MOHAMMED VI POLYTECHNIC UNIVERSITY | FACULTÉ DE GOUVERNANCE, SCIENCES ECONOMIQUES ET SOCIALES

A Crossed analysis of participations in labor and grain markets: Evidence from Malawi

A. Camara
L. Savard

CAHIER DE RECHERCHE
21-03

A Crossed analysis of participations in labor and grain markets: Evidence from Malawi

Alhassane Camara[1] and Luc Savard[2]

Abstract

This study contributes to the literature on the identification of factors shaping the decision to participate in the grain market in Africa. Unlike previous studies, we introduce labor market participation into the farm household model to highlight the heterogeneity in decision making. Empirically, we rely on an extension of Heckman's approach and introduce control functions to mitigate endogeneity problems related to the adoption of agricultural technologies. The results include that limited access to transportation infrastructure, by discouraging the supply of grain, constrains the household to experience an excess supply of labor; the use of agricultural technologies such as inorganic fertilizers encourages the production of marketable surpluses and the employment of external agricultural labor. However, in Malawi, price incentives may have a reverse effect on the choice of market regimes, even though the effect on production may be positive for households that are already participating. This study thus provides a better targeting when designing policies aimed at increasing the integration of smallholder farmers into both product and labor markets.

Code JEL: Q120, J430, Q130

Keywords : Farm Households, Agricultural Labor Markets, Agricultural Markets and Marketing.

[1] Corresponding author : École de Gestion, Université de Sherbrooke, Québec, Canada email alhassane.camara@usherbrooke.ca.
[2] FGSES, Université Mohammed VI Polytechnique, Rabat, Morocco.

I. Introduction

Agriculture is the main source of employment in developing countries and, as such, it plays a key role in the development strategy to reduce poverty and improve food security. The transformation of the agricultural sector is viewed therefore as an overarching requirement for achieving structural change in these economies.

Agricultural Commercialization appears to be an effective engine for achieving this transformation, as it improves farmers' incomes, allowing them to access agricultural technologies[3]. Yet agricultural marketing is not only about cash crops[4], as it is common in most low-income countries for smallholders to shift first to the marketing of food crops. In Africa for example, Carletto, Corral & Guelfi (2017), show that among farm households that sell part of their production, the majority produce[5] only food crops (79% in Uganda, 91% in Tanzania) even if sales are limited.

The early theoretical formulations of smallholder market participation decisions date back to the beginning of the 1980s through work by Strauss (1986) and Lopez, (1984)[6]. In these so-called farm household models, production and consumption decisions can be tied together under certain conditions, thereby jeopardizing the creation of marketable surpluses.

---

[3] See Poole (2017) for a review

[4] von Braun (1995) has already pointed out that commercialization can take several forms; it can take place upstream (product market) or downstream (input market). It concerns both cash crops and food crops.

[5] The market participation rates are 80%(Uganda), 68% (Tanzania), 90% (Malawi). In calculating the participation rates, the authors consider all products sold on the market (food and/or cash crops).

[6] Cited in Sadoulet and Janvry (1995)

Empirically, the pioneering work by Goetz (1992) uses the farm household model as a theoretical framework and focuses on the role of transaction costs and institutional factors when deciding (not) to participate in the market. He shows that reducing transportation costs, market information costs and improving access to technologies and productive assets improve the marketing of cereals in Senegal. Following the work by Goetz (1992), Key, Sadoulet, & Janvry, (2000) further develop and estimate a structural model to disentangle the effects of the proportional and fixed transaction costs when analyzing the supply response to price changes. Using data from Mexico in the corn sector, they find that proportional transaction costs are significant for net sellers, while fixed transaction costs are as relevant for net sellers as they are for net buyers.

Alene & al. (2008) point out that the previous studies implicitly overlook input market access constraints when analyzing output market participation. They argue that participation in the inorganic fertilizer market is as important as participation in the output market, since these factors are key to the production of the marketable surplus. Therefore, they examine the drivers of extensive and intensive margins in both the output and input markets. They find that the price of corn in Kenya has no influence on market entry, even though it is a determining factor of sales. On the other hand, price and institutional factors are crucial in the adoption and use of inorganic fertilizers.

Ouma & al.(2010) use an extended Heckman model to analyze extensive and intensive margins in the banana market both in Rwanda and Burundi. They find that transaction costs (geographic location, market information costs, travel time) are factors influencing participation. In addition, non-price factors such as land ownership, availability of labor, off-farm income, are also critical in the analysis of market participation.

In the same spirit as Alene & al.(2008), Mather, Boughton & Jayne (2013) assert that most prior studies have better justified the implementation of market access policies i.e reduction of transaction costs, price incentives, that ignore the constraints of productivity and access to land. They show empirically that in the countries under consideration -Mozambique, Zambia and Kenya-, market access is cheaper, and that emphasis should be placed on productivity policies to complement government efforts in reducing market access costs. In the same vein, Olwande & al.(2015) argues that increasing productivity is critical to scaling up marketing of maize, milk and cabbage in Kenya.

Burke, Myers, & Jayne, (2015) present a Triple Hurdle Model to analyze dairy market participation regimes in Kenya. Their contributions lie in the fact that they introduce the decision to become a producer before the decision to participate or not in the market. They find that households become net producers and net suppliers of milk when the chain of distribution is operational.

Handschuch & Wollni (2016) focus on the role of gender and membership in producer organizations when analyzing participation in the millet market in Kenya. They find that membership in producer organizations positively influences market participation, and this is particularly important for women who receive better prices as a result of their membership in these organizations.

Apart from Alene & al. (2008)[7], and Mather, Boughton, & Jayne (2013), most authors consider production to be a given. Therefore, the supply response to market incentives assumes the existence of enough production, which in turn is highly dependent on labor, especially in a setting where mechanization is quite limited. Yet, the household's labor supply is the result of a trade-off[8] that may (not) threaten farm production. Thus, including labor market in the analysis of agricultural commercialization allows for heterogeneity among households in their responses to the public policies.

The objective of this paper is to identify the factors that determine participation in the cereal market while considering the interaction with the rural labor market. The potential contributions of paper are as follows:

- we present an extension of the theoretical model by Key, Sadoulet, & Janvry, (2000) by adding the labor market[9]. This allows to gain heterogeneity among farmers when analyzing supply response to price changes.
- Empirically, we use an extension of Heckman's method to estimate the reduced form of the theoretical model. The extension consists in using the multivariate probit model at the first step. Indeed, the solution of the theoretical model has led to the identification of 9 market

[7]. The authors independently investigate the factors determining participation in the inorganic fertilizer and maize markets in Kenya.

[8] Trade-off between working on the family farm or off the farm. So far, Adjognon, Liverpool-Tasie & Reardon (2017).show that participation in the rural labor market by the farm household can be a source of funding for the purchase of agricultural inputs to boost productivity.

[9] Due to the limited sample size, Key, Sadoulet, & Janvry, (2000) were not be able to include labor market participation.

regimes and therefore the bivariate probit[10] is no longer adequate for the estimation. In addition, this extension includes control functions[11] to mitigate the endogeneity issues associated with technology adoption at the farm level. We check the robustness of results from the extension of Heckman model by using a nested logit which does not rely on simulated maximum likelihood estimation.

We are focusing on Malawi for our empirical application; it is a small country in South East Africa where agriculture accounts for about 30% of GDP, generates 64.1% of employment, and accounts for 80% of export earnings [12]. Malawi has a dual agriculture, with smallholder farms, which use most of the arable land[13], and estate plantations. The former are households that are mostly engaged in the production of food crops (maize in general). The latter, which account for 12% of agricultural land, are more specialized in cash crops (tobacco, cotton, tea and sugar) and are the main component of the country's exports.(Matchaya, 2012).

Agriculture is essentially rain-fed and the country is among the poorest in the world. Undernourishment and food insecurity are persistent and are exacerbated by the recurrence of climatic shocks (drought). Accordingly, the country's development program encourages measures to increase the productivity of the agricultural sector, while allowing for a greater market orientation (Aberman & Roopnaraine 2020).

[10] Ouma & al.(2010) used the bivariate probit
[11] See Wooldridge (2015)
[12] Ministry of Agriculture, Irrigation and Water Development (2016) source
[13] More than 72% of agricultural land is used by small farms. Matchaya (2012)

This paper is structured as follows: in section 1, we introduce the theoretical model, in section 2 we describe the empirical specification, data and results are presented in section 3 and finally, we provide concluding remarks in the last section.

II. Theoretical framework

The model includes the labor market into the analytical framework developed by Key, Sadoulet, & Janvry, (2000); it ignores intra-household decisions and risks. The decision maker (household) maximizes utility $u(c_v, c_m, c_l, z)$[14] from the consumption of produced crop (c_v), non-agricultural goods (c_m) purchased at the price p_m, and leisure (c_l). z stands for socio-economic characteristics of the household. $g\big(q_v\ \ x_f, \ell_d, s_o\ \Delta\big)$[15] is the production technology, q_v is the produced quantity of crop, x_f is inorganic fertilizers purchased[16] at market price p_f; s_o is the quantity of seeds from past production. ℓ_d is the labor employed in crop production; it is composed entirely or partly of family labor as the household can enter the market to supply the surplus $(m_l > 0)$ or to buy some when the family labor force is lacking $(m_l < 0)$. Also, the household can enter the output market to sell surplus $(m_v > 0)$ or to buy some to cover the household demand $(m_v < 0)$. Prices in the crop and labor markets are p_v and w respectively. When intervening in the crop market (v), households incur fixed (f) and proportional (p) transaction costs (t) as sellers (s) or buyers (b). : $t_f^{v,i=\{s,b\}}(\Omega)$, $t_p^{v,i=\{s,b\}}(\Omega)$. Conversely in the labor market (l), these costs are represented by:

[14]Monotonically increasing, strictly concave function with continuous partial derivatives.
[15] Concave and twice differentiable.
16 The household also uses organic fertilizers, to simplify, here we include them in Δ.

$t_f^{l,i=\{s,b\}}(\Omega)$ et $t_p^{l,i=\{s,b\}}(\Omega)$. Fixed transaction costs, which are independent of traded quantities, include, *inter alia*, information gathering costs, negotiation costs and monitoring costs. Information costs may include time spent searching for better prices (or compensation), time spent looking for potential buyers or sellers. Negotiation costs may include the time required to negotiate a contract, reach an agreement and set payment terms; the extent to which individuals (households) minimize these costs depends on their characteristics (education, skills, gender), the membership in network (associations, cooperatives), the quality of the goods/services. Monitoring costs may include the time spent or expenses incurred in enforcing contract clauses. Proportional costs are dependent on the quantities traded on the market, and include *inter alia*, transportation costs per unit of traded goods[17].

While transportation costs are observable, most components of fixed transaction costs are not directly observable in surveys. However, it is possible to identify the effects of the unobservable components of these costs[18] on household decisions using observables (Ω)[19]. As proportional transaction costs increase (decrease) market prices for buyers (sellers), fixed transaction costs are only incurred at market entry. Finally, when the household decides not to intervene in the food (labor) market, the rationale behind its behavior depends on shadow price $\tilde{p}_v$ $(\tilde{w})$.

The problem of representative household can be written as follows:

[17] Vakis, Sadoulet & Janvry (2003).
[18] Information gathering costs, negotiation costs.
[19] For example, distance to the nearest market, access to transportation infrastructure, access to transportation facilities, access to telecommunications services and equipment, level of education, experience, size of the rural market.

$$Max_{\boldsymbol{\Lambda}} = u(c_v, c_m, c_l, z);\ \boldsymbol{\Lambda} = \{c_v, c_m, c_l, q_v\ \ x_f, s_o, m_v, m_l\}$$

$$g(q_v\ \ x_f, \ell_d, s_o\ \Delta) = 0 \qquad (1)$$

$$E_v + q_v - m_v - c_v - \ s_o = 0 \qquad (2)$$

$$E_l - c_l - \ell_o = 0 \qquad (3)$$

$$\ell_o - \ell_o^a - m_l\delta_l^s = 0 \qquad (4)$$

$$\ell_d + m_l\delta_l^b - \ell_o^a = 0 \qquad (5)$$

$$\{(p_v - t_p^{v,s})\delta_v^s + (p_v + t_p^{v,b})\delta_v^b\}m_v + \{(w - t_p^{l,s})\delta_l^s + (w + t_p^{l,b})\delta_l^b\}m_l + T \ldots$$

$$= p_m c_m + p_f x_f + t_f^{v,s}\delta_v^s + t_f^{v,b}\delta_v^b + t_f^{l,s}\delta_l^s + t_f^{l,b}\delta_l^b \qquad (6)$$

$$if\ m_v > 0\ then\ \ \delta_v^s = 1 \text{ and } \delta_v^{\mathrm{b}} = 0;\ if\ m_v < 0\ \ then\ \ \delta_v^{\mathrm{s}} = 0 \text{ and } \delta_v^b = 1;$$

$$if\ \ m_l > 0\ then\ \delta_l^s = 1 \text{ and } \delta_l^b = 0;\ if\ \ m_l < 0\ \ then\ \ \delta_l^{\mathrm{b}} = 1 \text{ and } \delta_l^{\mathrm{s}} = 0;$$

E_v is the initial stock of food crops, E_l is the labor endowment, ℓ_o is the total labor supply, T stands for transfers, other income, and credits.

