African Journal of
Agricultural and Resource Economics
Volume 19, Number 3 (2024), pp 228–245

Effect of collective marketing on mango income among smallholder farmers in Mwala sub-county, Machakos County, Kenya

Veronica Wavinya Maingi*
Department of Agricultural Economics and Agribusiness Management, Egerton University, Egerton, Kenya. E-mail: veronicamaingi12@gmail.com

Magret Ngigi
Department of Agricultural Economics and Agribusiness Management, Egerton University, Egerton, Kenya. E-mail: mngigi@egerton.ac.ke

Raphael Gitau
Department of Agricultural Economics and Agribusiness Management, Egerton University, Egerton, Kenya. E-mail: raphael.gitau@egerton.ac.ke

* Corresponding author

Received: June 2024
Published: September 2024

DOI: https://doi.org/10.53936/afjare.2024.19(3).13

Abstract

The mango subsector is a major source of income for farmers in Kenya. However, due to marketing imperfections, smallholder mango farmers do may not be receiving a fair return on their income. This study examined the effect of collective marketing on mango income for 226 smallholder farmers in Mwala sub-county. The study employed an endogenous switching regression model to account for selection bias from observed and unobserved farmer attributes. The results reveal that participation in collective marketing significantly boosts mango income. In the counterfactual cases, participants in collective marketing would have realised USD 68.24 less income from mango if they had not participated. In contrast, non-participants would have earned about USD 167.63 more income if they had participated. The econometric results show that the gender, number of mango trees, farm size and access to market information significantly influence mango incomes. The study recommends encouraging more farmers to participate in collective marketing to increase their incomes from mango production

Key words: collective marketing organisations, endogenous switching regression, smallholder farmers, income, mango

1. Introduction

Mango (*Mangifera indicia* L.) is one of the most significant fruits grown in tropical and subtropical regions and is cultivated in over 100 countries. Globally, mango production reached approximately 59 million metric tons in 2022, with Africa contributing around eight million metric tons (FAOSTAT 2022). This growth reflects the rising demand for mangoes in both fresh and processed forms across international markets. Mangoes are highly valued for their nutritional benefits, being rich in vitamins A and C, dietary fibre and other essential minerals, making them a critical component of food security and a reliable income source for millions of farmers (Onyango *et al*. 2023). In Kenya, mango is the second most cultivated fruit after bananas, with production volumes increasing by 34.46%, from 564 009.6 metric tons in 2017 to 758 372.2 metric tons in 2022, driven by growing domestic and international demand (FAOSTAT 2022). The sector relies predominantly on smallholder farmers, who contribute 80% of the country's production, underscoring their critical role in enhancing rural livelihoods, reducing poverty and driving economic development (Onyango *et al*. 2023). The Eastern and Coastal regions dominate mango production, contributing 90% of the national output, with Makueni and Machakos counties emerging as key production areas (Wangithi *et al*. 2021; Githumbi 2022).

Despite these positive trends, smallholder farmers face numerous challenges that limit their profitability. These include high transaction costs driven by poor road networks, long distances to markets and expensive transportation, all of which reduce profit margins (Mutonyi 2019; Bien & Soehn 2022). Inadequate market infrastructure, including insufficient cold storage facilities and market structures, further constrains farmers' ability to access lucrative markets (Kiet *et al*. 2024). In addition, unreliable market information due to weak communication systems and inadequate extension services forces farmers to rely on intermediaries, who often exploit their position to offer unfavourable terms (Yankson *et al*. 2016; Bien & Soehn 2022). These factors expose farmers to low farm-gate prices, significant post-harvest losses of up to 40%, and income disparities even among those selling in the same markets (Bien & Soehn 2022; Kiet *et al*. 2024).

Collective marketing organisations (CMOs) have been promoted to address these challenges. By pooling resources and acting as a single seller, collective marketing organisations enable farmers to negotiate better prices, access larger markets, and benefit from economies of scale (Ma & Abdulai 2016; Mina *et al*. 2020). While studies have demonstrated the benefits of collective marketing for crops like bananas and avocados in Kenya (Fischer & Qaim 2012; Kwizerimana *et al*. 2023), evidence in the mango sector remains limited, with few studies focusing on the income effects of collective marketing organisations specifically in Mwala sub-county.

In the Mwala sub-county, mango farming is a vital source of income and sustenance for smallholder farmers, significantly contributing to their livelihoods. Farmers in this area market their produce through various channels, including direct sales, intermediaries and collective marketing organisations, each offering different returns (Mutonyi 2019). Despite the potential of collective marketing organisations to enhance incomes by reducing reliance on intermediaries, empirical evidence on their effectiveness remains sparse, particularly in the Mwala sub-county. In addition, there is a lack of understanding regarding the income disparities between farmers participating in collective marketing organisations and those marketing individually.

This study sought to fill this gap by examining the effect of collective marketing organisations on the income of smallholder mango farmers in Mwala sub-county. By focusing on income, a critical measure of farmers' well-being, this research provides valuable insights into the role of collective marketing organisations in improving smallholder livelihoods. The findings will inform policymakers and development practitioners seeking to design and implement strategies that enhance income generation and sustainability for smallholder mango farmers. By employing endogenous switching regression (ESR), this study accounts for unobserved factors influencing participation in collective marketing and the resulting income outcomes. This robust approach aims to provide more precise insights into the impact of collective marketing on mango farmers, filling important gaps in crop-specific research and improving the precision of the analysis. The study tested the following hypothesis: no significant relationship exists between participation in collective marketing organisations and income from mango sales.

2. Study methods

2.1 Study area

The study was conducted in Mwala sub-county in Machakos County, Kenya. The sub-county is the third largest sub-county in Machakos County, covering an area of approximately 1 071.9 km^2, with a population of 181 896 (Kenya Bureau of Statistics [KNBS] 2019). The sub-county is classified within the lower midland zones (LM3, LM4 and LM5) and experiences a semi-arid climate, characterised by annual rainfall ranging from 500 mm to 1 300 mm and mean temperatures between 18°C and 25°C (Kenya Ministry of Agriculture, Livestock and Fisheries [MoALF] 2018; Karuma *et al*. 2020). These climatic conditions provide an ideal environment for mango cultivation, as the crop thrives in areas with rainfall between 500 mm and 1 000 mm and temperatures within the optimal range of 20°C to 26°C (Griesbach 2003). The sub-county's soils, primarily loam and black cotton types, are deep and well-drained, aligning with the recommended soil conditions for mango farming (Griesbach 2003). Combined with the adaptability of mango trees to drought conditions once they are established, these agroecological factors make Mwala sub-county a prime area for mango production. Mwala is the leading mango-producing area in Machakos County, contributing significantly to the overall mango output (MoALF 2018). Mango farming has become a cornerstone of economic resilience for smallholder farmers, offering a reliable income source and a viable alternative to traditional drought-resistant crops such as sorghum, millet, pigeon peas, cowpeas and green grams (MoALF 2018; Karuma *et al*. 2020). The economic contribution of mango farming is substantial, as it generates approximately Ksh[1] 2.4 billion annually at the farm gate and accounts for 22% of farm household income in the Eastern region, highlighting its importance in improving livelihoods (Musyoka 2020).

