Doctoral Thesis

University of Trento

School of Social Sciences

Doctoral Program in Economics and Management

Essays on Farm Household Decision-Making: Evidence from Vietnam

a dissertation submitted to the doctoral school of economics and management in partial fulfillment of the requirements for the Doctoral degree (Ph.D.) in Economics and Management

Vu Minh Hien

November 2013

Advisors

Prof. Christopher L. Gilbert
University of Trento

Prof. Gabriella Berloffa
University of Trento

Evaluation Committee

Prof. Alexander Sarris
University of Athens

Prof. Giuseppe Folloni
University of Trento

Final Examination Committee

Prof. Guglielmo Chiodi
University of Rome "La Sapienza"

Prof. Riccardo Fiorentini
University of Verona

Prof. Emanuele Taufer
University of Trento

Acknowledgements

Foremost, I would like to express my sincere gratitude to Prof. Christopher L. Gilbert. I could not have asked for better guidance, support and encouragement.

I am also grateful to Prof. Gabriella Berloffa, who kept patient to provide me thoughtful insight and critical comments.

I also would like to thank Prof. Alexander Sarris, Prof. Giuseppe Folloni, Prof. Guglielmo Chiodi, Prof. Riccardo Fiorentini, and Prof. Emanuele Taufer, for serving on examination committee.

A special thank goes to all people in the program and university of Trento for a warm environment.

I am glad to mention my friends who have shared and spiced up my life in Trento.

I especially thank my family for their unconditional care and support.

Abstract

This thesis contains three studies which provide theoretical analysis and empirical evidence on the decision-making of farm households under shocks and imperfect markets in Vietnam.

The first study attempts to investigate the effects of the 2007-08 global food crisis on the investment, saving and consumption decisions of household producers by using the panel data of the Vietnam Household Living Standard Survey (VHLSS), covering 2006 and 2008. The results show that the high food prices had a positive effect on only fixed asset investments in the period of the crisis. When the price shocks are incorporated in the financial conditions, the findings reveal that the effects of household incomes, loans obtained and land sizes matter.

The second study uses the Vietnam Access to Resources Household Survey (VARHS) of 2010 to assess the determinants of chemical fertiliser adoption for rice cultivation, and effects on productivity and household welfare. The analysis implements both nonparametric (propensity score matching) and parametric (instrumental variables) approaches. The findings show determinants affecting decision of adoption differ from those affecting decision of adoption intensity. The results show unsurprisingly positive impact on outcomes, but focus on advantage of using parametric approach to estimate these impacts.

The third study employs a sub-sample from the 2008 VHLSS that is restricted to rural areas and to children from 10 to 14 years old to explore the relationship between farmland and the employment of children on their family's farm. The hypothesis is tested in three models (the Tobit, Heckit and double-hurdle models), in which the dependent variables are examined for two stages of decision-making, including the probability of participation and the extent of participation. Empirical evidence supports the hypothesis that child labour increases in land-rich households and decreases in land-poor households.

Keywords: Farm Households, Production, Shocks, Imperfect Markets

Contents

Abstract **3**
Contents **4**
Chapter 1: Introduction **7**
1.1. Farm household's decision-making in developing economy 7
1.2. The context of Vietnam 10
1.3. Objectives and structure of the thesis 16
References 18
Tables 21
Figures 22
Chapter 2: The effects of food crisis on productive investment, saving and consumption **27**
2.1. Introduction 27
2.2. Theoretical model 31
2.3. Context of the crisis 33
2.4. Descriptive statistics 36
2.4.1. Survey data 36
2.4.2. Production 37
2.4.3. Investment and household welfare 38
2.4.4. Price index 40
2.5. Empirical approach 41
2.6. Estimation results 43
2.6.1.The effects on investment, saving and consumption in general 43
2.6.2. The effects in relation to market imperfections 44
2.7. Conclusion 47
References 49
Tables 52
Figures 65
Appendix 73
Chapter 3: Agricultural technology adoption and natural disasters **76**
3.1. Introduction 76
3.2. Background of the case study and the objectives 78
3.2.1. Rice cultivation and chemical fertilizer adoption in Vietnam 78
3.2.2. Agro-ecological conditions and natural disasters 80
3.2.3. Literature reviews and objectives of the study 82
3.3. Methodology 83
3.3.1. The two-stage models 83

3.3.2. Problem of impact evaluation 85
3.3.3. Propensity score matching (PSM) with binary treatment 88
3.3.4. Generalized propensity score matching with continuous treatment 94
3.3.5. Instrumental variables method 98
3.4. Data and model specification 103
3.4.1. Natural conditions and shocks 103
3.4.2. Outcomes by treatment status 104
3.4.3. Observable characteristics by treatment status 106
3.4.4. Instrumental variables 107
3.5. Application and results 108
3.5.1. Heckman tests 109
3.5.2. Matching methods 111
3.5.3. Instrumental variables 118
3.5.4. Heterogeneity effects 119
3.6. Conclusion 120
References 123
Tables 126
Figures 181
Chapter 4: Farmland and child labour 190
4.1. Introduction 190
4.2. Literature review 192
4.3. Theoretical framework 194
4.4. Survey and descriptive statistics 197
4.4.1. Basic information in the survey 198
4.4.2. Land, labour and product markets 200
4.4.3. Child labour situation 202
4.4.4. Variable summary 204
4.5. Econometric specifications 205
4.5.1. The standard Tobit model 206
4.5.2. The generalized Tobit or Heckit 207
4.5.3. The double-hurdle model 208
4.5.4. Test for model appropriateness 209
4.6. Empirical results 210
4.7. Conclusion 214
References 215
Tables 225
Figures 238

Chapter 5: Concluding remarks 239
5.1. Contributions 239
5.2. Limitations and future research.. 242

Chapter 1

Introduction

1.1. Farm household's decision-making in developing economy

In developing countries, agriculture serves as a major source of income for the population. At macro level, agriculture ensures the food security of the country and is an important channel of foreign exchange earnings, which significantly contribute to the economic growth. In order to boost the agriculture sector and improve household welfare as well, it is ultimately essential to capture behavior at household level as the key unit in the sector. One of the underlying traits is that farm households constitute the so-called semi-commercialized economy, in which they produce partly for their own consumption and partly for sale in the market. Inherent to other characteristics of the low-income economies, where the markets are missing or incomplete and insecure under uncertainty, the decision-making patterns of farm households are not straightforward. Therefore, modeling agricultural households has been centered in understanding problems of the agricultural economy and evolved in a long tradition.

This section briefly reviews the major trend of literature on farm household production in developing economy setting. Toward the objective of the thesis, I focus on the recent works that investigate how farm households allocate resources in for production choices in relation with shocks and imperfect markets. With various directions in the literature, I group roughly into three main tendencies of agricultural household model (Mendola, 2007). This division basically relies on the similarity in the set of assumptions in each approach.

Profit-maximizing model

Schultz (1964) proposes a crucial hypothesis that farmers in developing countries "efficient but poor" which has been widely both influential and controversial. Schultz argues that the low income levels in developing countries are resulted from low productivity of the availability factors of production, and not from inefficiencies of their allocation. In this sense, the farmers behave to maximize the profit and in the condition of perfect competition. Thereby the principal policy implications are that the outsider experts (extension agents, farm advisers, etc.) could not help farmers increase their productivity by advising them to reallocate resources, but to invest in education to facilitate news factors and hence improve productivity.

The key contribution of Schultz (1964) is not about efficiency that attribute to long-term equilibrium, but about his recognition of farmers' willingness to adapt and innovate in response to changing economic conditions and opportunities (Ball & Pounder, 1996). On the other side, the hypothesis has drawn numerous criticisms which mostly imply that the poor in developing countries are characterized by particular features and dimensions. Myrdal (1968) argues that the people in developing countries are not primarily determined by motivation of costs and benefits. Recently, Ray (2006) develops the idea of an aspiration window. The window is formed of a cognitive world and may be multidimensional, i.e. some could aspire to a better material standard of living, but there are other aspirations: dignity, good health, recognition, etc. These aspirations may complement one another, or they may be mutual substitutes. E. Duflo (2006) also shows a tendency of new hypotheses incorporating insights of psychology to better understand economic decisions. By analyzing some empirical evidence, Duflo asserts that the poor do not always make choices that are in their best interest in the long run. Hence those arguments present a central contradict point at which farm households behave rationally toward benefit of outcome or other aspirations in the context of developing economies.

Utility-maximizing model

The utility maximization approach take into consideration of the dual-character of the farm household which the consumption and production decisions are interdependent. The pioneer effort is often credited to Chayanov's work in 1920s. He formulates the theoretical model of the peasant economy where he doubts about the profit-maximizing canons of either Marxist or neoclassical theories (Millar, 1970). Chayanov shows that the peasant farms are strictly family operated and the family labor allocation is found to be directly related to the ratio of consumers over workers in the family. The economic differences among peasants are more likely associated with the family size and composition rather than differential economic success. These features reflect an inconsistent approach to the hypothesis that farmers manage their farm to maximize benefits.

Becker (1965) further extends this trend to model household decision and allocation resource by considering household as both a producing and consuming unit. The household maximize utility through consumption of all available goods including household producing goods, market purchasing goods and leisure, and which is subjected to the constraint of full income. Based on the assumptions that markets are perfect and goods are tradable, prices are exogenous and production and consumption are decided independently. In this condition, allocation of family labor is connected to market-determined wage since leisure time and working time are independent. While income is associated with both production and consumption. This decision making process is referred as recursive or separable model.

Hence the recursive model is valid when markets are perfect. In the absence of labor market, for instance, the consumption and production decisions become dependent, and the separability property disappears. This fact pervasively occurs in developing countries where markets are incomplete or missing. Thereby de Janvry, Fafchamps, and Sadoulet (1991) develop a notable framework on modeling farm household behavior with missing markets. The household setting is still utility maximization of all commodities but bounded from constraints induced by missing markets. In this study, the authors show that in the situation of imperfect labor market and commodity market (i.e. food scarcity due to bad weather), households have to internally compensate by adjusting their labor allocation and commodity consumption. Then two set of decisions are linked through the endogenous price which satisfies the equilibrium between supply and demand. Vastly empirical studies provide evidence to support this approach.

Risk-averse model

The above approaches are criticized for ignoring the prospect of uncertainty and assume that households are risk-neutral. Since the agricultural households are inherently exposed to uncertainty of risk and shock events, such as natural hazards, market volatility, policy changes, etc. The effects of those uncontrollable factors play an important role in agricultural production. Motivated by these factors, another line in literature models the behavior of farm households in which uncertainty and risk are involved in the process of decision-making. Regarding the context of risk, expect utility, which was elaborated by von Neumann and Morgenstern (1944) from the initial work of Bernoulli (1738), has been the principle for decision-making theory. The early exposition of expected theory to agriculture can refer to the work of Dillon (1971). Derived from Bernoulli's principle, the study provides recognition of the personal nature of decision making in terms of belief and preferences, and that could represent the best possible to risky choice in agriculture. Newbery and Stiglitz (1981) contribute on commodity price stabilization issues, and explore the problems of risk in agriculture. In general, there are two approaches to capture the basic ideas of these works. First, the expected utility model allows household to make choices given preferences of outcomes and beliefs on possibility of occurrence. Second, the households are risk-averse, meaning that they prefer low risky choice than the high risky alternatives.

However, due to limited scope, the analysis of this thesis is relevant to the first two farm household approaches, which are profit-maximizing model and utility-maximizing model. The theoretical frameworks are developed and modified to appropriately explain and examine specific circumstance in each chapter. Therefore, the following part provides an overview of this case study of Vietnam that characterized by agricultural economy and its own features.

1.2. The context of Vietnam

An overview

Vietnam is located on the Indochina Peninsula in Southeast Asia. The country is bordered by China to the north, Laos to the northwest, Cambodia to the southwest and the East Sea to the east. It covers a total area of approximately 331,210 km^2. Vietnam is a country of tropical lowlands, hills, and densely forested highlands. Mountains account for 40 percent of the country's land area, and tropical forests cover around 42 percent, and arable land accounts for around 20 percent.[1] Due to a long shape, climate and seasons vary from locality to locality. Northern Vietnam has four distinctive seasons: spring, summer, autumn and winter. Southern Vietnam has only two seasons: dry and rainy. The monsoon season which brings high temperature, typhoons and heavy rain, is from June to November.

According to the census in April 2009, the population of the country is approximately 85.8 million of which male and female respectively account for 49.5 percent and 50.5 percent (GSO, 2009). The rural population comprises 70 percent. Over the period of 1999-2009, the population rose annually by 947,000 people, or equaled to 1.2 percent. Around 66 percent of population fall into group between ages 15 and 60 years. This suggests that Vietnam is young population.

Vietnam has 58 provinces and 5 cities, of which Hanoi is capital, and together with Hochiminh city are the most important cities. Those 64 provinces are categorized into 8 agro-ecological regions including: the Red River Delta, the Northeast, the Northwest, the North Central Coast, the South Central Coast, the Central Highlands, the Southeast and the Mekong River Delta. The country has 54 ethnic minority groups, and Kinh ethnic is the largest group and constitutes 86 percent of the population (GSO, 2009).

[1] http://www.indexmundi.com/vietnam/geography_profile.html

Economic history

Vietnam is currently classified as a developing economy. The Sixth Party Congress in December 1986 embarked an economic reform program known as *Doi Moi*, or renovation in Vietnam. This landmark has made a shift from a centrally planned economy to a socialist-oriented market economy. Then the economy has experienced rapid growth. Market forces were introduced and central planning was eliminated for all but essential commodities. Vietnam rejected a model of heavy industrialization in favor of sectors in which had a comparative advantage: agricultural commodities (rice, coffee, rubber, cashew, etc.), natural resources (oil and natural gas) exploitation, and labor-intensive manufacturing. The new mechanism has also accelerated establishment of private businesses and foreign investment, including foreign-owned enterprises. Efforts were also made to improve the managerial skills of government officials and quickly facilitate decentralizing planning. This was important and necessary for Vietnam since it lost the entire Soviet aid after the collapsing of USSR in 1991. Then Vietnam became a member of ASEAN and the United States removed its trade and aid embargo in middle of 1990s.

The reforms succeeded partially. The economy slowed down by the early-1990s. Substantial cuts of subsidies for state enterprises, the elimination of the central state price system, and higher devaluation of the Vietnamese currency resulted in a short-term economic recovery. Vietnamese authorities determined slow process of the structural reforms to refresh the economy and produce more competitive, export-driven industries.

Vietnam's economy has been followed the 1997 financial crisis in East Asia recession and the policy had been opted to emphasize macroeconomic stability rather than growth. While the country has shifted toward a more market-oriented economy, the Vietnamese government still has continued to control tightly over major sectors of the economy, such as the banking system, electricity, telecommunication, and areas of foreign trade (Vuong, 2010).

In 2006, Vietnam became the World Trade Organization's 150th member. Vietnam's access to the WTO has provided an important boost to Vietnam's economy in liberalizing reforms and expanding trade. However, WTO accession also brings serious challenges, requiring Vietnam's economic sectors to open the door to increase foreign competition.

The recent 2007-08 global food crisis and the followed global economic crisis had adversely affected on Vietnam’s economy, as almost other economies. The economic slowdown has suffered through three key channels: trade, investment, and capital mobility and financial market (Vo, 2008). With other underlying weaknesses, the crisis still has decelerated economic growth so far.

Figure 1.2 shows the annual rates of GDP growth and the GDP per capita over the period from 1985 to 2011. The trends of these indicators demonstrate corresponding economic situation that are described above. Overall, the annual growth rate of GDP increased since economic transform, and the average annual growth rate over that period was 6.7 percent (scale on the right axis). The economic went down in early of the 90s, during the 1997-98 Asia crisis and the recent global crisis, at an annual rate of around 5 percent. The average annual growth rates over the period 1992-97 and 2002-07 were remarkably high at 8.8 percent and 7.9 percent, respectively. In addition, the development of economy also is shown by constantly rising GDP per capita (at purchasing power parity). The GDP per capita was US$ 495 in 1985 and US$ 3412 in 2011, which increased by nearly seven times over this period (scale of GDP per capita on the left axis).

Structural change by economic sectors

The total economic growth was contributed by the all three sectors: agriculture, industry and services. Most of the fast-growing activities are using skill workers, foreign capital and technology (2008). However, the share of contribution to the growth differs across sectors. Figure 1.3 shows distribution of GPD in three economic sectors. The average annual growth rates of each sector were stable between the period 1990-99 and 2000-2010, at around 4 percent in agricultural sector, 10 percent in industrial sector and nearly 8 percent in services sector. Hence, industry has been the largest contributor for economic growth. By that fact, the economic structure has changed toward an increase in industry and services and decrease in agriculture. Over the three base years 1990-2000-2010, the shares of agricultural sector in total GDP were 31.8 percent, 23.3 percent and 16.4 percent. This decrease was mostly offset by an increase in industrial sector whose GDP were respectively 25.2 percent, 35.4 percent and 42.4 percent. While the proportion of services sector slightly declined from 43 percent in 1990 to around 41 percent in 2000 and 2010. Thus, industrial sector is not only the main force driving growth but also became surpassed services sector to be the main source of economic output.

Along with economic structure change, labour distribution has shifted toward that movement. Figure 1.4 shows distribution of labour market in the country corresponding to the period of Figure 1.2. In early of 1990s, the country was dominantly agrarian economy, which composed by 73 percent of those employed in agricultural sector, 11.2 percent in industrial sector and 15.8 percent in services sector in 1990. After two decades, the share of labour force in agricultural sector declined to 49.5 percent, and the share in industry and services sector rose up to 20.9 percent and 29.6 percent, respectively. Changes occurred more significantly during 2000s than 1990s. In industry sector, the increased share was attributed largely from manufacturing and construction activities.

Structural change by ownership sectors

As a centrally planned economy before the reform, all the economic activities was planned and controlled by government. There were two sectors: government and state-owned enterprises (SOEs, including co-operatives). Labour market also was monitored and allocated by respective administrative units. The reforms therefore had pushed a pressure to reshuffle the ownership structure toward a market-oriented economy, which basically was to reduce monopoly power and controlled areas of government in the economy. The privatization program was initiated in 1992. Approximately 2,600 SOEs were privatized over the period of 1992-2005 (Sjöholm, 2006). The number of SOEs dropped to 2,176 in 2007 (CIEM, 2007).

However, the fact is that the state still reserved large shares in privatized firms and maintained corporate governance and regulations. Contrary to experiences of transition economies in Eastern Europe, where the considerable reform in state sector was implemented as the ownership transferred to private stakeholders (Frydman, Gray, & Rapaczynski, 1999), the fundamental ownership structure has been insignificantly changed by the privatization. The private enterprises were approved for legal existence and were slowly emerged under the less favorable conditions during 1990s (Mac Millan & Woodruff, 1999). Whereas the SOEs enjoyed some competitive advantages over non-state sector, such as: land use and location, investment and access to credit, monopoly positions, etc. (Le, 1996). The state ownership still dominated the economy. This is shown in Table 1.1 that contribution of state sector (government and SOEs) to GDP reduced from 40 percent in 1995 to around 34 percent in 2010, while share of labor declined only one percent. Decrease of proportion in GDP's contribution from state and non-state sector (collective, private and household) was balanced by substantial increase from foreign investment sector, from around 6 percent in 1995 to 19 percent in 2010.

Agricultural sector and household welfare

As those facts, Vietnam nowadays is an agrarian country since around 70 percent of population living in rural areas and a half of labor force working in agricultural sector in 2010. Agriculture is the main source of employment and livelihood, especially in rural areas. Agricultural also plays an important role in ensuring food security, providing raw materials for industry, and earning foreign exchanges from export.

The remarkable achievements in agricultural sector also were attributed to market-oriented reforms. Before 1986, Vietnam was one of the five poorest countries in the world and suffered widespread food shortage. The first critical policy reform was initiated in agricultural sector (Glewwe, 1998). In 1987, price controls started to be removed for agricultural products and goods, and farm households were permitted to sell their products on the market with prices they want. In 1988, an important decree in Land Law issued that allowed households have the right to use agricultural land for 15 years or more, and households had to pay taxes for the plots they received. This essential step enabled households to take decision-making power on their production, and gradually left input and output markets liberalized. This change is referred as "one of the most radical land reforms in modern times", in which 80 to 85 percent of agricultural land area of the country were de-collectivized over a relatively short period (Ravallion & van de Walle, 2008). A further step was taken through the 1993 Land Law that (i) granted tenure of land for a longer period, i.e. 20 years for annual cropland and 50 years for perennial cropland; (ii) issued certificate of land use rights; (iii) and permitted the land transaction: i.e. transfer, exchange, mortgage and inherent. Another extension from the 1998 Land Law removed restrictions on size of landholdings and on the hiring of agricultural labour.

These fundamental changes, which considered farm households as the key unit of production, privatized land rights, and freed input and output markets, led to dramatic success in agricultural sector in particular, and in the economy in general. The average annual growth rate in agricultural sector achieved around 4 percent during the period 1990-2010 (Figure 3). Together with major reforms in investment and foreign trade implemented in late 1980s, the restrictions on internal and external trade have been relaxed. In 1992, Vietnam became the world's third largest rice exporter from its rice importer in mid-1980s (Glewwe, 1998), as rice is most important food in diet of local people and among exported crop. Figure 1.5 shows evolution of agro-food trade in Vietnam from 1997 to 2008. The value of agro-food exports had steadily risen. The values of agro-food imports also has been higher, however the net balance between exports and imports has increased gradually.

The dramatic progress of economy and, in particular, of agricultural activities apparently accompanied by increase in household income and decrease in poverty rate. Figure 6 illustrates the rate of poverty headcount over the period 1993-1998 which is estimated from six waves of the Vietnam national living standard household survey. The national poverty rates fell consistently from around 58.1 percent in 1993 to 37.7 percent in 1993 and 19.5 percent in 2004 and 14 percent in 2008. Like other developing countries, proportion of population below poverty line (based on living expenditure) and the proportion of absolute poor (based on food expenditure) are significantly higher in rural areas than in urban areas. In 1993, the poverty rates were 66.4 percent in rural areas and 22.1 percent in urban areas, which is more than two times higher. In 2008, the poverty rates declined to 17.7 in rural areas and 2.6 in urban areas. Since the population dominantly live in rural areas, the overall poverty reduction was contributed substantially from decline in rural areas.

Furthermore, progresses in household income and poverty reduction have been uneven not only between urban and rural areas, but also across regions. The remote and isolated areas, where infrastructure is under-developed, are impeded in accessing social and economic conditions. Also those areas with less favourable conditions for developing agricultural production face more difficulties. Figure 1.7 depicts the spatial distribution of the poverty rate across provinces and districts in 2009. The uplands regions appear to be the poorest, while river delta regions have lowest poverty rates. More specifically, poverty rate in provinces of the Northern Mountain region were around 60 to 70 percent, and in provinces of Central Highland region were around 40 percent. The rates in Red River Delta and Mekong River Delta regions, where Hanoi and Hochiminh city are located, were only around 10 percent. The geographical disparities hence widen the gap between rural and urban areas.

Although the whole country has gained in poverty reduction and living standard, the group of rural households or agricultural households are lagged behind due to various impediments. This lower progress has been theoretically and empirically documented and studied through various causes in literature. Relying on reports of Centre for International Economics (2002) and Poverty Working Group (2000), I briefly summary and add recent evidence about the major challenges as the following:

(1) Market access constraint: The farm households reply on market to sell their labor, products and to finance their investment. First, the physical location limits the possibility to assets these markets (Aksoy & Dikmelik, 2007; Tran, Hossain, & Janaiah, 2000; Van de Walle, 1996). The infrastructure system in rural areas are underdeveloped, especially the poorest regions in Northern Uplands and Central Highland regions. Second, lack of information also confines their ability to take advantage of market opportunities by growing incomes, expanding markets and diversifying product (World Bank, 2000). Third, the limited access to capital prevents the households from participating in the markets. With access to credit, the households are able to invest in the fix assets and to be active in the market (Dufhues & Buchenrieder, 2005; McCarty, 2001; B. D. Pham & Izumida, 2002; Ranjula Bali, Nguyen, & Vo, 2008).

(2) Social infrastructure constraint: First, education is essential factor for development and poverty reduction (Poverty Working Group, 2000). Many rural people, women and ethnic minorities have less opportunities to access education (Fahey, McCarty, Scheding, & Huong, 2000). The barriers are most likely to be: difficult transport to school, availability of school, high cost of schooling compared to their income, and heavy workloads. Second, the infectious diseases are prevalent in rural and remote people and health care is less accessed by the people in these regions (Castel, 2009; World Bank, 2001). The disease pattern is strong linked to geography, climate, socio-economic characteristics.

(3) External factors: First, the environmental shocks are the most relevant factors affecting the farm households and the poor (Conway, Turk, & Blomquist, 2001; Thomas, Christiaensen, Do, & Le, 2010). Due to of its location, Vietnam is frequently and severely hit by natural disaster, especially typhoons and floods. Second, the households lack of means and ability to cope with shocks, such as: illness, weather shocks, market fluctuation, changes in macro policy, etc. (Coxhead, Linh, & Tam, 2011; Fiona & Newman, 2011; K. T. Nguyen et al., 2012)

Therefore, the economy is growing from a low base and limited effective system, and need more effort to achieve sustainable and stable development, despite the fact that the economy has continued to expand.

1.3. Objectives and structure of the thesis

The evidence in previous section gives an overview of Vietnam's economy, in which agricultural sector appears to dominate the economy in terms of labor market. The majority of labor force in agriculture participate in the market as self employment that is referred as farm or agricultural households. Under the barriers of market and social constraints, farms households are unable to take advantage of the economic growth as in other sectors. In particular, they are vulnerable and hindered in an exposure to risk and shocks. Therefore, the thesis attempts to understand better the challenges that prevent farm households from generating higher and stable growth.

The analysis replies on the recent theoretical framework in microeconomics that adapted in developing economy setting. By modeling the existence of imperfect credit market, imperfect labour market, these frameworks allow us to apply for this case of farm households in Vietnam. Moreover, with a target to access the impact of risk and shock, the thesis uses data covering the period of the recent food and financial crisis from 2006 to 2010 which led to a slowdown in the economy and caused complex effects on farm households. Then econometric models, which are appropriate with assumptions in theoretical analysis and dataset, are employed to estimate the impacts. Thus, the thesis addresses three main issues organized in the next three chapters. Those chapters provide analysis and empirical evidence on decision-making of farm households under shocks and imperfect markets in Vietnam. The specific objectives are summarized as the flowing:

(1) Vietnam is one of few developing countries which is found to be overall better-off on welfare from the effects of the 2007-08 global food crisis. This chapter attempts to further investigate the decisions of household producers on investment, saving and consumption by

using panel data covering the period from 2006 to 2008. The impacts are estimated according to the assumption of imperfect credit market which are represented by initial income, access to loan and land endowment. The study will test effects of the price shocks on investment, saving and consumption; in case of the overall effect and in case of the price shocks incorporated with those credit constraint factors.

(2) Considered in the circumstance of the recent food crisis with rising global food demand, one of the key responses is to increase productivity under the resource scarcity (water, land, nutrients, and energy) by adopting technology and using input efficiently. In this case of Vietnam, rice is the most important crop for farmers and has contributed significantly for poverty reduction during the last two decades after the economic reforms. Accounting for more than 30 percent of total expenditure, chemical fertilizer adoption is an essential factor that effects on the productivity of crop. Moreover, Vietnam is harshly vulnerable to natural hazards, which negatively affect on farmers. Hence, the main objective of the study is to assess the impact of technology adoption and the degree of adoption on paddy cultivation in rural Vietnam in 2010 under different conditions of natural disasters.

(3) This chapter uses the sub-sample in rural areas and for children from 10 to 14 years old of the 2008 Vietnam household living standard survey. Although the descriptive data shows the child labor are likely to occur in the low-income households, the study aims to test another factor, farm land, which may potentially affect on child labor. Based on the theoretical model in which allows to track the effect of imperfect labor market, the study tests the relationship between child labor and farm size of household. Child labor variable is decomposed into two stages of decision: the probability to participate in economic activities, and the level of participation conditional on working.

Finally, the last chapter provides the concluding remarks that add evidence to the literature of this areas. This is accompanied with limitations that I failed to deal with, and suggestion the improvements for future search.

References

Aksoy, M. A., & Dikmelik, A. I. (2007). The Role of Services in Rural Income: The Case of Vietnam. *The World Bank Research Working Paper 4180*.

Ball, R., & Pounder, L. (1996). "Efficient but Poor" Revisited. *Economic Development and Cultural Change, 14*(4), 735-760.

Becker, G. S. (1965). A Theory of Allocation of Time. *Economic Journal, 75*, 493–517.

Castel, P. . (2009). *Vietnam Health Insurance: Use of Health Care Services by the Poor Efficiency and Equity Issues in the Province of Kon Tum*. Poverty Assessment. Vietnamese Academy of Social Sciences - Centre for Analysis and Forecasting.

Centre for International Economics, Canberra and Sydney. (2002). Vietnam Poverty Analysis. Australian Agency for International Development.

Cervantes-Godoy, D., & Dewbre, J. (2010, 29-30 November 2010). *Economic Importance of Agriculture for Sustainable Development and Poverty Reduction: The Case Study of Vietnam.* Paper presented at the Global Forum on Agriculture: Policies for Agricultural Development, Poverty Reduction and Food Security, OECD Headquarters, Paris.

CIEM, (Central Institute for Economic Management). (2007). Bao Cao Cua Chinh Phu Ve Co Phan Hoa Cac Doanh Nghiep Nha Nuoc (in English: Report of Goverment on Privatization of SOEs).

Conway, T. , Turk, C. H., & Blomquist, J. (2001). Vulnerability, Risk and Social Protection in Vietnam *first draft, prepared for the Poverty Task Force Vietnam*.

Coxhead, I. , Linh, Vu Hoang, & Tam, Le Dong. (2011). *Global Market Shocks and Poverty in Vietnam: The Case of Rice*. Staff Paper No. 559. University of Wisconsin-Madison, Department of Agricultural & Applied Economics.

de Janvry, A., Fafchamps, M., & Sadoulet, E. (1991). Peasant Household Behaviour With Missing Markets: Some Paradoxes Explained. *The Economic Journal, 101*(409), 1400-1417.

Dillon, J. L. (1971). An Expository Review of Bernoullian Decision Theory in Agriculture: Is Utility Futility? *Review of Marketing and Agricultural Economics, 39*(1), 3-80.

Dufhues, T., & Buchenrieder, G. (2005). Outreach of credit institutes and households' access constraints to formal credit in Northern Vietnam: Universitaet Hohenheim, Department of Agricultural Economics and Social Sciences in the Tropics and Subtropics. Discussion Paper No. 01/2005.

Duflo, E. . (2006). Poor but Rational? . In R. B. A. Banerjee, and D. Mookherjee (Ed.), *In Understanding Poverty* (pp. 367–378). Oxford: Oxford University Press.

Fahey, S. , McCarty, A., Scheding, J., & Huong, T. (2000). Strategic Framework for AusAID Support for the Education and Training Sector.

Fiona, W. , & Newman, C. (2011). *Income Shocks and Household Risk-Coping Strategies: Evidence from rural Vietnam*. Department of Economics, Trinity College Dublin.

Frydman, R., C. , Gray, M. Hessel, & Rapaczynski, A. (1999). When Does Privatization Work? The Impact of Private Ownership on Corporate Performance in the Transition Economies. *Quarterly Journal of Economics 114*(4), 1153-1191.

Glewwe, P. (1998). An Overview of Economic Growth and Household Welfare in Vietnam in the 1990s. In N. A. Paul Glewwe, David Dollar (Ed.), *Economic Growth, Poverty, and Household Welfare in Vietnam*. Washington, D.C., USA: The World Bank

Harvard Vietnam Program. (2008). Choosing success : the lessons of East and Southeast Asia and Vietnam's future : a policy framework for Vietnam's socioeconomic development, 2011-2020. *Asia Programs, Harvard University*.

Lanjouw, P. , Marra, M. , & Nguyen, C. . (2013). *Vietnam's Evolving Poverty Map Patterns and Implications for Policy*. Policy Research Working Paper 6355. The World Bank.