Equation (1) reflects the constraint related to the production technology. Equation (2) shows that the sum of production, endowments, and purchases equals the sum of consumptions, and sales. Equation (3) indicates that household allocates the labor endowment by using a portion of it for leisure and a portion for production. Then, total labor supply (equation 4) would be in turn split into labor employed in household farm activities (ℓ_o^a) and labor supplied in off-farm activities if household presents a surplus of family labor ($m_l > 0$ and $\delta_l^s = 1$). It should be reminded here that we implicitly assume that family labor and hired labor are perfect substitutes[20].

Equation (5) implies that farm labor demand would be met by family labor and hired labor if family labor is lacking; equation (6) indicates that household expenditures equals household income[21]. Expenditures include net food purchases, hired labor expenses, purchases of non-agricultural goods, and any fixed costs associated with participations in crop and labor markets. Resources include net sales of food products, off-farm income (including labor income), and other income, including cash crop sales.

Substituting successively, equations 3 to 5 can be expressed as one:

$$\ell_d + m_l(\delta_l^b + \delta_l^s) = E_l - c_l \qquad (7)$$

This equation implies that the household's labor supply first meets labor demand in household's farm activity, and then if the labor surplus exists[22], it is supplied outside of household's farm activity. On the other hand, if family labor supply is not enough, then the household purchases labor to meet the gap.

[20] Recently, Amare & Shiferaw (2017) support the substitution between farm and non-farm activities by the farm household. However, there is no consensus in the literature on this trade-off. Benjamin (1992); Pitt and Rosenzweig (1986) do not reject the substitution hypothesis while other authors (Eswaran and Kotwal, 1986; Fafchamps and Quisumbing, 1999; Frisvold, 1994; Jacoby, 1993) reject it. Cited in Ahituv and Kimhi (2002)

[21]In general, labor supply is made at the individual level; however, here we analyze the influence of household specific factor endowments and transaction costs on the choice of market regimes. Consequently, the focus is rather on the household as a unit of decision that aggregates the individual decisions of its members.. Sadoulet, De Janvry, & Benjamin (1998) also study the choice of labor market participation regime of farm households in Mexico, implicitly considering the household as a single decision-making entity.

[22]. Let's recall that the surplus labor offered by the household is used in non-agricultural activities if opportunities permit, otherwise it remains inactive outside the farm.

Let the Lagrangian associated with the above problem be:

$$\mathcal{L} = u(c_v, c_m, c_l, z) + \lambda g\big(q_v \;\; x_f, \ell_d, s_o\,\Delta\big) \ldots$$

$$+\zeta(E_v + q_v - m_v - c_v - \; s_o) \;\; \ldots.$$

$$+\,\kappa\big(E_l - c_l \; - \ell_d - m_l(\delta_l^b + \; \delta_l^s)\,\big) \;\ldots$$

$$+\,\theta\big[\{\big(p_v - t_p^{v,s}\,\big)\delta_v^s + \big(p_v + t_p^{v,b}\,\big)\delta_v^b\}m_v + \{\big(w - t_p^{l,s}\,\big)\delta_l^s + \big(w +$$

$$t_p^{l,b}\,\big)\delta_l^b\}m_l + T - p_m c_m - -p_f x_f - t_f^{v,s}\delta_v^s - \delta_v^b - t_f^{l,s}\delta_l^s - t_f^{l,b}\delta_l^b\big]$$

The first-order conditions can be written as follows:

$$u_{c_v} - \zeta = 0 \qquad (c1)$$

$$u_{c_m} - \theta = 0 \qquad (c2)$$

$$u_{c_l} - \kappa = 0 \qquad (c3)$$

$$\lambda g_{s_o} - \zeta = 0 \qquad (c4)$$

$$\lambda g_{x_f} - \theta p_f = 0 \qquad (c5)$$

$$\lambda g_{l_d} - \kappa = 0 \qquad (c6)$$

$$\theta\{(p_v - t_p^{v,s})\delta_v^s + (p_v + t_p^{v,b})\delta_v^b\} - \zeta = 0 \quad si \;\; m_v \neq 0 \quad (c7)$$

$$\tilde{p}_v = \frac{\zeta}{\theta} \;\; si \; m_v = 0 \qquad (c8)$$

$$\theta\{(w - t_p^{l,s})\delta_l^s + (w + t_p^{l,b})\delta_l^b\} - \kappa(\delta_l^b + \delta_l^s) = 0 \;\; si \; m_l \neq 0 \; (c9)$$

$$\widetilde{w} = \frac{\kappa}{\theta} \; si \;\; m_l = 0 \; (c10)$$

Maximizing with respect to the quantities traded on the markets (m_v, m_l) creates a discontinuity. Key, Sadoulet & Janvry, (2000) propose a two-step procedure to solve this problem: (i) solve the set of equations conditional on the participation regime; (ii) choose the alternative that gives the highest utility. By proceeding in this way, we can theoretically identify 9 participation schemes.

Unlike Key, Sadoulet & Janvry, (2000), we define a system of endogenous prices on the two markets ($\hat{p}_a$, $\widehat{w}$) characterizing the 9 participation schemes.

$$P^j = \left\{ \begin{array}{ll} \left(p_v - t_p^{v,s} ; w - t_p^{l,s}; \Phi\right) & \text{Sellers in both markets} \\ \left(p_v - t_p^{v,s} ; w + t_p^{l,b} ; \Phi\right) & \text{Sellers of outputs and buyers of labor} \\ \left(p_v - t_p^{v,s} ; \widetilde{w}; ; \Phi\right) & \text{Sellers of outputs and autarky in labor market} \\ & \\ \left(p_v + t_p^{v,b} ; w - t_p^{l,s} ; \Phi\right) & \text{Buyers of outputs and sellers of labor} \\ \left(p_v + t_p^{v,b} ; w + t_p^{l,b} ; \Phi\right) & \text{Buyers in both markets} \\ \left(p_v + t_p^{v,b} ; \widetilde{w}; \Phi\right) & \text{Buyers of outputs and autarky in labor market} \\ & \\ \left(\tilde{p}_v; w - t_p^{l,s} ; \Phi\right) & \text{Autarky in output market and seller of labor} \\ \left(\tilde{p}_v; w + t_p^{l,b} ; \Phi\right) & \text{Autarky in output market and buyer of labor} \\ (\tilde{p}_v; \widetilde{w}; \Phi) & \text{Autarky in both markets} \end{array} \right.$$

$$\Phi = \{p_f;\ p_m\}$$

Using above price system, we can define the indirect utility function that allows to assess the level of utility associated with each regime j

$$V^j = V\left(P^j, Y^j, z\right) \qquad (8)$$

Thus, we see that changes in the policy environment by altering prices and household incomes affect the utility levels; which explains the potential transitions between regimes after a policy shock. Conditional to the choice of participation scheme, we determine the supply functions $q_v^j(P^j, \Delta)$ to analyze intensive margins.

III. Empirical framework

The theoretical framework outlined in the above section demonstrates that the extensive margin depends on both fixed and variable transaction costs, whereas the intensive margin only depends on variable transaction costs.

An unbiased estimate of these margins is obtained using selection models - notably Heckman's approach[23]- - which is widely used in this literature. (Goetz ,1992, Alene & al., 2008, Ouma & al., 2010). However, it is important to note that Heckman's standard model is based on the univariate probit used in the selection equation.

Ouma & al. (2010) extended this approach by suggesting a bivariate probit in the estimation of the selection equation to consider the potential correlations between the alternatives. In the same vein, we build our estimates on an extension of Heckman's two-step approach. Unlike Ouma & al. (2010), the selection equation is estimated using a multivariate probit as we have more than two regimes[24].

1. Extensive Margins: A Multivariate Probit (MVP)

[23] The literature also indicates the use of hurdle (double or triple) model.

[24] Mussema, Kassa, Alemu, & Rashid, (2013) employ a similar approach but they use a multivariate Tobit to describe intensives margins.

The MVP model is derived from Cappellari and Jenkins (2003). Let us consider M equations describing alternatives (regimes) to be chosen:

$$y_{nj}^{*} = \beta_j X_{nj} + \varepsilon_{nj} \quad ; \quad j = 1, \ldots M \qquad (8)$$

$$y_{nj} = 1 \quad si \ \ y_{nj}^{*} > 0 \ ; \quad 0 \quad otherwise \qquad (9)$$

$\varepsilon_{nj} \sim MVN(0, \Omega)$ are error terms distributed as multivariate normal, each having a mean of zero, and the variance-covariance matrix Ω, where Ω has values of 1 on the main diagonal, and correlations $\rho_{jk} = \rho_{kj}$. In fact, the correlation between decisions to participate as a buyer or seller - for example in the grain market - can be explained by the fact that both decisions depend on the same factors underlying transaction costs. In addition, covariant shocks that affect this market can influence the decisions of both buyers and sellers. (Ouma & al., 2010)

The log-likelihood function is given by :

$$L = \sum_{n=1}^{N} w_n log\varphi(\mu_n; \ \Omega) \qquad (10)$$

w_n is an optional weight for observation n, φ is the multivariate standard normal distribution with arguments μ_n , Ω.

Due to the multiplicity of integrals in equation 10, the likelihood functional form is not analytically defined. Therefore, simulation methods are used to evaluate the likelihood and select the parameters that maximize the simulated likelihood (Cappellari et Jenkins, 2003).

2. Intensive margins

In following the work in Key, Sadoulet & Janvry (2000), we present a linear approximation of the production function.

$$q_v^j\left(P^j,\Delta\right) = P^j\beta + \Delta\gamma + \sigma^j\lambda^j$$

Where P^j is the price system for each alternative; this system is defined above; Δ stands for characteristics of the production system; β, γ and σ^j are parameters. λ^j is the inverse of the Mills ratio for alternative j.

$$\lambda^j{}_n = \frac{\phi(\beta_m X_{nj})}{\Phi(\beta_m X_{nj})} \text{ if } d^j = 1\,;$$

$$\lambda^j{}_n = \frac{-\phi(\beta_m X_{nj})}{\left(1-\Phi(\beta_m X_{nj})\right)} \text{ if } d^j = 0\,;$$

d^j is a binary variable indicating 1 if the alternative j is chosen.

For the alternative 1, according to the price system, the supply function can be written as:

$$q_v^1 = p_v\beta_p + w\beta_w + \Phi\beta_\varphi + M\beta_M^1 + \Delta\beta_\Delta + \sigma^1\lambda^1;\ \ \beta_M^1 = \alpha_b^l + \alpha_s^v$$

$$-t_p^{v,s} = \alpha_s^v M;\ -t_p^{l,s} = \alpha_s^l M$$

Thus, when β_M^1 is positive then the variable M (proxy of transaction costs such as equipment ownership) reduces variable transaction costs and improves the effective prices received.

