WPS7894

Policy Research Working Paper 7894

Heterogeneous Returns to Income Diversification

Evidence from Nigeria

Eleonora Bertoni
Paul Corral
Vasco Molini
Gbemisola Oseni

WORLD BANK GROUP
Poverty and Equity Global Practice Group
November 2016

Policy Research Working Paper 7894

Abstract

This paper investigates the impact of income diversification on farming households' welfare in Nigeria on two rounds of the Nigeria General Household Survey-Panel, namely the 2010/2011 and 2012/2013. The study finds that income diversification is the norm in Nigeria, with about 60 percent of farmers diversifying away from subsistence farming into non-farm activities and cash crops. In addition, using the panel of farmers interviewed before and after a severe drought that hit Northern Nigeria in particular in 2011, the study finds that diversification increased throughout Nigeria from 60 to 64 percent and in the North from 58 to 63 percent. The study postulates the existence of heterogeneous returns on diversification as a consequence of the drought, and estimates the returns through a non-parametric selection model using a local instrumental variable. The choice of this model is dictated by the necessity to account for both heterogeneous effects of diversification and selection bias related to households' decision to diversify. Overall, it is found that diversification positively affects consumption in Nigeria. However, who benefits the most is crucially determined by the initial conditions under which diversification is undertaken and the specific agro-climatic context in which households operate.

This paper is a product of the Poverty and Equity Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at ebertoni@worldbankgroup.org,pcorralrodas@worldbank.org, vmolini@worldbank.org, and goseni@worldbank.org.

Heterogeneous Returns to Income Diversification: Evidence from Nigeria

Eleonora Bertoni[a c], Paul Corral[a], Vasco Molini*[a], and Gbemisola Oseni†[b]

[a]Poverty-Global Practice, The World Bank, Washington, D.C., USA
[b]Development Data Group, The World Bank, Washington, D.C., USA
[c]Università di Torino, ITA

Key words: Income Diversification, Off-farm, Heterogeneous returns, Nigeria.
JEL classification: C14, D31, J43, Q12.

*Corresponding author: Vasco Molini. Telephone: +1 202 473 4891. E-mail: vmolini@worldbank.org
†We aknowledge comments from Carlo Azzarri and Alberto Zezza on previous versions of this paper.

1. Introduction

Although agriculture is the most prevalent income generating activity in many developing countries, a considerable portion of farming households do not rely solely on agriculture. If markets were complete, the rational behavior would be for households to focus on the activities where they have comparative advantage and that yield the highest returns. However, many farm households in developing countries are poor and have limited access to efficient credit, land, labor, and insurance markets. Thus households often engage in multiple income generating activities to buffer agricultural related shocks and ensure a more consistent income stream.

The belief that improvement of the lives of those residing in rural areas necessarily entails improvement of agriculture is not that uncommon (Escobal, 2001). Although the prevalence of diversification has already been proven empirically,[1] most policy that seeks to improve rural lives still tends to focus only on agriculture (Escobal, 2001). While it is still the case that many rural households in developing countries engage in agriculture, studies have shown that diversification has become more common and rural policies should reflect this. It is expected that income from sources outside of agriculture will become more prevalent, especially in Sub-Saharan Africa where agriculture may not be able to keep up with population growth (Haggblade et al., 2007). Diversification presents farming households an avenue to increase their consumption in the event that agriculture alone may not suffice. In any case, poor farming households often lack sufficient land to subsist only from agriculture (Banerjee and Duflo, 2007), and hence agriculture should just be one part of the policy agenda. Furthermore agriculture in Sub-Saharan Africa relies mostly on rainfall and thus is very susceptible to any weather shocks, and could leave farmers with unreliable income streams. Therefore diversification could be a path towards achieving consumption smoothing.

Despite the potential positive effects of diversification, households do not necessarily have equal access or benefit equally from diversification. Davis et al. (2009) mentions that access to activities outside of agriculture is not equally distributed, and that the rich may potentially benefit more from these activities. In a similar fashion, Dabalen et al. (2004) find in Rwanda that farm workers on average could considerably increase their income by participating in off-farm activities, but the gains from off-farm activities benefit mostly the non-poor.

The present paper contributes to the literature on two fronts. First, it attempts to address the above mentioned diversification heterogeneity applying a semi-parametric method which

[1] Davis et al., (2009); Winters et al., (2009), Davis et al., (2014).

allows to analyze both the differences resulting from observable characteristics as well as those from un-observable skills. Second, making use of a very rare data set for Sub-Saharan Africa - a panel of households interviewed before and after a serious climatic shock- it shows how heterogeneity is relevant also across time. Returns from observable characteristics but most importantly unobservable skills can vary substantially after a shock and standard statistical techniques fail to capture this effect.

Overall, it is found that diversification positively affects consumption in Nigeria. However, who benefits the most is crucially determined by the initial conditions under which diversification is undertaken and the specific agro-climatic context in which households operate. For example, in Nigeria, it is found that the severe drought that hit the country between 2011 and 2012 -just after the 2011 harvesting season and before the 2012 planting season (Bjerge and Fisket, 2016) - by reducing the opportunities to diversify in higher returns activities, has accentuated the heterogeneity of the returns on diversification. Figure 1 shows that in 2011 the country experienced a lower than average rainfall and higher than average temperatures resulting in a severe drought (Bjerge and Fisker 2016).

The paper is divided into six sections. The next section discusses some of the literature related to the topic of income diversification and farming. Section three provides background on Nigeria, specifically how the North and South of the nation differ, and discusses the data utilized. Section four elaborates on the methodology used, while section five discusses the results obtained. Finally, conclusions are presented in section six.

2. Diversification and agriculture

In recent literature,[2] the importance of off-farm[3] activities as a poverty reduction strategy is well documented. As much as 40-60 percent of rural incomes in Latin America, Asia, and Africa are obtained from rural non-farm employment and the sector outweighs migration and farm wage labor as a source of income and liquidity (Davis et al. 2009). Why farmers tend to diversify is also well documented and literature tends to classify the reasons into two groups: "push" and "pull" factors.[4] Households are "pushed" to diversify in order to

[2] Escobal, 2001; De Janvry et al., 2005; Reardon et al., 2007; Davis et al., 2009; Winters et al., 2010; Barrett et al., 2001.

[3] off-farm activities refers to any labor activity outside of the household's farm, including agricultural wage labor. Non-farm refers to activities which do not include agricultural wage labor (Gordon and Craig, 2001).

[4] See Barrett et al. (2001); Haggblade (2007); Davis et al. (2009); Reardon et al. (2007), among others.

minimize risk, due to low marginal productivity on the farm, to cope with shocks, or to counter liquidity constraints (Barrett et al., 2001; Reardon et al., 2007). When markets are perhaps missing, or imperfect, the household may have or want more than what the market is able to absorb or offer; it is under these circumstances that the farming decisions of the household are dependent on the household's endowments (Bardhan and Udry, 1999). These imperfect markets gives rise to liquidity constraints, low marginal productivity of labor on the family farm, and adoption of risk mitigating strategies. Thus, diversification entails taking up different activities with low correlation between each other in order to decrease the impact of a negative shock on consumption (Dercon, 2002). On the other hand, diversification which has accumulation objectives is often related to pull factors (Reardon et al., 2007). Households are "pulled" into diversification due to strategic complementarities between activities, and comparative advantage (Barret et al., 2001). Households involved in diversification driven by pull factors are usually accumulating assets; consequently, they see their incomes rise.

Studies finding a positive relationship between diversification and agricultural production are numerous. Off-farm income is found to have a positive and significant effect on agricultural production in rural Nigeria (Babatunde, 2012); Oseni and Winters (2009) find that in Nigeria, participation in non-farm activities has a positive and significant effect on the expenditure on agricultural inputs. Not all the studies examining the relationship between diversification and agricultural production found, however, positive links. Capital intensive activities which are more profitable tend to attract the rich, and often include activities demanding high human capital (Reardon et al., 2007). Consequently, these activities may potentially move households away from agriculture. Kilic et al. (2009) find that this is the case among rural households in Albania, where households prefer to move out of agriculture as opposed to investing in it. Pfeiffer et al. (2009) find for Mexico that diversification is negatively related to agricultural output; all else being equal, the greater the access to off-farm opportunities, the less likely the household will remain in agriculture.

As mentioned before, non-farm incomes can play an important role in reducing poverty but also income inequality. Results from a study conducted in the Hubei province of China show that in the absence of non-farm employment, rural poverty and inequality would be worse (de Janvry et al., 2005). The study also suggests that those farmers who remain in agriculture possess non-observable characteristics associated to higher productivity in agriculture. When linking off-farm activities to more general distributional issues, the question arises as whether returns of off-farm activities are homongenous along the income distribution or it can be postulated some sort of heterogenity. A corollary of this is clearly whether, as consequence of external shocks, the returns can modify and of how much.

