ELSEVIER
www.elsevier.com/locate/worlddev

World Development Vol. 39, No. 5, pp. 784–796, 2011

0305-750X/$ - see front matter

doi:10.1016/j.worlddev.2010.09.005

Supermarkets, Farm Household Income, and Poverty: Insights from Kenya

ELIZAPHAN J.O. RAO and MATIN QAIM *
Georg-August-University of Goettingen, Goettingen, Germany

Summary. — The expansion of supermarkets in developing countries may have important implications for poverty and rural development. While previous studies have compared farm profits between participants and non-participants in supermarket channels, wider income effects have hardly been analyzed. Moreover, most existing studies do not account for structural differences between the two groups. We address these issues by using endogenous switching regression and building on a survey of vegetable farmers in Kenya. Participation in supermarket channels is associated with a 48% gain in average household income, which also contributes to poverty reduction. To realize these benefits on a larger scale will require institutional support.

Key words — supermarkets, household income, sample selection, endogenous switching regression, Kenya, Africa

1. INTRODUCTION

The share of supermarkets in developing country food retailing has increased significantly in the recent past (Mergenthaler, Weinberger, & Qaim, 2009; Neven, Reardon, Chege, & Wang, 2006; Reardon, Timmer, Barrett, & Berdegue, 2003). This is affecting food consumers, but it also has far-reaching consequences for agricultural producers, because supermarket procurement channels are more integrated than traditional supply chains and have higher requirements in terms of product quality and consistency. Especially for perishable products, supermarket procurement often involves contractual arrangements with farmers. From a development policy perspective, it is particularly important to understand how poor rural households are affected. For farmers, participation in supermarket channels might be associated with market assurance as well as stability in volumes and prices, potentially entailing rising incomes. However, technical constraints and market imperfections might also lead to smallholder exclusion, which could result in increasing disparities and marginalization.

There is a growing body of literature analyzing such aspects in various developing countries (e.g., Neven & Reardon, 2004; Pingali, Khwaja, & Meijer, 2007; Reardon, Barrett, Berdegue, & Swinnen, 2009). Different studies have examined determinants of farmer participation in supermarket channels (Hernandez, Reardon, & Berdegue, 2007; Moustier, Tam, Anh, Binh, & Loc, 2009; Neven, Odera, Reardon, & Wang, 2009). There are also a few studies that have tried to assess economic effects by comparing gross margins for specific crops that are either supplied to supermarkets or traditional channels (Hernandez *et al.*, 2007; Neven *et al.*, 2009). While such comparisons may provide an indication of income effects, differences cannot be interpreted as net impacts of supermarket participation, because there may be other factors influencing the outcome that are not controlled for. Furthermore, simple gross margin analysis does not account for possible resource reallocation between different economic activities within the household occasioned by supermarket participation. To our knowledge, there are no previous studies that have looked more comprehensively into the impacts of supermarkets on household income and implications for poverty, as we do in this article.

A related strain of literature focuses on modernizing export supply chains for high-value foods and the increasing role of standards in international trade (Asfaw, Mithöfer, & Waibel, 2009, 2010; Bolwig, Gibbon, & Jones, 2009; Maertens & Swinnen, 2009; Minten, Randrianarison, & Swinnen, 2009; Miyata, Minot, & Hu, 2009; Warning & Key, 2002; Wollni & Zeller, 2007). Some of these studies also look at income and poverty effects in the small farm sector, mostly building on standard treatment models that account for non-random sample selection. This strain of literature is very relevant for our work, because the conditions in high-value export chains are often similar to those in supermarket channels. However, standard treatment models assume uniform impacts across different groups, whereas recent evidence suggests that there may be systematic differences between farmers supplying supermarkets and their counterparts in traditional channels (Hernandez *et al.*, 2007; Neven *et al.*, 2009). Household income structures are therefore likely to differ systematically, especially if participation in supermarket channels is determined by the same factors that affect income. In that case, assuming uniform impacts conceals inherent interaction between market channel choice and other factors influencing income, potentially leading to spurious conclusions.

We address these issues by using an endogenous switching regression model that treats marketing channels as regimes and thus allows for structural differences in income functions between farmers supplying supermarkets and traditional channels. Similar models have been used in other studies related to agriculture (e.g., Fuglie & Bosch, 1995), including for production function estimates in the context of supermarket developments (Hernandez *et al.*, 2007). We refine and extend the approach such that net household income and poverty effects of supermarket participation can be estimated. The empirical research builds on primary household level data from a survey of vegetable farmers in central Kenya. Even though the overall share of supermarkets in vegetable retailing is still relatively

* This research was financially supported by the German Research Foundation (DFG). Constructive comments from three anonymous referees of this journal are gratefully acknowledged. Final revision accepted: September 20, 2010.

small in Kenya, it is increasing rapidly (Neven & Reardon, 2004). Similar trends are also observed in other countries of Africa. Since many African smallholders are involved in horticultural production, there may be important effects for rural welfare and poverty.

The article proceeds as follows. In the next section, we present the analytical framework and estimation procedure. In Section 3, we describe the data and undertake some descriptive analyzes, while in Section 4 we present and discuss the estimation results. Section 5 concludes.