By proceeding in the same way for the other alternative, we obtain supply functions:

$$q_v^2 = p_v\beta_p + w\beta_w + \Phi\beta_\varphi + M\,\beta_M^2 + \Delta\beta_\Delta + \sigma^2\lambda^2;\ \ \ \beta_M^2 = \alpha_b^l + \alpha_s^v$$

$$-t_p^{v,s} = \alpha_s^v M;\ \ t_p^{l,b} = \alpha_b^l M$$

$$q_v^3 = p_v\beta_p + \Phi\beta_\varphi + M\,\beta_M^3 + \Delta\beta_\Delta^a + z\beta_z^a + \sigma^3\lambda^3;\ \ \ \beta_M^3 = \alpha_s^v;$$

$$-t_p^{v,s} = \alpha_s^v M;$$

$$q_v^4 = p_v\beta_p + w\beta_w + \Phi\beta_\varphi + M\beta_M^4 + \Delta\beta_\Delta + \sigma^4\lambda^4; \;\; \beta_M^4 = \alpha_s^l + \alpha_b^v$$

$$t_p^{v,s} = \alpha_b^v M; \;\; -t_p^{l,s} = \alpha_s^l M$$

$$q_v^5 = p_v\beta_p + w\beta_w + \Phi\beta_\varphi + M\beta_M^5 + \Delta\beta_\Delta + \sigma^5\lambda^5; \;\; \beta_M^5 = \alpha_b^l + \alpha_b^v$$

$$t_p^{v,s} = \alpha_b^v M; \;\; t_p^{l,s} = \alpha_b^l M$$

$$q_v^6 = p_v\beta_p + \Phi\beta_\varphi + M\,\beta_M^6 + \Delta\beta_\Delta^a + z\beta_z^a + \sigma^6\lambda^6; \;\;\; \beta_M^6 = \alpha_b^v;$$

$$t_p^{v,s} = \alpha_b^v M;$$

$$q_v^7 = w\beta_w + \Phi\beta_\varphi + M\beta_M^7 + \Delta\beta_\Delta^a + z\beta_z^a + \sigma^7\lambda^7; \;\; \beta_M^7 = \alpha_s^l$$

$$-\,t_p^{l,s} = \alpha_s^l M$$

$$q_v^8 = w\beta_w + \Phi\beta_\varphi + M\beta_M^8 + \Delta\beta_\Delta^a + z\beta_z^a + \sigma^8\lambda^8; \;\; \beta_M^8 = \alpha_b^l$$

$$t_p^{l,s} = \alpha_b^l M$$

$$q_v^9 = \Phi\beta_\varphi + \Delta\beta_\Delta^9 + z\beta_z^9 + \sigma^9\lambda^9;$$

By adding the error terms ε^j along with different production functions, we have the following system:

$$q_v^1 = p_v\beta_p + w\beta_w + \Phi\beta_\varphi + M\beta_M^1 + \Delta\beta_\Delta + \sigma^1\lambda^1 + \varepsilon^1 \qquad (13)$$

$$q_v^2 = p_v\beta_p + w\beta_w + \Phi\beta_\varphi + M\,\beta_M^2 + \Delta\beta_\Delta + \sigma^2\lambda^2 + \varepsilon^2 \qquad (14)$$

$$q_v^3 = p_v\beta_p + \Phi\beta_\varphi + M\,\beta_M^3 + \Delta\beta_\Delta^a + z\beta_z^a + \sigma^3\lambda^3 + \varepsilon^3 \qquad (15)$$

$$q_v^4 = p_v\beta_p + w\beta_w + \Phi\beta_\varphi + M\beta_M^4 + \Delta\beta_\Delta + \sigma^4\lambda^4 + \varepsilon^4 \qquad (16)$$

$$q_v^5 = p_v\beta_p + w\beta_w + \Phi\beta_\varphi + M\beta_M^5 + \Delta\beta_\Delta + \sigma^5\lambda^5 + \varepsilon^5 \qquad (17)$$

$$q_v^6 = p_v\beta_p + \Phi\beta_\varphi + M\,\beta_M^6 + \Delta\beta_\Delta^a + z\beta_z^a + \sigma^6\lambda^6 + \varepsilon^6 \qquad (18)$$

$$q_v^7 = w\beta_w + \Phi\beta_\varphi + M\beta_M^7 + \Delta\beta_\Delta^a + z\beta_z^a + \sigma^7\lambda^7 + \varepsilon^7 \qquad (19)$$

$$q_v^8 = w\beta_w + \Phi\beta_\varphi + M\beta_M^8 + \Delta\beta_\Delta^a + z\beta_z^a + \sigma^8\lambda^8 + \varepsilon^8 \qquad (20)$$

$$q_v^9 = \Phi\beta_\varphi + \Delta\beta_\Delta^9 + z\beta_z^9 + \sigma^9\lambda^9 + \varepsilon^9 \qquad (21)$$

The standard deviations of the coefficients are adjusted to account for the heteroskedasticity of errors terms due to the presence of inverse of the Mills ratios in the production equations.

3. Variables

In the selection equation, the dependent variables are binary, which take the value 1 if regime j is chosen by the household and 0 otherwise. According to the theoretical model, we thus have 9 binary variables. In the supply equations, the logarithm of the cereal quantity produced by the household is used as the dependent variable. So far, literature suggests that explanatory variables can be clustered into three categories (Mather, Boughton & Jayne, 2013): agro-ecological characteristics of the household's living area, market access, use of agricultural technology, prices and wages, socio-economic characteristics of the household and variables determining the production system.

3.1 Agroecological characteristics

In our estimation, we include dichotomous variables characterizing the agroecological zones in which the households live. In addition, we also use the frequency of droughts and floods during the agricultural season in order to control for differences in terms of climatic risks.

3.2 Market access

The distance to the nearest passable road is used as a proxy for the transportation costs. We also use the ownership of transportation equipment (bicycle, motorcycle, vehicle) because these assets allow the household to reduce search costs. However, it is likely that the ownership of such equipment may be endogenous at the individual (household) level as it may be the outcome of participation. The literature shows that this problem is largely ignored by most authors, usually due to the lack of enough and valid instruments. (Handschuch & Wollni, 2016; Mather et al.,2013; Ouma, & al., 2010; Alene et al. , 2008 ; Goetz, 1992). In this paper, to minimize this problem, we use community-level controls that can influence this variable, such as the lack of a public transportation station in the community/village and the proportion of households in the community that own a mode of transport . The idea is that these variables are more dependent on the characteristics of the community/village, which are beyond the control of individual households[25]. We also use the number of producer organizations and cooperatives at the village (or community) level as controls because they can enable households to access markets with lower research costs (networking). In the literature, extension services are highlighted as a channel for agricultural technology adoption. Therefore, we use the proportion of households at the village (community) level that have received and applied, during the agricultural season, advice from extension services on marketing strategies and seed use.

3.3 Use of agricultural technology

We include crop yields at the village or community level to reflect the combined effect of past investment and technology adoption decisions[26]. We also use the proportion of land under inorganic fertilizers. This variable can also be suspected of endogeneity, in that the farm household may use the income from the participation to purchase fertilizer in the market.

However, in the case of Malawi and in the sample in use, the inorganic fertilizers used are primarily from subsidies[27]. The proportion of purchases of inorganic fertilizers on the market is therefore relatively low; to this effect, the literature using cross-sectional data shows a crowding out effect of fertilizers acquired on the market by subsidized fertilizers (Takeshima & Nkonya , 2014 and see Ricker-Gilbert & Jayne, 2017). Since 2008, the allocation of subsidized fertilizer vouchers has

[25] Heltberg & Tarp (2002) use the same strategy to address the problem of endogeneity of agricultural technology adoption at the individual level. Also, Mather et al. (2013) determine the village level price of inorganic fertilizers, as an instrument of household fertilizer use.

[26] Heltberg & Tarp (2002) use the same strategy to address the problem of endogeneity of agricultural technology adoption at the individual level.

[27] The subsidy is equivalent to 2/3 or more of the market price, (see Arndt, Pauw, & Thurlow, 2016)

been based on household vulnerability, i.e. resource-poor households[28] are targeted as a priority(Ricker-Gilbert & Jayne , 2017). Therefore, we consider variables such as the dependency ratio within the household, age, and gender of the household head to control for these differences in coupon allocation. It should be remembered that these variables are also critical in the analysis of market participation.

Finally, to control for the endogeneity of coupon allocation, we use the control function approach. For this purpose, we rely on instruments such as the presence of a member of parliament in the community/village and household social capital measured by time spent living in the village.[29]. The control function approach is a two-step approach: (i) we regress the fertilizer use variable on the instruments; (ii) we introduce the residuals from this regression into the structural equation (equation 8). We adjust the standard deviations (boostrapping) of the coefficients in the 2nd step because the residuals are not directly observed. (Petrin & Train 2002). The significance test on the coefficient associated with the residuals reveals or not the endogeneity of the use of subsidized fertilizers.

We also consider the proportion of land under mechanization/animal traction to control for differences in the use of agricultural technologies. As previously stated, to solve the endogeneity problem, we use the control function approach. To do this, we use as an instrument a household

[28] According to the following criteria: age (old age), gender (women) and physical disability of the head of household

[29] Bezu et al. (2014) employ the same instruments in the case of Malawi to control for the endogeneity of subsidized fertilizer use.

wealth index[30] measured by a combination of the characteristics of the household's dwelling (roof, floor, wall, number of rooms in the dwelling, access to electricity). Handschuch & Wollni (2016).

3.4 Prices and wages

We include grain prices for buyers (sellers). These prices are obtained by dividing purchases (sales) by volumes. They are then adjusted by a spatial and temporal price index[31] to consider the time lag between purchase and sales periods. Indeed, during the harvest periods[32]-May to July-the price of maize (the main cereal product in Malawi) is relatively low; whereas it is relatively high during the January-March period, when farm households have already depleted their cereal stocks and most often act as buyers on the market (Ricker-Gilbert & Jayne, 2017).

Wages paid to hired labor are calculated by dividing the value by the number of days worked. Off-farm wages are estimated by calculating the daily remuneration of household members who have worked off the household's farm[33] as employees (permanent or casual[34]). These wages were also deflated by the price index.

3.5 Characteristics of the household and its production system

We consider, the gender of the head of household, his education, age, the proportion of women in the household, the dependency ratio within the household, the proportion of educated people as

---

[30] The index follows an ascending order from 0 to 5.

[31] This index uses February/March 2010 as a base. By dividing prices by this index, we hope to correct for temporal inflation differences.

[32] In general, sales occur in this period.

[33] These individuals did not work on the household farm throughout the year.

[34] In Malawi, this kind of work is widespread and is commonly referred to as Ganyu.

well as the size of the household. In addition, we include dichotomous variables indicating the size of the farm[35]. We also use the proportion of households at the village (community) level that have access to credit and the use of organic fertilizers (manure) on the farm.

3.6 Data and Descriptive Analysis

Data come from the Integrated Household Survey (IHS4, 2016/2017) conducted in Malawi. The IHS4 sample is nationally representative and was randomly selected from a two-stage draw. In the first, 780 primary units (enumeration area in the census) were randomly drawn with a probability inversely proportional to their size (number of households). In the second, 16 households were randomly drawn from each enumeration area, resulting in an initial sample of 12,480 households. However, technical difficulties led to some loss of surveyed units, leaving a final sample of 12,447 households.

Farm households are the main component of this sample. More than 78% of households are engaged in rain-fed agriculture[36] (November-May); about 96% of these farmers produce only or partly cereals (mainly maize). Here, we focus on cereal producers that are composed of net sellers[37] (13%), net buyers (14%), self-sufficient (69%), and net buyers and sellers[38]. (4%). On the labor

[35] Categorical variables constructed according to the different quartiles.
[36] Main growing season in Malawi. Moreover, in terms of area sown, the rainy season is by far the most important. About 8.8% of households engage in off-season (dry season) agriculture.)
[37] Net seller if they sell part of their cereal production and do not buy cereal for household consumption (remember that consumption data (purchases) are collected over the last 7 days before the interviewer's visit); net buyer if they buy cereal and do not sell part of their production, self-sufficient if they do not buy nor sell.
[38] The existence of this alternative on the cereal market may be due to the differentiation between the cereal produced by the household and that purchased on the market, or to the seasonal variation in prices according to Key, Sadoulet, & Janvry, (2000).

market, they are composed of net sellers[39](25%), net buyers[40] (14%), self-sufficient (53%), and net buyers and sellers[41] (8%). As we can see, there is a fourth alternative in both markets: that of buyer and seller. The theoretical model outlined above does not explain such market behavior, so we isolate and classify these households into a residual group (Regime 10). The final sample is thus 8,310 farm households producing grain during the main season. The nine (9) alternatives identified in solving the representative household maximization problem are obtained using cross-matching of alternatives in the two markets.

