Benin Journal of Statistics
ISSN 2682-5767
Vol. 2, pp. 47– 60 (2019)

Panel data models for studying the impact of climate change on certain agricultural crops using some test statistics

A. A. Edokpayi[1*], B. A. Oyejola[2], S. A. Adeh[3] and E. A. Iguodala[4]
[1,3,4]*Statistics Division, Nigerian Institute for Oil Palm Research, Benin City*; [2]*Department of Statistics, University of Ilorin, Ilorin*

Abstract. The issue of climate change is now a global concern and its effects on agriculture in Nigeria have been investigated severally using either time series or cross sectional data approach. The use of time series and cross sectional data approaches led to misspecification errors or omitted variable bias and this affect the efficiency of the model estimation parameters. As a remedy to estimation bias, this study applies the panel data approach. Three panel data models - pooled, fixed effects and random effects models were used to fit a balanced panel data consisting of seven crops for twenty-eight years' periods. The appropriate model was determined using some test statistics such as the F-Test, Breusch–Pagan Test and the Hausman's Specification. The analysis of the weather shows significant variation over time indicating the presence of climate change. Using the F- test, the pooled model is rejected which suggests the presence of unobserved heterogeneity. The Breusch-Pagan test also suggests the rejection of the pooled model while the panel effects exist in the panel. The panel model is able to accommodate heterogeneity better than the pooled model. The Hausman test however indicates that the fixed effects model is more appropriate for the study because the assumption that the individual effects are uncorrelated with the other regressors in the model was rejected. The study further reveals higher coefficients value for the pooled model and random effect model than the fixed effect model indicating that the values may have been overstated. The climatic factors were significant for the pooled and random effect models but not significant using the fixed effect model, thus revealing that the observed climate change have not significantly manifested so much on the selected crops in Nigeria. It was also observed from the study that the intercept coefficients for the different crops were individually highly significant; indicating the presence of individual heterogeneity. This suggests that different crops will be impacted differently. The study therefore affirms the presence of unobserved effects in the data and the fixed effects model accounts for this better than the polled and random effects model when estimating the impact of climate change on agriculture in Nigeria. Since the different crops are impacted differently, the study suggests crop specific and not country wide or national adaptation/mitigation policy

Keywords: climate change, panel data, panel data models, test statistics, agricultural crops.

Published by: Department of Statistics, University of Benin, Nigeria

1. Introduction

The issue of climate change is now a global concern and its effects are far reaching to people and Nations. Climate Change (CC) as stated by IPCC (2007) is a phenomenon that describes any change in the climate over time due to natural variability or as a result of human activities. Natural variability may be higher average temperature, changing rainfall pattern and rising sea levels. Human activities on the other hand could be due to the conversion of forests and grasslands for crops and posture, which result in significant release of greenhouse gases.

This rising concentration of greenhouse gases in the atmosphere is said to lead to higher temperature and increased precipitation over the next century. The changes in climate are predicted to have significant impact on economic activity Feres. et. al. (2011). One of the most significant ways that global climate change is predicted to affect economic activity is through its effects on agriculture, since temperature and precipitation are direct inputs to agricultural production. Already, in some

*Corresponding author. Email: edokpayiaustin@yahoo.com

areas, particularly dry regions, increased rain resulting from climate change is already having a devastating effects while in areas where the temperatures are already very high, further warming may make farming unprofitable (Feres, et al., 2011).

The Intergovernmental Panel on Climate Change (IPCC, 2007) forecasts that developing countries, like Nigeria, will continue to be affected by extreme weather variability such as temperature, severe water shortage, and flood-inducing rainfall events during the coming decades. The northern part of Nigeria is already at risk of desertification and drought while ocean surge, marine erosion and a receding coastline are problems in the south (Ayoade, 2002). Ogundele et al. (2011) revealed that the climate change has resulted to soil loss, plant nutrient loss, textural change, and increase in pests and diseases, and poor yield germination and a reduction of arable lands cultivation in Nigeria, resulting to declining food production in the Country.

In a Nigeria where agriculture is the second largest source of National wealth after oil, providing livelihood for more than 70% of the population and contributing to about 40% to National GDP Akor (2012), threat to agricultural productivity arising from climate change raised serious concern. Kayode (2010), Ayinde et al. (2011), Akpata (2012) and Bello (2012), have in the past provided information on the impact of climate change on agricultural production in the country. However, a unique feature of these studies is that they relied on either the cross-sectional approach or the time series approach. Using time series approach is very restrictive as it is only possible to estimate the impact of one crop at a time thereby neglecting information about other crops which are cultivated in the same environments. On the other hand, using cross sectional data makes it possible to estimate the impact for only one year. at a given period, and in a situation where you have data for thirty year periods, that would entail thirty cross sectional regression analysis.