On the heterogeneity across the income distribution, Dabalen et al. (2004), for example find that in Rwanda the benefits of diversification tend to diminish as it approaches the lower end of the distribution; authors posit that fixed costs and entry barriers into certain activities play a role in this results. Likewise, Reardon et al., (2007) find that differences in asset endowments between rich and poor households lead them to very different diversification paths. In general, poorer households, due to insufficient endowments, may pursue low risk diversification activities which offer low returns (Barrett et al., 2005) or initiate businesses that can only accommodate family members as employees (Banerjee and Duflo, 2007); and in many cases, this initiative is in response to a shock (ibid). On the other hand, richer households are often involved in low risk activities which present higher payoffs, but require more endowments (Reardon et al., 2007). Consequently, education, land ownings, and number of family members are found to be significant determinants of participation in off-farm activities[5].

This research purports to show that, in general, diversification is beneficial to participating households. The gains, however, are not expected to be the same for all parties involved as the returns to diversifying are potentially heterogeneous across the population and, for the same households, across different systemic conditions (e.g after a shock). This heterogeneity and how to model it is discussed more in depth in the methods section.

3. Country context and data

Nigeria is divided into 37 states, six geo-political zones[6] and, in addition, the country is also commonly divided into two major regions, the North, and the South. The majority of the Muslims in Nigeria are concentrated in the North, while Christians are more heavily concentrated in the South (Falola and Heaton, 2009). This divide between the North and South is not only evident in religion, but also in social-economic and agro-ecological characteristics (World Bank, 2016); the North of Nigeria is partly covered by the Sudan-Sahel ecological zone; consequently, the North suffers from worse climate variations compared to the South (Ekpoh et al., 2011). In the 1970s, oil, which is mainly found in the Niger-delta, in Southern Nigeria, became the most important commodity of the country (Falola and Heaton, 2008). While oil may be the main commodity for the country, agriculture is the main source of

[5]Escobal (2001), de Janvry et al. (2005), Oseni and Winters (2009).

[6]North West, North Central, North East, South West, South East, and South South.

employment for the population. About 60 percent of the country's population relies on agriculture for their livelihoods (Oseni and Winters, 2009). The majority of industries and cities of the country are located in the South, while the North is more rural and less densely populated (Oseni and Winters, 2009), thus, it is not surprising to find a socio-economic divide in the country. In addition, poverty is found to be much more prevalent in Northern Nigeria than it is in the South and concurrently, the majority of the country's middle class is concentrated in the Southern regions (Corral et al., 2015).

The analysis uses data from two rounds of the Nigeria General Household Survey Panel (GHS- Panel) conducted by the National Bureau of Statistics in 2010-2011 (round 1) and in 2012-2013 (round 2). The GHS-Panel which has a sample of 5,000 households is part of the Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) project which aims to improve agricultural statistics in Sub-Saharan Africa. The GHS-Panel provides information on household agricultural activities, household's human capital, other economic activities, and access to services and resources. Moreover, since the surveys were administered in two visits, namely: post-planting (Aug-Oct 2010; Sept-Nov 2012) and post-harvest (Feb-Apr 2011; Feb-Apr 2013), it contains data on farm and non-farm activities for the corresponding season. High resolution precipitation records for the 2001-2009 periods are derived from the National Oceanic and Atmospheric Administration (NOAA) climate prediction center. The latter are geo-referenced amounts for the households in the GHS-Panel sample.

The observational units for this analysis are Nigerian farming households that are present during both survey rounds,[7] in total we have 2,232 farming households [8]. The outcome of interest is a dummy variable that equals 1 if the household diversifies its sources of income. In order to construct this indicator, various modules from the GHS-Panel are relied upon, specifically, the labor and non-farm enterprise modules. These modules permit identification of individuals and the types of activities they are involved in. According to this information, a household is classified as diversified if any household member is involved in any labor activity outside of the household's farm. More in details, if any individual within the household reports working outside the household's farm, the household would be classified as diversified. Additionally, if anyone inside the household owns a non-farm enterprise which is in operation, the household is classified as diversified. Finally, if the

[7]Farming households are defined as all households who own land and reported planting crops during the post-planting season.

[8]Households for which land measures were not available were removed. Additionally households in Lagos and Abuja were also removed due to their urban nature.

household is involved in planting any commercial crops[9] that would also be considered a diversified household.

Table 1 displays the percentage of households involved in the different income generating activities, and how these activities are combined. There are four main activities which we consider: non-farm enterprises (NFE), non-farm wage labor, farm wage labor (outside one's household's farm), and commercial crop production. A number of preliminary findings are worth noting. First, the majority of Nigerian farming households diversify according to the aforementioned classification. On average, depending on the GHS-Panel survey round and location, 60-65 percent of households have a portfolio that is not solely based on food crop farming. Second, NFE diversification activities constitute the lion's share for all non-farm diversification activities. Diversification by pursuing commercial crop production (which is more prominent in the South than in the North) follows after NFE activities. The other two options are quite marginal. Households also combine more than one activity. The most common combination is commercial crop production and NFE ownership.

Table 1 also illustrates a sharp variation in the number of diversifying households between the two survey rounds; Northern households in particular increased their participation in NFE activities by close to 10 percent. A plausible explanation for this increase is the climatic shock faced by Nigeria in 2011, (Figure 1). The country, in particualr in the Northern area, faced severe droughts that started just after the Post-Harvest 2011 survey data were collected and lasted until a few months before the new survey round (Post Planting 2012) started (Bjerge and Fisker, 2016). The increase in diversification observed can be considered a a response to these unexpected climatic events.

The welfare measure chosen for the analysis is expenditures per capita. We expect diversified households to have higher per capita expenditures. The reason for this is that these households, as discussed above, should have lower income volatility, and should be less susceptible to shocks. Table 3 presents summary statistics by survey round, for the whole country, and by region. Diversified households have significantly higher levels of per capita expenditures in the first survey round. Additionally, it is evident that average expenditures in the North are considerably lower than those in the South in both survey rounds. Non-diversifying households in the South appear to have on average higher expenditures than Northern households which diversify.

Education of the household head, is significantly greater (by more than a year) among diversifying households. Interestingly, due to lower levels of education in the North, during

[9] Cotton, sesame, ginger, Arabic gum, cocoa, kolanut, palm oil.

both rounds, the mean years of education among non-diversifying households in the South is almost the same as that of diversifying households in the North. Education, as explained above, allows households to diversify in more lucrative activities.

Among diversifying households the proportion of male headed households is significantly greater than the propotion of male headed households among non-diversifying households. Male headed households are more likely to diversify than female headed ones. This likelihood is particularly evident in the South, where the number of female headed households in the sample is higher than the national average; among diversifiers, only 20 percent are female headed, while among non-diversifiers, the number increases to close to 30 percent. Additionally, diversified households have on average larger household sizes (more adult males, females, and dependents). In the North, diversifying households have on average more dependents as seen in both rounds where the diversified households have on average 4 dependents while the non-diversified households have 3. Family size is also a factor that may push households towards diversification. If we assume concave agricultural production, greater family sizes, all else being equal, lead to lower family labor productivity on the household's land, and could push households to seek outside options.

Diversifying households tend to reside in urban areas, or in proximity of urban areas. In the South, 26 percent of diversifying households reside in urban centers compared to around 11 percent for non-diversifying households. In the North, the difference is equally significant where about 5 percent of non-diversifying households and between 10 and 11 percent of diversifying households reside in urban areas. These figures support the argument that better infrastructure may pull households towards diversification, or at least present more lucrative opportunities. Furthermore, the infrastructural constraints[10] faced by diversifying households are on average significantly lower.

Overall, the data reveal that households who diversify are considerably different than those who do not. Households who diversify have on average greater years of education. Additionally, diversifying households have more household members on average, yet per capita expenditure is also greater. The following section discusses the proposed method used to analyze these two groups, and the impact of diversification on household's consumption.

[10]The index is constructed using principal-component analysis (Filmer and Pritchett, 2001). The index takes into consideration the time necessary to reach the closest bank, bus, market from the center of the village. Increasing values are used to denote distances larger than a certain threshold. Additionally, actual distances to road, market, and population centers are used. Therefore higher values of the index denote higher mobility constraints faced by the household.

