Heterogeneous Income Effects By Household Type in Rural Southern Ethiopia: A Quantile Regression Analysis

by Shiferaw Feleke, Adane Tufa, Steven M. Cole, Julius Manda, Tahirou Abdoulaye, Tesfamicheal Wossen, Arega Alene, and Victor Manyong

HETEROGENOUS INCOME EFFECTS BY HOUSEHOLD TYPE IN RURAL SOUTHERN ETHIOPIA: A QUANTILE REGRESSION ANALYSIS

Shiferaw Feleke[1], Adane Tufa[2], Steven M. Cole[1], Julius Manda[3], Tahirou Abdoulaye[4], Tesfamicheal Wossen[5], Arega Alene[2], and Victor Manyong[1]

Abstract
This paper applies a quantile regression model to explore household type−income relationships at different parts of the income spectrum *in rural southern Ethiopia. Data came from a representative sample of 154 households selected using a three-stage cluster sampling technique. Results show that income gaps between female-headed households (FHHs) and male-headed households (MHHs) are explained by differentials in access to extension access, with income effects being greater for the FHHs than for the MHHs. The plausible explanation is that FHHs may have received better extension services. Although both household types have access to extension services, differences in quality and intensity of the extension delivery could result in differential returns to extension access, leading to an income gap. Supervision of the quality and intensity of extension services is thus warranted to ensure equitable income distribution. Further, results have revealed non-uniform income effects across quantiles, suggesting that some households are more or less affected than others, depending on their relative income status in the population. These non-uniform income effects present important evidence for policymakers to avoid the 'one-size fits all' approach and tailor policy interventions to particular income groups.*

Keywords: *Income gap; Household type; Income; Inequality; Ethiopia*

Contributed Paper Selected for Presentation at the 31st ICAE, August 17-31, 2021 Online

1 International Institute of Tropical Agriculture (IITA), Dar es Salaam, Tanzania
2 IITA, Lilongwe, Malawi
3 IITA, Arusha, Tanzania
4 IITA, Bamako, Mali
5 IITA, Niarobi, Kenya

1. INTRODUCTION

Agriculture holds considerable potential to spark economic growth and reduce poverty in sub-Saharan Africa (SSA) [1-2]. In its world development report, the World Bank noted that growth in agriculture is, on average, at least twice as effective in reducing poverty as growth outside agriculture [3]. However, the poverty reduction effects of growth could be stifled if the achieved growth is accompanied by income inequality [4]. Growth can lead to income inequality without reducing poverty incidence when, for example, high-income groups increase their income even more. An increase in income inequality among society members is concerning, especially between women and men. Gender income inequality has far-reaching consequences for any society [5].

The literature on gender in developing countries can broadly be grouped into two strands (differences between men and women and differences between male- and female-headed households (MHHs and FHHs) [6]. The first strand recognizes intra-household differences, while the second strand recognizes inter-household differences.

Most empirical studies infer that FHHs are in a disadvantaged position compared to MHHs, citing a range of constraints associated with access to land, labour, education, credit, extension, input, and output markets [6-9]. Given their disadvantages in access to resources and services, FHHs have been associated with poverty. However, a growing number of studies have challenged the 'feminization of poverty' concept, citing several caveats such as the failure to consider the context (e.g., country), differences between different types of FHHs, the choice and use of equivalence scales in income transformation, the lack of consideration of economies of scale, and non-comparable or inconsistent measures of living standards [10]. Another caveat is using the classical linear regression model to establish the household type−income relationship. This technique, although helpful, does not reflect the income effects in all parts of the income distribution. It only estimates the mean effect, ignoring the heterogeneity across households. In economies where the income gap is widening, using the classical regression model could underestimate the impact on poor households. To better understand the heterogeneous income effects of household characteristics by household type, we opt to use quantile regression[6]. Unlike the classical linear regression, quantile regression allows parameter variation across all income quantiles, thus providing a more comprehensive picture of the income effects. Estimating the effects of household type and its interactions with other variables on household income distribution is still an active research area [11-12]. This paper contributes to the existing literature by assessing the household type-income relationships and identifying the sources and determinants of income inequality at different parts of the income distribution. The quantile regression is applied to cross-sectional data from a formal survey conducted in the Soddo-Zuria District of Southern Ethiopia.

The following section presents a conceptual framework, highlighting the hypothesized relationships between selected variables and income. Section three describes analytical approaches applied to measure and decompose overall income inequality by income source and identify the determinants of the income distribution. Section four describes the data, survey design, and measurement of the dependent and independent variables. The fifth section presents and discusses the significant findings of the study. The final section concludes the paper by summarizing the key findings and drawing policy implications thereof.

[6] Compared with classical linear regressions, quantile models are often less biased to skewed data, better fit to non-normal distributed data, yield more robust estimates, more robust to outliers and require much weaker assumptions for the distribution of the error term [26].

2. CONCEPTUAL FRAMEWORK

Household income (hereafter referred to as income) in the study area comes from on-farm (crop production and livestock production) and off-farm sources (off-farm employment). Crop production is the primary source of income (75%) coming from the sale of different crop outputs (maize, *tef*, potato, taro, wheat, barley, carrot, and red roots) produced in two cropping seasons – the *Mehir* (long rainy season) and *Belg* (short rain season). Livestock income comes from selling different animals (cattle, goats, sheep, and chicken) and animal products (milk, butter, and cheese) sales. The off-farm income sources include off-farm employment in wages and self-employment, payments from food-for-work programs, and other safety-net programs.

Three groups of factors − demographic, socioeconomic, and institutional − are hypothesized to influence household income distribution among rural communities in southern Ethiopia.

One of the factors hypothesized to affect income distribution in the study area is the age of the household head. The relationship between age of household head and income could be an inverted-U. With increasing age comes more experiences, leading to efficient resource management. However, older households beyond a certain age may not be enterprising and willing to take a risk to adopt new technologies. So, the income effect of the household head could be positive during their prime years of age and turn negative after the prime years.

The income effect of the education level of the household head is expected to be positive because education generally enhances one's ability to seek, gather, process new information and adopt new technologies, thereby contributing to higher productivity and income [13-15]. Besides, literate farmers tend to be market-oriented and have better access to product and labour markets and technical information, contributing to higher productivity and income. However, the educational level might be distributed unevenly between male household heads and female household heads, with the former tending to have higher education levels. So, the effect of education on income might be associated with household type.

