RESEARCH IN GLOBAL GOVERNANCE
No. 3 September 2018

The impact of non-farm diversification activities on economic well-being in rural Vietnam: A propensity score matching approach

Doctoral Degree Program Global Governance Program Enrollment in 2015
Do Ngoc Kien

【SUMMARY】

Using propensity score matching method and the 2014 Vietnam Access to Resources Household Survey, this study examines the determinants and welfare impact of the non-farm diversifying decision in rural Vietnam. The econometric estimation is implemented in two steps. First, we estimate the propensity score as the probability of participation in non-farm activities of households using a logistic model. We then match households with and without non-farm diversification based on these scores. The mean difference between two matched groups is the net impact. The results show that although the households specializing in the farming sector are enjoying a higher level of well-being, the participation in non-farm activities has a statistically positive impact on households' well-being. In our model, bonding social capital in terms of strong political ties, general trust and massive organization membership seemingly restricts the diversification process while communist party membership encourages the non-farm activities. Our discussion suggests that the farmers with abundant resources tend to reap the benefits from the farming sector and have less incentive to divert to non-farm activities.

【Key Words】 Propensity score, matching, non-farm diversification, social capital, social development

1. Introduction

The agricultural sector is the main source of income and employment in rural areas of Vietnam. According to the World Bank Development Indicators data, while the sector accounted for 18.1% of GDP of Vietnam, it employed around 42.8% of the total workforce in 2016. The poverty reduction effect of the agricultural sector is essential and much stronger than that of industrial and service sectors. The contribution of agricultural activities to households' economic well-being is crucial

The Graduate School Journal Committee Accepted: May 28, 2018

for those who are only involved in subsistence farming. Besides, the process of gradually liberalizing the economy and commercializing the agriculture affects the behavior of economic actors: when having a choice, farmers might either upgrade their agricultural production or diversify into non-farm economic activities. Davis (2004) defined the rural non-farm economic activities as ''comprising all those non-agricultural activities which generate income to rural households (including income in-kind and remittances), either through waged work or in self-employment''. Hence, if we divide all economic activities into 5 main categories: (1) job that paid, or wage; (2) farming production; (3) non-farm business activities such as trading and so on; (4) exploiting common resources in the common-resource areas and (5) housework, then any activity in (1), (3) and (4) can be named as non-farm. Non-farm activities could limit in rural areas or extend to urban-related ones depending upon the nature of the rural-urban interaction and its economic opportunities. Some households engage in rural entrepreneurship and trading services while others send their members to urban areas for salary seeking or jobs that paid. A household has a non-farm diversifying decision when it decides to be involved, partly or wholly, in non-farm economic activities. It would be worth noting that the concept of farming and agriculture here is used in general form, including activities in not only rice sector, but also forestry, and pisciculture. As a result, non-farm diversification does not include diversifying activities among these sectors, for example, from rice production to pisciculture or to fruticulture, which is termed ''among-farm'' diversifying activities.

In this paper, we examine the welfare impact of non-farm diversifying decisions in rural Vietnam using the propensity score matching method. Specifically, we estimate the mean difference in the economic well-being of the rural non-farm diversifying households and the counterfactual of themselves if they would remain solely specialized in farming production. This study contributes to the debate among Vietnamese policymakers whether a developing country with a solid foundation of agriculture like ours should specialize in the farming sector for gains from specialization or should diversify its economic activities to non-farm sectors for other welfare increase if any. The structure of this paper is as follows. In section two, we review the literature on non-farm diversifying decisions and its impacts. Section three introduces the propensity score matching method. We then set up the analytical framework and empirical model for the impact evaluation in section four. The fifth and final section closes the paper with some discussion and further research.

2. Literature Review

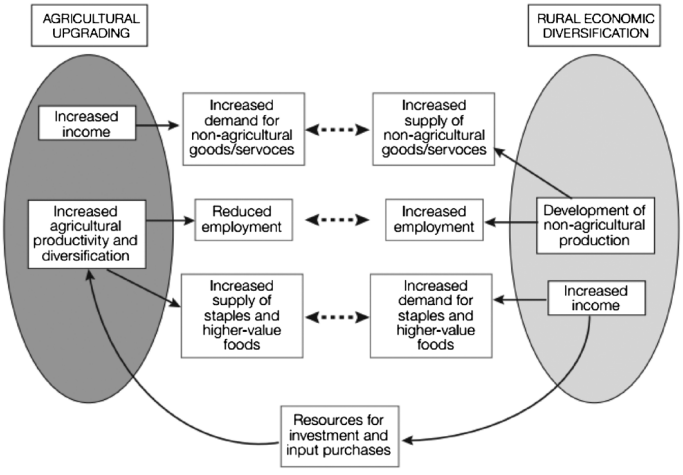
For decades, the literature of economic development has recognized a fundamental change in the source of growth in the developing countries, to be more specific, incomes from non-farm activities are becoming important in rural areas. The rise of non-farm activities in rural areas was noted and

widely discussed in the economic literature. A significant effort to understand the non-farm sector in rural areas was presented by Lanjouw and Lanjouw (2000). In this study, upgrading agriculture and diversifying to non-farm activities might be complimentary due to the dynamic interaction in economic activities. In a comprehensive study, United Nations Conference on Trade and Development [UNCTAD] (2014) also argued that agricultural upgrading and rural economic diversification did not play a zero-sum game. In contrast, upgrading agriculture and diversifying to non-farm activities might be complimentary due to the dynamic interaction in economic activities. The interaction and complementarity are visualized in Figure 1.