Le, D. D. (1996). Legal Consequences of State-Owened Enterprises Reform In N. J. F. Chee Yuen Ng, Frank Cong Hiep Huynh (Ed.), *State-Owned Enterprise Reform in Vietnam: Lessons from Asia*.

Mac Millan, J., & Woodruff, C. . (1999). Interfirm Relationships and Informal Credit in Vietnam. *Quarterly Journal of Economics, 114*(4), 1285-1320.

McCarty, A. (2001). Microfinance in Vietnam: A Survey of Schemes and Issues: DFID and State Bank of Vietnam.

Mendola, M. (2007). Farm Household Production Theories: A Review of "Institutional" and "Behavioral" Responses. *Asian Development Review, 24*(1), 49-68.

Millar, J. R. . (1970). A Reformulation of A. V. Chayanov's Theory of the Peasant Economy. *Economic Development and Cultural Change, 18*(2), 219-229.

Myrdal, G. (1968). *Asian Drama: An Inquiry into the Poverty of Nations*. New York: Pantheon.

Newbery, B. M., & Stiglitz, J. E. (1981). *The Theory of Commodity Price Stabilization: A Study in the Economics of Risk*. Oxford: Clarendon Press.

Nguyen, K. T. , Khuat, O. T. H., Ma, S., Pham, D. C., Khuat, G. T. H., & Ruger, J. P. (2012). Effect of Health Expenses on Household Capabilities and Resource Allocation in a Rural Commune in Vietnam. *PLOS ONE, 7*(10).

Pham, B. D. , & Izumida, Y. (2002). Rural Development Finance in Vietnam: A Microeconometric Analysis of Household Surveys. *World Development, 30*(2), 319–335.

Poverty Working Group. (2000). Attacking Poverty *Vietnam Development Report 2000*.

Ranjula Bali, S., Nguyen, V. S. , & Vo, V. T. . (2008). Microfinance and Poverty Reduction in the Mekong Delta in Vietnam. *African and Asian Studies, 7*(2-3), 191-215.

Ravallion, M. , & van de Walle, D. . (2008). *Land in Transition: Reform and Poverty in Rural Vietnam*. Washington, D.C., USA: Palgrave Macmillan and the World Bank.

Ray, D. (2006). Aspirations, Poverty, and Economic Change. In R. B. A. Banerjee, and D. Mookherjee (Ed.), *In Understanding Poverty* (pp. 409–421). Oxford: Oxford University Press.

Schultz, T. W. (1964). *Transforming traditional agriculture*. Yale University, Press, New Haven, CT.

Sjöholm, F. . (2006). *State owned enterprises and equitization in Vietnam.* Working Paper 228. Stockholm School of Economics, Sweden.

Thomas, T. , Christiaensen, L. , Do, Q. T. , & Le, D. T. . (2010). Natural disasters and household welfare: Evidence from Vietnam: The World Bank.

Tran, T. U. , Hossain, M. , & Janaiah, A. (2000). Modern Farm Technology and Infrastructure in Vietnam: Impact on Income Distribution and Poverty. *Economic and Political Weekly, 35*(52/53), 4638-4643.

Van de Walle, D. (1996). *Infrastructure and Poverty in Viet Nam*. Washington, D.C. : The World Bank

Vo, T. T. (2008). *Impacts of global financial crisis on Vietnam's economy and recommended policy responses.* Institue of Development Studies.

von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Vuong, Quan Hoang. (2010). *Financial Markets in Vietnam's Transition Economy: Facts, Insights, Implications*. Saarbrücken, Germany.

World Bank. (2000). Vietnam 2010 Pillars of Development *Vietnam Development Report 2001, Joint Report Asian Development Bank and UNDP*.

World Bank. (2001). Vietnam Growing Healthy: A review of Vietnam's Health Sector *Report No.22210-VN, June 2001*.

Tables

Table 1.1: Ownership structure in Vietnam

Panel A: Share of contribution to GDP, 1995-2010			
	1995	2000	2010
State (Government & SOEs)	40.18	38.52	33.74
Non-state	53.52	48.2	47.54
Collective	10.06	8.58	5.35
Private	7.44	7.31	11.33
Household	36.02	32.31	30.86
Foreign investment	6.3	13.28	18.72
Total	100	100	100

Panel B: Share of labour market, 1990-2010			
	1990	2000	2010
State (Government & SOEs)	11.6	11.7	10.4
Non-state	88.4	87.3	86.1
Foreign investment	0	1.0	3.5
Total	100	100	100

Source: Statistical yearbook of Vietnam and database online of General Statistics Office of Vietnam (GSO).

Figures

Figure 1.1: Vietnam regional map

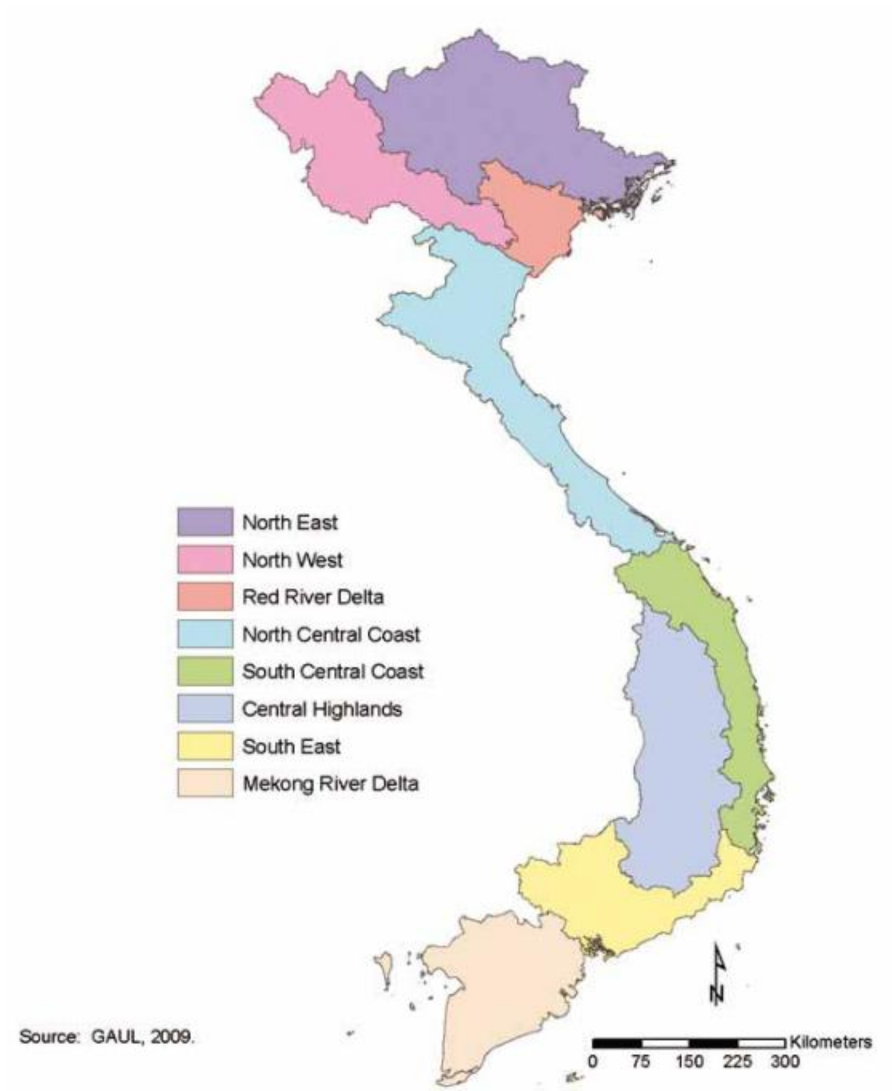
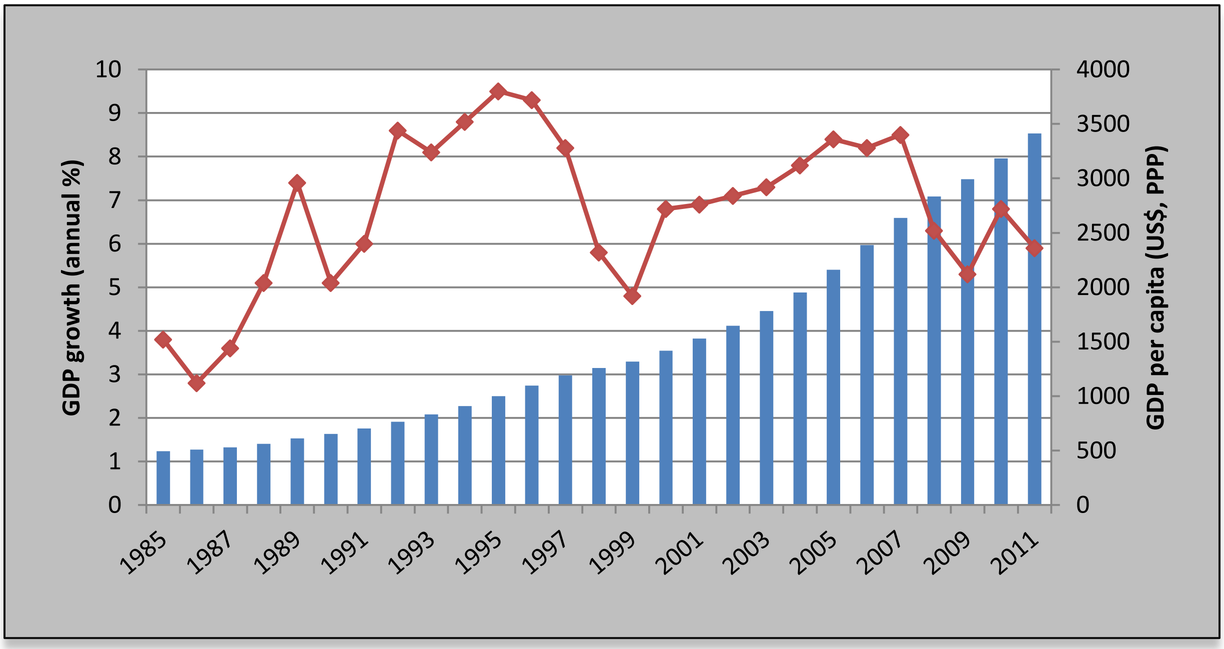


Figure 1.2: GDP and annual rates of GDP growth in Vietnam, 1985-2011

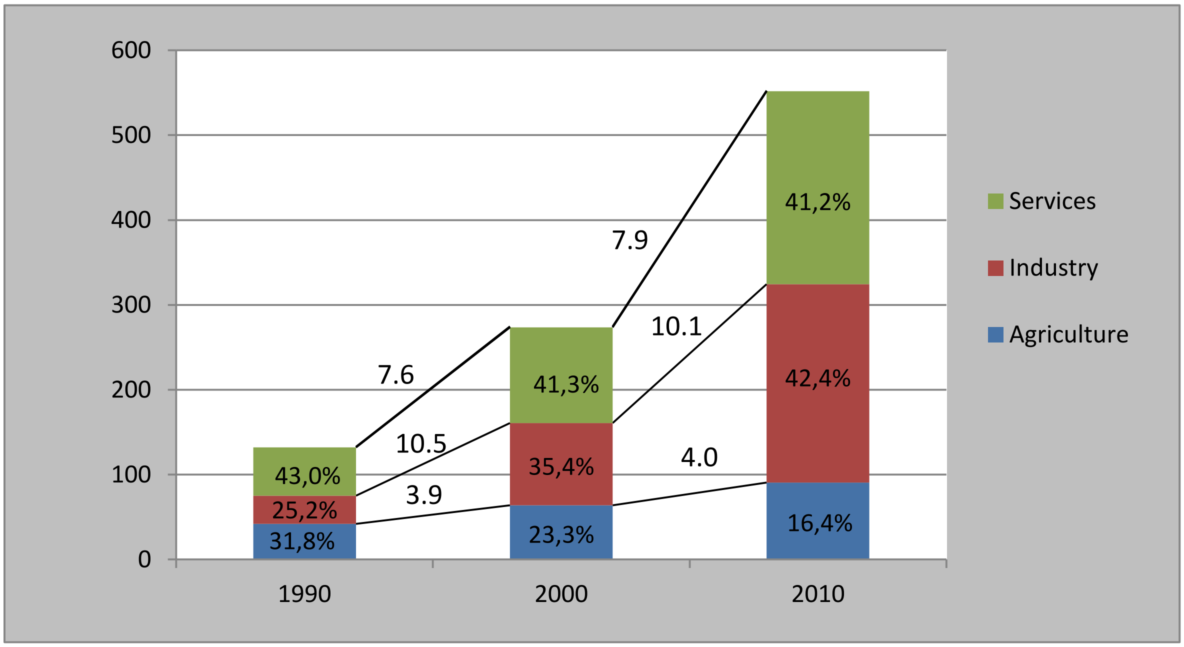


Source: World Development Indicators, World Bank

Note: (i) Annual percentage growth rate of GDP at market prices based on constant local currency.

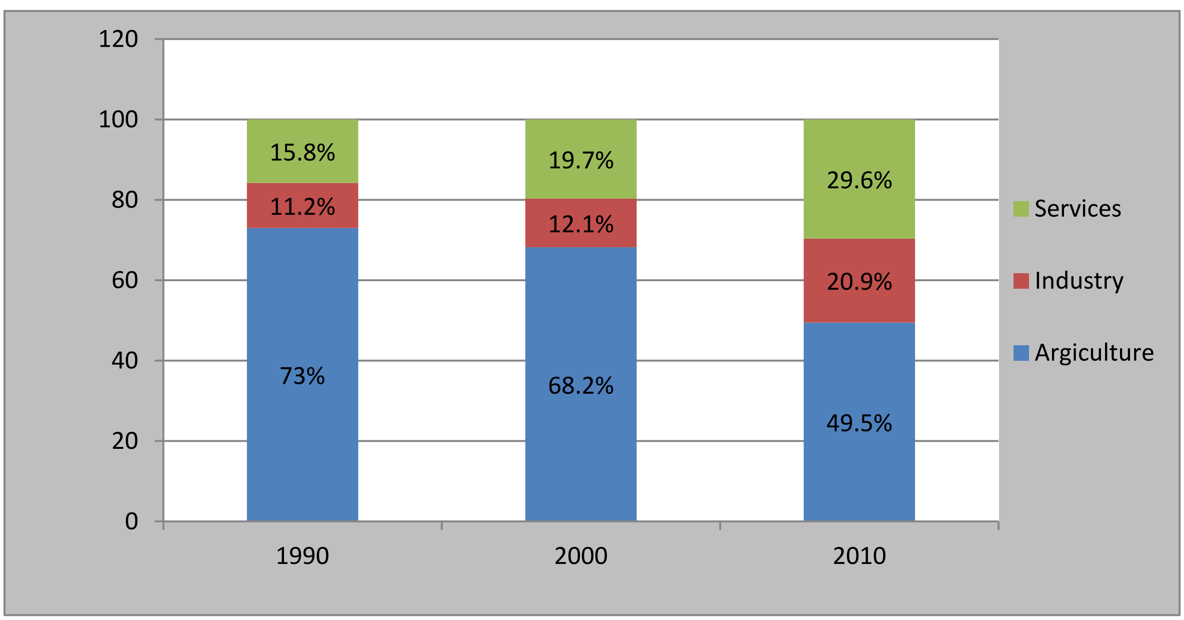
(ii) GDP per capita based on purchasing power parity (PPP). PPP GDP is gross domestic product converted to international dollars using purchasing power parity rates.

Figure 1.3: GPD (trillion VND, at constant 1994 prices), share of contribution to GDP, and average annual growth rate across sectors in Vietnam, 1990-2010



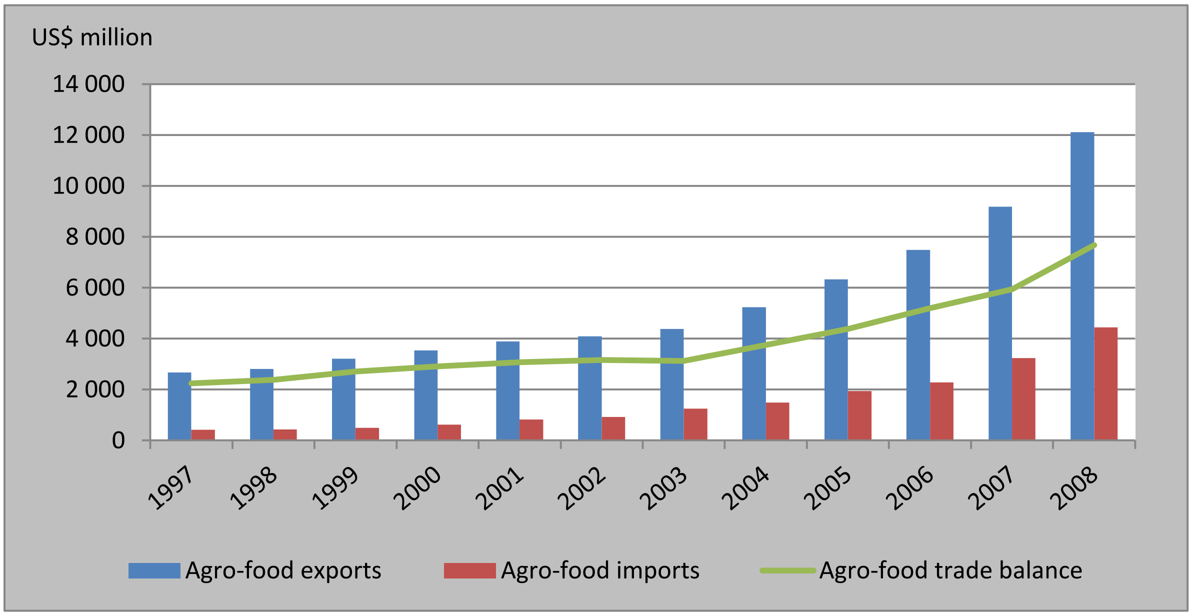
Source: Calculation from data of General Statistics Office of Vietnam (GSO).

Figure 1.4: Shares of labour market across economic sectors in Vietnam, 1990-2010



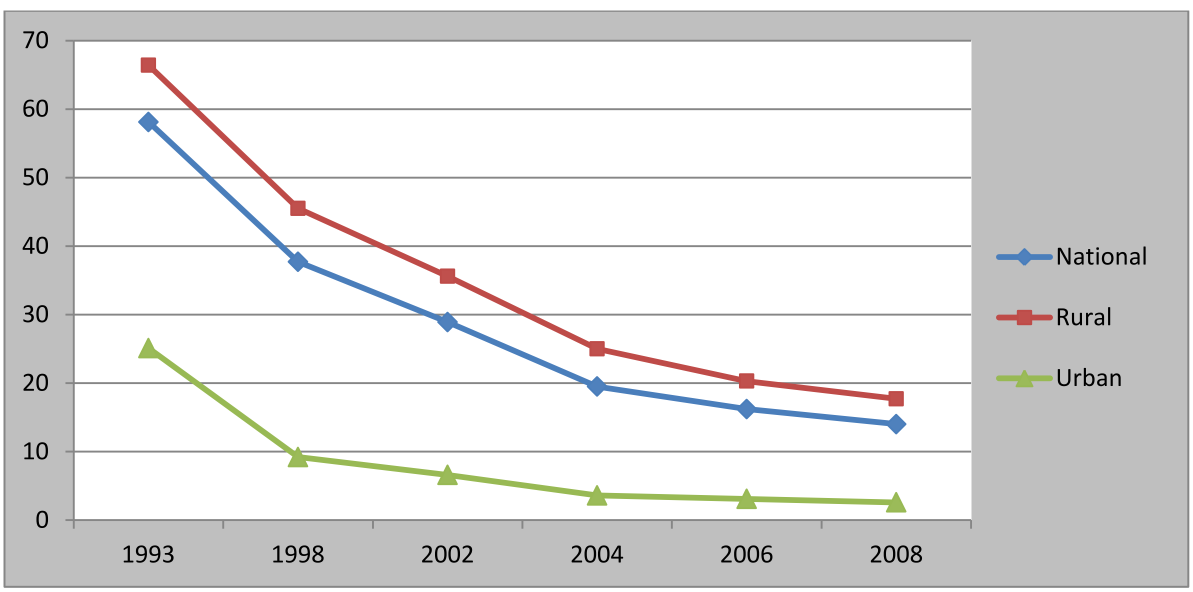
Source: General Statistics Office of Vietnam (GSO)

Figure 1.5: The agro-food trade in Vietnam, 1997-2008



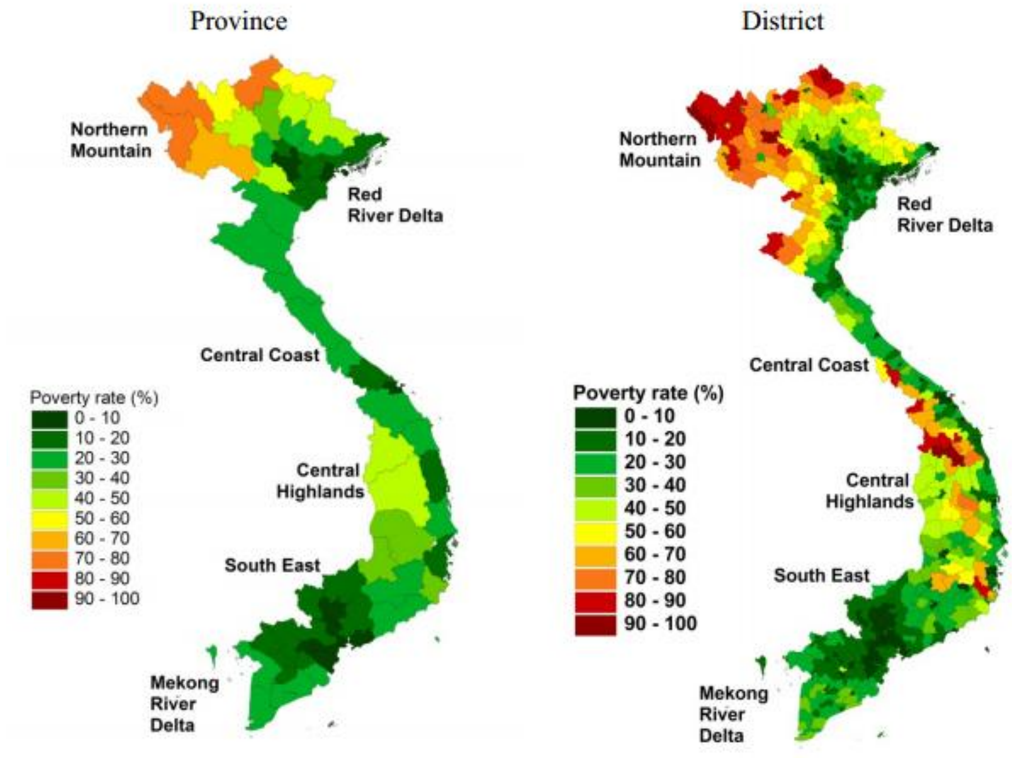
Source: Data of UN Comtrade in 2010 from Cervantes-Godoy and Dewbre (2010)

Figure 1.6: The poverty rate (%) in Vietnam, 1993-2008



Source: Calculation from the 1993, 1998 VLSS, and the 2002, 2004, 2006, 2008 VHLSS.
Note: The poverty line based on the standard of Vietnam General Statistic Office, and it was adjusted over this period.

Figure 1.7: The poverty rate of provinces and districts in Vietnam in 2009



Source: Estimation from the 2009 VPHC (Vietnam population and housing census) and the 2010 VHLSS by Lanjouw, Marra, and Nguyen (2013)

Chapter 2

The effects of food crisis on productive investment, saving and consumption

2.1. Introduction

An overview of the world food crisis

The global prices of all agricultural commodities have experienced a dramatic surge in 2007-2008. Although the food prices have still fluctuated since that period, the historical peak in mid-2008 has broadly drew attention to that situation and stimulated in measures to these problems. Figure 2.1 shows the monthly price indices of basic food commodities in the world market.[2] The staple foods index appears to show a higher rate of increase. The world benchmark of price index of cereals and oils peaked during the second quarter of 2008 at about 270 percent of its 2005 base level, remained around this level until the final quarter of 2008. The price indices for dairy and meat rose less, peaking at respectively 201 percent at the end of 2007 and 150 percent at September of 2008.

The claimed causes of the 2008 global food price crisis are numerous and relate to both the supply and demand sides of the market. The factors include: (1) rising biofuel demand in the U.S. and the EU which diverted food crops into energy uses (Mitchell, 2008; Trostle, 2008); (2) decline in productivity and stocks caused imbalance in supply-demand (Abbott, Hurt, & Tyner,

[2] The FAO food price index is a measure of the monthly change in international prices of a basket of food commodities. It consists of the average of five commodity (Meat, Dairy, Cereals, Oils and Fats, Sugar) group price indices, and weighted with the average export shares of each of the groups. These price indices are normalized at 100 percent in January 2005.

2009); (3) considerable growth in demand, especially from China and India (Headey & Fan, 2010); (4) the high energy and fertilizer prices which drove up the cost of production and distribution in agriculture (Mitchell, 2008); (5) climate and weather caused poor harvest in some main regions, particularly the 2006-2007 drought in Australia, Ukraine (Headey, 2011) ; (6) export restrictions of leading exporters and stockpiling of the importers (Dawe & Slayton, 2010; Headey, 2011; Mitchell, 2008); (7) speculation in financial markets (Gilbert, 2010; P. Timmer, 2010); (8) the depreciation of US Dollar relative to other major currencies which play an important roles in global agricultural markets (Abbott et al., 2009; Charlebois & Hamann, 2010).

The global food crisis had affected on both macro and micro level of related countries, particularly the ones that significantly participated in or depended on the world market (Benson, Minot, Pender, Robles, & Braun, 2008; FAO, 2008). (1) At the national level, high food prices impacted local commodity markets, local labour markets, government fiscal balances, external balances and political activity. (2) At the household level, the effects are mostly seen in changed real income and consumption. (3) At the individual level, the consequences occurred in nutrition, health care and education. The degree of impact depends on whether a country is net exporter or importer of resources and products, and whether a household is net buyer or seller of food commodities. These impacts finally transmit to individual level in households.

Low-income countries apparently are more vulnerable to volatile situation than high-income countries. This is because food is a smaller proportion of household budgets and agricultural is a smaller proportion of GDP in high income countries. For this reason, discussions and assessments of the impact of the food crisis focus mainly on the developing world. The IMF (2008) reports that the median 12-month rate of food price inflation of 120 non-OECD countries increased from 10 percent to 12 percent between December 2007 and March 2008, almost twice the rate of 2006. Although the increasing inflation was accelerated by both food and fuel factors, the average 2006 weight of food in CPI basket of 37 percent is far beyond this rate of fuel at around 7 percent. The World Bank (2009) also estimates the effect for developing countries that median inflation rose 12 percent in the first half of 2008. Subsequently, with the fall in commodity prices, inflation declined below 10 percent in the second half of 2008.

Surging food prices have generated severe effects for poor people in low-income countries. This is mainly because food expenditure often takes up more than 50 percent of people's disposable income in developing countries, or even 70 percent among the poor (IMF, 2008; World Bank, 2009). Since the majority of people in developing economies are involved in agricultural production, they are both consumers and sellers of agricultural commodities. Hence, the impacts of high food price on household welfare are mixed and depend on: (1) the relative share of commodities in production set and consumption basket, (2) the extent of price changes, (3) the availability of suitable substitute food items, and (4) the extent of compensation between price shocks and income changes (Aksoy & Isik-Dikmelik, 2008; Lustig, 2009).

The food crisis from the loser's side

In general, the recent empirical evidence reveals that the net food buyers are prevalent among poor households which implies that the food price spike led to an increase in poverty rates, despite the fact that many poor households are involved in agriculture and are net food sellers gain from high prices. M. Ivanic and Martin (2008) estimate first-order welfare changes of households for a sample of ten developing countries across Asia, Latin America, and Africa. They find the overall effect of food price increases on poverty to be negative. Wodon and Zaman (2008) find a similar effect in sub-Saharan Africa countries, as the negative impact on the net poor consumers surpasses the benefits to poor producers. Busjeet, Demombynes, and Sobrado (2008) also report an extremely rough estimate of the impact on poverty by using a price index for the poor in Central America countries.

Some other studies show the evidence of negative consequence on poverty or household welfare in specific countries, specifically: Ghana (Wodon, Tsimpo, & Coulombe, 2008), Mexico (Valero-Gil & Valero, 2008), Brazil (Ferreira, Fruttero, Leite, & Lucchetti, 2011), and the Philippines (Fujii, 2011).

The food crisis from the gainer's side

Among empirical analysis of developing countries, Vietnam and Indonesia are found to have exhibited poverty reduction during the food crisis. As the leading exporters of some agricultural commodities, the revenue of those products form a large part of GDP. In the case of Indonesia, the commodities accounted for one fourth of GDP and it is the largest producer of palm oil in the world. The rising commodity prices from 2003 to mid-2008 accelerated a growth in total exports around 14 percent per year this period. High commodities price also lifted total income by an average 1.2 percent of GDP in the 2004-2007 period. Rural poverty declined by 2.2 percent and urban poverty rate was unchanged with the result that the national poverty reduction was at 1.7 percent (World Bank, 2011). Motivated by these facts, Nose and Yamauchi (2012) recently estimated the impact of food prices on agricultural production by using the panel data in 2007 and 2010 in seven Indonesian provinces. They find that price shocks create an incentive to save and invest - wealthy farmers invested more in productive assets, while poor farmers increased financial saving and consumption.

While Vietnam is the second largest exporter of rice in the world. Rice also in the most important crops in agriculture. The agricultural sector accounted for 22 percent of GDP in 2007, and the growth rate in 2008 is 3.8 percent. The total revenue of total exports rose 29.5 percent between 2007 and 2008. In particular, the contribution of crude oil, industrial products and agricultural products to that growth rate respectively are 49.7 percent, 30 percent and 16.3 percent. Regarding empirical evidence, a number of studies show the overall positive impact of high prices on household welfare using different approaches. Applying a dynamic computable general equilibrium model, Thurlow et al. (2010) suggest that the 2008 food crisis increased employment and reduced poverty by favouring labour-intensive exports, especially in agriculture. Although the high prices resulted in most of households becoming worse off, the average loss of net buyers was lower than the average gain of net sellers (Vu & Glewwe, 2011). Looking at the decomposition, the greatest gainers appear in the quintiles 2 and 3 (D.T. Phung & Waibel, 2010). In the comparable of eleven developing countries, the median welfare changes declined in urban areas of all countries, but only in Vietnam overall welfare still increased in rural areas (Zezza et al., 2008).

Objective of this study

In recent literature, almost studies have quantified changes on household welfare in terms of income and expenditure with a little attention paid on the effects on agricultural production of households. Households who gain and lose from high price might react differently in terms of investment in agricultural production. In particular, the group of net food sellers or group of households which gains from increasing commodity prices may be expected to boost investment. Nonetheless, in developing economies, imperfect markets and risk aversion might cause the opposite decision on investment. Hence, the impact on agricultural investment of households could be ambiguous and still a puzzle.

Given this context, the paper aims to further investigate the response of Vietnamese households to commodity prices shock in relation to investment in agricultural production, consumption and saving. The short-term impact is investigated in order to understand how households use their income gained from price shocks in imperfect market conditions. The objective of paper is to examine these questions (Deaton, 1990; Paxson, 1991; Udry, 1995):

(1) Do the farmers have incentives to increase productive investment, or do they increase saving and consumption instead, if they gain additional income from positive price shocks?

(2) Is there any impact of financial constraints on such spending decisions?

2.2. Theoretical model[3]

Consider a household producer with two periods which produces entirely for the market. Household production follows the standard production function $f(k)$. Household consumption is c_1 and c_2. In the period 1, capital stock of household is k_1 that is from past decisions. Household chooses its period 1 borrowing level b_1 and net investment level $i_1 \geq 0$. For simplicity, we assume there is no production uncertainty. The product price is p_1 and p_2 in the two periods where p_1 is known but p_2 is uncertain at the time of the borrowing and investment decisions. The long run product price is p^* and we consider the situation of a price spike so $p_1 > p^*$. The budget constraints faced by the household in the two periods are:

$$c_1 + q(i_1 + \delta k_1) = p_1 f(k_1) + b_1 \quad (1)$$

$$c_2 + (1 + r)b_1 = p_2 f(k_2) = p_2 f(k_1 + i_1) \quad (2)$$

The rate of interest is r, the capital depreciation rate is δ and the capital cost is q (all constant). The capital rental is therefore $q(r + \delta)$. We suppose that the inherited capital stock k_1 is exactly that appropriate for the long run price stock p^*, i.e. $p^* f'(k_1) = q(r + \delta)$.