4. Methodology

Panel IV

The starting point of our analysis is assessing the impact of diversification on Nigerian farmers' welfare. Given this, we firstly take advantage of the panel nature of our data and implement a fixed-effects (FE) estimation of an individual-effects model with two error components of the form[11]:

$$Y_{it} = X'_{it}\beta + \alpha_i + \varepsilon_{it} \tag{1}$$

Where α_i is the time-invariant component of the error term and ε_{it} is the time-variant component. In the FE model, the regressors x_{it} are allowed to be correlated to the time-invariant component of the error (α_i), while assuming independence between x_{it} and the idiosyncratic error ε_{it} (Cameron and Trivedi, 2010). The outcome of interest y_{it} is per capita expenditure of household i at time t, where $t = 2$. The vector of regressors x'_{it} includes variables that might influence a household's level of per capita expenditure such as an indicator of whether the household is diversified; an infrastructural constraint index; some caracteristics of the household head such as gender, age, marital status and years of completed education; some variable describing the demographic composition of the household (i.e. the number of adult male and female and the number of dependants); the area of residence (whether urban or rural); the hectares of land owned by the household; and some distance measure between the household and major community sites (i.e. market, population center and border post).

When estimating the FE model, we have to take into account that the choice to diversify may be endogenous. Un-observables skills such as ability that lead a household to choose to diversify, may be conditioning the considered outcome (Angrist and Pischke, 2008). When this is the case, x_{it} will no longer be independent from the idiosyncratic error ε_{it}, resulting in incosistent FE estimators. An instrumental variable approach is, thus, required. We need to

[11] The Hausman test rejects the null that the random-effect (RE) estimator provides consistent estimates, so we use a FE model.

assume the existence of instruments z_{it}that are correlated with x_{it} and uncorrelated with ε_{it} in all periods. This assumption is confirmed by the result of the Wu-Hausman F-test which rejects the null hypothesis that the dummy for diversification is exogenous.

The instrument we propose is a measure of rainfall variability computed as the average of the mean yearly precipitation during the rainy season from 2001 up to 3 years before the survey interview for each household[12]. In building the rainfall measure we are taking into account the different timing of the rainy season according to the different climatic zones of the country, that is: June/July to September in the North, and March/April to October in the Central and Southern areas (Mustapha and Arshad, 2014).

Panel IV results

The results from the first-stage of the 2SLS estimation are reported in Table 4. The instrument we propose satisfies the relevance condition($Cov\left(z,x\right)\neq 0$) in that it is significantly negatively correlated with the decision to diversify. Moreover, the value of the Weak Identification F-statistics of 12.45 relieve concerns regarding a weak instrument bias [13]. The rationale behind choosing this instrument is that a farmer's decision to diversify taken at time t it is likely to be correlated to the "rainfall memory". Studies on farmers' memory of precipitations' patterns in West Africa and the Sahel confirm that it dates back to several years, especially if the event took place during the rainy season (Tschakert et al., 2010; Simelton et al., 2013). At the same time, a farmer's per capita expenditure at time t it is hardly affected by past rainfall events unless, of course, there was a big shock; in fact, according to the available data (table 2), with the notable exception of 2011, there is no sign of such an event.

Results from the second-stage of the 2SLS estimation reported in Table 5 suggest that income diversification is positively influencing Nigerian farmers' welfare. Nevertheless, as discussed in section 2, it is possible that agents benefit differently from diversifying and thus base their decisions on their anticipated returns from it (Radchenko, 2014). In the specific context of Nigeria, we believe that the events that took place in-between the two survey rounds, namely the 2011 droughts and to some extent -but this is very difficult to model and restricted to a

[12]The selection of 2001 as a starting point is dictated by the availability of the geo-referenced high resolution precipitation data from NOAA climate prediction center.

[13]Referring to a common threshold employed in literature (Staiger and Stock, 1977), we consider a value of the F-statistic above 10 from the test of joint significance of the instruments in the first-stage regression as necessary to conclude that our instrument is sufficiently strong.

few states.[14] The increase in Boko Haram's activity in the North East might have fostered differences in farmers returns to diversification. In a nutshell, these events could have pushed low-skilled farmers to diversify into non-farm activities already occupied by more skilled ones and resulting in very limited gains for them.

In the presence of heterogeneous responses to diversification on which individuals act (select into diversification), a variety of "effects" of diversification can be defined and conventional intuitions about instrumental variable estimators break down (Carneiro, Heckman, and Vytlacil 2000). A more detailed discussion of the inconsistency of the OLS and traditional IV estimators of the effect of diversification on the outcome of interest is presented further in this analysis.

For all these reasons, we take a step forward in our analysis and focus on modeling the major constraints to the identification of the effect of diversification on Nigerian farmers' wellbeing, which are:

1) Missing counterfactual (Blundell and Dias, 2009): we only observe per capita expenditure in one or the other state (diversified/non-diversified), not both simultaneously for each farmer.

2) Selection into diversification (decision to diversify is endogenous): farmers may self-select into diversification based on unobservables. These same unobservables may, in turn, have an impact on per capita expenditures so that the error terms in the per capita expenditure and selection equation can be correlated.

3) Essential Heterogeneity: agents base their decision to diversify on these unobservables as different farmers might benefit differently from diversification. The gap in per capita expenditures between diversifying and non-diversifying for a given farmer are individual-specific and depend on different returns to both observables (human capital; gender; physical constraints; etc.) and unobservables (ability; entrepreneurship; family background; shocks). Most importantly, since the returns to undobservables can differ across the distribution and can change over time as consequence of a shock as we document henceforth, using a panel Fixed Effect estimation leads to loose crucial information. The comparison of the two surveys taken separately and the use of non-parametric techniques captures rather than disregard this heterogeneity.

[14] The full escalation of the insurgency, with mass displacement and heavy fighting between insurgents and Nigerian Army starts de facto in June 2013, few months after our data collection. Proof that, during 2013, households were only marginally affected by the conflict is that in the third round of the panel survey (2015/2016), most of the households residing in Borno and Adamawa states visited in 2013 could not be re-interviewed due to the poor security situation.

Essential Heterogeneity

To identify the effect of income diversification on Nigerian farmers' welfare given these identification constraints, a non-parametric selection model is estimated by local instrumental variables (Heckman and Vitlacyl, 1999; Heckman and Vitlacyl, 2001; Carneiro et al., 2001). The methodology employed here accounts for heterogeneity and selection. It has been used extensively in the analysis of the returns to schooling, nonetheless, in this instance, it is applied to the analysis of diversification among farming households. The methodology which is described below is based on the methodology described in Heckman and Li (2004), Carneiro et al. (2011), and Radchenko (2014).

Households can be separated into those who diversify and those who do not diversify, $D_i = 1$, and $D_i = 0$ respectively. Given these two groups, the returns in each option are given by:

$$\ln Y_{i0} = \beta_0 X_i + U_{i0} \; if \; D_i = 0; \; E\left(U_{i0} | X_i\right) = 0 \tag{2}$$

$$\ln Y_{i1} = \beta_1 X_i + U_{i1} \; if \; D_i = 1; \; E\left(U_{i1} | X_i\right) = 0 \tag{3}$$

Y_{i0} is the returns for households who have chosen to not diversify, $Y_{i1}$is the returns for those households who have chosen to diversify. X_i is a vector of observable characteristics, that might affect the household's expenditure per-capita and includes household characteristics and,characteristics of the head of the household, among others. Finally, U_{i0} and U_{i1} are the error terms.

The household's choice to diversify is given by a standard discrete model of the form:

$$D_i^* = \theta Z_i - U_i \tag{4}$$

In equation (5) $D_i^*$is a latent variable, and may be considered a reduced form model for diversification. A household will choose to diversify if it expects positive returns given unobservables, where:

$$D_i = \begin{cases} 1 & if \; U_i < \theta Z_i \\ 0 & otherwise \end{cases} \tag{5}$$

The equation can be re-written as:

$$D_i = \begin{cases} 1 & if\ v_i < P(Z_i) \\ 0 & otherwise \end{cases} \tag{6}$$

Here Z is a vector of variables which may include some X, and are believed to affect the diversification decision[15], and $P(Z)$ indicates the probability that household i will diversify. Since U_i can have any density, then by construction $v_i \sim Uniform\,(0,1)$(Heckman and Li. 2004). If $P(Z_i) = F_u(\theta Z_i)$ and $v_i = F_u(U_i)$, then $F_u$is a cumulative distribution of U_i (Carneiro et al., 2011). The smaller v_i is, the more likely the household will diversify.

As noted above, the choice to diversify may be endogenously taken by households. Therefore, at least one instrument must be included in the selection equation (Carneiro et al., 2011). The instrument to be used is a variable denoting rainfall shocks.[16] The assumption is that if there are two households with similar characteristics, then, the rainfall shock measure is correlated to the outcome considered only through the effect it has on the choice to diversify.

