The following supplement accompanies the article

The role of risk in the context of climate change, land use choices and crop production: evidence from Zambia

Alessandro De Pinto*, Vincent H. Smith, Richard D. Robertson

*Corresponding author: a.depinto@cgiar.org

Climate Research 79: 39–53 (2019)

Supplement.

Text S1. We consider an agricultural production system in which crops compete for shares of a fixed area of land. A household is assumed to face a one period consumption budget B in which:

$$B = W(\boldsymbol{m}) + \Pi(\boldsymbol{p}, \boldsymbol{y}, \boldsymbol{c}) * A - C(\boldsymbol{q}, \boldsymbol{z}, \boldsymbol{h}) \quad \text{eq. A1}$$

where W indicates household wealth and $\boldsymbol{m}$ is a vector of household characteristics that affect its wealth, Π is the per-hectare net revenues function of a vector of crop market prices $\boldsymbol{p}$**,** a vector of crop yields $\boldsymbol{y}$, and of cost of using variable inputs $\boldsymbol{c}$ such as labor and fertilizers**.** A is the household landholdings, and C identifies the cost of quasi-fixed factors that must be allocated across production activities and can generate jointness in acreage decisions. C is a function of quasi-fixed inputs $(\boldsymbol{q})$ such as quantity of available farm tools and machinery, a vector of exogenous climate variables $(\boldsymbol{z})$**,** and household characteristics $(\boldsymbol{h})$ that can function as binding constraints on field activities such as the number of household member and the gender of the head of household.

Crop-specific net revenues are defined as $\Pi_j = a_j(p_j y_j - c_j)$ where a_j is the area allocated to $j = 1, \ldots\ldots, J$ crops. Each household faces a location-specific set of prices, yields and costs. Location and household will be later identified with subscripts l and n, respectively. We omit these subscripts for now for readability purposes. Note that $\sum_{j=1}^{J} a_j = A.$. Revenues, $R_j = \sum_j p_j y_j a_j$, are stochastic because output prices and yields have some distribution with finite means and variances $\left(\bar{p}_j, \Omega_{p_j}\right)$ and $\left(\bar{y}_j, \Omega_{y_j}\right)$, respectively, and their realized values are not known by farmers when land allocation decisions are made. Conversely, all production costs are assumed to be known with certainty by households at that time. The volatility in prices and yields implies that net revenues are also stochastics.

The household, who is assumed to be a price-taker, maximizes expected utility derived from farming available land holdings by allocating a_j to each available crop j:

$$\max_{a_j}\{EU[\Pi, \Omega, C, |W, A]\} \qquad \text{eq. A2}$$

where U is a utility function describing the impacts on the decision maker's utility of the level of profits and their volatility Ω, conditioned on wealth, including land endowments.

Thus, there is an optimal area allocation $a_j^* = F[\Pi, \Omega, C, W, A]$ to each crop that solves the farmer's utility maximization problem. We assume that $a_j^* = F[\Pi, \Omega, C, W, A]$ can be expressed as $a_j^* = A * f[\Pi, \Omega, C, W, A,]$. Wu & Segerson (1995) assume that their land use function is linear homogenous of degree one in area and therefore the function can be written removing area itself as a potential determinant of land allocation decisions and impose constant returns to acreages. However, removing area from the function arguments is not necessary except as a convenience. In fact, the homogeneity of degree one assumption could be dropped altogether. Retaining total farm size among the arguments of the optimal share function can account for differences in fixed costs of multiple activities or risk-spreading decisions available to farmers that manage farms of different sizes. Accordingly, for each crop j there is an optimal share allocation function:

$$s_j^* = \frac{a_j^*}{A} = f[\Pi, \Omega, C, W, A] \qquad \text{eq. A3}$$

Farm-level allocation shares (s^O), which are observed in the household survey, are a function of all crop restricted profits, depend on all crops' risk/returns profile, and on the cost of quasi-fixed inputs. The probability that a farmer chooses to grow a particular crop j is given by $P_j = Pr(U_j > U_i \forall\, j \neq i)$ and the expected area allocated to that crop is given by $A * P_j$. Therefore, the expected share s of farmland allocated to crop j is $s_j = \frac{1}{A}[A * Pr(U_j > U_i \forall\, j \neq i)]$. This means that the share allocated to a crop j is equal to its probability P_j to be chosen by a farmer. Acknowledging that $U = V + \boldsymbol{\xi}$ where V represents a knowable-by-all component of the utility function while $\boldsymbol{\xi}$ is known to the farmer but unobserved by the researcher and assuming that $\boldsymbol{\xi}$ has an *iid* Type 1 EV distribution, then:

$$s_j = \frac{exp\{\boldsymbol{E}[\boldsymbol{V}_j]\}}{\sum_{i=1}^{J} exp\{\boldsymbol{E}[\boldsymbol{V}_i]\}} \qquad \text{eq. A4}$$

We then can rewrite the optimal share allocation function for a crop as:

$$s_j^* = \frac{exp[s(\Pi_j, \Omega_j, C, W, A)]}{\sum_{i=1}^{J} exp[s(\Pi_j, \Omega_j, C, W, A)]} \qquad \text{eq. A5}$$

Equation A3 explicitly accounts for the influence of profit variability on land allocation decisions by farmers who account for both expected returns and the volatility of those returns in their land use decisions. To operationalize the measure of revenues variability in the estimated model, the variable Ω_j is constructed using the standard formula to compute the variability of the product of two stochastic independent variables yield and price[1]:

$$\Omega_j = E(y_j^2) * E(p_j^2) - [E(y_j p_j)]^2 \qquad \text{eq. A6}$$

Wu & Segerson (1995) assume that the optimal share function is linear in parameters while Chavas & Holt (1990) use a first order expansion to linearize the optimal acreage function. Following this literature, we assume that the optimal share function for each crop can be approximated by a linear in parameters combination of explanatory variables such that $\ln\left(\frac{s_{jln}}{s_{0ln}}\right) = \boldsymbol{\beta_j X_{jln}} + \boldsymbol{\xi_j}$, where $\boldsymbol{X_{jln}}$ is a vector of explanatory variables $(R_{jl}, c_{ln}, \Omega_{jl}, q_n, z_l, h_n, m_n, A_n)$, where l identifies the location and n the n^{th} household, $\boldsymbol{\beta_j}$ is a vector of parameter to be estimated, $\boldsymbol{\xi_j}$ an error term and the subscript 0 in s_0 indicates a reference crop.

The parameters $\boldsymbol{\beta_j}$ are estimated using a pseudo-maximum likelihood as proposed by Mullahy (2015). We refer the interested reader to Mullahy (2015) for the calculation of the score equation and the parameters' asymptotic variance. Given s_j^*, the quasi-log-likelihood function to be maximized with respect to the parameters $\boldsymbol{\beta_j}$ is:

$$L = \sum_{n=1}^{N} \sum_{j=1}^{J} s_{nj}^{O} \log s_j^*(\boldsymbol{X}, \boldsymbol{\beta})$$

We estimate two model types. The first is a standard multinomial logit model; the second is a two-level nested multinomial logit model (see Figure 2). In both models, the probability of a crop being chosen is interpreted as the share of the available land to be allocated to the crop (Theil 1969, Berry 1994, Greene 2003). Nested multinomial logit models are estimated sequentially under assumptions analogous to the multinomial logit model with respect to $\boldsymbol{\xi}$, but with the error terms for crop shares correlated within each nest but uncorrelated among

[1] Due to data limitations the covariance between prices and yields is assumed to be equal to 0 both at the household and at the country level. In our specific case we assume that the prices are de-linked from the year-to-year weather conditions.

nests. The explanatory variables are partitioned with some used to choose among the nests and the others to choose among the options within each nest.
Under these assumptions, the probability that household *n* chooses alternative j ($j \in k$) can be derived from the product of two multinomial logit probabilities (McFadden 1977, Train 2003); that is,