The descriptive analysis of the sample presented in Table 1 shows that Regime 9 (40%) is by far the most prevalent, followed by Regime 7 (20%); that is, the majority (60%) of these grain-producing households do indeed not produce enough to sell on the market, and partly have a labor surplus. On the other hand, if the household decides to participate in both markets, Regime 3 is relatively more present (6%). The least present alternatives are respectively regimes 1, (2%) 4 (3%) and 5 (3%). On average, buyer prices are much higher than seller prices[42] (more than 2 times). Indeed, this reflects the presence of variable transaction costs and margins[43] that create a price band around market price. On the labor market, the farm wage is relatively lower than the non-

[39] These are households in which some of the members did not work on the family farm in the last 12 months. This set also includes inactive persons and individuals employed in non-farm activities (livestock, fishing, non-farm household businesses, casual off farm work, permanent off farm work). In addition, these households did not purchase labor to perform agricultural tasks during the rainy season. As a result, they are a net supplier of labor.
[40] These households do not provide work outside the family farm, however they increase the family labor force by hiring labor for ploughing, sowing, weeding, harvesting, during the rainy season.
[41] In the labor market; this may be due to the difference between labor used in agricultural production and labor used in off-farm activities.
[42]It should be remembered that these prices are specific to households and therefore include variable transaction costs. see Renkow, Hallstrom, & Karanja, (2004).
[43]Margins are also responsible for these price differentials.

farm wage - outside the family farm -; this gap may reflect differences in productivity between family farms and other sectors of activity existing in rural areas.

On average, it takes up to 10.5 km to reach the nearest road, while only 40% of households have at least one transportation equipment (usually bicycles). More than half of the households (69%) are male headed, on average under 44 years of age, and most often without primary education (76%). There are 4 people in the household and women are in the majority (54%); the dependency ratio[44] and the proportion of people with at least primary education are 25% and 13% respectively. Over 51% of these households live below the poverty line and 18% suffer from extreme poverty[45].

The average cultivated area is 0.5 ha, revealing that these producers are generally small-scale family farmers. More than half of the planted area uses inorganic fertilizers (56%). However, the use of agricultural equipment/animal traction remains relatively low, with 10% of cultivated land subject to mechanization. Annual production is 0.519 tons of cereals; nearly 41% of these producers produced other crops[46] in addition to cereals. There are 21% of producers that use organic fertilizers; the proportion of farmers who have received credit in a community is about 23%. In addition, 5% and 31% of households living in community reported having received and

[44] Proportion of persons less than 5 years old and more than 64 years old in the household.
[45] The dichotomous variables characterizing poverty status are taken from the IHS4 Consumption aggregate module. The national poverty line at 2016 prices is adjusted by spatial price differences.
[46] These crops include tubers, fruits and horticultural products. On the other hand, the proportion of those producing cash crops (tobacco, cotton, groundnuts) is very low. This is not surprising given that this sample is made up of small family farmers; and where more than half of these households are poor.

applied training in marketing and seed use[47], respectively. While 43% of households report having been negatively affected by a drought, only 5% report having been adversely impacted by floods.

This is probably due to the fact that more than 85% of these producers are in the warm tropical agroecological zones. On average, there is one producer organization in the communities, while the presence of cooperatives is relatively limited. Finally, about 50% of the communities in which these households live do not have a public transportation infrastructure, forcing residents to support higher transaction costs.

[47] These trainings are offered by the agricultural extension services

Table 1 : Descriptive statistics

Variables	Description	Mean	SD
prix_s	Seller's price in local currency	260.33	288.94
prix_b	Buyer's price in local currency	612.98	198.74
wag_s	Daily wage for off-farm labor in local currency	1237.57	936.39
wag_b	Daily wage for hired labor in farm activity in local currency	1033.86	737.47
	Transaction costs		
equipT	1 if the household owns at least one mode of transportation (bicycle, motorcycle, vehicle), 0 otherwise	0.40	0.49
dist_road	Distance to the nearest passable road (km)	10.53	11.23
tprc	1 if the community has a bus station, 0 otherwise	0.50	0.50
nb_ag_coop	number of cooperatives in the community	0.28	1.59
nb_farm_group	number of producer organizations in the community	1.23	2.31
	Characteristics of Households		
HH_sex	1 if male head of household, 0 otherwise	0.69	0.46
HH_ag	Age of household Head	44.31	16.30
educ_HHd	1 if the head of the household has at least primary education, 0 otherwise	0.24	0.43
sh_fem	Proportion of women in the household	0.54	0.23
dep_ratio	Dependency ratio	0.25	0.24
educ_d	Proportion of people with at least primary education	0.13	0.22
hhsiz_est	Household size	4.44	1.92
poor	1 if the household is poor, 0 otherwise	0.51	0.50
upoor	1 if the household is extremely poor, 0 otherwise	0.18	0.38
producteur_d_other	1 if the household grows crops other than cash crops, 0 otherwise	0.41	0.49
share_sup_mecan	Proportion of land under mechanization	0.10	0.32
share_sup_engrais	Proportion of land under inorganic fertilizers	0.56	0.46
share_sup_manure	Proportion of land under organic fertilizers	0.21	0.38
area_ha_ajd	Total land area (ha)	0.55	0.50
prod_kg_cereal	Annual cereal production (kg)	518.70	633.04
prod_kg_othr	Annual production of other crops (kg)	259.73	603.01
pcredic_com	Proportion of households with access to credit at the community level	0.23	0.17
p_ag_ext_Marketing	Proportion of households, having implemented the recommendations of extension services on marketing	0.05	0.09
p_ag_ext_semences	Proportion of households, having implemented the recommendations of extension services on seed use	0.31	0.20

natu_choc1	1 if the household was negatively affected by a drought during the campaign, 0 otherwise	0.43	0.50
natu_choc2	1 if the household was negatively affected by a flood during the campaign, 0 otherwise	0.05	0.21
	Regimes		
regm1	1 net seller on both markets, 0 otherwise	0.02	0.15
regm2	1 net seller on the grain market and net buyer on the labor market, 0 otherwise	0.03	0.18
regm3	1 net seller the cereal market and autarky on the labor market, 0 otherwise	0.06	0.24
regm4	1 net buyer on the grain market and net seller on the labor market, 0 otherwise	0.03	0.18
regm5	1 net buyer in both markets, 0 otherwise	0.03	0.16
regm6	1 net buyer in the grain market and autarky in the labor market, 0 otherwise	0.05	0.23
regm7	1 autarky on the grain market and net seller on the labor market, 0 otherwise	0.20	0.40
regm8	1 autarky on the grain market and net buyer on the labor market, 0 otherwise	0.06	0.24
regm9	1 autarky on both markets, 0 otherwise	0.40	0.49
regm10	1 if the household is not classified in one of the regimes defined above, 0 otherwise	0.11	0.31
	Agro-ecological zones		
AEZ1	1 if the household is in a tropical-hot/semi-arid agro-ecological zone, 0 otherwise	0.55	0.50
AEZ2	1 if the household is in an agro-ecological Tropical-Hot/Subhumid zone, 0 otherwise	0.30	0.46
AEZ3	1 if the household is in a tropical-cold/semi-arid agro-ecological zone, 0 otherwise	0.11	0.32
AEZ4	1 if the household is in an agro-ecological Tropical-cold/subhumid zone, 0 otherwise	0.03	0.18
Observation		8310	

IV. Results

We proceed in three steps: i) we first present the factors influencing the choice of market regimes (extensive margins), ii) second, we outline the determinants of the supply functions (intensive margins) of the respective regimes, iii) third, we perform a robustness check

1. Extensive margins

Using the control function approach to solve the problems of endogeneity of agricultural technology adoption (fertilizer use and mechanization) the results of the first stage regression are presented in Table 2. They show that the instruments used are significant in explaining the endogenous variables. Indeed, the social capital of the household, and the presence of a member of parliament in the village/community determine fertilizer use. In addition, the level of household wealth positively influences the use of mechanization on the farm. The associated Fisher statistic shows that these regressions are globally significant.

Table 2 : First step regression results in the control function approach

	share_sup_engrais	share_sup_mecan
Variables	b/se	b/se
nb_year_hh	0.016**	0.011***
	0.005	0.003
capso1_c	0.037**	-0.056***
	0.013	0.01
house_index	0.043***	0.011**
	0.005	0.004
equipt_c	-0.144***	0.104***
	0.025	0.021
dist_road	-0.004	0.095*
	0.044	0.041
tprc	-0.018	-0.028***
	0.011	0.007

nb_ag_coop	-0.002	0.003
	0.002	0.002
nb_farm_group	0.004	0.005**
	0.003	0.002
price	-0.004*	-0.002
	0.002	0.002
wag	-0.001*	0
	0.001	0.001
share_sup_manure	-0.017	0.004
	0.013	0.01
rdt_c	-0.010***	-0.001
	0.001	0.001
Head_homm_d	0.025*	-0.008
	0.011	0.008
educ_HHd	0.013	-0.011
	0.016	0.012
age_head	-0.001	-0.001*
	0	0
sh_fem	0.004	0.004
	0.023	0.016
dep_ratio	-0.035	0.028
	0.022	0.017
educ_d	0.076*	0.006
	0.035	0.025
pcredic_com	-0.033	0.019
	0.029	0.024
p_ag_ext_Marketing	0.037	-0.005
	0.053	0.046
p_ag_ext_seed	-0.203***	0.028
	0.031	0.021
prd_c_othr	0.385***	-0.088***
	0.018	0.016
fsiz1	-0.185***	-0.040***
	0.013	0.01
fsiz2	0.008	-0.009
	0.013	0.01
fsiz3	0.031**	-0.011
	0.011	0.01
natu_choc1	0.022	0.065***
	0.011	0.008
natu_choc2	-0.061**	0.047*

	0.022	0.019
AEZ1	-0.060**	-0.166***
	0.021	0.023
AEZ2	0.013	-0.120***
	0.02	0.023
AEZ3	-0.031	-0.098***
	0.022	0.024
hhsiz_est	-0.003	-0.003
	0.003	0.002
_cons	0.510***	0.205***
	0.039	0.031
N	8310	8310
p	0	4E-163
F	0	0

Instruments : nb_year_hh, capso1_c, house_index
nb_year_hh : log of years spent in the village
capso1_c: 1 if a member of parliament resides in the community, 0 otherwise
house_index: wealth index measured by the characteristics of the house: roof, floor, wall, access to electricity, number of occupied rooms. The index goes from 0 to 5 with 5 being the highest level.
*5% ** 1% ***0.1%

The determinants of the extensive margin are presented in Table 3; they show the determinants of the probability of membership in each of the participation regimes compared to the others. The likelihood ratio test shows that adding the residuals (control function) from the first step improves the likelihood relative to a specification that would ignore endogeneity. Moreover, these residuals (resi_eng, resi_mecan) are significant in regimes 5, 8, 9 respectively. The control function approach is then adequate. Furthermore, the test of independence of the alternatives shows that the correlation coefficients between regimes are jointly and significantly different from zero. Consequently, the estimates would be inefficient if the equations had been estimated separately; the multivariate probit model is therefore adequate to describe the membership of the alternatives.

Our results[48] (see column 9) show that market participation, i.e. the membership in at least one of the two markets, is mainly determined by productivity factors[49]; this result support the view that for products as important to household welfare as cereals, when production is low, the household would prioritize food security before marketing. In the case of Malawi, Aberman & Roopnaraine (2020) in a qualitative study highlight the importance for households to consider their food security in their decisions to sell or not sell the maize that is produced.

In column 9, results indicate that the use of inorganic fertilizers reduces the probability of falling into regime 9 -autarky in both markets-; this is mainly due to the existence of a potential productivity gain linked to the use of this input. In the same vein, when the household resides in a locality where the rate of access to credit is high, and/or the characteristics[50] of the community (village) allow for higher crop yields, the likelihood of the household exiting the market is low. The rationale behind these results lies in the fact that access to credit allows producers to cover purchases of agricultural inputs on the one hand, and on the other hand, higher crops yields allow for better productivity that may or may not free up family labor. Furthermore, when the household receives and applies advice from extension services on the use of improved seeds, the probability of remaining in autarky is relatively low, due to the resulting productivity gains.

[48] We consider only those variables that are significant at least at the 5%.

[49] The literature reviewed by Barrett (2008) and the results of work by Mather, Boughton, & Jayne, (2013) reveal that the determinants of productivity can be decisive complements in the implementation of prices and transaction cost reduction policies. These factors are also determinants of shadow prices (wages).