Another major methodological challenge with the use of cross sectional or time series approaches is that they fail to account for other variables that also affect crop yield but not included in the regression equation especially when such omitted variable(s) correlates with any regressors included in the model. Using cross sectional and time series approach, such omitted variables is usually lump up in the disturbance term. When this happens, it violates the classical regression assumptions that the error term should not correlated with any of the explanatory variables. Not accounting for such correlated omitted variables led to misspecification errors or omitted variable bias and this affect the efficiency of the model estimation. Another concern is the inability to model the differences or heterogeneity (individual differences) among crops. In cross – sectional and time series regression analysis, the uniqueness of each crops are usually ascribed to the disturbance term, in that case assuming the different crops to be the same Failure to account for heterogeneous quantities in the model may introduce bias into the model estimators. The implication of these problems is that one may be accepting what should have been rejected and in a situation where policies are based on such econometric analyses, policy/decision makers would have been misled.

These inadequacies have led some researchers to suggest the use of a panel data approach to take into account the problem of omitted variables and control of individual heterogeneity. As observed by Park (2009), panel data approach can be helpful for the correction of the bias generated by omitted variables and heterogeneity in the classical cross – sectional regression. Also Elhorst (2003) noted that Panel data are generally more informative and contain more variation and less collinearity among the variables and results in a greater availability of degrees of freedom and hence increase efficiency in the estimation than the typical cross section or time series data approaches.

Since Deschenes and Greenstone (2007), pioneering work, Panel data model has become reasonably common in climate change impact research. However, one major challenge is the lack of consensus on which panel data model to use. While some researchers have used the pooled model, others have argued that pooling may not be justified when parameters of time series regression at the individual level vary considerably across samples. Some papers have equally questioned the homogeneity assumption of the pooled model and have shown that heterogeneous models are less biased than the traditional homogeneous models. Even the heterogeneous models, there is the question of whether to use the fixed effects model or the random effects model.

The present study empirically addresses the questions of whether or not to pool in the context of modeling the impact of climate change on crop yield in Nigeria. Secondly, when pooling fails, in other word, when heterogeneity is present, is it appropriate to use fixed effects or random effects model to account for this heterogeneity? More specifically, the study intends to investigate and determine the appropriate panel data model specification for estimating the impact of climate on crop yield in

Nigeria.

2. Materials and method

Panel data like other aspect of econometrics study, uses regression analysis as one of the statistical tool to formulate, describe and evaluate models. Generally, the starting point in panel data analysis is the specification of panel data model. A panel data model is of the form

$$y_{i,t} = \beta_0 + \beta_1 X_{it} + \beta_2 X_{it} + \cdots + \beta_k X_{it} + \alpha_i + \mu_{it} \tag{1}$$

where y_{ij} the depended variable, β regression coefficients, X_{it} is the a vector of explanatory t variables, i is the individuals, t is time periods, k is the number of independent variables, μ_{it} is the time varying idiosyncratic error and α is the unobserved time invariant individual i effect. The variable α is called the unobserved effects consisting of variables that have impact on the dependent variable but are not directly observable.

But if we assume that the unit effects are all equivalent, that is, $\alpha_i = \alpha_k$ for all i and k, then equation (1) reduces to the pooled model (Lioyed et al., 1997). The pooled model ignores α_i and can easily be estimated by fitting a linear regression to the full dataset, ignoring information about how observations are grouped into units. The pooled model does not account for unobserved heterogeneity.

When unobserved effects are included in the regression, the researcher is faced with whether to use fixed effects or random affects models to account for these unobserved effects, Robert (2009). In fixed effects model, there is significant differences between countries (crops/districts) but no significant temporal effects. That is, the model examines group differences in intercepts, and assuming the same slope and constant variance across entities or subject (Park, 2011). Since a group (individual–specific effect) is time invariant and considered a part of the intercept, α_i is allowed to be correlated to other regressors. The advantages of fixed effects (FE) specification are it can allow the individual and/or time specific effects to be correlated with the explanatory variables. The disadvantages are that, firstly, the number of unknown parameters increase with the number of sample size. In the case when T or N is finite, it introduces the classical incidental parameter problems. Secondly, the fixed effect does not allow the estimation of the coefficients that are time invariant.

A random effect model estimates variance components for group (or time) and error, assuming the same intercepts and slopes. α_i is a part of the errors and thus should not be correlated to any regressors, otherwise, a core OLS assumption is violated. Major advantages with this type of specification are, firstly, the numbers of parameter stay constant when samples sizes increases, secondly, it allows the derivation of efficient estimators that makes use of both within and between (group) variations and finally, it allows the estimation of the impact of time -invariant variables. Its disadvantages are that if α_i is correlated with X_{it} or if there is a fundamental difference among individual units, X_{it}, Y_{it} cannot be viewed as random draw from a common distribution. In that case, the random effect model is mis-specified and the resulting estimator is biased.