The household maximizes the sum of current and future utilities over i_1 and b_1

$$\max_{i,b} \{u(c_1) + \beta E u(c_2)\} \quad (3)$$

where $\beta \leq 1$ is a discount factor.

The first order condition with respect to net investment i_1 is

$$-qu'(c_1) + \beta E[p_2 u'(c_2)] f'(k_2) = 0 \quad (4)$$

The first order condition with respect to borrowing b_1 is

$$u'(c_1) - \beta(1 + r)E[u'(c_2)] = 0 \quad (5)$$

Substitution of $u'(c_1)$ gives

$$q(1 + r)E[u'(c_2)] = E[p_2 u'(c_2)] f'(k_2) \quad (6)$$

We can expand

[3] I thank Christopher Gilbert for helping me to develop this part.

$$f^{'}(k_2) = f^{'}(k_1) + f^{''}(k_1)i_1 = \left(1 - \rho\frac{i_1}{k_1}\right)f^{'}(k_1) \qquad (7)$$

where $\rho = -\frac{k_1 f''(k_1)}{f'(k_1)}$ captures the extent of decreasing returns to capital.

The price in the period 2 is expressed as:

$$p_2 = E[p_2|p_1] + \varepsilon_2 = \lambda p_1 + (1-\lambda)p^* + \varepsilon_2 p^* \qquad (8)$$

Then we have:

$$E[p_2\, u^{'}(c_2)] = \{(1 + \lambda\pi_1)E[u^{'}(c_2)] + E[\varepsilon_2 u^{'}(c_2)]\}p^* \qquad (9)$$

where $\pi_1 = \frac{p_1 - p^*}{p^*}$ is the proportional excess of the period 1 price over the long run price. Making these substitutions, we obtain:

$$\left(1 - \rho\frac{i_1}{k_1}\right)p^* f'(k_1)\left\{(1 + \lambda\pi_1) + \frac{E[\varepsilon_2 u'(c_2)]}{E[u'(c_2)]}\right\} = (1 + r)q \qquad (10)$$

and using the condition for the initial capital stock

$$\left(1 - \rho\frac{i_1}{k_1}\right)\left\{(1 + \lambda\pi_1) + \frac{E[\varepsilon_2 u^{'}(c_2)]}{E[u^{'}(c_2)]}\right\} = \frac{(1 + r)}{(r + \delta)} \qquad (11)$$

We can transform to obtain:

$$\frac{i_1}{k_1} = \frac{1}{\rho} - \frac{(1 + r)}{\rho(r + \delta)\left\{(1 + \lambda\pi_1) + \frac{E[\varepsilon_2 u^{'}(c_2)]}{E[u^{'}(c_2)]}\right\}} \qquad (12)$$

It shows that the higher the period 1 price above the long term average (π_1) and the greater the perceived likely persistence of these high prices is, the higher the investment-capital ratio i_1/k_1 is.

From the fact that farmers in developing countries often face with market imperfection. Indonesia or Vietnam also is not an exception (Nose & Yamauchi, 2012).[4] Overall, poor farmers or households have less possibility to obtain credit in formal market and with lower amount of

[4] Distribution of access to credit in terms of share and amount in Indonesia and Vietnam is shown in Table 2.14.

credit. Hence the theoretical framework aims to consider a situation in which credit market is constrained.

Suppose households face a borrowing constraint at $\bar{b}$ such that for $b > \bar{b}$ they pay a higher interest rate $r + \tau$. The period 2 budget constrain becomes:

$$c_2 + \left(1 + r + \tau.1\left(b_1 > \bar{b}\right)\right) b_1 = p_2 f(k_1 + i_1) \qquad (13)$$

Optimal investment becomes:

$$\frac{i_1}{k_1} = \frac{1}{\rho} - \frac{\left(1 + r + \tau\left(b_1 > \bar{b}\right)\right)}{\rho(r+\delta)\left\{(1 + \lambda\pi_1) + \dfrac{E[\varepsilon_2 u'(c_2)]}{E[u'(c_2)]}\right\}} \qquad (14)$$

Since investment is declining in the rate of interest, this will give rise to a corner solution at $b_1 = \bar{b}$ where investment will be constant over a range of values of $\lambda\pi_1$. Over this range, household pays to borrow $\bar{b}$ at the unconstrained interest rate *r* but not to increase borrowing and be forced to pay the higher rate ($r + \tau$).

Moreover, the imperfect credit market lead to liquidity constraint. We assume that the interest rate for saving is constant, but the interest rate for obtaining credit depends on initial welfare or asset. The interest rate is expected higher for smaller farm household in term of income and landholding which could serve as collateral or deposit. In other words, the less-capitalized households bear higher extra cost τ for access to credit. So the households with lower income, smaller farm tend to decrease investment.

The equation (5) refers to consumption smoothing. Farmers will borrow to invest. If they can not borrow, they will reduce consumption to invest. The appendix gives more detail of explanation.

2.3. Context of the crisis

The economy in this period overview

Table 2.1 shows the economic structure of Vietnam in 2007. The industrial sector contributes the largest component in GDP accounting for 42 percent; followed by services and agriculture at 36 percent and 22 percent, respectively. However, labor shares are in the reverse order. More than half of labour force works in the agricultural sector, at 54 percent; while this rate is

26 percent in the service sector and is only 20 percent in the industrial sector. Crops account for the largest share of the labour force, at nearly 37 percent. These composition ratios therefore suggest that the farmers make up the majority of the labour force but generate the lowest share in GDP. Within the industrial sector, mining, which is mainly oil extraction, is the most productive area in which adding 11 percent to GDP through merely 1 percent of employment.

Out of total export revenue, the share of raw agricultural production is more than 7 percent. Take into account of processing production, total agriculture-related value constitutes more than 20 percent. Clothing and textiles are the leading sources in export accounting for nearly 26 percent. Crude oil also creates a substantial proportion of 19 percent. Looking within sectors, the ratio of export to total value of output is around 21 percent for agricultural production, 33 percent for agro-processing, 69 percent for clothing and textiles, and as high as 83 percent for crude oil.

Figure 2.1 illustrates the growth rates across sectors. The economy maintained a growth rate of 7-8 percent over the period from 2000 to 2007. The industrial sector enjoyed the highest rate, at 10 percent. Services and agriculture achieved around 8 percent and 4 percent respectively. At the beginning of commodity prices shock stage in 2008, the total growth rate decreased arising from a sharp decline in industry and services. Only the agricultural sector managed to attain a higher growth rate by taking advantage of rising prices during 2008, before all sectors have moved to a downward trend as the result of the 2008-09 financial crisis.

The price shocks and transmission

As an agriculturally dependent country and a leading rice exporter, the rise in international commodity prices might cause significant effects to both producers and consumers at household-level. The extent to which price volatility is transmitted to domestic market depends on the basket of export commodities relative to world market and the trade policies of the country. The prices in the world market turn to affect the households through the relative price of traded and non-traded commodities that households face in the market (Benson et al., 2008). The world markets directly influence the former and indirectly influence on the latter. The indirectly impacts occur as consumers substitute toward non-traded products with relatively lower prices and producers substitute toward the traded products with relatively higher prices.

In Vietnam, rice is an essential source for agricultural exports, as well as being the principal staple of local consumption. Figure 2.3 shows the price of rice exports between Vietnam and Thailand, which reflects the similar trend. As being a major world export of rice, Thailand is selected to observe the comparable movement of prices during crisis. All the prices are normalized to 100 percent in 2005. The figure illustrates the export price index of 5% broken rice in Vietnam and Thailand on the base year 2005. Rice price started to increase from the beginning of 2007 and then surged from January 2008. Both countries hit the peak in the second quarter of 2008 (May) the extent of the rise being higher in Vietnam than in Thailand, at 322 percent and 300 percent correspondently. Then prices both declined significantly in the second half of 2008.

In the world market, the factors driven high price of rice were somehow different to other agricultural commodities. According to the argument of Dawe and Slayton (2010), the causes of rice price spike were not from market fundamentals. First, rice production and stocks were irrelevant to the turmoil in the world rice market. Since the growth in the world rice production was similar to the rate of population growth in Asia in 2005-07. Second, rice market was less closely connected with other cereals markets, since maize markets were associated with bio-fuel policies and wheat markets were associated with bad weather. Third, an increase in prices of fuels and other cereals cumulative pressures on the policy decision of the major exporters and importers. Additionally, with a smaller share in the world market, as well as in the futures markets, and the important role of government, the price of rice markets were more volatile than maize and wheat market (P. Timmer, 2009). Therefore, policies and panic are claimed to be the fundamental factors induced significant and more rapid price increase on the world rice market. The price spike of the world rice market was largely triggered by export restrictions of India and Vietnam and stockpiling of the Philippines in 2007-08 (Dawe & Slayton, 2010). In 2007, India and Vietnam were the second and third largest rice exporters, and the Philippines was the largest rice importer in the world rice market. The price uncertainty was also contributed by interventions of others exporters and importers.

In case of Vietnam, the government tightly regulates both international and domestic rice price. The government has monitored the volume of rice export as it entered the international market (Nielsen, 2003). The annual of export quota is adjusted in response to the domestic crop output. Normally, the annual quota is approved in late summer. Thus, by the end of July 2007, the quantity of rice export was determined, and no further supplement of rice export in end of 2007 was anticipated (Dawe & Slayton, 2010). In 2008, in the atmosphere of the crisis, the government banned rice exports from late March to June in order to guarantee the domestic food security (H. N. Pham, 2010). Moreover, the relevant ministries also stabilized local market during the high price episode through adjusting tax, credit, and preferential interest rate in the local market.

In the domestic market, Figure 2.3 shows that the retail price of ordinary rice and farm gate price of paddy consequently rose upward but lagged behind the international movements both in terms of timing and extent. The local prices obtained peak in the third quarter of 2008, at

224 percent for retail price and 218 percent for farm gate price. This trend explicitly reveals that the international price was partially transmitted to local market. Thus, the intervention of government in an attempt to mitigate the adverse effects of global demand on domestic prices played considerable role.

In addition, Figure 2.4 shows the price indices of cereals and vegetables, which are the major crops behind rice, also experienced an increase along. The other staple commodities include maize, cassava, potato and sweet potato. The price of cereals peaked to around 200 percent of its 2005 level in October where it remained until the end of the year, even slightly increasing during 2009. It is possible that this is the consequence of substitution effects from both demand and supply sides as discussed above.

At the sub-national level, price volatility affected to different regions to different degrees depending on local supply and demand characteristics. In 2007, the Mekong River Delta region provided 51 percent of total rice production, followed by the Red River Delta region with 18 percent (GSO [5]). The mountainous North West and Central Highland accounted for only 1.6 percent and 2.4 percent of the total. These two regions have high poverty rates compared to other regions, 49 percent and 29 percent in 2006, respectively (VHLSS 2006). However, the mountainous and highland regions provide relatively high amount of maize which is the second important staple food in Vietnam. In 2007, the production of maize is 25 percent in the Central Highland, 18 percent in the North East, and 13 percent in the North West (GSO). Thus, these regions with lower rice production could offset the volatility through their reliance on substitute food. Figure 2.5 depicts the rice price indices across regions. The North West and the North East exhibit the lowest degree of price volatility. The Red River Delta, the Mekong River Delta and the South East, where the large proportion of people reside in urban areas and are net food consumers, experienced the highest increase in price. Therefore, the effect of price shocks was transmitted through a variety of mechanisms with differing impacts across regions.

2.4. Descriptive statistics

2.4.1. Survey data

After the economic transition in 1986, Vietnam conducted two important household surveys: Multi-Purpose Household Survey (MPHS) and Vietnam Living Standard Survey (VLSS) during the 1990s. The two surveys are partly overlapped, which motivated General Statistics of Vietnam to

[5] http://www.gso.gov.vn

merge MPHS and VLSS to become a new Vietnam Household Living Standard Survey (VHLSS). The VHLSS has been implemented biennially from 2000 to 2010, i.e. including four waves in 2002, 2004, 2006 and 2008 based on the same master sample, and with technical assistant from UNDP and the World Bank.

Based on the relevant documents and manual instruction of Phung & Nguyen (2006) and GSO (2008), the survey could be briefly summarized as the following. The interviews in each survey were conducted from May to November in each year. Sampling was at three levels: communes/wards at the first stage, census enumerate on areas (EA) at the second stage and households at the third stage. At the first stage, the sample was selected from the master frame designed for four waves of the VHLSS in this period which included 3,063 communes/wards from 1999 Population Census. At the second stage, wards and communes were partitioned into EAs and three EAs in communes/wards selected. Only one EA constitutes for each wave of survey and the two others are used for the sequential rotated waves. At the third stage, a sample of households was selected systematically with twenty households in each rural EA and ten households in each urban EA. This is technically a three-stage design (including the selection of households), but it is operationally equivalent to a two-stage design since only one EA is selected within each commune for each wave of survey. The sample is rotated 50 percent from one wave to the successive wave of the VHLSS based on the master sample. More specifically, the current survey keeps 50 percent of households in the previous survey, and randomly selects another 50 percent of households from EAs which are different to the ones used in the previous survey, as mentioned at the second stage.

The paper relies on panel data of the VHLSS in 2006 and 2008. As this designed method, 50 percent of households in the VHLSS 2006 were retained in 2008. Out of 9,189 households in each round of the 2006 and 2008 VHLSS, 4,104 households in 2006 were re-interviewed in 2008. The farm households involved in crop cultivation are extracted from this panel to analyze.

2.4.2. Production

In the survey, the agricultural crops are divided into five categories: (i) rice, (ii) other staples/ starchy and vegetables, (iii) annual industrial crops (soybean, peanut, sugar cane, tobacco, etc.) and perennial industrial crops (tea, coffee, rubber, coconut, cashew, etc.), (iv) fruit crops, and (v) crop by-products (straw, thatch, starchy stems, sugarcane leaves, etc). Due to the limited availability of price indices, which is not contained in the survey, we confine attention in this paper examines to the first and second group of crops: rice, other staples and vegetables. These commodities are the main food of consumers in Vietnam and comprise the majority share of revenue of total crops. In the 2006 VHLSS, rice, other staples and vegetables contributed 80 percent out of total crop production (Table 2.2). It is therefore reasonable to capture the price volatility experienced by Vietnamese households by these three groups of commodities: rice, other staples (maize, cassava, sweet potatoes, potatoes), and vegetables. In the panel data, 2,531 households in the 2006 VHLSS reported production of those crops, but only 2,303 households continued to cultivate crops in 2008, and 168 households started cultivation in 2008. The analysis focus on those households who reported production of these crops in both 2006 and 2008.

Table 2.2 shows how the proportion of crops produced varies across regions. In the total sample, rice revenue accounts for 73 percent, whereas other staples and vegetables account for 15 and 13 percent respectively. Rice is responsible for a larger share in the North River Delta and the Mekong River Delta regions, and a lower share in mountainous and highland regions. The shares of other staples are correspondingly are higher in mountainous and highland regions.

2.4.3. Investment and household welfare

Table 2.3 presents the summary statistics of sample used in the two surveys. I confine attention to the 2,303 households that continue to produce the three crops under consideration (rice, other staple and vegetables) in 2006 and 2008. Details of fix asset investment are shown in Table 2.4.[6] The assets bought for crop cultivation are extracted from two sources of question. First, the small and non-durable agricultural tools and minor repair of assets are extracted in expenditure of total crop during the last 12 months. Second, the remaining of equipments and machineries are picked up from fix assets of household which have value over VND 500,000 (around US$ 31)[7] at time purchased.[8]

Average cultivated land areas increased from 2.09 acres in 2006 to 2.32 acres in 2008. Total expenditure of crops and expenditure per acre of land also increased. This increase may be caused by rising investment or higher input prices, in particular fuels and fertilizers. As a consequence, total crop revenues increased but revenue per acre of land decreased.[9] On the

---

[6] Distribution of fix asset investment is shown in Figure 2.8 and Table 2.13.

[7] Use the average exchange rate in 2007 from the World Bank online database: USD/VND = 15,994

[8] In particular, the second source is what household reported about total available fix assets and durable goods with information of the time of purchase (month, year), value at the time of purchase, quantity, and the percentage of household ownership. Then I picked up the ones related to crop cultivation and purchased in 2007-08 only, calculated the total value of household, and deflated with CPI to the base time January 2006 as other values.

[9] Distribution of increase in crop revenue is shown in Figure 2.9.

other hand, the revenue per capita increased 26 percent in 2008 making a significant contribution to household welfare.

To obtain more insight on financial flows of households over the period of the price shocks, we break down the different sources of income and expenditure for three groups of households. The group " continuing households" refers to households continued to cultivate during period of 2006-08, the group "out households" refers to those who stopped cultivating in 2008, and the group "in households" contains those who started cultivating in 2008. Table 2.7 summaries the total income and expenditure. Since the inflation in 2008 rose to a peak over the crisis, the nominal figures exaggerate the real increase. All values therefore are deflated at constant price in January 2006.

Total income is divided into three categories. The first source of income comes from: wages/ salaries, farm and nonfarm, and business activities. The second source of income is from: remittances, social welfare allowances (jobless, policy household, disaster, etc), pension, income from insurance, subsidies, and interests (savings, bonds, loans, shares). The third source of income comes from: (i) selling means of production (working cattle, reproductive pigs, machines, equipment, workshops), houses, assets, exchanging lands, etc; (ii) selling gold, silver, precious stone, jewelry; (iii) withdrawal from savings, stocks, obtaining debts, etc; (iv) borrowing on interest, advance payment; and other uncategorized items.

Total expenditure also is grouped in four categories. The first category includes expenditure on foods and drinks. The second category consists the non-food living expenditure: (i) daily non-food (expenditure on housing, electricity, water and garbage; soap, shampoo, newspapers, flowers, entertainment, etc.); (ii) annual non-food (clothes, other garments, household items, internet, travel fee, services, etc.); (iii) health and education fee. The third category is expenditure for repairing and purchasing fix assets, durable goods, and house, land. The final category of expenditure includes: (i) other spending that is considered as expenditure: legal and administrative fees and taxes; contribution/support to funds/donations; wedding, funeral, entertainments; and others; (ii) other spending that is not considered as expenditure: debt repayment, incomplete big investment, and other uncategorized types.

Then saving is summation of residual income (difference between total income and total expenditure) and other spending for saving purpose, including: (i) lending, contributing to revolving credit groups, buying shares and bonds; (ii) buying gold, silver, gemstone, foreign exchange; (iii) depositing in savings accounts; (iv) buying life and life security insurance.

Overall, the data shows that the total income significantly increased in 2008 compared to 2006. The increase is highest for "out households", followed by "continuing households" , and lowest for “ in households”. Increases in income are attributable to all three income categories, except for the second category of “continuing households” and “ in households”. The total expenditure of these groups of households also increased , and highest increase is for “out households”. Increases in the third and fourth categories of spending of group “out households” are remarkable greater than these other two groups of household. Hence, with highest income and expenditure in 2008, this group “in households” had lowest saving, and the group “in households” obtained the greatest saving. Especially, saving in all groups of household decreased in 2008 compared to 2006. The data also shows that loans obtained in “in households in” is greatest in 2008; and an increase in loans obtained in 2008 compared to 2006 is highest in “continuing households”, but is negative in “out households”. Thus, an access to credit may be linked to the motivation to increase investment in cultivation.

2.4.4. Price index

The price data are not in the VHLSS but come from another source of the Vietnam General Statistic Office. Due to the absence of real price, we have to use a price index instead of measuring the level of prices.

Retail price index at national and regional level

Deriving from the test of Deaton (1990), Table 2.5 shows summary statistics for the retail price indices from three crop groups during the two surveys for May 2006 to November 2008 that the analysis uses (details in the followed section). Retail price indices are normalized to be 100 at January 2005. These indices reflect level of increase in nominal prices, and note that they are not real prices. The coefficients variation show the extent of variability in relation to the mean of the sample. Rice price indices appear to have greatest volatility 0.29, following by other staples at 0.22 and vegetables at 0.19. The coefficients of autocorrelation are tested at lags of one and six months to reflect the seasonal variation. The first-order autocorrelation are high for all groups, at 0.99 for other staples, and 0.98 for rice and vegetables. The autocorrelation still remains high for six months lags, especially other staples.

Price index at household level

I use the monthly retail price index to measure the degree of price shocks at household level. There are two limitations in the use of these price indices. First, the relevant price for producers is the farm-gate price index since this would capture more precisely the effect on the production decision of farmers. Because only the farm-gate price index for rice is available, I have to employ the retail price index. Nonetheless, the degree of fluctuation in the margin between retail and farm gate price indices is fairly small, as shown in Figure 2.3 for rice. Second, the retail price indices are disaggregated by three groups of crops. The retail price indices are available at national level for all three groups, but at regional level only for rice and other staples.[10] Since the average revenue of vegetables accounts for a small fraction of crops and fluctuates less, it may be acceptable to combine this national index with the regional indices of these two groups.

In order to construct the index of price shocks at household level, this index/variable has to capture the price change with the weight of each crop in the basket corresponding to the household's production. The price index at household level therefore is adjusted by the share of three type of crops in used sample. The retail price indices in each year (2006 and 2008) are identified in regional level, denoted by *j*, except vegetables at national level, and the crop groups (γ_1=rice, γ_2=other staples, γ_3=vegetable). The retail price index is then transformed to household level (*i*) by multiplying the difference of retail price index between 2006 and 2008 with shares of crops of a household (g). Shares of crops are measured by two indicators: share of total output value and share of sole/ bartered value.[11] This is shown as:

$$\Delta PI_{ij} = \sum_{\gamma=1}^{3} g_{ij\gamma}(PI_{j\gamma,2008} - PI_{j\gamma,2006}) \quad (17)$$

2.5. Empirical approach

In this section, the theoretical framework developed in section 2 will be translated into an econometric model which can be estimates using the VHLSS data. Household behaviour (investment, saving, expenditure) is expected to be influenced by the level of food price shocks alone, and conditional on financial constraints.

Investment, saving and expenditure

Initially, the model is estimated in first differences using the panel data for 2006 and 2008. Differencing eliminates time-invariant fix effects at household level. These unobserved factors could be correlated with the price weights to share of crops and lead to bias in estimating the effects of the price spike on investment or expenditure decisions.

[10] The difference in distribution of national and regional price indices is shown in Figure 2.7.

[11] According to the survey, these crops were collected with different information. Rice crop has details of output value on: total value, lost, sold or bartered, retained as breeds, used as food, used as food for cattle or poultry, used as gift, lending or input of business, payment, etc., left for future consumption. However, other staples and vegetables have details on: total value, sold/bartered, retained for consumption.

$$\Delta lnk_{ij} = \beta \Delta PI_{ij} + \Delta \varepsilon_{ij} \quad (18)$$

where Δlnk_{ij} is expenditure on agricultural assets in 2007-2008 period, and $\Delta\varepsilon_{ij}$ is uncorrelated to ΔPI_{ij} due to the assumption $E[\Delta\varepsilon_{ij}|\Delta PI_{ij}] = 0$, then the first-difference estimator is unbiased and consistent.

Moreover, the time-varying factors at regional level are also aware of potential effect. The model includes a dummy variable u_j captures the regional-specific effects on investment of each region *j*. Inclusion of this dummy variable can account for the different characteristics of regions vary in terms of natural endowments, crop structure, food consumption behaviour, etc. For example, the highland and mountainous regions are suitable for industrial crops than rice and staple crops as in delta regions (Table 2.2). This lack of homogeneity might induce the households in different regions adjust to price shocks in different ways.

The effect of price index volatility on investment is estimated by the first-difference model:

$$\Delta lnk_{ij} = \beta \Delta PI_{ij} + u_j + \Delta \varepsilon_{ij} \quad (19)$$

The impact of price shocks on investment is estimated by the coefficient β. If increased price shocks raise investment, the coefficient will be positive, $\beta > 0$. Otherwise, the price shocks may decrease investment, $\beta < 0$; or leave it uninfluenced. The standard error is clustered at area (urban and rural) and provincial level.[12]

The same procedure is applied for saving and expenditure. Since the expenditure may vary depending on items, the model tests for total expenditure and separately food expenditure. Replacing the investment dependent variable in equations (18) and (19) by the difference of saving and expenditure during 2006 and 2008, the estimation function are:

$$\Delta ln\ (saving)_{ij} = \beta \Delta PI_{ij} + u_j + \Delta \varepsilon_{ij} \quad (20)$$

$$\Delta ln(expenditure)_{ij} = \beta \Delta PI_{ij} + u_j + \Delta \varepsilon_{ij} \quad (21)$$

[12] The first-difference regression assumes that the residuals are independent. The sample contains households covering all 8 regions, 64 provinces of both urban and rural areas in Vietnam. Heterogeneity across regions is controlled by dummy variables. Hence, it is possible that the impact of price shock may not independent within each province, and this could cause the residuals that are dependent within province and within urban or rural areas also. The cluster option indicates that the households are clustered within area-province and correlated within area-province but may not correlated across area- province. The correlations within area- province are checked graphically and statistically, and reveal that correlation coefficient values (between price shock and investment) are fairly diversified.

Imperfect credit market

According to theoretical analysis, given a positive income shock, wealthier producers will invest more due to their greater availability of financial sources, whereas the less wealthy producers will save or spend more. An alternative possibility is that poor producer takes advantage of their higher marginal return on capital relative to that of rich producers. If poor producers gain in terms of income from the price shocks, they will invest more in production. Therefore, the decision of the producer depends on which effect dominates. The model examines the effect of price shocks on investment, saving, and expenditure conditional on initial wealth conditions. Initial wealth is captured by both disposable income and loan obtained. We use the total income and loan obtained over the last 12 months in 2008 representing for the initial wealth. The wealth variable (A_0) enters the model both linearly and through an interaction term. Applying this to equation (14) we get:

$$\Delta lnk_{ij} = \beta_1 \Delta PI_{ij} + \beta_2 \Delta PI_{ij} * A_{0ij} + \beta_3 A_{0ij} + u_j + \Delta\varepsilon_{ij} \quad (22)$$

In addition, when land serves as collateral, small producers may find themselves credit constrained and reduce their investment. Although the theoretical model does not allow us to distinguish the effects of credit constraints or increasing return to scale, the model examines the total effect of price shocks on investment conditional on land size. Similarly to the initial wealth, the land size in 2006 variable (L_0) enters linearly and as an interaction term with the price index.

$$\Delta lnk_{ij} = \beta_1 \Delta PI_{ij} + \beta_2 \Delta PI_{ij} * L_{0ij} + \beta_3 L_{0ij} + u_j + \Delta\varepsilon_{ij} \quad (23)$$

2.6. Estimation results

2.6.1.The effects on investment, saving and consumption in general

This section uses the theoretical and empirical models to test the hypothesis that households increased investment, or alternatively that households increased saving and consumption. As shown in the theoretical analysis, households gain income from positive price shocks and this leads to adjustment of their spending allocation. By taking advantage of the higher prices, households can boost investment in order to increase farm profit. In that analysis we assumed that households were risk-neutral. They prefer to increase investment in the absence of credit constraints to undertake the required amount of investment in relation to high future expected price. Instead, households may leave investment unchanged and increase saving or consumption if credit constraints impedes their ability to expand the cultivation. Households also can adjust their expenditure baskets depending on price of different goods. The model therefore also distinguishes the impact on total expenditure and food expenditure.

Table 2.8 reports the first-difference estimation of equations (19, 20 and 21) to evaluate the effects of overall the price spike. Panel A shows result of price index measured by share of total output revenue, and Panel B shows result of price index measured by share of sole output revenue. The impacts on the investment, saving and food expenditure are positive, with an exception of total expenditure. The effects of both price indices are quite similar. The first column shows statistically significant positive effect of price shocks on investment at less than the 1 percent level. This implies that households which experienced higher level of price shocks expanded production by increasing fixed assets investment. The effect of the price shocks on saving is positive, but not statistically significant. Hence, there is no evidence that households are influenced by precautionary motives to save in order to smooth consumption. In addition, estimation reveals that the price shocks have a negative effect on total expenditure and a positive effect on food expenditure, but again the effects are not statistically significant.

2.6.2. The effects in relation to market imperfections

This section considers the effects of the price shocks on decisions in relation to investment, savings and consumption decisions incorporating with financial conditions (income, loan and land) levels as a measure of financial constraint. There are two options to measure those initial financial variables in terms of timing, in 2006 or in 2008. Each wave of survey collected the information over the past 12 months. For instance, household income in the 2006 VHLSS survey was the income that households obtained during previous 12 months until the day they were interviewed in 2006. We use financial conditions in 2006 as a proxy for initial financial variables. This is in line with the period of price shocks used between the two waves of survey. The followed results are estimated from the equations (22, 23).

Income

In the theoretical model, less wealthy households may decrease investment or leave it unchanged and increase savings and consumption due to borrowing constraints. They also may increase investment due to the higher marginal return to capital. Financial wealth is firstly measured by income per capita in 2006. Table 2.9 shows the estimates of price shocks conditional on income per capita in 2006 as initial financial situation. The interaction term between income and price shocks has significant positive parameters on investment and saving, and significant negative parameter on total expenditure. This parameter is negative on food expenditure but not statistically significant. This shows that the effect of price shocks has a positive effect on investment conditional on available income in 2006. To compare the magnitude of the effect, we need to compute the marginal effect. From the estimated model in Panel A, the marginal impact of price shocks during 2006-08 on investment is: $\frac{\partial \Delta k}{\partial \Delta PI} = -4.185 + 0.574 * A$, where the mean of total income per capita in 2006 is 8.65, hence the marginal effect for mean income is 0.78. This effect is interpreted as, at mean of income, investment increases 78 percent point for every one unit increase in price index. Analogously, the marginal effect of the price shock index measured by share of sole revenue is: $\frac{\partial \Delta k}{\partial \Delta PI} = -1.620 + 0.296 * A$, and the effect is 0.94. As in the general case, this finding shows that the more market-oriented households are likely to response more to positive price shocks.

Access to loans

Table 2.10 shows the estimated results of price shock impact conditional on value of loan obtained. The parameters of the interaction term between the loan variable and price shocks are statistically significant positive for only investment. Based on the parameters of the direct effect in Panel A, the interaction term and mean of loan obtained at 3.286, $\frac{\partial \Delta k}{\partial \Delta P} = 0.503 + 0.111 * L$, we can compute the effect to be equal to 0.868. Similarly, the effect of the price index measured by sales value in Panel B is 0.979 and higher than the price index of total value in Panel A . This implies that households those which had access to higher loan are likely to increase investment, and households with more market-oriented productions respond more. Thus, as in the case of income, households with higher loans have more incentive to increase investment, but not saving. The effect of loans has a lower magnitude of parameter. This could be explained as value of loan accounts for only a fraction of total income. The household may use alternative financial sources, for example, from remittances, social welfare, etc. which could constitute a larger share in total income of household than loan obtained (as shown in descriptive statistic, Table 2.7).

Farmland

As referred above, the survey only provides information on land used to grow crops in the previous 12 months. This could be lower or higher than the actual area landholdings. To match with the income and loan variables above, the analysis test for land used in 2006. As in the empirical framework, the farmland is a factor of production and can become a financial source when it serves as collateral for obtaining credit. The larger the producer is, the higher possibility he has access to a greater value of credit. Thus, the theoretical model predicts that the smaller producers have less incentive to increase investment due to the borrowing constraint when facing positive income shocks. Instead, they will increase saving and consumption.

Table 2.11 shows that the results of the impact of price shock when interacted with farm land. The land used in linear variable and interaction variable have statistically significant effect on investment only with the price index of total value in Panel A. Interaction term of land size and price shock negatively affects on investment. With the mean of land size in 2006 is 2.2, the marginal effect of the price shock on food spending conditional on mean of land size, $\frac{\partial \Delta \mathrm{k}}{\partial \Delta PI} = 0.801 - 0.216 * Land$, is 0.326. Therefore, the larger farmer tends to decrease investment in response to higher price shocks.