What's more, UNCTAD (2015) pointed out three main motivations for households when they engage in non-farm activities: entrepreneurship by choice, entrepreneurship by necessity and risk management. ''Entrepreneurship by choice'' means that some households with productive endowments choose to engage in more profitable non-farm activities in higher entry barrier markets for maximizing their economic well-being. ''Entrepreneurship by necessity'' denotes that some households with disadvantageous endowments find themselves insufficient in agriculture production and must engage in low return non-farm activities with low entry barriers for their survival. ''Risk management'' means that other households find agricultural production risky and prone to external shocks, especially in remote areas, thus they self-insure against risks through a variety of coping strategies of diversifying income sources. Therefore, non-farm diversifying activities is a strategy to reduce the vulnerability of rural households. As reviewed by Davis (2006), this motivation might dominate in transition economies or in an economy with a structural change.

In terms of inequality, the net effect of non-farm diversification might be positive or negative.

Figure 1: Complementarity of agricultural upgrading and rural economic diversification



Source: UNCTAD (2014)

Income inequality might increase due to non-farm diversifying because the unequal distribution of endowment determines the unequal distribution of economic opportunities and future incomes. The richer households tend to own more productive assets that would generate more income in the future when they extend to non-farm high return activities. Worse still, for the poorer households who do not own productive assets for generating sustained income, the demand for income security from an external shock forces them to engage in non-farm low return activities for diversifying their income. In other words, the poorer households sacrifice gains from specialization for coping with external shocks while the richer households sacrifice specialization for more gains. However, the increase in non-farm economic activities might also induce more investment and inputs supply for agricultural upgrading. It is important to note that the extension to non-farm activities leads to the increase in labor productivity in agriculture as illustrated in the macro-perspective by UNCTAD (2015) withia sample of 26 least developed countries or in the micro-perspective by Anang (2017). As a result, non-farm diversification might be the pro-poor strategy. The difficulty is that, at least from its phenomenon, it is impossible to observe the motivation of the non-farm diversification. For example, rural-urban migration is a strategy of rural households for diversifying their source of income. For some families, rural-urban migration is an unproductive activity for income security. For others, it is essential for reducing poverty, especially in the regions with poor climatic conditions. The welfare effect of non-farm diversification might be positive or negative and should be answered empirically (Lanjouw and Lanjouw, 2000).

In a study for the case of Hubei province in China, De Janvry et al. (2005) estimated the participation in the non-farm equation by the OLS model and income equation by the probit model using observed data of about 7,000 households. Using the predicted value from these estimations, they simulated the counterfactual of income distribution without the non-farm participation and compared with the fact. Their data suggested that around 72% of rural households obtained non-farm incomes, accounting for about 36% of their total income. Their results showed that non-farm incomes contributed significantly to improving the total household incomes, reduced poverty, and mitigated inequality.

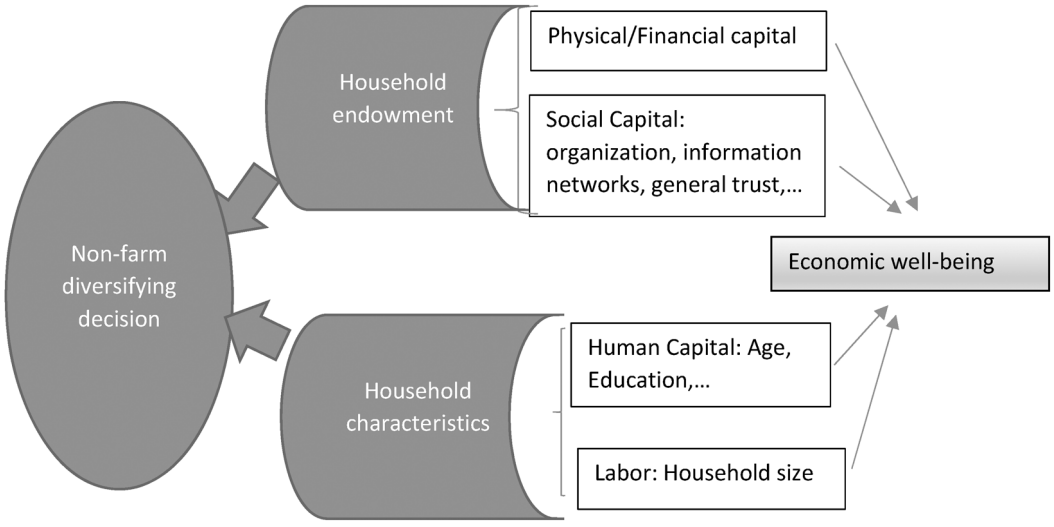
De Walle and Cratty (2004) explored this issue using the Vietnam Living Standard Survey (VLSS) in 1993 and 1998. They found a high association between non-farm diversification and economic well-being in the cross-section data. However, the correlation disappeared when they moved from the cross-section data to the panel data. The direct causal link between these two variables did not exist and what they found from the cross-section data is the common association with other factors. This viewpoint was supported by a case study of the village of Palanpur, Uttar Pradesh, India in Himanshu et al. (2013). The authors found that the nonfarm sector appeared to be breaking

