Factors affecting the income of rural households in the Midlands and the Northern Mountains, Vietnam: Evidence from quantile regression

Abstract

To study the factors affecting the income of rural households in the Northern Midlands and Mountainous regions, Vietnam using both OLS regression methods and quantile regression at the same time, the analysis was based on the tabular dataset of 2,630 households filtered from the Vietnam Rural Household Resource Access Survey (VARHS) 2010 – 2018. The OLS regression method aims to determine the factors affecting the average income of households, the quantile regression method aims to determine the factors affecting the income of rural households on different percentiles. The results show that gender, education and savings have a significant impact on household income, with gender disparities particularly pronounced in the lower income segments. In addition, the role of ethnicity becomes more important in high-income segments, suggesting that access to resources and opportunities varies between social groups.

Key word: Quantile regression, income, rural household, Northern Mountains, Vietnam,…

1. Introduction

Vietnam's rural areas have always been considered extremely strategically important in the country's development process, considered an inseparable core part in the process of promoting sustainable socio-economic development, not only in terms of the proportion of residential areas but also in terms of positive socio-economic contributions associations in our country in the current period. The Northern Midlands and Mountains are regions with a particularly important position in socio-economic development. According to the General Statistics Office (2023), the Northern Midlands and Mountains region has nearly 7 million ethnic minorities living in a concentrated area, accounting for more than 50% of the country's ethnic minority population, which is the region with the lowest average income in the country. In 2020, the average income of the whole region was 2.7 million VND/person/month, by 2023 the per capita income has increased from 2.7 million VND to 3.44 million VND/person/month but is still considered an area in the "poor core" of the country.

In the world and Vietnam, there have been quite a lot of studies on factors affecting household income. The results of the study show that the identification of factors affecting the income of rural households plays an important role in proposing recommendations to improve household income and living standards. In Vietnam, research on household income is mainly concentrated in small areas such as a province or a district, there are few studies in large areas, especially in the Northern midland and Mountainous provinces where the economy and income are the lowest in the country. On the other hand, previous studies have used the OLS regression method to analyze factors affecting income, but this method has not clarified the differences in the impact between different groups of households with different incomes, the quantile regression method allows to consider factors that affect the income of households on each percentile different. Therefore, the author chooses the topic *"Factors affecting the income of rural households in the Northern Midlands and Mountains, Vietnam: Evidence from quantile regression"* which is extremely important and necessary.

2. Literature review

The authors' studies (Spring & South, 2011; Imam & Associates, 2018; Sien, 2021; Fadipe et al., 2014) Household income is influenced by key factors such as age, gender, education level, land area, social capital, assets, employment, electricity use, and income from the non-commercial sector. The education level of the head of the household, the ethnic composition of the head of the household, the rate of dependency, the gender of the head of the household, the employment status and marital status of the head of the household are also significant determinants of household income (Biyase& Zwane, 2018). Similarly, the research results of Duyen (2014), Loan et al. (2015), and Tuyen (2015) show that factors such as gender, education level, land area, loans, and the number of employees have an impact on the income of farmers.

Research by Artha and Dartanto (2018), Ogutu and Qaim (2019), Ravindra et al. (2020) found a direct relationship between household size and household income. The ethnic composition of the head of the household is also considered the main determinant of the income and poverty of the household. Education level and household size have an impact on household income (Farzam et al., 2021). Dassanayake et al. (2015) studied the heterogeneity of household structure and income in Zimbabwe and South Africa, suggesting that female-headed households do not have lower incomes than male-headed households. Income disparities among female-headed households are significantly related to the number of adult males present and their complementarity with children living in households.

Human capital has long been defined as the asset of each country, human capital is directly related to economic growth, and this relationship can be measured by the amount of money invested in human education (Scully, 2002). An increase in human and material capital can reduce inequality and help distribute income more equitably (Shahpari& Davoudi, 2014). Tuyen (2015) has studied socio-economic factors that determine household income among ethnic minorities in the Northwest Mountains as the poorest region in Vietnam, the study has confirmed the role of education in improving household income. Teame& Woldu (2016) studied the factors influencing the income diversification of rural households in Zoba Maekel, Eritrea, confirming human capital; social capital has a positive impact on the total income of households.