2.1 *Empirical model specification*

In the pooled or constant coefficients modeling approach, the heterogeneity (individual differences) among the crops is ignored and all observations are pooled together. This model type subsumed the individuality of each subject in the disturbance term μ_{it}. In this case, the model assumed that the regression coefficients are the same for the seven crops. That there is no distinction between the crops – one crop is as good as the other crop and therefore the model estimates a common constant and slope coefficient for all cross sectional individual units (crops). For this study, we specify the pooled model as follows.

$$y_{i,t} = \beta_0 + \beta_1 R_{it} + \beta_2 R_{it}^2 + \beta_3 T_{it} + \beta_4 T_{it}^2 + \beta_5 R_{it} T_{it} + \beta_6 C_{it} + \beta_7 P_{it} + \mu_{it} \tag{2}$$

where $i = 1, 2, \cdots, 7$ and $t = 1, 2, \cdots, 28$ is the cross sectional unit (different crops) and t is the time period (measured in years). Y_{it} is the crop output for the ith crop at time period t, and R_{it}

and T_{it} are the vectors of climate variables (rainfall and temperature, respectively) and $R_{it}T_{it}$ is the interaction term for the ith crop at time period t. β_0 is the intercept and β_i is the slope coefficients which reflects parameter estimates for all independent variables.

Quadratic terms for weather variables are included in the specification to account for non-linear weather effects on crop yields. The interaction term, $R \times T$, between the weather variables are used to determine the potential effects of one weather variable given the effect of the other variable. μ_{it} is the stochastic disturbance term that accounts for other variables such as fertilizer, improved crop variety, crop management, weeds, pests and disease control and soil quality which were omitted in the model. C_{it} is the credit to farmers for ith crop at time period t and P_{it} is the prices of the ith crop for time period t.

The linear terms in the above model represent the marginal value of climate at the mean while the squared terms are representing the shape of the relationship between climate and crop yields.

According to Mendelsohn (2001), a positive coefficient indicates a U shape and the negative coefficients reflect a hill shape relationship. A hill shape relationship between climate variables and crops indicates that as the climate variables increases, crop yields increases to a certain point (Maximum), increasing climate variable beyond this points reduces crop yields. On the other a U shape relationship shows that crop yields will decrease as climate variable rise to reach a certain point (Minimum), and then both crop yields and climate variables will increase.

In the fixed effects model specification, we assume that the slope coefficients are constant but the intercept varies across individual units (crops). The subscript i is included on the intercept to indicate the differences of the intercept across the seven crops. In this case the individual (crop) intercepts do not vary with time but vary across individual crops as units. In other words, the intercepts are time invariants. For this study, the fixed effects specification is of the form

$$y_{i,t} = \beta_0 + \beta_1 R_{it} + \beta_2 R_{it}^2 + \beta_3 T_{it} + \beta_4 T_{it}^2 + \beta_5 R_{it} T_{it} + \beta_6 C_{it} + \beta_7 P_{it} + \alpha_i + \mu_{it} \tag{3}$$

where all parameters are as defined above, α_i, which is added to the equation, is called the unobserved or a heterogeneity effect. This additional term (α_i) accounts for other variables such as soil quality which contribute to crop yields, varies among crops, but is constant over time for a given crop while μ_{it} is the idiosyncratic error or time varying error and represents unobserved factors that change over time and affects crop yield. We can rewrite equation (3) as:

$$y_{i,t} = \alpha_1 + \alpha_2 D_{2i} + \cdots + \alpha_7 D_{7i} + \beta_2 X_{2it} + \cdots + \beta_k X_{kit} + \mu_{it} \tag{4}$$

where α_1 is the intercept for crop 1, which is the benchmark, and $\alpha_2, \cdots, \alpha_7$ are the differential intercepts for the remaining six crops. In this case, all the heterogeneity is subsumed in the intercept values and the estimated intercepts for each subject or crop in this case represent the subject-specific characteristics.

For the random effects model, we assume that the intercept of each cross sectional unit is drawn from a distribution that is centered on the mean intercept. Since α_i is not directly observable, we can consider it as random and include it in the error term μ_{it}, and thereby consider the composite error term

$$V_{it} = \alpha_i + \mu_{it}$$

Thus the random effects model becomes

$$y_{i,t} = \beta_0 + \beta_1 R_{it} + \beta_2 R_{it}^2 + \beta_3 T_{it} + \beta_4 T_{it}^2 + \beta_5 R_{it} T_{it} + \beta_6 C_{it} + \beta_7 P_{it} + \alpha_i + V_{it} \tag{5}$$

2.2 *Model testing and selection criterion*

Three tests are used in panel data analysis. They are: the F test which compares Pooled vs Fixed effects models, the Lagrange Multiplier (LM) test that compares Pooled vs Random Effects models and Housman Test which compares fixed versus random effects model.