down long-standing barriers to economic mobility among the poorest in rural India. Economic mobility might be the factor determining both income and diversifying activities. Rahut et al. (2015) used the 2012 Bhutan Living Standard Survey for estimating the determinants of non-farm participation decision, and the determinant of non-farm income with education is an endogenous variable. They presented that rural non-farm activities comprised 60.7% of rural household income in Bhutan, and this contribution increased with higher income and education levels. They confirmed that education was important for both participation in and generating income from the non-farm sector. Education might also be the factor that influences both income and non-farm diversification.

In a recent study, Newman and Kinghan (2015) examined the welfare effect of diversification in Vietnam. Including household fixed effects, time lags for wealth and income in the models, they regressed real consumption per capita with dummy variables of farm and several types of non-farm activities. The study found that the non-farm diversifying households enjoy about 10%–23% higher consumption per capita. However, this study lacked a solid theoretical framework while suffering from the selection bias in empirical analysis because the expected consumption in the period $t+1$ would influence the income and wealth in the period t through future expectation and motivation. In this case, only including time lags in the model did not solve the selection bias. As a consequence, the welfare impact could be inflated to the high estimates of 10%–23%.

In brief, the economic literature provides a mixed description of the relationship between non-farm diversification and economic well-being of households. However, the survey of literature is

Figure 2: The conceptual framework for diversifying decision



Source: Author's

relatively consistent in emphasizing several points in the following.

First, non-farm diversifying decision of a household depends significantly upon its characteristics and endowment. Second, the economic well-being of a household relies on the inputs or factor of production that it owns, including physical or financial capital, labor, human capital and social capital. We can integrate these two groups into one consistent framework in Figure 2.

3. Propensity score matching method

Impact evaluation theory points out several challenging in examining an economic decision due to the weakness of constructing counterfactual without randomized trial control. In particular, for examining the impact of non-farm diversifying, it is risky to simply compare the outcome of the diversifying households and non-diversifying ones. Gertler et al. (2016, p. 59) explained that ''selection bias will occur when the reasons for which an individual participates in a program are correlated with outcomes, even in absence of the program''. Hence, a rural household might decide to diversify into non-farm activities *because* she is rich (or poor). If we compare these two groups, we might simply compare a rich group and a poor group. The difference between them is the impact of not only non-farm diversifying but also many other factors and characteristics of households. As a result, in theory as well as in practice, ''ensuring that the estimated impact is free of selection bias is one of the major objectives and challenges for any impact evaluation'' (Gertler et al., 2016, p. 59).

From the regression analysis perspective, it would be biased to estimate the effect of participation (denoted as τ) into a program (D_i, the status of enrollment into the program, where $D_i = 1$ if enrolled, $D_i = 0$ if not enrolled) on outcome Y_i by comparing the mean outcome of the linear regression model $Y_i = \beta_0 + \beta_1 D_i + \beta_2 X + \varepsilon_i$ (where X is the vector of controlled, observed confounding variables, ε_i is error term) because without randomized trial control, the correlation between error terms ε_i and D_i is not equal to zero, D_i is not exogenous and τ is biased..

Impact evaluation theory introduces the matching method to overcome this obstacle. In principle, the matching method enables us ''to identify the set of non-enrolled individuals that look most similar to the treated individuals, based on the characteristics that [we] have available in [our] data set. These matched non-enrolled individuals then become the comparison group that [we] use to estimate the counterfactual'' (Gertler et al., 2016, p. 143). Technically, the matching method uses statistical techniques to construct an artificial comparison group: for every possible household diversifying to non-farm activities, it attempts to find a household of non-diversifying that has the *most similar characteristics possible*. However, when we increase the number of characteristics (or dimensions of matching), we may face with the so-called ''curse of dimensionality'': data may not

contain a good match for most of the program participants who are enrolled. Rosenbaum and Rubin (1983) developed an extension of the matching method for facilitating this challenge. They argued that under the certain assumptions, with a vector of pre-treatment characteristics of household X and defining the propensity score of X, $P(X)$, as the conditional probability of receiving a treatment given X, then matching on $P(X)$ is as good as matching on X. In particular, $P(X) = Pr(D = 1 | X) = E(D | X)$ where $D = \{0,1\}$ is the indicator of exposure to treatment. This propensity score is a real number between 0 and 1 that summarizes the influence of all the observed characteristics on the probability of participating in non-farm diversification. When the propensity score has been computed for all households, then the households that diversifying into non-farm activities can be matched with households in the pool of non-diversifying ones that have the closest propensity score. These closest households become the comparison group and are used to produce an estimate of the counterfactual. The average difference in economic well-being variable between the diversifying households and their matched ones produces the estimated impact of the participation behavior. Given a population of units, if the propensity score $p(X_i)$ is known, then the average treatment effect on the treated (ATT) can be estimated as follows:

$$\tau = E_i(Y_{1i} - Y_{0i} | D_i = 1) = E_{Xi}[E_i(Y_{1i} - Y_{0i} | D_i = 1,\ p(X_i))]$$
$$= E_i[E_{Xi}(Y_{1i} | D_i = 1,\ p(X_i)) - E_{Xi}(Y_{0i} | D_i = 0,\ p(X_i)) | D_i = 1]$$