Tobit model to deal with zero investment

The results report significant and consistent effect of price shock on investment and also in relation with financial conditions. We now further focus on investment equation by other empirical test. We observe in this sample that there are around 11 percent of households who did not spend on fix assets. Therefore, the Tobit model is employed to control the effect of households with zero investment against the case of households with positive investment.

Table 2.12 shows estimated effects of price shock on investment. All tests still take into consideration of regional dummy variables and area-province clusters. In both Panel A and B, price shocks have positive effects on investment, and interaction terms with income and loan have positive effect on investment, except for the case of land with negative effect. The results are all similar to those of first-difference tests in terms of magnitude and sign.

The effects in 3D space charts

In order to visualize the estimation results with interaction term, I illustrate the effects on investment of income, loan obtained, and land in hyper-plane graphs. Based on the statistically significant results, the effect on dependent variables are computed from estimated parameters of independent variables. A three dimensional surface chart is created on three-axis space which represents for these three variables in each regression function. Scale of each axis falls into the distribution (min and max values) of independent variable that axis represents. Colours on the surface chart are associated with changes of the value or the effect on dependent variable and specified in the data band. The lighter the colour indicates the higher value of effect. The lines between data bands and lines within data bands corresponds to gridlines on the chart walls and chart floor.

Figure 2.6A shows the effect of investment on fix asset investment conditional on income. As we can see, the surface of the chart in the region of lowest income moves to downward in investment axis along increasing price index. Whereas, the surface of the chart in the region of higher income constantly converts to upward trend and appears to be steepest at highest income. We can alternatively navigate that, at the lowest price index, the area of the chart goes down along higher income, which shows a decreasing trend of investment. In contrast, at the highest price index, the area of surface goes up along higher income, and hence shows the revert effect of investment. Figure 2.6B depicts the effect on investment of the price shocks conditional on loan obtained. In this case, the chart shows that change in effect on investment of higher price index is less significant between lowest and highest loan. This also is referred as the lower parameter of interaction terms of loan than that of income. In contrast, Figure 2.6C illustrates the negative effects of the price shocks on investment conditional on land size, respectively.

2.7. Conclusion

This study has investigated the short-term impact of the food price crisis on productive investment, saving and expenditure decisions in Vietnam by using panel data covering 2006 and 2008 and in conjunction with regional food price indices. The empirical results show a positive impact of higher food prices on only fix asset investment over the crisis period. The impact of the price shocks are positive on saving and negative on total expenditure, but both are not statistically significant. Moreover, when the price shocks are incorporated with financial condition, the findings reveal that the effects of household income, loan obtained and land size matter. Higher-income households tended to invest more in response to higher level of the price shocks and those with higher loan obtained also tended to invest more. Whereas the higher-income households and larger-landholding households respectively were likely to decrease total expenditure and food expenditure. Higher-income households also saved more to respond with higher level of the price shocks. This implies that low-income households tend to be more constrained as the result of imperfect financial markets and hence are unable to increase productive investment as much as they might have wished.

Based on this empirical evidence, the study shows the important role of capital in investment decision in production with higher expected profit. The poor farm households are under-capitalized and subject to be more vulnerable. The positive food price shocks not only caused the negative impact of the poor (since they are likely to be net food consumers) but also broaden the gap with the rich farmers. The implication therefore suggests that mitigating the credit constraint can both reduce poverty and increase the investment in productive assets. Hence the efficiency in input-allocation can lead to higher profit and smoother consumption. This binding can be relaxed through mechanisms or policies that provide and assist external financial sources for farmers. Consequently, the investment could further accelerate the long-term benefit and push upward agricultural production. [13]

[13] The preview of data in Table 2.13 reports the significant increase in income by an increase in production investment during global crisis in short-term.

References

Abbott, C. P. , Hurt, C. , & Tyner, W. E. . (2009). What's driving food price? *Farm Foundation Issue Report*. Oak Brook, IL, USA: Farm Foundation.

Aksoy, M. A., & Isik-Dikmelik, A. (2008). Are low food prices pro-poor? Net food buyers and sellers in low-income countries *Policy Research Working Paper 4642*: The World Bank.

Arndt, C., Garcia, A. , Pham, H.H. , & Thurlow, J. . (2009). *A Provisional 2007 Social Accounting Matrix for Vietnam*. Department of Economics, University of Copenhagen, and Central Institute of Economic Management: Hanoi.

Benson, T. , Minot, N., Pender, T. , Robles, M., & Braun, J. v. . (2008). Global food crisis: Monitoring and Assessing Impact to Inform Policy Responses: IFPRI Issue Brief 55, Washington, D.C. .

Busjeet, G., Demombynes, G., & Sobrado, C. (2008). Food prices and the poor in Latin America *LAC Poverty and Gender Group, The World Bank, Washington, D.C.*

Charlebois, P. , & Hamann, N. . (2010). *The Consequences of a Strong Depreciation of the U.S. Dollar on Agricultural Markets*. Agriculture and Agri-Food Canada.

Dawe, D. , & Slayton, T. (2010). The World Rice Market Crisis of 2007–2008. In D. Dawe (Ed.), *The Rice Crisis*. London: Earthscan.

Deaton, A. (1990). *Saving in Developing Countries: Theory and Review.* Paper presented at the Proceedings of The World Bank Annual Conference on Development Economics.

FAO. (2008). *Growing demand on agriculture and rising prices of commodities: An opportunity for smallholders in low-income, agricultural-based countries?* Paper prepared by the Trade, Markets, and Agricultural Development Economics Division of FAO for the Round Table, 31st session of IFAD's Governing Council.

Ferreira, F. H. G., Fruttero, A., Leite, P., & Lucchetti, L. . (2011). Rising Food Prices and Household Welfare: Evidence from Brazil in 2008 *Policy Research Working Paper 5652*: The World Bank.

Fujii, T. (2011). Impact of food inflation on poverty in the Philippines *Research Collection School of Economics. Paper 1315*.

Gilbert, C. L. . (2010). How to understand high food prices? *Journal of Agricultural Economics, 61*(2), 398–425.

GSO. (2008). *Operational Handbook: Vietnam Household Living Standard Survey 2008*. Hanoi, Vietnam

Headey, D. (2011). Rethinking the global food crisis: The role of trade shocks. *Food Policy, 36*(2), 136-146.

Headey, D., & Fan, S. (2010). *Refections on the Global Food Crisis*. IFPRI. Washington D. C., U.S.A.

IMF. (2008). Food and Fuel Prices - Recent Developments, Macroeconomic Impact, and Policy Responses. In International Monetary Fund Report (Ed.), *Prepared by the Fiscal Affairs, Policy Development and Review, and Research Departments*.

Ivanic, M., & Martin, W. (2008). Implications of higher global food prices for poverty in low-income countries *Policy Research Working Paper Series 4594, The World Bank.*

Lustig, N. (2009). Coping with rising food prices: Policy dilemmas in the developing world *Working Paper Number 164*. Washington, D.C.: The Center for Global Development.

Mitchell, D. (2008). A note on rising food prices. Washington, D.C.: World Bank.

Nielsen, C. P. . (2003). Vietnam's Rice Policy: Recent Reforms and Future Opportunities. *Asian Economic Journal, 17*(1), 1-26.

Nose, M. , & Yamauchi, F. . (2012). Long-term Impacts of Global Food Crisis on Production Decisions: Evidence from Farm Investments in Indonesia *Policy Research Working Paper 6065*: The World Bank.

Nuryartono, N. . (2006). *Impact of Smallholders' Access to Land and Credit Markets on Technology and Land Use Decisions: The Case of Tropical Forest Margins in Central Sulaewsi-Indonesia*: Cuvillier Verlag Gottingen.

Paxson, C. (1991). Using weather variability to estimate the response of savings to transitory income in Thailand. *American Economic Review, 8*(1).

Pham, H. N. (2010). The Vietnamese Rice Industry During the Global Food Crisis. In D. Dawe (Ed.), *The Rice Crisis* (pp. 219-232). London: Earthscan.

Phung, D. T., & Nguyen, P. . (2006). Vietnam Household Living Standard Survey (VHLSS) 2002 and 2004: Basic Information.

Phung, D.T., & Waibel, H. (2010). Food Price Crisis, Poverty and Welfare in Vietnam: An Ex-post Decomposition Analysis *Working Paper. Leibniz Universität Hannover, Institute of Development and Agricultural Economics, Germany*.

Thurlow, J., Tarp, F., McCoy, S. , Nguyen, M. H. , Breisinger, C. , & Arndt, C. . (2010). The Impact of the Global Commodity and Financial Crises on Poverty in Vietnam *New Directions in Development Economics. Working Paper No. 2010/98*: UNU-WIDER, Helsinki.

Timmer, P. (2009). *Did Speculation Affect World Rice Prices?* Agricultural and Development Economics Division of the Food and Agriculture Organization of the United Nations.

Timmer, P. . (2010). Did Speculation Affect World Rice Prices? In D. Dawe (Ed.), *The rice Crisis*: Earthscan.

Trostle, R. (2008). Global agricultural supply and demand: Factors contributing to the recent increase in food commodity prices. Washington, D.C.: United States Department of Agriculture.

Udry, C. (1995). Risk and saving in northern Nigeria. *The American Economic Review, 85*(5), 1287-1300.

Valero-Gil, J. N., & Valero, M. (2008). The effects of rising food prices on poverty in Mexico. *Agricultural Economics, 39*(2008), 485–496.

Vu, L. , & Glewwe, P. . (2011). Impacts of rising food prices on poverty and welfare in Vietnam. *Journal of Agricultural and Resource Economics, 36*(1), 14-27.

Wodon, Q., Tsimpo, C., & Coulombe, H. . (2008). Assessing the impact on poverty of rising cereal prices: The case of Ghana *Policy Research Working Paper 4740*: The World Bank.

Wodon, Q., & Zaman, H. (2008). Rising food prices in Sub-Saharan Africa: Poverty impact and policy responses *Policy Research Working Paper 4738*: The World Bank.

World Bank. (2009). Global economic prospects 2009: Commodities at a crossroads. Washington, D.C.

World Bank. (2011). Boom, Bust and Up Again? Evolution, Drivers and impact of Commodity Prices: Implications for Indonesia *A Trade Development Report.*

Zezza, A., Davis, B. , Azzarri, C. , Covarrubias, K. , Tasciotti, L. , & Anriquez, G. . (2008). The impact of rising food prices on the poor *ESA Working Paper No 08-07*: FAO (Food and Agriculture Organization of the United Nations).

Tables

Table 2.1: Structure of Vietnam economy in 2007

	Share of total (%)				Export intensity*	Import intensity**
	GDP	Employment	Exports	Imports		
Total GDP	100	100	100	100	100	100
Agriculture	22.1	53.9	7.6	2	21.3	8.4
Crops	13.4	36.6	4.6	1.4	23.3	10.3
Livestock	2.7	8.5	0.4	0	7.4	1.6
Forestry/Fishing	5.9	8.8	2.6	0.6	23.4	8.2
Industry	41.7	19.9	76.1	85.4	38.5	48.7
Mining	10.9	0.9	19	0.6	82.9	14.9
Manufacturing	20.1	13.3	57.1	84.9	40.7	57.8
Agro-processing	5.8	4.1	12.5	4.3	33.5	20.3
Textiles/clothing	3.7	2.2	25.8	15.3	68.6	62.2
Wood/paper	1.4	0.8	2.9	3	34.9	44.2
Fuel/chemicals	2.8	2.3	3.2	22.5	21.8	74.7
Metals/machinery	4.4	3	11.9	38.5	36.8	70.4
Other	2.8	1.4	1.4	2.7	8.1	17.6
Others	10.7	5.7	0	0	0	0
Services	36.2	26.1	16.3	12.6	22	20.5

Source: Calculation of Thurlow et al. (2010) using the 2007 social accounting matrix Arndt et al. (2009)
Note: (*) Export intensity is the share of exports in gross domestic output. (**) Import intensity is the share of imports in total demand of input.

Table 2.2: Share of crop production in 2006

	Rice	Other staples	Vegetables	Total sample*	Total sample / Total survey
Red River Delta	83.13	6.69	10.18	100	85.54
North East	65.40	20.04	14.57	100	77.95
North West	56.14	30.53	13.33	100	82.55
North Central Coast	72.41	15.95	11.64	100	73.97
South Central Coast	75.36	15.36	9.28	100	81.25
Central Highlands	48.38	33.12	18.50	100	63.40
North East South	58.52	29.18	12.30	100	73.16
Mekong River Delta	83.52	1.88	14.60	100	84.89
Total	72.52	14.87	12.61	100	79.85

Source: Calculation from the 2006 VHLSS.
Note: (*)Total sample refers to crops used in the paper, which are excluded perennial, industry, fruits from total crops collected in the survey.

Table 2.3: Basic information of farm households

Panel A: Sample from VHLSS 2006								
				Quantiles				
Variable	N	Mean	S.D.	Min	0.25	Median	0.75	Max
Land (acres)	2,531	2.09	3.55	0	0.65	1.26	2.17	72.03
Crop expenditure (total)	2,285	4,339	9,553	2	1,271	2,256	3,820	190,264
Crop expenditure (per area)	2,285	2,154	2,794	82	1,400	1,961	2,443	93,474
Crop income (total)	2,303	5,583	8,065	-29,462	1,953	3,866	6,190	182,852
Crop income (per area)	2,303	3,740	11,523	-85,684	2,150	2,860	3,699	492,241
Crop income (per capita)	2,303	1,262	1,706	-9,821	499	896	1,462	36,570
Household size	2,303	4.52	1.66	1	4	4	5	15
Age of hh head	2,303	47.77	13.01	19	38	46	56	89
Education of hh head	2,303	2.69	0.83	1	2	3	3	5

Source: Calculation from the 2006 VHLSS.
Note: All the values are deflated to price in January 2006

Panel B: Sample from VHLSS 2008								
				Quantiles				
Variable	N	Mean	S.D.	Min	0.25	Mdn	0.75	Max
Investment *	2,053	461	4147	1	26	55	118	142041
Land (acres)	2,303	2.32	5.08	0.00	0.72	1.28	2.24	158.08
Crop expenditure (total)	2,282	5,418	16,751	1	1,336	2,458	4,384	533,775
Crop expenditure per land	2,282	2,498	5,837	23	1,514	2,156	2,807	188,799
Crop income (total)	2,303	6,930	17,118	-47,607	2,228	4,268	7,236	606,523
Crop income per land	2,303	4,048	7,636	-160,618	2,387	3,306	4,266	194,793
Crop income (per capita)	2,303	1,652	4,846	-9,521	560	1,047	1,756	202,174
Household size	2,303	4.39	1.71	1	3	4	5	14

Source: Calculation from the 2008 VHLSS.
Note: All values are deflated to price in January 2006. (*) Asset investment is decomposed into groups of items purchased in the table below.

Table 2.4: Investment on fixed assets, 2007-08 (VND 1,000)

					Quantiles			
Variables	Obs.	Mean	S.D.	Min	0.25	Median	0.75	Max
Agricultural tools*	2036	87	156	1	24	47	95	3,591
Minor repair assets**	314	124	310	2	22	49	100	3,344
Rice milling machine	9	4,570	2,528	865	2,847	3,472	6,636	8,329
Harvesting machine	13	6,189	12,702	266	530	641	8,742	46,418
Pesticide sprayer	7	6,820	15,034	230	1,062	1,235	1,878	40,896
Tractor	4	35,379	56,194	883	3,609	10,710	67,148	119,211
Tractor plough	22	12,737	14,998	1,272	4,695	8,671	15,432	72,055
Cart	20	962	508	412	596	897	1,097	2,481
Pump	42	2,828	4,542	177	733	1,461	3,194	26,780
Power generator	3	3,312	2,642	581	581	3,498	5,856	5,856
Total	2060	463	4,141	1	26	55	119	142,041

Source: Calculation from the 2008 VHLSS.
Note: All values are deflated to price at January 2006. (*)Agricultural tools include small and non-durable tools such as: sickles, shears, shovels, jackknifes, etc. (*) and (**) are expenditure of production which spent during last 12 months. The other items are fixed assets which were bought during 2007-2008 and have value over 500,000 VND at time purchased.

Table 2.5: Variation of monthly commodity price indices, 2006-08

	Coefficient of variation	Autocorrelation	
		1 month	6 months
Rice	0.29	0.98	0.71
Other staples	0.22	0.99	0.93
Vegetables	0.12	0.98	0.76

Source: Calculation from data of GSO
Note: Period during the two rounds of survey, from May 2006 to November 2008

Table 2.6: Variable description

1	Price volatility	The sum of change in price indices of 3 crops (rice, other staples and vegetables) during the period between the two surveys 2006-2008 and weighted by revenue shares of each crops.
2	Investment	Log of total expenditure on fixed assets during 2007-2008 (table 4)
3	Saving	Log difference of saving amount (total income minus total expenditure) between 2006 and 2008.
4	Living expenditure	Log difference of food, nonfood and other annual consumption between 2006 and 2008.
5	Fix/Durable expenditure	Log difference of spending on repaired and purchased fix assets, durable goods, and house/land between 2006 and 2008.
6	Other expenditure	Log difference of debt repayment, lending, buying gold, or silver, deposit, buying insurance, big investment, other kind of fee, tax, donation, etc. between 2006 and 2008
7	Total income	Log disposal income in 2006 and 2008
8	Income 1	Log income from wages, farm and non-farm, and business in 2006 and 2008
9	Income 2	Log income from remittances, social allowance, pension, insurance, subsidies, interest of saving account, bonds, etc. in 2006 and 2008
10	Income 3	Log income from selling means, leasing equipment, withdrawing account, borrowing, advanced payment, etc. in 2006 and 2008
11	Loan	Log of loan obtained in 2006 and 2008
12	Land 2006 and Land 2008	Size of land used for cultivating crops (acres) in 2006 and 2008.

Note: Variables (except 1 and 12) are adjusted in per capita and deflated in January 2006.

Table 2.7: Financial sources of households in 2006 and 2008 (VND 1,000)

	2006			2008		
	Continuing HH[a]	Out HH[b]	In HH[c]	Continuing HH	Out HH	In HH
Total income	7,305	9,528	11,115	8,386	12,406	11,141
Income 1	4,850	6,007	5,848	5,499	7,008	6,578
Income 2	1,094	2,014	2,847	965	2,305	1,811
Income 3	1,360	1,506	2,419	1,921	3,093	2,752
Total expenditure	6,217	8,050	9,067	7,818	12,442	9,648
Food	2,103	2,546	2,629	2,450	2,884	2,975
Non-food	1,829	2,720	2,813	2,257	2,864	3,110
Fix/durable	1,284	1,185	1,736	1,700	3,158	1,408
Other spending	1,472	2,196	2,688	2,094	4,226	3,066
Saving= Inc. - Exp.	1,579	2,308	2,466	1,093	1,284	2,281
Loan	990	1,782	1,731	2,138	1,224	2,364
Obs.	2,303	225	168	2,303	225	168
% of obs.	85.28	8.51	6.21	85.28	8.51	6.21

Source: Calculation from the 2006 and 2008 VHLSS.
Note: [a, b, c] respectively refer to households continued to cultivate crops during the period of 2006-08, households stopped cultivating in 2008, and households started cultivating in 2008. The values are per capita and deflated in January 2006.

Table 2.8: Estimated effect of price shocks, 2006-08

Panel A: Price index measured by share of output revenue				
	Investment	Saving	Total exp.	Food exp.
Price Index	0.903***	0.368	-0.008	0.043
	(0.224)	(0.309)	(0.065)	(0.035)
Intercept	2.051***	-1.809***	0.194	0.081
	(0.419)	(0.570)	(0.120)	(0.064)
Regional dummies	Yes	Yes	Yes	Yes
Area-Province clusters	Yes	Yes	Yes	Yes
N	2,303	2,303	2,303	2,303
R-squared	0.040	0.008	0.004	0.024

Panel B: Price index measured by share of sole revenue				
	Investment	Saving	Total exp.	Food exp.
Price Index	1.014***	0.083	-0.033	0.033
	(0.114)	(0.227)	(0.043)	(0.024)
Intercept	1.752***	-1.299***	0.242***	0.095**
	(0.240)	(0.449)	(0.087)	(0.046)
Regional dummies	Yes	Yes	Yes	Yes
Area-Province clusters	Yes	Yes	Yes	Yes
N	2,303	2,303	2,303	2,303
R-squared	0.079	0.008	0.005	0.024

Note: ***, **, * denote the level of significance at 1%, 5%, and 10%. Standard errors are in parentheses and clustered at the area (rural/urban) and provincial level. The seven regional dummies are included but not reported.

Table 2.9: Estimated effect of price shocks and income by first-difference model

Panel A: Price index measured by share of output revenue				
	Investment	Saving	Total exp.	Food exp.
Price Index (PI)	-4.185**	-2.177	0.266	-0.312
	(1.696)	(3.753)	(0.814)	(0.392)
Total income * PI	0.574***	0.277	-0.037	0.038
	(0.196)	(0.431)	(0.094)	(0.045)
Total income	-1.060***	-1.019	-0.228	-0.194**
	(0.358)	(0.774)	(0.172)	(0.081)
Intercept	11.443***	7.274	2.244	1.811**
	(3.127)	(6.765)	(1.484)	(0.713)
Regional dummies	Yes	Yes	Yes	Yes
Area-Province clusters	Yes	Yes	Yes	Yes
N	2,303	2,303	2,303	2,303
R-squared	0.045	0.014	0.078	0.064

Panel B: Price index measured by share of sales revenue				
	Investment	Saving	Total exp.	Food exp.
Price Index (PI)	-1.620*	-0.803	0.072	-0.206
	(0.960)	(2.125)	(0.632)	(0.237)
Total income * PI	0.296***	0.089	-0.018	0.024
	(0.107)	(0.240)	(0.073)	(0.027)
Total income	-0.514**	-0.696	-0.267*	-0.171***
	(0.207)	(0.432)	(0.142)	(0.052)
Intercept	6.345***	4.952	2.642**	1.629***
	(1.876)	(3.853)	(1.226)	(0.460)
Regional dummies	Yes	Yes	Yes	Yes
Area-Province clusters	Yes	Yes	Yes	Yes
N	2,303	2,303	2,303	2,303
R-squared	0.082	0.013	0.081	0.064

Note: ***, **, * denote the level of significance at 1%, 5%, and 10%. Standard errors are in parentheses and clustered at the area (rural/urban) and provincial level. The seven regional dummies are included but not reported.

Table 2.10: Estimated effect of price shocks and loan by first-difference model

Panel A: Price index measured by share of output revenue				
	Investment	Saving	Total exp.	Food exp.
Price Index (PI)	0.503**	0.227	0.002	0.048
	(0.250)	(0.380)	(0.076)	(0.044)
Loan * PI	0.111***	0.054	-0.006	-0.001
	(0.040)	(0.074)	(0.011)	(0.007)
Loan	-0.198***	-0.021	-0.007	0.005
	(0.074)	(0.140)	(0.020)	(0.012)
Intercept	2.773***	-1.800**	0.235	0.060
	(0.474)	(0.716)	(0.144)	(0.079)
Regional dummies	Yes	Yes	Yes	Yes
Area-Province clusters	Yes	Yes	Yes	Yes
N	2,303	2,303	2,303	2,303
R-squared	0.046	0.013	0.014	0.025

Panel B: Price index measured by share of sales revenue				
	Investment	Saving	Total exp.	Food exp.
Price Index (PI)	0.785***	-0.185	-0.008	0.043
	(0.138)	(0.304)	(0.059)	(0.031)
Loan * PI	0.059***	0.074	-0.007	-0.002
	(0.020)	(0.051)	(0.009)	(0.004)
Loan	-0.109***	-0.065	-0.004	0.008
	(0.040)	(0.108)	(0.018)	(0.008)
Intercept	2.182***	-1.029*	0.254**	0.064
	(0.292)	(0.612)	(0.122)	(0.059)
Regional dummies	Yes	Yes	Yes	Yes
Area-Province clusters	Yes	Yes	Yes	Yes
N	2,303	2,303	2,303	2,303
R-squared	0.083	0.013	0.015	0.025

Note: ***, **, * denote the level of significance at 1%, 5%, and 10%. Standard errors are in parentheses and clustered at the area (rural/urban) and provincial level. The seven regional dummies are included but not reported.

Table 2.11: Estimated effect of price shocks and land by first-difference model

Panel A: Price index measured by share of output revenue				
	Investment	Saving	Total exp.	Food exp.
Price index (PI)	0.801***	0.207	-0.005	0.070*
	(0.210)	(0.357)	(0.077)	(0.041)
Land * PI	-0.216**	0.046	0.004	-0.019
	(0.080)	(0.152)	(0.022)	(0.015)
Land	0.570***	-0.032	-0.011	0.033
	(0.168)	(0.310)	(0.045)	(0.028)
Intercept	1.886***	-1.637**	0.196	0.039
	(0.392)	(0.665)	(0.141)	(0.073)
Regional dummies	Yes	Yes	Yes	Yes
Area-Province clusters	Yes	Yes	Yes	Yes
N	2,303	2,303	2,303	2,303
R-squared	0.107	0.010	0.005	0.027

Panel B: Price index measured by share of sales revenue				
	Investment	Saving	Total exp.	Food exp.
Price index (PI)	0.837***	-0.007	-0.028	0.050**
	(0.113)	(0.241)	(0.049)	(0.025)
Land * PI	-0.027	-0.003	-0.002	-0.024
	(0.070)	(0.160)	(0.021)	(0.018)
Land	0.183	0.069	0.001	0.043
	(0.142)	(0.328)	(0.043)	(0.035)
Intercept	1.806***	-1.265***	0.239**	0.074
	(0.227)	(0.474)	(0.096)	(0.047)
Regional dummies	Yes	Yes	Yes	Yes
Area-Province clusters	Yes	Yes	Yes	Yes
N	2,303	2,303	2,303	2,303
R-squared	0.130	0.010	0.005	0.028

Note: ***, **, * denote the level of significance at 1%, 5%, and 10%. Standard errors are in parentheses and clustered at the area (rural/urban) and provincial level. The seven regional dummies are included but not reported.

Table 2.12: Estimated effect of price shocks by Tobit model

Panel A: Price index measured by share of output revenue				
	Investment			
Price index	1.074***	-4.726**	0.610**	0.985***
	(0.271)	(2.006)	(0.293)	(0.258)
Income*PI		0.654***		
		(0.233)		
Income		-1.224***		
		(0.430)		
Loan*PI			0.130***	
			(0.047)	
Loan			-0.235***	
			(0.087)	
Land * PI				-0.243***
				(0.090)
Land				0.631***
				(0.187)
Intercept	1.963**	12.786***	-0.593	1.329**
	(0.637)	(3.738)	(0.514)	(0.575)
Regional dummies	Yes	Yes	Yes	Yes
Area-Province clusters	Yes	Yes	Yes	Yes
N (censored)	2,303 (250)	2,303 (250)	2,303 (250)	2,303 (250)
Log pseudolikelihood	-4670	-4664	-4595	-4595

Note: ***, **, * denote the level of significance at 1%, 5%, and 10%. Standard errors are in parentheses and clustered at the area (rural/urban) and provincial level. The seven regional dummies are included but not reported.

Panel B: Price index measured by share of sole revenue				
	Investment			
Price index	1.207***	-2.338*	0.922***	1.025***
	(0.160)	(1.317)	(0.178)	(0.156)
Income*PI		0.399***		
		(0.150)		
Income		-0.720**		
		(0.294)		
Loan*PI			0.075***	
			(0.027)	
Loan			-0.142***	
			(0.054)	
Land * PI				-0.243***
				(0.090)
Land				0.631***
				(0.187)
Intercept	1.748***	8.137**	2.289***	1.329**
	(0.431)	(2.634)	(0.469)	(0.575)
Regional dummies	Yes	Yes	Yes	Yes
Area-Province clusters	Yes	Yes	Yes	Yes
N	2,303 (250)	2,303 (250)	2,303 (250)	2,303 (250)
Log pseudolikelihood	-4618	-4614	-4613	-4595

Note: ***, **, * denote the level of significance at 1%, 5%, and 10%. Standard errors are in parentheses and clustered at the area (rural/urban) and provincial level. The seven regional dummies are included but not reported.

Table 2.13: Preview of survey data

Panel A: Log of increase in crop income across quantiles of investment					
Quantiles of investment	Q1	Q2	Q3	Q4	Q5
Increase in crop income	-0.237	-0.065	0.152	0.204	0.316

Panel B: Log of investment across quantiles of price volatility					
Quantiles of price volatility	Q1	Q2	Q3	Q4	Q5
Investment	3.247	3.825	3.933	3.814	3.573

Note: *** significant at the 1%, ** significant at the 5 %, * significant at the 10 %. Standard errors are in parentheses and clustered at the provincial level. The dummy variables are included in the test but not reported.

Table 2.14: Distribution of households' access to credit market in formal and informal markets

Panel A: Indonesia				
		Poorest	Poor	Less poor
Formal				
	% access	13	17	34
	Amount (US$)	3	26	154
Informal				
	% access	78	81	58
	Amount (US$)	11	13	19

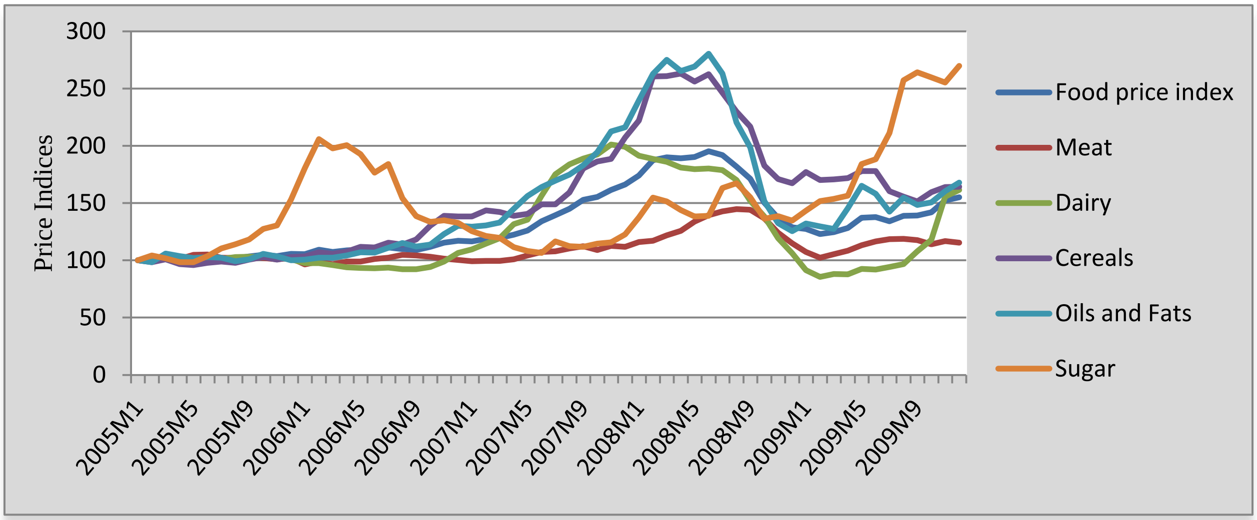
Source: Data is from survey in two districts of Central Sulawesi Province, Indonesia where the poverty rate is relatively high (46.1 % in 2004) compared to other districts (Nuryartono, 2006).
Note: Data is in poor group of the survey.

Panel B: Vietnam				
		Low	Medium	High
Formal				
	% access	29	35	43
	Amount (US$)	329	509	1,509
Informal				
	% access	19	20	19
	Amount (US$)	210	368	973

Source: Calculation from the 2006 VHLSS.
Note: Data of the survey covers the whole country.

