Determinants of food production in Sub Saharan Africa:
The impact of policy, market access and governance

Federica Di- Marcantonio1, Cristian Morales-Opazo2, Jesus Barreiro-Hurle 3 and Mulat Demeke 4

1 JRC- European Commission: email: federica.di-marcantonio@jrc.ec.europa.eu
2 FAO- ESA email: cristian.moralesopazo@fao.org
3 European Commission email: jesus.barreiro-hurle@ec.europa.eu
4 FAO- ESA email: mulat.demeke@fao.org

Paper prepared for presentation at the EAAE 2014 Congress
'Agri-Food and Rural Innovations for Healthier Societies'

August 26 to 29, 2014
Ljubljana, Slovenia

Abstract

This study examines the relationship between policy, market access, country governance indicators and food production in 41 African countries. Based on a cross-country panel sample, a fixed-random effect models is employed to test the hypothesis that beyond agricultural inputs and macroeconomic reforms other exogenous factors could foster food production. Our findings show that improving food-agricultural inputs enhance production, while conflicts, food aid and geographic location such as landlocked countries negatively affect food production. Exogenous factors influencing production response include rainfall, market access, and education. Both governance and education can indirectly improve food production by enhancing growth, through investment in infrastructures, and human capital.

Keywords : food , production, market, governance

1- Introduction

This study tries to provide some additional evidence on the role of different factors in explaining increased domestic food availability. Thus, in addition to the classical input factors approach, we also consider the role of public governance stability and market access as variables that could influence the evolution and the development of food production. The rest of the paper is organized as follows: section two presents the methodological basis of our estimations, which is followed by a description of the data sources; in section three the empirical results of the different models and specifications are presented; some conclusive remarks are presented in the final section.

2- Methodology

In order to investigate the impact of policy, market access and governance on food production in SSA, and following Barrios et al (2006), we specified a simple empirical production function where for any country i at time t agricultural output is a function of a number of inputs grouped into five main categories:

$$\begin{aligned}\log(Y_{it}) = b_1 &+ \underbrace{b_2 log(Land_{it}) + b_3 log(Irrigation_{it}) + b_4 log(Fertilizer_{it})}_{Agricultural\ Input} + \\ &+ \underbrace{b_5 log(Animals\ for\ Transport\ LU_{it}) + b_6 log(Tractors_{it}) + b_7 log(Labour_{it})}_{Agricultural\ Input} + \\ &+ \underbrace{b_8 (Urbanization_{it}) + b_9 (Landlocked_i) + b_{10} log(Telephone\ Line_{it})}_{Market\ Access} + \\ &+ \underbrace{b_{11} log(Export_{it}) + b_{12} log(Cereal\ Aid_i) + b_{13} (Policy\ Score_{it}) + b_{14} (Inflation_{it})}_{Macroeconomic\ Enviroment} + \\ &+ \underbrace{b_{15} (School\ Enrollment_{it})}_{Human\ Capital\ Access} + \underbrace{b_{16} (Rain\ Dummy_{it}) + b_{17} log(Battle_{it})}_{Environmental\ Factors} + h_t + \eta + e_{it}\end{aligned} \quad [1]$$

Where *Y* is the food total output measured as net food production value in constant 2004-2006 international dollars. Independent variables such as *Agricultural Inputs and Human Capital are directly linked to production, while Market Access, Macroeconomic Environment,* and *Environmental factors* represent the business, policy and external environment within which production takes place.

3. Data sources

We used a macro approach with country-level data for 41 SSA countries. For each model the number of countries included was selected to maximize the number of countries that allowed obtaining a balanced panel data set by reducing the number of countries and/or the length of time series for the period 1968 to 2008. The data is drawn from various cross-country datasets. The first main dataset is obtained from FAOSTAT, which provides aggregate indicators of the depend variable, production inputs, in addition to measures of external food aid (CERAID), urbanization (URB), active population in agriculture (LABOUR) and agriculture export (EXP). A second set of variables incorporates figures from the World Bank's World Development Indicators (2010) and include: inflation (INF) school enrolment (SCHOOL), Battle-related deaths (BATTLE), Telephone lines (per 100 people) (TELINE), Combined Polity Score (POLSCO). The effect of climate condition, added a dummy variable (RAIN) calculated as the deviance from the annual average country precipitation, was taken from (NCAS, 2008). We also include a dummy variable to capture whether the country has a direct access to the sea. The LANDLOCK takes one if a country is landlocked and takes zero otherwise. These show that there is considerable variation in all of the variables used.