2.2.1 F test

The hypothesis to be tested is

$$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_7$$

versus

$$H_1 : \alpha_1 \neq \alpha_2 \neq \alpha_3 \neq \alpha_4 \neq \alpha_5 \neq \alpha_6 \neq \alpha_7$$

that is, H_0 : cross-sectional heterogeneity does not exist versus H_1 : there is presence of cross-sectional heterogeneity. The test statistic is defined as

$$\frac{(e'e_{OLS} - e'_{FE}/n - 1)}{e'e_{FE}/nT - n - k} = \frac{\left(R^2_{FE} - R^2_{OLS}/n - 1\right)}{(1 - R^2_{FE})/(nT - n - k)} \sim F(n - 1, nT - n - k) \tag{6}$$

where n is the number of groups, T is the total number of time periods, $N = nT$ is the total number of observations, k is the number of regressors and $e'e$ is the sum of squared error for both the pooled and the fixed effects models. R^2_{FE} is the coefficient of determination of the fixed effects model and R^2_{OLS} is the coefficients of determination of the pooled OLS model.

If the observed F statistic is greater than the critical F, then we reject the null hypothesis and conclude that significant difference exists among the crops and therefore the fixed effects is more appropriate for the study. But if the null hypothesis is not rejected, it therefore means that there is no difference among the seven crops. In that case, the pooled model will be appropriate for the study.

2.2.2 Breusch-Pagan LM test

The Breusch –Pagan Lagrange multiplier (LM) test is for testing the significance of random effects. The hypothesis to be tested is

H_0: there is no random effects versus H_1: there is random effects

The test statistics is defined as

$$LM_\mu = \frac{nT}{2(T - 1)} \left[\frac{T^2}{e'e} - 1\right]^2 \sim \chi^2_{(1)} \tag{7}$$

where n is the number of groups, T is the total number of time periods, $N = nT$ is the total number of observations, k is the number of regressors and $e'e$ is the sum of squared error. The LM statistic is distributed as chi-square distribution with one degree of freedom. If the null hypothesis is rejected, we conclude that there is a significant random effect in the panel data, and that the random effect model is able to deal with heterogeneity better than the pooled model.

2.2.3 Hausman's specification test

The hypotheses is to test if there is significant correlation between unobserved individual specific (α_i) random effects and the regressors (X_{it}) and this is defined as:

$$H_0 : Cov(X_{it}, \alpha_i) = 0 \text{ versus } H_1 : Cov(X_{it}, \alpha_i) \neq 0.$$

The hypotheses can be modified as follow:

$$H_0 : (\beta_{FE} – \beta_{RE}) = 0 \text{ versus } H_1 : (\beta_{FE} – \beta_{RE}) \neq 0.$$

The Hausman statistic is defined as;

$$H = (\hat{\beta}_{FE} - \hat{\beta}_{RE})' \left[\hat{V}(\hat{\beta}_{FE}) - \hat{V}(\hat{\beta}_{RE})\right]^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{RE}) \tag{8}$$

where $\hat{V}_S$ denotes estimates of the true covariance metrics. The Hausman test is a kind of Wald χ^2 test with $k-1$ degrees of freedom, where k is the number of regressors. Use random effects unless test rejects orthogonality conditions between α_i and X_{it}. Rejection means that the random effects assumption fails and fixed effects should be used.

2.2.4 Panel data model estimation

For the pooled model, we used the ordinary least square method (OLS) to estimate all 196 observations by stacking the 28 observations of each crops one after the other. For the fixed effects model, the parameters $\alpha_1, \alpha_2, \cdots, \alpha_N$ and β can be estimated by OLS. The implied estimator for β's is referred to as the Least Square Dummy Variable (LSDV) estimator. If we use OLS procedure under random effects specification, a key feature of the random effects model will be ignored. This is because the unobserved is in the composite error in each time period. Because the composite errors are serially correlated across time, the OLS standard errors ignore this correlation, and will be incorrect as well as other test statistics. The generalized least square can be used to estimate models with autocorrelation, when the variance structure among groups is known and the feasible generalized least square (FGLS) method is used when the variance structure is not known.

3. Results and discussion

This section presents trend analysis of rainfall and temperature distributions, test of hypotheses and the empirical estimation of model and analysis of results. A panel data comprising of seven cross-sectional units and twenty-eighty ears' periods was used for the study. The seven crops consist of Groundnut, Cotton, Coconut, Shea nut, Oil palm, Cocoa, and Rubber. State level annual mean rainfall and temperature data were collected from the Nigerian Meteorological Agency for the periods 1981 – 2009. Data used for the study were drawn from 2011 Statistical Bulletin of the CBN (Central Bank of Nigeria).

Table 1.: Pattern of temperature distributions between 1981 – 2009 across Nigeria

Variables	Means	Constant	Coefficients	R^2
Groundnut	29.97	34.117	0.2859	0.5329
Cotton	25.93	38.701	0.454	0.4305
Coconut	30.33	23.198	0.0888	0.0244
Shea nut	32.12	33.844	0.2422	0.4245
Oil palm	27.86	128.951	0.075	0.3372
Cocoa	26.51	26.112	0.0275	0.0388
Rubber	27.22	27.441	0.0148	0.0133