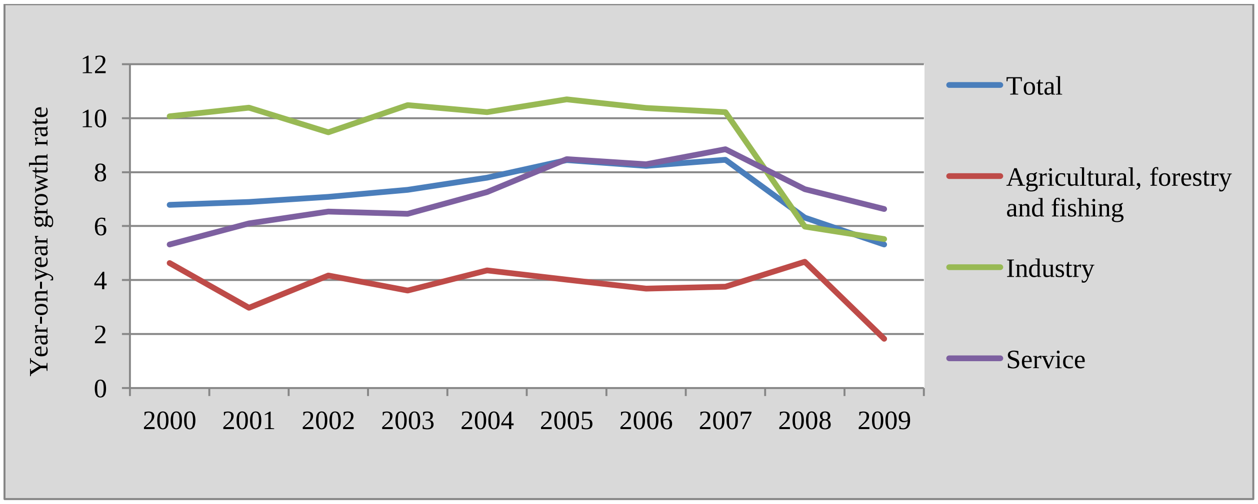
Figures

Figure 2.1: International FAO food price indices during 2005-2009



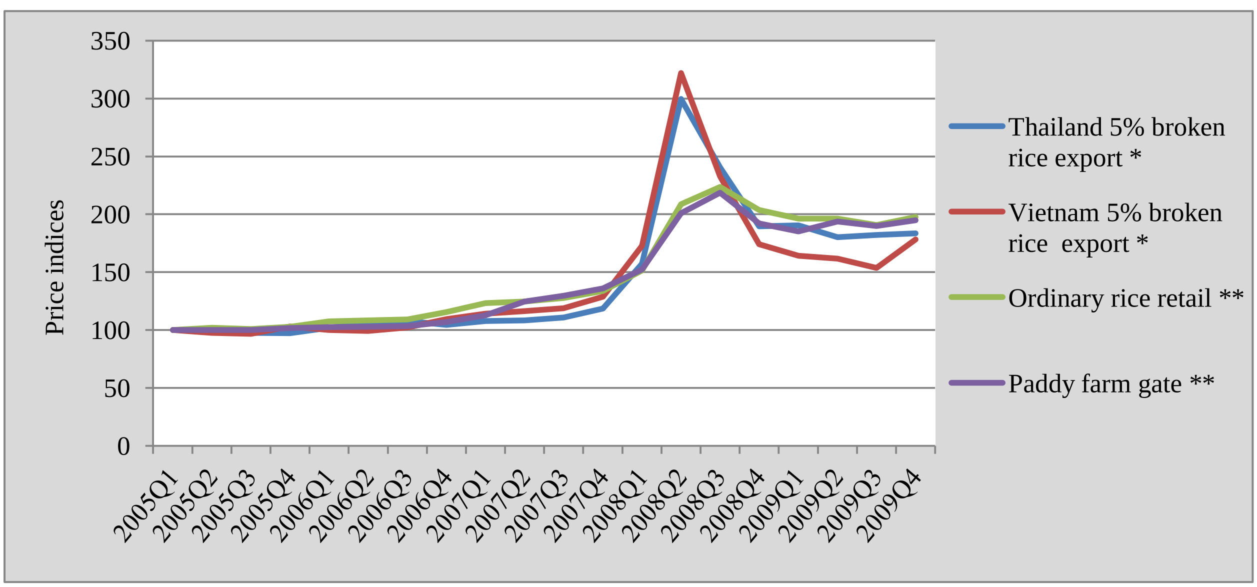
Source: FAO's database.
Note: The FAO food price index is a measure of the monthly change in international prices of a basket of food commodities. It consists of the average of five commodity group price indices, and weighted with the average export shares of each of the groups. Similarly, each of the followed five group (Meat, Dairy, Cereals, Oils and Fats, Sugar) commodities is computed from average prices of a basket of that specific group and also weighted by world average export trade shares. These price indices are normalized at 100 percent in January 2005.

Figure 2.2: Growth rates by sectors in Vietnam, 2000-2009



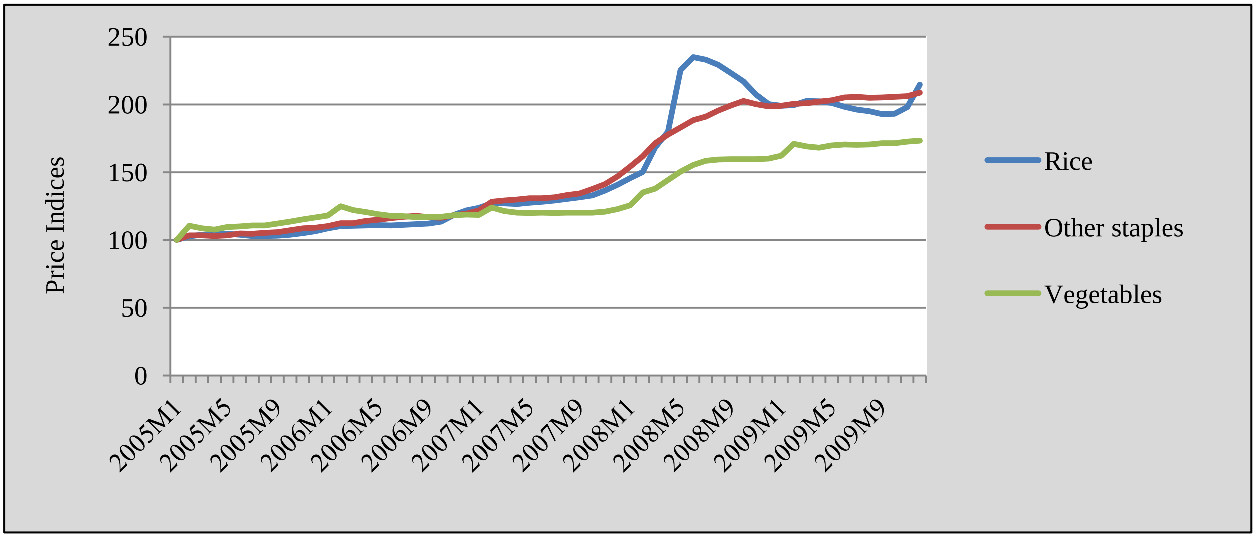
Source: Statistical yearbook of Vietnam 2009, GSO.

Figure 2.3: Price indices of rice in Thailand and Vietnam, 2005-2009



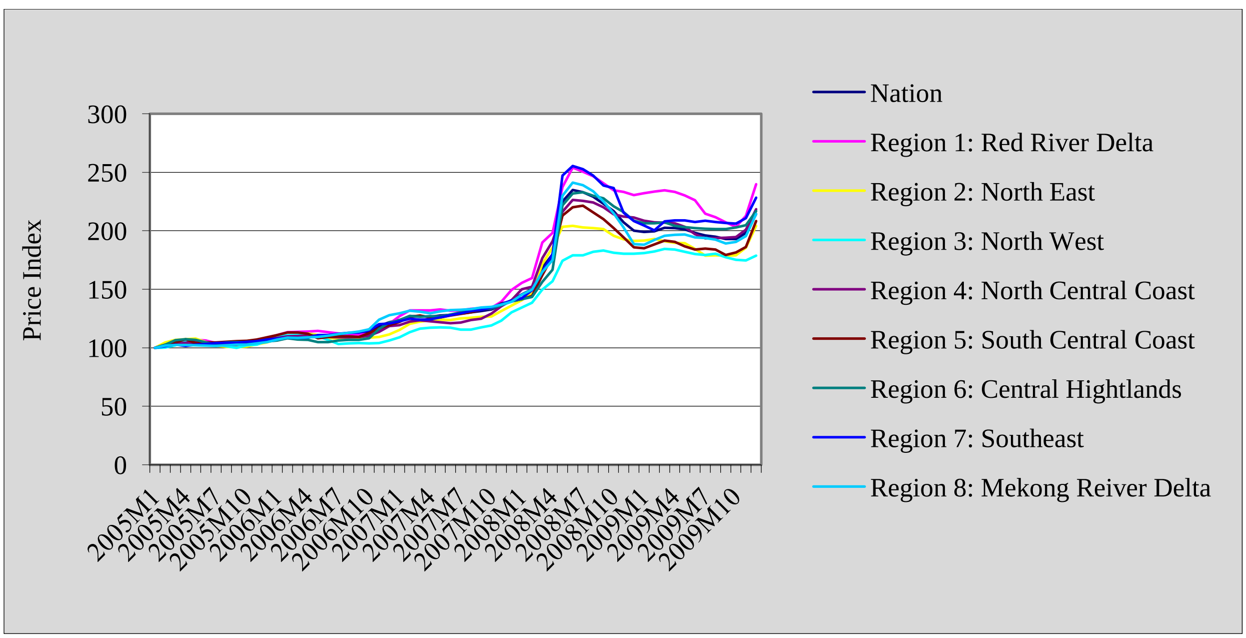
Source: (*) Calculation from data of Ministry of Finance, (**) calculation from data of GSO.
Note: (i) Ordinary rice retail and paddy farm-gate price indices are from Vietnam.
(ii) The prices are monthly available, except for farm gate price of paddy (unhusked rice), so all prices are converted to quarterly prices for comparable purpose. Prices are normalized at 100 percent in January 2005.

Figure 2.4: Monthly retail price indices of main crops in Vietnam, 2005-2009



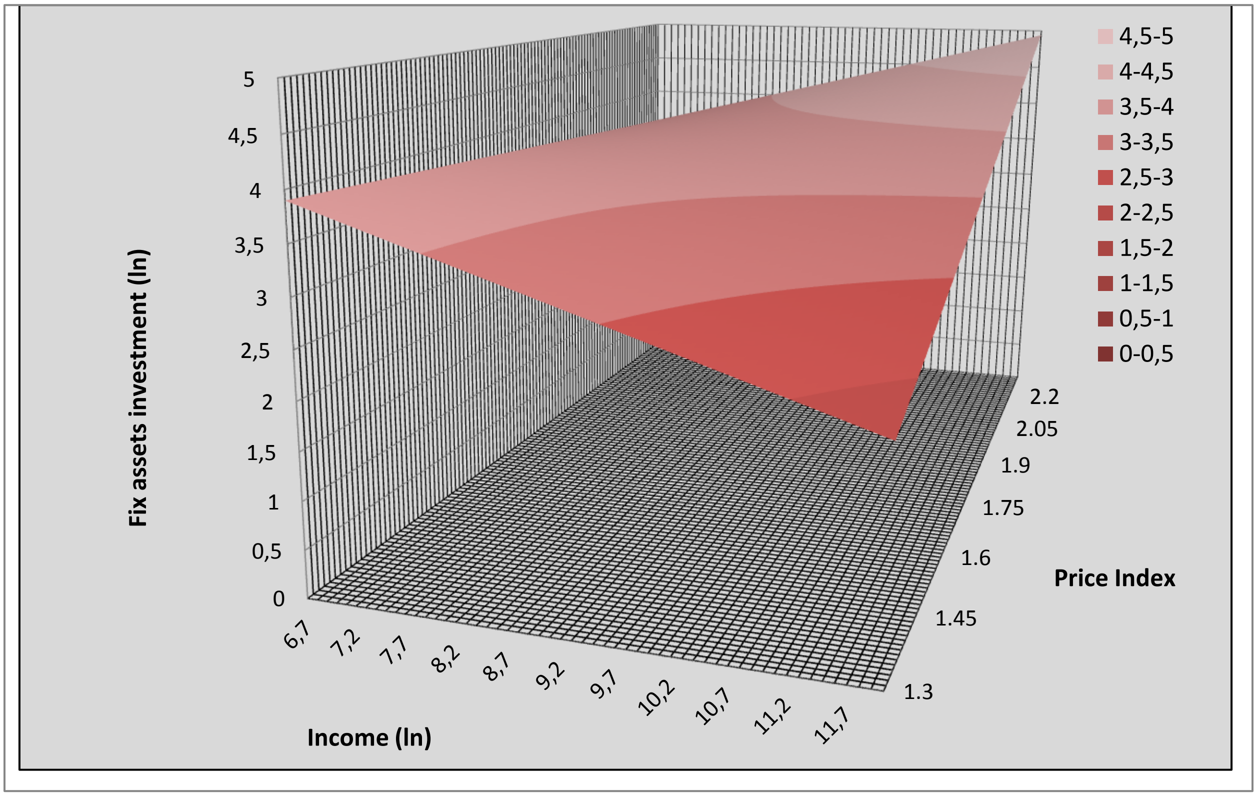
Source: Calculation from data of GSO.
Note: The price indices are normalized at 100 percent in January 2005.

Figure 2.5: Retail price indices of ordinary rice across regions in Vietnam, 2005-2009



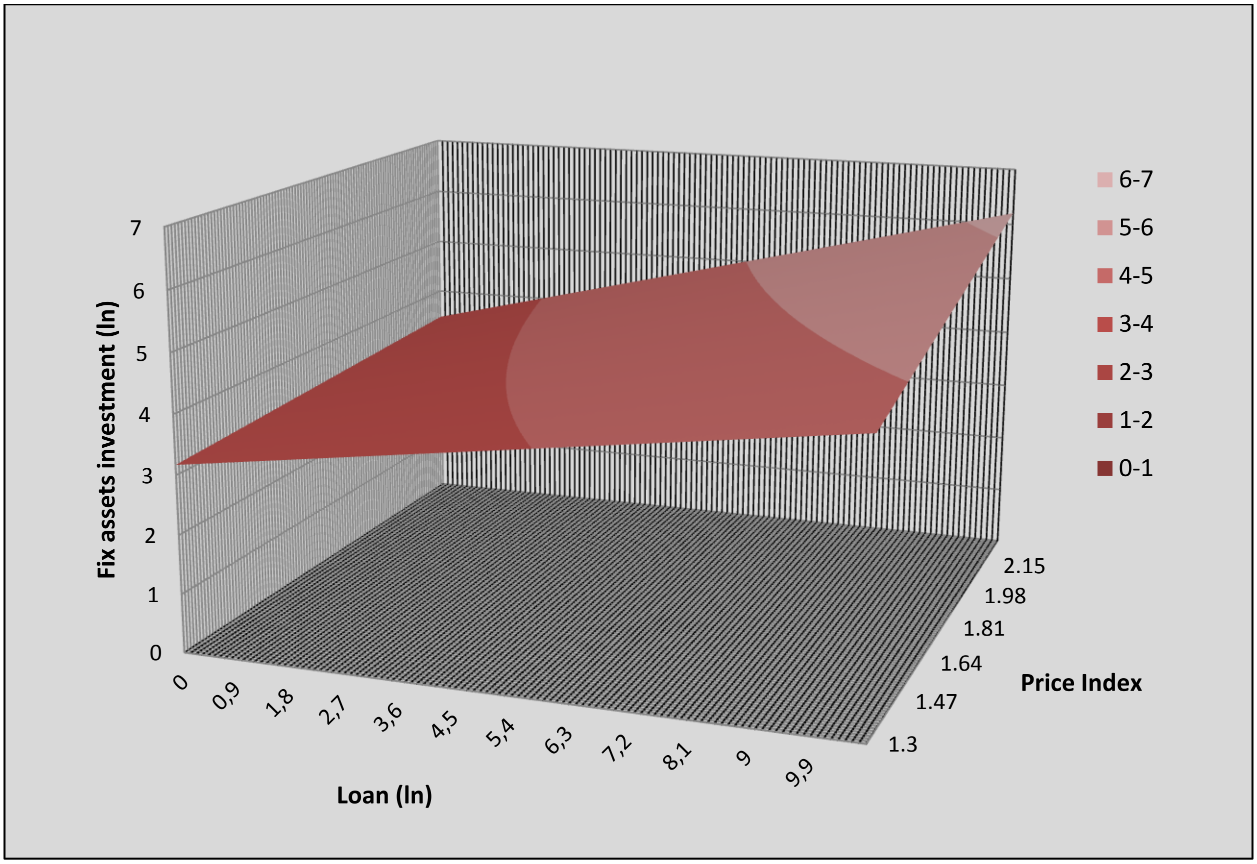
Source: Calculation from data of GSO
Note: The price indices are normalized at 100 percent in January 2005.

Figure 2.6A: Effect on investment of price shocks conditional on income



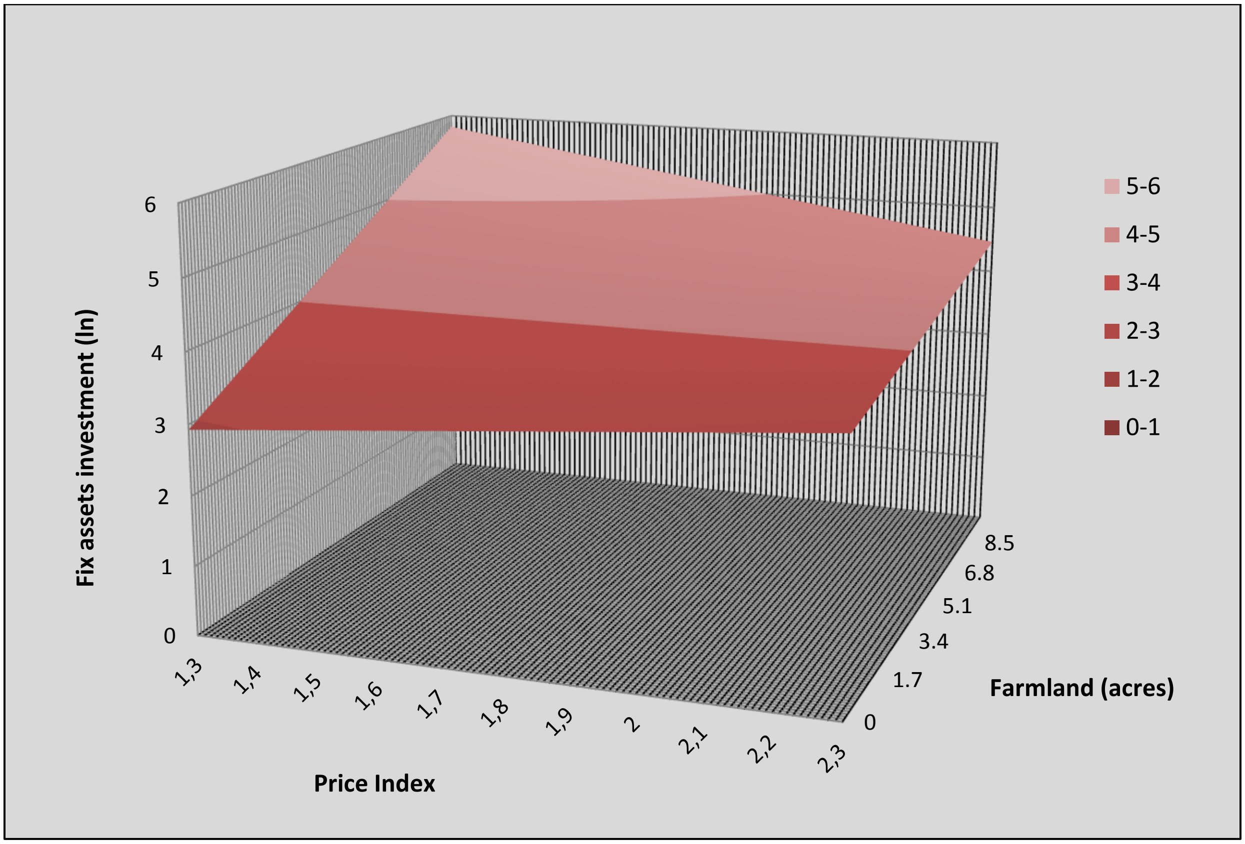
Note: The graph based on estimates of column 1, Panel A, Table 2.9. The scales of Income and Price Index fall into their distribution, [6.7; 12] and [1.3; 2.3], respectively.

Figure 2.6B: Effect on investment of price shocks conditional on loan obtained



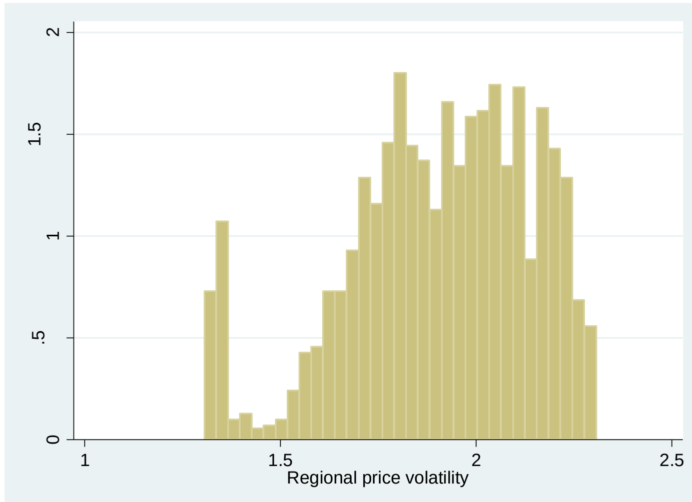
Note: The graph replies on estimate of column 1, Panel A, Table 2.10. The scales of Loan and Price Index fall into their distribution, [0; 10.7] and [1.3; 2.3], respectively.

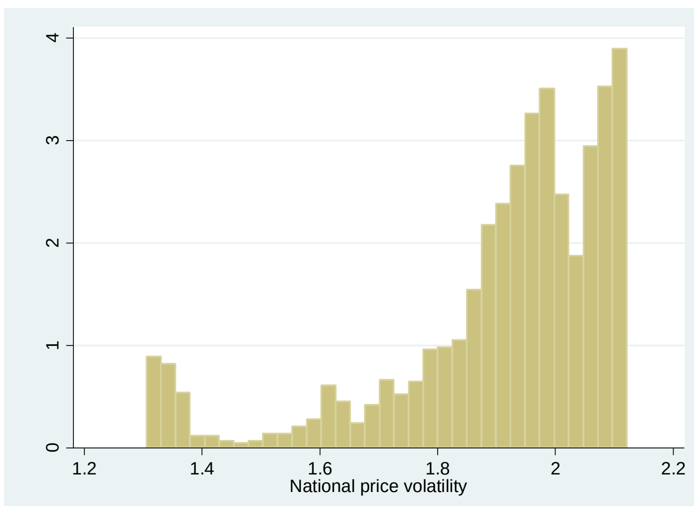
Figure 2.6C: Effect on investment of price shocks conditional on farmland



Note: The graph bases on estimates of column 1, Panel A, Table 2.11. The scales of Farmland and Price Index fall into their distribution, [0; 10] and [1.3; 2.3], respectively.

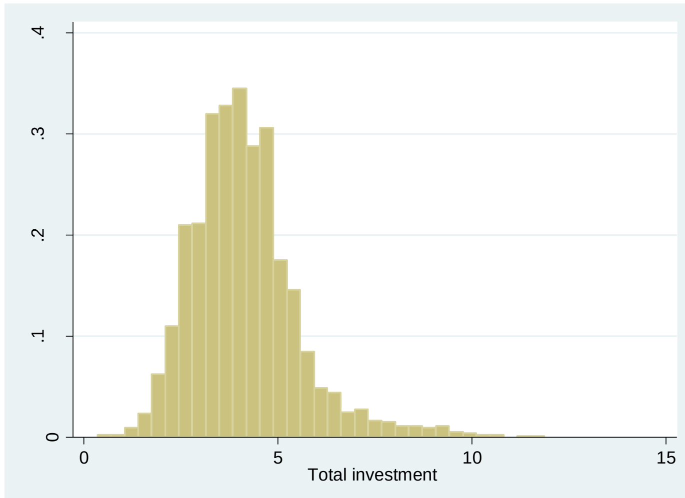
Figure 2.7: **Distribution of price volatility index, 2006-08**





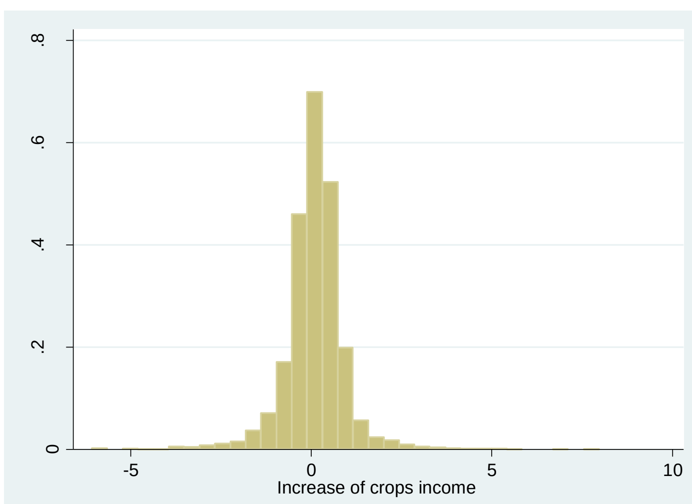
Note: The analysis uses the regional price index (vegetables price still are at national level) that looks more symmetric and more fluctuated than national price index. However, at the lowest level of regional index (on the left), the distribution appears unusual higher. Those two groups include households grew only vegetables that lead those household keep staying at the stable price index. N=2,303

Figure 2.8: Investment in agricultural assets, 2006-08



Note: Log of total expenditure. N=2,060

Figure 2.9: The difference in log income from crop between 2006 and 2008



Note: N=2,303

Appendix [14]

Now suppose the household is unable to borrow. The consequence is that investment must be financed by current saving. Equations (1) and (2) become

$$c_1 + q\left(i_1 + \delta k_1\right) = p_1 f\left(k_1\right) \tag{1'}$$

$$c_2 = p_2 f\left(k_2\right) = p_2 f\left(k_1 + i_1\right) \tag{2'}$$

Gross investment must be non-negative: $i_1 + \delta k_1 \geq 0$. The household's maximization problem is

$$\max_{i_1} \left\{u\left(c_1\right) + \beta u\left(c_2\right)\right\} \tag{3'}$$

There is now a single first order condition which is precisely that given by equation (4):

$$qu'\left(c_1\right) \geq \beta E\left[u'\left(c_2\right)p_2\right] f'\left(k_2\right) \quad \perp \quad i_1 + \delta k_1 \geq 0 \tag{4}$$

The left hand side of inequality (4) is the utility cost of foregone period 1 consumption. The right hand side is the discounted utility benefit of the additional period 2 consumption. If the right hand side exceeds the left hand side, the household would like to reduce investment but this is impossible if gross investment is zero.

To make the investment level explicit, we need to expand $f'\left(k_2\right)$ around k_1. Doing this

$$f'\left(k_2\right) = f'\left(k_1 + i_1\right) \approx f'\left(k_1\right) + f''\left(k_1\right) i_1 = \left(1 - \lambda i_1\right) f'\left(k_1\right)$$

where $\theta = -\dfrac{k_1 f''\left(k_1\right)}{f'\left(k_1\right)}$, the curvature of the production function. Substitution into equation (4) gives

[14] This part was helped by Prof. Christopher Gilbert.

$$i_1 = \max\left(\frac{1}{\theta}\left[1 - \frac{qu'(c_1)}{\beta f'(k_1)E\left[u'(c_2)p_2\right]}\right], -\delta k_1\right) \qquad (5')$$

To proceed further, consider a steady state in which at the normal price level to be p* in which $k_1 = k^*$, $i^* = 0$ (no net investment) and $q = \beta p^* f'(k^*)$, i.e. the marginal cost of investment is equal to its discounted marginal revenue product. Let $c_1^* = p^* f(k^*) - \delta k^*$ be the period 1 consumption level at this price. Then

$$\begin{aligned} u'(c_1) &= u'\left(c^* + (p_1 - p^*) f(k^*) - qi_1\right) \\ &\approx u'(c^*) + \left[(p_1 - p^*) f(k^*) - qi_1\right] u''(c^*) \\ &= \left[1 - \rho \frac{(p_1 - p^*) f(k^*) - qi_1}{c^*}\right] u'(c^*) \end{aligned}$$

where $\rho = -\frac{c^* u''(c^*)}{u'(c^*)}$, the household's coefficient of relative risk aversion.

Similarly,

$$\begin{aligned} u'(c_2) &= u'\left(c^* + (p_2 - p^*) f(k^* + i_1)\right) \\ &\approx u'(c^*) + (p_2 - p^*) f(k^* + i_1) u''(c^*) \\ &= \left[1 - \rho \frac{(p_2 - p^*) f(k^* + i_1)}{c^*}\right] u'(c^*) \end{aligned}$$

Writing, as previously, $p_2 = \lambda p_1 + (1 - \lambda) p^* + \varepsilon_2 p^*$, we can evaluate

$$\begin{aligned} u'(c_2) &\approx \left[1 - \rho \frac{\left[\lambda (p_1 - p^*) + \varepsilon_2 p^*\right] f(k^* + i_1)}{c^*}\right] u'(c^*) \\ &\approx \left[1 - \rho \frac{\left[\lambda (p_1 - p^*) + \varepsilon_2 p^*\right] (1 + \varphi i_1) f(k^*)}{c^*}\right] u'(c^*) \end{aligned}$$

where $\varphi = \frac{k^* f'(k^*)}{f(k^*)}$. If the production function is Cobb-Douglas, $f(k) = Ak^\alpha$, then $\varphi = \alpha$ and $\theta = 1 - \alpha$. This allows evaluation of $E[u'(c_2)p_2]$. Ignoring quadratic terms in $(p_1 - p^*)$ and ε_2

$$E[u'(c_2)p_2] = p^* \left[1 + \lambda(p_1 - p^*) - \rho\lambda(1 + \varphi i_1) \frac{(p_1 - p^*) f(k^*)}{c^*} \right] u'(c^*)$$

Returning to equation (5')

$$\begin{aligned} \frac{qu'(c_1)}{\beta f'(k_1) E[u'(c_2)p_2]} &\approx \frac{q}{\beta p^* f'(k^*)} \frac{\frac{c^*}{p^*} - \rho\lambda(1 + \varphi i_1) f(k^*) \frac{p_1 - p^*}{p^*}}{\frac{c^*}{p^*} + \lambda c^* \frac{p_1 - p^*}{p^*} - \rho\lambda(1 + \varphi i_1) f(k^*) \frac{p_1 - p^*}{p^*}} \\ &\approx \left(1 - \lambda \frac{p_1 - p^*}{p^*}\right) \frac{q}{\beta p^* f'(k^*)} \end{aligned} \qquad (6')$$

Using the fact that $\frac{q}{\beta p^* f'(k^*)} = 1$, it follows that the period 1 investment level satisfies

$$i_1 = \max\left(\frac{\lambda}{\theta} \frac{p_1 - p^*}{p^*}, -\delta k_1 \right)$$

For positive price shock, and provided the shock is at least to some extent persistent ($\lambda > 0$) the inequality reduces to

$$i_1 = \frac{\lambda}{\theta} \frac{p_1 - p^*}{p^*} \qquad (7')$$

so investment responds positively to the shock even if households cannot borrow.

Chapter 3

Agricultural technology adoption and natural disasters

3.1. Introduction

The recent food crisis still has given rise to broader concerns that need to be solved in the future. Given scare resources (land, water, nutrients and energy) but increasing population and demand of food, one of the key responses to which attention should be paid is agricultural productivity growth (M. Ivanic & Martin, 2010). Looking back over the last century, technology changes have played a fundamental role in generating rapid growth in agriculture (Cochrane, 1993; Schultz, 1964). In particular as the result of the Green Revolution, which occurred from the 1940s and spread rapidly throughout Asia over the period 1965-1990, production and productivity in agriculture have increased remarkably through the expansion of research and technological transfer initiatives. For developing Asia countries, the movement has led to unprecedented development in agriculture that has contributed to a substantial reduction in poverty and to rapid economic growth (Hazell, 2009). Across Asian countries, cereal yields grew significantly. For instance, over the period 1965-1982, the annual growth rate of yield was 4.1 percent for wheat, 3.5 percent for maize and 2.5 percent for rice; correspondingly the annual growth rates of production were about 5.4 percent, 4.6 percent and 3.3 percent (Rosegrant & Hazell, 2000).