The observed per-capita expenditures are the following:

$$\begin{aligned} Y_i &= D_i \ln Y_{i1} + (1 - D_i) \ln Y_{i0} && (7) \\ &= D_i\left(\beta_1 X_i + U_{i1}\right) + (1 - D_i)\left(\beta_0 X_i + U_{i0}\right) && (8) \\ &= \left[\left(\beta_1 - \beta_0\right) X_i + \left(U_{i1} - U_{i0}\right)\right] D_i + \beta_0 X_i + U_{i0} && (9) \\ &= \delta_i D_i + \gamma_0 X_i + U_{i0} && (10) \end{aligned}$$

where $\delta_i = \left[\left(\beta_1 - \beta_0\right) X_i + \left(U_{i1} - U_{i0}\right)\right]$ is the heterogeneous return to diversifying for each household i. This δ_i is the parameter of interest. As it can be seen, the parameter has two sources of heterogeneity. Observed heterogeneity comes in the form of $\left(\beta_1 - \beta_0\right) X_i$, and unobserved heterogeneity in the form of $\left(U_{i1} - U_{i0}\right)$. Therefore, if $\left(\beta_1 \neq \beta_0\right)$, there may be observed heterogeneity, additionally, if $E\left[X|D\right] \neq E\left[X\right]$, there may also be heterogeneity (Radchenko, 2014). The other source of heterogeneity comes from unobservables $\left(U_{i1} - U_{i0}\right)$. If the considered household is able to discern what its idiosyncratic gains are from diversifying this will play a role in their decision to diversify and thus $U_{i1} \neq U_{i0}$ (Heckman and Li, 2004).

[15]The variables included are those which have been noted to be relevant by Escobal (2001), Lanjouw (2001), de Janvry et al. (2005), and Oseni and Winters (2009).

[16]This is discussed in the data section.

The objective is to estimate the average treatment effect (ATE) of diversifying ($E[\delta_i|X_i]$), as well as the average treatment effect on the treated (diversified households) ATT, and on the non-treated (non-diversified households), TNT. The potential endogeneity mentioned above comes from the possibility that ($E[U_{1i}|D_i = 1] - E[U_{0i}|D_i = 0]$) does not equal zero, therefore, ordinary least squares estimates of ATE will be biased:

$$\begin{aligned}
\hat{\beta}_{ols} &= E(Y_i|X_i,\, D_i = 1) - E(Y_i|X_i,\, D_i = 0) && (11)\\
&= E(\beta_1 - \beta_0)\, X_i + (E[U_{1i}|D_i = 1] - E[U_{0i}|D_i = 0]) && (12)\\
&= \bar{\beta}(X) + E(U_{1i} - U_{0i}|D_i = 1) + E(U_{0i}|D_i = 1) - E(U_{0i}|D_i = 0) && (13)
\end{aligned}$$

Where $E(U_{1i} - U_{0i}|D_i = 1)$ is the sorting on the gain and comes from one's anticipation of the gains from diversifying, and $E(U_{0i}|D_i = 1) - E(U_{0i}|D_i = 0)$ is the selection bias (difference in non-diversification unobservables for thise who do diversify and those who do not). Furthermore, traditional instrumental variable methods will also fail to identify ATE, ATT, and TNT. Finding a variable which is correlated to the choice to diversify, and at the same time is unrelated to the error term in Equation (11), is not sufficient to obtain the parameters of interest (Heckman and Li, 2004). This last statement is made clear by rearranging Equation 11:

$$\begin{aligned}
Y_i &= (\beta_1 - \beta_0)\, X_i D_i + \beta_0 X_i + [(U_{1i} - U_{0i})\, D_i + U_{0i}] && (14)\\
&= (\beta_1 - \beta_0)\, X_i D_i + \beta_0 X_i + \varepsilon_i && (15)
\end{aligned}$$

If an instrumental variable is to be used, it would instrument for the choice to diversify, D. The instrument would be correlated to D, and it should not be correlated to ε. However, this latter requirement will unlikely be met since ε is a function of D. Nevertheless IV could work under some specific circumstances[17]. The first instance where IV would work is when $(U_{1i} - U_{0i}) = 0$, i.e. there is no heterogeneity and selection bias. The second instance is the case when there is unobserved heterogeneity, but it is independent from the choice to diversify, i.e. $(U_{1i} \neq U_{0i})$ but $(U_{1i} - U_{0i})$is independent from D.

In the presence of both selection and heterogeneity, OLS and IV will both convey inconsistent estimates of the effect of diversification on the outcome of interest, and do not identify any

[17] Assuming the instrument is related to D, but not to U_0.

relevant treatment effect. Given this, in order to obtain ATE, ATT, and TNT it is necessary to first estimate the mean return to diversification through the marginal treatment effect (MTE). The marginal treatment effect gives a simple characterization of heterogeneity in returns by measuring the returns of diversification for households with different levels of observables (X), and unobservables (v) (Carneiro et al., 2011). The MTE may be obtained via local instrumental variables (Heckman and Li, 2004). Taking expectations of Equation 15:

$$\begin{aligned} E\left[Y|X,P(Z)=v\right] &= E\left[\left(\beta_1-\beta_0\right)X_iD_i+\beta_0X_i+\left[\left(U_{1i}-U_{0i}\right)D_i+U_{0i}\right]\right] && (16)\\ &= \left(\beta_1-\beta_0\right)Xp+\beta_0X+\left(U_1-U_0\right)p+U_0 && (17)\\ &= \left(\beta_1-\beta_0\right)Xp+\beta_0X+K(p) && (18)\end{aligned}$$

The difference between intercept terms of the equation cannot be estimated separately from $K(p)$ (Carneiro et al., 2011). Therefore, if we define the intercept terms to be α_1 and α_0,[18] we get:

$$K(p)=\left(\alpha_1-\alpha_0\right)P+E\left(U_1-U_0|P(Z)=P\right)P+E\left(U_0|P(Z)=p\right) \tag{19}$$

In this context $K(p)$ is a function of p which can be estimated via the partially linear regression estimator illustrated by Robinson (1988)[19]. Finally MTE is:

$$\begin{aligned} MTE\left(X_i=x,v_i=P_i=p\right) &= \frac{\partial E\left[Y_i|X_i=x,P_i(Z_i)=p\right]}{\partial p} && (20)\\ &= \left(\beta_1-\beta_0\right)X+\frac{\partial K(P)}{\partial P} && (21)\end{aligned}$$

A simple test for the heterogeneity coming from unobservables is whether $\frac{\partial K(P)}{\partial P}$ is flat or not (Carneiro et al. 2011). If it is flat it implies that heterogeneity from unobservables is not relevant in the model.

Finally, in order to obtain ATE, ATT, and TNT the MTE must be integrated across the entire population (Radchenko, 2014):

[18]In Equation 14, α_0 intercept is considered within β_0X

[19]For a detailed description on how $\frac{\partial K(P)}{\partial P}$ is obtained refer to Carneiro et al., (2011), or Radchenko (2014).

$$ATE(X) = \int_0^1 MTE(x,v)dv = (\beta_1 - \beta_0)\, x + \int_0^1 MTE^v dv \tag{22}$$

where MTE^v is the portion corresponding to unobservables.

The effect of diversifying for those who do diversify is equal to the treatment on the treated:

$$TT(X) = \int_0^1 MTE(x,v) w_{tt} dv = (\beta_1 - \beta_0)\, x + \int_0^1 MTE^v w_{tt} dv \tag{23}$$

where w_{tt} weighs more heavily those observations which are more likely to be diversified and is equal to:

$$w_{tt} = \frac{P(P(Z) > v|X = x)}{\int P(P(Z) > v|X = x)dv} \tag{24}$$

Finally, the effect of diversifying on those who do not diversify is equal to the treatment on the non-treated:

$$TNT(X) = \int_0^1 MTE(x,v) w_{tnt} dv = (\beta_1 - \beta_0)\, x + \int_0^1 MTE^v w_{tnt} dv \tag{25}$$

where w_{tnt} weighs more heavily those observations who are less likely to be diversified and is equal to:

$$w_{tnt} = \frac{P(P(Z) < v|X = x)}{\int P(P(Z) < v|X = x)dv} \tag{26}$$

All treatment effects shown are decomposed into a component due to observables, and a component due to unobservables. Finally, the standard errors are estimated using bootstrap methods following Radchenko (2014).

5. Results

Selection model

As indicated in the previous section, the first step of the semi-parametric methodology employed is estimating the probability of a household diversifying. The probability is modeled through a logit model (Maddala, 1983). Figures 2 (2010/11 round) and 3 (2012/13 round) present the propensity scores for diversifying and non-diversifying households. As is evident from the graphs there is almost full common support in our sample. The results for this model are presented in Table 5, with columns labeled accordingly for each round of the GHS-Panel survey. The values reported are the marginal effects from the logit model.

The characteristics of the head of the household, such as marital status and years of completed education, are significantly related to the diversification choice. The head's years of education is positive and significantly related to the probability of a household diversifying in both years. This is possibly related to education allowing households to overcome barriers to entry, and possibly have access to a broader range of options on which to diversify. The gender of the head is significant (1%) only for 2012/13, while the head's age does not appear to be significant in both models.