4. Estimation

Before proceeding to the analysis, however, we first tested for the presence of a stochastic trend, i.e. whether our data was stationary. In this regard, we applied Fisher unit root tests for panel data according to Maddala and Wu (1999) as well as Im et al (2003) tests (hereafter IPS), which improves the Levin et al (2001) test (hereafter LL) by relaxing the assumption of a common ρ: the IPS test runs a separate unit test for each of the units and computes the mean of the t-statistic of each independent Augmented Dickey-Fuller test. IPS fits only balanced data with the same number of observations per unit. The Fisher test of Maddala and Wu (1999) and the IPS test are directly comparable. Since our data are only slightly unbalanced, we computed both statistics. For most of the variables in our sample, the null hypothesis of a unit root is clearly rejected.

The estimation strategy started with the simplest model, the Pooled OLS, which does not take into account the panel structure of the data. In order to capture country-specific effects we estimated two panel data models, one with cross-section fixed effects and another with random effects. The decision on which model to choose between the two models was made based on the Hausman test (Hausman, 1978). The results show that fixed effects are the preferred models across all specifications except the first one (Tables 2 to 4). Notwithstanding, the use and appropriateness of fixed-effects (FE) regression model required further investigation on a number of statistical properties. First, to deal with cross-sectional correlation we look at Pasaran statistics, which test for cross-section independence in the residuals (De Hoyos and Sarafidis, 2006). As shown in tables 2 to 4 the results of this test rejects the null hypothesis in all the three cases (the null hypothesis of Pasaran test is that of cross- sectional independence). Thus, we can conclude that there is spatial correlation in our data. However, the fixed effect model is estimated assuming the homoskedasticity of the residuals. When heteroskedasticity is present, the standard errors of the estimates will be biased and one should compute robust standard errors correcting for the possible presence of heteroskedasticity (Arellano, 2003). To deal with this issue, we calculate a modified Wald statistic for groupwise heteroskedasticity (Franklin, 2005) in the residuals of a fixed effect regression model.

All the above tests suggest that in order to estimate a Fixed-Effects model we need to take into account spatial correlation and panel heteroskedasticity. As a consequence, following Beck and Katz (1995) we computed a panel corrected standard errors (PCSE). Further tests for heteroskedasticity and cross-sectional dependence show that all problems are present. All models have been estimated using the Robust Variance in order to correct for heteroskedasticity and to obtain robust standard errors.

However, the presence of time-invariant variables in a panel data regression model poses another problem for the analysis. This is also shown in the Table 3 in annex, where the ratio of the between variance, in some cases, is more than six times higher than the within variance. To handling the longstanding problem of time-invariant variables in an FE model Plümper and Troeger (2007) have recently proposed an estimator, labeled FE vector decomposition (FEVD), a three-stage model which improves on the efficiency of LSDV in the FE model and along the way solves the problem of non- identification of the coefficients on time-invariant variables in this model.

All models were estimated in the five different specifications. First, we set up the baseline model (M1), in which the independent variables included six agricultural inputs; two market access (land locked, rate of urbanization); two macroeconomic factors (export and cereal aid), and two environmental aspects (rain dummy and battle). The policy score and school enrolment entered in model M2, while in model M3 another macroeconomic variable was included (inflation). All variables, except urbanization, landlocked, school enrolment and policy score were used in logarithmic form.

5. Results

In this section, we present the results of our estimation strategy. Starting with the results of our basic estimator and its possible variations, we also present the results of robustness tests that we use. Tables 2-4 display the results of the different models for the different specifications. We started our analysis by estimating the first model (M1). This model does not include policy score, school enrolment, battle and inflation as explanatory variables but contains all 41 SSA countries as shown in Table 2. In Column 1 we reported results of regressions based on the OLS estimator, while columns 2-5 show results for FE/RE, PCSE and FEVD estimators. As can be seen, for all our estimators, inputs turn out to be statistically significant determinants of agricultural output with the expected sign, thus providing some support for our model. However, it appears that the FEVD procedure allows us to improve the significance of the coefficients of our inputs variables, though maintains constant the size of their impact.

Comparing the individual coefficients (R^2) across the different macro groups, one notices that in general the coefficient on the log tractors is more than double the value of those on irrigation and labour, however animals for transport remains the most important input factor. This may be because we are unable to control for the quality of these inputs, which are likely to be higher in relatively richer SSA nations. In contrast, we found that the food production output is substantially lowered in landlocked countries and those who have experienced long term conflicts. Furthermore, as for the overall impact of agricultural exports, we also find a positive impact on food production due to the increasing rate of urbanization and rain. The other control variables are also correlated with the dependent variable as expected and generally statistically significant across the different models. Indeed, as shown in columns 2 and 4, food aid by deterring local producers negatively impacts the overall food output. However, by accounting for heteroskedasticity, cross-sectional dependence and serial correlation (column 4) we noted how some variables become not significant (log land, logbattle and references periods) thought the sign continue to be consistent with the FE model.