2.2 Sampling procedure

The study targeted smallholder mango farmers, focusing on both participants and non-participants in collective marketing initiatives. While approximately 17 676 households in the sub-county engage in mango farming (Musyoka 2020), the study specifically targeted those with a history of selling surplus mangoes. This criterion ensured that the sample consisted of households actively participating in mango-marketing activities, aligning with the study's objective. According to data from the ward agricultural officers and farmer organisation officials, 677 households – comprising 243 participants and 434 non-participants – met this criterion and formed the basis for the sampling process, as shown in Table 1.

A multi-stage sampling technique was used to obtain the required sample size. The first stage involved purposively selecting Mwala sub-county because of its high mango production compared to other sub-counties in Machakos County. The second stage involved the purposive selection of Masii and Mwala wards among the six wards of Mwala sub-county, as these mango farmers are well organised.

[1] Kenyan shilling

Two farmer groups and two cooperative societies were selected from these wards based on their active engagement in collective marketing initiatives. These included the Mwala Fruit Growers' Society, the Masii Horticultural Cooperative Society, the Mango Farmers Integrated CBO, and the Mwala Farmer Group Society. These are instrumental in supporting farmers by facilitating the collective sale of mangoes and establishing direct links between farmers and buyers. This ensured active participation in the marketing of mangoes during the study period. For the non-participants, one sub-location was randomly selected from each ward for the study, after which, three villages from each sub-location were chosen randomly. In the third stage, respondents were selected using systematic random sampling. The first respondent was randomly selected from a list of smallholder mango farmers provided by the ward agricultural officers and operating farmer organisation officials. Additional respondents were then selected by taking every Kth item from the list. K represented the interval after every two and one farmers for non-participants and participants respectively until the desired sample size was reached (i.e. for participants, every second farmer was chosen (skipping one), and for non-participants, every third farmer was chosen (skipping two each time)).

The required sample size was determined using Yamane's (1967) formula, which is suitable for a known population:

$$n = \frac{N}{1+N(e)^2}, \tag{1}$$

where n is the desired sample size, N is the estimated population of mango farmers, one is a constant, and e is the error term. According to Kothari (2014), an error of less than 10% is acceptable. Therefore, the study used a permissible error of 0.0543 (5.43%). The sample size was then calculated as:

$$n = \frac{677}{1+677(0.0543)^2} = 225.576 \sim 226. \tag{2}$$

The computation resulted in a total sample of 226 respondents. To ensure balanced representation, one non-participant was selected for each participant, leading to a target of 113 participants and 113 non-participants; hence, the targeted sample comprised 113 respondents who were participants and 113 who were non-participants. The distribution was then done using probability proportional to the population of mango households in each ward.

Table 1: Distribution of the sample size

Ward	The population of collective marketing participants	Proportion	Sample size	Estimated population of non-participants	Proportion	Sample size	Total sample size
Masii	140	0.58	65	220	0.51	58	121
Mwala	103	0.42	48	214	0.49	55	105
Total	**243**	**1**	**113**	**434**	**1**	**113**	**226**

Data was collected between May 2023 and June 2023 by trained enumerators through face-to-face interviews using a semi-structured questionnaire. The questionnaire was administered using the KOBO collect tool and pre-loaded on Android smartphones. During this process, the study adhered to the necessary ethical considerations. All respondents were informed about the objectives and procedures of the study. Respondents were also informed of their voluntary participation and were reassured of their anonymity and confidentiality. This comprehensive approach ensured the integrity and validity of the research process. The variables of the study, along with their units of measurement, are defined and captured in Table 2. The choice of variables was guided mainly by previous studies on collective action in smallholder farming, the researcher's knowledge, and intuition.

Table 2: Description of variables used in empirical models and prior expectations

Variable	Description	Measurement	Expected sign
Dependent variables			
CMP	Collective marketing participation	1 = yes, 0 = otherwise	+/-
Income	Mango income	Kenya shilling	
Independent variables			
Gender	Sex of household head	1 = male, 0 = otherwise	+
Education	Education level of household head	Years	+
Household size	Household size of the respondent	Number of members in the household	+
Market Information	Access to market information	1 = yes, 0 = otherwise	+
Age	Age of the household head	Years	+/-
Farm experience	Experience in mango production	Years	+
Farm size	Farm size under mango production	Acres	+
Total trees	Total mango trees	Number of trees	+
Market distance	Distance to the nearest input/output market	Kilometres	+
Extension	Frequency of extension services	Number of contacts with extension agents	+
Neighbour membership	Respondent's neighbours in CM	1 = yes, 0 = otherwise	+
Extension distance	Distance to the extension service	Kilometres	+

2.3 Empirical framework

The endogenous switching regression method was employed to analyse the effect of participation in collective marketing on mango income. In non-experimental research, various econometric approaches, such as the endogenous switching regression method (ESR), two-stage least squares and propensity score matching (PSM), can be used. While two-stage least squares do not fully account for self-selection, which may result in biased estimates, PSM estimates treatment effects by matching participants and non-participants based on observed covariates. However, PSM does not address unobserved heterogeneity, which can still lead to biased results if relevant unmeasured factors influence the likelihood of participation. Despite this limitation, PSM is still widely used, as it can provide consistent estimates when the assumption of no measured confounding holds (Rosenbaum & Rubin 1983).

Conversely, the ESR model is specifically designed to address endogeneity and sample selection bias, allowing for the simultaneous estimation of both the selection and outcome equations. By using the full information maximum likelihood (FIML) method, ESR accounts for both observed and unobserved heterogeneities, providing more robust estimates of the treatment effects of participation in collective marketing (Maddala 1983; Lokshin & Sajaia 2004; Wossen *et al*. 2017). However, ESR also has limitations, such as the assumption that the error terms in both the selection and outcome equations follow a trivariate normal distribution with a covariance matrix and a mean vector of zero. This assumption may not always hold in practice, potentially limiting the model's robustness in specific contexts. Given the nature of the current study, the ESR model was more appropriate because, unlike PSM, which relies solely on observed covariates, the ESR model addresses endogeneity and selection bias by accounting for unobserved factors that may influence participation. By simultaneously estimating the selection and outcome equations, the ESR model provides reliable and unbiased estimates and offers counterfactual insights, making it ideal for capturing the impact of participation in collective marketing.