As expected, higher infrastructural constraints limit the opportunities to diversify for farming households (e.g. if the household is very distant from a bus stop or a market center, the opportunity cost of diversifying may be too high for them to do so). Thus, it seems to be the case that improved infrastructure may "pull" households towards diversification, or at least provide more options. The excluded instrument used for the participation equation is statistically significant. An increase in the average of the mean yearly precipitation during the rainy season from 2001 up to 3 years before the survey interview for each household is negative and significantly related to diversification. Households who experienced higher than average rainfall during the rainy season in 2001-2007 (round 1) and 2001-2009 (round 2) were less likely to choose to diversify. Rainfall shocks potentially "push" households into diversification in order to cope with the shocks itself and to ensure smooth consumption patterns. In this instance, a positive rainfall shock seems to make households more likely to rely solely on agriculture, a finding which is consistent with the hypothesized effect of the 2011 drought previously discussed.

As mentioned previously, one of the potential reasons of households' decision to diversify is due to decreasing marginal labor productivity in agriculture. Supporting this argument, we find that an increase in family members significantly increases the likelihood of diversification, provided that the additional member is part of the work force. Otherwise, as suggested by the negative sign on the dependent variable, an additional member in dependency age will trigger the opposite effect, all else being equal. However, the amount of land owned by the household is not significantly related to the likelihood of diversifying. This is probably due to the method of land acquisition in Nigeria where most households acquire land through family or community distribution as opposed to outright purchase. Additionally, we do not control for the quality of land owned which may prove to be more relevant.

The results suggest that households with more educated heads are more likely to diversify[20]. Additionally households with more members which are part of the labor force, all else being equal, are significantly more likely to diversify. The sign and significance of the infrastructure index provides evidence that households that reside in areas with better than average infrastructure are more likely to diversify. Finally, the chosen instrument is also significant. The sign and significance of the coefficients for rainfall variability, gives credence to the idea that households may seek to diversify due to the uncertainty faced in rain-fed agriculture.

Treatment effects - TT, ATE, TNT

The main outcome of the chosen methodology is presented in Table 6, and Figures 2 and 3.

In 2010/2011 returns from diversification are positive for everyone, especially for those who do diversify (Table 6 column 1); the average per capita expenditure anticipated gain among those diversifying (TT) is 82 percent, while it is 37 percent for those who do not diversify (TNT) and 60 percent for the average randomly selected farmer (ATE). Also, returns from diversification given observables are positive for everyone, especially for those already diversifying (Table 6 column 2).

Although at the margin everyone would have benefited from diversification, these results suggest that those not diversifying choose mainly on the basis of their observable characteristics. This reveals a potentially segmented market, one where agents who would expect a gain from participation do not participate (Radchenko, 2014). Diversifying appears to be not so much a voluntary decision but a segment for the highly educated and “less constrained” farmers which can successfully exploit diversification opportunities as a consequence of observable factors (e.g. household’s composition and human capital, initial endowments, infrastructural

[20]However, education is potentially endogenous in this instance.

constraints, etc.). The 2010/2011 results suggest, therefore, the presence of barriers to entry into diversification that clearly prevent more constrained farmers from doing so. Figure 4 graphically displays this finding. Recalling equation (5) from the methods section, lower values of unobserved heterogeneity v correspond to a higher likelihood of diversification due to unobservables (Radchenko, 2014). MTE^vorthogonal to unobserved heterogeneity v implies no heterogeneity in returns due to unobservables that is, at the margin, diversified and non-diversified farming households would have benefited the same from diversification. Figure 4 graphically represents the homogeneity of the returns given unobservables showing that the MTE^v is constant and positive along the propensity score distribution (except for the tails, corresponding to farmers having particularly high or low propensity to diversify).

The comparison of the MTE^vshapes in Figure 4 and Figure 5 highlights the changes in the returns on unobservables occurred between the two survey rounds. More in details, while the MTE^v in 2010/2011 was flat, in 2012/2013 it becomes negative and declining in v. This implies that a) based on unobservables, the decision to diversify in 2012/2013 would yield highly negative marginal returns to everyone; b) households with lower probability to diversify (higher values of v), if deciding to diversify, would be the most penalized by this decision: total returns from diversification are not significantly different from zero for those who do not diversify (Table 6 column 4). The total anticipated gains are now positive only for the average farmer and for those who do diversify; a result that is exclusively driven by the returns on observables. In this instance, the gains from diversification are mainly driven by observable factors and competitive mechanisms are driving the allocation of farmers in the different economic activities.

We argue that the 2011 drought and, to a lesser extent, the incipient conflict may have left many households struggling. As shown in Table 1, more households were pushed into diversification in the second survey round. This diversification was mainly towards low-risk and low-returns activities such as non-farm enterprises (in particular in the North). Farming households sought to cope with the shock by massively entering into non-farm activities, thus saturating the market. This saturation led to a decline in the returns to unobservables (e.g. skills, entrepreneurial ability, connections, etc.) since households were possibly forced to opt for activities which, under normal conditions, they would not consider.

The higher returns on observables in column 5 (Table 6) also suggests that having strong endowments in times of crisis helps to buffer shocks quite substantially. More precisely, having proper infrastructure to access opportunities and having sufficiently able members in the household to engage in activities outside of farming becomes important. All in all, the returns on observable override the unobservable losses, but only households which diversify are better off and have positive total returns. Therefore, in times of crisis diversification becomes an important factor, yet, not every household is able to benefit from it; households choosing to not diversify during round 2 of the survey, would not have gained from diversifying.

This study has uncovered important results that would have been disregarded by the panel analysis alone. Firstly, to estimate the impact of income diversification in the presence of identification constraints such as selection and essential heterogeneity, the use of a non-parametric estimation is warranted. Second, we can expect returns on diversification to be different across an income distribution and, in presence of a shock that occurs between two survey rounds, between the two survey rounds. Third and most important, a shock can modify the returns on observables but also those on unobservables; in the case presented here the differences in returns were mainly driven by changes in the return on unobservables that a standard estimation process would have not captured.

Semi-parametric model results

In Table 7 the semi-parametric results for the natural log of per-capita expenditure are presented. As indicated in the methods section, a selection of control variables is interacted with the probability of diversifying obtained from the logit estimation (presented in Table 5). The coefficients on these interactions are presented in the column under $(\beta_1 - \beta_0)$. These coefficients provide the difference in returns between diversified and non-diversified farming households. The column under β_0 reports results for non-diversifying farming households.

Years of education in 2010/11 is significant and positive for non-diversifying households, while the difference in coefficients between groups is not significant. In 2012/13 years of education is not significantly different from zero for non-diversifying households, while for diversified households it is significantly greater than that of the non-diversified. Nevertheless, the variable in all likelihood is endogenous and thus the coefficient is potentially biased (Carneiro et al., 2011). Additionally, only the education of the household head is considered, which may not be the best indicator since perhaps younger household members are the ones working off the farm while the head tends to work on the household's farm.

The gender of the household head is slightly significant for non-diversifiers in 2012/13. Increases in household members is negative and significantly related to per capita expenditure for the group of non-diversifying households, while it yields positive returns for those who do diversify for both rounds. Finally, the effect of land holdings per capita is possibly absent for both groups, most likely as we do not control for land quality.

The results imply that diversifying households may obtain higher consumption due to observable factors. Especially due to households' head characteristics and the household's composition. non-diversifying households would face decreased per-capita expenditures if the number of household members increased, all else being equal. This effect is reversed for diversifying households, most likely due to the link between more members and the likelihood of diversifying.

7. Conclusions

The study began with the premise that diversification among Nigerian farmers is predominant: about 60% of farmers do diversify away from subsistence farming into non-farm activities and cash crops. Using a panel of farmers interviewed before and after a severe drought that hit in particular Northern Nigeria in 2011, we find that diversification increased substantially: throughout Nigeria from 60 to 64 percent and in the North from 58 to 63 percent.

In the course of the analysis we investigate this change in household activities and postulate that diversification brings considerable gains to those involved, and may bring gains to the population at large. However, we find support to the argument that the conditions under which a household diversifies matter substantially: gains are heterogeneous throughout the income distribution and, after the drought, returns on unobservable characteristics turn form positive to negative.

Before the shock, overall gains from diversification tend to be higher among those already diversifying and lower but positive among those not diversifying. This difference is mainly driven by differences in returns to observable characteristics such as human capital, household's composition and access to infrastructures; returns on unobservables characteristics such as ability and entrepreneurship are always positive and rather homogeneous along the propensity score (except for the tails corresponding to farmers having higher or low propensity to diversify).

On the other hand, after the shock, we find that those who are the least likely to diversify, due to the changes in the returns on unobservables, would not obtain any gain by diversifying. We purport that the drought and, possibly, the intensification of the conflict in the North of the country, may have damaged many agricultural crops, pushing many low skilled households to opt for activities which, under normal conditions, they would not consider. In this way, many activities that had few barriers to entry would become saturated. The most common method of diversification among Nigerian farming households is on non- farm enterprises. These household enterprises are often characterized by low productive activities that require limited skills (Haggblade, 2007); nevertheless, these are often the only viable solution given the lack of alternatives (Fox and Sohnensen, 2012).