In addition, access to credit can significantly increase the ability to meet household financial needs such as improved purchase and use of agricultural inputs, access to credit capital that affects the welfare of households (Diagne et al., Social capital theory acknowledges that network relationships help solve social problems (Bourdieu, 1986). Research by Sun et al. (2014) suggests that social capital also affects their ability to improve their income, providing them with job opportunities and advancement.

3. Methods and data

3.1. Methods

The study uses the quantile regression method to analyze the determinants of household income. Previous studies have often used the conventional least squared (OLS) method, to study the determinants of household income (Eshetu, 2020). OLS regression focuses only on the conditional average estimation of the dependent variable (e.g., income average). Quantile regression, on the other hand, provides information at different percentiles of distribution, such as low-income groups (10% percentile) or high-income groups (90% percentile) (Koenker& Bassett, 1978).

3.2. Data

The study was conducted based on data from the Vietnam Rural Household Resource Access Survey (VARHS) in the period 2010 - 2018. From this dataset, the author filtered out the data of households with incomes in the Northern Midlands and Mountains, and filtered out the necessary variables for his research model. After filtering the data, the observations with no income and the observations with insufficient data are removed, and the data is used for the research model.

Research Model

The research model on factors affecting the income of rural households in the Northern Midlands and Mountains of Vietnam is as follows:

$$Ln(Thunhap)_i = \beta_i + \beta_i Gioitinh_i + \beta_i TuoiCH_i + \beta_i TDHV_i + \beta_i Dantoc_i + \beta_i Honnhan_i + \beta_i Hongheo_i + \beta_i TylePT_i + \beta_i Dientichdat_i + \beta_i Sovonvay_i + \beta_i Tietkiem_i + \beta_i KC\_duongnhua_i + \beta_i Socusoc_i + \beta_i Sodonvithamgia_i + \beta_i Chiphi\_tg_i + \beta_i Quanhe\_CQ_i + \beta_i Nhocay_i + \varepsilon_i \quad (1)$$

Table 1. Descriptive statistics

Observed variables	Smallest Value	Greatest Value	Average value	Standard deviation
Gender head of household	0	1	0.179	0.383
Age head of Household	21	90	55.769	13.102
Educational level of the household head	0	12	5.442	4.287
Ethnic households	0	1	0.502	0.500
Marriage	0	1	0.838	0.368
Poor households	0	1	0.200	0.400
Dependency Rate	0	100	41.703	28.806
Production land area	3	12	8.257	1.170
Loan amount	0	15	3.188	4.714
Thrifty	0	1	0.900	0.301
Distance from the house to the asphalt road	0	200	2.466	8.621
Number of shocks	0	3	0.706	0.923
Number of participating nits	0	3	1.194	0.458
Cost of participation	0	3,384	46.325	118.759
Relations with the government	0	1	0.227	0.419
Rely on	0	1	0.063	0.243
Y: Average household income (thousand VND/year)	1,920	18,200,000	91,040	366,398
Sample size = 2,630				

(Source: Author's calculation)

Table 1 provides descriptive statistics of the variables in the research model, helping us to have an overview of the characteristics of households in the Northern Midlands and Mountains such as: the gender of the head of the household is a variable with a value from 0 to 1, the average value is 0.179, This shows that male heads of households only account for about 17.9%. Next is the age of the head of the household with an average age of 55.57 and a

standard deviation of 13.102, indicating that the head of the household is mainly middle-aged, but there is a diversity of ages. The dependency ratio has an average value of 41,103 with a standard deviation of 28,806, indicating a relatively high proportion of dependents relative to the total number of household members. The average annual income is 91,040 thousand VND (about 91 million VND/year) with a standard deviation of 366,398, showing a huge difference in income between households. The lowest value is 1,920 thousand VND and the highest is 18,200,000 VND, proving that there are some households with very low or very high incomes compared to the common ground.