As in the case with education, land size is expected to exhibit a positive relationship with income. Past studies indicate that access to land is positively associated with higher incomes [16]. The larger the land, the less risk-averse the farmer might be and the higher propensity to adopt new technologies and generate more income [17]. However, land size might also be distributed unevenly between MHHs and FHHs, with the former owning a larger share, suggesting that the effect of land size could be associated with household type.

The effect of labour is also expected to be positive in that high labour availability increases the propensity to adopt labour-intensive technologies, enhancing the capacity to generate more income. However, labour availability in the household might also be distributed unevenly between MHHs and FHHs, suggesting that the effect of labour could be associated with household type.

As in the case with labour, we expect a positive income effect for off-farm employment through its direct income contribution. Although off-farm employment may take some labour away from farm operations, it may not negatively impact agricultural production because the land is scarce in the study area. In farmland-constrained areas, the contribution of off-farm employment to income can be significant. The off-farm income can also enhance the farm household’s financial capacity to purchase farm seeds, thereby boosting productivity and income.

The institutional characteristics (e.g., access to seeds, access to extension, and access to credit) are expected to affect income positively. Access to seeds increases income directly through increased crop productivity. However, FHHs may have less access to production technology than MHHs [e.g., 18-20]. Likewise, access to extension increases income through increased productivity [21]. However, FHHs in developing countries tend to have less access to extension services [e.g., 9] due to their disadvantages of relatively low levels of education, smaller farms, and the fact that extension workers often are men [22]. Farmers who get access to extension services are more likely to adopt improved technologies and apply agronomic practices, leading to increased income [13-15]. In the study district, farmers get access to extension services through contacts with development agents and participation in training, field days, on-farm research trials, and demonstrations.

Smallholder farmers face limited access to formal credit markets [19]. They cannot purchase seeds in the absence of access to credit. Since farmers with access to credit can buy seeds more readily than those with no credit access, we expect access to credit to affect income positively.

Given FHHs' disadvantages with access to these institutional services, income effects associated with these services could be correlated with household type. FHHs' disadvantages with the above resources and services constitute the source of income gap by household type. Whether there is an income gap between FHHs and MHHs due to one or more of the household characteristics presented above is an empirical question. Further, whether the hypothesized relationships stipulated above hold at different parts (lower, median, or upper tail) of the income distribution is another empirical question. For example, education is expected to affect income positively. But, does it affect all parts of the income distribution (low-income, middle-income, and high-income households) equally? This hypothesis is relevant because a college degree may be more valuable for high-income households as their jobs might require higher learning, but it goes beyond most low-income households' needs. Also, as education level tends to be unevenly distributed between female and male household heads, the income effect of education is likely to differ based on the household type (MHHs vs. FHHs). This paper tests if (i) the effects of different variables such as education vary depending on the household type (i.e., if the effect of the variables is household type-neutral or equal between the MHHs and FHHs); (ii) the effects of variables are likely to vary across different income groups (low-income, middle-income, and high-income households).

3. ANALYTICAL APPROACH

The study applies standard methods of income inequality measurement, decomposition, and regression analysis. These standard methods include the Gini coefficient and quantile regression. The Gini coefficient is applied to explore the structure of income inequality by income source. The regression models are applied to assess the heterogonous income effects of household characteristics by household type.

3.1. Inequality measurement and decomposition by income sources

The overall income inequality can be measured using the commonly used index of income inequality – Gini coefficient – [23] given as:

$$G = \frac{Cov[y, F(y)] \times 2}{\bar{y}}, \qquad (1)$$

where G is the Gini coefficient; $Cov[y, F(y)]$ is the covariance between income y and cumulative distribution of $F(y)$; $\bar{y}$ is mean income.

Given that the knowledge of overall inequality may be insufficient to target public policies appropriately [24], we decompose overall inequality by income source. Decomposing the overall income inequality by source and determining the proportion of overall inequality accounted for by each source could shed some light on the potential entry points for policy interventions [25]. For example, it can help determine which income sources are primarily responsible for the highest inequality level.

The contribution of each income source to overall income inequality (as measured by the Gini coefficient) can be given as:

$$\boldsymbol{C_k} = \frac{S_k G_k R_k}{G} \qquad (2)$$

where C_k is the contribution of source k to overall inequality; S_k is the share of income from source k in total income; G_k is the Gini coefficient for source k; R_k is the Gini correlation of income from source k with the distribution of total income.

The partial derivative of the overall Gini with respect to a percentage change in source *k* can be given by:

$$\psi_k = S_k(G_k R_k - G) \qquad (2.1)$$

The marginal effect of source *k* relative to the overall Gini effects of the change in income from each source (crop, livestock, and off-employment) on overall inequality can be determined by dividing Eq. (2.1) by the Gini coefficient as:

$$\epsilon_k = \frac{S_k G_k R_k}{G} - s_k \qquad (2.2)$$

Eq. (2.2) measures the change in overall inequality due to a 1% increase in income from source k across all households who earn income from the respective source.

3.2. Regression analysis

Starting with the classical linear regression, we model the log of annual income per capita as a function of household type and covariates.

$$LnY_i = \alpha + H_i'\beta + X_i'\gamma + \varepsilon_i\ , \qquad (3)$$

To test the null hypothesis of no income gap between FHHs and MHHs due to household characteristics, we include an interaction term for each household characteristic by household type and estimate a pooled regression using the full sample as:

$$LnY_i = \alpha + H_i'\beta + X_{ij}'\gamma_j + H_i X_{ij}'\lambda_j + \varepsilon_i \qquad (3.1)$$

where LnY_i is log of annual income per capita for household i=1,2,…,N; H_i denotes household type; $X_{.j}$ =1,2,…K is a vector of K independent variables (Table 1); β and γ are parameters representing the main and interaction effects; ε_i is the error term associated with income.

The classical linear regression (Eq. 3) is extended to the quantile model by [26]. Let $\{(y_i, x_{ij})\}$ for $i = 1, \ldots, N$ be independent and identically distributed sample for distribution of (y_i, x_i) where $y \in R$ is the regressand and $x \in R^k$ is used as a regressor vector. Also, let $F_{y|x}$ be the conditional distribution of y given x. The τ^{th} conditional quantile function of y given x is defined by:

$$Q_\tau(y|x) = \inf\{y|F_{y|x}(y|x) \geq \tau\} \qquad (4)$$

The linear quantile regression of equation (3) can thus be written as:

$$Q_\tau(y|x) = x'\beta(\tau), \qquad (5)$$

where $\tau \in (0{,}1)$ is a fixed and known quantity of interest; $\beta(\tau)$ is a $kx1$ vector of unknown regression coefficients.