Policy interventions should focus on ways to improve the productivity of non-farm enterprises. This may include increased vocational training or apprenticeships. Improved infrastructure and access to credit could also be beneficial to non-farm enterprises. This does not mean agriculture should be disregarded altogether, since it still may provide the bulk of income for many households. Also, poor landless households are often ignored and the gains they obtain from interventions in agriculture are possibly limited to lower food prices. Any drastic shift away from agriculture, thus, may have detrimental effects on the poor due to increased food prices. Rural policies, therefore, should not focus on agriculture issues alone but be holistic in nature in order to address the different activities farming households engage in, especially in time of need.

References

Akpoveta, V. O., Osakwe, S. A., Ize-Iyamu, O. K., Medjor, W. O., & Egharevba, F. 2014. Post Flooding Effect on Soil Quality in Nigeria: The Asaba, Onitsha Experience. Open Journal of Soil Science, 2014.

Angrist, Joshua D and Jörn-Steffen Pischke. 2008. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton university press.

Babatunde, Raphael O. 2012. "Assessing the Effect of Off-Farm Income Diversification on Agricultural Production in Rural Nigeria."

Banerjee, Abhijit V and Esther Duflo. 2007. "The Economic Lives of the Poor." The journal of economic perspectives: a journal of the American Economic Association, 21(1), 141.

Bardhan, Pranab and Christopher Udry. 1999. Development Microeconomics. Oxford University Press.

Barrett, Christopher B; Mesfin Bezuneh and Abdillahi Aboud. 2001. "Income Diversification, Poverty Traps and Policy Shocks in Côte D'ivoire and Kenya." Food policy, 26(4), 367-84.

Barrett, Christopher B; Mesfin Bezuneh Clark; Daniel C Clay and Thomas Reardon. 2005. "Heterogeneous Constraints, Incentives and Income Diversification Strategies in Rural Africa." Quarterly journal of international agriculture, 44(1), 37-60.

Bjerge, Benedikte and Peter Fisker. 2014. "How do Nigerian Small-scale Farmers React to Droughts?". Mimeo.

Blundell, Richard and Monica Costa Dias. 2009. "Alternative Approaches to Evaluation in Empirical Microeconomics." Journal of Human Resources, 44(3), 565-640.

Carneiro, Pedro; James J Heckman and Edward Vytlacil. 2001. "Estimating the Return to Education When It Varies among Individuals," Paper presented as the Economic Journal Lecture at the Royal Economic Society meeting, Durham, England.

Carneiro, Pedro Manuel; Michael Lokshin; Cristobal Ridao-Cano and Nithin Umapathi. 2011. "Average and Marginal Returns to Upper Secondary Schooling in Indonesia."

Corral, Paul; Vasco Molini and Gbemisola Oseni. 2015. "No Condition Is Permanent: Middle Class in Nigeria in the Last Decade." World Bank Policy Research Working Paper, (7214).

Dabalen, Andrew; Stefano Paternostro and Gaëlle Pierre. 2004. The Returns to Participation in the Non-Farm Sector in Rural Rwanda. World Bank Publications.

Davis, Benjamin; Paul Winters; Thomas Reardon and Kostas Stamoulis. 2009. "Rural Nonfarm Employment and Farming: Household-Level Linkages." Agricultural Economics, 40(2), 119-23.

Davis, Benjamin; Stefania Di Giuseppe and Alberto Zezza. 2014. "Income Diversification Patterns in Rural Sub-Saharan Africa: Reassessing the Evidence." World Bank Policy Research Working Paper, (7108).

De Janvry, Alain; Elisabeth Sadoulet and Nong Zhu. 2005. "The Role of Non-Farm Incomes in Reducing Rural Poverty and Inequality in China." Department of Agricultural & Resource Economics, UCB.

Dercon, Stefan. 2002. "Income Risk, Coping Strategies, and Safety Nets." The World Bank Research Observer, 17(2), 141-66.

Ekpoh, I. J., & Nsa, E. (2011). Extreme climatic variability in North-western Nigeria: an analysis of rainfall trends and patterns. Journal of Geography and Geology, 3(1), p51.

Escobal, Javier. 2001. "The Determinants of Nonfarm Income Diversification in Rural Peru." World Development, 29(3), 497-508.

Falola, Toyin and Matthew M Heaton. 2008. A History of Nigeria. Cambridge University Press.

Ferreira, Francisco HG and Peter Lanjouw. 2001. "Rural Nonfarm Activities and Poverty in the Brazilian Northeast." World Development, 29(3), 509-28.

Filmer, Deon and Lant H Pritchett. 2001. "Estimating Wealth Effects without Expenditure Data—or Tears: An Application to Educational Enrollments in States of India*." Demography, 38(1), 115-32.

Fox, Louise and Thomas Pave Sohnesen. 2012. "Household Enterprises in Sub-Saharan Africa: Why They Matter for Growth, Jobs, and Livelihoods." World Bank Policy Research Working Paper, (6184).

Gordon, Ann and Catherine Craig. 2001. Rural Non-Farm Activities and Poverty Alleviation in Sub-Saharan Africa. Natural Resources Institute, University of Greenwich.

Haggblade, Steven. 2007. "Alternative Perceptions of the Rural Non Farm Economy." Transforming the rural non-farm economy: opportunities and threats in the developing world, 262-319.

Heckman, James J and Xuesong Li. 2004. "Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China in 2000." Pacific Economic Review, 9(3), 155-71.

Heckman, James J and Edward J Vytlacil. 1999. "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects." Proceedings of the national Academy of Sciences, 96(8), 4730-34.

Kilic, Talip; Calogero Carletto; Juna Miluka and Sara Savastano. 2009. "Rural Nonfarm Income and Its Impact on Agriculture: Evidence from Albania." Agricultural Economics, 40(2), 139-60.

Marc, A., S. Verjee, and S. Mogaka. 2015. The Challenge of Stability and Security in West Africa. Africa Development Forum Series. Washington, DC: World Bank. doi:10.1596/978-1-4648-0464-9. License: Creative Commons Attribution CC BY 3.0 IGO.

Miguel, Edward; Shanker Satyanath and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." Journal of political Economy, 112(4), 725-53.

Minot, Nicholas; Michael Epprecht; Tran Thi Tram Anh and Le Quang Trung. 2006. Income Diversification and Poverty in the Northern Uplands of Vietnam. Intl Food Policy Res Inst.

Mustapha, A. B., and Arshad, Adzemi Mat. 2014. “Coping Strategy of Rainfall Variability due to Changing Climate in Dry Sub-Humid Region of Borno State of Nigeria”. Journal of Biology, Agriculture and Healthcare, 4(16), 122–126.

Oseni, Gbemisola; Paul Corral; Markus Goldstein and Paul Winters. 2014. "Explaining Gender Differentials in Agricultural Production in Nigeria." World Bank Policy Research Working Paper, (6809).

Oseni, Gbemisola and Paul Winters. 2009. "Rural Nonfarm Activities and Agricultural Crop Production in Nigeria." Agricultural Economics, 40(2), 189-201.

Pfeiffer, Lisa; Alejandro López-Feldman and J Edward Taylor. 2009. "Is Off-Farm Income Reforming the Farm? Evidence from Mexico." Agricultural Economics, 40(2), 125-38.

Radchenko, Natalia. 2014. "Heterogeneity in Informal Salaried Employment: Evidence from the Egyptian Labor Market Survey." World Development, 62, 169-88.

Reardon, T. (1997). Using evidence of household income diversification to inform study of the rural nonfarm labor market in Africa. World Development, 25(5), 735-747. doi:http://dx.doi.org/10.1016/S0305-750X(96)00137-4

Reardon, Thomas; Julio Berdegué; Christopher B Barrett and Kostas Stamoulis. 2007. "Household Income Diversification into Rural Nonfarm Activities." Transforming the rural non-farm economy: opportunities and threats in the developing world, 115-41.

Reuters, “Nigeria Floods Kill 137, Displace Thousands,” 2012.

Robinson, Peter M. 1988. "Root-N-Consistent Semiparametric Regression." Econometrica: Journal of the Econometric Society, 931-54.

Stampini, Marco and Benjamin Davis. 2009. "Does Nonagricultural Labor Relax Farmers' Credit Constraints? Evidence from Longitudinal Data for Vietnam." Agricultural Economics, 40(2), 177-88.

Winters, Paul; Timothy Essam; Alberto Zezza; Benjamin Davis and Calogero Carletto. 2010. "Patterns of Rural Development: A Cross-Country Comparison Using Microeconomic Data." Journal of Agricultural Economics, 61(3), 628-51.