The study used a two-stage treatment framework to model the effect of participation in collective marketing organisations on mango income. In the first stage, farmers' decision to participate in a collective marketing organisation was modelled and estimated using a probit model. In the second stage, the relationship between the outcome variable and the participation decision and a set of explanatory variables was assessed using the ordinary least squares (OLS) model with selectivity correction. The observed outcome of participation in a collective marketing organisation was modelled following a random utility formulation. Considering the i^{th} farm, the household faces a decision on whether or not to participate in the collective marketing organisation. Let U_j represent the benefits to the farmer from participation in a collective marketing organisation, and let U_m represent the benefit stream from not participating. The farmer will participate in a collective marketing organisation if

$$D_i^* = U_j - U_m > 0. \tag{3}$$

The net benefit, D_i^*, that the farmer derives from participating in a collective marketing organisation is a latent variable determined by observed characteristics (Z_i) and the error term (ε_i):

$$D_i^* = Z_i\, \gamma + \varepsilon_i, \text{ with } D_i = \{1 \; if \; D_i^* > 0, 0 \; otherwise\}, \tag{4}$$

where D_i is a binary indicator variable that equals 1 for household i in the case of participation in a collective marketing organisation, and 0 otherwise; γ denotes a vector of parameters to be estimated; ε_i is the error term that explains the unobserved characteristics of household i; and Z_i is a vector of observable factors such as household- and farm-level factors.

The outcome function conditional on participation can be written as an endogenous switching regime model:

$$\text{Regime 1: } Y_{1i} = \beta_1 X_{1i} + \mu_{1i} \text{ if } D_i = 1, \text{ and} \tag{5}$$

$$\text{Regime 2: } Y_{2i} = \beta_2 X_{2i} + \mu_{2i} \text{ if } D_i = 0, \tag{6}$$

where Y_i denotes the income for both the participant and non-participant regimes; X is a vector of covariates; β is a vector of unknown parameters to be estimated; and μ is the random error term. For the ESR to be identified, the Z variables in the participation equation need to contain a selection instrument in addition to those generated by the non-linearity of the selection model. A valid instrument is required to influence farmers' decisions to participate in collective marketing, but does not affect outcomes. In this study, distance to the market, distance to the extension office and neighbour membership were employed as the identifying instruments, and a simple falsification test was executed to validate the choice of these instruments. Previous studies have shown that a farmer's choice of participation or membership in collective action is positively and significantly influenced by neighbour membership and distance to the nearest output market (Ma *et al*. 2016; Mojo *et al*. 2017).

One of the key assumptions of ESR is that the error terms in the selection equation and the regime equations have a trivariate normal distribution, with zero mean and covariance matrix of the following form:

$$\text{Cov}\ (\varepsilon_i, \mu_1, \mu_2) = \begin{bmatrix} {}^{\delta 2}_{\varepsilon} & \delta_{\varepsilon_1} & \delta_{\varepsilon_2} \\ \delta_{\varepsilon_1} & {}^{\delta 2}_{1} & . \\ \delta_{\varepsilon_2} & . & {}^{\delta 2}_{2} \end{bmatrix}, \tag{7}$$

where ${}^{\delta 2}_{\varepsilon}$ is the variance of the error term in the participation Equation (4), ${}^{\delta 2}_{1}$ and ${}^{\delta 2}_{2}$ are the variances of the error term in the outcome functions (5) and (6), and δ_{ε_1} and δ_{ε_2} represent the covariance of ε_i, μ_1 and μ_2. The covariance between μ_1 and μ_2 is not defined, since Y_{1i} and Y_{2i} are not observed simultaneously. One of the important implications of the error term structure is that, conditional on the sample selection, the expected values of μ_1 and μ_2 are non-zero because the error term of Equation (4) is correlated with the error term of the outcome functions (5) and (6). Following Heckman (1979), the inverse Mills ratio for collective marketing-organisation participation and non-participation is computed as follows:

$$E[u_{1i}|D_i = 1] = \delta_{\varepsilon_1} \frac{\Phi(Z_{i\propto})}{\phi(Z_{i\propto})} \tag{8}$$
$$= \delta_{\varepsilon_1}\lambda_{1i}$$

and

$$E[u_{2i}|D_i = 0] = \delta_{\varepsilon_2} \frac{\Phi(Z_{i\propto})}{1-\phi(Z_{i\propto})}, \tag{9}$$
$$= \delta_{\varepsilon_2}\lambda_{2i}$$

Where $\Phi(.)$ is the standard normal probability density function, $\phi(.)$ is the standard normal cumulative density function, and $\lambda_{1i} = \frac{\Phi(Z_{i\propto})}{\phi(Z_{i\propto})}$ and $\lambda_{2i} = \frac{\Phi(Z_{i\propto})}{1-\phi(Z_{i\propto})}$ are the inverse Mills ratio computed from Equation (4). To correct for selection bias in the ESR model, the inverse Mills ratio for collective-action participants (λ_{1i}) and non-participants (λ_{2i}) and the covariance terms δ_{ε_1} and δ_{ε_2} are then included in the outcome equations, as follows:

$$Y_{1i} = \beta_1 X_{1i} + \delta_{\varepsilon_1}\lambda_{1i} + \varepsilon_{1i} \text{ if } D_i = 1 \text{ and} \tag{10}$$

$$Y_{2i} = \beta_2 X_{2i} + \delta_{\varepsilon_2}\lambda_{2i} + \varepsilon_{2i} \text{ If } D_i = 0, \tag{11}$$

where λ_{1i} and λ_{2i} are the selectivity correction terms used to control for selection bias caused by unobserved attributes; and ε_{1i} and ε_{2i} are the random error terms with conditional zero means. To simultaneously estimate the selection Equation (4) and the outcome equations (5) and (6), the study employed the full information maximum likelihood (FIML), as proposed by Lokshin and Sajaia (2004).