Table 1 shows the temporal and spatial distribution of temperature within the study periods. In some cropping area, particularly in the northern part of Nigeria represented by (Groundnut, Cotton and Shea), the highest average temperature was 32.12^0c while the lowest was 25.93^0c. The computed Temperature (R^2 for these northern crops were 0.53%, 0.48% and 0.42% in Groundnut, Cotton and Shea growing areas of northern Nigeria respectively within the periods, showed increase in temperature in these areas. The computed temperature (R^2) values for Coconut, Oil palm, Cocoa, and Rubber were 0.02%, 0.33%, 0.03% and 0.01 % respectively, thus indicating that temperature increase was generally higher in Northern Nigeria than the South. The results (see Table 1), also show that the mean temperature values were different indicating spatial variations across the country. Generally, the result shows temperature increase across the country and this finding agree with earlier report by Obioha (2007) and (IPCC, 2008) were they reported observed temperature increase in

Table 2.: Pattern of rainfall distribution in areas suitable for the cultivation of crops used in the study

Crops	Mean	Constant	Coefficients	R^2
Groundnut	439.51	993.14	38.181	0.8448
Cotton	1170.60	902.59	31.7070	0.8009
Coconut	581.58	2740.4	108.28	0.442
Shea nut	442.83	1285.3	48.531	0.7548
Oil palm	1096.00	2021.5	70.66	0.5191
Cocoa	1789.40	18367.8	33.385	0.0078
Rubber	2407.50	1780.5	43.257	0.0107

Nigeria and globally respectively. There is occurrence of rainfall variability in the country across the cropping areas under study (Table 2). The average rainfall was particularly high in the southern part of the country. The northern part of the country dominated by Groundnut, Cotton and Shea cultivation experienced low average rainfall within the study periods. The computed R^2 values of 0.84%, 0.80%, 0.42%, 0.75%, 0.52%, 0.01% and 0.0% for Groundnut, Cotton, Coconut, Shea, Oil palm, Cocoa and Rubber respectively shows high variability in the northern part of the country than the southern parts. The coefficients (see Table 2) which represent the amount of change in rainfall reported a higher value for the Southern parts than the Northern parts of the country. This result agrees with Odjugo (2007) where he showed that there was incessant increase in rainfall in most coastal areas and a drastic declines and erratic in rainfall amount and duration in the interior of the semi- arid regions of Nigeria. General findings showed that the average rainfall for 1981 (base year) was higher than the average rainfall for the periods covered in the study. This declining rainfall is an evidence of climate change. From Table 3, the F-test conducted to shows if significant difference

Table 3.: Cross-sectional heterogeneity among the seven crops

Var.	G. Nut	Coconut	Sheatree	Cotton	Oilpalm	Cocoa	Rubber	Prob.Value
Crops	1.7249E6	1.52721E5	3.1969E5	2.8365E5	6.4083E5	2.1894E5	2.1579E5	≤0.001
Rainfall	4.3951E2	1.1706E3	5.8158E2	4.4283E3	1.0960E3	1.7894E3	2.4075E3	≤0.001
Squared Rainfall	3.0568E3	3.1531E6	5.4023E5	2.7800E5	1.8289E6	3.2578E6	3.6620E5	≤0.001
Temperature	29.9714	25.9357	30.3321	32.1179	27.8636	26.5107	27.2268	≤0.001
Squared Temperature	9.1021E2	6.7639E2	9.2665E2	1.0628E3	7.7751E2	7.0376E2	7.4237E2	≤0.001
Rainfall X Temperature	1.3985E4	2.8752E4	1.8501E4	1.5381E4	3.0735E4	4.7403E4	6.5187E4	≤0.001

existed among the cross sectional units (crops) revealed highly statistically significant values for all the cross sectional units (crops) indicating that significant differences exists among the crops. In this case the null hypothesis (homogeneity among the cross sectional units) rejected while the alternative hypothesis of (heterogeneity among the cross sectional units) is accepted. From the results (Table 3) we conclude that, there is significant variation among the crops indicating the presence of unobserved heterogeneity. Therefore, using homogeneous model (pooled model) will result in heterogeneity bias and that a heterogeneous model (fixed effect model) will be better model for the study data. Table

Table 4.: Breusch and Pagan and Lagrangian Multiplier tests

Parameters	Var.	Sd.=Sqrt.(var.)
Breusch and Pagan	4.44e+11	666141.6
Lagrangian Multiplier	1.31e+11	361631.5
U	0	0

Test Var.(u)=0 with χ^2 value = 652.30 and the p-value = 0.000.

4 reports the results of the Breusch and Pagan Lagrangian Multiplier test which contrasts a random effect model with the Pooled model. Based on the χ^2 (652.30) and a ($P \leq 0.05$) values, the null hypothesis is rejected, we therefore conclude that there is a significant random effect in the panel data, and that the random effect model is able to deal with heterogeneity better than the pooled model.

Table 5.: Hausman test

Variables	Fixed model coefficients	Random model coefficients	Differences between the two models
Rainfall	-521.847	-363.3438	-158.5032
Squared rainfall	-0.0045361	-0.0422437	0.0377076
Temperature	-113780.6	300766.7	-414547.3
Squared Temperature	872.8444	-5186.586	6059.4304
Rainfall x Temperature	19.15582	12.60898	7.14684

The Wald $\chi^2(7)$ is 53.93 (prob. $\geq$ 0.000).