$P_{nj} = P_{nk} * P_{nkj|k}$, where

$$P_{nk} = \frac{\exp\,(\boldsymbol{\beta}_k \boldsymbol{X_n} + \lambda_k I_{nk})}{\sum_{k'=1}^{K} \exp\,(\boldsymbol{\beta}_k \boldsymbol{X_n} + \lambda_k I_{nk})} \qquad \text{eq. A7}$$

and $\boldsymbol{X}_n$ is a vector of explanatory variables $(c_n, q_n, h_n, m_n, A_n)$ recorded at the household level, wealth and assets and $\boldsymbol{\beta}_k$ is vector of coefficients for $\boldsymbol{X}_n$, and where

$$P_{nkj|k} = \frac{\exp\,(\boldsymbol{\beta}_{kj} \boldsymbol{X}_{nj} / \lambda_k)}{\sum_{j' \in k} \exp\,(\boldsymbol{\beta}_{kj} \boldsymbol{X}_{nj} / \lambda_k)} \qquad \text{eq. A8}$$

and $\boldsymbol{X}_{nj}$ is a vector of explanatory variables $(R_{jl}, c_l, \Omega_{jl}, z_l)$ including crop-specific revenues and revenue variability recorded at the district level and $\boldsymbol{\beta}_{kj}$ is a vector of coefficient parameters specific to crop *j*. I_{nk}, often referred to the inclusive value of nest k, is defined as $I_{nk} = \ln\left(\sum_{j \in k} \exp\,(\boldsymbol{\beta}_{kj} \boldsymbol{X}_{nj} / \lambda_k)\right)$. I_{nk} is called the inclusive value or inclusive utility for alternative *k* in the first level. The inclusive value links the two levels of the nested logit model by bringing information from the bottom level into the upper level. In essence, $\lambda_k I_{nk}$ measures the expected value or utility to individual *n* of the alternatives available in particular nest. Equation A7 defines the marginal probability of choosing any alternative in nest *k* and equation A8 the conditional probability of choosing alternative *j* given that any alternative in nest *k* is chosen. We refer to the marginal probability as the upper-level model and to the conditional probability as the lower-level model, reflecting their relative positions in the hierarchy structure in shown in Figure 2.

Table S1. Parameter estimates for the Multinomial Logit model specification, (reference category "Others"). Significance codes: (***) 0.001; (**) 0.01; (*) 0.05

	Maize	**Millet**	**Cassava**	**Sorghum**	**Groundnuts**	**Beans**
Labor Costs	0.007635	-0.002503	-0.015430***	0.009592	0.030348***	-0.009298
Fertilizer Price	-0.000535	-0.000795*	0.002879***	-0.000126	0.000067	-0.001487
Farm Size	-0.149839***	-0.190442***	-0.288800***	-0.273203***	-0.140441***	-0.096833**
Female Head HH	-0.014339	-0.329084	-0.266493	0.183843**	0.232643	0.093139
Household Members	0.011308	0.013404	-0.013733	-0.000636	0.019042	0.02305
Value of Farm	0.000346**	-0.005883	-0.001625	0.000169	0.000283	0.000223

Assets						
Livestock	0.004198	0.008885	-0.004953	-0.00365	0.005589	0.004684
Off-farm revenues	0.252177	-0.935579	0.793245***	0.520654	-0.26068	-0.167617*
Rain Median	0.000474***	-0.000498	-0.000356***	-0.000672	0.000591	0.000302
Rain Interquartile Spread	-0.001162*	0.000183*	0.002813***	0.000387	0.000662	0.004819***
Temp. Median	-0.023936***	0.029311***	0.036948	-0.03295	0.012594	-0.021873
Temp. Interquartile Spread	0.006231	0.010766**	-0.012526***	0.030770**	-0.003366	0.001505
Distance median	0.170992*	0.896663***	0.612805***	1.26633	-0.337964	-0.013788
Distance Interquartile	0.459361	1.061110**	-1.322250***	0.860236***	-1.41625	0.405273
Off-farm Revenues	0.003033*	0.000527**	-0.000156	0.000261	0.000177*	-0.000602
Revenue Volatility	-0.000897	-0.001353*	-0.001043**	-0.000124	-0.000016*	-0.000055
Log-Likelihood	-5399.212					

Table S2. Parameter estimates for three Nested Logit model specifications. Significance codes: (***) 0.001; (**) 0.01; (*) 0.05