4. Analytical Framework

4.1. Data

In this study, we employ the Vietnam Access to Resources Household Survey (VARHS) dataset. VARHS data is a comprehensive survey data implemented by the Central Institute for Economic Management (CIEM) of Vietnam, the Institute of Labor Science and Social Affairs (ILSSA) of Vietnam and the Institute of Policy and Strategy for Agriculture and Rural Development (IPSARD) of Vietnam under the co-support of the University of Copenhagen since 2002 and the United Nations University since 2010. After the first wave in 2002, since 2006 the survey has been implemented bienially in 12 provinces with an increasing sample from 2,324 in 2006 to 3,648 in 2014. Of the total samples, about 2,660 households have been traced since 2006. The survey presented detailed information about rural households for understanding their behavior, opportunities, and constraints. In this paper, we employed the 2014 VARHS dataset. We ignored all missing observations in any variable of interest in the 2014 dataset. We finally obtained a dataset with a sample of 2,956 households.

For the variable of well-being, we refer to Schaffner (2014, p.18) that defined a person's well-being as ''a summary assessment of how good or bad her life circumstances are, paying attention at

Table 1: Non-farm diversifying activities in Vietnam 2014

Economic activities generating income *Sample of N=2956*	Number of households	Percentage
Completely specialized in farming	1,161	39.28%
Involving in a job that paid	1,198	40.53%
Involving in non-farm business activities	412	13.94%
Involving in exploiting common resources in the common areas	999	33.80%

Source: Author's calculation from VARHS 2014 sample dataset

a minimum to the quantities and qualities of the goods and services she consumes, the activities to which she allocates her time, and her hopes and fears regarding the future''. From this understanding, it is impossible to measure economic well-being directly. Instead, ''the best we can do is make educated guesses about well-being effects by examining what happens to indicators of well-being, which are imperfect measures or correlates of well-being'' (Schaffner, 2014, p. 28). For these indicators, Schaffner (2014) suggested considering household income per capita, consumption expenditure per capita, direct assessment such as food consumption and asset measures. In this study, we employed four indicators: food expenditure of household, food expenditure per capita of household, the income of household and income per capita of household.

For non-farm diversification, in VARHS, the economic activities were asked for all household members in five main categories if they are involved in any: (1) Job that paid, or wage (2) Agriculture production of the household, or farming in general (3) Non-farm business activities such as trading and so on (4) Exploiting common resources in the common areas and (5) Housework. We ignored the housework in this study and regarded items (1), (3) and (4) as non-farm diversification capturing by the dummy variable *diversifying*. In our samples, we have 1,161 households that specialized in farming activities (non-diversifying household), while 1,795 households engaged in non-farm diversification activities.[1] An overview of non-farm diversifying activities in Vietnam is described in Table 1 while the mean comparison of food expenditure of diversifying and non-diversifying households is reported in Table 2.

Interestingly, Table 2 reports that on average the non-farm diversification household has a statistically lower level of economic well-being compared with those specialized in agriculture in terms of food expenditure (4.12%), food expenditure per capita (7.85%), income (7.86%) and income per capita (12.42%). Of course, these are simply the mean differences. The findings suggest

[1] Because the sample of specialized in farming households is less than the sample of non-farm diversifying households, we use non-farm diversifying households group as a pool of matching selection. Our quasi-intervention in this case should be, in fact, ''specialization in farming''

Table 2: Mean difference between diversifying and non-diversifying households

	Specialized in farming (thousand VND)	Non-farm diversifying (thousand VND)	Difference (thousand VND and percentage)	
Food expenditure (latest 4 weeks)	1,571.27	1,506.55	−64.71*	−4.12%
Food expenditure per capita (latest 4 weeks)	392.21	361.44	−30.77***	−7.85%
Income (12 months)	105,222.35	96,954.68	−8,267.67**	−7.86%
Income per capita (12 months)	26,218.77	22,962.01	−3,256.76***	−12.42%
Number of households	1161	1795		

*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Source: Author's calculation from VARHS 2014 sample dataset

that in our samples, on average, a household specializing in farming is enjoying a statistically higher level of economic well-being and this *does not* imply that non-farm diversification leads to a lower level of economic well-being. For estimating the difference in economic well-being *due to* non-farm diversifying activities, we set up empirical models in the next section.

4.2. Empirical model

To estimate the impact of diversifying activities, this paper employed the propensity matching method for capturing the difference in nature of households. The econometric estimation is implemented in the two following steps. First, we estimate the probability of participation in non-farm economic activities of rural households by a logit model for constructing propensity scores. We then match households with and without non-farm diversification based on these scores. The mean difference between two matched groups is the net impact.

