produce entirely at random. There are likely certain unobservable household characteristics (e.g., entrepreneurship, psychological factors) that influence their decision to produce these types of crops (**C**$_i$,) and that **C**$_i$ is endogenous as it is correlated with the error term of Equation (3).

To try to mitigate these problems, we follow Greene (2010) and 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 a particular type of crops (namely, cash or food crops) an agricultural household has made. In addition, crop choice is likely to be an endogenous determinant of poverty and technical efficiency.

The treatment effects model estimates the effect of an endogenous binary treatment, $\boldsymbol{C}_{it}$ (the crop choice in a binary case at time *t*), on a continuous, fully observed outcome variable, $\boldsymbol{Y_{it}}$ (in this case technical efficiency or poverty in separate models); conditional on vectors of explanatory variables, $\boldsymbol{X}_{it}$ and $\boldsymbol{Z}_{it}$(which would include exclusion restrictions). This can be modelled in the following way.

$$Y_{it} = \beta \boldsymbol{C}_{it} + \eta \boldsymbol{X}_{it} + \mu_i + \nu_{it} \tag{5}$$

In this case, β represents the parameter of interest as the average net effect of being treated on the outcomes, μ_i is the unobservable time fixed effect and ν_{it} is the error term. However, since $\boldsymbol{C}_{it}$, the crop choice, is endogenous, we would need to model the selection into treatment or the farmer's crop choice following Greene (2010). Further technical details of the treatment effects model are shown in Appendix 1.

4. Data

General description of data

For this analysis, the Nigeria General Household Survey-Panel (GHS-Panel) for 2010/2011 (Wave 1) and 2012/2013 (Wave 2) is used, which is the 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.[6] The panel covers all the 36 states of the country including the Federal Capital Territory, Abuja. The survey method was based on 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 is thus nationally representative. The total number of EA is 500.

For the GHS-Panel, 5,000 households were randomly 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 households in the panel sample where 10 households were randomly selected in each of 500 EAs.

Descriptive Statistics

[6] We have used the first two waves of the available four waves in GHS as most of the households were revisited in the second wave and the attrition bias is negligible as not all the initial households were revisited. The use of the first two surveys would minimise the attrition bias. However, future study should use Waves 3 and 4 (in 2015/16 and 2018/19) to see if our findings remain unchanged by correcting the attrition bias.

Table 4 presents descriptive statistics of some variables used in this study for Wave 1.[7] The mean age of the household heads in the sample is about 50 years and about 89% of the agricultural households are headed by males. In addition, the sample is almost 90% made up of households in the rural areas and 75% of the household heads are married. With regards to educational status, about 47 per cent of the sample are literate and can at least read or write, and the average length of time in formal education is about 4 years. The mean household size in the sample is about 6 individuals with averages of about 1 adult male, 2 adult females, 2 dependant males and 2 dependant females.

[Table 4 to be inserted]

5. Results

Agricultural Productivity in Nigeria

Table 5 shows the results of the crop productivity estimation of agricultural households in Nigeria, using the SFA with Greene's (2010) correction for sample selection bias regarding the decision over whether cash or food crops are chosen. The result of the production function (based on Equation (3)) shows that all the covariates, i.e., the logarithms of input terms, the squared logarithms of input terms, and the cross-interactions of the logarithms of input terms are statistically significant, except $ln(Land)$. Further, land, labour, and inputs - including seeds, fertilizer, and equipment among others - all positively contribute to the household-level agricultural outputs with stronger positive effects from the squared logarithm terms or the cross-logarithm terms. For instance, a 1% increase in land leads to a 2.36% increase in outputs without considering the effects from the squared- or cross-log terms. However, as the land increases, the output increases more than proportionally as a positive and significant coefficient estimate of $ln^2(Land)$ suggests. This will be further accelerated for a higher level of labour or inputs as implied by the cross-log terms. This implies that the large landholders can achieve a high level of outputs per unit of labour or inputs and their productivity in terms of *economic efficiency* or *per-unit output* is larger than that of smallholders. It should be noted, however, this does *not* imply a higher level of technical efficiency denoting the extent to which the observed level of output is close to the maximum feasible output given the observed levels of land, labour and inputs.[8]

[7] Appendix 2 provides descriptive statistics for Wave 2.
[8] See Ruttan (2002) on the evidence of how and why agricultural technical efficiency deteriorated, while productivity (or Total Factor Productivity) increased from 1965 to 1995 in the international context.

[Table 5 and to be inserted]

A similar relationship is observed between labour and outputs or inputs and output. However, the estimated coefficient of $ln(Labour)$ is positive and not statistically significant, while $ln^2(Labour)$ and cross-log terms are significant. On the other hand, the estimated coefficients of both $ln(Other\ Inputs)$ and $ln^2(Other\ Inputs)$ are statistically significant. Overall, the results suggest that the total output is getting higher and higher once the farmer increases the input levels proportionally. This does not necessarily imply the farmers use the inputs optimally as, for example, the technical efficiency tends to decrease as the land size increases as discussed later (see Table 6).

σ_v^2 in the middle panel of Table 5 shows the estimate of the variance of ν_{it}, the idiosyncratic error term in Equation (3), and σ_u^2, the variance of u_{it}, the technical inefficiency error term, where $u_{it} \sim^{iid} N^+(\mu,\ \sigma_u^2)$ in which μ is the estimate of μ, which is the mean of the truncated-normal distribution. γ is the estimate of σ_u^2/σ_S^2 where σ_S^2 is the estimate of $\sigma_S^2 = \sigma_v^2 + \sigma_u^2$. Due to the restrictions on γ, the optimization is parameterised in terms of its inverse logit, and this estimate is reported as $logit(\gamma)$. Likewise because σ_S^2 must be positive, the optimization is parameterised in terms of $ln(\sigma_S^2)$. The log-likelihood value for this model based on the trans-log specification is higher (-987.40) than that for the Cobb-Douglas model (-1719.32). The LR test result suggests that the trans-log specification should be used.