Table 5 report the results of the Hausman test performed to test the appropriateness of employing either Fixed Effects Model (FEM) or Random Effects Model (REM). The result shows that the coefficients of the fixed effects model for the climatic and other variables were considerable higher than the corresponding random effect model. The Wald $\chi^2(7)$ value is 53.93, while the p value = 0.000 indicating a rejection of the random effect model. This shows that the fixed effects model is better able to deal with heterogeneity than the random effects model.

Table 6.: Estimated pooled, fixed and random effects models

Variable	Pooled model		Fixed effect model		Random effects model	
	Coeff.	Std err	Coef	Std err	Coef	Std err
Constant	-3719007	3098146	4129486	2050776	-3719007	3098146
Rainfall	-363.343	845.0628	-521.847	547.79	-363.3438	845.0628
Squared Rainfall	-0.042***	0.016123	-0045361	0.0100808	-0.042***	0.016123
Temperature	300766.7	189936.7	-113780.6	1252858	300766.7	189936.7
Squared temperature	-5186.59*	2816.859	872.8444	1870.22	-5186.58*	2816.859
Rainfall XTemperature	12.0089	31.87238	19.15582	20.71	12.00898	31.87238
D2(cotton)			-1827142***	110917.7		
D3(coconut)			-1291236***	106040.5		
D4(shea nut)			-1275978***	111150.1		
D5(oilpalm)			-1120681***	108007.9		
D6(cocoa)			-1524117***	113217.7		
D7(rubber)			-1447203***	116697.7		
F	7.70		36.90			
R^2	0.36		0.72		0.22	
Adjust R^2	0.19		0.70			
SEE	361631.46		598061.15		361631.48	
D.W.	0.402		0.407			
MSE	(0.000636		0.000036			

Table 6 compares estimates of the pooled model, fixed effects model and the random effects models. The estimated pooled model shows that both precipitation and temperature has significant negative effect (U) relationship with crop yield. The negative squared terms indicates that crop yield will decrease as the climate variables increases, after a certain point (minimum) both crop yields and climate variables will increase. The significance of the quadratic terms suggests that the relationship between climate factors and crop yield is non-linear. These findings aggress with the findings Deressa et al (2005) that climate change has significant non-linear relationship with net revenue or crop yield. The estimated pooled model shows that the control variables (loan and price) were highly significant.

An examination of the statistical criterion shows the R^2(0.22) and Durbin Watson statistics (0.402) of the pooled model to be very low. The R^2 suggests that the independent variables could only explain 22 percent of the total variation in crop yield indicating that other variables not included in the model accounted for the remaining 78 percent. The low Durbin Watson statistics used to test for the presence of autocorrelation the data suggests that perhaps there is autocorrelation. The presence of autocorrelation suggests that the estimated parameters are inefficient as the standard error is under estimated. A low D.W. could also be attributed to specification error (Gujarati, 2009). Another problem with the pooled model is that it does not distinguish between the crops (unobserved effects or individual heterogeneity). That is it does it tell us whether the response of crop yields to the climate variables over time is the same for all the crops. In other words, by lumping together different crops at different times, we camouflage the heterogeneity (individuality or uniqueness) that may exist among the crops. Another way of stating this is that the individuality of each crop is subsumed in the disturbance term . As a consequence, it is quite possible that the error term may be correlated with some of the regressors included in the model. For example, the soil types have great influence on crop yield but since this was not observable, it was subsumed in the disturbance term and it is very likely that the soil type correlated heavily with climatic factors as the soil type of every region is influenced by the climatic factors of the locations. If that is the case, the estimated coefficients for the pooled model are biased as well as inconsistent.

An examination of the results of the fixed effects model and random effects model shows that in linear terms, the climatic variables were not significant for the two models. But for the quadratic terms, just as the pooled model, the random effect estimation shows that squared rainfall was significant at 5% while the squared temperature was significant at 1%. A most noticeable difference among the estimated models is that the signs, size of the coefficients and the significance of the parameters were not generally the same. In the pooled model, and random effects estimations, the climate parameters particularly the squared terms were significant, while the fixed effect panel data model, individual climate parameters were not significant, thus suggesting that the observed increase in temperature and precipitation has not yet manifested on the yields of the selected crops in Nigeria within the study periods. Also the magnitude of the coefficients for the pooled model and the random effects models were almost three times higher than that of the fixed effects model. This again suggests that the pooled and the random effects models may have over stated the impact of climate change on the selected crops in Nigeria.