[50] These characteristics can be related to soil quality, past investments in the adoption of agricultural technologies, etc

The smaller the size of the farm, the more likely the household is to be in autarky[51] ; the higher the use of organic fertilizers, the greater the likelihood of preferring to stay out of the markets. This is because organic fertilizers would provide less productivity gain relative to the use of inorganic fertilizers, which is widespread in the Malawi context. Household size is also critical when it decides to move out of autarky. Indeed, this variable is a determinant of the shadow price (wage), which depends on productions and demands within the household. When household size is large, the likelihood is high that the household has a surplus of family labor, which can be employed in non-family farm activities. Producing other crops (horticulture and tubers, fruits) in addition to cereals increases the likelihood of being in autarky. In fact, the production of these crops reveals that the household's supply is relatively diversified; and the consequence is that this diversification can intensify labor use - for example, the production of horticultural crops is relatively labor intensive - thus reducing the household's surplus labor and limiting cereal production to a level that satisfies only home-consumption. Education of the head of household also plays a positive role in market participation because of its effect on reducing fixed transaction costs related to information or negotiation. When the household resides in a community/village where there is a high proportion of households with transportation facilities, it is likely that the household participates in the market as either a buyer or seller. This result is because members of these communities are relatively wealthier and therefore able to bear transaction costs.

[51] However, when the size is relatively too small, area<=0.2ha (fsiz1), it is possible that the household has an extra family labor that can be used in non-farm activities.

In columns 1 to 8, we present results on heterogeneities within the participant group. Our results show that the use of inorganic fertilizers increases the probability of generating a marketable surplus in cereal production and hiring labor (Regime 2); its positive effect on regime 5 may appear counter-intuitive insofar as the increase in productivity resulting from the use of this factor should reduce the dependence (purchase of grain) on the market. But on closer look, the story may be quite different. Indeed, the increase in production may be insufficient in the sense that the household's production capacity remains limited by the area under cultivation. As a reminder, most of this sample is composed of small producers. Consequently, it is possible that a marginal increase in inorganic fertilizer use increases the chance that a randomly selected household will be a net buyer in both markets[52].

The positive effect of the use inorganic fertilizer on production encourage the employment of an external labor. Indeed, this result can be observed through its negative effect on membership in Regime 7 while it is positive on membership in Regime 8[53].

The average yield observed in the community/village is a proxy that combines past investment decisions, agricultural technology adoption and agroecological conditions; its positive effect on the decision to participate as a net seller of cereals (Regime 2) and net buyer of labor is also explained by the associated productivity gains; its effects are similar to those observed with inorganic fertilizer.

[52] As the theoretical model indicates, households belonging to this regime can finance purchases through other sources of income, i.e., transfers, income from non-farm businesses, and income from livestock.
[53] As a reminder, in regimes 7 and 8, the household is self-sufficient on the grain market.

Regarding transaction costs, our results show that they are determinant for participation in regimes 5, 7. The longer the distance from the nearest road, the higher the transportation costs, thus discouraging the decision to sell cereals and constraining households to supply their extra labor (Regime 7).

The presence of producer organizations in the village encourages participation in Regime 1; this result may be explained by the fact that these organizations strengthen the capacity of their members, with the aim of taking advantage of market opportunities as sellers. To this end, they can be channels for the dissemination of agricultural technology and information that can improve productivity and the generation of marketable surplus of cereals and labor. In communities where there is more transportation equipment ownership; they may be relatively better provided with infrastructure/roads. This may result in lower mobility costs for goods and services and thus facilitate participation in markets (regime 5).

According to the theoretical model, an increase in the market price, *ceteris paribus*, would increase the effective price received by the net seller and decrease the price paid by the net buyer. As a result, higher prices are expected to encourage (discourage) participation in regimes where the household is a net seller (net buyer) of cereals. However, our estimates deviate from the predictions of the theoretical model; that is, higher prices are associated with a greater probability of being a net seller of cereals (Regimes 4 to 6), regardless of the status in the labor market. A plausible explanation would be that higher cereal prices could be interpreted by households as a looming shortage of staple food and therefore they would be more encouraged to store it[54]. It is important to highlight the food shortages are common in Malawi, the signs of which can be seen in cereal price movements. Indeed, two major food crises have occurred in the country, leading to a famine in 2001/2002 and a food emergency in 2005 (World Bank, 2007).

Symmetrically, higher prices lead to a lower probability of being a net seller (Regimes 1 to 3); this result also remains independent of status in the labor market. This result is explained by the high importance of cereals (especially maize) in the income of these households, which are most often poor; they may therefore have a strong preference for these staple foods, which limits any possibility of substitution between these products and others. Mather, Boughton, & Jayne, (2013) find a similar result in a comparative study between Mozambique, Zambia and Kenya. Indeed, they find that in poor areas of Mozambique, the probability of selling maize - a staple food in this country bordering Malawi - decreases in the face of an increase in expected price.

Wages have a positive influence on the probability of selling labor and buying cereals (Regime 4); however, it reduces the likelihood of being a net buyer of labor when the household is in autarky on the cereal market (Regime 8). The larger the size of the farm, the greater the likelihood that the household will have a marketable surplus of grain (Regimes 2, 3). In contrast, when the farm size is small, the likelihood of the household being a net buyer of cereals increases (Regimes 4 to 6); this effect remains independent of the position in the labor market. In analyzing membership in regimes 7 and 8, it appears that when the area under cultivation is small, the household is likely to

[54] This point is also made by Goetz (1992), who finds a similar result in the cereals market in Senegal.

have a surplus of labor. When the household is headed by a female, it is more likely that the household prefers not to sell the grain and sell the extra family labor in the market (Regime 7); therefore we conclude that for cereals that are important for household food security, women are more benevolent than men.

The dependency ratio within the household indicates its endowment of productive labor; when it is relatively large, the household is constrained both in crop production and in the supply of off-farm labor. The results for regimes 7 and 8 confirm this intuition. Indeed, the higher the dependency ratio, the lower the probability that the household has a surplus of labor (Regime 7), whereas the probability that it is a net buyer of labor is higher (Regime 8).

Furthermore, the higher the proportion of educated individuals in the household, the lower the likelihood that the household will have a surplus of labor in crop production (Regimes 4 and 7). This result is plausible since households with a high proportion of educated individuals have a large portion of the family labor employed off the farm. Sadoulet, De Janvry, & Benjamin (1998) implicitly show a similar result by arguing that skilled individuals in the farm household are more oriented towards off-farm activities.

Communities with higher access to credit are more likely to have sellers (Regime 1) or buyers (Regimes 5, 6, 8). In fact, credit can be used to purchase agricultural inputs that encourage the adoption of agricultural technologies, on the one hand, and on the other hand, it can enable the household to meet its consumption needs during the lean season.

The implementation of the marketing strategy advice discourages the household from being in autarky (Regime 7). In the same direction, training offered by agricultural extension services on seed use encourages participation as a seller (Regimes 2 and 3). Comparing Regimes 7 and 8, we observe that the use of these seeds according to the instructions received, encourages the use of hired labor. This effect remains the same when the household is a net buyer of grain (regime 5).

The production of crops other than cereals, diversifies the household's supply and can intensify the use of inputs and labor, forcing the households to use exclusively and entirely its labor endowment (comparison between Regimes 8 and 9). The recurrence of droughts, by reducing food production, reduces the likelihood that the household generate a surplus of grain (regime 7) while forcing it to adopt survival strategies such as selling surplus labor to off-farm activities. This effect is similar to the effect of flooding on the choice of Regime 9. The size of the household also reflects its labor endowment; our results indicate that the larger the size of the household, the higher (lower) the likelihood that the household is a net seller (buyer or in autarky) in the labor market regardless the status in the grain market.

Table 3 : Extensive margin estimation

	Regime1	Regime2	Regime3	Regime4	Regime5	Regime6	Regime7	Regime8	Regime9
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
equipt_c	-0.434	0.502	-0.284	0.441	0.931**	-0.089	-0.104	0.174	-0.502***
	0.350	0.297	0.248	0.288	0.308	0.236	0.182	0.236	0.132
dist_road	-0.383	0.038	-0.378	0.333	-0.775	-0.330	0.597**	-0.415	
	0.521	0.436	0.372	0.381	0.498	0.328	0.231	0.340	
tprc	0.060	-0.041	-0.018	-0.099	-0.064	-0.010	-0.005	0.117	
	0.105	0.091	0.074	0.091	0.099	0.072	0.053	0.072	
nb_ag_coop	-0.050	-0.010	-0.020	0.028	0.008	0.010	-0.011	0.012	-0.005
	0.052	0.025	0.025	0.016	0.023	0.015	0.014	0.013	0.010
nb_farm_group	0.055*	0.007	0.012	0.006	-0.041	-0.007	0.013	-0.018	0.007
	0.023	0.020	0.017	0.020	0.024	0.016	0.012	0.016	0.009
price	-0.213***	-0.029*	- 0.209***	0.050***	0.093***	0.072***			
	0.024	0.013	0.017	0.013	0.013	0.010			
wag	0.010	-0.007		0.009*	-0.011		0.005	-0.016***	
	0.006	0.006		0.004	0.006		0.003	0.004	
share_sup_mecan	-0.549	0.314	-0.230	-1.328	-1.328	-0.069	-1.282	1.237	0.917
	1.493	1.274	1.057	1.186	1.286	1.022	0.751	1.022	0.489
share_sup_engrais	-1.273	1.341*	-0.584	1.097	4.561***	0.547	-0.847*	2.248***	-2.757***
	0.788	0.643	0.535	0.654	0.693	0.516	0.401	0.508	0.289
share_sup_manure	-0.268	-0.109	0.006	0.029	0.106	0.090	-0.021	-0.071	0.102*
	0.146	0.122	0.093	0.108	0.117	0.086	0.066	0.091	0.051
rdt_c	-0.002	0.024**	-0.001	-0.007	0.038***	-0.019	-0.023**	0.020*	-0.045***
	0.011	0.008	0.009	0.013	0.009	0.013	0.008	0.009	0.004
Head_homm_d	-0.085	0.175	0.163	0.130	-0.105	0.053	-0.251***	0.032	0.070
	0.122	0.117	0.091	0.112	0.116	0.085	0.062	0.086	0.049
educ_HHd	0.135	0.053	-0.014	0.211	-0.010	-0.019	0.075	0.060	-0.149*
	0.166	0.141	0.121	0.135	0.140	0.113	0.091	0.112	0.067
age_head	-0.007	-0.004	-0.005	-0.007*	-0.006*	-0.006*	0.000	0.006***	0.002
	0.004	0.003	0.003	0.003	0.003	0.003	0.001	0.001	0.001
sh_fem	0.274	0.078	0.035	0.201	-0.165	0.012	0.035	0.093	-0.076
	0.262	0.201	0.161	0.227	0.208	0.157	0.132	0.155	0.092

dep_ratio	-0.143	-0.090	-0.071	-0.402	0.377	0.018	-0.741***	0.372*	0.078
	0.295	0.192	0.154	0.265	0.212	0.158	0.145	0.147	0.088
educ_d	-0.624	0.076	0.013	-0.699*	0.243	-0.002	-1.145***	0.081	0.070
	0.432	0.262	0.236	0.323	0.250	0.219	0.222	0.208	0.132
pcredic_com	0.674*	0.139	0.119	0.019	0.827***	0.385*	-0.161	0.473*	-0.616**
	0.289	0.255	0.212	0.248	0.248	0.196	0.157	0.197	0.118
p_ag_ext_Marketing	0.550	0.520	0.731	0.172	-0.770	-0.373	-0.655*	0.291	0.038
	0.523	0.445	0.375	0.506	0.570	0.433	0.320	0.377	0.235
p_ag_ext_seed	0.088	0.787**	0.446*	-0.156	0.941***	-0.094	-0.384*	0.441*	-0.560***
	0.322	0.270	0.224	0.273	0.276	0.218	0.170	0.217	0.120
prd_c_othr	0.605	-0.200	0.113	-0.334	-1.833	-0.034	-0.024	-0.794**	1.034***
	0.438	0.371	0.309	0.364	0.383	0.296	0.227	0.294	0.163
fsiz1	-0.343	-0.303	-0.436**	0.739***	0.825***	0.319*	0.087	-0.009	-0.407***
	0.210	0.173	0.148	0.178	0.187	0.141	0.109	0.139	0.079
fsiz2	-0.137	-0.556**	-0.207*	0.384**	-0.048	0.166	0.187*	-0.340***	0.205***
	0.147	0.128	0.103	0.129	0.136	0.101	0.076	0.095	0.057
fsiz3	0.026	-0.375**	-0.043	0.232	-0.068	0.014	0.148*	-0.243**	0.191***
	0.136	0.115	0.097	0.129	0.134	0.105	0.074	0.089	0.057
natu_choc1	-0.194	-0.208	-0.154	0.120	0.124	-0.016	0.240***	-0.137	0.036
	0.143	0.121	0.100	0.115	0.123	0.095	0.071	0.094	0.051
natu_choc2	-0.120	0.033	0.139	0.128	0.336	0.104	-0.015	0.186	-0.323***
	0.231	0.208	0.158	0.183	0.207	0.151	0.117	0.152	0.092
AEZ1	-0.085	0.307	0.045	-0.391	-0.123	-0.194	-0.386**	0.414	-0.036
	0.330	0.275	0.230	0.262	0.273	0.218	0.171	0.223	0.123
AEZ2	0.162	0.261	0.223	-0.453*	-0.358	-0.128	-0.484**	0.138	0.187
	0.284	0.237	0.199	0.227	0.229	0.187	0.152	0.192	0.112
AEZ3	0.287	0.312	0.269	-0.275	-0.471	-0.366	-0.244	0.221	-0.086
	0.284	0.242	0.200	0.243	0.272	0.207	0.157	0.197	0.115
hhsiz_est	0.199***	-0.133***	-0.119***	0.179***	-0.080**	-0.057**	0.295***	-0.097***	-0.189***
	0.026	0.025	0.021	0.023	0.026	0.019	0.014	0.019	0.011
resi_eng	1.371	-1.144	0.641	-0.905	-4.299***	-0.429	0.680	-2.147***	2.651***
	0.789	0.643	0.537	0.654	0.691	0.516	0.401	0.509	0.289
resi_mecan	0.373	-0.068	0.376	1.376	1.182	-0.041	1.217	-1.118	-1.035*