We have found that the overall technical efficiency averages about 64.3% as suggested by the estimate of $ln(\sigma_S^2)$. This is lower than other estimates that have been obtained by more crop-specific studies (e.g., 89% by Adeyemo et al., 2010; 81% in Ebong et al., 2009; 65% in Onyenweaku and Ohajianya 2009; and 68% in Amaza et al., 2005). Given that technical efficiency represents how effectively inputs produce the output in comparison with the maximum output level which could be achieved by the same set of inputs, our estimates suggest that there is room for improvement in productivity given the current levels of inputs and technology.

Using the econometric results of the trans-log model and following the methodology put forward by Chen et al. (2009), we have derived the elasticities and marginal products of factors of production as presented in the bottom panel of Table 5. Elasticities are evaluated at the geometric means of the inputs and output (ibid., 2009, p. 159). These estimates suggest that labour and land have large coefficient estimates of around 0.4 and the estimate of 'other inputs' is only 0.2. That is, on average, if land or labour increases by 1%, the output increases on average 0.4%, while the output will increase only by 0.2% in response to a 1% increase in 'Other Inputs'. The estimates of the marginal product show the extent to which the total output increases in response to a unit increase in land, labour or 'Other Inputs'. The results are only indicative as they are evaluated at means, but they suggest that overall outputs tend to increase in response to the increase in either labour, land or 'Other Inputs'.

Table 6 shows the variation in technical efficiency across the sample by gender and age of the household head as well as household land size based on the first wave.[9] It indicates that male-headed households in the sample are more efficient than female-headed households with average technical efficiency of 64% as opposed to 60%. As expected, the most 'technically efficient' age of the head of the household ranges between 20 and 60 years with a technical efficiency estimate of 75%. The results of Table 6 can be associated with Schultz's (1964) hypothesis of "the efficient small farmer". It is noted that Schultz formulated the hypothesis that small-scale farmers in developing countries were "poor-but-efficient"[10], implying that they made the best decisions in allocating their scarce resources by responding to price incentives. Consistent with the Schultz view, the technical efficiency is found to reduce as land size increases. Furthermore, most of the households with the land size below 10 hectares fall within the 50%-75% range of technical efficiency, while the share between 25-50% is the largest for the large landholders with the land size above 10 hectares. However, as we discussed earlier, small-holders are 'economically inefficient' as a statistically significant coefficient estimate of $ln^2(Land)$ in Table 5 implies that as land size increases, the output tends to increase more than proportionally. It should be noted that, given that land size remains almost the same between the rounds, the land-output relationship primarily depends on the cross-sectional relationship.

[Table 6 to be inserted]

Impact of crop choice on technical efficiency and poverty

This sub-section reports the results of the treatment effects model to estimate the determinants of crop choice (i.e., choosing cash crop over food crop) and hence the impact of this choice on technical efficiency and mean per capita consumption expenditure (MPCE). Here the low level of MPCE implies poverty. In essence, we will test whether the productivity and welfare differences between the two groups of farmers with different crop choices (i.e., cash and food

[9] We have obtained similar results for the second wave. They will be furnished on request.

[10] Here poverty is defined in terms of land-holding given that the smallholders tend to be poor in many developing countries, including Nigeria.

crops) are significantly different from zero after controlling for household characteristics and addressing the endogeneity associated with the farmers' crop choice.

The results are reported in Tables 7 and 8. Column (1) of both tables shows the results of the first stage selection into the treatment equation, determining the probability of being treated (growing cash crops). However, since these are drawn from probabilistic functions and not from linear probability modelling, the coefficients cannot be interpreted as probabilities, but indicate the direction of the effect and its statistical significance. Column (2) indicates the results of the impact equation in the second stage, showing the average treatment effect on the treated (i.e., the households choosing cash crops) in comparison to the counterfactuals where the same households chose food crops, rather than cash crops given the observable household characteristics and unobservable household fixed effects.

[Tables 7 and 8 to be inserted]

The exclusion restrictions used for the equation are the amount of free input used in production, the amount of non-farm income the household possesses and the amount of seed used from the previous growing season. The instruments are strong as the F statistic for excluded instruments is 29.04. On the contrary, for the consumption expenditure equation, only the free input and previous year's seeds are used because non-farm income is directly related to household expenditure. These variables were positive and significant in determining participation in growing export-oriented crops and tubers or roots. The instruments are strong with the F statistic for excluded instruments equal to 25.02.

For the use of the previous year's seeds variable, the data show that the greater the amount of primary input like seeds that were saved from the previous year, the more likely it would be for that agricultural household to plant the cash crop in the next growing season. The amount of free agricultural input received is positive and significant at the 1% significance level in both regressions. This indicates that at the point where farmers decide on the crop to produce, there is scope to influence their decisions by the amount of free agricultural inputs they are given. The positive parameter estimate implies that the more inputs received, the more likely the households would choose to produce cash crops. This is because cash crops in general require a greater initial investment where free inputs act as a buffer to reduce the costs, or risks, of planting the crops.

Other major significant determinants of choosing cash crops include the land size, the regions in which the household resides, the size of the household and the gender of the household head. If a household owns a large area of land, it is more likely to grow cash crops.

The regions are 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 the crop produced. The size of the household is significant and positive. This indicates that the larger a household is, the more likely they are to plant cash crops. This is possibly because a larger household can devote more hours of labour necessary for cash-crop production.

On the impact of the choice on technical efficiency, if a household adopts cash crops, given the observable household characteristics and unobservable household fixed effects, its technical efficiency will be higher by 0.026 on average once the sample selection bias regarding the crop choice is taken into consideration.

Table 8 indicates, however, that the selection of cash crops has a significant negative effect on the log of mean household expenditure per capita (MPCE). The estimated coefficient is negative, implying that the farmer who has grown cash crops has a lower MPCE on average. This implies that, if a food-crop farmer chooses to grow cash crops, the expenditure of the household headed by him/her is likely to be lower. This would increase overall poverty by making a non-poor household poor or a poor household poorer. It should be noted that this estimate is based on the methodology taking into account sample selection bias and the current household characteristics. Even if cash crops would potentially increase productivity, unless the cash-crop producers are supported by policies that would help them grow new crops, it would not make sense for the food-crop farmers to switch to cash crops given the current conditions, as this switch would make them potentially poorer.