The estimation of the β_τ's varies depending on the quantile. The $\beta_\tau{}'s$ at the median τ =0.5 are estimated by minimizing the absolute deviations [27]. However, at other parts, they are obtained by solving a linear programming problem of minimizing asymmetrically weighted absolute residuals as

$$\hat{\beta}(\tau) = \underset{b \in R^k}{Min} \frac{1}{n} \sum_{i=1}^{n} \rho_\tau (y_i - x_i' b), \qquad (6)$$

where $\rho_\tau(z) = |\tau - I(z < 0)|.|z|$ is the called the check function and $I(.)$ is the indicator function ($\rho_\tau(z) = z(\tau - 1)$ if $z < 0$ and $\rho_\tau(z) = z\tau$ otherwise).

Note that in general, the coefficients $\beta(\tau)$ vary with τ. The regression coefficient β_τ associated with an explanatory variable is interpreted as the marginal change in the τ^{th} conditional quantile of the dependent variable (household income) corresponding to the marginal change in the variable. Comparisons of β_τ across different quantiles allow inferring the effects of an individual variable at different points in the income distribution. In our context, quantile regression compares how the income of the median household of a quantile (e.g., second quartile) responds to changes in its determinants relative to the response in the income of any other household below or above that specific quantile (e.g., the first quartile).

Based on the test of the null hypothesis of no income gap between the FHHs and MHHs due to the household characteristics using Eq. 3.1, we identify the interaction terms that need to be retained and then estimate the final pooled quantile regression using the full sample as:

$$Q_\tau[LnY|X, H, X] = \alpha_\tau^0 + H'\beta_\tau^h + X_k{}'\beta_\tau^k + (H.X_k)'\boldsymbol{\beta_\tau^{xh}} + u_\tau \,, \qquad (7)$$

where LnY is the log of income; $Q_\tau[Y|X]$ is the τ^{th}conditional quantile income; H is a household type; β_τ^h is a parameter to be estimated for the main effects of household type in each quantile; β_τ^x is a vector of parameters to be estimated for the main effects of independent variables in each quantile; X is the vector of k independent variables other than household type (Table 1) that are hypothesized to explain the variation in income; β_τ^{xh} represents a slope parameter for independent variables representing the difference in log income between the τ^{th} quantile of the income distribution for the

MHHs and the corresponding τ^{th} quantile for the FHHs and u_τ is the error term associated with income in each quartile.

4. Data, study design, and measurement of variables

Primary data for this study came from a formal survey conducted in Sodo-Zuria district, Southern Ethiopia. A three-stage, randomized sampling procedure was applied in the selection of the sample households. Firstly, three *Kebeles*[7] were randomly selected from the list of 31 *Kebeles* in the district. These are *Bosa Kacha, Delbo Atwaro,* and *Delbo Wogene.* Secondly, 11 villages were randomly selected from the list of villages in the three selected *Kebeles* – four villages from *Delbo Atwaro* and *Delbo* Wogene, each, and three villages from Bosa Kacha. Finally, 154 households (94 MHHs and 60 FHHs) were randomly selected – 13 households per village from four selected villages of Delbo Atwaro; 12 households per village from four selected villages of Delbo Wogene, and 18 households per village from three selected villages of Bosa Kacha.

Data were gathered on household income disaggregated by source, demographic, socioeconomic, and institutional characteristics using a standardized questionnaire administered to the sample in 2014/15. The data were collected by trained enumerators using tablets equipped with *Surveybe* software and exported to Stata Version 14 (Stata Corp, College Station, TX, USA) for statistical analysis. In the regression analysis, annual farm household income earned from crop production, livestock production, and off-farm employment constitutes the dependent variable. The sample households' demographic, socioeconomic, and institutional characteristics constitute nine (four continuous and five discrete) independent variables (Table 1).

[7] The smallest administrative unit

Considering the small sample size, we tried to see how the values of some key socioeconomic variables, such as household size, education, and farm size, fare with those in a larger survey such as the Ethiopian Rural Socioeconomic Survey (ERSS) conducted in southern Ethiopia in 2011/12. The ERSS was implemented by the Central Statistical Authority of Ethiopia (CSA) in partnership with the World Bank Living Standards Measurement Study (LSMS) team as part of the Integrated Surveys on Agriculture Program.

Table 1: Description and measurement of variables

Variables	Code	Description and measurement
Annual household income per capita	Income	Annual income calculated as the sum of the self-reported household's agricultural production (crop plus livestock) and off-farm employment incomes, measured in Ethiopian Birr
Household type	HHtype	HHtype =1 if the household head is a man; otherwise, HHtype =0
Age	Age	Age of household head measured in number of years
Education	Education	Education measured in years of schooling
Cultivated farmland size	Land	Number of hectares dedicated to farm production
Labour	Labour	Measured in number of adult household members
Access to off-farm employment	Employment	Employment =1 if the household head had access to off-farm employment in the year preceding the survey; otherwise, employment =0
Access to seeds	Seeds	Seeds=1 if the household had access to improved seeds of any crop in their villages in the year preceding the survey; otherwise, Seeds =0
Access to extension	Extension	Extension=1 if the household was visited by an extension agent in the year preceding the survey; otherwise Extension=0
Access to credit	Credit	Credit=1 if the household had access to credit in the year preceding the survey; otherwise Credit=0

According to this regional survey, the rural household size in the Southern Region of Ethiopia is 5.3, compared to 5.5 members in our sample survey. The literacy rate in the Southern Region of Ethiopia is 55.2%, compared to 60% in our sample. The average field size of a rural household in the Southern Region is 0.88 ha, compared to 0.80 ha in our survey. Also, the size of a field, according to the regional socioeconomic survey, is 0.07 ha, compared to 0.11 ha allocated to a sweet potato field in our survey. The closeness of the values of these key variables between the two surveys testifies a careful consideration of the sampling frame and the use of the random sampling method in the study. It is well known that representativeness has to do more with the sampling frame than the sample size.