Estimating the logit model

The logit model is constructed upon the conceptual framework in Figure 2 to ensure that it includes only variables that simultaneously affect the participation decision and the outcome variable (Caliendo and Kopeinig, 2005, p.6). The conceptual framework reflects both the characteristics and the endowment of the household. For *characteristics* of the household, we include age and education of the head of households. These variables could be seen as the proxies for human capital (age for experience and education for skills). The gender of the head is also included, as argued in Davis (2004). For *household endowment*, we focus on three types of household resources: (1) Labor, proxied by the household size; (2) Physical capital or assets, proxied by the value of total durable assets (in nature logarithm form), (3) social capital in four dimensions: Organizations (the number of household's members participating in important organizations including the communist party and other massive ones); Informal network (the number of people that households can ask

Table 3: Empirical models for non-farm diversifying decision

VARIABLES: Diversifying	Logit Model Odds ratio	Logit Model β	Probit Model β
Gender of the head	1.060	0.0584	0.0319
Household size	1.062**	0.0599**	0.0370**
Age of the head of household	0.861***	−0.149***	−0.0906***
Age square	1.001***	0.00125***	0.000756***
Education of the head	0.800***	−0.224***	−0.133***
Education square	1.011***	0.0114***	0.00675***
Number of Communist Party members	1.357***	0.305***	0.183***
Number of other massive organizations members	0.925*	−0.0775*	−0.0481*
Informal network	0.994	−0.00615	−0.00393
General trust	0.681***	−0.384***	−0.230***
A household member is a leader	0.684**	−0.380**	−0.227**
A relative is a leader	1.166	0.154	0.0993
A Friend is a leader	0.931	−0.0714	−0.0468
The total value of the durable asset (logarithm form)	0.937**	−0.0652**	−0.0410**
Constant	497.4***	6.209***	3.780***
Observations	2,956	2,956	2,956

*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Source: Author's estimation

for help in case of cash need); general trust; and political connection (whether a household member is a leader, a relative person is a leader or a friend is a leader).

Unlike the OLS regression model, the logit regression model helps avoid the case in which the predicted value (or propensity score) would be less than 0 or greater than 1. In the logistic regression model, we estimate the logistic transformation of the probability that the households with certain characteristics would diversify into non-farm activities. The probability of the households with observed characteristics for diversifying is given to each household as a propensity score. The results are reported in Table 3. The first column is the independent variable. We report logistic regression model in terms of odds ratio in column 2 and in terms of the traditional coefficients in column 3. Column 4 presents the estimation from the probit model for comparing purpose. The coefficients estimated by the logit regression model and probit regression model are reasonably similar, implying that our estimation is consistent. Besides the comparison with the probit model, we also employ several diagnostics for the logistic model. We test the model specification using the link test. The variable *_hat* with z-value = 7.68 is significant, implying that our model is a good fit, while variable *_hatsq* with z-value = −0.45 is not significant, indicating that the model is correctly specified. As recommended by Mehmetoglu and Jakobsen (2017, p.178), we also check the overall goodness-of-fit test to see if the observed values on the dependent variable match the expected values, both for number of covariate patterns [Pearson $\chi^2(2941) = 2961.06$, p-value = 0.394 > 0.05]

and for a set number of groups of 10 [Hosmer-Lemeshow $\chi^2(8) = 3.12$, p-value $= 0.927 > 0.05$]. The results show that we cannot reject the model, or our model fits the data well. For testing for multicollinearity, we regress that dummy variable diversifying with dependent variables and check variance inflation factor (VIF). Except for the variables with square terms in the model, all other tolerance values (1/VIF) are greater than 0.2, which is the ''rule-of-thumb'' level to conclude that there is no evidence of multicollinearity in the model.

Table 3 shows that almost all independent variables in our logistic regression model statistically explain the decision of non-farm diversification in rural Vietnam, except for the gender of the head of households and two types of social capital (i.e. informal network and weak political connection). While the coefficient magnitude is difficult to explain in the logistic model, the sign of the coefficients matters.

From our estimation, the household size has a positive relationship with the probability of non-farm diversifying. With more labor, it is easier for the household to expand its activities beyond farming. The age and education of the head of the family have a negative relation with the diversification. It means, on average, the younger and less educated tends to diversify more than the older and the more educated people, or the more experienced and knowledgeable households tend to specialize in farming activities. These results are different from the findings by De Walle and Cratty (2004) and many other studies which argued that high education enhances the probability of non-farm diversifying. The squared coefficients of both variables are positive, which means that the speed of change is quicker for the younger. This might imply that the younger and less educated do not want to work in the farming sector, regardless of the productivity of this sector.

Households with more communist party members are more willing to divert to non-farming activities than households with fewer party members. In contrast, households with more members of other massive organizations are less likely to diversify compared to households with less membership. One explanation is that the massive organizations, such as farmer unions, women organizations and so on, in rural Vietnam have a close relation with the farming sector.

General trust has a negative relation to the diversifying decision. Those who trust their rural communities tend to stay at farming rather than diversify into non-farm activities. It might be reasonable because farming is somehow a communal activity. The rich and politically powerful households are more likely to stay in farming activities. These results imply that farming is still attractive and might be productive for one who possesses good knowledge and assets for utilization.