The FIML was used to estimate the treatment and heterogeneity effects of the treatment to obtain the parameters of the endogenous switching regression model (ESR). The ESR model enables the computation of four comparable expected outcomes: the actual expected income/outcomes of farm households that participate in a collective marketing organisation (Equation 12) and those that do not participate (Equation 13), and the income in the counterfactual scenario, that is, the outcomes of participating farmers had they not participated (Equation 14) and that of non-participating farmers had they participated (Equation 15), as follows:

$$E(Y_{i1}|\, D_i = 1) = \beta_1 X_{1i} + \delta_{\varepsilon_1}\lambda_{1i} \tag{12}$$

$$E(Y_{i2}|\, D_i = 0) = \beta_2 X_{2i} + \delta_{\varepsilon_2}\lambda_{2i} \tag{13}$$

$$E(Y_{i2}|\, D_i = 1) = \beta_2 X_{2i} + \delta_{\varepsilon_2}\lambda_{2i} \tag{14}$$

$$E(Y_{i1}|\, D_i = 0) = \beta_1 X_{1i} + \delta_{\varepsilon_1}\lambda_{1i} \tag{15}$$

The treatment effect on the treated (TT) is then calculated as the difference between (12) and (14), while the effect of the treatment on the untreated (TU) is given by the difference between (13) and (15), as shown below:

$$\mathrm{TT} = E(Y_{i1}|\, D_i = 1) - E(Y_{i2}|\, D_i = 1) \tag{16}$$

$$\mathrm{TU} = E(Y_{i2}|\, D_i = 0) - E(Y_{i1}|\, D_i = 0) \tag{17}$$

3. Results and discussion

3.1 Preliminary diagnostics of variables in the econometric analysis

This section presents the results of the preliminary diagnostics conducted on the variables included in the econometric analysis. These diagnostics aimed to identify and address potential statistical issues, specifically multicollinearity and heteroskedasticity. The tests were performed to ensure the reliability and validity of the variables used in subsequent econometric model analysis. The multicollinearity test was done for continuous variables using the variance inflation factor (VIF).

3.1.1 Multicollinearity test

Multicollinearity is when two or more independent variables in a regression model are highly correlated. According to Yang and Wu (2016), this problem arises when two or more predictor variables in the regression model have a perfect relationship. For categorical variables, multicollinearity happens when one category or a combination of categories can be precisely predicted from others. This situation complicates the interpretation of individual variable effects on the dependent variable, resulting in unstable coefficient estimates. This study obtained a mean VIF of 1.22 (Table A1 in the Appendix). A VIF value should be below the standard cut-off threshold of 10, or a more restrictive criterion of less than five (Hair *et al*. 2011). Thus, the results indicate a small degree of multicollinearity among the selected study variables, as indicated by 1/VIF > 0.2 and VIF < 10.

3.1.2 Heteroscedasticity test

Heteroscedasticity occurs whenever the variance of the unobserved factors changes across different population segments. Two tests can detect heteroscedasticity: the White test and the Breusch-Pagan test. However, the White test is preferred over the Breush-Pagan test because it can detect both linear and non-linear forms of heteroscedasticity by considering the magnitude and direction of change (Uyanto 2022). This study tested heteroscedasticity, skewness and kurtosis using the White test, with a null hypothesis that the residuals are homoscedastic. The results indicate a chi-square of 0.47; therefore, we failed to reject the null hypothesis at 95% and conclude that the residuals are homoscedastic, indicating that the variance of the error term is constant across different levels of the independent variables. The skewness and kurtosis tests yield a p-value of 0.00 and 0.08, respectively, suggesting a marginal departure from normality in the distribution of the error term. However, the overall, resulting p-value is not statistically significant at the conventional levels: 0.05, 0.1 and 0.01 levels of significance (Table A2 in the Appendix).

3.2 Demographic and socio-economic characteristics of participants and non-participants

The results in Table 3 show that the average age of the household heads in the study area was 57 years, with those participating in collective marketing organisations being significantly older than non-participants (t-test 4.24, $p < 0.01$). Respondents had an average of 13 years in mango production. Notably, those participating in collective marketing had significantly more farming experience than their counterparts who marketed individually (t-test = 3.71, $p < 0.01$). An explanation for the significant difference could be that older farmers face greater risks and obstacles in the evolving market environment than younger ones, making collective marketing more appealing and advantageous. Regarding market accessibility, the average distance that mango farmers travelled to the nearest output market was approximately 1.7 kilometres. Interestingly, non-participants travelled significantly further than participants (t-test = 4.09, $p < 0.01$).

Table 3: Socio-economic characteristics of household heads by participation in CMOs

	Pooled (n = 226)	Participants (n = 113)	Non-participants (n = 113)	
Variable	**Mean (Std dev.)**	**Mean (Std dev.)**	**Mean (Std dev.)**	**t-test**
Age	56.90 (13.44)	60.55 (12.99)	53.25 (12.92)	-4.24***
Education	9.17 (3.15)	9.35 (3.15)	9.01 (3.16)	-0.76
Experience	12.92 (7.10)	14.63 (7.50)	11.22 (6.25)	-3.71***
Household size	4.76 (1.95)	4.95 (2.06)	4.57 (1.81)	-1.47
Distance to the market	1.73 (1.24)	1.40 (1.16)	2.06 (1.28)	4.09***
Extension	3.00 (1.75)	3.84 (1.07)	2.17 (1.89)	-8.18***
Size under mango	2.03 (2.49)	2.14 (1.49)	1.92 (3.20)	-0.66
Mango trees	30.69 (18.19)	37.35 (17.69)	24.04 (116.21)	-5.90***
Number of p trees	28.35 (15.16)	33.11 (15.15)	23.59 (13.66)	-4.97***
Dummy variables	**Percentage of farmers**			χ^2
Gender	73.66 26.34	72.57 27.43	74.77 25.23	0.21
Access to market info (% yes)	80.97	53.55	35.71	5.35*
Neighbour membership (% yes)	87.60	57.07	42.93	31.95***

Notes: *, ** and *** indicate significance at 10%, 5% and 1% respectively; standard deviations in parentheses.

On average, farmers had three contacts with extension officers in the previous production season. However, those participating in collective marketing organisations had significantly more frequent access to these services, as indicated by a t-test value of 8.18 ($p < 0.01$). Access to extension services enhances farmers' access to improved technologies and important production information that is necessary for their various enterprises (Okello *et al*. 2017). In terms of production, the average number of mango trees in the studied population was approximately 30. The participants had a significantly higher number of mango trees than the non-participants (t-test = 5.90, $p < 0.01$). This trend was also observed in the number of productive mango trees, with participants having significantly more than the non-participants (t-test = 4.96, $p < 0.01$). This could be associated with better access to knowledge on agronomic practices, and the greater sharing of knowledge or even resources among the participants than among their counterparts.