5.0. RESULTS AND DISCUSSIONS

5.1. Descriptive analysis of household characteristics

Table 2 compares the descriptive statistics of the household characteristics between FHHs and MHHs. About 40% of the sample households reported as FHHs, while the remaining reported as MHHs. The results indicate significant differences in some household characteristics between FHHs and MHHs. The MHHs have relatively older, and better-educated household heads and also have a larger household size and more access to seeds compared to FHHs. In contrast, FHHs have larger farmland than MHHs. These results suggest that FHHs and MHHs are systematically different. In the face of such systematic differences between FHHs and MHHs, it will be erroneous to attribute an observed income gap to household type. The income gap could well be due to the systematic advantages or association of one or more of the observed resources with household type.

Table 2: Descriptive statistics of household characteristics between FHHs and MHHs

Variables	FHHs	MHHs	Difference
Age	43.4 (1.9)	47.7 (1.4)	-4.3* (2.3)
Education	2.2 (0.4)	4.4 (0.4)	-2.2*** (0.6)
Land	0.68 (0.10)	0.52 (0.03)	(0.17)* (0.09)
Labour	4.4	5.5	-1.1***

	(0.2)	(0.2)	(0.3)
Employment	0.77	0.71	0.05
	(0.05)	0.05	(0.07)
Seeds	0.78	0.92	-0.142***
	(0.05)	(0.03)	(0.055)
Extension	0.70	0.72	-0.02
	(0.06)	0.05	(0.07)
Credit	0.35	0.32	0.03
	(0.06)	(0.05)	(0.08)

***, **, * significant at 1 %, 5% and 10% level, respectively; and figures in parenthesis are standard errors

Figure 1 compares the household income's kernel density estimates between FHHs and MHHs, showing higher estimates for FHHs. FHHs have 22.6% higher income than MHHs. When calculated according to the sample's quartiles, the income shares were 6.1%, 19.3%, and 43.2% at the 25th, 50th, and 75th quartiles (Table 3). These results mean that the first or lowest quartile (the poorest) accounted for about 6% of the population's income while the third quartile had about 43.2% of the income. Comparing the income shares between the FHHs and MHHs shows that the first and third quartiles of the MHHs control a relatively larger share than the FHHs. For example, the third quartile of the MHHs controls 52.8% of the income, compared to 39.6% by FHHs.

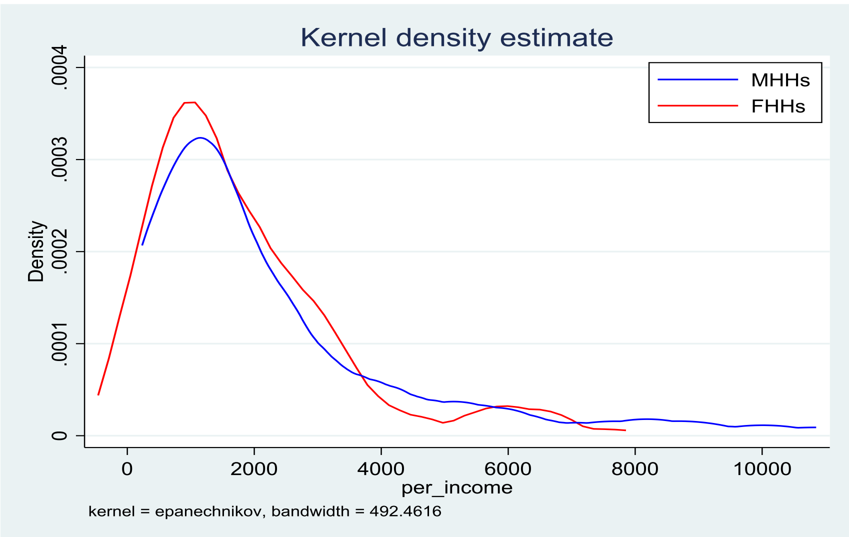

Figure 1: Household income's kernel density by household type

Comparing the different quartiles' income levels shows that the FHHs had income 3.4%, 8.5%, and 11.8% higher than the MHHs at first, second, and third quartiles. The interquartile range of the FHHs is 1.15 times higher than that of the MHHs. These differences underscore the importance of using a quantile regression rather than a mean regression.

Table 3: Income shares and levels at different quartiles

Quartile	Income shares (%)			Income levels		
	FHHs	MHHs	FHHs<MHHs	FHHs	MHHs	FHHs>MHHs
First	5.7	8.3	2.6	761.8	736.7	3.4%
Second	18.2	23.2	5.0	1507.2	1389.3	8.5%
Third	39.6	52.8	13.2	2870.8	2568.0	11.8%
Fourth				10851.0	7362.0	47.3%

5.2. Income differentials by source and household type

Table 4 presents the descriptive statistics of income from all sources by household type. Both the MHHs and FHHs earn the most income from crop production, followed by off-farm employment and livestock production. Although FHHs make more income than MHHs from all three income sources, the difference was not statistically significant. Overall, the per capita income for FHHs is ETB 2243 compared to ETB 1828 for the MHHs. However, FHHs have highly dispersed per capita income levels.

Table 4: Summary statistics of income by source for FHHs and MHHs

Income sources	FHHs	MHHs	Difference
Crop production	1198 (169)	945 (96)	253 (181)
Livestock production	228 (64)	193 (48)	34 (79)
Off-farm employment	817 (178)	690 (122)	126 (209)
All	2243 (292)	1828 (155)	414 (303)

*Note: Standard deviations in parentheses.

The Gini coefficient for the full sample in the study area is 0.447, consistent with developing countries. The Gini coefficient for the FHHs is 0.48 compared to 0.39 for the MHHs, suggesting more inequality among the FHHs relative to the MHHs.