Technological innovations can be classified into two groups: those that are embodied in capital goods or products and those which are disembodied (Sunding & Zilberman, 1999). Examples for embodied innovations are: mechanical innovations (adoption in tractors and combines), biological innovations (new seed varieties), and chemical innovations (new fertilizers and pesticides); while examples of disembodied innovations could be: agronomic innovations (new farm management practices) and informational innovations. Many disembodied innovations take the form of practical knowledge and can be shared by users. Since the difficulty in measuring the benefit from that disembodied innovations, investment and research activities are less expanded. The Green Revolution was mainly driven by a package of modern inputs including: improved seeds, irrigation, fertilizers and pesticides, which are embodied innovations.

A large body of literature has developed the theoretical framework in relation to technological adoption. This theory focuses on the behavior of individual farmers who decide whether or not and the degree of adoption. The parameters associated with adoption models include adoption costs, input prices, cost of alternatives, product prices, wealth, risk aversion, etc. In conjunction with these conceptual frameworks, a number of empirical studies have evolved to investigate the adoption decision and process of adoption. The results can support or contradict theoretical studies depending on specific sample. The main factors affecting adoption that have been widely tested are farm size, risk and uncertainty, human capital, labor availability, credit constraint, tenure, supply constraint, aggregate adoption over time (Gershon, Just, & Zilberman, 1985).

This paper focuses on risk and uncertainty factors, particularly the objective risk of natural disasters; rather than more subjective factors such as the unexpected performance of technology adoption itself. Following Schultz (1964), we suppose that the farmers are rational profit maximizers, so they make the decision of whether to adopt and the level of adoption depending on their anticipation of profit. In developing countries where almost farmers are vulnerable to shocks, the technology adoption might easily have a negative impact as the consequence of weather fluctuations. Rosenzweig and Binswanger (1993) find that a one standard deviation decrease of weather risk could increase profits by up to 35 percent in the lowest quintile of a sample in India. Esther Duflo, Kremer, and Robinson (2008) observe from an experiment in Kenya that farmers switch back and forth between using and not using fertilizer from season to season. These studies show that the returns of technology adoption are not always profitable once we take risky events into account. This study therefore aims to evaluate the returns of technology adoption taking into consideration of natural shocks.

There has been concern about confounding issue when assessing the effects of technology changes on outcomes. The possibility of bias appears since differences in outcomes between adopters and non-adopters could be affected by other factors as well as by technology adoption. In an experimental context, randomization reduces selection bias whereas in observational studies, the assignment of technology adoption is not randomized. This study overcomes that methodological problem in evaluating the effect of technological changes by implementing a non-parametric method, namely propensity score matching (PSM). This technique, which has been widely applied to estimate the impact of a program, policy, or treatment, was first proposed by Rosenbaum and Rubin (1983).

This study explores the empirical impact of chemical fertilizer adoption in paddy cultivation using data from a survey conducted in rural Vietnam. As a developing country in Asia, Vietnam is the second-largest rice exporter (after Thailand). Vietnam is also exposed to frequent natural disasters which strike farmers hard, especially poor farmers. In this context, the sample could be expected to exhibit importantly disparate impacts across different conditions of natural disasters. The paper uses the terms user, adopter, participant or treated unit, interchangeably.

The paper structure is organized in seven parts. The following section 2 introduces the current context of Vietnam. Section 3 explains the methodological framework and its application. Section 4 summarizes general information of the survey and descriptive statistics of the data. Section 5 presents results of PSM approach versus OLS and IV approach. The concluding remarks are discussed in section 6.

3.2. Background of the case study and the objectives

3.2.1. Rice cultivation and chemical fertilizer adoption in Vietnam

Vietnam is an agriculturally based economy. The agricultural sector involved 52 percent of total labor force of the economy and accounted for 28 percent of total GPD in 2009 (GSO). In the agriculture sector, rice is the most important crop. The area for growing rice is cultivated in 7.4 million hectares and accounted for 74 percent of total area for crops in 2009. Although the share of protein supply of average Vietnamese decreased from 63 percent in 1990 to 38 percent in 2010, it remains the second-highest source behind vegetables (FAOSTAT). The reason is that increasing income results in less consumption of rice.

Following the institutional transition of in 1986, the central planning economy became a market-oriented economy with liberalization affecting all sectors. The reforms created favorable conditions for rice industry such as a contract-based system, longer land lease period, a free market in land use and privatization of input and output markets. After nearly two decades, from being a major rice importer, the country became one of the major rice producer and exporter. In 2010, compared to the top rice exporter Thailand with 8.9 million tonnes, Vietnam remained in the second position achieving 6.8 million tonnes (FAOSTAT).

The development of the rice industry is displayed in Figure 3.1. Production slowly increased from 1961 to before the reforms, from around 9 million tonnes to 15 million tonnes in 1986. Since that time, output has risen sharply and reached nearly 40 million tonnes in 2010. Over the period from 1988 to 2010, the annual growth rate achieved 4.3 percent. This trend also appeared in yield and cultivated area for rice. Yield increased from 2.9 t/ha in 1998 to 5.3 t/ha in 2010. Those values for paddy cultivated area were 5.7 million hectares and 7.4 million hectares. The expansion of the area mainly was derived by the conversion from single to double or even triple cropping per year with the use of shorter rice varieties. The growth rates of production, yield and cultivated area have been risen sustainably but at lower rates during the 1990s.

As in many other countries, one of key driven factors of increasing production and yield is technological adoption. The principle innovations are modern varieties, greater fertilizer adoption, greater crop intensity increase and irrigation improvement. On the other hand, the farmers in Vietnam are characterized by small scale land holdings in relation to a large population. The population in 2011 was over 87 million, nearly three times more than 50 years previously. Hence, under the constraints of resource scarcity, enhancing productivity of rice is the crucial issue. Among the above-mentioned technical changes, fertilization and nutrient management have contributed most in increasing productivity.

In spite of a long tradition of using organic fertilizers, chemical fertilizers are more efficient and used more widely. According to the Vietnam household living standard survey (VHLSS) in 2010, which sampled through the entire country, an average of 60 percent (by weight) of chemical fertilizers used by households is devoted to paddy/rice out of all crops. Within the paddy crop, the households spend an average of 36 percent of total expenditure on chemical fertilizers, and only 8 percent for organic fertilizers. Vietnam is also a major exporter of chemical fertilizers. Nonetheless, the availability of fertilizers from local production is far insufficient to meet the demand of farmers and crops depend heavily on foreign suppliers.

Figure 3.2 shows the ratio between imported and consumed fertilizers by the main kinds of nutrient in Vietnam from 2002 to 2010. The important chemical fertilizers for crops include these kinds of nutrients: nitrogen, phosphate, and potash. The import data are based on the trading database (FAOSTAT). Consumption is on an apparent basic and is obtained as the difference between total imported and local produced values subtracting export values. It is clear that 100 percent of consumed potassium fertilizers are imported. For nitrogen fertilizers this decreases from 96 percent in 2002 to 58 percent in 2010. Although both quantities of imported and consumed nitrogen declined (from raw data), the extend of decline in import is greater than in consumption. The share of phosphate fertilizers fluctuates around 60 percent during this period. Nonetheless, the quantity imported of potash is lower than nitrogen and phosphate fertilizers; for instance, 298 thousands tonnes, 631 thousand tonnes and 310 thousands tonnes in 2010, respectively. Thus, the rice producers are significantly dependent on the international market of fertilizers.

3.2.2. Agro-ecological conditions and natural disasters

The country is divided into economic eight regions (Figure 3.4). These correspond to different agro-ecological conditions. The distribution of paddy cultivation area is distinct for each region. Also, the decision and the degree of adoption vary across regions since the different characteristics of environment (soil, weather) and other cultivating practices. Figure 3.3 displays the paddy cultivation features across six regions from data of Vietnam General Statistic Office (GSO). Using a slightly different classification of eight economic regions, the crops and conditions in North East and Northern West regions, and in North Central Coast and South Central Coast regions are somewhat similar, then they are merged into combined Northern Mountains and Central Coast regions, respectively. The characteristics in each region are illustrated by the maps in term of soils, slopes, temperature, rainfall, and flooding distribution throughout of country (Figure 3.6, 3.7, 3.8, 3.9). They are briefly summarized as the follows:

(1) Red River Delta: Because of its rich nutrient fluvial soils, the ancient Viet people settled here and cultivated wet rice over 3000 years ago. The urbanization and industrialization have turned it to be the most densely populated region at 949 person/km^2 comparing to 265 person/km^2 on average of the country in 2011 (GSO), and smallest landholding farmers. The region obtains the highest yield of the country at 5.9 tonnes/ha. The advantageous conditions arising out of an improved irrigation system has allowed farmers to intensify production to obtain crops to twice or three times per year. The region is the second largest “rice bowl” of the country and contributes 17 percent of total production.

(2) Northern Mountains: This mountainous and hilly area makes up 9 percent of total paddy land in the country. The region is inhabited predominantly by ethnic minorities. Shifting cultivation is based on slash-and-burn and long fallow. The degraded soil and dry are unfavorable conditions for rice growing. The yield is 4.6 tonnes/ha, the second lowest among others. Production accounts for 7.7 percent in the total output of the country.

(3) Central Coast: The region features denuded, moderately-eroded hills and mountains and sandy land. The cultivated areas mostly concentrated in the small delta along the coastline. Some locations are extremely dry in the dry season. The yield however is, lower than only two largest delta regions, around 5 tonnes/ha. The share of paddy land is 16 percent and of production is 15 percent, out of total.

(4) Central Highlands: The conditions in this region is unfavorable for rice. Instead, with the mild temperature and humidity, many industrial crops are intensively cultivated such as coffee, rubber, coffee, tea, etc. The ratio of paddy cultivation is only nearly 3 percent and contributes 2.6 percent for total production of the country.

(5) Southeast: Similar to Central Highlands region, the physical and weather conditions are more suitable for industrial crops, corn, and high-value vegetables than for rice (Young, Wailes, Cramer, & Nguyen, 2002). The economic return is also higher than from a rice crop. The paddy sown area takes up 3.9 percent and the production is 3.3 percent. The yield of this region is lowest in the country, only 4.5 tonnes/ha.

(6) Mekong River Delta: Located in a low elevation area, the highest flow of Mekong river combine with high rainfall in the wet season causing annual flood while the lowest flow months coincides with dry season that often leads to water shortage when the agricultural demand is highest (Nebitt, Johnston, & Solieng, 2004). About 40 percent in this region is acid sulfate soils and seasonal saline soils (Young et al., 2002). The environmental conditions are not good as in the Red River Delta and this results in a lower yield at around 5.5 tonnes/ha. Nonetheless, the proportion of cultivated area is about 53 percent and of production is 54 percent out of total which is greatly higher than other regions. Thus, this region is the largest-rice producing region in the country.

Located in the tropical monsoon area of South East Asia, Vietnam is one of the most hazard-prone areas in the region. Based on the long-term climatic risk index, Vietnam is placed in the sixth position of countries that was most affected by the weather events during the period 1991-2010 (Harmeling, 2011). The indicators of that index include the death toll, deaths per inhabitant, total value loss, total value loss per unit GPD, and number of events. The most frequent and most devastating weather events to Vietnam are typhoons/cyclones and floods. The storm season occurs from May to December and hits the north part of the country from May through June, then moves down to the South from July to December. Since most of the population reside along the coastline and in the low-lying delta region, disasters damage to their livelihoods heavily.

According to the Ministry of Agriculture's Central Committee for Flood and Storm Control, around 80 to 90 percent of the population is affected by the typhoons. The damage threatens both people living along the coastline and living in the upland areas who are exposed to subsequent flashfloods resulting from the typhoons' heavy rains (The World Bank, 2009). Moreover, Vietnam has 2,360 rivers which are steep and short. Heavy rainfall in their basins produces intense and short duration floods. According to the report, an average of six to eight typhoons or tropical storms strike Vietnam each year with more frequent occurrences in the northern and central coastal region earlier in the season.

3.2.3. Literature reviews and objectives of the study

In the recent literature, there are many studies about the impact of natural disasters and climate change on Vietnam. Almost of all studies adopt a macroeconomic perspective, with only a small number focusing on microeconomic context. One reason may be limited data available at the household level. Among the microeconomic studies, Thomas et al. (2010) measure the natural disasters and hazards at disaggregated geographical levels from primary meteorological weather station data, storm tracks and satellite observations. These weather indicators are then matched with the panel data of VHLSS in 2002, 2004, and 2006 in the corresponding geographical areas. The study find that riverine floods cause welfare losses up to 23 percent and storm reducing welfare by up to 52 percent in cities with a population over 500,000. Households are better able to cope with the effects of droughts, largely by irrigation. Another paper of Yu, Zhu, Breisinger, and Nguyen (2012) also constructs rainfall and temperature data recorded in 25 weather stations across the country which are used as the indicators of climate change and natural shocks. This data set is combined with VHLSS for 2004 and 2006 to examine the effects on agriculture and rural poor people, and their adaptation behavior in terms of adjustment to input usage. The results show that rice production is affected by climate change. The finding additionally suggests that expanding irrigation and agricultural intensification are the key components in maintaining productivity level and dealing with climate change.

In these circumstances, the farmers in Vietnam are faced with serious challenges. First, rate of productivity growth in the rice industry has slowed over the most recent decade. Second, the substantial losses and damage caused by disasters impact frequently and severely on farmers. For those reasons, technology adoption in agriculture is an important means to overcome such problems. The paper therefore intends to add additional empirical evidence about the impact of using technology given natural risk conditions. Utilizing across section data collected in 2010 in rural areas of Vietnam and the natural disaster losses reported by households, the study attempts to address fours main questions:

(1) Which are the characteristics that make households use chemical fertilizers?

(2) Whether and to what extent do the chemical fertilizer users obtain more beneficial outcomes than non-users in terms of rice productivity and household welfare?

(3) How do these results vary depending whether or not the households are exposed in natural disaster shocks?

(4) Are there differences in these results according to the econometric methodology employed (matching, OLS and IV)?

3.3. Methodology

3.3.1. The two-stage models

This section investigates determinants affecting on decision of using chemical fertilizer. The effects of relevant factors on technology adoption have been widely documented and studied. This study uses the limited dependent-variable models to test the household behavior on chemical fertilizer adoption. The most familiar type is Tobit model which has been applied in a wide range of consumer demand studies using micro-data. However, the Tobit model bases on the restrictive assumption, we therefore alternatively employ Heckman's two-stage sample selection to understand better the behavior of farm household on decision of adoption. The two-stage method includes the first stage deals with decision of adoption and the second stage deals with the extent of adoption.

Tobit model

The Tobit model is a traditional approach to deal with data with many zeros, which estimates the relationship between non-negative dependent variable and an independent variable (Tobin, 1958). The marginal effect of an independent variable is treated conditional on limited fraction of dependent variable which takes positive values as in this case.

The model is defined as:

$$h_i^* = x_{1i}\,\beta_1 + v_i \;; \qquad v_i \sim N(0, \sigma^2) \qquad (1)$$

$$h_i = \begin{cases} h_i^* \; if \; h_i^* > 0 \\ 0 \; otherwise \end{cases} \qquad (2)$$

where h_i^* is a latent unobserved endogenous variable which presents the optimal working hours; h_i is the corresponding observed variable which measures actual hours worked; x_{1i} and β_1 are vectors of independent variables and their parameters, respectively; v_i is a homoskedastic and normally distributed error term. The condition (2) implies that the observed the amount of fertilizer adoption is positive continuous if the positive fertilizer used is desired, and zero otherwise. Due to the non-negative values of fertilizer used, dependent variable h_i is censored at zero. This means that the observed zero on the dependent variable can be either "true" zero (i.e. individual deliberate choice) or censored zero (i.e. data collection methods, certain circumstances). Using maximum likelihood method, the likelihood function of standard Tobit is:

$$lnL = \sum_{0} \ln [1 - \emptyset(\frac{x_{1i}\ \beta_1}{\sigma})] + \sum_{+} \ln [\frac{1}{\sigma}\ \varphi(\frac{h_i - x_{1i}\ \beta_1}{\sigma})] \qquad (3)$$

Where "0" denotes the zero observations (fertilizer used h_i is zero) in the sample and "+" indicates the positive observations (fertilizer used h_i is positive); $\emptyset(.)$ and $\varphi(.)$ denotes standard normal cumulative distribution function and standard normal probability density function, respectively.

Heckit model

There is an argument pointing out an inadequacy of the Tobit model in solving that problem. In the Tobit model, the choice of being censored (adoption) and expected value conditional on un-censored (level of adoption) are determined by the same factors. The model considers dependent variable to be censored at zero but ignores the source of zeros, in which could be caused by deliberate household's decision or certain circumstances (e.g. financial conditions, characteristics of demographic (Newman et al, 2003; Martinez-Espineira, 2006).

Heckman (1979) proposes the two-stage estimation procedure to test the sample with zero observation (non-adopter). Heckman pointed out that estimation on selected subsample results in selection bias. The first stage is Probit estimation and the second stage is censoring estimation on selected subsample. In other words, the first stage estimates the probability of observed positive outcome or adoption decision. The second stage estimates the level of adoption conditional on observed positive values. The model assumes that these two stages are affected by different sets of independent variables, which is different from Tobit model. Another extended point in the Heckit model is that all zero observations are assumed to be derived from respondent's deliberate choices.

The first step estimates the adoption decision and the second step estimates for level of adoption. According to Heckman (1979) and Flood and Gråsjö (1998), the Tobit model is modified as:

The adoption decision:

$$d_i^* = x_{2i}\ \beta_2 + u_i\ ; \qquad u_i\ \sim\ N(0,1) \qquad (4)$$

$$d_i = \begin{cases} 1\ if \quad d_i^* > 0 \\ 0\ otherwise \end{cases} \qquad (5)$$

The level of adoption decision:

$$h_i^* = x_{1i}\ \beta_1 + v_i\ ; \qquad v_i\ \sim\ N(0,\sigma^2) \qquad (6)$$

$$h_i = \begin{cases} h_i^* \; if \; d_i = 1 \\ 0 \;\; if \; d_i = 0 \end{cases} \qquad (7)$$

In this model, x_{1i} and x_{2i} are vectors of explanatory variables in two stages of decision. Hence, the model assumes that the decisions of adoption and level of adoption are affected by separated sets of factors. As in Tobit model, β_1 and β_2 are corresponding vectors of parameters; d_i^* is a latent variable that denotes binary censoring; d_i is the observed value representing the adoption decision. The observed fertilizer used equals to unobserved latent value when a positive fertilizer used is reported; otherwise it takes the value zero. The error terms u_i and v_i are assumed to be independently distributed. This assumption implies that there is no relationship between the two stages of decision.

However, Heckman (1979) assumes that the two error terms are correlated and the first stage dominates the second one. Thus the error terms follow the bivariate normal distribution:

$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix}\right] \qquad (8)$$

Where ρ is correlation coefficient of the error terms. The domination assumption means that if the household reports positive fertilizer used, this is the intentional purpose of the household. In other words, the adoption is a deliberate choice. Then the model is estimated by Probit for the decision of adoption and standard OLS for the positive fertilizer used. The log-likelihood function for this approach is:

$$lnL = \sum_0 \ln\,[1 - \emptyset(\frac{x_{2i}\,\beta_2}{\sigma})] + \sum_+ \ln\,[\emptyset(\frac{x_{2i}\,\beta_2 + \frac{\rho}{\sigma}(h_i - x_{1i}\,\beta_1)}{\sqrt{1-\rho^2}} \;\; \frac{1}{\sigma} \;\; \varphi(\frac{h_i - x_{1i}\,\beta_1}{\sigma})] \;\; (9)$$

If the error terms are independent, ρ=0, the equation (9) is simplified as:

$$lnL = \sum_0 \ln\,[1 - \emptyset(\frac{x_{2i}\,\beta_2}{\sigma})] + \sum_+ \ln\,[\emptyset(x_{2i}\,\beta_2) \;\; \frac{1}{\sigma} \;\; \varphi(\frac{h_i - x_{1i}\,\beta_1}{\sigma})] \qquad (10)$$

3.3.2. Problem of impact evaluation

The definitions of impact evaluations have been described in varying ways. The World Bank (2008) states “impact evaluations compare the outcomes of a program against a counterfactual that shows what would have happened to beneficiaries without the program. Unlike other forms of evaluation, they permit the attribution of observed changes in outcomes to the program being evaluated by following experimental and quasi-experimental designs". The OECD (2002) states that impact evaluation analyze "positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended". These definitions have been widely used and somewhat modified, in a large number of researches, studies and reports.

The basic ideas of those definitions are almost the same, although they are described in different ways. In what follows, the main idea from those different descriptions is rephrased. In general, impact evaluation aims to assess the probabilities of outcomes which are directly affected by an intervention (or treatment). The treatments could be laws, regulations, educational programs, financial subsidies, medical drugs, technologies, etc. The participants could be individuals, households, firms, districts, countries, etc. Impact evaluation looks for directly at cause and effect questions.

In short, the issue can be understood in the following manner: the participants are exposed to a treatment or different levels of a treatment. The treatment effect can be measured by comparing the outcome of the same participant when exposed and not exposed to the treatment. Nonetheless, no agent can participate in more than one treatment at the same time. Thus, only outcomes of the participants exposed to the treatment are observed. Outcomes of treated agents in the absence of treatment cannot be observed. Since the counterfactual is unobservable, the impact must be evaluated through a reference to a comparison or control group. Therefore, the main challenge of counterfactual method is to find a similar group of nonparticipants that is comparable with the group of participants.

By seeking an appropriate comparison group, the counterfactual approach attempts to cure potential problems that arise from impact evaluations. These include confounding factor, selection bias, and heterogeneity (Khandker, Koolwal, & Samad, 2010; Ravallion, 2008; White, 2006).

(1) Confounding issue refers to omitted factors or unobserved factors that are correlated with both exposure to the treatment and the potential treatment effect. These factors could result in the spurious attribution of effects to the treatment. This is analogous to omitted variable bias.

(2) Selection bias is a special case of confounding, in which omitted factors influence the non-randomly selection of participants from the population. For example, those farm households that decide to adopt a new technology (mechanization) are more likely to have better complementary conditions (large cultivated land, irrigated land). Again these factors

simultaneously affect the probability of adoption and the productivity of crops. The endogeneity of treatment may cause simultaneity bias.

(3) Heterogeneous effects occur when the participants have diverse characteristics. For example, the households in different geographically areas may exhibit different socio-economic characteristics. Thus, each specific socio-economic type could affect on potential outcome.

There are several approaches controlling these problems. Each method imposes different assumptions and constructs the suitable framework to evaluate the impact of treatment (Khandker et al., 2010). The widely used approaches are briefly described as following:

(1) The randomized evaluation method allocates randomly treatments across subjects. The counterfactual information is captured through randomized trials or experiments which ensures similarity between the control and treatment group.

(2) The propensity score matching method generates statistical similar groups by estimating the probability of participating in the treatment conditional on the observed characteristics which are independent of the treatment. These estimated probabilities are used to match the treated and non-treated groups.

(3) The double-difference (difference in differences, DD) method compares the outcomes before and after treatment of the treated and non-treated groups. The DD method clears out the bias from time-invariant unobserved heterogeneity. This method is only available for panel data.

(4) The instrumental variable (IV) method identifies variables which correlate with the treatment decision but do not correlate with unobserved characteristics of outcomes. These variables can be used as instruments for the endogenous treatment variable. When used with panel data it can control for time-varying selection bias.

(5) Regression discontinuity methods exploit the near randomness of treatment selection around the cutoff for treated and non-treated groups. This method allows both observed and unobserved heterogeneity but throws away a large proportion of the sample information.

Randomized experiments are not normally in empirical studies in economics since they are costly and give rise to ethical issues. Faced with cross-sectional non-experimental data, this study adopts the propensity matching method to tackle problem of impact evaluation. The paper applies propensity score methods for the cases of both binary treatment and continuous treatment. The analysis further compares the result of this non-parametric approach with that from ordinary least squares and instrumental variable approaches. The following sections introduce the theoretical aspects of these methods.

3.3.3. Propensity score matching (PSM) with binary treatment

3.3.3.1. The basic framework

The counterfactual framework was initially proposed by Rubin (1974) and Rosenbaum and Rubin (1983). Let T be a binary treatment indicator, where $T = 1$ indicates the actual participation in a treatment, and $T = 0$ otherwise. In principle, each agent in the population might have either outcomes depending on whether or not he was treated. Suppose the outcomes of an agent received a treatment ($T = 1$) is Y_{i1}, and outcome without treatment ($T = 0$) is Y_{i0}. The treatment effect is $Y_{i1} - Y_{i0}$. Because each agent whether was or was not treated either Y_{i1} or Y_{i0} is observed. The impact evaluation of the treatment therefore faces missing data problem.

According to Rosenbaum and Rubin (1983), the average treatment effect on the population is the difference in the outcomes with and without treatment:

$$ATE = E[Y_{i1} - Y_{i0}] \qquad (11)$$

where ATE is the expected effect of treatment on a random agent drawn from population. Heckman (1997) argues that a treatment might target a certain group of population that relates to treatment purposes. In this sense, the interesting treatment effect is that of the agents who actually received the treatment. This is referred as the average treatment effect on the treated:

$$ATT = E[Y_{i1}|T = 1] - E[Y_{i0}|T = 1] \qquad (12)$$

Hence, ATT is the average difference in the outcomes with and without treatment of agents received the treatment. Which treatment effect is more relevant will depend on the specific circumstances and available data (Heckman, LaLonde, & Smith, 1999).

As mentioned, an agent either was or was not treated. In equation (12), the counterfactual mean of an agent who received the treatment, $Y_{i0}|T = 1$, cannot be observed. We therefore have to use substitute agents to estimate this counterfactual. In case of the randomized experiment, the treatment is independent with outcomes, the unobserved counterfactual, $Y_{i0}|T = 1$, could be replaced by the outcome of an untreated agent, $Y_{i0}|T = 0$. In this way, the ATT is rewritten as:

$$E[Y_{i1}|T = 1] - E[Y_{i0}|T = 0] = ATT + E[Y_{i0}|T = 1] - E[Y_{i0}|T = 0] \qquad (13)$$

$$E[Y_{i1}|T=1]-E[Y_{i0}|T=0]=ATT+B \qquad (14)$$

In non-randomized contexts, the factors affecting treatment assignment also affect outcomes, or the treatment decision may be affected by the expected outcomes, or both. In these cases, confounding and selection bias problems arise as pointed out in previous section. Therefore, the treatment is likely to be correlated with the potential outcomes, and the unobserved counterfactual is not equal to the expected outcome of untreated agent. The second term in equation (4), $B=E[Y_{i0}|T=1]-E[Y_{i0}|T=0]$, reflects the bias of estimates, so-called selection bias. To overcome this problem and estimate a true ATT, appropriate methods must be devised to clear out the selection bias, B = 0. The validity of these methods depends on whether the assumptions they make are valid. Rosenbaum and Rubin (1983) propose the propensity score matching method should ensures a reasonable substitute for unobserved counterfactual. The idea of matching approach is to find a control group of nonparticipants who have similar observed characteristics to those in the group of participants. Participants are matched with observationally nonparticipants conditional on their characteristics, and the average difference of the outcomes between the two groups is the average treatment effect. Datasets may include a very large number of characteristics so finding a close match in a general sense will be difficult. What the authors showed that the treatment probability is a sufficient statistic for matching in the context of measuring treatment effects. This reduces a multidimensional matching problem to a unidimensional problem. Corresponding to the problems, the authors introduce assumptions to identify a valid treatment. The two necessary conditions are unconfoundedness and common support.

Assumption of unconfoundedness

This assumption, first mentioned by Rosenbaum and Rubin (1983), is also known as ignorability of treatment, or selection on observables (Heckman & Robb, 1985), and conditional independence (Lechner, 1999). This states that the potential outcomes *Y* are independent on the treatment decision *T*, given observable characteristics *X* which are not affected by that treatment.

$$[Y_{i1};Y_{i0}] \sqcup T|X \qquad (15)$$

where ⊔ denotes independence. Given the controls X, the probability of treatment is independent of outcomes. Essentially, we have sufficiently many controls to account for treatment. The condition implies that the pre-treatment characteristics affecting the treatment decision and outcomes are simultaneously observed.

Unconfoundedness is strong assumption, and it is determined by the characteristics of the treatment or program. If there are unobserved factors that correlate with treatment assignment and potential outcome, the conditional dependence assumption is violated. Otherwise, if the assumption holds, the matching method is similar to that with a randomly assigned treatment in experimental process. This assumption is un-testable, so the theoretical and empirical literature provide guidance to select the pretreatment variables.

Assumption of common support or overlap

This condition states that each agent has a positive probability of being participant or non-participant given the certain covariates (Heckman et al., 1999), as the following expression:

$$0 < Pr(T = 1|X) < 1 \qquad (16)$$

This assumption ensures the overlap in the distribution of covariates. To achieve the efficiency, the sample should be large enough to generate sufficient set of predictors that matches in common region and leads to valid estimation of casual inference. If a nonparticipant lies out of the range in support region, it needs to be dropped out to guarantee the similarity with the treated group. If these two assumptions hold, the treatment assignment is referred to strong ignorability (Rosenbaum & Rubin, 1983).

When the treatment is identified under such assumptions, the valid ATT can be estimated conditional on covariates *X*. If *X* is a high dimensional vector, this may cause the large-scale problem for estimation. Therefore, Rosenbaum and Rubin (1983) suggest the balancing score method which is applicable for large sample. The propensity score is defined as the conditional probability of receiving a treatment given observed characteristics or observed covariates X :

$$P = p(X) = \Pr(T = 1|X) \qquad (17)$$

The authors prove that the matching method conditioning on *p(X)* is efficient as conditioning on X if the sample satisfies the key assumptions. This is because the biases from observable factors are swept out by conditioning on the propensity score. The balancing propensity score allows investigator to identify those agents with similar characteristics in the relevant sense. In other words, agents with similar propensity score should have similar distribution of covariates X, regardless treatment decision.[15] Then the expected difference of the observed outcomes is:

[15] This statement is presented and proven more specifically in Theorem 2 by Rosenbaum and Rubin (1983). Initially, suppose balancing score $b(X)$ is a function of covariates X. The most trivial balancing score is $b(X) = X$. The balancing score $b(X)$ that we need to identify is the "finest" propensity score $p(X)$. The Theorem states that: $b(X)$ is a balancing score, $X \sqcup T|b(X)$, if and only if $b(X)$ is finer than $p(X)$ in the sense that $p(X) = f\{b(X)\}$ for some function f (proof is shown in the paper). Theorem implies that "if subclasses or matched treated-control pairs are

$$E[Y_{i1}|T=1, p(X)] - E[Y_{i0}|T=0, p(X)] = E[Y_{i1} - Y_{i0}|p(X)] \qquad (18)$$

Since the strong ignorable treatment decision holds, the agents with the same propensity score *p(X)* but different treatment status can act as counterfactual for each other. The average treatment effect is estimated from the ATE conditional on *p(X)* and ATT conditional on *p(X)* through a two-step sampling process.[16] It is noted that ATE conditional on *p(X)* and ATT conditional on *p(X)* become identical since this process sweeps out selection bias (in equation (4)). By averaging ATT across distribution of *p(X)* in the entire population, we have the effect on the population:

$$E_{p(X)}\{E[Y_{i1}|p(X), T=1] - [Y_{i0}|p(X), T=0]\} = E_{p(X)}\{E[Y_{i1} - Y_{i0}|p(X)]\} = E[Y_{i1} - Y_{i0}] \; (19)$$

where $E_{p(X)}(.)$ is expectation over the distribution of *p(X)* in the entire population. The average treatment effects in the equation (19) and in the equation (11) are equal. The implementation of the propensity matching method will be presented in the next section.

3.3.3.2. Implementation

In this case of binary treatment, the average treatment effect is estimated in three steps.

Step 1. We first estimate the propensity score by modeling the participation in treatment. In this case of binary treatment, the propensity score is the probability of participantion versus nonparticipation given the observed characteristics. This can be carried out by logit or probit models which give the similar results:

---

homogenous in both $b(X)$ and certain components chosen of X, it is still reasonable to expect balance on the other components of X within these refined subclasses or matched pairs". An important advantage of this Theorem is that such procedure of classifying and matching is not only for $p(X)$ but for other function of X as well, e.g. estimates of average treatment effect in subpopulation defined by components of rural and urban areas.