Most households (74%) were male-headed, with 75% of these households being non-participants. Among the interviewed mango farmers, 91.59% were older than 35, indicating low involvement by the youth in mango production and marketing. Regarding institutional factors, access to market information was high (80.97%), with farmers in collective marketing organisations reporting higher access (53.55%) compared to non-participants (35.71%). Access to information on mango markets is important because it improves farmers' awareness of market dynamics and trends, facilitating more informed decisions and helping them better negotiate with buyers. With access to this information, smallholders can weigh the pros and cons of the available market options and ways of approaching them (Mango *et al*. 2017). In addition, the majority of households (87.60%) had neighbours who were members of collective marketing organisations, with a more significant proportion among participants (57.07%) than non-participants (42.93%). This underscores the importance of social networks in influencing participation.

3.2.1 Endogenous switching regression model

The endogenous switching regression model examines how participation in collective marketing influences income from mango. The selection equation was estimated using the probit model in the first stage. Accounting for selection bias was done in the second stage. Income from mango was calculated as the total annual revenue from the sale by households of both mango fruit and mango products, such as firewood and charcoal, measured in Kenya shillings. Mango income and other factors were controlled for in this model. These factors included sex, education level, farming experience, the main occupation of the household head, number of extension contacts, farm size under mango, distance to the market, distance to the nearest extension services, total number of mango trees, neighbours' membership of an agricultural collective, household size and access to market information. Neighbours' membership of an agricultural collective, distance to the nearest extension services, and distance to the nearest market were used as instruments in the model.

3.2.2 Tests of endogeneity and validity, and strength of the instruments

Several tests were carried out to use the model effectively, including endogeneity tests, over-identifying restrictions, and weak instruments (see Appendix tables A3, A4 and A5). Durbin and Wu-Hausman tests were conducted to test for the presence of endogeneity, yielding Durbin (score) chi^2 (1) = 3.88 ($p = 0.04$) and Wu-Hausman F (1 215) = 3.76 ($p = 0.04$). The null hypothesis of exogeneity was rejected ($p < 0.05$), providing evidence of endogeneity in the model. The strength of the instruments was also tested, with a partial R-square of 93%, indicating that the model was fit and that the null hypothesis – namely that the instruments were weak – was rejected (F-statistics < 0.01). Furthermore, the Sargan and Basmann tests were used to test for over-identifying restrictions and, in both tests, the null hypothesis of valid instruments was rejected ($p > 0.1$). This suggests that the instruments used in the model were not statistically different from the actual values of the endogenous variables, and that they were appropriate for addressing endogeneity.

3.2.3 Results of the endogenous switching regression parameter estimates

Table 4 presents the results of the endogenous switching regression model. As indicated previously, the FIML approach estimates the selection and outcome equations jointly. The first column presents the selection equation representing the determinants of participation in collective marketing organisations. The outcome equations that describe the effect of collective marketing on mango income for both the participants and the non-participants are presented in the second and third columns, respectively. The likelihood ratio test of independence of the selection and outcome equations (LR test of independence equations: Chi^2 (2) = 6.14, $p < 0.05$) indicates that there is a correlation between participation in collective marketing organisations and mango income. To test for selectivity bias in the model, the study examined the covariance coefficients of the two regimes.

The results show that the correlation coefficient for the non-participants (/r1) was negative (-0.39) and statistically significant ($p < 0.01$). The negative and significant correlation suggests that there is sample selectivity, hence implying that there was endogeneity. A further implication is that participants differed from non-participants and that there may not be similar effects on them if they were to participate, as supported by Lokshin and Sajaia (2004). The results also show that the coefficients of variance of error terms were positive and significant ($p < 0.01$), indicating that participation in collective marketing organisations contributed to improved mango income. Farm households that participate have more returns, which improves mango income compared to the situation for non-participants. Therefore, these results confirm the appropriateness of the endogenous switching regression model in addressing the unobserved characteristics of the participants and non-participants.

Table 4: Endogenous switching regression model estimates of the selected outcome

Model specification			**FIML endogenous switching regression**			
	Selection equation		**Outcome equation**			
	Participation (1/0)		**Participants = 1**		**Non-participants = 0**	
Variable	**Coef.**	**Std. err.**	**Coef.**	**Std. err.**	**Coef.**	**Std. err.**
Education	0.04	0.04	4.93	6.25	-0.39	2.25
Gender	-0.60**	0.29	47.52	42.75	-19.45**	10.70
Household size	0.11*	0.06	1.02	9.84	0.17	3.77
Age	0.01	0.01	0.09	1.66	-0.07	0.59
Extension	0.25***	0.07	-5.48	17.74	-0.64	4.18
Experience	0.39***	0.07	-13.63	10.61	8.43	5.76
Market Information	0.80**	0.33	117.07**	59.43	37.59**	17.39
Farm size	0.06**	0.03	3.24	4.70	4.42**	1.73
Mango trees	0.02**	0.01	2.66**	53.88	1.55***	0.47
Distance to the market	-0.11***	0.03				
Distance to extension office	0.07**	0.03				
Group membership	0.32**	0.17				
_cons	-3.79***	0.83	-71.55	165.11	-25.74	39.20
/lns0	9.14***	0.05				
/lns1	10.19***	0.05				
/r0	0.01	0.25				
/r1	-0.39**	0.25				
sigma0	66.86	3.51				
sigma1	190.54	10.06				
rho0	0.01	0.25				
rho1	-0.37	0.21				
LR test of independent equation: $Chi^2(2) = 6.14^{**}$ $p = 0.04$						

Note: The income equation was jointly estimated with the equation for participation in collective marketing organisations. Sigma presents the square root of the variance of the error terms. In contrast, /r presents the correlation coefficients of the error terms of the selection and outcome equations, as represented in equations (5) and (6). *, ** and *** indicate significance at 10%, 5% and 1% respectively.

The results from the selection equation (Table 4) reveal several factors influencing smallholder farmers' participation in collective marketing organisations. The sex of the household head was significant at the 5% level, implying that female-headed households were more likely to participate in collective marketing than male-headed ones. These results align with those of Mango *et al*. (2017), who also found higher participation rates among female-headed households. Fischer and Quaim (2012) attribute this to women's greater susceptibility to exploitation by farm-gate traders, making them perceive more benefits from collective bargaining. However, these findings contrast with those of Ma *et al*. (2022), who observed that male farmers were more likely to join banana cooperatives and participate in collective marketing.