	1.490	1.278	1.059	1.186	1.287	1.021	0.751	1.023	0.489
_cons	-1.485	-2.366	0.160	-3.432	-4.275	-1.680	-0.982	-3.035	2.063
	0.648	0.526	0.434	0.543	0.563	0.411	0.327	0.417	0.228
N	8310								
Wald test									
Prob>Chi2(280)	0								
LR-test: Correlation									
chi2c(36)	2846.17								
Prob > Chi2 (36)	0								
LR Test endogeneity									
chi2(18)	177.005								
Prob>chi2(18)	0.00								

*5% ** 1% ***0.1%

fsiz1 =1 if area<=0.2ha, 0 otherwise; fsiz2 =1 if 0.4ha>area>0.2ha, 0 otherwise; fsiz3 =1 if 0.71ha>area>0.4ha, 0 otherwise; fsiz4 =1 if area>=0.71ha, 0 otherwise. NB: According to our theoretical model, prices, wages, and transaction cost measures are absent in equations where the household is autarky in one or both markets. However, we retain ownership of transportation equipment in equation 9 because it is also a measure of wealth. In addition, the presence of producer organizations and cooperatives also influences the adoption of productivity-enhancing agricultural technology. Therefore, they are kept in equations where the household is in autarky.

2. Intensive margin

Estimates of supply functions of the regimes are shown in Table 4. It is important to remember that the variables used to respect an exclusion restriction in Heckman's approach are those that determine the fixed transaction costs[55] i.e. education of the household head, proportion of educated individuals in the household, presence of producer organizations or cooperatives in the community/village, and whether or not the household receives marketing advice from extension services. , The theoretical model shows that the extensive margin depends on both fixed and variable transaction costs, whereas the intensive margin depends only on variable costs. First, it is important to note that the variables used to respect an exclusion restriction, yield relatively satisfactory results since the inverse of the Mills ratio is significant in 3 out of the 9 equations[56].

For Regime 1, farm size, crop yield in community/village, and household size are determinants of grain supply. In theory, these variables affect productivity at the household level. Indeed, the smaller the size of the farm, the lower the household's productive capacity. On the other hand, when the household is in a village where crop yields are high, it is likely that the household will generate marketable surpluses.

Regarding Regime 2, the price of cereals seems to be a determining factor. In the whole sample (column 10), the application of seed use advice has a positive effect on grain production. However, for Regime 2, which hires external labor, this positive effect is mitigated by the fact that the

---

[55] Heltberg & Tarp (2002) used the same restriction strategy

[56] Regimes 2, 3 and 5. For the last two, the significance is at the 10%.

application of these improved seeds would require the use of other factors[57] (complementarity in production) that would increase production costs, which could limit the increase in output.

For Regime 3, the increase in production may be due to higher prices and/or intensive use of inorganic fertilizers; the smaller the size of the farm, the lower the production; the higher dependency ratio reduces the household's productive labor force, and thus production. As in Regime 2, household size influences production.

The gender appears to play a key role in determining supply in Regime 4. In Regime 5, prices on the one hand and crop yields in the village/community on the other influence grain supply; household size and the experience of the household head (age) also play a positive role in this regime.

The size of the farm increases the supply of grain while a higher dependency ratio limits production in Regime 6. Transaction costs related to labor mobility are important in determining farm supply in Regime 7. The higher these costs, the lower the actual wage received by the household with an off-farm labor supply; following the budget constraint, the consequence is a decrease in the household's off-farm income, thus reducing household's expenditure (including purchases of productive inputs). In addition, inorganic fertilizer use, farm size, gender (male) and experience of the household head also play a positive role.

[57] Fertilizers, external labor, pesticides, etc...

Farm size and crop yield in the village appear to be the main determinants of grain supply in Regime 8. In autarky (Regime 9), grain supply is positively influenced by the use of inorganic fertilizers, farm size, experience (age), gender (male) of the household head, and access to transportation facilities (at village level). However, producing other crops in addition to cereals may put pressure on the use of factors (land, labor, etc.) that may limit cereal production. Moreover, the use of mechanization (only) does not always seem to correlate with an increase in production. There are two potential explanations for this observation: (i) it may be that land under mechanization is not allocated solely to cereal production, since more than 41% of farmers also produce other crops in addition to cereals; it is therefore possible that the household's total production increases; (ii) for this group, the use of mechanization may be associated with the purchase of external labor, showing the complementarity between the factors[58] - traction/machinery and labor - during the agricultural season. In this case, the promotion of mechanization of agricultural activities to increase production should be accompanied by incentives from decision makers to enable the household to hire external labor.

[58] Wineman & Jayne (2020) also underline the hypothesis of complementarity. However, Liu et al (2016) point out that a possible substitution exists between these two factors during the soil preparation phase.

Table 4 : Estimation of the supply function

	Regime1	Regime2	Regime3	Regime4	Regime5	Regime6	Regime7	Regime8	Regime9	All sample
	1	2	3	4	5	6	7	8	9	10
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
equipt_c	-0.011	-0.024	0.519	-0.41	0.258	0.296	0.157	0.727	0.567**	0.278***
	0.625	0.627	0.274	0.784	0.569	0.362	0.203	0.375	0.19	0.071
dist_road	-0.738	0.34	-0.271	-1.13	0.58	-0.458	-0.721***	0.082		-0.632***
	0.63	0.604	0.428	0.656	0.785	0.627	0.216	0.64		0.111
tprc	-0.09	-0.053	0.071	0.111	0.204	0.014	0.061	0.015		0.021
	0.124	0.112	0.083	0.131	0.216	0.109	0.049	0.138		0.025
price	-0.123	0.064*	0.058**	0.02	0.442*	-0.004				0.007
	0.088	0.029	0.02	0.062	0.203	0.1				0.005
wag	0.012	0.009		0	-0.001		0.002	0.003		0.002
	0.01	0.01		0.009	0.007		0.003	0.012		0.001
share_sup_mecan	2.698	-2.246	0.369	0.772	0.94	0.071	-0.234	-2.036	-2.396***	0.445
	1.952	2.183	1.466	2.875	2.044	1.658	0.792	1.343	0.654	0.399
share_sup_engrais	0.137	-0.734	1.600**	0.914	2.23	0.87	1.491**	3.071	2.399**	1.828***
	1.282	1.464	0.579	3.299	1.249	1.098	0.533	1.883	0.828	0.263
share_sup_manure	-0.028	0.387	-0.095	-0.059	-0.191	0.081	0.092	0.176	0.063	0.022
	0.157	0.214	0.09	0.14	0.179	0.147	0.058	0.11	0.046	0.028
rdt_c	0.059*	-0.011	0.065	-0.009	0.036**	0.054	0.027	0.052**	0.061	0.042***
	0.03	0.035	0.037	0.037	0.014	0.032	0.025	0.02	0.054	0.005
Head_homm_d	0.001	0.029	0.032	0.295*	-0.306	0.055	0.183***	0.199	0.098*	0.164***
	0.116	0.305	0.101	0.145	0.28	0.118	0.047	0.107	0.048	0.024
age_head	-0.006	0.012*	0.002	0.006	0.016*	0.01	0.004*	0.007	0.003***	0.003***
	0.004	0.005	0.003	0.006	0.007	0.007	0.002	0.004	0.001	0.001
sh_fem	-0.128	-0.005	0.072	-0.246	-0.246	0.11	0.078	-0.041	0.003	0.003
	0.297	0.361	0.177	0.283	0.304	0.205	0.146	0.188	0.072	0.05
dep_ratio	-0.667	0.308	-0.573**	0.429	-0.154	-0.542*	-0.107	0.242	-0.089	-0.161***
	0.37	0.307	0.191	0.452	0.422	0.266	0.196	0.352	0.069	0.048
pcredic_com	0.031	0.144	0.204	0.627	-0.681	0.041	0.246	-0.021	0.241	0.146*
	0.448	0.396	0.262	0.659	0.363	0.602	0.155	0.44	0.187	0.063
p_ag_ext_seed	-0.233	-1.533*	-0.235	0.175	-0.883	0.162	0.35	0.066	0.374	0.178*

	0.336	0.779	0.198	0.623	0.602	0.291	0.194	0.453	0.192	0.086
prd_c_othr	0.96	0.008	-0.203	0.636	-0.176	-0.116	-0.129	-0.858	-0.751*	-0.168
	0.674	0.528	0.33	1.37	0.572	0.398	0.24	0.727	0.348	0.128
fsiz1	-0.773**	0.127	-0.356	-0.596	0.438	-1.086*	-0.699***	-0.552**	-0.555***	-0.633***
	0.294	0.334	0.251	0.573	0.482	0.461	0.108	0.187	0.152	0.053
fsiz2	-0.826***	-0.024	-0.366*	-0.202	-0.123	-0.692**	-0.672***	-0.890***	-0.678***	-0.761***
	0.179	0.547	0.182	0.18	0.275	0.249	0.07	0.251	0.076	0.035
fsiz3	-0.355**	0.242	-0.166	0.057	-0.342	-0.517***	-0.332***	-0.563**	-0.491***	-0.437***
	0.137	0.334	0.146	0.165	0.229	0.144	0.076	0.196	0.073	0.032
natu_choc1	-0.232	0.29	-0.082	-0.286	0.095	-0.179	-0.074	0	-0.022	-0.212***
	0.16	0.266	0.135	0.217	0.225	0.15	0.068	0.145	0.047	0.036
natu_choc2	-0.166	-0.111	-0.104	0.181	-0.483	0.171	-0.151	-0.1	0.104	-0.065
	0.315	0.321	0.189	0.423	0.401	0.246	0.104	0.28	0.133	0.048
AEZ1	0.844*	-0.028	0.4	0.268	-0.29	-0.105	0.191	-0.351	-0.388**	0.265***
	0.376	0.467	0.305	0.491	0.44	0.447	0.18	0.384	0.145	0.073
AEZ2	0.465	-0.455	0.064	-0.22	-0.703	-0.279	-0.034	-0.713**	-0.606***	-0.052
	0.336	0.389	0.26	0.51	0.375	0.347	0.148	0.232	0.135	0.057
AEZ3	1.020**	0.118	0.484*	0.964	-0.094	0.45	0.416**	-0.063	0.029	0.529***
	0.364	0.406	0.233	0.588	0.406	0.651	0.144	0.293	0.116	0.056
hhsiz_est	0.220**	0.256*	0.125*	0.02	0.305*	0.028	-0.005	0.011	0.106*	0.035***
	0.082	0.123	0.059	0.065	0.122	0.085	0.034	0.065	0.052	0.005
resi_eng	0.025	0.43	-1.307*	-0.974	-2.243	-0.477	-1.131*	-2.836	-1.987*	-1.508***
	1.312	1.318	0.586	3.16	1.235	0.962	0.527	1.815	0.811	0.264
resi_mecan	-2.978	1.593	-0.605	-0.838	-1.328	-0.137	0.196	2.075	2.404***	-0.48
	1.98	2.122	1.443	2.747	2.142	1.651	0.8	1.328	0.678	0.399
imr1	0.604									
	0.428									
imr2		-1.892*								
		0.965								
imr3			0.517							
			0.308							
imr4				-0.097						
				0.732						

imr5					-2.061					
					1.053					
imr6						-0.583				
						1.595				
imr7							-0.177			
							0.162			
imr8								0.544		
								0.779		
imr9									-0.413	
									0.378	
educ_HHd										0.014
										0.04
nb_ag_coop										-0.030***
										0.007
nb_farm_group										-0.008
										0.006
educ_d										0.146
										0.081
p_ag_ext_Marketing										0.737***
										0.139
_cons	3.638**	9.932***	3.081**	4.38	5.933***	6.094	4.695***	3.469	4.720***	4.291***
	1.259	2.824	1.18	3.215	1.152	3.413	0.418	2.656	0.326	0.168
N	221	267	525	291	249	489	1552	526	3264	8310
p	3.30E-91	1.80E-208	3.37E-77	2.14E-87	8.78E-23	4.31E-68	2.70E-157	2.91E-66	0	0
F										

*5% ** 1% ***0.1%

3. Robustness check

In this section, we explore another approach to check the robustness of our results in the estimation of the selection equation. Indeed, the extension of Heckman's approach proposed in this paper relies on a Multivariate Probit (MVP) to empirically estimate the selection equation. However, for reasons to do with the evaluation of the likelihood function of MVP (see Cappellari & Jenkins, 2003), the estimation is based on simulation techniques, whose effectiveness depends on the number of simulations carried out (number of random draws).