Finally, Tables 9 and 10 report the results of the impact of commercialization and its interactions with the categorical choice variables, namely cash crop or food crop, on technical efficiency and poverty. In each table, Columns 1 and 2 show the results for commercialization without any interaction terms based on a Fixed-Effects (FE) model and a Correlated Random Effects (CRE) model. On the other hand, Columns 3 and 4 show the impact of commercializing the cash crops, and Columns 5 and 6 for that of commercializing food crops. The results in Columns 1 and 2 of Table 9 show that the household index of commercialization is not a statistically significant determinant of technical efficiency, but it is significant for poverty. This is somewhat surprising because one might expect that the more commercialized a farm household is, the better its technical efficiency would be due to the monetary incentives in producing the most output possible with the given amount of inputs. However, the incentives to the household head of increasing technical efficiency to keep his family fed may be greater than the incentives from doing so for the sake of the possible monetary value of his goods. Our results thus imply that if the government is interested in increasing technical efficiency, it should prioritise food security over commercialization. The result in Columns 1 and 2 of Table

10 that commercialization is negatively associated with MPCE implies that, if poverty alleviating policy is the main policy concern, commercialization alone is not sufficient.

[Tables 9 and 10 to be inserted]

A few interesting results emerge when the crop choice variables are interacted with the index of commercialisation. For instance, if food crops are commercialised, the technical efficiency tends to be higher (Columns 5 and 6 of Table 9), while similar results are not found for cash crops (Columns 3 and 4). We observe in Table 10 that, if the cash crop is commercialised, MPCE tends to be higher (Columns 3 and 4), while if the food crop is commercialised, MPCE tends to be lower (Columns 5 and 6). It is conjectured that the commercialisation of cash crops can potentially reduce poverty as this would bring higher income through market-based transactions. However, the commercialization of food crops would reduce the self-consumption at home and could potentially make the household poorer, or food insecure. Hence, if the government adopts an agricultural policy of commercialising the agricultural products, it should pay attention to the type of crops with respect to their differential impact on household poverty.

6. Conclusion

The present study examines the arguments on whether or not smallholder farmers in Nigeria who produce certain types of crops (cash crops versus food crops) experience any technical efficiency and welfare differences, and 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, we re-examined the old arguments surrounding whether smallholders are indeed "efficient but poor" (Abler and Sukhatme, 2006). We have carried out stochastic frontier analysis (SFA) with Greene's (2010) correction for sample selection about crop choices, and have found that smallholders are generally efficient in their allocation of resources. We have found that access to free inputs, non-farm income, the use of seeds from the previous growing season, the larger owned land, household size, gender and the different regional differences were the main determinants of choosing cash crops. The results of SFA imply the large farmers – which are typically cash-crops growing farmers is productive, though not technically efficient. So even if there is an economic incentive to switch to cash crops, the results of the selection equation imply a high hurdle for opting to grow cash crops (e.g., acquiring larger land, using more inputs). Consistent with this observation, we have also econometrically found that the adoption of cash crops by food-crop producing households will not reduce poverty, although it will improve technical efficiency marginally. Poverty is reduced only if cash crops are commercialised.

Our results provide a few important policy implications. Implications of our research may differ depending on the national poverty alleviation strategy of the government, but our findings show why far less effort would be needed to lift these groups of farmers out of poverty. The econometric results suggest that agricultural household crop choices are not random, but can be predicted by socioeconomic factors. This means that there are factors that could influence the eventual choice of the crop planted. If the government wishes to promote cash crop production, the policies helping farmers purchase inputs at lower prices (e.g., microcredit programmes or subsidies for poor farming households) would be useful in this context. However, the government should also be aware that a switch to cash crops without adjustment of farmers' factor endowments can be poverty-increasing. Shultz's argument for investment in human capital, particularly education (Schultz, 1961) is still valid in the present context to help poor farmers invest in new technologies and escape from poverty. The agricultural extension could be utilised to get more people within areas of comparative advantage to switch to these high productivity crops to improve their welfare. Educating farmers on the marketing opportunities for their products, and concomitant greater commercialization would also have positive welfare effects.

References

Adeyemo, R., Oke, J. T .O., and Akinola, A. A. ‘Economic efficiency of small scale farmers in Ogun State, Nigeria’, *Tropicultura,* 28.2, (2010) pp. 84-88.

Aigner, D., Lovell, CA K., and Schmidt, P, ‘Formulation and estimation of stochastic frontier production function models’’ *Journal of Econometrics,* 6.1 (1977) pp. 21-37.

Abler, D.G. and Sukhatme, V.A., 2006. The" efficient but poor" hypothesis. *Review of Agricultural Economics*, 28(3), pp.338-343.

Amaza, P. S., Kamara A. Y. and Helsen J. ‘Determinants of food crop production and technical efficiency in Guinea Savannahs of Borno State, Nigeria’, *African crop science proceedings* 7, (2005) Available at: https://www.africabib.org/rec.php?RID=Q00041758 (last accessed 6 September 2021).

Bravo-Ureta, B. E., Solís, D., López, V. H. M., Maripani, J. F., Thiam, A., and Rivas, T. ‘Technical efficiency in farming: a meta-regression analysis’, *Journal of Productivity Analysis,* Vol. 27(1), (2007) pp. 57-72.

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

Chen, Z., Huffman, W.E. and Rozelle, S., 2009. Farm technology and technical efficiency: Evidence from four regions in China. *China Economic Review*, 20(2), pp.153-161.

Coelli, T. J. and Battese, G. E. ‘Identification of factors which influence the technical inefficiency of indian farmers’, *Australian Journal of Agricultural and Resource Economics,* Vol. 40 (2), (1996) pp. 103-128.

Commission for Africa (2010) Annual report 2010 : Economic Commission for Africa. UN. https://repository.uneca.org/handle/10855/3100 (last accessed 23 April 2022).

Delarue, J., Mesple-somps, S., Naudet, J., and Robilliard, A. ‘The Sikasso Paradoxe: Cotton and Poverty in Mali’, Development Institutions and Analysis Working Paper. 2009. Available at: http://entretiens.iamm.fr/pluginfile.php/1909/mod_resource/content/2/Paradoxe%20of%20Sikasso_summary%20on%20statistics.pdf (last accessed 23 April 2022).