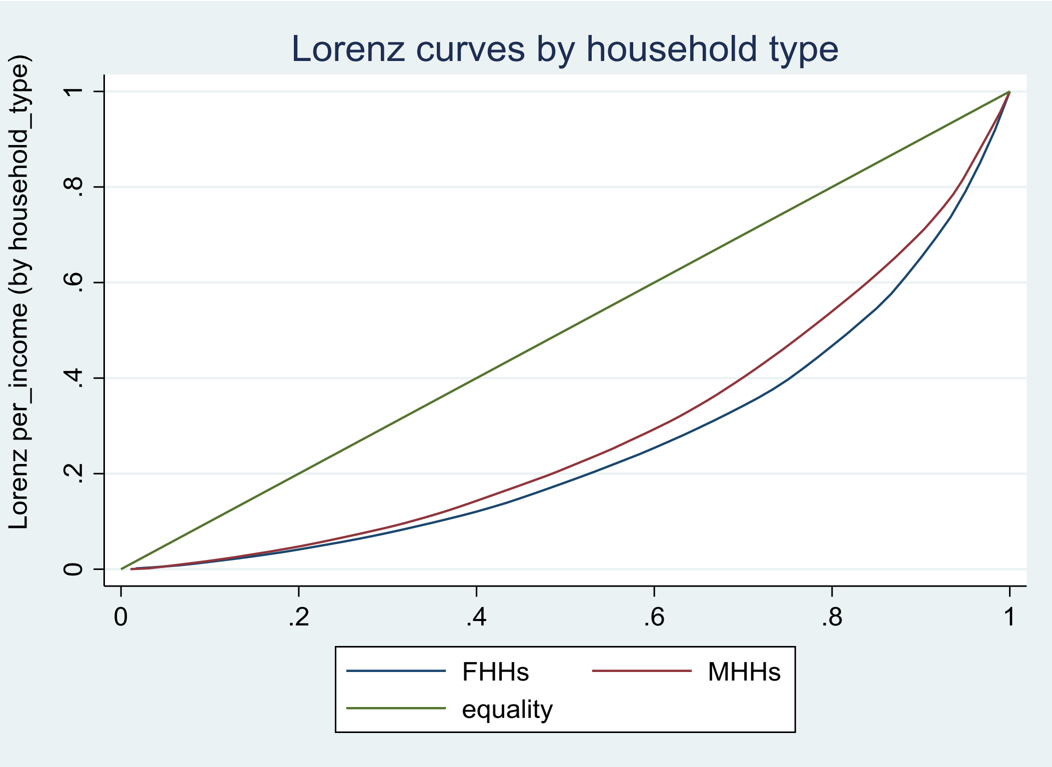

Figure 2: Lorenz curve by household type

Table 5 shows that crop production accounts for the largest share of overall income inequality (44.9%), closely followed by off-farm employment (43.5%). Livestock production contributes the least (11.6%). In terms of income share, crop production accounts for 52.4% of the total income, followed by employment (37.2%) and livestock production (10.4%). This finding implies that a given percentage increase in crop income will significantly affect overall inequality than the same percentage increase from the other two income sources. A 1% increase in crop income across all households earning income from crop production would lower the Gini coefficient by 0.076% (Table 5). However, the same percentage increase in income from off-farm employment and livestock production across all households who earn income from such activities would raise the Gini coefficient by 0.064% and 0.012%, respectively.

Consistent with the full sample results, crop production decreases income inequality while off-farm employment and livestock production increase overall inequality for both MHHs and FHHs. However, there is a difference in the magnitude of the income elasticity of inequality between them. A 1% increase in crop income would decrease income inequality by 0.109% for the MHHs, compared to 0.042% for the FHHs. In contrast, a 1% increase in off-farm income and livestock income for the MHHs would increase income inequality by 0.094% and 0.015%, compared to 0.032% and 0.010% for the FHHs.

Table 5: Gini decomposition by income sources for FHHs, MHHs, and full sample

Income source	Gini coefficient (G_k)			Share in overall income Inequality (C_k)			Impact of a 1 % change on income inequality (ϵ_k)		
	All	FHHs	MHHs	All	FHHs	MHHs	All	FHHs	MHHs
Crop	0.511	0.521	0.495	0.449	0.493	0.408	-0.076	-0.042	-0.109
Livestock	0.819	0.801	0.830	0.116	0.111	0.120	0.012	0.010	0.015
Off-farm	0.692	0.672	0.704	0.435	0.396	0.472	0.064	0.032	0.094
Total	0.447	0.479	0.417	1.00	1.00	1.00	0.00	0.00	0.00

5.3. Income differentials by household characteristics net of household type

Table 6 provides the income effects of individual household characteristics net of household type effect. For example, after controlling for household type, households with access to off-farm employment earn more income than those with no access, suggesting that access to off-farm employment positively influences income, irrespective of household type. Results also show within the MHHs, land size is significantly associated with income, with households having more than the average farmland making higher income, compared to those with less than average land.

However, within the FHHs, land size is not significantly associated with income. Extension access has also differing results between FHHs and MHHs. While FHHs with access to extension earn more income than those with no access, MHHs with access to extension earned no statistically significant income.

Table 6: Mean income per capita (ETB) between different subgroups within the FHH and MHH types

Variables	FHHs			MHHs		
	No	Yes	Difference	No	Yes	difference
Age (middle aged+?)	1949	2653	704	2110	1620	-489
Education (high schooled+?)	2222	2278	55	1684	1914	231
Land (farm size > 0.5 ha)	2158	2388	230	1668	2226	558*
Labour (household size >5?)	2327	2174	-153	2368	1673	-694**
Employment access	1133	2580	1448**	1192	2085	893***
Seed access	1902	2560	658	1735	1892	157
Extension access	1254	2666	1412**	2017	1756	-261
Credit access	2307	2122	-185	2060	1334	-726**

*** (**)* significant at 1 %, 5% and 10% level, respectively

Contrary to expectation, MHHs with no access to credit earned more income than those with access to credit. Likewise, FHHs with no access to credit also earned more income than those with access to credit. However, the difference was not statistically significant. Note that this section assessed the difference in income between households with access and no access to resources and services, controlling for household type.

5.4. Income differential based on household type

Using Columns 3 and 6 of Table 6, we assess whether or not there is an income gap between subgroups with equal access to resources and services without keeping constant the effects of other characteristics. In this case, we net out the effect of the difference in access to resources and services and see if there are income gaps between FHHs and MHHs (income gap based on household type) who reported having access to a given service. For example, do MHHs, who reported having access to improved seeds, make more income than FHHs who reported having access to the same resource? The purpose of this analysis is to draw implications if the returns to various resources and services are associated with the household type (e.g., higher among MHHs than FHHs or vice versa). Comparing the income levels of FHHs and MHHs with high school education and above, access to land and labour resources, off-farm employment, seeds, credit, and extension shows that the FHHs do better than the MHHs, implying that the returns to these resources and services could be higher among the FHHs than the MHHs. While these descriptive results provide important insights, regression analysis is warranted, controlling the effects of other observed factors.