From the estimated fixed effects model, the MSE (0.000036) values which show the extent the slope and the intercept are correctly estimated reported a much lower values for the fixed effects model than for the OLS model (0.000636) indicating that the estimated fixed effects model is more efficient. Also standard error of estimate which is a measure of the likely error that may occur if the model is used for estimation also report a lower value for the fixed effects model (361631) than the pooled model (598061). The R^2 values which explain how much variation in crop yields accounted for by the climatic factors, reported (0.22) and (0.36) for the pooled model and the fixed effect model respectively, indicating again that the fixed effects model was able to account for variation in crop yield better than the pooled model. The differential intercept coefficients D2(-1827142), D3(-1291236), D4(-1275978), D5(-1120681), D6(-1524117), D7(-1447203) and D1(-4129486) were highly significant indicating the presence of individual heterogeneity, thus suggesting that perhaps the seven crops exhibit individual uniqueness or that the effect of climatic factors on the different crops are not the same.

4. Conclusion

Temporal and spatial analysis of key weather variables shows strong evidence of climate change in Nigeria. The quantification of the impact of this change in climate on some selected crops in Nigeria using three panel data models shows the presences of unobserved effects and that the fixed effects model was able to remove this unobserved effect better than the random effect model. The study also revealed that climate parameters were not significant suggesting that the observed increase in temperature and precipitation have not manifested significantly on production of some selected crops in Nigeria within the study periods. The estimated intercept coefficients for the fixed effects model were all significant indicating the presence of individual heterogeneity. It also suggests that the different crops are likely to be impacted differently by climate change the study therefore conclude that for efficient estimation of the impact of climate change. Therefore, the study concludes that for efficient estimation of the impact of climate change on crop production in Nigeria, the fixed effect model is better than the random effect model and the pooled model. Given the presence of individual heterogeneity, crop specific adaptation/mitigation policy would be more appropriate than national policy for the control of the impact of climate change on crop yield in Nigeria.

References

Akor, G. (2012). Climate Change and Agriculture: The Nigerian Story. Conference Presentation, FES, Ghana, Acrra.

Apata, T. G. (2010). Effects of global climate change on Nigerian agriculture: an empirical analysis. CBN Journal of Applied Statistics, **2**(1).

Ayinde, O. E., Muchie, M. and Olatunji, G.B. (2011). Effect of climate change on agriculture production in Nigeria: a co-integration model approach.

Ayoade, J. O.(2002). Introduction to Agro-Climatology. Vantage Publishers, Ibadan, Nigeria.

Bello, O. B., Ganiyu, O. T., Wahab, M.K.A., Afolabi, M.S., Oluleye, F., Igs, A., Mahmud, J., Azeez, M. A., Abdulmalig, S.Y. (2012). Evidence of climate change impact on agriculture and food security in Nigeria. International Journal of Agriculture and Forestry, **2**(2): 49–55.

Deressa,T, Hassan, R. and Poonyth, D. (2005). Measuring the impact of climate change on South African agriculture: the case of sugar cane growing regions. Agrekon, **44**(4).

Deschanes, O. and Greenstone, M. (2007). The economic impacts of climate change: evidence from agricultural output and random fluctuations.

Elhorst, J.P. (2003). Specification and estimation of spatial panel data models. International Regional Science Review, **26**: 244–268.

Feres, F., Reis, E. and Speranza, J. (2011). Assessing the impact of climate change on the Brazilian agricultural sector. A publication of the Institute for Applied Economic

Gujariti, D. (2009). Basic Econometrics (5th ed.). New York: The McGraw – Hill.

IPCC (2009). Climate Change impacts, adaptation and vulnerability. Contribution of working Group 11 to the fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK.

Kayode, A.O (2010), The effect of weather on cocoa production in different agro-ecological Zones in Nigeria. World Journal of Agricultural Sciences, **6**(5): 609–614.

Lioyed,T., Oliver, M. and Osei, R. (2001). Problems with pooling in panel data for developing countries: the case of aid and trade relationships. CREDIT Research Paper 01/14, School of Economics, University of Nottingham.

Mendelsohn, R. (2000). Measuring the Effect of Climate on Developing Country Agriculture: Two Essays on Climate Change and Agriculture. FAO Economic and Social Development Paper 145, Edwaed Elgar Publishing, Chattenhen, UK.

Obioha, E. (2007). Climate change population drift and violet conflict over land resources in North Eastern Nigeria. J. HumEcol., **23**(4): 311–324.

Odjugo, P.A.O. (2007). The impact of climate change on water resources: global and regional analysis. The Indones J. Georg. **39**: 23–41

Park, H. M. (2011). Practical Guides to Panel Data Modeling: A Step by Step Analysis using Stata. http://www.iuj.ac.jp/faculty/kucc625.