Farm size under mango production was significant at the 5% level, implying that farmers with large mango farms were more likely to participate in collective marketing organisations than those with smaller farms. These results are consistent with those of Mojo *et al*. (2017), Ma *et al*. (2019) and Gemechu *et al*. (2024), who state that farmers with larger orchard sizes tend to be more likely to join agricultural cooperatives to obtain the expected benefits. However, Twaya (2018) found no significant impact of land size on participation in farmer-based organisations. Access to extension services also positively influenced participation at the 1% level of significance, consistent with the findings of Mina *et al*. (2020). The distance to the nearest mango market negatively influences participation in a collective marketing organisation, and the results were significant at the 1% level. This implies that farmers located further away from the market are less likely to participate. These results contradicted the expectations of the current authors, namely that an increased distance increases the possibility of the farmer participating in the collective marketing organisation. It also contradicts those of Abdul-Rahaman and Abdulai (2020) and Ahmed and Mesfin (2017), who argue that farmers who live further away from the market are more likely to participate in collective marketing organisations.

In addition, distance to extension services had a positive and significant effect on participation at the 5% level, echoing the findings of Kiprop *et al*. (2020). Neighbour membership in an agricultural group, farming experience in mango production and market information also influenced participation in collective marketing organisations.

The results of the effect of participation in collective marketing organisations on income are presented in columns 4, 5,6 and 7 of Table 4 for the farmers who participated and those who did not participate. Mango income was calculated based on net revenue from household sales of mango fruits between January and April 2023. The estimates revealed that the sex of the household head, total number of mango trees, farm area under mango, and access to market information significantly affected mango income. The coefficient of the sex of the household head was negative and significantly different from zero for non-participants. This suggests that, conditional on collective marketing participation, being female is associated with lower mango income among non-participants. This may be explained by cultural norms and gender roles restricting women's mobility and participation in markets, limiting their direct engagement with buyers and market actors and compelling them to rely on intermediaries who often offer less-favourable terms, leading to low income.

Access to market information exerts a positive and significant effect on mango income for both collective marketing participants and non-participants at the 5% level of significance. This implies that having access to market information boosts mango income. For the participants, the results suggest that collective marketing strategies enhance the value derived from market information, allowing participants to make more informed decisions and potentially secure better prices for their mangoes. Although non-participants experience an increase in mango income, the effect is negligible compared to participants in collective marketing organisations. This implies that, while market information is beneficial, it alone is insufficient to unlock the full potential of market opportunities. Non-participants often lack the collective bargaining power and market reach that participants possess, which limits their ability to negotiate better prices and access more lucrative markets. The total number of mango trees was found to positively and significantly influence mango income among the participants, at a 5% level of significance for participants and a 1% level of significance for non-participants. These results are consistent with those of Zhang *et al*. (2014), who found that the number of apple trees had a positive and significant impact on household income for cooperative members.

3.2.4 Average expected mango income, treatment and heterogeneity effects

Table 5 shows the results of expected mango income, treatment and heterogeneity effects.

Table 5: Mean treatment effects on mango income

Sub-sample	Participation status	Mango income (USD)	Treatment effects (USD)
Participants	Participation (n = 113)	(a) 199.64 (81.08)	TT = 68.24***
	Non-participation (n = 113)	(b) 131.28 (50.79)	
Non-participants	Participation (n = 113)	(c) 250.29 (109.63)	TU = 167.63***
	Non-participation (n = 113)	(d) 82.62 (42.26)	
Heterogeneity effects		BH1 = -50.63	TH = -99.45
		BH2 = 48.75	

Notes: Standard deviations are in brackets in the third column; BH1/2 = base heterogeneity 1/2; TT = treatment on the treated; TU = treatment on the untreated; TH = transactional heterogeneity; *** = significant at the 1% level; the exchange rate is approximately 1 USD = 129.50 Kenya shillings.

Cells (a) and (d) represent the observed income values from mango production for the participants in the collective marketing organisation and the non-participants, respectively. Cell (b) represents the income outcome for participants if they had decided not to participate. In contrast, cell (c) represents the income outcome that non-participants would realise if they had chosen to participate. Based on the results, the observed mango incomes of participants and non-participants are USD 199.64 and USD 82.62, respectively. A comparison of the observed outcomes for the participants and the non-participants would mean that participants would realise USD 116.99 more income from mango production than non-participants. However, this simple comparison could be misleading, because these groups have unobserved heterogeneous characteristics.

The last column of Table 5 presents the treatment effects of participation in collective marketing organisations on mango income. In the counterfactual case (b), if mango farming households that participated had not participated, they would have realised about USD 131.28, which is USD 68.24 less. On the other hand, in case (c), non-participants would earn USD 250.29 if they participated in collective marketing organisations, resulting in USD 167.63 more than their actual income. The results imply that participation in collective marketing organisations improves mango income among smallholder mango farmers. This is because farmers can reduce transaction costs associated with marketing by pooling their resources. In addition, economies of scale enable them to negotiate and secure better prices with larger buyers like processing companies, thereby reducing post-harvest losses compared to those marketing individually. These results are consistent with evidence from Fischer and Qaim (2012), Mutonyi (2019), Kwizerimana *et al*. (2023) and Olumeh and Mithöfer (2024), namely that participation in collective action increased income among banana, avocado, mango and baobab farmers and collectors, respectively.

In addition, the last row of Table 5, which adjusts for the potential heterogeneity in the sample, shows that participating farming households would still earn USD 48.75 more income from mango than non-participants in the counterfactual. However, if non-participants had decided to participate, they would have earned USD 50.63 more in mango income than the participants, indicating that non-participants have greater potential for income gains. The results indicate that, for each participation status, the counterfactuals are higher than the actual incomes for the two groups. Lastly, the negative transitional heterogeneity effect of USD -99.45 suggests that non-participants stand to gain more from participation than current participants, and would benefit significantly from collective action.

3.2.5 Contribution of mango revenue to total household income

The study further examined how mango revenue contributes to overall household income. This analysis is essential for assessing the economic impact of mango farming on smallholder farmers. Table 6 provides a detailed breakdown of the percentage contribution from various income sources to total household income, comparing participants and non-participants in collective marketing organisations.

Table 6: Contribution of various sources of revenue to total household income

	Non-participants	Participants	Pooled mean
Mango income (%)	59.52	52.67	56.06
Other agricultural income (%)	6.14	6.19	6.16
Salaried income (%)	14.86	16.92	15.89
Casual labour income (%)	7.68	2.90	5.25
Pension and rental income (%)	1.05	6.47	3.81
Remittances (%)	4.02	5.72	4.88
Business income (%)	6.73	9.13	7.95

Note: Percentages represent the mean contribution of each income source to total household income.