5.5. Results of the linear regression analysis of income by household type

Table 7 presents the difference in parameter estimates of the independent variables included in the linear regression [Eq. 3] between FHHs and MHHs. We also ran a pooled regression by including an interaction term for each independent variable (Table A1). The pooled regression model estimates the differential income effects while at the same time increases the degrees of freedom. The coefficents of the interaction terms represent a slope parameter for independent variables representing the income gap between the FHHs and MHHs. The purpose of running two separate regressions by household type or a pooled regression with interaction terms included in it is to see if the income gap based on household type is intrinsic or just due to the difference in household characteristics. In other words, does the hypothesized household type-income differential disappear if we control for the variations in the observed variables? If not, which independent variables have income effects that are dependent on household type?

The OLS estimation results disaggregated by household type indicate that extension has differential income effects (Table 7). The income effects of extension for the FHHs are 0.626 compared to 0.098 for the MHHs. The difference in extension parameter estimate between the two groups is statistically significant [$X^2 = 3.46$; $p < 0.1$].

Table 7: Test of the difference in OLS parameter estimates by household type

Variables	FHHs	MHHs	Chi-square
Age	0.023 (0.028)	-0.034 (0.032)	1.85
Age-sq.	-0.000 (0.000)	0.000 (0.000)	2.18
Education	0.033 (0.031)	0.046 (0.025)	0.11
Land	0.267 (0.213)	0.289 (0.167)	0.01
Labour	-0.216 (0.064)	-0.240 (0.051)	0.09
Employment	0.619 (0.230)	0.896 0.184	0.79
Seeds	0.317 (0.206)	0.311 (0.161)	0.00
Extension	0.627 (0.206)	0.098 (0.195)	3.48*
Credit	-0.035 (0.229)	-0.457 (0.156)	2.31
Intercept	6.677 (0.769)	8.471 (0.809)	0.00

* Significant at 10% level, respectively; figures in parenthesis are standard errors

The pooled OLS estimation results confirm that extension has differential income effects depending on the household type. Therefore, we retained the interaction term for extension by household type [*Ext.×HHtype*][8] in the final estimated linear and quantile regression model[9].

Table 8 presents the ordinary least square (OLS) parameter estimates from the classical linear regression model (Eq. 3] and quantile regression model [Eq. 7] estimates in the 2nd column and 3-5th columns, respectively. The quantile regression results reveal that the quantile parameter estimates assume the same signs as the OLS estimates except for age, but their magnitude varies across quantiles. Some variables that exhibit statistically significant effects at the mean level (OLS estimates) have no statistically significant effects at other levels, such as the lower, or median, or upper-income levels (quantile estimates). For example, after controlling other variables, the main effect of household type [*HHtype*] and its interaction with extension [*Ext.×HHtype*] does not exhibit statistical significance at the lower-income level (τ=0.25). However, both the main effect of household type and its interaction effects with extension are statistically significant at the mean level, middle (τ=0.50) and high-income (τ=0.75) levels*.* This means that access to extension does not matter for lower-income FHHs and MHHs. But, it does for both FHH and MHHs at the middle – and high-income levels.

Education is positively and significantly related to household income at the mean level, middle and upper tail of the income distribution. The magnitude of the contribution of

---

[8] To confirm the necessity of including an extension by household type interaction term in the pooled regression (Eq. 7), we run a likelihood ratio test to test if $\beta_{\tau}^{eh} = 0$. Results indicate that the interaction term is statistically different from zero [X^2 (1) =3.67; p =0.055], confirming that it should be retained.

[9] The estimated model was tested for multicollinearity and heteroskedasticity. The multicollinearity test is based on the variance inflation factor (VIF) and condition index (CI) values. The VIF and CI computed for all the predictors in the model using a weighted least square method of linear regression are below 10, which is well within the acceptable range, suggesting a lack of multicollinearity. The problem of multicollinearity is present if the value of VIF is greater than 10 [28] and the CI is equal to or greater than 30 [29]. The test for heteroskedasticity was conducted using the Breusch-Pagan /Cook-Weisberg test, which indicates that heteroskedasticity is not a problem as evidenced by the failure to reject the null hypothesis of homoscedasticity that the error variances are all equal [X^2 *(1)=0.48; p* $>X^2$ *(1)= 0.49*].

education is highest among high-income households. An additional year of schooling results in 5.4% more income for high-income households, compared to 4.7% earned by middle-income households. On average, an additional year of schooling results in 3.7% more income, which is consistent with existing empirical evidence on schooling's return.

All things being equal, land size has a positive and marginally significant effect on household income. A 1-ha increase in land size is associated with a 37%, 27%, and 35% income increase among low, middle- and high-income households, respectively. On average, a 1-ha increase in land size is associated with a 26% increase in income.

After controlling other variables, labour availability measured by the number of adult household members is negatively and significantly associated with income across all parts of the income distribution – poor, middle, and high-income households, suggesting that the net income effect of an additional household member is negative across all income groups. An additional household member is associated with a 24%, 26%, and 30% income reduction among the lower, middle- and high-income households. On average, an additional household member is associated with a 23% decrease in income.

Access to off-farm employment and seeds exhibit positive and statistically significant income effects across all parts of the income distribution – poor, middle, and high-income households. This shows that employment and access to seeds benefit the households across the entire income distribution. Off-farm employment results in much more significant income effects than the other factors. Off-farm employment more than doubles the income among the lower- and middle-income households, while it nearly doubles the income among the high-income households. On average, a household with access to off-farm employment earns 120% higher income than otherwise. Access to seeds is positively associated with income, leading to a 45% increase among low-income, 40% among middle income, and 36% among high-income households. On average, access to seeds is associated with a 37% increase in income.