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

Govereh, J., Jayne, T. S., and Nyoro, J. ‘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, Working Paper 39,* 1999. Available at:

https://www.researchgate.net/publication/238075353_Smallholder_Commercialization_Interlinked_Markets_and_Food_Crop_Productivity_Cross-Country_Evidence_in_Eastern_and_Southern_Africa (last accessed 6 September 2021).

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

Greene, W. H. 'A stochastic frontier model with correction for sample selection', *Journal of Productivity Analysis* Vol. 34, (2010) pp. 15–24.

Grote, U., Fasse, A., Nguyen, T. T., & Erenstein, O. (2021). Food security and the dynamics of wheat and maize value chains in Africa and Asia. Frontiers in Sustainable Food Systems, 4, 617009.

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

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

Jiang, N., and Sharp, B. 'Technical efficiency and technological gap of New Zealand dairy farms: A stochastic meta-frontier model', *Journal of Productivity Analysis*, Vol. 44 (2015) pp. 39-49.

Loveridge, S., Nyarwaya, J. B., and Shingiro, E. 'Decaffeinated? Situation, Trends and Prospects for Smallholder Coffee Production in Rwanda', FSRP and DSA, Ministry of Agriculture, Livestock and Forestry 34pp. 2002. Available at: https://ageconsearch.umn.edu/record/55385/files/farmer_attitudes_coffee.pdf (last accessed 6 September 2021).

Lundahl, M., 1987. 'Efficient but Poor',—Schultz', theory of traditional agriculture. *Scandinavian Economic History Review*, 35(1), pp.108-129.

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

Martey, E., Wiredu, A. N., Etwire, P. M., and Kuwornu, J. K. (2019). The impact of credit on the technical efficiency of maize-producing households in Northern Ghana. *Agricultural Finance Review*, 79(3), 304-322.

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. '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. R. ‘Economic Efficiency and Frontier Techniques’, *Journal of Economic Surveys*, 18(1), (2004) pp. 33–77.

Obidiegwu, J. and Akpabio, E. ‘The geography of yam cultivation in southern Nigeria: Exploring its social meanings and cultural functions’, *Journal of Ethnic Foods*, Vol. 4(1), (2017) pp. 28-35.

Onyenweaku, C.E. and Ohajianya D.O. ‘Technical Efficiency of Rice Farmers in Ebonyi State, Nigeria’, *Journal of Agriculture and Food Sciences* Vol. 3(2), (2009) pp. 175–184.

Pitt, M. M., and Lee, L. F. ‘The measurement and sources of technical inefficiency in the Indonesian weaving industry’, *Journal of Development Economics*, Vol. 9(1), (1981) pp. 43-64.

Rahman, S. ‘Resource use efficiency under self‐selectivity: the case of Bangladeshi rice producers’, *Australian Journal of Agricultural and Resource Economics*, Vol. 55(2), (2011) pp. 273-290.

Ruttan, V.W., 2002. Productivity growth in world agriculture: sources and constraints. *Journal of Economic Perspectives*, 16(4), pp.161-184.

Schultz, T.W., 1961. Investment in human capital. *The American Economic Review*, 51(1), pp.1-17.

Schultz, T. W. *Transforming traditional agriculture* (Chicago, USA: University of Chicago Press, 1964),

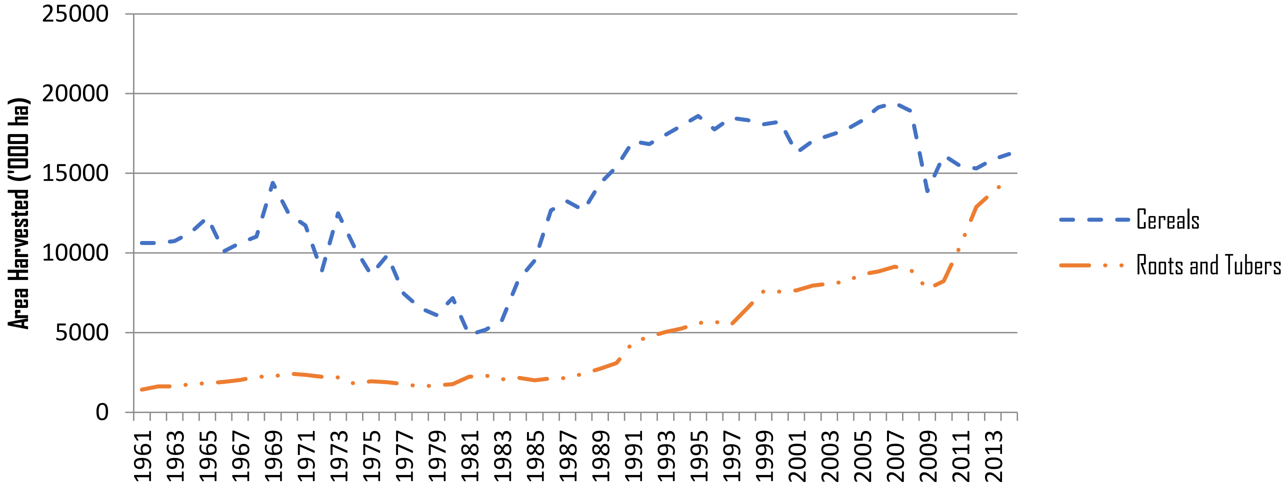
Street, A. ‘How Much Confidence Should we Place in Efficiency Estimates?’, *Health Economics*, Vol. 12(11), (2003) pp. 895-907.

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.

Wooldridge, J. M. Quasi-Maximum Likelihood Estimation and Testing for Nonlinear Models with Endogenous Explanatory Variables. Department of Economics, Michigan State University, 2012. Available at: http://econ.msu.edu/faculty/wooldridge/docs/qmle_endog_r3.pdf (last accessed 6 September 2021).