We next included our school enrolment and policy proxy, as shown in the second set of regressions of Table 3. The first point to note is that the inclusion of these variables has little effect on the coefficients of the other explanatory variables with the exception of agricultural land, which almost doubled in both the FE and FEVD regressions. More importantly, however, our empirical results supported the hypothesis related to these two variables although they lose level of significance among the different models. As shown in regressions (2), the coefficients of policy score are positive and highly significant. This indicates that a country with a better democracy infrastructure will produce more agricultural output with the same amounts of agricultural inputs. However, this variable turned to be not significant in regression 4 (PCSE) where agricultural inputs, export and rain dummy seems to be the main determinants of agricultural output. On the contrary the figure for cereal aid is -0.004, which indicates that an increase in food aid of 1% will lower the overall output by 0.004%, given the same amounts of agricultural inputs. Estimates of including the inflations and telephone line variables in the empirical equation are shown in the third set of regressions (M3) as reported in Table 4. Differently form the previous two additional variables, the inclusion of an extra market access indicator (telephone line) significantly affect food output by increasing the impact of some key inputs. In contrast, inflation appears not to have played a significant role in variations in food output in SSA countries.

Table 2. Base model results (M1)

Variable	OLS	FE	RE	PCSE	FEVD
Log Food Net Production Value (constant 2004-2006 1000 I$)					
Agricultural Inputs					
Log land	.026**	.061**	.053**	.040	.061***
Log irrigation	.088***	.042***	.039***	.091***	.043***
Log fertilizer	.006	.012***	.013***	.011**	.012***
Log animals for transport in LU	.063***	.368***	.345***	.102***	.368***
Log tractors (# of tractors)	.009	.084***	.081***	.038***	.084***
Log labour (active pop.in Agric.)	.550***	.042	.150***	.557***	.042***
Market access					
Landlocked (dummy)	-.046	(omitted)	-.222*	-.054	-.206***
Urbanization	.009***	.0052***	.005***	.011***	.005***
Macroeconomic Environment					
Log export	.229***	.061***	.065***	.079***	.061***
Log Food Aid	-.027***	-.005***	-.005***	-.004**	-.005**
Environmental Factors					
Rain dummy	.014	.029***	.030***	.025***	.029***
Conflict	-.001	-.009***	-.008***	-.002	-.009***
Periods					
p1 (1968-1983)	-.004	-.082***	-.065***	.008	-.082***
p3(1994-2008)	.053*	.165***	.154***	.005	.165***
h hat					1***
_cons	5.341***	5.921***	5.571***	5.867***	5.998***
N	1640	1640	1640	1640	1640
r2	.925	.794		.978	.989
r2_a	.925	.787			

Number of Observations	1640	
Number of Countries	41	
Period covered	1968-2008	
Fixed effects model		
Joint Significant Tests	F(39, 1587)	237.41
R2 (within)	0.7940	
R2 (between)	0.7329	
R2 (overall)	0.7359	
Random effects model		
Joint Significant Tests	Wald $Chi^2(14)$=	6182.73
R2 (within)	0.7818	
R2 (between)	0.7859	
R2 (overall)	0.7855	
*FGLS: xtpcse**		
Joint Significant Tests	Wald $Chi^2(14)$=	7390.25
Rho	.8632431	
[1] Hausman Test for the choice between fixed or random effects Model	$Chi^2(12)$=	-698.84
	Prob>Chi^2=	-
	Chosen model:	-
[2] Pesaran's test of *H0: cross-sectional independence*	23.725	
	Prob	0.000
[3] Modified Wald test for in fixed effect regression model	chi2 (40) =	2764.18
H0: sigma(i)^2 = sigma^2 for all	Prob>chi2 =	0.0000
[4] Modified Wald test for Wooldrige Test		
H0:no first-order autocorrelation	F(1,39)=	57.970
	Prob>F=	0.000

Table 3. Model results (M2)