Tables and graphs

Table 1: Diversification strategies among farming households

	Nigeria		North		South	
	2010-2011	*2012-2013*	*2010-2011*	*2012-2013*	*2010-2011*	*2012-2013*
(i) NFE	33.67	40.83	37.59	47.79	26.49	28.76
(ii) Non-farm wage	2.03	2.74	1.79	1.98	2.47	4.04
(iii) Farm-wage	2.50	1.09	3.12	0.82	1.35	1.55
(iv) Comm. crops	5.50	4.20	2.85	1.85	10.34	8.27
(i ∩ ii)	3.22	3.11	3.60	2.77	2.54	3.70
(i ∩ iii)	0.78	0.62	0.98	0.52	0.42	0.79
(i ∩ iv)	6.56	7.55	4.10	5.25	11.06	11.53
(ii ∩ iii)	0.39	0.16	0.35	0.12	0.45	0.24
(ii ∩ iv)	0.56	0.46	0.17	0.00	1.28	1.24
(iii ∩ iv)	0.52	0.12	0.56	0.00	0.44	0.32
(i ∩ ii ∩ iii)	2.44	0.98	2.15	0.95	2.99	1.04
(i ∩ ii ∩ iv)	0.19	0.75	0.05	0.30	0.45	1.52
(i ∩ iii ∩ iv)	0.32	0.66	0.15	0.06	0.63	1.71
(ii ∩ iii ∩ iv)	0.04	0.09	0.00	0.15	0.10	0.00
(i ∩ ii ∩ iii ∩ iv)	0.20	0.41	0.00	0.08	0.57	0.97
Farming only	41.09	36.24	42.54	37.34	38.43	34.33
Total	100	100	100	100	100	100
Observations	2,232	2,232	1,445	1,445	787	787

Note: (i) refers to households who diversify only on NFE. For example, (i ∩ ii ∩ iv) refers to the percentage of farming households who diversify by having a NFE, engaging in non-farm wage labor, and commercial cropping. Values are obtained using survey weights.

Figure 1: Temperature and precipitation patterns, Nigeria 2000-2013. Source: Bjerge and Fisker 2016.

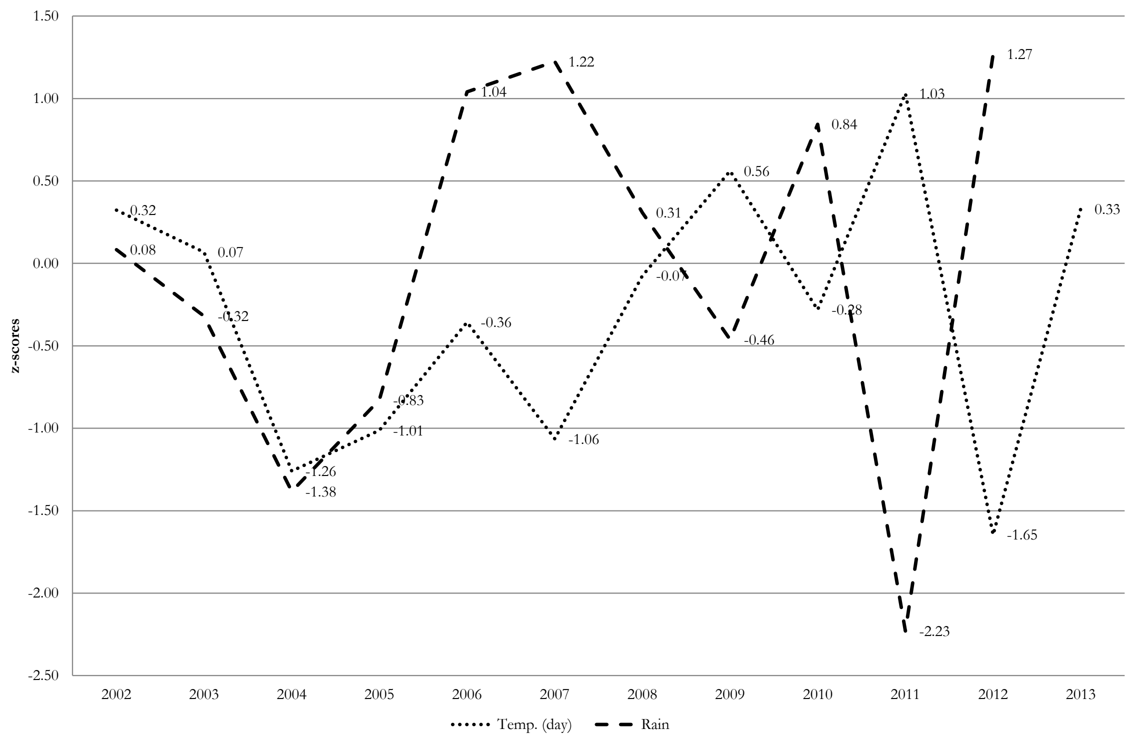

Table 2: Summary statistics, North and South by survey round

	2010/11									2012/13								
	Nigeria			North			South			Nigeria			North			South		
	Non-diversified	Diversified		Non-diversified	Diversified		Non-diversified	Diversified		Non-diversified	Diversified		Non-diversified	Diversified		Non-diversified	Diversified	
Per capita expenditure	72,062	76,956	**	60,092	64,757	**	96,620	98,121		96,309	101,544		80,725	87,009		125,861	125,741	
Harvest value (Naira)	132,333	146,250		135,961	162,606	***	124,890	117,872		155,971	184,148	**	168,752	214,241	***	131,734	134,054	
Harvest per capita	30,693	27,423		27,871	25,891		36,482	30,080		33,361	31,723		33,981	32,298		32,185	30,766	
Male head	0.88	0.92	***	0.96	0.98	**	0.72	0.82	***	0.86	0.91	***	0.96	0.98	*	0.68	0.80	***
Head's age	50.97	49.86		47.42	46.46		58.26	55.78	**	54.09	52.91		50.84	49.55		60.26	58.49	
Head is married	0.84	0.91	***	0.92	0.97	***	0.69	0.82	***	0.80	0.90	***	0.90	0.96	***	0.62	0.80	***
Head's education	3.73	4.88	***	3.31	4.46	***	4.58	5.61	***	3.28	4.83	***	2.80	4.31	***	4.19	5.70	***
Household's size	5.68	6.68	***	6.10	7.47	***	4.80	5.31	**	5.71	6.92	***	6.24	7.77	***	4.72	5.49	***
N. of dependents	2.91	3.39	***	3.21	4.01	***	2.30	2.32		2.73	3.29	***	3.11	3.96	***	2.03	2.17	
Urban	0.06	0.16	***	0.05	0.11	***	0.10	0.26	***	0.07	0.16	***	0.04	0.10	***	0.12	0.26	***
Rainfall shock	184.19	176.93	***	183.23	176.21	***	186.17	178.19	***	178.42	170.06	***	178.23	170.45	***	178.78	169.41	***
Infra. Index	0.11	-0.08	***	0.20	(0.15)	***	-0.07	0.05		0.14	-0.08	***	0.15	-0.09	***	0.11	-0.06	**
Land ownings (Ha)	1.52	1.90		1.40	1.57		1.76	2.48		0.97	1.13	*	1.08	1.23	*	0.76	0.96	
Dist to road (km)	18.67	16.50	***	21.75	19.84		12.36	10.70	**	7.76	7.54		8.54	8.71		6.28	5.59	
Dist to pop center (km)	26.12	25.20		31.74	31.19		14.60	14.79		23.72	21.48	***	27.71	25.89	*	16.14	14.13	***
Dist to market (km)	69.48	71.40		71.44	71.77		65.46	70.75		68.67	72.89	**	70.21	72.48		65.76	73.56	**
Dist to border (km)	304.51	294.44		226.77	225.92		463.99	413.32	***	324.93	292.20	***	250.26	227.11	**	466.53	400.55	***
Observations	949	1,283		622	823		327	460		851	1,381		555	890		296	491	
Total Obs.	2,232			1,445			787			2,232			1,445			787		

Note: All values are obtained using survey weights. *p<0.1; **p<0.05;***p<0.01

Table 3: First-stage results of Panel IV FE estimation

Dependent variable:	*Diversified*
Rainfall Shock	-0.00394***
	(0.00136)
Infrastructural constraint index	-0.00540
	(0.0122)
Female headed	0.122
	(0.0884)
Head's age	0.00262
	(0.00562)
Head's age (sq)	-1.77e-05
	(4.97e-05)
Head is married	0.0491
	(0.0463)
Head's years of education	0.00154
	(0.00314)
Household's size	0.0501***
	(0.00934)
Dependents	-0.0170*
	(0.00939)
Urban	-0.131
	(0.223)
Nat. log of hectares per capita	0.0201*
	(0.0107)
Nat. log of hectares per capita (sq)	0.000258
	(0.00157)
Dist. to market (km)	0.00944**
	(0.00467)
Dist. to pop. center (km)	0.000717
	(0.000581)
Dist. to border (km)	0.000286
	(0.000532)
State FE	Yes
Observations	4,464
Number of households	2,232
Weak identification test (F-stat)	12.45
Endogeneity test (p-value)	0.00002