Park, H. M. (2009). Linear Regression Models for Panel Data using SAS, Stata, LIMDEP and SPSS. A Study Guide of the Indiana University Information Technology Services. http:/www.indiana.edu/stamathn/stat/all/panel

Wolfram, S. and Roberts, M. (2008). Estimating the impact of climate change on crop yields: the importance of nonlinear temperature effects. A National Bureau of Economic Research (NBER), Working Paper. http://www.nber.org/paper/w13799

Appendix I

Stata codes for generating the results

```
xtset id year
xtsum y x1 x2 x3 x4, x5, x6, x7
pwcorr y x1 x2 x3 x4, sig
estat vif
xtserial y x1 x2 x3 x4, x5, x6, x7
reg y x1 x2 x3 x4, x5, x6, x7
xtreg y x1 x2 x3 x4, x5, x6, x7, fe; est sto fe
xtreg y x1 x2 x3 x4, x5, x6, x7, re; est sto re
estat hottest
estat ovtest
xtdolshm y x1 x2 x3 x4, x5, x6, x7
pescadf x, lags(1)
pescadf d.x, lags(1)
```

Appendix II

OLS regression for pooled model

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.472[a]	.223	.194	598061.15074	.223	7.703	7	188	.000	.402

a. Predictors: (Constant), rxt, squaredtemp, price, loan, squaredrain, temperature, rainfall

b. Dependent Variable: cropyield

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	19286890577344.484	7	2755270082477.783	7.703	.000[a]
	Residual	67243302323624.860	188	357677140019.281		
	Total	86530192900969.340	195			

a. Predictors: (Constant), rxt, squaredtemp, price, loan, squaredrain, temperature, rainfall

b. Dependent Variable: cropyield

Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
1 (Constant)	-3719003.386	3098144.583		-1.200	.231		
rainfall	-363.344	845.063	-.867	-.430	.668	.001	983.416
squaredrain	-.042	.016	-.199	-2.620	.010	.719	1.392
temperature	300766.480	189936.650	1.610	1.584	.115	.004	250.107
squaredtemp	-5186.582	2816.858	-1.732	-1.841	.067	.005	213.967
loan	50.686	10.253	.357	4.944	.000	.793	1.261
price	-27.952	9.830	-.193	-2.844	.005	.899	1.113
rxt	12.009	31.872	.754	.377	.707	.001	967.793

a. Dependent Variable: cropyield

Appendix III

Estimated results for the fixed effects model

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.851[a]	.725	.705	361631.46595	.725	36.897	13	182	.000	.407

a. Predictors: (Constant), D7, price, D6, D2, D5, loan, rxt, D4, squaredrain, squaredtemp, D3, temperature, rainfall

b. Dependent Variable: cropyield

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	62728721176761.080	13	4825286244366.236	36.897	.000[a]
	Residual	23801471724208.270	182	130777317165.979		
	Total	86530192900969.340	195			

a. Predictors: (Constant), D7, price, D6, D2, D5, loan, rxt, D4, squaredrain, squaredtemp, D3, temperature, rainfall

b. Dependent Variable: cropyield

Coefficients[a]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
1 (Constant)	4129502.697	2050775.568		2.014	.046		
rainfall	-521.849	547.787	-1.245	-.953	.342	.001	1130.165
squaredrain	-.005	.010	-.021	-.450	.653	.672	1.488
temperature	-113781.673	125285.796	-.609	-.908	.365	.003	297.626
squaredtemp	872.860	1870.214	.291	.467	.641	.004	257.964
loan	39.607	6.583	.279	6.017	.000	.703	1.422
price	-.404	6.670	-.003	-.061	.952	.714	1.401
rxt	19.156	20.715	1.202	.925	.356	.001	1118.073
D2	-1827142.227	110917.733	-.962	-16.473	.000	.443	2.258
D3	-1291236.439	106040.492	-.680	-12.177	.000	.485	2.064
D4	-1275978.072	111150.129	-.672	-11.480	.000	.441	2.267
D5	-1120681.417	108007.927	-.590	-10.376	.000	.467	2.141
D6	-1524117.401	113217.703	-.803	-13.462	.000	.425	2.352
D7	-1447202.844	116697.741	-.762	-12.401	.000	.400	2.499

a. Dependent Variable: cropyield

Appendix IV

Estimated results for the random effects model

Random-effects GLS regression		Number of obs =	196
Group variable (i): entity		Number of groups =	7
R-sq: Within = 0.1291		Obs per group: min =	28
Between = 0.5428		avg =	28.0
Overall = 0.2229		max =	28
Random effects u_i ~ Gaussian		Wald chi2(7) =	53.92
corr(u_i, X) = 0 (assumed)		Prob > chi2 =	0.0000

Cropyield	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
RAIN	-363.3438	845.0628	-0.43	0.667	-2019.636	1292.949
RAIN2	-.0422437	.0161233	-2.62	0.009	-.0738447	-.0106427
TEMP	300766.7	189936.7	1.58	0.113	-71502.43	673035.8
TEMP2	-5186.586	2816.859	-1.84	0.066	-10707.53	334.3562
RXT	12.00898	31.87238	0.38	0.706	-50.45973	74.47769
LOANS	50.68623	10.25284	4.94	0.000	30.59104	70.78142
PRICE	-27.9519	9.830025	-2.84	0.004	-47.2184	-8.685404
_cons	-3719007	3098146	-1.20	0.230	-9791261	2353247
sigma_u	0					
sigma_e	361631.48					
rho	0	(fraction of variance due to u_i)				