Table 4 presents the result of an impact evaluation of non-diversifying activities using the matching method with several propensity score matching techniques in both logistic model and probit model. Our estimation results are consistent when all matching techniques produce very similar

Table 4: Impact of non-diversifying on economic well-being

Variable Sample	Specialized ('000 VND)	Diversifying ('000 VND)	Difference ('000 VND)	Difference (%)
Unmatched food expenditure	1571.27	1506.55	−64.71	−4.12
Unmatched food expenditure per capita	392.21	361.44	−30.77	−7.85
Unmatched income	105,222.35	96,954.68	−8,267.67	−7.86
Unmatched income per capita	26,218.77	22,962.01	−3,256.76	−12.42
ATT in the logistic model: Food Expenditure				
Kernel	1466.54	1506.55	40.01	+2.73
Kernel (biweight)	1465.83	1506.55	40.73	+2.78
Local linear regression	1459.63	1506.55	46.92	+3.21
Radius with caliper $\varepsilon=0.25$ $\sigma_P=0.0334$	1465.02	1506.55	41.54	+2.84
ATT in the probit model: Food Expenditure				
Kernel	1466.51	1506.55	40.05	+2.73
Kernel (biweight)	1465.68	1506.55	40.87	+2.79
Local linear regression	1459.37	1506.55	47.19	+3.23
Radius with caliper $\varepsilon=0.25$ $\sigma_P=0.0333$	1464.87	1506.55	41.68	+2.85
ATT in the logistic model: Food expenditure per capita				
Kernel	355.90	361.44	5.54	+1.56
Kernel (biweight)	355.60	361.44	5.84	+1.64
Local linear regression	355.22	361.44	6.23	+1.75
Radius with caliper $\varepsilon=0.25$ $\sigma_P=0.0334$	355.43	361.44	6.01	+1.69
ATT in the probit model: Food expenditure per capita				
Kernel	355.71	361.44	5.73	+1.61
Kernel (biweight)	355.39	361.44	6.05	+1.70
Local linear regression	355.24	361.44	6.20	+1.69
Radius with caliper $\varepsilon=0.25$ $\sigma_P=0.0334$	355.31	361.44	6.13	+1.73
ATT in the logistic model: Income				
Kernel	94,863.30	96,954.68	2,091.38	+2.20
Kernel (biweight)	94,774.73	96,954.68	2,179.95	+2.30
Local linear regression	94,214.50	96,954.68	2,740.18	+2.91
Radius with caliper $\varepsilon=0.25$ $\sigma_P=0.0334$	94,758.75	96,954.68	2,195.93	+2.32
ATT in the probit model: Income				
Kernel	94,859.33	96,954.68	2,095.34	+2.21
Kernel (biweight)	94,769.00	96,954.68	2,185.67	+2.31
Local linear regression	94,193.31	96,954.68	2,761.36	+2.93
Radius with caliper $\varepsilon=0.25$ $\sigma_P=0.0334$	94,779.35	96,954.68	2,175.32	+2.30
ATT in the logistic model: Income per capita				
Kernel	22744.93	22962.01	217.08	+0.95
Kernel (biweight)	22719.23	22962.01	242.78	+1.06
Local linear regression	22725.72	22962.01	236.29	+1.03
Radius with caliper $\varepsilon=0.25$ $\sigma_P=0.0334$	22719.15	22962.01	242.86	+1.07
ATT in the probit model: Income per capita				
Kernel	22742.13	22962.01	219.88	+0.97
Kernel (biweight)	22716.11	22962.01	245.90	+1.08
Local linear regression	22733.52	22962.01	228.49	+1.01
Radius with caliper $\varepsilon=0.25$ $\sigma_P=0.0334$	22723.30	22962.01	238.71	+1.05

Source: Author's estimation

outcome. We also employed several balancing tests for measured covariates with our matched data, which was not shown here fore brevity. We found that we had similar covariate distributions and that there was no statistically significant difference in means between treatment and control groups for all individual covariates. These diagnostics imply that our estimations are efficient.

As mentioned in section 4.1, without matching, we observe that on average the households who diversify to non-farm activities enjoy a 4.12% lower in food expenditure, 7.85% lower in food expenditure per capita, 7.86% lower in income and 12.42% lower in income per capita compared with ones who specialize in agriculture. However, after matching, the non-farm diversification activities have a positive impact on all well-being indicators. Depending upon the model and well-being indicators, the net impact of diversifying is about 3% in food expenditure, 1.7% in food expenditure per capita, 2.3% in income and 1% in income per capita.

5. Conclusion, discussion and further research

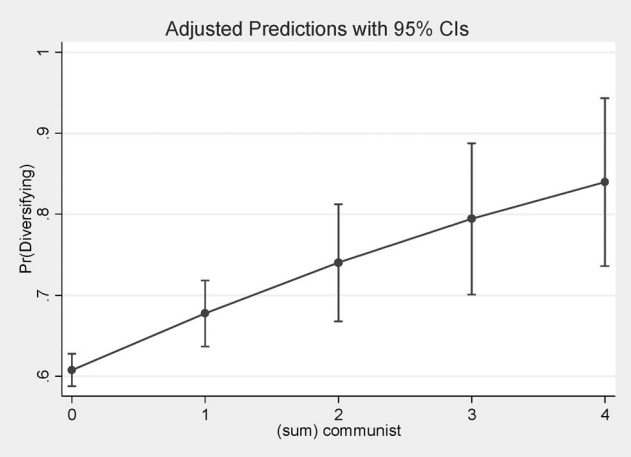
While the impact of non-farm diversifying activities is ambiguous in the literature, our study provides evidence that the diversification helps increase the economic well-being of rural households in Vietnam. Our result rejects the hypothesis that the country should focus its resources in the farming sector for the gain from specialization. The findings implicitly indicate that the strategy of ''leaving the farm without leaving the hometown'' inspired by the government might be a pathway for rural development in Vietnam. For this reason, the government should continue promoting this strategy in the next phase of the rural development strategies.