Robust standard errors clustered at the household-year level in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 4: 2SLS results of Panel IV estimation

Dependent variable:	*ln pc exp.*
HH diversifies	4.164***
	(1.081)
Infrastructural constraint index	0.00498
	(0.0363)
Female headed	-0.542*
	(0.293)
Head's age	-0.0189
	(0.0179)
Head's age (sq)	0.000161
	(0.000158)
Head is married	-0.364**
	(0.151)
Head's years of education	0.00379
	(0.00991)
Household's size	-0.264***
	(0.0664)
Dependents	0.0507
	(0.0375)
Urban	0.425
	(0.753)
Nat. log of hectares per capita	-0.0998***
	(0.0385)
Nat. log of hectares per capita (sq)	-0.00300
	(0.00477)
Dist. to market (km)	-0.0490**
	(0.0220)
Dist. to pop. center (km)	-0.00461**
	(0.00188)
Dist. to border (km)	0.00361**
	(0.00172)
State FE	Yes
Observations	4,464
Number of households	2,232

Robust standard errors clustered at the household-year level in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5: Selection equation, Logit MFX estimates

	2010/11	*2012/13*
Dependent variable:	*Diversified*	*Diversified*
Rainfall shock	-0.00187***	-0.00153**
	(0.000621)	(0.000682)
Infrastructural constraint index	-0.0632***	-0.0493***
	(0.0164)	(0.0151)
Female headed	-0.0671	-0.142***
	(0.0530)	(0.0467)
Head's age	0.00633	0.000336
	(0.00524)	(0.00468)
Head's age (sq)	-7.69e-05	-8.04e-06
	(4.88e-05)	(4.13e-05)
Head is married	0.108**	0.180***
	(0.0493)	(0.0469)
Head's years of education	0.00662**	0.0125***
	(0.00259)	(0.00247)
Household's size	0.0347***	0.0309***
	(0.00705)	(0.00783)
Dependents	-0.0188**	-0.0155*
	(0.00883)	(0.00926)
Urban	0.0777	0.0463
	(0.0504)	(0.0478)
Nat. log of hectares owned	-0.00279	0.00634
	(0.00982)	(0.0117)
Nat. log of hectares owned (sq)	-0.00153	0.00120
	(0.00226)	(0.00308)
State FE	Yes	Yes
Observations	2,232	2,232

Standard errors clustered at village level in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

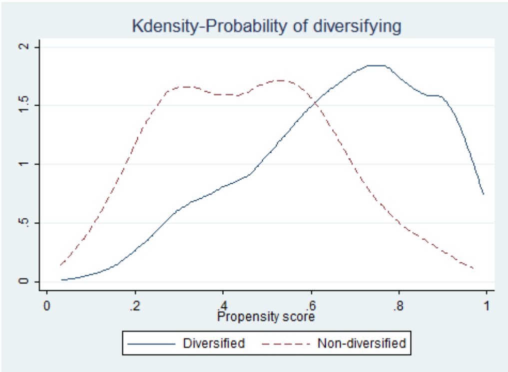

Figure 2: Common support (2010/11)

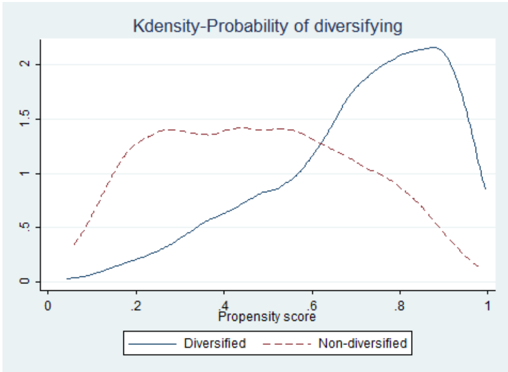

Figure 3: Common support (2012/13)

Table 6: Treatment effects for 2010/11 and 2012/13

	2010/11			2012/13		
	(1)	(2)	(3)	(4)	(5)	(6)
	Total	Observables	Unobservables	Total	Observables	Unobservables
TT	0.825***	0.368***	0.457***	1.511***	2.929***	-1.419***
	(0.118)	(0.012)	(0.118)	(0.117)	(0.016)	(0.117)
ATE	0.595***	0.296***	0.298***	0.771***	2.840***	-2.070***
	(0.073)	(0.009)	(0.073)	(0.103)	(0.012)	(0.102)
TNT	0.368***	0.202***	0.165	0.101	2.696***	-2.595***
	(0.116)	(0.013)	(0.115)	(0.180)	(0.019)	(0.180)

Bootstrapped standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

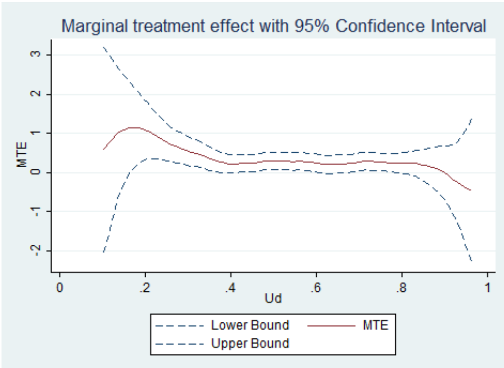

Figure 4: MTE^v with 95% CI-Expenditure (2010/11)

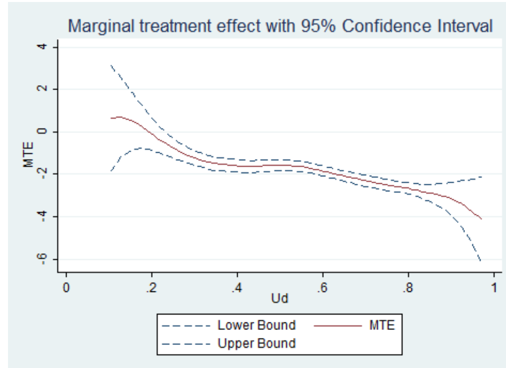

Figure 5: MTE^v with 95% CI-Expenditure (2012/13)

Table 7: Semi-parametric estimates

	2010/11		2012/13	
Dependent variable:	*ln pc exp.*		*ln pc exp.*	
	Return from non-diversifying	Gap in divers.-non divers. returns	Return from non-diversifying	Gap in divers.-non divers. returns
	β_0	$\beta_1 - \beta_0$	β_0	$\beta_1 - \beta_0$
Female headed	0.196	-0.470	0.297*	-0.157
	(0.157)	(0.291)	(0.156)	(0.227)
Head's age	-0.00918	0.0133	-0.0363**	0.0725***
	(0.0127)	(0.0229)	(0.0145)	(0.0228)
Head's age (sq)	9.23e-05	-0.000121	0.000329**	-0.000625***
	(0.000120)	(0.000221)	(0.000131)	(0.000208)
Head is married	-0.314**	0.417	-0.264*	0.00698
	(0.141)	(0.274)	(0.140)	(0.224)
Head's years of education	0.0197***	0.000735	-0.0104	0.0301**
	(0.00737)	(0.0110)	(0.00989)	(0.0142)
Household's size	-0.143***	0.124***	-0.144***	0.110***
	(0.0277)	(0.0375)	(0.0242)	(0.0310)
Dependents	-0.00308	-0.0551	-0.0785***	0.0486
	(0.0303)	(0.0426)	(0.0271)	(0.0370)
Urban	-0.148	0.288	-0.240	0.359*
	(0.157)	(0.228)	(0.161)	(0.217)
Nat. log of hectares per capita	0.0692	0.0356	0.138**	-0.130
	(0.0507)	(0.0763)	(0.0692)	(0.106)
Nat. log of hectares per capita (sq)	0.00499	0.00329	0.0186	-0.0192
	(0.00835)	(0.0130)	(0.0116)	(0.0164)
Dist. to market (km)	-3.61e-05	-0.00146	-0.00143	0.00104
	(0.000985)	(0.00154)	(0.00100)	(0.00141)
Dist. to pop. center (km)	0.000577	-0.00227	-0.00578*	0.00175
	(0.00181)	(0.00310)	(0.00308)	(0.00455)
Dist. to border (km)	0.00104**	-0.00124***	0.00153***	-0.000941**
	(0.000410)	(0.000476)	(0.000416)	(0.000448)
State FE	Yes	Yes	Yes	Yes
Observations	2,232	2,232	2,232	2,232
R-squared	0.508	0.508	0.425	0.425

Standard errors clustered at village level in parentheses.*** p<0.01, ** p<0.05, * p<0.1