We use the nested logit model (NL); which is an extension of the multinomial logit model (ML). It is close to the multinomial probit with correlated errors (MP) and the multinomial mixed logit (MLM). It partially relaxes the hypothesis of independence of irrelevant alternatives (IIA) [59], and presents an analytical form of the likelihood function that can be maximized without using simulation techniques (Appendix A).

Results[60] for the Nested Logit model are presented in Table 5. The IIA test shows that the hypothesis of independence between the utilities can be rejected. The coefficient of independence between the alternatives in Group 1 (Tau) [61] is positive. These results confirm those of the MVP

[59] This hypothesis implies that the utilities associated with the different alternatives are independent (IIA)

[60] In order not to lose information by removing households that are both buyers and sellers in the same market (11% of the sample), we first made a slight reclassification of households in this group. In the cereal market, they were classified as net buyers; this was because households in Malawi are usually net buyers of cereals (Ricker-Gilbert & Jayne 2017). In the labor market, they were classified as net suppliers because on these small farms there is often a surplus of labor due in part to low productive capacity (small area under cultivation), or the seasonality of rainfed agriculture (De Janvry, Duquennois, & Sadoulet (2020).

[61] This is a condition of validity for the use of the NL. However, the coefficient is greater than 1, which implies that the estimated model remains consistent with the theoretical framework of the random utility model for a subset of the

which showed that efficiency is gained when the correlation between the alternatives is considered. In addition, the introduction of control functions to correct for endogeneity of fertilizer and mechanization use improves the quality of the estimates. (see the likelihood ratio test).

Our estimates (Part A of Table 5) indicate that when the crop yield in villages/communities is large, the household is highly likely to be in the participant group given productivity gains. The smaller the size of the farm, the lower the likelihood of being in the participant group[62]. The education of the household head, access to credit, use of improved seeds, and access to transportation facilities play a positive role in the choice of market participation. Producing other crops in addition to cereals reduces the likelihood of intervening in the market. These results confirm those observed with the MVP on the determinants of membership in Regime 9.

In Part B of Table 5, the determinants of regime choice are compared to regime 9 (with autarky as reference) are presented. As in the MVP, it appears that the price of grain negatively influences the probability of being a net seller; on the opposite, it increases the likelihood of being a net buyer of grain, regardless of labor market status. Consistent with the MVP results, we observe that wage increases enhance the likelihood of being a net seller of labor. In line with the results of the MVP, the use of inorganic fertilizers improves the likelihood of having a marketable grain surplus (Regime 2); it also encourages the use of hired labor (Regime 2 and 8). Finally, in both specifications, the size of the household positively influences the choice of regimes in which the

---

values taken by the explanatory variables (Train, 2003). The coefficient of independence is set at 1 for group 2, which is degenerate.

[62] However, when the size is relatively too small, area<=0.2ha (fsiz1), it is possible that the household has a surplus of labor that can be used in non-farm activities.

household is a net seller of labor, and negatively affects the choice of regimes in which it is a net buyer of labor; these effects are not affected by the position in the grain market. Moreover, in the NL model, we clearly observe that mechanization would encourage the employment of hired labor (Regime 8).

Table5 Estimation of nested logit

Part A	
	b/se
dist_road	-0.958**
	0.358
equipt_c	0.905***
	0.266
tprc	-0.023
	0.081
nb_ag_coop	0.021
	0.019
nb_farm_group	-0.020
	0.018
rdt_c	0.179***
	0.029
Head_homm_d	-0.186
	0.091
educ_HHd	0.364**
	0.133
age_head	-0.008**
	0.003
sh_fem	0.112
	0.175
dep_ratio	-0.246
	0.171
educ_d	-0.545*
	0.260
pcredic_com	1.076***
	0.233
p_ag_ext_Marketing	-0.237
	0.464

p_ag_ext_seed	0.891***
	0.248
prd_c_othr	-2.194***
	0.338
fsiz1	0.801***
	0.161
fsiz2	-0.301**
	0.110
fsiz3	-0.315**
	0.109
natu_choc1	-0.110
	0.106
natu_choc2	0.498**
	0.177
AEZ1	0.428
	0.255
AEZ2	-0.225
	0.222
AEZ3	0.400
	0.232

Part B	Regime1	Regime2	Regime3	Regime4	Regime5	Regime6	Regime7	Regime8
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
price	-1.283***	-0.118	-2.985***	0.443***	0.509***	0.445***		
	0.187	0.125	0.284	0.059	0.064	0.057		
wag	0.086***	-0.065	0.000	0.106***	-0.058	0.000	0.043	-0.179
	0.025	0.048	0.000	0.019	0.040	0.000	0.013	0.044
share_sup_mecan	-6.252**	1.084	-2.754	-3.248	1.799	0.361	-1.476	4.380*
	2.423	2.569	1.953	1.959	2.302	1.907	1.376	1.925
share_sup_engrais	4.296***	8.358***	2.843**	5.320***	11.061***	5.868***	2.209**	9.201***
	1.204	1.362	0.972	0.954	1.242	0.933	0.716	1.039
share_sup_manure	-1.224	-1.025	-0.085	-0.272	0.209	0.378	0.128	-0.258
	0.629	0.696	0.420	0.436	0.519	0.390	0.219	0.430
hhsiz_est	1.282***	-0.503***	-0.498***	1.090***	-0.428***	-0.294**	1.205***	-0.518***

		0.124	0.144	0.107	0.093	0.122	0.097	0.069	0.105
resi_eng		-3.406**	-7.378***	-2.621**	-5.252***	-10.049***	-5.230***	-2.738***	-8.634***
		1.214	1.330	0.976	0.955	1.204	0.921	0.710	1.022
resi_mecan		6.312**	0.276	3.239	3.425	-1.829	-0.883	1.403	-3.829*
		2.417	2.593	1.956	1.950	2.301	1.890	1.372	1.935
_cons		-11.240	-9.076	4.292	-16.700	-13.377	-9.598	-9.769	-7.464
		1.379	1.394	0.962	1.311	1.313	0.911	0.727	0.915
Tau(G1)	2.815								
	0.251								
Tau(G2)	1								
N	74790								
ll	-12518.24								
p	0.00								
LR test IIA									
chi2	181.087								
Prob>chi2	0.00								
LR Test endogeneity									
chi2	197.11								
Prob>chi2	0.00								

*5% ** 1% ***0.1%

Conclusion

The transition from semi-subsistence agriculture to a marketed system in which producers are fully integrated into the market is seen as a catalyst in the structural transformation of economies whose growth is strongly influenced by agriculture. This argument is based on the belief that it allows small producers to participate in more remunerative value chains, thus giving the poor a way out of poverty and ensuring food security.

The literature on the determinants of participation in the food market in sub-Saharan Africa can be summarized in two waves. The first one has focused on price and market access factors to explain the constraints small producers face (Goetz, 1992, as a pioneer); the second one stresses that price and market access policies are necessary but not sufficient and therefore identifies productivity constraints that limit the effects of price and market access policies. (Alene et al., 2008; Barrett, 2008; Mather, Boughton, & Jayne, 2013). This study is part of the second wave of the literature; unlike the others, it introduces the rural labor market into the theoretical model developed by Key, Sadoulet, & Janvry, (2000) to assess the response of small producers to the incentives of labor/cereal markets and productivity shocks. We use Malawi data to conduct model estimation through an extension of Heckman's approach to the multivariate case and the introduction of control functions to minimize endogeneity problems. We first find that reducing the costs associated with the mobility of factors and goods through access to transportation infrastructure is important in promoting participation in grain and labor markets. Second, the presence of producer organizations would stimulate the household's market orientation (as net sellers in both markets) as it enables its members to minimize transaction costs and a better diffusion of agricultural technology. Third, through its potential effect on productivity, the use of inorganic fertilizers encourages the production of a marketable surplus and the employment of hired labor. Fourth, training from extension services on marketing strategies and use of improved seeds can be key tools to increase participation in markets. Fifth, price incentives can be counterintuitive, i.e. produce effects contrary to those expected in promoting market orientation even though the effect on sales may be positive for those who are already participants.

These results highlight a set of instruments for better targeting in the implementation of policies aimed at transforming the rural areas of developing countries through the integration of small farms into both food and labor markets. As such, it can be a powerful source of support to reduce migration to the urban areas where job opportunities are becoming more limited in the context of Malawi[63].

Finally, this empirical application is based on hypotheses that could be relaxed in possible extension. Hence, the proposed theoretical model does not include intra-household decisions in the use of resources. In addition, it would also be interesting to investigate the determinants of the entries and exits of a regime in a dynamic setting. Indeed, few authors explore drivers of market participation in a dynamic framework (Andersson et al., 2015;. Michelson, 2017; Romero Granja & Wollni, 2018) but these applications are not exempt from problems, namely the limited representativeness of the sample ; there focus on one crop or one market segment (export) while

[63] De Janvry, Duquennois, and Sadoulet (2020)

the producer can simultaneously participate in several market segments allowing him/her to have different market positions depending on the crop nature. Considering possible participation in two markets in a dynamic context would be a nice extension to our contribution.

References

Aberman, Noora-Lisa, and Terry Roopnaraine. 2020. “To Sell or Consume? Gendered Household Decision-Making on Crop Production, Consumption, and Sale in Malawi.” *Food Policy* ,1-15 Accessed June 29, 2020. https://doi.org/10.1007/s12571-020-01021-2.

Adjognon, Serge G., Lenis Saweda O. Liverpool-Tasie, and Thomas A. Reardon. 2017. “Agricultural Input Credit in Sub-Saharan Africa: Telling Myth from Facts.” *Food Policy* 67 (February): 93–105. https://doi.org/10.1016/j.foodpol.2016.09.014.

Ahituv, Avner, and Ayal Kimhi. 2002. “Off-Farm Work and Capital Accumulation Decisions of Farmers over the Life-Cycle: The Role of Heterogeneity and State Dependence.” *Journal of Development Economics* 68 (2): 329–53. https://doi.org/10.1016/S0304-3878(02)00016-0.

Alene, Arega D., V. M. Manyong, G. Omanya, H. D. Mignouna, M. Bokanga, and G. Odhiambo. 2008. “Smallholder Market Participation under Transactions Costs: Maize Supply and Fertilizer Demand in Kenya.” *Food Policy* 33 (4): 318–28. https://doi.org/10.1016/j.foodpol.2007.12.001.

Amare, Mulubrhan, and Bekele Shiferaw. 2017. “Nonfarm Employment, Agricultural Intensification, and Productivity Change: Empirical Findings from Uganda.” *Agricultural Economics (United Kingdom)* 48: 59–72. https://doi.org/10.1111/agec.12386.

Andersson, Camilla I M, Christine G K Chege, J O Rao, and Matin Qaim. 2015. “following up on smallholder farmers and supermarkets in kenya.” *American Journal of Agricultural Economics* 97(4), 1247-1266. Accessed December 29, 2018. https://doi.org/10.1093/ajae/aav006.

Arndt, Channing, Karl Pauw, and James Thurlow. 2016. “The Economy-Wide Impacts and Risks of Malawi’s Farm Input Subsidy Program.” *American Journal of Agricultural Economics* 98 (3): 962–80. https://doi.org/10.1093/ajae/aav048.