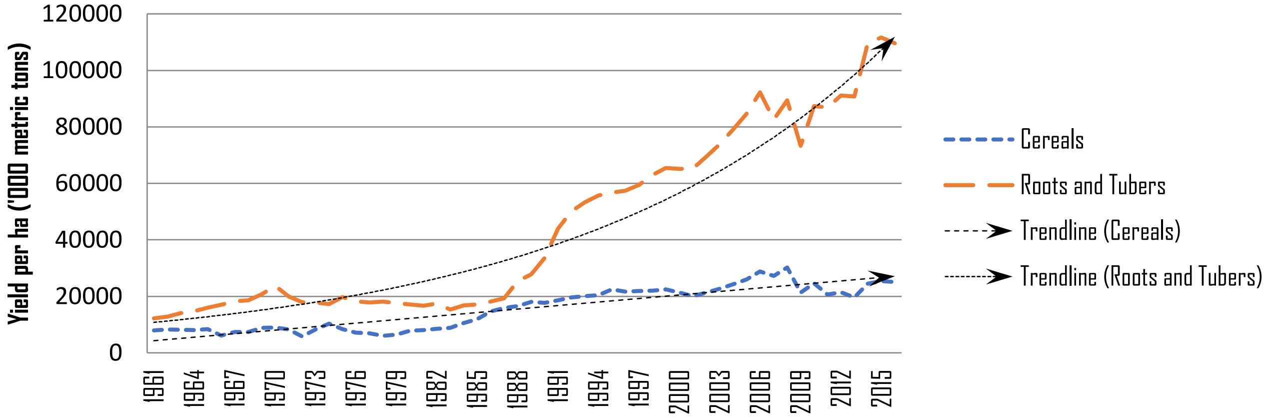
World Bank. (2021), World Development Indicators (WDI), https://databank.worldbank.org/source/world-development-indicators (last accessed 6 September 2021).

Figure 1: Time Trend of Area Harvested for Cereals, Roots and Tubers in Nigeria



Source: Authors' drawing from FAOSTAT database, 2016 database

Figure 2: Time Trend of Yield/Ha for Cereals, Roots and Tubers in Nigeria



Source: Authors' drawing from FAOSTAT database, 2016 database

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

Table 2: List of cash-crops

Crops (**C_1**)	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 (Beni-seed)	127.60	0.36	0.35
Total	44790.94	11.06	7.14

Source: Author's calculation based on the Nigerian LSMS data for 2011 and 2013.

Table 3: List of food crops

Crop	% of sample (wave 1)	% of sample (wave 2)
Yam	21.51	23.17
Maize	8.07	7.30
Rice	2.90	2.74
Cocoyam	1.49	1.71
Groundnuts	1.79	1.45
Potatoes	0.58	0.64
Ginger	0.08	0.08
Total	36.42	37.09

Source: Authors' calculation based on the Nigerian LSMS data for 2011 and 2013.

Table 4: Descriptive Stats (for Wave 1)

Variable	Mean	Std. Dev.	Min	Max
Primary output is cash-crop	0.11	0.32	0	1
Primary output is food-crop	0.35	0.47	0	1
Household Commercialization Index	48.22	7.36	0	80.40
ln(Total Food Auto-Consumed in HH)	10.75	1.21	1.78	13.94
ln(output)	10.98	1.72	0	15.59
ln(land)	8.89	1.73	0	13.04
ln(labour)	4.26	5.30	0	16.73
ln(Other Inputs)	7.01	4.41	0	14.25
Age of HH Head	50.09	15.10	16	110
Marital Status of HH (Married=1)	0.75	1.71	0	1
Religion of HH Head (Christian=1)	0.53	0.55	0	1
Gender of HH Head	0.89	0.31	0	1
Number of adult males in household	1.36	0.93	0	11
Number of adult females in household	1.54	0.89	0	7
Number of dependent males in household	1.69	1.62	0	16
Number of dependent females in household	1.51	1.47	0	11
Household size	6.11	3.13	1	31
Literate (Can read and write=1)	0.47	0.49	0	1
Years of education of HH Head	3.89	3.24	1	13
Rural	0.89	0.32	0	1
Mean per capita expenditure (MPCE) in naira	448408.6	290725.4	33907.57	2975185

Source: Authors' calculation based on the Nigerian LSMS data for 2011 and 2013.

Table 5: Maximum likelihood estimates of the Stochastic Frontier Analysis model with Greene (2010) correction for sample selection bias

	Translog Production Function	
	Coefficient	SE
Constant	9.066	1.649
$ln(Land)$	2.363***	0.677
$ln(Labour)$	0.590	0.510
$ln(Other\ Inputs)$	2.220***	0.047
$ln^2(Land)$	0.282*	0.122
$ln^2(Labour)$	0.273***	0.086
$ln^2(Other\ Inputs)$	0.065*	0.040
$ln(Land)\ ln(Labour)$	0.366***	0.015
$ln(Land)ln(Other\ Inputs)$	0.794***	0.014
$ln(Labour)ln(Other\ Inputs)$	0.398***	0.014
Year dummy	-0.140*	0.073
σ_S^2	2.695	0.030
γ	0.163	0.002
σ_u^2	0.232	0.180
σ_v^2	3.705	0.152
$ln(\sigma_S^2)$	0.643***	0.019
$logit(\gamma)$	-1.133***	0.160
μ	19.387	22.131
Statistics		
No. of obs.	5192	
No. of groups	3045	
Log likelihood (Trans-log)	-987.4	
Log likelihood (Cobb-Douglas)	-1719.32	
LR-stat for $H_{o:}$ The two ratios are not different	1002.13***	
Decision	Trans-log preferred	

Elasticities and Marginal Products of Factors of Production

	Coefficient	SE
Elasticities		
Land	0.390***	0.100
Labour	0.409***	0.071
Other Inputs	0.201***	0.071
Marginal Partial Product		
Land	223.02	22.44

Labour	14.71	1.870
Other Inputs	1.420	0.460

Notes: 1. ***, **, * represents significance at 1%, 5%, and 10% alpha respectively .