Table 8: OLS and quantile regression parameter estimates of income

Variables	Average	Lower-income	Middle-income	High-income
		τ=0.25	τ=0.50	τ=0.75
HHtype	0.465** (1.94)	0.713* (1.73)	0.582** (2.26)	0.497** (2.05)
Age	-0.000 (-0.01)	-0.014 (-0.28)	0.0118 (0.31)	0.014 (0.39)
Age-sq.	-0.000 (-0.34)	0.000 (0.05)	-0.000 (-0.46)	-0.000 (-0.29)
Education	0.037** (1.95)	0.0338 (1.06)	0.0471* (1.79)	0.054* (1.90)
Land	0.263** (1.94)	0.369 (1.61)	0.273* (1.69)	0.347* (1.57)
Labour	-0.233*** (-5.34)	-0.241*** (-3.46)	-0.263*** (-5.07)	-0.302*** (-5.02)
Employment	0.789*** (5.09)	0.856*** (3.81)	0.778*** (3.75)	0.666*** (3.01)
Seeds	0.317** (2.45)	0.374 (1.56)	0.338** (1.97)	0.308* (1.77)
Extension	0.630*** (3.00)	0.319 (1.05)	0.554** (2.11)	1.007*** (3.07)
Ext.×HHtype	-0.511* (-1.74)	-0.340 (-0.65)	-0.605* (-1.69)	-0.802** (-2.11)
Credit	-0.286** (-2.09)	-0.223 (-0.99)	-0.225 (-1.41)	-0.376** (-2.17)
Constant	7.230*** (11.79)	7.162*** (6.19)	7.005*** (7.58)	7.404*** (8.36)
N	154	154	154	154

***, **, * significant at 1 %, 5% and 10% level, respectively; figures in parenthesis are *t*-statistics.

Unlike the case with access to seeds and off-farm employment, the extension has differential income effects at different parts (the mean level, middle- and upper tail) of the income distribution depending on the household type. For example, at the mean level, the effect of extension is given by [0.630-0.511*household type]. Since FHHs and MHHs are coded 0 and 1, respectively, the extension coefficients for FHHs and

MHHs are 0.630 and 0.119. These results suggest that FHHs with access to extension earn 87.8% [exp. (0.630)-1] more income than those FHHs with no access. In contrast, MHHs with access to extension make just 12.6% [exp. (0.119)-1] more income than those MHHs with no access. Similarly, FHHs at the upper level (τ=0.75) with access to extension earn 2.73 times more income than those FHHs with no access. In contrast, MHHs with access to extension make just 22.7% more income than those MHHs with no access at the median level (τ=0.5). These findings suggest that access to extension presents more benefits to the FHHs than to the MHHs. The plausible explanation for the differential income effects based on household type is that although both household types have access to extension, the members in the FHHs may have implemented the advice by extension agents properly. In contrast, members in the MHHs may have ignored them, resulting in differential returns.

Although extension has a positive income effect at the mean and upper-income levels, it does not have differential income effects at the lower level, as evidenced by the lack of statistical significance for the interaction term. The effect of extension among the low-income households is independent of household type (i.e., household type-neutral), suggesting that FHHs in those income groups may get as much returns from the extension services as the MHHs.

The quantile regression results suggest that had we applied the classical linear regression instead of the quantile regression to identify income determinants, we would have made misleading conclusions. For example, the quantile results indicated no significant relationship between extension and income in the low-income household category. However, results at the mean level, middle- and upper-end income distribution, show significant income effects. Using quantile regression enabled us to identify income determinants for different income categories such as low-income, middle income, and high-income households.

6.0. CONCLUSION AND IMPLICATIONS

The paper has achieved two objectives using data from a formal survey of 154 households selected using a cluster sampling technique in the Sodo-Zuria district, Southern Ethiopia. First, the paper measured the income inequality within FHHs and MHHs and assessed the contribution of different income sources (crop, livestock, and off-farm employment) to the overall income inequality. Results show that income is relatively more evenly distributed within the MHHs than the FHHs. Crop production represents an inequality-decreasing source. In contrast, livestock production and off-farm employment represent inequality-increasing sources of income, suggesting that crop production-focused interventions are preferred to address the issue of income inequality.

Second, using quantile regression, the paper tested if the income effects of different variables vary across three income groups (low-income, middle-income, and high-income households) depending on the household type. The paper identified the sources of income differential between FHHs and MHHs and determinants of income inequality among FHHs and MHHs.

Among nine variables considered in the model, only extension has income differential effect based on household type at the middle- and upper-income levels, with more benefits accruing to the FHHs than the MHHs. This might be because the middle- and upper-income FHHs have received complete and better extension services than MHHs. So, although both groups reported having access to extension services, the difference in the quality and intensity of the delivered services could lead to different returns to extension services. This finding implies that while addressing differentials in access to extension services between FHHs and MHHs is necessary, it is not sufficient to reduce income inequality. It is imperative to address the differentials in returns to those services through supervision of the quality and intensity of extension delivery at the farm level.

In contrast to extension, other factors such as age, education, labour, land, seeds, and off-employment, while contributing to the overall income gap, did not result in an income gap between the two household types. In other words, there is no income gap between the FHHs and MHHs depending on their differences in age, education, credit, labour, land, seeds, and off-employment. For example, on average, literate FHHs earn as much income as literate MHHs. This findig suggests that age, education, credit, labour, land, seeds, and employment are not considered contributors to income inequality associated with household type. Note, however, that these factors are contributing to overall income inequality.

Regression results reveal that the magnitudes of contribution vary depending on the quantile being considered. Some income groups are more affected by certain factors than other groups. For example, the income effect of education is highest among high-income households. Some parts of the income distribution (income groups) are not affected by certain factors, while other income groups are affected. For example, extension did not matter for lower-income households but mattered a lot among high-income households. These non-uniform income effects present important evidence for policymakers to avoid ‘one-size fits all’ approach and tailor policy interventions to particular income groups.