Bezu, Sosina, Girma T. Kassie, Bekele Shiferaw, and Jacob Ricker-Gilbert. 2014. “Impact of

Improved Maize Adoption on Welfare of Farm Households in Malawi: A Panel Data Analysis." *World Development* 59: 120–31. https://doi.org/10.1016/j.worlddev.2014.01.023.

Burke, William J., Robert J. Myers, and Thomas S. Jayne. 2015. *A Triple-Hurdl Emodel of Production and Market Participation in Kenya's Dairy Market*. *American Journal of Agricultural Economics*. Vol. 97. https://doi.org/10.1093/ajae/aav009.

Cappellari, Lorenzo, and Stephen P Jenkins. 2003. "Multivariate Probit Regression Using Simulated Maximum Likelihood." *The Stata Journal*. Vol. 3.

Carletto, Calogero, Paul Corral, and Anita Guelfi. 2017. "Agricultural Commercialization and Nutrition Revisited: Empirical Evidence from Three African Countries." *Food Policy*. https://doi.org/10.1016/j.foodpol.2016.09.020.

"Document of the World Bank Malawi Poverty and Vulnerability Assessment Poverty Reduction and Economic Management 1 Africa Region." 2007.

Goetz, Stephan J. 1992. "A Selectivity Model of Household Food Marketing Behavior in Sub-Saharan Africa." *American Journal of Agricultural Economics* 74 (2): 444–52. https://doi.org/10.2307/1242498.

Handschuch, Christina, and Meike Wollni. 2016. "Traditional Food Crop Marketing in Sub-Saharan Africa: Does Gender Matter?" *Journal of Development Studies* 52 (3): 343–59. https://doi.org/10.1080/00220388.2015.1068289.

Janvry, Alain De, Claire Duquennois, and Elisabeth Sadoulet. 2020. "Labor Calendars and Rural Poverty: A Case Study for Malawi." http://iresearch.worldbank.org/PovcalNet.

Key, Nigel, Elisabeth Sadoulet, and Alain De Janvry. 2000. "Transactions Costs and Agricultural Household Supply Response." *American Journal of Agricultural Economics* 82 (2): 245–59. https://doi.org/10.1111/0002-9092.00022.

Matchaya, Greenwell C. 2012. "Estimating effects of constraints on food security in malawi:

policy lessons from regressions quantiles." *Applied Econometrics and International Development*. Vol. 12.

Mather, David, Duncan Boughton, and T S Jayne. n.d. "Explaining Smallholder Maize Marketing in Southern and Eastern Africa: The Roles of Market Access, Technology and Household Resource Endowments." *Food Policy* 43:248-266 Accessed June 20, 2020. https://doi.org/10.1016/j.foodpol.2013.09.008.

Michelson, Hope C. 2017. "Influence of Neighbor Experience and Exit on Small Farmer Market Participation." *American Journal of Agricultural Economics*. https://doi.org/10.1093/ajae/aaw097.

Mussema, Rehima, Belay Kassa, Dawit Alemu, and S Rashid. 2013. "Analysis of the Determinants of Small-Scale Farmers' Grain Market Participations in Ethiopia: The Contribution of Transaction Costs." *J. Agric. Sci* 23: 75–94.

Olwande, John, Melinda Smale, Mary K. Mathenge, Frank Place, and Dagmar Mithöfer. 2015. "Agricultural Marketing by Smallholders in Kenya: A Comparison of Maize, Kale and Dairy." *Food Policy* 52: 22–32. https://doi.org/10.1016/j.foodpol.2015.02.002.

Ouma, Emily, John Jagwe, Gideon Aiko Obare, and Steffen Abele. 2010. "Determinants of Smallholder Farmers' Participation in Banana Markets in Central Africa: The Role of Transaction Costs." *Agricultural Economics* 41 (2): 111–22. https://doi.org/10.1111/j.1574-0862.2009.00429.x.

Petrin, Amil, and Kenneth Train. 2002. "Omitted Product Attributes in Discrete Choice Models."

Poole, Nigel. 2017. *Smallholder Agriculture and Market Participation*. https://doi.org/10.3362/9781780449401.

Renkow, Mitch, Daniel G. Hallstrom, and Daniel D. Karanja. 2004. "Rural Infrastructure, Transactions Costs and Market Participation in Kenya." *Journal of Development Economics* 73 (1): 349–67. https://doi.org/10.1016/j.jdeveco.2003.02.003.

Ricker-Gilbert, Jacob, and T. S. Jayne. 2017. "Estimating the Enduring Effects of Fertiliser Subsidies on Commercial Fertiliser Demand and Maize Production: Panel Data Evidence from Malawi." *Journal of Agricultural Economics* 68 (1): 70–97. https://doi.org/10.1111/1477-9552.12161.

Romero Granja, Cristina, and Meike Wollni. 2018. "Dynamics of Smallholder Participation in Horticultural Export Chains: Evidence from Ecuador." *Agricultural Economics (United Kingdom)*. https://doi.org/10.1111/agec.12411.

Sadoulet, Elisabeth, and Alain De Janvry. 1995. "Quantitative Development Policy Analaysis," 1–438. https://doi.org/10.2307/1243800.

Sadoulet, Elisabeth, Alain De Janvry, and Catherine Benjamin. 1998. "Household Behavior with Imperfect Labor Markets." *Industrial Relations* 37 (1): 85–108. https://doi.org/10.1111/0019-8676.731998036.

Takeshima, Hiroyuki, and Ephraim Nkonya. 2014. "Government Fertilizer Subsidy and Commercial Sector Fertilizer Demand: Evidence from the Federal Market Stabilization Program (FMSP) in Nigeria." https://doi.org/10.1016/j.foodpol.2014.04.009.

Tobergte, David R., and Shirley Curtis. 2013. *Applied Choice Analysis*. *Cambridge University Press*. Vol. 53. https://doi.org/10.1017/CBO9781107415324.004.

Train, Kenneth E. 2003. *Discrete Choice Methods with Simulation*. *Discrete Choice Methods with Simulation*. Vol. 9780521816. https://doi.org/10.1017/CBO9780511753930.

Vakis, Renos, Elisabeth Sadoulet, and Alain De Janvry. 2003. "Measuring Transactions Costs from Observed Behavior: Market Choices in Peru." *Working Paper, GRADE*, no. October: 20. https://escholarship.org/uc/item/7p81h66q.pdf.

Wineman, Ayala, and Thomas S. Jayne. 2020. "Factor Market Activity and the Inverse Farm Size-Productivity Relationship in Tanzania." *Journal of Development Studies* 00 (00): 1–22. https://doi.org/10.1080/00220388.2020.1797686.

Woldeyohanes, Tesfaye, Thomas Heckelei, and Yves Surry. 2017. "Effect of Off-Farm Income on Smallholder Commercialization: Panel Evidence from Rural Households in Ethiopia." *Agricultural Economics (United Kingdom)* 48 (2): 207–18. https://doi.org/10.1111/agec.12327.

Wooldridge, Jeffrey M Wooldridge. 2015. "Econometrics," no. April 2013

Zhou, Shadreck, Isaac J. Minde, and Bruce Mtigwe. 2013. "Smallholder Agricultural Commercialization for Income Growth and Poverty Alleviation in Southern Africa: A Review." *African Journal of Agricultural Research* 8 (22): 2599–2608. https://doi.org/10.5897/AJAR11.1040.

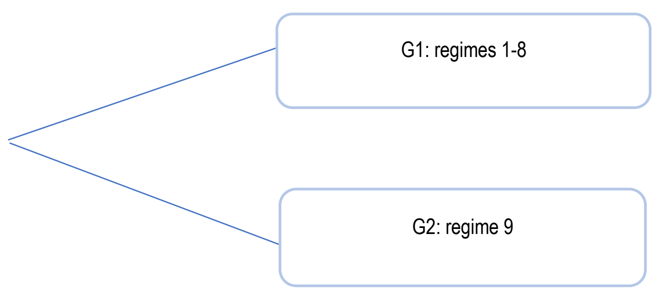
Annex A: NESTED LOGIT

The Nested Logit pools together regimes with similarities. Within each group thus formed, the alternatives are assumed to be correlated and there is independence between groups. The probability of choosing an alternative is derived by taking the product of the probability of choosing the group and the probability of choosing the alternative conditional on the group chosen. The NL model used here presents the pooling in Figure 2. Group 1 (G1) includes regimes 1 to 8 and Group 2 (G2) includes only the regime where the household is inactive in markets.

We isolate group G2 by the following criteria i) households belonging to regime 9 (autarky) alone constitute 40% of our sample; ii) in theory, households in autarky suffer from factor endowment constraints that prevent them from engaging in initiatives to take advantage of market opportunities; iii) households in group G1 are similar in that they bear idiosyncratic transaction costs in markets. (one or both markets).

Hensher, Rose & Greene (2005) point out that, contrary to what one might think, the clustering presented in the NL models is not based on a theory that the decision is sequential, i.e. the decision maker chooses the group (node), then the alternative in the group. Mainly, these authors draw attention to the fact that the only purpose of clustering is to distinguish only similarities between alternatives to circumvent the IIA hypothesis in standard ML. In this section, we maintain that the NL presented is based on the model described above; that is, the household chooses from among the 9 regimes, the alternative that provides the most utility.

Figure 1: NL



Let V_{nj} be the utility that an individual household derives from choosing regime j in node n :

$$V_{nj} = \alpha_j X_{n,j} + \beta_n Z_n + \epsilon_{nj}, \quad j = 1, \dots .9; n = 1, 2, 3 \quad (8)$$

X is the matrix of variables that defines the characteristics of the alternatives (j); it includes variables that vary according to the household and the alternatives. We consider prices, wages, technology use (inorganic fertilizer use and mechanization) as the main components of the matrix X; in other words, the associated coefficients (α_j) are specific to each regime.

In literature on applied NL models (see Heiss, 2002) Z_n includes the socio-economic characteristics of the household (age, education of the head of household, ...); β_n is the matrix of associated coefficients; ϵ_{nj} are the error terms; its cumulative distribution is :

$$\exp\left(- \sum_{n=1}^{G} \left(\sum_{j} e^{\epsilon_j/\tau_n} \right)^{\tau_n} \right)$$

This distribution is of the GEV (Gumbel Extreme Value) type; it is a generalization of the standard logit model (ML). While in the standard logit, the error terms are independent, in the nested logit (NLM), they are correlated within each group (node) (Train 2003):

$$\mathrm{Cov}(\epsilon_j, \epsilon_m) \neq 0 \;\; \text{if } j \neq m$$

τ_n measures the coefficient of independence within each group. It is assumed to be constant between alternatives within the same group::

$$\tau_n = \sqrt{1\text{-}\rho_n} \quad \in (0, 1]$$

ρ_n is the coefficient of correlation within the group *n*.

For a household, the probability of choosing alternative j conditional on the choice of group n, $P(j|n)$, the probability of being in group n, $P(n)$, and the probability of choosing alternative j, $P(j)$ can be expressed as follows :

$$P(j|n) = \frac{\exp(\alpha_j X_{n,j})}{\sum_m \exp(\alpha_m X_{n,m})} \quad (9)$$

$$P(n) = \frac{\exp(\beta_n Z_n + \tau_n I_n)}{\sum_k \exp(Z_n \beta_n + \tau_n I_n)} \; (10)$$

$$I_n = \ln\left\{\sum_m \exp\left(\alpha_m X_{g,m}/\tau_n\right)\right\} \quad (11)$$

$$P(j) = P(n)^*P(j|n) \quad (12)$$

Equation (10) shows that the choice of the node is not only influenced by the variables in the matrix Z_n, but also those in the matrix $X_{n,j}$through the inclusivity parameter I_n.

For N households having made their respective choices, i is the household indicator, and y_{inj} is a dichotomous variable revealing choice j in node n, of household i. The logarithm of likelihood is written as follows

$$\log l = \sum_i \sum_n \sum_j y_{inj} \log P_{ij} = \sum_i \sum_n \sum_j y_{inj} \log\{P_{in}{}^*P_{ij|n}\} \quad (13)$$

Estimation consists in finding the model parameters that maximize this function.

However, we can highlight that there is a difference in the specifications of the MVP and NL models; in the latter, we assume that the variables determining the choice of group should not technically and directly enter into the variables explaining the choice of alternatives within the groups. That is, the same explanatory variable cannot enter the group selection and alternative choice equations at the same time; this does not mean that there is independence between the two. Moreover, in the NL, the choices of alternatives are made relative to the reference regime (regime 9), which is not the case in the MVP.