The results reveal that mango revenue is a significant source of household income, contributing 56.06% on average across the entire sample. Notably, mango income constitutes a higher proportion of total household income for non-participants (59.52%) compared to participants (52.67%). Further, the results reveal that participants exhibit greater income diversification, with higher contributions from salaried income (16.92%) and other business ventures (9.13%) compared to non-participants. Non-participants, in contrast, show higher reliance in terms of casual labour income (7.68%), indicative of less stable income sources. These findings underscore the vital role of mango farming in rural household economies in Mwala sub-county, particularly for non-participants. However, they also highlight the potential benefits of promoting income diversification and strengthening the mango value chain to enhance income stability and resilience among smallholder farmers.

4. Conclusions and recommendations

This paper has investigated the factors influencing farmers' participation in collective marketing organisations and the effect of participation on mango income using the ESR method. In doing so, the paper provides both methodological and empirical contributions. Methodologically, the paper goes beyond the case studies on the effect of collective action on income that are prevalent in the literature. It provides a rigorous econometric analysis that accounts for endogeneity, thereby providing an unbiased causal impact of collective marketing organisations on income. The findings demonstrate the significant potential of collective marketing to enhance income for smallholder mango farmers. The results indicate that participating households earn significantly higher income than non-participants, even after accounting for confounding factors. Moreover, non-participating households would have significantly increased their income had they participated in collective marketing. These results highlight the transformative role of collective marketing in boosting income and reducing poverty among smallholder farmers. Expanding participation in collective marketing is therefore essential for improving livelihoods and fostering sustainable agricultural development. To achieve this, policies should focus on building farmers' capacity to understand the benefits of collective marketing and support them in overcoming market barriers. This can be facilitated by designing policies that facilitate easy access to market information. Government, in collaboration with agricultural organisations and ICT departments, can establish platforms that provide real-time market data. This will enable farmers to make informed decisions about when and where to sell their produce, thereby improving their bargaining power.

Among the limitations of the study is that it finds that participation in collective marketing organisations positively affects mango producers' income. However, using a binary variable (1 for participants, 0 for non-participants) does not account for variations in participation levels. This limits the study's ability to explore how different degrees of involvement affect outcomes. Future research should examine different levels of participation to optimise the benefits of collective marketing.

Acknowledgements

The authors would like to thank the team of enumerators involved in collecting the data. We also thank all the respondents for taking the time to participate in this study. This study was funded by the Africa Economic Research Consortium (AERC) through a scholarship to the corresponding author.

References

Abdul-Rahaman A & Abdulai A, 2020. Farmer groups, collective marketing and smallholder farm performance in rural Ghana. Journal of Agribusiness in Developing and Emerging Economies 10(5): 511–27. https://doi.org/10.1108/JADEE-07-2019-0095

Ahmed MH & Mesfin HM, 2017. The impact of agricultural cooperatives membership on the well-being of smallholder farmers: Empirical evidence from eastern Ethiopia. Agricultural and Food Economics 5: 6. https://doi.org/10.1186/s40100-017-0075-z

Bien J & Soehn I, 2022. Unlocking the Kenyan mango value chain. Master's policy analysis paper, John F Kennedy School of Government, Harvard University, Cambridge MA, United States. https://www.hks.harvard.edu/sites/default/files/degree%20programs/MPAID/files/Bien%2C%20James%20%26%20Isabella%20Soehn_SYPA.pdf

FAOSTAT, 2022. Food and agriculture data. Food and Agriculture Organization of the United Nations. https://www.fao.org/faostat/

Fischer E & Qaim M, 2012. Linking smallholders to markets: Determinants and impacts of farmer collective action in Kenya. World Development 40(6): 1255–68.

Gemechu A, Jaleta M, Zemedu L & Beyene F, 2024. Impact of membership of seed-producer cooperatives on commercialisation among smallholder farmers in the central highlands of Ethiopia. African Journal of Agricultural and Resource Economics 19(1): 54–84. https://doi.org/10.53936/afjare.2024.19(1).4

Githumbi R, 2022. Assessment of drivers of postharvest losses and factors influencing adoption of loss reduction practices along the mango value chain in Embu, Machakos and Nairobi counties, Kenya. Doctoral dissertation, University of Nairobi, Nairobi, Kenya.

Griesbach J, 2003. Mango growing in Kenya. Nairobi, Kenya: World Agroforestry Centre (ICRAF).

Hair JF Jr, Hult GTM, Ringle CM & Sarstedt M, 2017. A primer on partial least squares structural equation modeling (PLS-SEM). Second edition. New York NY: Sage.

Heckman JJ, 1979. Sample selection bias as a specification error. Econometrica 47(1): 153–61.

Karuma AN, Gicheru PT & Gachene CKK, 2020. Financial returns of maize and bean production under selected tillage practices in semi-arid area of Mwala sub county, Kenya. Asian Journal of Agricultural Extension, Economics & Sociology 38(10): 11–23.

Kenya Ministry of Agriculture, Livestock and Fisheries (MoALF), 2018. Climate risk profile for Machakos County. Kenya County Climate Risk Profile Series. Nairobi: MoALF. https://hdl.handle.net/10568/96283

Kenya National Bureau of Statistics (KNBS), 2019. 2019 Kenya Population and Housing Census: Volume II – Distribution of population by administrative units. https://housingfinanceafrica.org/app/uploads/VOLUME-II-KPHC-2019.pdf

Kiet HVTT, Sang MV, Anh VTN, Truong TTK, Le TXC, Dao PTB & Nguyen LTK, 2024. Identifying opportunities and challenges of horticulture production: A case study of Tuong-Mango value chain in the southern Vietnam. Journal of Infrastructure, Policy and Development 8(3): 3078. https://doi.org/10.24294/jipd.v8i3.3078

Kiprop E, Okinda C, Wamuyu S & Geng X, 2020. Factors influencing smallholder farmers participation in collective marketing and the extent of participation in improved indigenous chicken markets in Baringo, Kenya. *Asian Journal of Agricultural Extension, Economics & Sociology* 37(4): 1–12.

Kothari CR, 2004. Research methodology: Methods and techniques. Second revised edition. New Delhi: New Age International Publishers.

Kwizerimana S, Mugwe J & Nigat B, 2023. Impact of collective marketing participation on farmers' income: Evidence from smallholder avocado farmers of Murang'a County, Kenya. Social Sciences & Humanities Open 8(1): 100614.

Lokshin M & Sajaia Z, 2004. Maximum likelihood estimation of endogenous switching regression models. The Stata Journal: Promoting Communications on Statistics and Stata 4(3): 282–9.