Variable	OLS	FE	RE	PCSE	FEVD
Log Food Net Production Value (constant 2004-2006 1000 I$)					
Agricultural Inputs					
Log land	.052***	.104***	.081***	.031	.104***
Log irrigation	.083***	.051***	.051***	.116***	.051***
Log fertilizer	.000	.008	.008	.012*	.008**
Log animals for transport in LU	.027*	.340***	.312***	.067***	.340***
Log tractors (# of tractors)	.010	.101***	.091***	.030**	.101***
Log labour (active pop.in Agric.)	.574***	.011	.150***	.582***	.011
Market access					
Urbanization	.007***	.008***	.008***	.012***	.008***
landlocked (dummy)	-.072*	(omitted)	-.193	-.002	-.194***
Macroeconomic Environment					
Log export	.246***	.045***	.055***	.084***	.045***
Log Food Aid	-.035***	-.003*	-.003*	-.004*	-.003
Policy Score	.003	.007***	.007***	.002	.007***
Human Capital					
School enrolment, primary (%gross)	5.66E-03	.001**	.001**	.001	.001***
Environmental Factors					
Rain dummy	.012	.030***	.030***	.026***	.030**
Log Battle	.007*	-.007***	-.007***	-.001	-.007***
Periods					
p1 (1968-1983)	-.039	-.063***	-.046***	.018	-.063**
p3(1994-2008)	.032	.109***	.101***	-.007	.109***
h_hat					1***
_cons	5.454***	6.110***	5.711***	6.037***	6.189***
N	1258	1258	1258	1258	1258
r2	.906	.779		.981	0.989
r2_a	.905	.769			

Number of Observations	1258				
Number of Countries	34				
Period covered	1975-2008				
Fixed effects model					
Joint Significant Tests	F(36, 1206)	241.83			
R2 (within)	0.7788				
R2 (between)	0.6387				
R2 (overall)	0.6458				
Random effects model					
Joint Significant Tests	Wald	4201.53			
R2 (within)	0.7754				
R2 (between)	0.7350				
R2 (overall)	0.7370				
*FGLS: xtpcse**					
Joint Significant Tests	Wald	5213.14			
Rho	.8535554				
[1] Hausman Test for the choice between fixed or random effects Model	$Chi^2(12)=$ $Prob>Chi^2=$ Chosen model:	493.01 0.000 **Fixed effects model**			
[2] Pesaran's test of cross sectional independence *H0: cross-sectional independence*	 Prob	17.203 0.000			
[3] Modified Wald test for groupwise heteroskedasticity in fixed effect regression model *H0: sigma(i)^2 = sigma^2 for all i*	chi2 (37) = Prob>chi2 =	3084.68 0.0000			
[4] Modified Wald test for groupwise heteroskedasticity Wooldrige Test for autocorrelation *H0:no first-order autocorrelation*	 F(1,36)= Prob>F=	 59.826 0.000			

Table 4: Model results (Model 3)

Variable	OLS	FE	RE	PCSE	FEVD
Log Food Net Production Value (constant 2004-2006 1000 I$)					
Agricultural Inputs					
Log land	.042**	.087***	.068***	.044	.087***
Log irrigation	.084***	.081***	.086***	.111***	.081***
Log fertilizer	.015	-.0002	-.0006	.012*	-.0002
Log animals for transport in LU	.032*	.306***	.278***	.065**	.306***
Log tractors (# of tractors)	.027**	.087***	.077***	.018	.087***
Log labour (active pop.in	.527***	-.017	.187***	.589***	-.017
Market access					
Urbanization	.008***	.011***	.011***	.011***	.011***
Landlocked (dummy)	-.117**	(omitted)	-.123	-.015	-.124***
Log Telephone line	-.087***	.027*	.027	.053*	.027***
Macroeconomic Environment					
Log export	.254***	.051***	.063***	.098***	.051***
Log Food Aid	-.036***	-.001	-.0004	-.003	-.000
Policy Score	.006*	.008***	.007***	.003	.008***
Inflation	.0001	1E-03	3E-03	4E-03	1.22E-06
Human Capital					
School enrolment, primary	.001	.001***	.001***	.001	.001***
Environmental Factors					
Rain dummy	-.008	.034***	.034***	.024**	.034***
Log Battle	.018***	-.008***	-.007**	-.001	-.008***
Periods					
p1 (1968-1983)	-.041	-.039*	-.018	.020	-.039
p3(1994-2008)	.065	.072***	.057***	-.015	.072***
h hat					1***
cons	5.328***	6.861***	5.880***	5.907***	6.913***
N	1008	1008	1008	1008	1008
r2	.903	.755		.984	.990
r2_a	.901	.742			