[16] Following the previous footnote, this step is also proven more specifically in Theorem 3 by Rosenbaum and Rubin (1983). The Theorem states that it is treatment assignment is strongly ignorable given covariates X, then it is strongly ignorable given any balancing score $b(X)$. We therefore can estimate treatment effects given balancing scores and ignorable treatment assignment. Suppose a specific value of the vector of covariates X is randomly sampled from the entire population of agents. That is both treated and control agents together. Then a treated agent and a control agent both are found having this value for the vector of covariates. In this two-step sampling process, the expected difference in outcomes is $E_X\{E[Y_{i1}|X, T=1] - [Y_{i0}|X, T=0]\}$, where E_X is expectation with respect to the distribution of X in the entire population of agents. If treatment assignment is strongly ignorable, that expected difference in outcomes equals to $E_X\{E[Y_{i1} - Y_{i0}|X]\} = E[Y_{i1} - Y_{i0}] = ATE$. Now suppose a value of a balancing score $b(X)$ is sampled from the entire population of agents, and then a treated agent and a control agent are sampled from all agents having this value of $b(X)$. In this sense, the equation (18) follows the Theorem 3, and similarly gives the equation (19). For simplicity, I keep using only $p(X)$ which is equal to $b(X)$ if it satisfies the required assumptions.

$$P = p(X_i) = \Pr(T_i = 1|X_i) = \Phi(\beta' X) \qquad (20)$$

where $\Phi(.)$ is normal Cumulative Distribution Function of the standard normal distribution that contains all the covariates. The PSM will be bias if the pre-treatment characteristics which determine the participation decision are not included in the function. The reasons could be poor-quality data or a lack of understanding the feature of the treatment. In the other hand, over-specification (too many X variables) might result in higher standard errors for the propensity score. Thus, The investigator needs to balance parsimony against bias reduction.

Step 2. This step defines the region of common support where the distribution of propensity score for treatment and control groups overlap. Nonparticipants who fall outside the range of common support are dropped. If the dropped nonparticipations are systematically different from the retained ones in control group, selection bias might appear. A balancing test can be used to examine whether the average propensity score and the mean of X are the same for each quintile of propensity score distribution. Although the balancing test implies the comparable propensity distribution in treated and control groups, they might not similar if misspecification occurs.

Step 3. The last step is to match participants and nonparticipants by the estimated propensity score. The mean difference of outcome in these two matched groups is the impact of the treatment. There are several matching algorithm methods based on weighting for each matched participant and nonparticipant set, including:

(1) Near-neighbor matching. This method is one of most widely used which assign each treated unit to the non-treated unit with the closest propensity score. The number of selected neighbor unit could be more than one, and also bases on the sample size and propensity distribution. With the option of matching with replacement, one non-participant can be matched for different participants. In contrast, matching without replacement allows an untreated unit can be used only once.

(2) Caliper or radius matching. This procedure imposes a certain maximum range for propensity score. It eliminates the problem that some nearest neighbours may nevertheless be quite distant. The poorly-matched agent pairs are cleared out and quality of matching improves. Nonetheless, if too many agent pairs are dropped this might lead to increase variance of estimates.

(3) Stratification or interval matching. This method divides the common support of the propensity score into different strata or intervals. Then taking the mean difference of the outcome between treated and control observations within each stratum, we obtain the impact

of treatment. The weighted average of those strata impact yields the overall treatment impact by taking the ratio of participant in each stratum as the weights

(4) Kernel and local linear matching. This approach is a nonparametric matching procedure that use an appropriate weighted average of all agents in the control group to generate the counterfactual outcome. This matching has an advantage of lower variance because more information is used. In return, the problem of poor-matched might appear. Thus, the proper imposition of commons support region is necessary for better quality of matching.

In short, all approaches have itself advantage and disadvantage. The choice of estimator depends on data structure, and characteristics of covariates and treatment. All the methods however should give similar results asymptotically since the similar since they become closer to compare exact matches with growing sample size (Jeffrey, 2000). Hence, checking the results across matching approaches serves to indicate the robustness of the impact evaluation. In case the different approaches yield diverse results, the specification need to be further examined.

3.3.3.3. Balancing test for common support condition

The fundamental idea of matching method depends on the quality of balanced propensity score between treated and control groups which are identified by the common support region. Violation of the common support condition can lead to substantial selection bias for causal inference. If many control observations are different from treated observations, the subsequent comparison of outcome is inappropriate. Therefore, it is important to test for validity of the overlap region to ensure the comparability between treated and control groups.

There are several methods to check the validity of propensity score adjustment. This study applies the suggestion of Rosenbaum and Rubin (1985) that evaluates the distance between marginal distribution of covariates. This indicator is referred as “standardized difference in mean” or “standardized bias” (SB). The SB is measured by the ratio of the difference of means in treated and control groups divided by the square root of variances in these two groups averaged across treated and nontreated agents. The SB is computed for each covariate or observable pre-treatment characteristic before and after matching procedure, in the followingmanner :

$$SB_{before} = 100\frac{\overline{X_1} - \overline{X_0}}{\sqrt{\frac{V_1(X) + V_0(X)}{2}}} \qquad (21)$$

$$SB_{after} = 100 \frac{\overline{X_{1M}} - \overline{X_{0M}}}{\sqrt{\frac{V_{1M}(X) + V_{0M}(X)}{2}}} \qquad (22)$$

where V_1 and V_0 are variances of observable characteristic X in treated and control groups. The subscription *M* denotes for sample after matching selection in common support region. Although there is no specific SB value for validity of balancing property, the empirical studies considered the figures of under 5 percent as acceptable.

Another approach is to check for the balancing property by the t test for each covariate used to estimate propensity score. The test is applied for covariates before and after matching. If the t-statistics show that the two groups of participants and non-participants are significantly different before matching and indifferent after matching, then the matching method helps to improve the balance for the sample.

If the results reveal considerable lack of balance, the matching method needs to be altered or the pre-treatment characteristics should be re-checked. Smith and Todd (2005) state that the observable characteristics of matching method should influence both participation decision and outcomes, but should not affected by the treatment, although there is little guide available on how to select the conditional variables. There is a trade-off effect between the plausibility of unconfoundedness condition and the variance of the estimates. When using the full specification, bias will arise from adoption of a wide bandwidth for the common support region. On the other hand, when using a minimal specification, the plausibility of unconfoundedness condition could be violated. Hence, the suitable pre-treatment variables and balancing test should be iterated until achieving the convinced balance. This results in considerable latitude for the investigator to select the results which best suit his theoretical predilections.

3.3.4. Generalized propensity score matching with continuous treatment

3.3.4.1. The basic framework

To evaluate the impact when the treatment is a continuous variable, we use an approach that extends the basic framework of the binary treatment case. The propensity score in this case is called generalized propensity score (GPS). As the binary treatment case, adjusting observed covariates by GPS could avoid the selection bias associated with these covariates. This analysis follows the approach developed by Hirano and Imbens (2005).

Analogously, we start with a random sample drawn from the population. For each agent, there is a set of potential outcomes *Y(t),* where *t* belongs to an interval $[t_0, t_1]$ and refers to as the unit-level treatment. The interested effect is estimated through the average dose-response function, $\mu(t) = E[Y(t)]$. Given in certain observed covariates *X*, and treatment value received, $T \in [t_0, t_1]$, the potential outcomes of agent observed is shown as *Y = Y(T).*

Assumption of weak unconfoundedness

The GPS also requires the unconfoundedness assumption:

$$Y(t) \sqcup T|X, \qquad for\ all\ t \qquad (23)$$

This condition however is relaxed from the one of Rosenbaum and Rubin (1983) in the binary treatment case. The authors mentioned this as weak unconfoundedness, because it does not require joint conditional independence of all potential outcomes, but only each value of the treatment. Hirano and Imbens (2005) define the GPS as:

$$R = r(T, X) \qquad (24)$$

where $r(t, x) = f_{T|X}(t|x)$, that is the conditional density of the treatment given the observed covariates. The property of balancing score is similar to the case of standard propensity score above. Thus, with a similar GPS, the decision of receiving treatment level, *T = t,* is independent of the observed covariates.

GSP can remove selection bias in two steps. Under the weak independence assumption, the decision of treatment is random and uncorrelated with the potential outcome within each strata identified by GPS, *r(t, X)*. We have:

$$E[Y(t)|T = t, r(t, X) = r] = E[Y(t)|r(t, X) = r] \qquad (25)$$

From each treatment value of an agent, we also have:

$$E[Y(t)|T = t, R = r] = E[Y(t)|T = t, r(T, X) = r] = E[Y(t)|T = t, r(t, X) = r] \qquad (26)$$

Substituting the function (25) to (26), we obtain the expected outcome conditional on given GPS interval *r(t, X)*:

$$\beta(t, r) = E[Y(t)|r(t, X) = r] \qquad (27)$$

This is the first step in which conditional expectation of the outcome is estimated by a function of two scalar variables that are treatment level *t* and GPS interval *r(t, X).* In the second step, the average dose-response function is estimated by averaging the conditional expectation function over GPS across each level of treatment.

$$\mu(t,r) = E_{r(t,X)}[\beta(t,r)] = E_{r(t,X)}[Y(t)|r(t,X) = r] = [Y(t)] \qquad (28)$$

This emphasizes that the estimation of dose-response function (18) averages over the GPS at interested treatment level, $r = r(t,X)$, rather than averaging over GPS, $R = r(T,X)$. In this way, the regression function $\beta(t,r)$ does not have a casual interpretation. However, the dose-response function corresponding to $\mu(t,r)$ presents the average treatment effect at each level of treatment and can be used to compare the effect at different levels of treatment. Hence, the function $\mu(t,r)$ has a casual interpretation.

3.3.4.2. Implementation

In this case, the effect of treatment also can be undertaken in three steps.

Step 1. As in the binary treatment case, the first step is to estimate the score $r(t,x)$ from the conditional distribution of the treatment given observed characteristics. The distribution is assumed to have a normal distribution:

$$T_i|X_i \sim N(\beta_o + \beta_1' X_i, \sigma^2) \qquad (29)$$

The parameters of the distribution are estimated by maximum likelihood, and then used to calculate the GPS:

$$\widehat{R_\iota} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(T_i - \beta_o - \beta_1' X_i)^2\right] \qquad (30)$$

Step 2. In the second step computes the conditional expectation function of outcomeY_i, given T_i andR_i. The polynomial approximations of order three or less is used. This study applies the quadratic function:

$$E[Y_i|T_i, R_i] = \alpha_0 + \alpha_1 T_i + \alpha_2 T_i^2 + \alpha_3 R_i + \alpha_4 R_i^2 + \alpha_5 T_i R_i \qquad (31)$$

These parameters are estimated by ordinary least square using the calculated GPS R_i in the previous step. As emphasized, there is no direct interpretation for the estimated parameters in this equation.

Step 3. Finally, the last step consists of averaging estimated regression function over the score function evaluated at the desired level of the treatment.

$$\mathrm{E}[\widehat{\mathrm{Y(t)}}] = \frac{1}{N}\sum_{i=1}^{N} \widehat{\alpha_0} + \widehat{\alpha_1} t + \widehat{\alpha_2} t^2 + \widehat{\alpha_3}\hat{r}(t,X_i) + \widehat{\alpha_4}\hat{r}(t,X_i)^2 + \widehat{\alpha_5} t\hat{r}(t,X_i) \qquad (32)$$

In particular, we estimate the average potential outcome for each level of the treatment that we are interested in, in order to get the estimate of the entire dose-response function. This analysis uses the bootstrap method to obtain standard errors which take into account estimate of the GPS and the parametersα.

3.3.4.3. Balancing test

Similar to the case of the binary treatment, a balancing test can be used to assess the comparability of the covariates as adjusted by matching GPS. While the approach of binary case involved comparison of covariate means between treated and control observables before and after matching, the procedure for continuous treatment is more complicated.

Followed the approach proposed by Hirano and Imbens (2005), the test examines the balance on both estimated GPS and treatment level. First, the set of continuous treatment values is divided into *K* treatment intervals (*k=1, 2, …K*). Within each treatment interval, the GPS are computed at representative points (e.g., the mean, median, or another percentile) of treatment variable for each observation. Second, each interval *k* continuously is divided into *m* blocks by the quintiles of GPS. Note that those blocks are categorized conditional on the distribution of GPS within each *k* treatment interval, and denoted as $B_1^{(k)}, B_2^{(k)}, \dots, B_m^{(k)}$. Third, within each block $B_j^{(k)}, (j = 1, 2, \dots m)$, we calculate the mean difference of each covariate *X* between observations that belong to treatment interval *k* and observations that are in the same block GPS interval B_j, but belong to another treatment interval *k*. This step is carried out for all block intervals.

Finally, using weight average given by the number of observations in each block GPS interval, we combine all *m* differences in means of those blocks. Then the test statistics (the Student's *t* statistics) are the functions of this weight average to examine the balancing property. The value of test statistics (the highest value of Student's t statistics) is compared with reference to the critical value to check the level that balancing property is supported by the data. The less the extent that these differences in means are statistically different from zero, the better balancing level those covariates are. As the binary case, if the results suggest imbalance in the propensity scores in the sample, we should re-check the specifications, the division of treatment intervals and GPS intervals.

3.3.5. Instrumental variables method

3.3.5.1. Basic framework of IV method

To evaluate treatment effect of a single cross-sectional data, the instrumental variables (IV) estimator is frequently applied. IV estimation could be an alternative solution when unconfoundedness is suspected to have failed. The IV gives consistent estimates by viewing treatment assignment as endogenous. Based on discussions of Blundell and Dias (2002), Wooldridge (2002), Wooldridge (2005) and (Heckman, Urzua, & Vytlacil, 2006), this section compares the IV and matching approaches.

It is said that the IV methods can be very effective for estimating average treatment effect if a good instrument for treatment is available (Wooldridge, 2002). Otherwise, "the cure can be worse than the disease" (Bound, Jaeger, & Baker, 1995).

The problem is rewritten in parametric regression method as the following.

$$Y_0 = g_0(X) + \varepsilon_0 \qquad (33)$$

$$Y_1 = g_1(X) + \varepsilon_1 \qquad (34)$$

where subscriptions (0, 1) still denote for variables of agent without and with treatment, and error terms, ε_0 and ε_1, have zero mean. The function *g(X)* is a set of covariates, *g(X)* = $\delta + \beta X$. The average treatment effect conditional on covariates *X* is:

$$\alpha_X = ATE(X) = [g_1(X) - g_0(X)] + (\varepsilon_1 - \varepsilon_0) \qquad (35)$$

When the observed outcome Y is expressed in linear form:

$$Y = T[g_1(X) + \varepsilon_1] + (1 - T)[g_0(X) + \varepsilon_0] \qquad (36)$$

$$Y = g_0(X) + T[g_1(X) - g_0(X)] + \varepsilon_0 + T(\varepsilon_1 - \varepsilon_0) \qquad (37)$$

If the error terms of Y_0 and Y_1 are not affected by treatment assignment, i.e. meaning that the treatment effect is homogeneous across agents, the equation (37) becomes:

$$Y = g_0(X) + \alpha_X T + \varepsilon_0 \qquad (38)$$

If unconfoundedness holds, the treatment *T* is independent of error term ε conditional on covariates *X*. Then we can estimate the treatment effect, α_X, as the parameter of binary treatment variable *T* by the OLS regression (28). In particular, the ATE is, $\widehat{\alpha_X} = g_1(X) - g_0(X)$, constant.

Therefore, there are two conditions of homogeneity and unconfoundedness for unbiased and consistent OLS estimator. We now explain the IV method to fix on failure of these two assumptions. Firstly, in the case of failure of the unconfoundedness condition, the treatment is correlated with error term, and leads to inconsistent estimation by OLS. We need to find one excluded instrumental variable which helps to isolate the correlation between endogenous treatment variable *T* and error termε. For simplicity, I suppose a single instrumental variable. The instrumental variable, which is absent from outcome equation, must satisfy two assumptions:

(1) Exogeneity assumption: Given a set of covariates *X*, the instrumental variable *Z* must be uncorrelated with the error terms. This is called as exclusion restriction.

$$[(\varepsilon_0, \varepsilon_1) \sqcup Z|\, X] \qquad (39)$$

(2) Relevance assumption: The instrumental variable must be correlated with the endogenous treatment variable.

$$E(T|Z, X) \neq E(T|X) \qquad (40)$$

These two assumptions imply that the instrument variable *Z* affects the participation decision but does not have any relationship with factors affecting the outcomes. In other words, IV does not have impact on outcome through the unobservable factor, but it affects the outcome through the treatment variable *T*. Now the treatment effect under the standard instrumental variable estimator is defined as the ratio of the covariances (Durbin, 1954). The treatment effect uses only a fraction of the variation in *T* that is correlated with *Z:*

$$\widehat{\alpha_{IV}} = \frac{cov(Y, Z)}{cov(T, Z)} \qquad (41)$$

The IV absorbs the correlation between treatment *T* and error term ε. The parameter of treatment variable is estimated consistently through the IV procedure when the two assumptions of relevance and exogeneity hold. If the number of IVs is greater than the number of endogenous variables, we call the estimation method 2SLS

Secondly, bias and inconsistency may also arise if the impact of treatment is heterogeneous. The heterogeneity could be caused by observable or unobservable variables. The effect of treatment varies across agents for specific treatment status. This is shown as the error terms are different, $\varepsilon_1 \neq \varepsilon_0$. So in equation (37) the error term presents the relationship with treatment variable, which is: $\varepsilon_0 + T(\varepsilon_1 - \varepsilon_0)$. To avoid the heterogeneity from self-selection bias, Blundell and Dias (2002) impose another assumption which isolates the different characteristics from the participation decision. In particular, the agents do not use the information on the idiosyncratic factors of the treatment effect when deciding about participation. The idiosyncratic factors indicate their different characteristics that assist for greater gains from treatment if they participate. If the assumption holds, the agent participates the treatment without the prior information (but the researcher has all of that in covariates X) and the decision is based on the treatment effect of a specific group conditional on covariates X (without that prior information).

$$E[(\varepsilon_{i1} - \varepsilon_{i0})|X, Z, T] = E[T_i(\alpha_{iX} - \alpha_X)|X, Z] = 0 \qquad (42)$$

where $\alpha_X = E(\alpha_{iX})$, and we put an subscript i to indicate the error term of a certain agent in the population. Then we can identify the average treatment effect α_X, and ATE is identical with ATT, $E(\alpha_{iX}) = E(\alpha_{iX}|T = 1)$.

However, in fact, the treatment decisions are most likely to be non-random. Agents always participate in a treatment because of the expected gains from treatment which might rely on the advantage of their own characteristics. Agents will therefore be aware of the gains arising out of idiosyncratic factors. An example in Blundell and Dias (2002) is the distance to the treatment place used as an instrument. The variable is unlikely to be related with outcomes such as earnings or employment opportunities. This is considered as idiosyncratic factor of the treatment. However, agents are likely to participate if they reside closer to treatment place, since they incur lower travel cost, even if they expect lower gains from treatment. This selection process causes a correlation between the idiosyncratic gain from treatment α_{iX} and instrumental variable *Z*. Then the correlation is presented as:

$$\varepsilon_{i1} - \varepsilon_{i0} = (Y_{i1} - Y_{i0}) - \alpha_X = \alpha_{iX} - \alpha_X \qquad (43)$$

Then the error term under heterogeneity of treatment effect now is rewritten:

$$\varepsilon_i = \varepsilon_{i0} + T_i(\alpha_{iX} - \alpha_X) \qquad (44)$$

where T is identified by *Z*, and Z depends on the gain of $\alpha_{iX} - \alpha_X$. In this sense, we cannot estimate the average treatment effect with only two basic assumptions of IV (exogeneity and relevance assumptions). In other words, the standard instrument variables do not identify any meaningful treatment effect (Blundell & Dias, 2002; Heckman et al., 2006). The problem normally relates to the features of the regions that identify the set of covariates X. Then changes in Z affect the participation decision but are uncorrelated with the expected gains from treatment. Thus, to deal with heterogeneous population, IV might be not as effective as matching method. The PSM method, using different techniques, can eliminate observations which are highly unbalanced in terms of their characteristics. This may substantially diminish sample heterogeneity.

3.3.5.2. Implementation

The 2SLS method is processes in two stages.

(1) The first stage is to regress treatment T on the instrumental variables Z and covariates X and get the fitted values $\hat{T}$. This is part of treatment T that is uncorrelated with the error term ε. The residual may instead be correlated with error term ε, this is why the treatment T is potentially endogenous.

(2) The second stage is to regress Y on $\hat{T}$ and covariates X to obtain β_{IV}.

As the theoretical framework, if the composite residual has zero mean and is uncorrelated with fitted values $\hat{T}$ and exogenous variables X, the regression provides unbiased and consistent OLS estimates. But if the residual mean is different from zero, the effect of treatment is heterogeneous. Therefore, we have to check for validity of those three assumptions in IV approach: endogeneity, relevance, and homogeneity.

3.3.5.3. Tests for validity

Test for endogeneity

The 2SLS estimator will be more efficient than OLS if the suspected explanatory variable is endogenous. Otherwise, the estimates of 2SLS could produce large standard errors. It is necessary to have a test for endogeneity of the treatment variable. In the equation (28), we suspect a single endogenous variable is treatment T. So if the treatment T is proved to be uncorrelated with the error term, we should estimate the equation (28) by OLS. Based on the Hausman (1978) proposed a test that compares OLS and 2SLS estimates and checks whether the differences are statistically significant. The test is done in two steps:

(1) Similar to the above implementation, we first regress the treatment T on Z and other exogenous covariates X to obtain error term, $\hat{\vartheta}_i = T - \hat{T}$. The residual captures unobserved variables influencing treatment, but not captured by treatment and other exogenous covariates X.

(2) Then we regress outcome Y on covariates X and error term $\hat{\vartheta}$. Using the t-test, if the coefficient on the residual $\hat{\vartheta}$ is statistically different from zero, the unobserved variables jointly influencing on treatment and outcome Y are significant. The null hypothesis that error terms $\hat{\vartheta}$

and ε are uncorrelated is rejected. In other words, the treatment T is endogenous and 2SLS is more efficient than OLS.

Test for weak instruments

A major concern in relation to the IV method is to find good instrumental variables. An IV is considered weak when it fails to satisfy the two conditions, relevance and exogeneity or exclusion restriction.

The relevance condition requires that the IV be significantly correlated with the endogenous treatment variable. This helps to decrease the standard error of the IV estimate and the impact of endogenous treatment T could be predicted more accurately. In the first stage regression of 2SLS, if the coefficients on the excluded instrumental variables Z are statistically different from zero, the relevance condition hold. Since the model has a single endogenous variable, we simply use t statistic if there is one IV, and F statistic if there are more than one IV. The null hypothesis for the case of multiple IV is that all coefficients on excluded IVs are jointly equal to zero. If the null hypothesis is rejected, there is at least one excluded IV satisfied.

The exclusion restriction indicates that IV must be uncorrelated with the error term ε of the outcome equation. This cannot be tested because the error term ε is unobserved. However, if the model has more than one excluded IV, it could be examined whether at least one of them uncorrelated with the structural error. The test refers as overidentification test. The number of overidentifying restrictions is the number of extra excluded IVs. If each endogenous variable has one excluded IV, there is no overidentifying restriction that needs to be tested. If there are more than one IV for the single endogenous variable, we can test for overidentifying restrictions. The test is proceeded in two steps:

(1) Obtain the 2SLS residual $\hat{\varepsilon}$ by estimation of 2SLS structural equation.

(2) Regress residual $\hat{\varepsilon}$ on all exogenous variables Z and X to obtain R-squared, R^2. Use the null hypothesis that all IVs are uncorrelated with the residuals, $NR^2 \sim \chi_q^2$, where q is the number of excluded IVs minus the number of endogenous variables, or the number of overidentifying restrictions. If the calculated Chi-square statistic is greater the critical value at specific significant level, the null hypothesis is rejected. This indicates that at least one excluded IV is correlated with error term or not endogenous.

3.4. Data and model specification

The paper uses the cross-sectional data from Vietnam Access to Resources Household Survey (VARHS) conducted in June, July, and August of 2010. This survey with two previous rounds, in 2006 and 2008, were carried out in collaboration between University of Copenhagen in Denmark and the Institute for Labor Studies and Social Affairs in Vietnam. The surveys aim to investigate changes in household welfare in rural areas of 12 provinces, out of total 64 provinces. In particular, they are: Lao Cai, Dien Bien, Lai Chau, Phu Tho, Ha Tay, Nghe An, Quang Nam, Khanh Hoa, Dak Lak, Dak Nong, Lam Dong, and Long An. The surveys collected data in detail about agricultural production, employment, income and partial food expenditure, credit, risk and shock; along with basic information of household members and demographic. The 2010 VARSH was implemented in 3,208 households. Due to missing data and relevant information, the sample used in this analysis contains 2,280 observations or farm households who cultivated paddy.

3.4.1. Natural conditions and shocks

Figure 3.4 shows the location of the 12 provinces which provide data. These are spread throughout the country. They belong to seven economic regions, out of total eight regions in Vietnam. Hence, the agricultural land of households is characterized by diverse types of land as well as weather conditions, which substantially affect crop yield and therefore fertilizer adoption. The provinces in North East and North West regions, namely Lao Cai, Dien Bien and Lai Chau, are mountainous with sloping land and poor soils. The other three provinces Dak Lak, Dak Nong and Lam Dong are also located in hilly area. Instead Ha Tay and Phu Tho in Red River Delta region, and Long An in Mekong River Delta region are located in the two plains which form the two largest rice producing regions of the country.

The survey collected information on land where problems were experienced in planting. Table 3.2 shows the percentage of land area by six problematic types that include: erosion, dry land, low-lying land, sedimentation, landslide, stony soils or clay, and uncategorized type of problematic land. The first column is the ratio with respect to total land area. Dry land exhibits the highest ratio , at around 29 percent. The other types of problem accounts for much lower proportions, from 3 to more than 5 percent. Land without problems is approximately 42 percent of the total. However, farmers typically grow more than one crop. Especially if farmers manage forest or pasture land which are normally much larger than the area used for crops, they might underestimate the ratios of problematic land. Although the information on land with problem is available only for total land, not for paddy land, we can compare these two areas. The second column is the ratio between total unfertile land and paddy land that households hold. Dry land is nearly equal to land without problem, about 92 percent and 97 percent; followed by eroded land, low-lying land, and land subject to landslides, with ratios of 21 percent, 17 percent and 14 percent respectively.

Regarding the uncertain risk from weather fluctuations, the maps in Figure 3.8 show that the average annual temperature increases but fluctuates less as we move from the North to the South. Correspondingly, the three provinces in Northern mountain region have disadvantageous temperature conditions. Instead, the coastal provinces in the middle and the South Center of Vietnam, which contain Quang Nam and Khanh Hoa, face higher rainfall level and also a greater possibility of suffering coastal flooding. The other two River Delta regions are in the plains area with its intensive river system that is the main source of water and which periodically enriches the land. On the other hand, these conditions turn to be harmful when riverine and coastal floods arise. Such phenomena are most frequent in Ha Tay in the North and Long An in the South.

This analysis uses the information on natural disaster reported by households that their plots (land) suffered in the period starting from July 2008 with the exact month and year for each incident. In order to make natural disasters consistent with the outcomes, the sample uses only shocks that households suffered in the previous twelve months. The natural disasters are classified into seven types. Based on the similar consequences, those types are further aggregated into three groups. The objective of this is to generate a large enough sample size to investigate the impact in each group of disaster. The first group includes floods, storms or typhoons, and landslides; the second group is drought and the last group includes pests or plant diseases, insects and animals problems. The household could be hit by more than one disaster, so Table 3.2 reports the numbers that households suffered one disaster exclusively and with other types simultaneously. The raw data indicates that the farmers are exposed in highly possibility of natural shocks. The sample reveals that more than 30 percent of all households hit by at least one natural disaster event in the previous twelve months. The proportion of households that experienced floods, storms and landslides is around 12 percent, while the proportion experiencing drought is about 8 percent, and that for plant/insects diseases is nearly 14 percent. Considering those disaster events that happened exclusively to farmers, the numbers decreases around 1 or 2 percent for each group of disasters.

3.4.2. Outcomes by treatment status

Table 3.4 relates to chemical fertilizer adoption. Fertilizer adoption is measured by the total expenditure spent on fertilizers purchased for rice production. The purchases cover urea, NPK, phosphate, etc; but exclude organic fertilizers. The units are thousand VND per square meter.

The data show that the majority of respondents adopt chemical fertilizers for their paddy cultivation. Out of total 2,280 observations, only 267 farmers cultivate rice without using chemical fertilizers accounting for around 12 percent of the total. The distribution of users is slightly right skewed with mean 0.37 and median 0.33 units.

Base on the distribution of fertilizer users, we track the data on outcomes by three categories of fertilizer-using status: non-users, below average users and above average users. The data is observed through four categories of natural shock status: without disasters, flood/ storm/ landslide disasters, drought disasters, and plant disease/ insect disasters. Since different natural disasters may cause different levels of harm, the analysis investigates the impact of each shock type separately.

The outcomes variables measuring crop productivity are rice yield and rice revenue. Rice revenue is the total value of rice produced per unit land area, which includes both value for sale and the value retained for household use. The outcomes capturing household welfare are monthly per capita income from agricultural activities and monthly per capita income. It is acknowledged that these indicators may fail to adequately capture the effect of treatment. First, although rice production is the main crop of farmers in Vietnam, agricultural income also arises from crops and livestock. The income from rice production solely is a closer proxy for evaluating impact. However, the survey aggregates some items of expenditure of rice production in common with other crops which makes it impossible to compute the income (revenue minus expenditure) from rice separately. So we have to use revenue indicator instead of income indicator. Second, the same problem for total income variable for those households for which rice growing accounts for only a small fraction of total household income. The survey also reports information on food expenditure which might be a better indicator for household welfare changes. However, this variable relates to expenditure information for only the previous four weeks while household food expenditure may fluctuate substantially over the year.

Table 3.5 presents summary statistics and statistical tests on the difference in means for the continuous variables between groups of non-users and users, and between groups of below average and above average users. About 50 percent of households belong to the below average group and 37 percent of households are in above average group. In the whole sample, the outcomes of those four indicators show significant positive differences between users and non-users, and among users, at less than 1 percent level.

Moreover, we observe the diversity in the outcomes across categories of disasters. For example, the households suffered flood/ storm/ landslide shocks exhibit a negative difference in rice revenue between users and non-users, and in agricultural income within user group. In group of households which did not experience disaster shocks, the outcomes show little difference across fertilizer usage status, except for total income among users. The outcome variables for non-users appear lower than for users resulting in higher values of difference between users and non-users than between groups of users. Farmers experienced drought show the positive incomes differences but insignificant. Therefore, the descriptive data suggests differing impacts of chemical fertilizer adoption on outcomes conditional on the level of adoption and type of disaster shock.

3.4.3. Observable characteristics by treatment status

Table 3.6 reports descriptive statistics on the variables that will be used to estimate the propensity scores. All variables are tested to examine whether the mean of Observable characteristics differ significantly across groups. Overall, the differences are more significant between users and non-users than among the user groups, both in terms of observable characteristics and the magnitude of the differences. The features of household structure are significantly different in both comparison groups, except for the number of members at the ages from 16 to 25, the number of members at the ages from 26 to 60 between users and non-users, and gender headship among users. The non-users reports fewer members working for wages or salary, fewer workers hired outside for paddy cultivation, but more workers in household agricultural or business activities. The workers in household business and hired workers within the group of users are broadly equal. Moreover, the non-users are more likely to report being constrained by credit access and much lower assistant receipted [17]. The demographic (ethnic minority) characteristics and distance to the center of commune are also associated with the status of fertilizer use.