4. Results and discussion

In order to determine the accuracy of the research model and estimate the regression, ensuring the reliability of the research results, the author has carried out testing steps such as: the suitability of the model, multi-collinear, variance of variable error and autocorrelation phenomenon.

Table 2. Regression results of factors affecting household income

Independent variables	OLS	FIVE	REM
Gender head of household	0.181*** [3.45]	**0.162** [1.58]	0.160** [2.56]
Age head of Household	0.002* [1.74]	0.0123*** [4.55]	0.003** [2.26]
Educational level of the household head	0.027*** [7.78]	0.010*** [2.72]	0.022*** [6.50]
Ethnic households	0.175*** [4.92]	**0.021** [0.14]	0.189*** [4.27]
Marriage	0.249*** [4.57]	**0.11** [1.32]	0.220*** [3.58]
Poor households	-0.429*** [-11.74]	-0.259*** [-6.25]	-0.367*** [-9.94]
Dependency Rate	-0.005*** [-9.85]	-0.005*** [-6.27]	-0.005*** [-9.09]
Production land area	0.066*** [4.49]	0.048* [1.66]	0.062*** [3.55]
Loan amount	0.006** [2.15]	**-0.003** [-0.82]	**0.002** [0.81]
Thrifty	0.463*** [9.99]	0.363*** [8.22]	0.414*** [9.62]
Distance from the house to the asphalt road	**-0.001** [-0.69]	**-0.000** [0.02]	**-0.000** [-0.56]
Number of shocks	-0.055*** [-3.50]	-0.035** [-2.29]	-0.048*** [-3.23]
Number of participating nits	0.186*** [6.21]	0.089*** [3.10]	0.133*** [4.74]
Cost of participation	0.000***	0.000***	0.000***

	[6.09]	[3.36]	[5.06]
Relations with the government	0.130*** [3.88]	**0.037** [1.14]	0.089*** [2.85]
Rely on	0.253*** [4.30]	0.145* [1.92]	0.242*** [3.87]
Constant	9.458*** [54.51]	9.536*** [30.60]	9.591*** [47.56]
N	**2.630**	**2.630**	**2.6300**

(Source: Author's calculation)

Table 3, the regression results of table data by the OLS method after multi-collinear testing, autocorrelation testing, change error variance testing and overcoming autocorrelation, change error variance shows that the 15 explanatory variables of the xtgls model are meaningful at 1% except for the significant loan variable. at 5% and 01 variable explaining no statistical significance is the distance from the house to the asphalt road.

Although quantile regressioncan be estimated from the 0.01 percentile to the 0.99 percentile with any jump, due to the software's processing time and speed limitations, the study only performed regressions at the basic percentiles 0.1 – 0.25 – 0.5 – 0.75 – 0.9. The regression results at these basic percentiles are enough as a basis for the topic to make analyses and recommendations, the higher the percentile corresponds to the higher the average income of the head of household.

Regression results by least squares method (xtgls) and quantile regressionmethod (10%, 25%, 50%, 75% and 90%) are presented side-by-side in the same result table to show the advantages of quantile regression.

Table 3: Results of OLS regression model and quantile regression

Independent variables	**xtgls**	**Quantile regression**				
		Q10	**Q25**	**Q50**	**Q75**	**Q90**
Gender head of household	0.181*** [4.21]	0.296*** [3.15]	0.312*** [4.66]	0.137** [2.22]	0.109* [1.81]	**0.084** [0.92]
Age head of Household	0.002*** [2.10]	**0.001** [0.27]	**0.001** [0.45]	0.003* [1.69]	0.004*** [2.72]	**0.002** [1.03]
Educational level of the household head	0.021*** [8.45]	**0.006** [0.89]	0.023*** [5.22]	0.032*** [7.59]	0.034*** [8.30]	0.036*** [9.51]
Ethnic households	0.215*** [7.26]	**0.091** [1.42]	**0.071** [1.57]	0.201*** [4.79]	0.260*** [6.36]	0.266*** [4.30]
Marriage	0.188*** [4.23]	0.583*** [5.95]	0.295*** [4.23]	0.196*** [3.04]	0.138** [2.20]	**0.093** [0.98]
Poor households	-0.376*** [-15.47]	-0.310*** [-4.73]	-0.391*** [-8.36]	-0.446*** [-10.32]	-0.456*** [-10.83]	-0.478*** [-7.50]