2. The result is based on Equation (3). σ_S^2 is the estimate of the sum of σ_u^2, the variance of u_{it}, the technical inefficiency error term, and σ_v^2, the variance of v_{it}, the idiosyncratic error term. γ is the estimate of σ_u^2/σ_S^2, showing the estimated proportion of the inefficiency component in the total variance in the aggregate error term. μ is the estimate of the mean of the technical inefficiency error term, where $u_{it} \sim^{iid} N^+(\mu,\ \sigma_u^2)$.

3. Elasticities are evaluated at the geometric means of the inputs and output; Standard errors are calculated using the delta method;

Source: Authors' calculation based on the Nigerian LSMS data for 2011 and 2013.

Table 6: Technical Efficiencies of different segments of the population by the characteristics of the household heads (from Wave 1) based on SFA with Greene (2010) correction for sample selection bias

	Male	Female	Age (<20)	Age (20-60)	Age (>60)	Land size (<1ha)	Land size (1-5ha)	Land size (5-10ha)	Land size (>10ha)
Technical Efficiency (<25%)	5%	13%	10%	1%	6%	6%	8%	7%	7%
Technical Efficiency (25-50%)	19%	37%	20%	6%	10%	19%	20%	12%	45%
Technical Efficiency (50-75%)	66%	50%	69%	69%	64%	70%	67%	67%	38%
Technical Efficiency (>75%)	10%	0%	1%	24%	20%	5%	5%	14%	10%
Overall Average Technical Efficiency	64%	60%	63%	75%	64%	64%	64%	61%	60%

Source: Authors' calculation based on the Nigerian LSMS data for 2011 and 2013.

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

	C_1 – Farmer chose a cash crop	
	Selection	Impact
	(1)	(2)
Crop Choice		0.026***
		(0.001)
Age of HH Head	0.015	0.0003
	(0.990)	(0.0006)
Age Square of HH Head	-0.055	-0.0000
	(0.22)	(0.0000)
Education of HH Head	0.0007	-0.0433***
	(0.000)	(0.0028)
HH Size	0.149***	-0.440***
	(0.008)	(0.280)
Sex of HH Head	0.527***	0.857***
	(0.054)	(0.055)
Rural	-0.004	0.065
	(0.007)	(0.048)
Female Share	0.000	-0.000
	(0.000)	(0.001)
Married	0.354***	0.000
	(0.002)	(0.001)
Region1 (NW)	0.167	-0.008*
	(0.209)	(0.004)
Land Size	0.541***	1.793***
	(0.099)	(0.821)
Farm Machinery owned	0.020***	0.002
	(0.001)	(0.002)
Region2 (NC)	2.340***	0.166***
	(0.711)	(0.004)
Region3 (SW)	-1.207***	-0.197
	(0.112)	(0.187)
Region4 (SE)	2.131***	-0.001***
	(0.209)	(0.000)
Region5 (SS)	2.903***	0.032
	(0.901)	(0.022)
Free Inputs[#]	0.877***	
	(0.199)	
Non-farm income[#]	0.107*	
	(0.067)	

Previous year's seeds#	0.902***	
	(0.081)	
Constant	-16.384***	7.588***
	(0.495)	(0.018)
F-Stat, excl. instruments	29.04	
p-value	0.000	
N	2422	2422
Time Dummies	Yes	Yes

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; # Exclusion restrictions; F-stat below 10 indicates weak instruments

Source: Authors' calculation based on the Nigerian LSMS data for 2011 and 2013.

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

	C_1 – Farmer chose a cash crop	
	Selection (Probit)	Impact
	(1)	(2)
Crop Choice		-0.530**
		(0.009)
Age of HH Head	-0.007	0.008
	(0.019)	(0.007)
Age Square of HH Head	0.001	0.000
	(0.00)	(0.00)
Education of HH Head	-0.055	0.085
	(0.095)	(0.037)
HH Size	0.101*	0.150***
	(0.008)	(0.008)
Sex of HH Head	0.537***	-0.274**
	(0.054)	(0.096)
Rural	-0.22	0.011
	(0.34)	(0.057)
Female Share	-0.00007	-0.029***
	(0.000)	(0.006)
Married	0.177	-0.075***
	(0.163)	(0.017)
Land Size	0.816***	0.065***
	(0.018)	(0.003)
Farm Machinery owned	0.191***	0.191***
	(0.020)	(0.020)
Region1 (NW)	0.560*	-0.118*
	(0.270)	(0.052)
Region2 (NC)	1.266***	-0.221***
	(0.257)	(0.056)
Region3 (SW)	1.276***	-0.038
	(0.289)	(0.087)
Region4 (SE)	1.277***	-0.239***
	(0.263)	(0.061)
Region5 (SS)	2.471***	-0.087
	(0.263)	(0.080)
Free Inputs#	0.677***	
	(0.023)	
Previous year's seeds#	0.420*	

	(0.10)	
Constant	-2.706***	11.084***
	(0.619)	(0.235)
F-Stat, excl. instruments	25.02	
p-value	0.000	
N	2422	2422
Time Dummies	Yes	Yes

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; # Exclusion restrictions; ; F-stat below 10 indicates weak instruments

Source: Authors' calculation based on the Nigerian LSMS data for 2011 and 2013

Table 9: Results of Impact of Crop Commercialization with Crop Choice on Technical Efficiency: Fixed Effects (FE) Model and Correlated Random Effects (CRE) Model