2. ANALYTICAL FRAMEWORK AND ESTIMATION PROCEDURE

Participation in supermarket channels can be viewed as a binary choice decision problem by farm households that try to maximize utility or net returns.[1] Utility is determined by a set of variables Z, which influence the ability and the cost of adjusting to a market option with new requirements (such as acquiring information and implementing new market standards). Variables in Z also determine the relative returns that a farmer can earn from supermarket and traditional market channels. Thus, Z can include farm, household, and contextual characteristics such as assets, measures of human capital, income sources, credit access, and other variables capturing farmers' risk attitudes. Specific determinants of transaction costs, such as physical infrastructure and transportation costs, can also affect household marketing preferences and opportunities (Dorward, 2001).

The probability that farmers participate in supermarket channels is therefore determined by a comparison of the expected utility of participation, I_S^*, against the expected utility of supplying traditional markets, I_t^*. In making this comparison farmers evaluate both the benefits and costs of adjustment. Farmers will participate in supermarket channels only if $I_S^* > I_t^*$, implying that the potential benefits outweigh the constraints. However, I_S^* and I_t^* are latent variables; what is observed is actual participation in supermarket channels, I, with $I = 1$ if $I_S^* > I_t^*$ and $I = 0$ if $I_S^* \leqslant I_t^*$. Participation in supermarket channels can therefore be represented as follows:

$$I = Z\alpha - v, \tag{1}$$

where α is a vector of parameters, and v is an error term with zero mean and variance σ^2. Since farmers are heterogeneous in their characteristics, not all of them will participate in supermarket channels. For those who do, participation is expected to result in higher farm returns and household incomes.

(a) *Modeling income effects*

Household income is determined by various socioeconomic factors. For farm households, income is usually influenced by returns from agricultural production, which depend on asset ownership and capacity to produce and market efficiently. Hence, participation in certain market channels may directly influence household income. We hypothesize that supermarket channel participation has an important positive effect on household income, due to market assurance, higher and more stable output prices, and better access to inputs and technologies. In order to evaluate income effects, we build on a model commonly used in the impact assessment literature:

$$y = X\beta + \gamma I + u, \tag{2}$$

where y is household income, X is a vector of farm, household, and contextual characteristics, and I is the participation dummy. Thus, the coefficient γ captures the impact of supermarket participation on household income. However, because farmers self-select into the group of participants, this coefficient may be biased. Especially when more efficient farmers, whose incomes are higher anyway, are more likely to participate in supermarket channels, the income effect would be overestimated. In order to correct for such bias, Heckman selection or instrumental variable approaches could be used. Yet, these approaches still assume that the income functions would differ only by a constant term between participants and nonparticipants. In reality, differences between the groups may be more systematic, that is, there may be interactions between marketing channel choice and the other income determinants captured in X. Maertens and Swinnen (2009) have used propensity score matching, which can deal with structural differences, but only to the extent that these differences are based on observables. When there are unobserved factors that simultaneously influence farmers' marketing decisions and household incomes, such as individual skills, ability or motivation, then propensity score matching may still result in biased estimates.

An approach that can account for systematic differences across groups is switching regression (Maddala, 1983). A switching regression model treats market channels as regime shifters; this can be represented as follows:

$$\begin{aligned} y_s &= X\beta_s + u_s, \\ y_t &= X\beta_t + u_t, \\ I^* &= Z\alpha - v, \end{aligned} \tag{3}$$

where y_s and y_t represent household income for supermarket and traditional channel suppliers, respectively, and I^* is a latent variable determining which regime applies. β_s and β_t are sets of parameters to be estimated. While the variable sets X and Z are allowed to overlap, proper identification requires that at least one variable in Z does not appear in X. Note that in a cross-section sample y_s and y_t are only partially observed: y_s is only observed for the subsample of supermarket suppliers, and y_t for the subsample of farmers supplying traditional channels. What is totally observed is a single variable y_i defined as follows:

$$y_i = \begin{cases} y_s & \text{if } I^* > 0 \\ y_t & \text{if } I^* \leqslant 0 \end{cases} \quad \text{and } I = \begin{cases} 1 & \text{if } I^* > 0 \\ 0 & \text{if } I^* \leqslant 0 \end{cases}. \tag{4}$$

In Equation. (3), u_s, u_t, and v are residuals that are only contemporaneously correlated; they are assumed to be jointly normally distributed with a mean vector 0, and covariance matrix as follows:

$$\sum = \begin{pmatrix} \sigma_s^2 & \sigma_{st} & \sigma_{sv} \\ \sigma_{st} & \sigma_t^2 & \sigma_{tv} \\ \sigma_{sv} & \sigma_{tv} & \sigma^2 \end{pmatrix}, \tag{5}$$

where $var(u_s) = \sigma_s^2, var(u_t) = \sigma_t^2, var(v) = \sigma^2, cov(u_s, u_t) = \sigma_{st}$, $cov(u_s, v) = \sigma_{sv}$, and $cov(u_s, v) = \sigma_{sv}$. The variance of v is set to one, since α is estimable only up to a scale factor (Greene, 2008; Maddala, 1986). In addition, $\sigma_{st} = 0$, since y_s and y_t are never observed together.

The switching model outlined so far accounts for observed systematic differences between farmers in the two market channels. When there are unobserved factors that matter, there will be correlation between the error terms of the regime equations and the selection equation. Estimates of the covariance terms can therefore provide a test for endogeneity. If $\sigma_{sv} = \sigma_{tv} = 0$, there is exogenous switching, but if either σ_{sv} or σ_{tv} is non-zero, then we have a model with endogenous