References

1. World Bank Group, (2016). *Global Monitoring Report 2015/2016: Development Goals in an Era of Demographic Change.* Washington, DC: World Bank. DOI: 10.1596/978-1-4648-0669-8. License: Creative Commons Attribution CC BY 3.0 IGO.
2. FAO, (2015). *The State of Food and Agriculture. Social protection and agriculture: breaking the cycle of rural poverty*. ISSN 0081-4539 Food and Agriculture Organization of the United Nations, Rome, 2015.
3. World Bank (2008). World development report 2008: Agriculture for development. The World Bank, Washington, DC, 2008
4. Geda, A., Shimeles, A. and Weeks, J., (2009). Growth, poverty and inequality in Ethiopia: Which way for pro-poor growth?” *Journal of International Development* 21(7):947-970
5. UNDP (United Nations Development Program). (2015). Inclusive growth and inequalities in the context of structural transformation, Ethiopia, Issue Brief, UNDP
6. Klasen, S., Lechtenfeld, T., and Felix Povel, F. (2011). What about the Women? Female Headship, Poverty and Vulnerability in Thailand and Vietnam. Courant Research Center: Poverty, Equity and Growth – Discussion Papers 76, Courant Research Centre PEG.
7. World Bank (2012). World development report 2012: Gender equality and development. The World Bank, Washington, DC, 2012
8. AfDB (African Development Bank). (2005). Empowering African women: An agenda for action. Gender and Social Development Monitoring Division, Quality Assurance and Results Department, AfDB, Abuja, Nigeria
9. FAO (2011). Women in Agriculture: Closing the Gender Gap. Rome, 2011 gender gap
10. Milazzo, A. and van de Walle, D. (2015). Milazzo, A. and van de Walle, D. (2015). Women Left Behind? Poverty and Headship in Africa. Policy Research Working Paper 7331. World Bank Group. Development Research Group Human Development and Public Services Team, June 2015
11. Firpo, S., N. Fortin and T. Lemieux (2009), “Unconditional Quantile Regressions,” Econometrica, Vol. 77, pp. 953-973, May.
12. Fournier and Koske, 2016). (2012), “The determinants of earnings inequality: evidence from quantile regressions”, OECD Journal: Economic Studies, Vol. 2012/1. http://dx.doi.org/10.1787/eco_studies-2012-5k8zs3twbrd8
13. Ghimire, R., Wen-chi, H., and Shrestha, R.B., 2015. Factors affecting adoption of improved rice varieties among rural farm households in central Nepal. Rice Science 22(1): 35−43
14. Kassie M, B. Shiferaw, and G. Muricho. (2011). “Agricultural technology, crop income, and poverty alleviation in Uganda.” World Development 39 (10): 1784–1795.
15. Asfaw, S., Shiferaw, B., Simtowe, F. and Lipper, L., 2012. Impact of modern agricultural technologies on smallholder welfare: Evidence from Tanzania and Ethiopia. Food Policy, 37(3): 283–295.
16. Mignouna, B., Manyong, V.M., Rusike, J., Mutabazi, S., and Senkondo, M., 2011. Determinants of adopting Imazapyr-resistant Maize technology and its impact on household income in Western Kenya. AgBioforum, 14(3): 158-163

17. Asfaw, A. and Admassie, A. (2005). The role of education on the adoption of chemical fertiliser under different socioeconomic environments in Ethiopia. Agricultural Economics, 30 (3)|: 215-228. https://doi.org/10.1111/j.1574-0862.2004.tb00190.x
18. Mwangi, M., and Kariuki, S. (2015). Factors determining adoption of new agricultural technology by smallholder farmers in developing countries. Journal of Economics and Sustainable Development, 6(5)
19. Mohamed and Temu (2008). Access to credit and its effect on the adoption of agricultural technologies: The case of Zanzibar. African Review of Money Finance and Banking
20. Feleke and Zegeye 2005). Feleke, S., and Zegeye, T., 2005. Adoption of improved maize varieties in Southern Ethiopia: Factors and strategy options. Food Policy 31, 442-457
21. Wossen, T., Abdoulaye, T., Alene, A., Haile, M., Feleke, S., Olanrewaju, A. and Manyong, V. (2017). Impacts of extension access and cooperative membership on technology adoption and household welfare. Journal of Rural Studies 54: 223-233.
22. World Bank. (2014). Levelling the field: Improving opportunities for women farmers in Africa, 2015. Retrieved on February 18, 2021 from https://openknowledge.worldbank. org/bitstream/handle/10986/17790/860390WP0WB0ON0osure0date0March0180.pdf?sequence=1.credit
23. Lerman, R. I., and Yitzhaki, S. 1985. Income inequality effects by income source: A new approach and applications to the United States” Review of Economics and Statistics 67: 151-156.
24. Bellù, L.G., 2006. Inequality analysis: The Gini index. Agricultural Policy Support Service, Policy Assistance Division, FAO, Rome
25. Haughton, J. and Khandker, S. 2009. Handbook on poverty and inequality. World Bank, Washington DC
26. Koenker and Bassett (1978). Koenker, R., and Bassett, G. S. 1978. Regression quantiles,” Econometrica 46, 33-50
27. Buchinsky, M.,1998. Recent advances in quantile regression models: A practical guideline for empirical research. Journal of Human Resources 33 (1): 88-126
28. Myer, R.H., and Montgomery D., 1995. Response Surface Methodology: Process and Product Optimization Using Designed Experiments, New York, NY: Wiley
29. Belsley, D., Kuh, E. and Welsch, R., 1980. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity” New York: John Wiley and Sons

Table A1: Pooled linear regression parameter estimates of income

Variables	Model 1 (Without interaction)	Model 2 (With interaction)
HHtype	0.119 (0.85)	1.794 (1.35)
Age	-0.005 (-0.20)	0.023 (0.72)
Age-sq.	-0.000 (-0.23)	-0.000 (-1.06)
Education	0.036* (1.97)	0.0332 (1.07)
Land	0.272** (2.00)	0.267 (1.22)
Labour	-0.235*** (-5.82)	-0.216*** (-3.52)
Employment	0.782*** (5.58)	0.619*** (2.65)
Seeds	0.311** (2.39)	0.318 (1.50)
Extension	0.329** (2.31)	0.628*** (2.74)
Credit	-0.275** (-2.09)	-0.0356 (-0.17)
Age×HHtype	—	-0.0578 (-1.12)
Age-Sq.×HHtype	—	0.000581 (1.22)
Educ.×HHtype	—	0.0128 (0.32)
Land×HHtype	—	0.0225 (0.08)
Labour×HHtype	—	-0.0244 (-0.29)
Employ.×HHtype	—	0.276 (0.94)
Seed×HHtype	—	-0.00619 (-0.02)
Ext.×HHtype	—	-0.530* (-1.78)
Credit×HHtype	—	-0.421 (-1.56)
Constant	7.601*** (12.35)	6.677*** (7.48)
N	154	154

Likelihood-ratio test: LR $X^2(9)$ = 9.98

Assumption: model 1 nested in model 2: Prob > X^2 = 0.3520