	FE	CRE	FE	CRE	FE	CRE
	(1)	(2)	(3)	(4)	(5)	(6)
commercialization	0.012	0.018				
	(0.038)	(0.038)				
Cash crop*commercialization			-0.03	-0.112		
			(0.026)	(0.113)		
Food crop*commercialization					0.046*	0.070**
					(0.203)	(0.325)
Age of HH Head	0.134***	0.008	0.096***	0.008	0.096***	0.008
	(0.027)	(0.007)	(0.027)	(0.007)	(0.027)	(0.007)
Age Square of HH Head	-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 of HH Head	-0.022*	-0.300**	-0.022*	-0.300**	-0.022*	-0.300**
	(0.013)	(0.096)	(0.013)	(0.096)	(0.014)	(0.096)
Education of HH Head	0.096	0.067	0.096	0.067	0.096	0.067
	(0.095)	(0.037)	(0.095)	(0.037)	(0.095)	(0.037)
HH Size	0.747***	0.152***	0.747***	0.152***	0.747***	0.152***
	(0.045)	(0.008)	(0.045)	(0.008)	(0.045)	(0.008)
Rural	0.01	0.019***	0.01	0.019***	0.01	0.019***
	(0.35)	(0.01)	(0.35)	(0.01)	(0.35)	(0.01)
Female Share	-0.022	-0.050	-0.022	-0.050	-0.022	-0.050
	(0.22)	(0.041)	(0.22)	(0.041)	(0.22)	(0.041)
Married	2.02e-05	0.358	2.02e-05	0.358	2.02e-05	0.358
	(1.81e-05)	(0.041)	(1.81e-05)	(0.041)	(1.81e-05)	(0.041)
Farm Machinery owned	0.210***	0.988***	0.210***	0.988***	0.210***	0.988***
	(0.0270)	(0.0178)	(0.0270)	(0.0178)	(0.0270)	(0.0178)
Land Size		0.209***		0.201***		0.208***
		(0.0210)		(0.0210)		(0.0210)
Region1 (NW)		-0.808***		-0.808***		-0.808***
		(0.280)		(0.280)		(0.280)
Region2 (NC)		0.766***		0.766***		0.766***
		(0.316)		(0.316)		(0.316)
Region3 (SW)		-0.001		-0.001		-0.001
		(0.00)		(0.00)		(0.00)
Region4 (SE)		-0.299		-0.299		-0.299
		(0.270)		(0.270)		(0.270)
Region5 (SS)		-0.087		-0.087		-0.087
		(0.080)		(0.080)		(0.080)
Constant	10.23***	11.095***	10.23***	11.095***	10.23***	11.095***

	(0.326)	(0.229)	(0.326)	(0.229)	(0.326)	(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

Source: Authors' calculation based on the Nigerian LSMS data for 2011 and 2013

Table 10: Results of the Impact of Crop Commercialization with Crop Choice on Poverty (log MPCE): Fixed Effects (FE) Model and Correlated Random Effects (CRE) Model

	FE	CRE	FE	CRE	FE	CRE
	(1)	(2)	(3)	(4)	(5)	(6)
Commercialization	-0.140***	-0.060***				
	(0.026)	(0.008)				
Cash crop*commercialization			0.022*	0.0188***		
			(0.00766)	(0.006)		
Food crop*commercialization					-0.088*	-0.232***
					(-0.021)	(0.022)
Age of HH Head	0.088***	-0.001	0.088***	-0.001	0.088***	-0.001
	(0.027)	(0.003)	(0.027)	(0.003)	(0.027)	(0.003)
Age Square of HH Head	-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 of HH Head	-1.640	0.080***	-1.740	0.080***	-1.740	0.080***
	(1.025)	(0.015)	(1.025)	(0.015)	(1.025)	(0.015)
Education of HH Head	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	0.00001	-0.056***	0.00001	-0.056***	0.00001	-0.056***
	(0.000)	(0.008)	(0.000)	(0.008)	(0.000)	(0.008)
Farm Machinery owned	0.779***	0.816***	0.779***	0.816***	0.779***	0.816***
	(0.0270)	(0.0178)	(0.0270)	(0.0178)	(0.0270)	(0.0178)
Land Size		0.191***		0.191***		0.191***
		(0.0210)		(0.0210)		(0.0210)
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.

Source: Authors' calculation based on the Nigerian LSMS data for 2011 and 2013

Appendix 1

Households' crop choice, $\boldsymbol{C}_{it}$, in Equation (5) can be written as:

$$\boldsymbol{C}_{it}{}^{*} = \gamma \boldsymbol{Z}_{it} + \varepsilon_{it} \qquad \text{(A1)}$$

The selection into treatment $\boldsymbol{C}_{it}{}^{*}$ in this model is a function of ε_{it}, which is correlated with ν_{it}, the error term in the outcome equation of $\boldsymbol{Y}_{it}$ above. Thus, $\boldsymbol{C}_{it}{}^{*}$ is 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 $\boldsymbol{Z}_{it}$ and a random component ε_{it}. The relationship between the observed $\boldsymbol{C}_{it}$ and the latent $\boldsymbol{C}_{it}{}^{*}$ can be defined in this way:

$$\boldsymbol{C}_{it} = \begin{cases} 1, & if \quad \boldsymbol{C}_{it}{}^{*} < 0 \\ 0, & if \quad \boldsymbol{C}_{it}{}^{*} \geq 0 \end{cases} \qquad \text{(A2)}$$

The problem here is that estimating Equation (A1) directly by OLS would only be consistent if there is no correlation between ν_{it} and ε_{it} (notationally, this correlation is represented by ρ; so ideally, we want $\rho = 0$) (Greene, 2008). But in this case, ρ is not zero, thus a different method would have to be used to estimate the coefficients consistently.

Formally, if we assume that the binary data ($\boldsymbol{C}_{it}$) have been generated by an underlying normal distribution, the expected conditional outcome of productivity and poverty (Y_{it}) could be written in this way:

$$\begin{aligned} E[Y_{it}|\boldsymbol{C}_{it},\boldsymbol{X}_{it},\boldsymbol{Z}_{it}] &= \eta \boldsymbol{X}_{it} + \beta \boldsymbol{C}_{it} + \mu_i + E[v_{it}|\boldsymbol{C}_{it},\boldsymbol{X}_{it},\boldsymbol{Z}_{it}] \\ = \eta \boldsymbol{X}_{it} + \beta \boldsymbol{C}_{it} + \mu_i & \\ &+ \left[\rho_1 \sigma_{v_1} \{\phi(\gamma \boldsymbol{Z}_{it})/\Phi(\gamma \boldsymbol{Z}_{it})\}|\boldsymbol{C}_{it},\boldsymbol{X}_{it},\boldsymbol{Z}_{it}\right] P(\boldsymbol{C}_{it} = 1|\boldsymbol{X}_{it}) \qquad (A3) \\ &+ \left[\rho_0 \sigma_{v_0} \{-\phi(\gamma \boldsymbol{Z}_{it})/1 - \Phi(\gamma \boldsymbol{Z}_{it})\}|\boldsymbol{C}_{it},\boldsymbol{X}_{it},\boldsymbol{Z}_{it}\right][1 - P(\boldsymbol{C}_{it} = 1|\boldsymbol{X}_{it})] \end{aligned}$$