Number of Observations	1008	
Number of Countries	28	
Period covered	1981-2008	
Fixed effects model		
Joint Significant Tests	F(35, 955)	233.51
R2 (within)	0.7553	
R2 (between)	0.5982	
R2 (overall)	0.6044	
Random effects model		
Joint Significant Tests	Wald $Chi^2(18)$=	2934.73
R2 (within)	0.7488	
R2 (between)	0.7448	
R2 (overall)	0.7448	
*FGLS: xtpcse**		
Joint Significant Tests	Wald $Chi^2(18)$=	4914.88
Rho	0.8430392	
[1] Hausman Test for the between fixed or random Model	$Chi^2(12)$=	155.78
	Prob>Chi^2=	0.000
	Chosen model:	**Fixed effects model**
[2] Pesaran's test of cross sectional independence		10.792
H0: cross-sectional	Prob	0.000
[3] Modified Wald test for groupwise heteroskedasticity in fixed effect regression	chi2 (46) =	2671.90
H0: sigma(i)^2 = sigma^2	Prob>chi2 =	0.0000
[4] Modified Wald test for groupwise heteroskedasticity Wooldrige Test for autocorrelation		
H0:no first-order	F(1,35)=	55.130
	Prob>F=	0.000

However, though it is not possible to theoretically assess the direction of the bias, we clearly see the empirical difference between the coefficients of Table 2 (M1) and those of Table 3 and 4.

Translating the results in policy actions, it means that once we control for agricultural inputs, infrastructures, macroeconomic and environmental variables; the parameter estimates for the other variables “explain” the remaining direct effect.

Thus, by looking at our key variables, five important messages can be derived from our estimates.

Despite the progress of developing countries in regions like Asia and Latin America, many SSA households are food insecure and the continent is the largest receiver of food aid in the world. Food productivity is, and will be, one of the most important challenge of the world. But this challenge takes another crucial face in SSA countries.

The roles of investment in agriculture and the expenditure on agricultural research and development together with the macroeconomics reform during the last three decades were keys factors to the food security in this region of the world. The situation is not the same across all the countries of the region, and the different incentives and disincentives are playing different roles, but some common factors can influence the performance in food and agriculture.

Incentives and sustainability are two of the appropriate words that we should use if we think in improving food productivity next years. Our study demonstrates empirically that considering the same amount of inputs democratic countries produce more food. It is clear that governance is a clear factor that can influence in food performance in SSA countries. Democracy and good governance at local, regional, and world level are crucial. If we consider that good governance could imply low level of conflict that will be useful in terms of food production. Our study proves something well knows: countries under conflicts in SSA have more problems in terms of food security.

We realize during the analysis that landlocked countries are facing the most difficult part of this challenge and we insist in the necessity to improve technology transfer to these countries.

As agricultural output grows, SSA farmers should become more commercialized and be focused on satisfying the demands of consumers so that their enterprises can be remunerative and provide them with decent livelihoods. Due to the smallness of their operations, farmers in

SSA require effective policies and institutions to support them. Most commodities face an unfavorable market and policy environment, the former being more important than the latter (MAFAP, 2013).

This line of research could continue with in more detailed analysis by country where we can use another variables than can explain the evolution of the food production.

References

Arellano, M. (2003). "Modelling Optimal Instrumental Variables For Dynamic Panel Data Models," Working Papers wp2003_0310, CEMFI.

Barrios, S; Ouattara, B and Strobl, E. (2008): The impact of climatic change on agricultural production: Is it different for Africa? Food Policy 33 287–297.

Beall, J. (2010). Urbanization and Development. Published to Oxford Scholarship Online: January 2011.

Beck, N; Katz, J. (1995). What to do (and not to do) with time-series cross-section data. The American Political Science Review 89, 634–647.

Im K., Pesaran H. and Y. Shin [2003] Testing for Unit Roots in Heterogeneous Panels. Journal of Econometrics, 115, 53-74.

IMF. (2003). The IMF's Enhanced Structural Adjustment Facility (ESAF): Is It Working? Report.

Keefer, P and Knack, S. (1997). Why don't poor countries catch up? A cross-national testof an institutional explanation. Economic Inquiry 35, 590–602.

Knack, S and Keefer, P. (1995). Institutions and economic performance: cross-country tests using alternative institutional measures. Economics and Politics 7, 207– 227.

Levin, A., C, Lin and C.-S.J. Chu. (2002). "Unit Root Tests in Panel Data: Asymptotic and Finite-sample Properties", Journal of Econometrics, 108(1): 1-24.

Lio, M., & Liu, M-C. (2008). "Governance and agricultural productivity: a cross-national analysis," Food Policy, Vol. 33, pp. 504-512.

Maddala, G.S. and S. Wu. (1999). "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test", Oxford Bulletin of Economics and Statistics, 61(0): 631-652.

Pl¨umper, T. and Troeger, V. (2007). "Efficient Estimation of Time- Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects", Political Analysis, 15(2): 124-139.