Dependency Rate	-0.006*** [-14.12]	-0.003*** [-3.51]	-0.005*** [-8.16]	-0.006*** [-9.27]	-0.005*** [-9.03]	-0.005*** [-4.97]
Production land area	0.052*** [4.37]	0.108*** [4.09]	0.068*** [3.59]	0.064*** [3.68]	0.056*** [3.32]	**0.033** [1.27]
Loan amount	0.004** [2.18]	**0.001** [0.18]	**0.000** [0.11]	**0.006** [1.72]	0.006* [1.90]	0.016*** [3.10]
Thrifty	0.386*** [12.70]	0.404*** [4.85]	0.421*** [7.11]	0.425*** [7.76]	0.419*** [7.85]	0.474*** [5.87]
Distance from the house to the asphalt road	**-0.001** [-1.14]	**0.000** [0.05]	**0.000** [0.05]	**-0001** [-0.59]	**-0.002** [-1.09]	**-0.002** [-0.62]
Number of shocks	-0.046*** [-4.51]	-0.065** [-2.31]	-0.070*** [-3.50]	-0.053*** [-2.88]	-0.050*** [-2.77]	**-0.044** [-1.63]
Number of participating nits	0.125*** [6.11]	0.237*** [4.42]	0.184*** [4.81]	0.184*** [5.20]	0.153*** [4.45]	**0.070** [1.34]
Cost of participation	0.001*** [5.26]	0.001*** [4.19]	0.001*** [4.53]	0.001*** [3.89]	0.001*** [7.26]	0.001*** [5.37]
Relations with the government	0.063*** [2.83]	0.158*** [2.63]	0.099** [2.31]	0.070* [1.76]	0.107*** [2.79]	0.176*** [3.02]
Rely on	0.285*** [7.28]	0.322*** [3.05]	0.317*** [4.21]	0.275*** [3.95]	0.176*** [2.60]	0.177* [1.73]
Constant	9.838*** [71.79]	8.071*** [25.93]	9.191*** [41.46]	9.571*** [46.68]	9.986*** [49.99]	10.646*** [35.22]
N	**2,630**	**2,630**	**2,630**	**2,630**	**2,630**	**2,630**
*, **, ***: meaning 10%, 5%, 1%						

(Source: Author's calculation)

The results of Table 3 show that, for the OLS regression method, the gender of the head of household is statistically significant at 1%, which is related to the household income. For quantile regression, the gender of the head of the household has the strongest impact on the distribution of income in the 10% percentile, if the head of the household is male, the average income will increase compared to the average income of the female head and then tend to decrease in the 25% percentiles. 50% and 75% This shows that women often have difficulties in accessing resources such as capital, land, technology, and women's social role may be bound by family responsibilities, leading to less time and opportunities to participate in economic activities. The results of this estimate are similar to the research results of Datt et al. (2000), Fadipe et al. (2014), Imam et al. (2018). However, at the 90% percentile, the gender of the head of the household is not statistically significant. This may explain that, at the 90% percentile, the household sample usually includes households with superior incomes, are

susceptible to exceptions, or sources of income that do not properly reflect the role of gender, such as income from sources other than agriculture or financial investment.

The education level of the head of household is statistically significant at 1% and the regression coefficient is marked (+), so there is a covariate relationship with the income of the household in both models. The results of the quantile regression model show that, except for the 10% percentile variable of education level, the average household income gradually increases at the high percentile. This also shows the importance of education level in the income of rural households in the Northern Midlands and Mountains in particular, Vietnam in general, education has a stronger impact on high-income groups, because they are able to take advantage of knowledge to expand economic activities, such as business or investment.