Thus, the expected outcomes for farmers with different crop choices have been disaggregated. The expected outcome for a particular crop choice (the crop choice "1") would be:

$$E[Y_{it}|\boldsymbol{C}_{it},\boldsymbol{X}_{it},\boldsymbol{Z}_{it}] = \eta \boldsymbol{X}_{it} + \beta \boldsymbol{C}_{it} + \mu_i + \left[\rho_1 \sigma_{v_1} \{\phi(\gamma \boldsymbol{Z}_{it})/\Phi(\gamma \boldsymbol{Z}_{it})\}|\boldsymbol{C}_{it},\boldsymbol{X}_{it},\boldsymbol{Z}_{it}\right]$$

(A4)

And the expected outcome/ for the other crop choice (or the crop choice "0") would be:

$$E[Y_{it}|\boldsymbol{C}_{it}, \boldsymbol{X}_{it}, \boldsymbol{Z}_{it}] = \eta\boldsymbol{X}_{it} + \mu_i + \left[\rho_0\sigma_{v_0}\{-\phi(\gamma\boldsymbol{Z}_{it})/1 - \Phi(\gamma\boldsymbol{Z}_{it})\}|\boldsymbol{C}_{it}, \boldsymbol{X}_{it}, \boldsymbol{Z}_{it}\right] \quad \text{(A5)}$$

Here, $\rho_1\sigma_{v_1}$ represents the covariance between v_i and ε_i for farmers with the crop choice "1", $\rho_0\sigma_{v_0}$ represents the covariance between v_{it} and ε_{it} for those with another crop choice (the crop choice "0"), $\phi(\gamma\boldsymbol{Z}_{it})$ is the marginal probability of the standard normal distribution at $\gamma\boldsymbol{Z}_{it}$ and $\Phi(\gamma\boldsymbol{Z}_{it})$ is the cumulative distribution function of the standard normal distribution at $\gamma\boldsymbol{Z}_{it}$. 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:

$$E[Y_{it}|\boldsymbol{C}_{it} = 1, \boldsymbol{X}_{it}, \boldsymbol{Z}_{it}] - E[Y_{it}|\boldsymbol{C}_{it} = 0, \boldsymbol{X}_{it}, \boldsymbol{Z}_{it}] = \beta + bias\ from\ selection \quad \text{(A6)}$$

In this case, it is expected that there is a positive bias on the OLS estimates (that it overestimates the impact of the crop choice "1" 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:

$$lnL_{it}\begin{cases} ln\Phi\left\{\frac{\gamma\boldsymbol{Z}_{it}+(Y_{it}-\eta\boldsymbol{X}_{it}-\beta)\rho/\sigma}{\sqrt{1-\rho^2}}\right\} - \frac{1}{2}\left(\frac{Y_{it}-\eta\boldsymbol{X}_{it}-\beta}{\sigma}\right)^2 - \ln\left(\sqrt{2\pi\sigma}\right), & \boldsymbol{Z}_{it} = 1 \\ ln\Phi\left\{\frac{-\gamma\boldsymbol{Z}_{it}-(Y_{it}-\eta\boldsymbol{X}_{it})\rho/\sigma}{\sqrt{1-\rho^2}}\right\} - \frac{1}{2}\left(\frac{Y_{it}-\eta\boldsymbol{X}_{it}}{\sigma}\right)^2 - \ln\left(\sqrt{2\pi\sigma}\right), & \boldsymbol{Z}_{it} = 0 \end{cases} \quad \text{(A7)}$$

So in reduced form, there are two stages of regression; the first stage is the regression to estimate the probability for a farmer choosing to grow a type of crop, conditional on $\boldsymbol{Z}_{it}$; the inverse mills ratio was computed from the residuals and used in the second stage – an impact regression of the $\boldsymbol{X}_{it}$ and the IMR as an extra regressor to deflate the selection bias on productivity and poverty. The $\boldsymbol{Z}_{it}$ 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. Instruments that will satisfy the 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 seed received by the farmer and used in planting.

For the continuous crop choice variable (**C**$_3$) and its interactions, a Fixed Effects (FE) model or a Correlated Random Effects (CRE) model is will be used to address 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). However, a limitation of the FE model is that we are unable to use the time-invariant variables. This can be an issue when important variables affecting productivity such as gender are time-invariant. On the other hand, The CRE model can address endogeneity due to unobserved time-invariant factors with time-invariant variables (Wooldridge, 2010; Sheahan et al., 2013).

Appendix 2

Descriptive Stats for Wave 2

Variable	Mean	Std. Dev.	Min	Max
Primary output is cash-crop	0.18	0.30	0	1
Primary output is food-crop	0.38	0.47	0	1
Household Commercialization Index	50.02	7.35	0	79.80
ln(Total Food Auto-Consumed in HH)	10.73	1.21	1.78	13.94
ln(output)	10.77	1.57	0	16.59
ln(land)	8.76	1.55	0	15.04
ln(labour)	4.04	5.26	0	15.73
ln(inputs)	7.83	4.40	0	17.25
Age of HH Head	51.92	15.11	16	112
Marital Status of HH (Married=1)	0.73	1.70	0	1
Religion of HH Head (Christian=1)	0.51	0.54	0	1
Gender of HH Head	0.87	0.30	0	1
Number of adult males in household	1.33	0.92	0	12
Number of adult females in household	1.56	0.89	0	9
Number of dependent males in household	1.73	1.62	0	17
Number of dependent females in household	1.57	1.47	0	11
Household size	6.03	3.15	1	31
Literate (Can read and write=1)	0.47	0.49	0	1
Years of education of HH Head	3.90	3.24	1	13
Rural	0.89	0.32	0	1
Mean per capita expenditure (MPCE) in naira	210161	324383.1	4062.246	7494811

Source: Authors' calculation based on the Nigerian LSMS data for 2011 and 2013.