This study also provides some interesting perspectives on the role of social capital in economic policy formulation. First, the communist party membership is one of the important factors that affect the non-farm diversifying activities of the household. A household with more communist party members tends to have statistically higher probability to engage in non-farm diversification. Figure 3 illustrates this viewpoint using estimated probabilities when a non-farm diversifying household has from 0 to 4 communist party members given other variables fixed at their means.

This result at a glance seems strange. The Vietnamese communist party had a tradition of being against private economic activities and discouraged its members to do business in the past, especially before the Doi-moi in 1986. The different approach in this matter newly introduced in the 10th National Congress of the party in 2006 when most representatives agreed that party members should be able and encouraged to pursue private economic activities.

One might argue that the communist party membership is a kind of social network so the membership might be assumed to facilitate the transmission of information about the market. Therefore, the party membership might have an advantage in doing business with more market informa-

Figure 3: Estimated probabilities of diversifying with communist members



Source: Author's estimation

tion. However, by definition, the communist party is not a channel for economic information transmission. Indeed, CIEM et al. (2015, p. 325) reported that the major sources of market information came mainly from relatives, friends, neighbours (70.2%), local market (59.1%) and television (60%) while market information from other groups and massive organizations accounted only for 14%.

Does our result reflect the more dynamic of the communist party members in the business matter since the 10th national congress? To test this hypothesis with our dataset, we check the rate of households with communist member(s) engaging and not engaging in each type of non-farm activities. For job paid, this rate is 0.93. For exploiting common resources such as hunting or fishing in the common areas, this rate is 0.414. However, for non-farm business activities, this rate is 0.141. The lowest rate in non-farm business activities does not support our hypothesis about the more dynamic of the communist party members. Instead, the highest rate in the job paid activities supports the view that the communist party has its root in the worker class, so the probability that we observe a household with communist party members engaging in a job paid (one type of non-farm diversifying) might be higher. In other words, the positive relationship between the number of party members in the household and probability of non-farm diversifying results from the worker-class nature of the Vietnamese communist party instead of any party membership's information advantage or the effect of the 10th national congress.

The negative relationships generated from the general trust, massive organization membership such as women associations or farmer unions and having a family member being the leader imply that these types of social capital tend to keep the households specializing in agricultural activities.

These results support the hypothesis of the downside of bonding social capital widely discussed in the social capital literature (Portes, 2014). Excessive trust in the community and strict bonding to homogenous groups might become a social liability to the members and they may miss many better economic opportunities that have weak ties with the community or groups. In other words, while we cannot find out "the strength of the weak ties" (both variables that a relative is a leader and a friend is a leader are insignificant), we have evidence of the weak of the strong ties.

The negative impact of age and education to non-farm diversifying contrasts with several results in the economic literature (Davis (2006), Rahut et al. (2015) and De Walle and Cratty (2004)). De Walle and Cratty (2004) employed the 1993 VLSS and 1998 VLSS of Vietnam and pointed out that both age and education had a strong effect on the non-farm diversification. We explain this different result between the 2014 VARHS survey data and the 1993/1998 VLSS survey data by the dissimilar viewpoint between the parent and children generation.

Another way to discuss our results is that farming, as a kind of community activities, is seemingly considered productive. The farmers with abundant resources (social capital such as more social network, high social trust, strong political connection to leaders; human capital such as experience and education; physical capital or assets) can still well benefit from this sector, so they have lower motivation for diversifying into non-farm activities.

The quantitative methods used in this study cannot clearly show the underlying reasons. Hence, further research should be complemented with qualitative methods such as the case study or ethnographic method for in-depth understanding of the diversification into non-farm activities. The combination of qualitative methods with the propensity score matching method in the impact evaluation with a quasi-experiment intervention will be helpful in testing theoretical conditions and assumptions of the quantitative analysis. The qualitative method might also help understand the policy impact on the variable of interest, such as the viewpoint and policy of the communist party on rural and agriculture. Besides, a more concrete and formal measurement of social capital should be used in future research. We also recommend exploring the endogenous nature of social capital with other determinants in a new conceptual framework. Finally, when the data is updated and publicly available, one can integrate the 2014 VARHS and 2016 VARHS into a single panel dataset for exploring the combination between difference-in-differences (DID) and propensity score matching method. In this case, the models can deal with all time-invariant unobservable variables, such as policy instruments that do not change between 2014 and 2016, especially when the communist party's policies on rural areas, agriculture and farmers are supposed to have a significant effect. This technique will provide a better estimation and from that result, one can check the problem of causality noted in De Walle and Cratty (2004).