The estimated results from the quantileregression method also show that the heads of households who are Kinh ethnic groups have an income level compared to the heads of households of other ethnic groups and gradually increase from the 50% percentile to the 90% percentile. This can be explained because the Kinh people often own more land and assets more clearly, making it easier for them to invest and produce than other people. In addition, the literacy rate and education level of Kinh people are usually higher, making them able to apply new techniques and improve labor productivity. Moreover, ethnic minority groups are more vulnerable to economic fluctuations than Kinh households. This study gives similar results to the research results of Fadipe et al. (2014), Tran Quang Tuyen (2015), Tran Dinh Thao et al. (2022). However, at the 10% percentile, 25% of ethnic variables are not statistically significant, possibly because these households often have more difficulty accessing resources and economic opportunities. Therefore, ethnic differences may not make a significant impact in the interpretation of income. Meanwhile, in higher segments (50%, 75%, 90%), ethnic differences may become more pronounced because these households have more opportunities for development and investment. In summary, the difference in the statistical significance of the household ethnicity variable between income fractions reflects the complexity of the relationship between ethnicity and income, which depends on the economic and social conditions of each group.

The results from Table 3 show that the saving variable is statistically significant in both models and the value gradually increases from 10% to 90%. Savings often reflect the financial capacity and financial management of the household. Households with higher savings tend to have higher incomes, which is more evident in higher income levels. As income increases, so does the household's ability to save. Households with better savings can invest in education, health, and other economic activities, which in turn generate higher incomes. This may

explain why the savings variable makes sense at higher levels, where households tend to invest more in development resources. This is similar to previous studies on the role of savings in raising household income (Kabeer, 2015). Kabeer's research shows that households with higher savings can invest in education and health, which in turn raises their incomes.

5. Conclusion and recommendations

The study analyzed the factors affecting the income of rural households in the Northern Midlands and Mountains of Vietnam, based on data from 2,630 observations and applied two models of OLS regression and quantile regression. The results show that household income is significantly influenced by a number of factors such as gender, education level, savings, and ethnicity of the head of the household. Specifically, income disparities between the sexes still exist, with men often earning higher than women. This emphasizes the urgent need to implement gender equality policies to improve the economic status of women in the community. Furthermore, education was identified as the most important factor in raising income, suggesting that improving access to education is necessary for human resource development. Savings also play an important role, with higher savings helping households increase their income and financial stability. In addition, the role of ethnic heads of households has become more pronounced in high-income segments, suggesting that ethnic minority groups may face more barriers to accessing economic opportunities.

From the results of the analysis of factors affecting the income of households in the Northern Midlands and Mountains, Vietnam has made a number of recommendations to improve incomes as well as living standards and economies for households as follows:

Firstly, strengthen vocational education and training: expand the general education network in remote and remote areas to increase access for children. Provide free or reduced-fee vocational training programs for local workers, focusing on occupations that are suitable for natural conditions and actual needs (such as agricultural product processing, handicrafts, or community tourism services). Develop vocational training programs that are tailored to local labour market needs, helping to upskill both men and women. Ensure that all children, especially children from ethnic minority households, have access to quality education.

Secondly, capital and microcredit support: developing microcredit programs with preferential interest rates, creating conditions for poor and near-poor households to invest in production and business. Encourage the establishment of credit savings groups in the community, creating revolving capital. The Vietnam Bank for Social Policies and microcredit institutions have experience in implementing these models in mountainous areas. The government only needs to increase promotion and support supervision.

Thirdly, promoting social relations and the role of the government: strengthening the relationship between the people and the government through agricultural extension programs, technical training, singing the role of hamlet heads, village heads and social organizations in supporting people's economic development.

REFERENCES

Artha, D. R. P., &Dartanto, T. (2018). The multidimensional approach to poverty measurement in Indonesia: Measurements, determinants and its policy implications. *Journal of Economic Cooperation & Development*, *39*(3), 1-38.