Ma W & Abdulai A, 2016. Does cooperative membership improve household welfare? Evidence from apple farmers in China. Food Policy 58: 94–102.

Ma W, Abdulai A & Goetz R, 2018. Agricultural cooperatives and investment in organic soil amendments and chemical fertiliser in China. American Journal of Agricultural Economics 100(2): 502–20.

Ma W, Zheng H, Zhu Y & Qi J, 2022. Effects of cooperative membership on the financial performance of banana farmers in China: A heterogeneous analysis. Annals of Public and Cooperative Economics 93(1): 5–27. https://doi.org/10.1111/apce.12326

Maddala GS, 1983. Methods of estimation for models of market with bounded price variation. International Economic Review 24(2): 361–78.

Mango N, Makate C, Lundy M, Siziba S, Nyikahadzoi K & Fatunbi AO, 2017. Collective market participation for improved income among smallholder farming households: A case of Balaka Innovation Platform in Malawi. African Crop Science Journal 25(1): 97–108.

Mina CS, Jimenez CD & Catelo SP, 2020. Assessing the impact of collective marketing on farm income: The case of Calamansi Farmers in Oriental Mindoro, Philippines. Philippine Journal of Social Sciences and Humanities 25: 45–57.

Mojo D, Fischer C & Degefa T, 2017. The determinants and economic impacts of membership in coffee farmer cooperatives: recent evidence from rural Ethiopia. Journal of Rural Studies 50: 84–94. https://doi.org/10.1016/j.jrurstud.2016.12.010

Musyoka JK, Isaboke HN & Ndirangu SN, 2020. Farm-level value addition among small-scale mango farmers in Machakos County, Kenya. Journal of Agricultural Extension 24(3): 85–97.

Mutonyi S, 2019. The effect of collective action on smallholder income and asset holdings in Kenya. World Development Perspectives 14: 100099.

Okello DO, Feleke S, Gathungu E, Owuor G & Ayuya OI, 2020. Effect of ICT tools attributes in accessing technical, market and financial information among youth dairy agripreneurs in Tanzania. Cogent Food & Agriculture, 6(1): 1817287.

Olumeh DE & Mithöfer D, 2024. Impact of collective action on household welfare: Empirical evidence from baobab collectors in Malawi. Annals of Public and Cooperative Economics 95(2): 385–411. https://doi.org/10.1111/apce.12448

Onyango K, Bolo P, Ndiwa A, Wanyama R & Chege CG, 2023. A rapid agroecological mango value chain analysis in Kenya. CGIAR Initiative on Agroecology. https://hdl.handle.net/10568/138772

Rosenbaum PR & Rubin DB, 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70(1): 41–55.

Twaya GW 2018. Effect of participation in farmer-based organisations on the profitability of pigeon pea (*Cajanus cajan*) enterprise in Mulanje district, Malawi. Doctoral dissertation, Egerton University, Njoro, Kenya.

Uyanto SS, 2022. Monte Carlo power comparison of seven most commonly used heteroscedasticity tests. Communications in Statistics – Simulation and Computation 51(4): 2065–82. https://doi.org/10.1080/03610918.2019.1692031

Wangithi CM, Muriithi BW & Belmin R, 2021. Adoption and dis-adoption of sustainable agriculture: A case of farmers' innovations and integrated fruit fly management in Kenya. Agriculture 11(4): 338. https://doi.org/10.3390/agriculture11040338

Wossen T, Abdoulaye T, Alene A, Haile MG, Feleke S, Olanrewaju A & Manyong V, 2017. Impacts of extension access and cooperative membership on technology adoption and household welfare. Journal of Rural Studies 54: 223–33.

Yamane, T. (1967). Statistics: An introductory analysis. Second edition. New York: Harper and Row.

Yang TM & Wu YJ, 2016. Examining the socio-technical determinants influencing government agencies' open data publication: A study in Taiwan. Government Information Quarterly 33(3): 378–92. https://doi.org/10.1016/j.giq.2016.05.003

Yankson PWK, Owusu AB & Frimpong S, 2016. Challenges and strategies for improving the agricultural marketing environment in developing countries: Evidence from Ghana. Journal of Agricultural & Food Information 17(1): 49–61.

Zhang M, Kagatsume M & Yu J, 2014. Market channel choice and its impact on farm household income: A case study of 243 apple farmers in Shaanxi province, China. Japan Agricultural Research Quarterly 48(4): 433–41. https://doi.org/10.6090/jarq.48.433

Appendix

Table A1: Results for the multicollinearity test using VIF

Variable	VIF	1/VIF
Age of household head (years)	1.43	0.70
Extension service frequency (number of contacts)	1.21	0.82
Farming experience (years)	1.38	0.73
Education level of household head (years in school)	1.20	0.84
Distance to the market (in kilometres)	1.04	0.96
Total number of mango trees	1.33	0.75
Household size (number of household members)	1.10	0.91
Total land size under mango farming (in acres)	1.21	0.83
Distance to the extension services (in kilometres)	1.12	0.89
Mean VIF	**1.22**	

Table A2: Results of tests for heteroscedasticity, skewness and kurtosis residuals

	Chi^2	df	p
Heteroscedasticity	224.00	223	0.47
Skewness	51.66	27	0.00
Kurtosis	2.98	1	0.08
Total	278.64	251	0.11
$chi^2(223) = 224.00$			
Prob > $Chi^2 = 0.47$			

Note: Chi^2 = Chi-square; df = degrees of freedom; p value = level of significance

Table A3: Tests of endogeneity

Tests of endogeneity
H0: Variables are exogenous
Durbin (score) $Chi^2(1) = 3.88$ (P = 0.04)
Wu-Hausman F (1 215) = 3.76 (P = 0.04)

Table A4: Testing for weak instruments

Variable	R-squared	Adjusted R-squared	Partial R-squared	F (2 215)	Prob > F
Membership ~p	0.94	0.94	0.93	1 492.25	0.00
Minimum eigenvalue statistic = 1 492.25					
Critical values H0: Instruments are weak	# of endogenous regressor: 1 # of excluded instruments: 2				

2SLS relative bias	5%	10%	20%	30%
	(not available)			
	10%	15%	20%	25%
2SLS size of nominal 5% Wald test	19.93	11.59	8.75	7.25
LIML size of nominal 5% Wald test	8.68	5.33	4.42	3.92

Table A5: Testing the validity of the instruments

Tests of overidentifying restrictions
Sargan (score) $Chi^2(1) = 1.66$ (p = 0.20)
Basmann $Chi^2(1) = 1.59$ (p = 0.21)