References

Anang, B. T. (2017). Effect of non-farm work on agricultural productivity: Empirical evidence from northern Ghana. WIDER Working Paper 2017/38. Retrieved from https://www.wider.unu.edu/sites/default/files/wp2017–38_0.pdf

Caliendo, M. & Kopeinig, S (2015). Some practical guidance for the implementation of propensity score matching. IZA Discussion Paper No. 1588. Retrieved from http://ftp.iza.org/dp1588.pdf

CIEM, IPSARD and ILSSA (2015). *Characteristics of the Vietnamese rural economy: Evidence from a 2014 rural household surveys in 12 provinces of Vietnam*. Hanoi: Hong Duc. Retrieved from http://www.ciem.org.vn/Content/files/danida/Sach_DDKTNT%20-%20final.rar

Davis, R. J. (2004). The rural non-farm economy, livelihoods and their diversification: Issues and options. Working Paper. Natural Resources Institute, Chatham, UK. doi: 10.2139/ssrn.691821

Davis, R. J. (2006). Rural non-farm livelihoods in transition economies: emerging issues and policies. *Journal of Agricultural and Development Economics, 3*(2),180–224. Retrieved from http://www.fao.org/3/a-ah756e.pdf

De Janvry, A., Sadoulet, E.& Zhu, N. (2005). The role of non-farm incomes in reducing rural poverty and inequality in China. http://are.berkeley.edu/～esadoulet/papers/RNF-Nong.pdf

De Walle, V. D. & Cratty, D. (2004). Is the emerging non-farm market economy the route out of poverty in Vietnam? *Economics of Transition*, 12: 237–274. doi:10.1111/j.0967–0750.2004.00178.x

Gertler, P., Wilde, S. M., Premand, P., Rawlings, L., & Vermeersch, C. (2016). *Impact Evaluation in Practice*. Washington, DC: World Bank Group. doi:10.1596/978–1–4648–0779–4

Grootaert, C., Van Bastelar, T. (2002). *Understanding and Measuring Social Capital: A Multidisciplinary Tool for Practitioners*. Directions in Development. Washington, DC: World Bank.

Guo, S., & Fraser, M. W. (2015). *Propensity Score Analysis: Statistical Methods and Applications* (2nd ed., Vol. 11, Advanced Quantitative Techniques in Social Sciences Series). Los Angles, London, New Delhi, Singapore, Washington DC: Sage Publications.

Himanshu, Lanjouw, P., Murgai, R. & Stern, N (2013). Non-farm diversification, poverty, economic mobility and income inequality: A case study in village India. *Agricultural Economics 44*, 461–473. doi: 10.1111/agec.12029

Khai, L. D., Kinghan, C., Newman, C. & Talbot, T (n.d.). Non-farm income, diversification and welfare: Evidence from rural Vietnam. Working Paper. Retrieved from http://www.ciem.org.vn/Portals/1/CIEM/IndepthStudy/VARHS12_diversification_welfare_final.pdf

Lanjouw, J. O. & Lanjouw, P. (2000). The Rural Non-Farm Sector: Issues and Evidence from Developing Countries. *Agricultural Economics, 26*(1), 1–23. Doi:10.1016/S0169–5150(00)00104–3

Liu, X. (2016). *Applied ordinal logistic regression using Stata*. Los Angles, London, New Delhi, Singapore, Washington DC: Sage Publications

Mehmetoglu, M. & Jakobse T. G. (2017). *Applied statistics using Stata: A guide for the social sciences*. Los Angeles, London, New Delhi, Singapore Washington DC, Melbourne: Sage Publications

Newman, C. & Kinghan, C. (2015). Economic Transformation and the Diversification of Livelihoods in Rural Viet Nam. WIDER Working Paper 2015/064. Retrieved from https://www.wider.unu.edu/sites/default/files/wp2015–064.pdf

Rahut, D. B., Jena, P. R., Ali, A., Behera, B., & Chheri, N. B. (2015). Rural Nonfarm Employment, Income, and Inequality: Evidence from Bhutan. *Asian Development Review, 32*(2), 65–94. doi:10.1162/adev_a_00052

Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70(1)*, 41–55. doi: 10.2307/2335942

Portes, A. (2014). Downsides of social capital. *Proceedings of the National Academy of Sciences of the United States of America, 111*(52), 18407–18408. doi: 10.1073/pnas.1421888112

Schaffner, J. (2014). *Development Economics: Theory, empirical research, and policy analysis*. Wiley

United Nations Conference on Trade and Development (2014). *The Least Developed Countries Report 2014: Growth with structural transformation: A post-2015 development agenda*. Retrieved from http://unctad.org/en/PublicationsLibrary/ldc2014_en.pdf

United Nations Conference on Trade and Development (2015). *The Least Developed Countries Report 2015: Transforming Rural Economies*. Retrieved from http://unctad.org/en/PublicationsLibrary/ldc2015_en.pdf

United Nations University (2017). *Characteristics of the Vietnamese rural economy: Evidence from a 2016 rural household survey in 12 provinces of Viet Nam*. Retrieved from https://www.wider.unu.edu/sites/default/files/Publications/Report/PDF/VARHS16-report-2017.pdf