Barnard, H., & Turner, C. (2011). Poverty and ethnicity: A review of evidence. *York: Jrf*.

Biyase, M., & Zwane, T. (2018). An empirical analysis of the determinants of pverty and household welfare in south Africa .*Journal of developing areas*, 115-130.

Bourdieu (1986). The Forms of Capital. In: Richardson, J. G., ed. 1986. *Handbook ofTheory and Research for the Sociology of Education*. NewYork: Greenwood.

Dassanayake, W., Luckert, M. K., & Mohapatra, S. (2015). Heterogeneity of household structures and income: Evidence from Zimbabwe and South Africa. *Journal of policy modeling*, *37*(4), 668-692.

Datt, G., Simler, K. R., Mukherjee, S., & Dava, G. (2000). Determinants of poverty in Mozambique: 1996-97.

Diagne, A., Zeller, M., & Sharma, M. P. (2000). Empirical measurements of households'access to credit and credit constraints in developing countries: methodological issues and evidence (No. 583-2016-39550).

Dinh, T. T., Mai, N. N., Yen, N. N., Chung, N. V., & Thanh, H. N. T. (2022). Factors affecting household income in some northwestern provinces of Vietnam. *Journal of Economics and Development*, (305 (2)), 69-78.

Duyen, N. L. (2014). Factors affecting the income of farmers in An Giang.

Eshetu, F. (2020). Determinants of Rural Households Food Insecurity in Southern Ethiopia. *Asian Food Science Journal*, *15*(1), 29-43.

Fadipe, A. E. A., Adenuga, A. H., & Lawal, A. (2014). Analysis of income determinants among rural households in Kwara State, Nigeria. *Trakia Journal of Sciences*, *4*, 400-404.

Farzam, F. M., YÜCEL, R., & AKÇAKALE, C. Ş. (2021). The determinants of household income in Afghanistan: case study of Samangan province. *Journal of Humanities Insights*, *5*(03), 59-66.

Imam, M. F., Islam, M. A., & Hossain, M. J. (2018). Factors affecting poverty in rural Bangladesh: An analysis using multilevel modelling. *Journal of the Bangladesh Agricultural University*, *16*(1), 123-130.

Kabeer, N. (2015). Gender, poverty, and inequality: a brief history of feminist contributions in the field of international development. *Gender & Development*, *23*(2), 189-205.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, 33-50.

Loan, C. T. K., & Huong, N. V. (2015). The effect of resources on the income of farmers in Thanh Hoa province: a case study in Tho Xuan and Ha Trung districts. *Journal of Economics and Development, vol*., *13*, 1051-1060.

Ogutu, S. O., & Qaim, M. (2019). Commercialization of the small farm sector and multidimensional poverty. *World Development*, *114*, 281-293.

Ravindra, N. P., &Minuwanthi, R. W. W. K. (2020). Determinants of poverty: Is age non-linearly related with poverty? Evidence from Sri Lanka. *International Journal of Asian Social Sciences*, *10*(4), 181-192.

Scully, GW. 2002. Economic Growth and Income Inequality, Public Choices, Vol. 115:3, pp. 299-312.

Shahpari, G., & Davoudi, P. (2014). Studying effects of human capital on income inequality in Iran. *Procedia-Social and behavioral sciences*, *109*, 1386-1389.

Sien, L. V. (2021). Factors affecting the income of rice farming households in Tan Hong district, Dong Thap province. *Journal of Finance-Marketing Research*, 66-78.

Sun, L., Liu, J., & Chen, B. (2014). Analysis of Social Capital's Effect on Income of Poor Households: A Case Study in Sichuan Province. *Asian Agricultural Research*, *6*(1812-2016-144172), 36-40.

Teame, G. T., & Woldu, T. Y. M. (2016). Factors affecting rural households' income diversification: case of Zoba Maekel, Eritrea. *American journal of business, economics and management*, *4*(2), 7-15.