	Nested Logit with all variables	**Nested Logit without controls for risk**	**Nested Logit without controls for field operation costs**
Upper Nest (reference category: Group 1, maize and "others")			
Labor Costs			
Group 2	-0.003046 ***	--0.000511**	-0.005526***
Group 3	0.026649 *	0.023846**	0.014565
Fertilizer Price			
Group 2	0.004933	0.005097	0.001808***
Group 3	0.002437 ***	0.001825***	-0.001846
Farm Size			
Group 2	-0.179874 ***	-0.167964***	-0.165529***
Group 3	-0.050802 **	-0.056015**	-0.056126*
Number of Members in Household			
Group 2	-0.009271***	-0.005423	-0.012596
Group 3	0.019445	0.018328	0.013285
Female Head HH			
Group 2	-0.217669 **	-0.208448**	-0.207268
Group 3	0.198515	0.203970	0.181667
Value of Assets			
Group 2	-0.001231 ***	-0.001046***	-0.001452***
Group 3	0.000053	0.000043	0.000048
Livestock			
Group 2	0.001269	0.002237**	-0.000695
Group 3	0.002765	0.002743	0.001709

Off-farm revenues			
Group 2	0.238620	0.207553	0.259905
Group 3	-0.584476*	-0.479717	-0.379683
Lower Nest (reference category: "others")			
Rain median			
Maize	0.000390 ***	0.000613***	-
Millet	0.0390338 **	0.041679	-
Cassava	0.041684 ***	0.044534	-
Sorghum	0.036790 **	0.039772	-
Groundnuts	0.008831 **	0.001422***	-
Beans	0.008086 ***	0.000599*	-
Rain Interquartile Spread			
Maize	0.000824 *	0.000639**	-
Millet	0.117854 ***	0.062892	-
Cassava	0.123993	0.069416	-
Sorghum	0.115646 *	0.061315*	-
Groundnuts	0.015429	0.000975	-
Beans	0.021475 **	0.002739*	-
Temp. median			
Maize	-0.023216 ***	-0.021862***	-
Millet	-1.354110	-0.777925	-
Cassava	-1.364410	-0.780989	-
Sorghum	-1.402100 *	-0.817151	-
Groundnuts	-0.245897 *	-0.024217**	-
Beans	-0.273013 ***	-0.031752***	-
Temp. Interquartile Spread			
Maize	-0.025320 ***	-0.016085***	-
Millet	-1.464670	-0.176709	-
Cassava	-1.481650	-0.210474	-
Sorghum	-1.489680	-0.212163	-
Groundnuts	-0.304083	-0.014417	-
Beans	-0.323074 *	-0.001058	-
Distance median			
Maize	0.127354	0.110921*	0.149458
Millet	20.582800 *	9.537920	6.297250
Cassava	20.438600	10.061900	6.202450**
Sorghum	20.861300 *	10.510700	6.966810
Groundnuts	0.664508	0.088618	0.802590
Beans	0.776178*	0.143671	1.332120
Distance Interquartile			
Maize	0.924546 ***	0.994347**	0.997035*
Millet	82.253300	52.893400	37.815100
Cassava	81.791900	53.669800	39.247900
Sorghum	82.873200	54.688200	40.618100
Groundnuts	11.554300 **	1.304490**	5.616900*
Beans	11.830800 **	2.757320**	8.909470
Revenue			
Maize	0.008682	0.006015***	-0.000052
Millet	0.001045 **	0.002009*	0.009060**
Cassava	0.000246	-0.003871	0.026368**
Sorghum	0.008123 ***	0.008347	0.004693
Groundnuts	0.007077 ***	0.001422**	0.001114
Beans	0.000731 *	0.005911**	0.004171
Revenue Volatility			

Maize	-0.006293	-	0.000238
Millet	-0.186662 ***	-	-0.020953**
Cassava	-0.002826	-	-0.005620***
Sorghum	-0.007016 **	-	-0.006186
Groundnuts	-0.004296 ***	-	-0.000240*
Beans	-0.008643 *	-	-0.000082**
Inclusive Value Parameters			
Group 1	13.1013***	15.0177**	81.9411**
Group 2	0.06351	0.1178	1.2095**
Group 3	0.41604**	4.464780**	7.1810**
Log-Likelihood	- 5,398.466	5,413.153	-5,469.799