Related to land characteristics, the *t*-test results show insignificant differences across groups for both paddy land and the proportion of land without problem out of total land area (which relatively measures the quality of land). The plots of fertilizer users are more dependent on public irrigation system than non-users. The likely reason is that almost non-users reside in Northern West mountain area (Table 3.18) where the public irrigation is underdeveloped. The other farmers in coastal and plain regions are favored by extensive public irrigation for their crops. According to the data of this sample, the proportions of irrigated area in total land for paddy cultivation take up 27 percent, 75 percent and 87 percent in non-users, below average

[17] Money or goods from persons who are not members of your household such as relatives living elsewhere friends or neighbours (private transfers); or from public institutions, e.g. insurance money or social assistance (public transfers.

users and above average users, respectively. This variable exhibits a severe in propensity score matching, so it will not be used in the model and reported in the table.

The farmers also were asked to report the main type of seed cultivated. There are three categories of seeds used: hybrid seed, ordinary improved varieties, and old local varieties (with about 1 percent of uncategorized seed).[18] Each type of seeds has different advantages and disadvantages that are appropriate for different agro-ecological, socio-economic areas. For example, based on the empirical result from a survey covering nine provinces in Vietnam (Hossain, Tran, & Janaiah, 2003), hybrid rice is adopted less in Central Highland and in the south than in the north. Compared to inbred varieties (either ordinary improved or old local varieties), the study finds that the hybrid seeds generate higher return but also involves in higher costs. Recently, about 86 percent of the farmers expressed their willingness to continue hybrid rice cultivation due to of higher profitability. Only 14 percent of that sample indicated that they would discontinue hybrid cultivation in the next seasons because of lower yield, lower price and lower profitability. Therefore, the type of seed used may differ across provinces. In this sample, the non-users are most likely to adopt ordinary varieties, and use less hybrid seeds. In the group of users, higher chemical fertilizer use is associated with a higher ratio of hybrid seed adoption and lower improved seed adoption. Use of the old local seed appear does not differ across fertilizer users and non-users. The descriptive statistics reveals that the adoption of hybrid seed is expanding in Vietnam, as noted in Hossain et al. (2003) .

3.4.4. Instrumental variables

According to the two key requirements for an IV, we need to find an appropriate set of variables (or a single variable) that are correlated with decision on treatment participation and correlated with outcome only through the treatment participation. The geographical indicator could be considered for the choice of IV as it is much likely to affect fertilizer adoption. However, different regions also characterized by different agro-ecological conditions that contribute to a higher productivity by types of soil, temperature, rather than only fertilizer; and socio-economic conditions that contribute to a higher income. Such correlations will cause correlation between the candidate IV and error term. For instance, in this case study, the majority of non-adopter live in mountainous regions. Since the agro-ecological conditions don

[18] According to the manual of VARSH 2010, those types of seed are described as: (i) Hybrid seeds are produced by artificially cross-pollinated plants. It is necessary to buy new hybrid seeds every season. (ii) Ordinary improved varieties cover modern, improved seeds for which it is possible to store harvested rice and use it as seeds in the next season. Hence, with this type of seeds it is not necessary to buy new seeds every season. Hence, with this type of seeds it is not necessary to buy new seeds every season. (iii) Old local varieties are the seeds that have not been improved by modern methods. These are varieties that have typically been grown in the area for many years.

not favor paddy cultivation, production and productivity cannot be as high as other regions, even if comparable amounts of chemical fertilizer are adopted.

We propose the two variables as instruments: the availability of information about fertilizer received by household and the availability of crop insurance. Information on the first one of these is acquired from a question that asks whether the household received any information on and assistance in fertilizer utilization provided by extension workers who visited the household in the previous 12 months. These visits might have taken place at the request of the household. Although requested visits form 15 percent of total visits. The household is assumed to be more likely to adopt fertilizer when it has more information. So if we find a correlation between information assistance and outcomes, this can only arise through adoption. Table 3.6 shows that the share of household that received information in group of non-users was 69 percent, while that ratio in group of users was 91 percent. The gap between non-users and users is 32 percent and differs from zero at 10 percent level of significance. The gap between two groups of users is 3 percent.

The second IV derives from the question which asks how much households would receive if the insurance against lost and damage crops were available. This availability of insurance appears to have a positive impact on fertilizer adoption, especially for households who frequently suffered from disasters. The binary treatment model applies the dummy IV of crop insurance, and the continuous treatment applies the continuous one. Crop insurance may affect other kinds of technology adoption, such as irrigation, varieties, soil management, etc., and this may complicate interpretation of the results. Also, crop insurance contributes directly to household welfare if there is loss or damage in its production. We will check this through test for IV validity. The descriptive data in Table 3.6 show that the share of household with crop insurance in the group of non-users, below average users, and above average are 10 percent, 39 percent, and 47 percent, respectively. The difference of non-users versus users, and below average users, and above average are statistically significant at 1 percent level. Similarly, the differences in term of value of crop insurance also are significant at 1percent. The following section exhibits result and analysis in detail.

3.5. Application and results

First, we investigate the determinants which affect the decisions of farm-household on chemical fertilizer adoption. The first stage of Heckman model evaluates whether or not farmers used fertilizers, and the second stage evaluates intensity of adoption. We test for both cases, with and without consideration of natural disasters. The explanatory variables are the expected factors that have been tested in empirical studies and shown in Table 3.6.

Second, we continue to explore the effect of adoption on outcomes by matching methods. The variables used in the model are pre-treatment characteristics as in Heckman tests. On the basic of the significance tests discussed in section 5, some variables appear to differ substantially between the groups of users and non-users. This is likely to generate unbalanced propensity scores. To ensure that the propensity scores relate to adoption decisions and outcomes, we experiment with different specifications with the objective of achieving satisfactory balancing test outcomes. In line with other empirical studies, the model can be augmented with interaction terms and higher order terms. Therefore, in order to achieve a better balance in the propensity score, highly significant variables are used interaction terms and higher order terms, including: education of head, ethnicity, paddy land area, ratio of unproblematic land, and three kinds of seed varieties. In particular, the education of head is interacted with ethnicity; the paddy land area is interacted with the share of unproblematic land, and with seed varieties. The square of the education of the household head is also included into in the model. In addition, the values of loan and assistance receipt are transformed into log term to reduce the skewness. On the basic of this specification, the application of matching method and results are discussed below in the two cases of binary and continuous treatment.

Third, for comparability with matching method, we also report results of using OLS and IV methods. The regressions take both binary treatment and continuous treatment as the intervention variable corresponding to all categories of outcomes and types of disaster as in the matching method. All the pre-treatment variables, which are used for measuring the propensity scores, become the explanatory variables in the regressions.

3.5.1. Heckman tests

We test for four models in which the determinants in the second stage of decision are selected differently. The first model test for the same explanatory variables in both stages of decision making, and the other models chooses only the most relevant and significant explanatory variables in the second stage. Table 3.7 shows results of the whole sample and Table 3.8 shows results with taking consideration of natural disasters as dummy explanatory variables.

The impacts of most of the statistically significant variables appear with expected sign. Education of household head, numbers of household labours in agriculture, outside labours and assistance receipt have positive impacts on adoption decision, whereas numbers of member from 10 to 25 years old, minority ethnic households, male headship, numbers of household labours in business, distance to committee, and land independent on public irrigation have negative effects on adoption decision. The effects between two stages of decision-making are fairly different. In model 2, we keep the same explanatory variables in the second stage, household members from 10 to 25 years old, education, ethnicity, male headship, outside labours, assistant receipt become statistically insignificant; while labours for salary, loan obtained and land independent on irrigation have positive effects on level of adoption. In the model 2, 3 and 4, we keep the significant variables in the second period. Then education appear to be positively significant. Moreover, households who suffered disasters are more likely to adopt fertilizers. Effects of different type of natural shocks vary in both stages. Flood/ storm disasters have positive effects in both stages, while disease/ insect shocks positively affect in the first stage, and drought positively affect in the second stage.

The Maximum Likelihood Ratio Test refers to the similarity of the decision making mechanisms of adoption and level of adoption, in term of the existence of significant correlation between the error term of the two stages. According to the results of LR test, the error terms of the two stages are found to be zero correlation, except model 4, based on the 10 percent significance level of the chi-squared distribution. These findings fail to reject the hypothesis that the decision of whether or not to adopt fertilizer does not affect the decision of adoption intensity.

The model 1 shows the lowest correlation between the two stages as it includes all the same covariances in the second stages. The directions and level of significance of determinants are fairly different between adoption and level of adoption equations. When we drop the insignificant variables in the level of adoption equation, the correlation of increases. This refers that there is selection bias and Heckit model is appropriate choice. Moreover, the natural disasters show statistically significant effects in both stages. The flood and storm have significant and positive effect in both stages, but disease and insects have positive effect in the first stage and drought has positive effect in the second stage.

As theoretical framework, the Heckit model is identified if the same independent variables in the first stage appear in the second stage. The identification occurs from distributional assumptions about residuals and not because of variation in the independent variables. This means that non-linear relationship in the first stage causes identification. So identification is one of the problem to select the independent variables in the second equation, and we may have inaccurate estimates in level of adoption. We need to find the independent variables that effect adoption decision but not level of adoption. In general, we can either exclude variables in the first equation, or add variables in the second equation, although this might not be correctly theoretically motivated. Therefore, identification issue is one of limitations for this approach.

3.5.2. Matching methods

Binary treatment

In this case, we estimate the mean difference of outcomes between group of farmers who use chemical fertilizers and group of farmers who do not use chemical fertilizers. To reveal the different effects of natural disaster shocks, the matching method is employed in turn for the whole sample, for the group of farmers who do not experience disasters, and for the group of farmers suffered from natural disasters. In the sample, the number of non-adopters is 276 households, and the number of non-adopters experienced a specific disaster is less than 10 percent, which is too small to be checked by matching method. Thus, we do not distinguish between households which experienced different types of disasters but instead consider just households suffered from at least one sort of disaster. We use the Stata package developed by Leuven and Sianesi (2003).

The logistic model is employed to predict the probability of adopting chemical fertilizers with a set of households and crop characteristics. This step could equally be employed using probit model giving the similar results. Table 3.9 shows the results of estimated propensity scores from logit regression. The first column is estimate of the whole sample, the second is for subsample of households who did not suffer disasters, and the third is for households suffered at least one kind of disaster during the previous twelve months. The role of propensity scores is just as a means to balance the observed distribution of covariates between users and non-users. Although the estimated parameters of the logit model are not the objective of this step, they show a numbers of characteristics that significantly affect on the decision of adoption and are similar as in the Heckman tests. These impacts are statistically insignificant for the subsample of households who suffered disasters. The results are quite similar for the whole sample and subsample of households without disasters with only a slightly differences in magnitude. These estimated parameters are less statistically significant for the subsample of households with disasters.

Using the estimated propensity score in the above model of the whole sample and the subsample, we generate the new samples of matched observations between users and non-users by four alternative matching algorithm techniques. The common support condition is imposed in all matching techniques. To implement this condition, the users whose propensity scores are higher than the maximum or less than the minimum propensity score of non-user are dropped from the matched sample. Then the means of the two groups, users and non-users, remaining in the common support region are compared to obtain the difference. This difference measures the effect of treatment, here chemical fertilizer adoption.

Table 3.10 presents the average effect of adoption on rice productivity (yield, revenue) and household welfare (agricultural income and total income) for different categories of households in Panel A, B, C. The first column is the result of sample before matching. The remaining four columns shows the result from the four matching techniques. The treatment assignment displays the numbers of observations which are left in common support region and used to evaluate the difference of means. In the nearest neighbor (NN) method with replacement in the second column, the user is matched with closest non-user in term of propensity score. Since the number of non-users is lower than of users, each non-user is replicated several times to match with users. The second method, we continue to use nearest neighbor method combining with non-replacement condition (NNN). This method performs that only one non-user is matched to one user in each pair. The third technique, we employ the caliper matching that imposes the certain distance for any match (CN). In this method, the observations which lie outside of caliper are dropped. These estimates apply caliper 0.01, which means that all matches not equal to within 0.01 standard deviations of each covariate are dropped. The last method is kernel-based matching (KBM) by which a pair of matched observations is identified by kernel-weighted average condition. The KBM method requires selection of the kernel function and bandwidth (or smoothing parameter). The selection of bandwidth is more important (Pagan & Ullah, 1999). There is a tradeoff between variance and bias which is caused by selection of bandwidth. Higher bandwidth leads to a smoother density function and hence lower variance. On contrary, higher bandwidth causes greater bias estimate. The model employs Epanechnikov kernel and 0.06 bandwidth, as the default in this Stata package.

The detailed propensity scores before and after matching are reported in Table 3.11, 3.12, and 3.13 for those three categories of households. In these tables, we also show the t-test, between users and non-users, before matching and after matching. The bias reduction refers to how much the difference in propensity score between the two groups decreases after matching. The results demonstrate that almost of the observable characteristics achieve more balanced propensity scores after matching. This is indicated by a higher p-value of the t-test. Moreover, we also report the propensity score distribution in kernel density estimation. The figure below each panel illustrates the kernel distribution of sample before matching and after matching. We can observe that the distribution of propensity scores is more balanced after matching, and the difference is clearer when the standardized bias is smaller. However, this is not always the case, for instance in the case households suffered disasters in KBM technique (Panel D, Table 3.15), and the mean of standardized bias is lowest, compared to other matching techniques. But the kernel distribution of propensity score in KMB appears less balanced than that in NNN and CN. Therefore, tests of the balancing property should be checked together with t-test and bias reduction in each pre-treatment variable, together visual examination of the kernel distribution of propensity scores.

In Panel A, Table 3.10, the results show that the rice yield obtained by the group of fertilizer adopters is significant higher than that obtained by the group of non-adopters. Specifically, the NN method estimates the effect of adoption is 0.056 (kg/m^2) higher than without adoption. Similarly, the effect in NNN, CN, and KBM techniques are 0.103 (kg/m^2), 0.093 (kg/m^2), and 0.067 (kg/m^2). This is the average difference of rice yield between the similar pairs of household that adopted chemical fertilizers and those that did not adopt. Compared to the result of the sample before matching, the impacts appear lower after matching. The overestimate of before matching result may be contributed by other factors that also are favorable of higher yield, such as soil quality, irrigation, credit, etc. The balancing tests show the evidence of reduced bias level after matching. In detail, bias before matching is 27.95 percent, while the lowest bias after matching is 5.25 percent in the KN method. This standardized bias is almost equivalent to the level of validity which is commonly sought of less than 5 percent. The NN, NNN and KBM give a higher bias level, at around 8 percent. Hence, the result from CN method provides the "best quality" after matching.

The effect of fertilizer adoption on rice revenue is significantly positive before matching. In the CN technique, the impact is around 0.3 (thou. VND/m^2). The effect is slightly smaller in the KBM, greater in the NNN, whereas the effect is not significant in the NN. The relatively significant levels of these matching techniques are reflected correspondingly by the standardized bias test. The comparable results of chemical fertilizer adoption on rice revenue are shown similar as yield, in term of matching methods. Compared with the effect before matching, the magnitude after matching is halved which is in line with the effect on yields.

Concerning household welfare, the effects of fertilizer adoption on income from agriculture are significantly positive and equal to around 0.27 on the basic of the CN method. Since income is measured logarithmically, this can be interpreted as implying that the average income from agriculture in the group of adopters is 27 percent point higher than the group of non-adopters. The impact after matching of agricultural income is slightly greater than before matching which is 23 percent point. However, the effect of adoption on total income is lower than on agricultural income, and approximately 17 percent point. The estimates prior to matching are much higher, at 50 percent. The difference could be understood as the total income of household includes other sources of income rather than income from rice production. This is reasonable if the effect before matching is overestimated. The real implication is that adoption is much more likely for richer households, biasing the unmatched estimates.

Panel B shows the estimates for the households which did not suffer disasters. The balancing tests for NNN and CN matching methods display the lowest standardized bias with around 5.6 percent and 5.8 percent, respectively. The effects of fertilizer adoption are all significantly positive in these two techniques whereas the effects in NN and KBM are found to be less statistically significant and are associated with the higher standardized bias of around 9 percent. The estimates prior to matching also are higher than estimates after matching, except for agricultural income. Compared to the whole sample, the adoption effects in the group of households without disasters is found to be higher on yield and revenue, but lower on incomes. So these results may reflect that the positive effect between chemical fertilizer users and non-users is higher on the productivity if farmers did not suffer from natural disaster loss or damage.

Panel C reports the estimates for households that suffered disasters. The magnitude of the effects is smaller than in these two above groups of households, and also less statistically significant reflecting the much lower number of observations. The balancing test show a high standardized bias value in all matching technique. The lowest standardized bias is 11 percent in KBM, and the highest is 23 percent in NN matching. The most important reason is likely to be the small number of non-users, only 55 households. This factor makes it more difficult to find a similar fertilizer-using household and so increases the bias.

Continuous treatment

This case is different from the binary case in two main respects. First, the average potential effect is evaluated in terms of the extent chemical fertilizer used. In this context, the estimates relate only to fertilizer users. Second, the sub-sample of fertilizer users is sufficiently large enough to allow us to evaluate separately the impact on groups of households suffered a particular type of disaster. We investigate how the dose-response of fertilizer use affects on outcomes distinguishing between different groups of households. The groups of fertilizer users are further separated into five categories containing respectively the whole sample, households which did not suffer any disasters, households which suffered flood or storm or landslide disasters, households which suffered drought, and households which suffered plant disease or insect disasters.

The application of continuous treatment follows the implementation described in the previous section by using the Stata package of Bia and Mattei (2008). The generalized propensity score can be estimated by OLS, in which the values of chemical fertilizer used are regressed on pre-treatment characteristics. The model specifications are the same as in the binary case. Since the regression results show that that data depart from normality irrespective of whether the data are measured in levels or in logs, we use the Box-Cox transformation for chemical fertilizer used to improve the normality of variables. [19] The Box-Cox regression finds the maximum likelihood estimates of the Box-Cox transform. The estimate of GPS is reported in Table 3.15. Then the distributions across all categories of households satisfy the assumption of normality at 5 percent or 10 percent level of significance. As the binary case, the estimates of regression in this step only accounts for computing GPS, and does not have impact on interpretation of the treatment effect.

In addition, the treatment range or the amount of fertilizer used is divided into two intervals by the sample mean. To test the balancing property (BP), the estimated GPS is divided into four blocks in the group of households without suffering disasters, and three blocks in each group of households suffered from disasters. The number of GPS block in the sub-sample of households without disaster shocks is greater than the other groups due to a larger sample size. The GPS blocks correspond to the distribution of GPS percentiles. Explicitly, all of four GPS blocks conditional on median of the two treatment intervals are categorized by 25th, 50th, 75th and 100th centile of GPS distribution. Similarly, three GPS blocks in the group of household suffered disaster shocks are 25th, 75th and 100th centile of GPS distribution. Within each block GPS, the mean difference is compared between two intervals of the treatment: above and below average in terms of fertilizer adoption. There are four mean differences in the sub-sample of households who did not suffered disasters and three mean differences in each group of households who suffered from disasters. To compute the mean t-statistics, the corresponded four and three blocks are weighted by the number of observations. Table 3.16 reports the values and distributions of the GPS evaluated at the representative point of each treatment intervals.

The last rows of Table 3.16 and Table 3.17 show the t-statistic test of balancing property. According to the guidance of Bia and Mattei (2008), the "order of magnitude" interpretations of the test is classified into 5 levels: from supported evidence to decisive evidence against balancing property.[20] The results show that the *t* test for the whole sample is 1 percent; the group of households suffered flood/storm/landslides is significant at 10 percent level and of other groups is significant at 5 percent level. This implies that the means of covariates used for estimates are indifferent at 5 and 10 percent level of significance across groups of low and high

---

[19] The Box-Cox transformation is defined as (Box & Cox, 1964): $y_i^{\lambda} = \begin{cases} \frac{y_i^{\lambda}-1}{\lambda}, & if\ \lambda \neq 0 \\ \log(y_i), & if\ \lambda = 0 \end{cases}$

[20] (i) If *t* test is significant at less than 1 percent: Decisive evidence against BP. (ii) If *t* test is significant from 1 percent to less than 5 percent: Strong to very strong evidence against BP. (iii) If *t* test is significant from 5 percent to less than 10 percent: Moderate evidence against BP. (iv) If *t* test is significant from 10 percent to less than 20 percent: Very slight evidence against BP. (v) If *t* test is significant at greater 20 percent: Evidence supports BP.

adoption, given the estimated GPS. Comparing these figures with the Bia and Mattei (2008), the evidence against the balancing property may be considered as only slight or moderate.

In Table 3.17, the expected outcome indicators are calculated using a flexible function of the treatment variable to estimate GPS according to equation (31). The parameters in polynomial function are estimated by OLS using observed quantity of chemical fertilizer applied and the estimated GPS. Hirano and Imbens (2005) note that the parameters do not allow direct causal inference. They only test whether all parameters involving the GPS are equal to zero. This can be interpreted as a test whether the covariates generate any bias. Table 3.14 shows the estimated result across four groups. The significant parameters of GPS indicate that they are relevant and matter in reducing bias of response function. The linear estimated GPS, The quadratic estimated GPS and whereas the interaction term between GPS and treatment appear highly significant in the whole sample and in the group of households without shocks. These variables have less statistically significant effects in the groups of households who suffered from disasters.

Finally, using these parameters, we further estimate the mean difference of outcome at each treatment level. Then the entire dose-response equation is mean-weighted using then estimated parameters and estimated GPS at each specific treatment level as expressed in equation (32). The bootstrap procedure (100 replications) is applied to compute the standard errors and confidence intervals. Table 3.14 reports the estimated effect of fertilizer adoption on outcomes. The impact evaluation is specified in average of dose-response function and at certain level of treatment. The treatment level is the degree of chemical fertilizer used which belongs to a set of evenly spaced values *t0, t1, ..., t10*. The levels of treatment at *t0* and *t10* are the min and max values of chemical fertilizer used. The second column is the average dose-response computed at each level of treatment or at each level of chemical fertilizer adoption. The marginal casual effect in the third column reflects the variation of treatment on the variation of outcome. This also is referred as the derivative of dose-response function. We employ the treatment gap of 0.1 to estimate the marginal potential outcome. In particular, the marginal effect of treatment can be interpreted as how outcome changes if treatment increases by 0.1 units.

The estimated treatment effects are graphed in Figure 3.5 together with the corresponding 95 percent confidence intervals for both dose-response function and marginal casual effect function. On average, the level of chemical fertilizer adoption is highest in households suffered plant diseases, followed by households suffered flood, households without disasters and household suffered drought; which equals to 0.396, 0.393, 0.359 and 0.352, respectively. The degree of fertilizer adoption does not correspond monotonically with rice yield. The households suffered plant diseases/insects turn out to have the lowest rice yield while the highest yield is from households who suffered flood, then households without disaster and households suffered drought. This order of effect on rice revenue is similar, except the magnitude of impact in households without natural shocks is a bit lower than in households with drought. Thus, the effect of chemical fertilizer impact most on household that experienced plant diseases or insect problems.

In order to investigate the effect of different levels of fertilizer adoption, we check marginal effect over the wide-range of treatment. In all groups of households, the marginal effect increases productivity at low level of fertilizer adoption, but decreases at higher intensity of adoption. However, the confidence intervals are very wide at high levels of application. It is not clear that the downturn is statistically significant. Observed at similar levels of treatment, the rice yield in group of households which did not experience disasters is relatively higher than households suffered disasters. Overall, the treatment effect functions appear to be consistent with a declining marginal impact and in some cases zero marginal impact at higher levels of application.

About household welfare, we observe the effect of adoption on agricultural income and total income. The average dose-response on household welfare has a quite opposite order comparing to the effect on productivity. The households suffered flood have lowest average agricultural income and total income. The households suffered plant diseases have the highest agricultural income, and households suffered drought have the highest total income. Furthermore, the results by treatment levels show the change in the effect more clearly. The higher level of fertilizer used has a negative relationship to agricultural income for households without disasters, except for the treatment level t0 and t1. This marginal casual effect could be explained by the idea of returns to scale. The farmers in Vietnam are constrained by small farm size that lead to a lower degree of increased output when input increases. The empirical evidence also supports this trend in which there is negative relationship of mechanical adoption with credit constraint and farm size (Alviar, 1972; Greene, 1973; Pingali, Bigot, & Binswanger, 1987; Weil, 1970).

On the other hand, the intensity of fertilizer used has a positive impact on agricultural income for households suffered plant diseases or insects, except at the lowest treatment level t0. The negative marginal impact of fertilizer adoption also are occurred at only t3 and t4 level of households suffered drought, and t0, t4, t5, and t6 of households suffered flood. For total income, the marginal effect appears negative with lower adoption level in households without disasters, but with higher adoption level in households suffered flood while the negative relationship of marginal effect is shown in the middle level of treatment in households suffered flood and plant diseases. Differently from the effects discussed in the previous section, agricultural income and total income might fail to capture the effect of fertilizers since the household welfare could come from other agricultural sources and other kinds of income, rather than only from rice cultivation.

3.5.3. Instrumental variables

Table 3.18 reports the estimates from using OLS and IV methods for the case of binary treatment and for three categories of sample as in matching method: the whole sample the group which was free from disaster, and the group which suffered at least one disaster. Overall, the effect of chemical fertilizer adoption on outcomes is positive and statistically significant at less than 1 percent. However, the effects in IV estimates are two or three times as high as those in OLS estimates. As in matching method, the magnitude of treatment effect on yield and revenue is highest in group of households without disasters, followed by group of all sample, and group of households with disaster.

We also report checks the on the validity of these two methods. Panel B of Table 3.18 shows the result for the 2SLS estimates. The IVs, fertilizer information and crop insurance, appear individually to correlate with fertilizer adoption at 1 percent level of significance in all three categories of household. The weak IV test is designed for two key conditions of IV method. First, the F test reveals that these two IVs are jointly different from zero, and hence satisfies the relevance condition of IV. Second, the null hypothesis in the Sagan test states that the excluded instruments are correctly excluded from the estimated equation, i.e. that at least one excluded instrument is uncorrelated with the structural error. The test statistic has the chi-squared distribution. Rejection casts doubt on the validity of the instruments. The result shows that only in the group of household with disaster, and all groups in revenue outcome, satisfy the overidentifying condition.

The Hausman statistic is computed as a test of endogeneity by comparing the IV estimates and OLS estimates. The null hypothesis states that an OLS estimator of the same equation would yield consistent estimates. A rejection of the null indicates that endogenous regressors' effects on the estimates are meaningful, and IV techniques are required. The result shows that all regressions in 2SLS strongly reject the null hypothesis implying that the IV method is preferred to OLS. Thus, the evidence favors IV estimates than OLS estimates.

We now turn to the case of continuous treatment. Table 3.19 reports OLS and IV estimates across five groups of households depending on their disaster status. As in the binary case, the effect of chemical fertilizer adoption is higher in IV estimates than in OLS estimates. The effects of treatment on yield and revenue are higher in the group of households without disaster is higher than in the whole sample. These effects are lower in the groups of households suffered from disaster, or statistically insignificant. The effects of treatment on agricultural income and total income are positive, but inconsistent across the groups of households. Again all the regressions satisfies two first assumption of IV and thus the tests also favor IV estimates over OLS estimates.

3.5.4. Heterogeneity effects

Results obtained from both methods are almost similar in terms of direction (positive or negative). Nonetheless, in the case of binary treatment, the effects of chemical fertilizer adoption shows a greater impact in IV method than in PSM method. In the case of continuous treatment, although matching method estimates over different intervals and IV method estimates on average of the sample, we find somewhat non-linear relationship in matching method in comparison with statistically significant positive impact in IV method. As pointed out in the framework analysis, we suspect the flaw of IV method arise from heterogeneity. More specifically, households are likely to use or use more fertilizer with partial or full knowledge of gain in expected outcomes as they live in areas with advantage agro-ecological conditions for cultivation. The higher estimated effects of adoption on outcomes from IV method could be additionally gained by such advantages. In contrast, households may not use or use less fertilizer as the awareness of their less favorable and risky (e.g. natural disasters) conditions. The adoption decision therefore is made with the knowledge of gain or loss to adoption after controlling observable characteristics, and IV cannot identify treatment effects. Moreover, in this situation, households attain relevant information of fertilizer and buy insurance for crop loss/ damage usually induced by the awareness of their expected gain or loss outcomes. Instrumental variables hence violate the basic assumption although the tests for validity show supporting results. On the other hand, we notice that it could also violate the unconfoundedness assumption. However, the matching method helps to diminish this heterogeneity problem by estimating comparable outcomes over households with similar observable characteristics, and even dropping the highly different households.

We further investigate in a subsample of households characterized by specific features that may cause heterogeneity. The descriptive statistic in Table 3.20 shows that more than 80 percent of non-adopters reside in two mountainous provinces where the agro-ecological conditions are unfavorable for crops. To carry out a quick test, we examine the effects of adoption looking only at these two provinces. In Table 3.21, the effects of adoption from PSM method are found considerably smaller and less statistically significant in this subsample than in the whole sample. The decline of treatment effect seems to be an extraction of the additional impact from adoptions in the regions with more favorable agro-ecological conditions. Thus in the analysis with four kinds of matching techniques, the nearest-neighbor non-replacement matching and caliper matching are found to have lower degree of bias. These techniques impose a higher level of similarity and hence eliminate poorly matched observations which are likely lead to such overestimation.

3.6. Conclusion

In the circumstance of the recent food crisis, the challenge of meeting rising global demand for food needs to be paid more attention. One of the key responses is to increase productivity under resource scarcity (water, land, nutrients, and energy) by adopting technology and using input efficiently. This is a case study of Vietnam where rice is the most important crop for farmers and has contributed significantly to poverty reduction over the past two decades following the transition to a liberalized economy. Accounting for more than 30 percent of total expenditure on inputs, chemical fertilizer adoption is an essential factor that affects crop yields. Moreover, Vietnam is one of the most frequently and worst affected area for natural disasters, which negatively affect farmers. Hence, the main objective of the study has been to investigate the determinants and to assess the impact of technology adoption and the degree of adoption on paddy cultivation in rural Vietnam under different conditions of natural disaster incidence. Additionally, the analysis attempts to implement the propensity matching method in comparison with IV method for impact evaluation, and points out the advantage of matching over IV for this case.

The effects of technology adoption on the productivity and household welfare are not straightforward. The main difficulty is that those agents who are most likely to benefit from adopting a treatment (here fertilizer application) will also be the most likely adopters. The estimated effects of adoption are therefore likely to exceed those that would apply to a randomly selected adopter. Due to these concerns over selection bias, the analysis in this thesis has applied the propensity matching method to deal with this problem. The idea of this approach aims to imitate a randomized experiment by comparing the outcomes of units that applied technology versus to those that did not. The analysis implements two approaches of matching method to evaluate the impact of fertilizer adoption. First, the matching method for binary treatment examines the effect of chemical fertilizer adoption on outcomes between users and non-users. Second, the matching method for continuous treatment explores the effect of the degree of chemical fertilizer adoption among users. The tests are applied for all groups of households under different situation of natural shocks.

First, the findings show that determinants affecting decision of adoption differ from those affecting decision of adoption intensity. Second, the results exhibit that the adoption of chemical fertilizers has a significant positive impact on both productivity and household welfare. The level of impact is higher for households without suffering natural shocks on cultivation. Although the impacts are unsurprised, this shows overestimated impacts between before matching versus after matching. The effects are complex when take into consideration of natural disasters. The impacts of level of fertilizer adoption are non-linear across groups of households. In term of productivity, the rice yield and revenue are lower for households that experienced natural disasters. These effects show that yields increase until reaching a turning point of certain level of fertilizer adoption, and then start to decrease. However, the turning points are different for households that suffered particular forms of natural disaster. The group of households that experienced drought has lowest value of turning point, and the highest point appears in group of households that did not suffer disasters.

The movement of effect on household welfare differs between agricultural income and total income. The effect of increased application shows a positive marginal impact for agricultural income for households that suffered disasters, but for a negative impact form household without disasters. The finding indicates that higher intensity of fertilizer adoption significantly improves the income from agricultural source when households face disasters to their paddy crop whereas the effect on total income increase for lower intensity of adoption and decrease for higher intensity of adoption in the groups of households suffered disasters.

From a methodological perspective, the IV method is usually employed for dealing with endogenous and selection bias issues. Instead of that, we apply the non-parametric propensity score matching. This has a couple of advantages. First, in the context of the sample employed in this thesis, technology adoption depends in large measure on regional conditions. The IV method appears infeasible to estimate under the significant heterogeneity. The matching method overcomes this issue by generating counterfactual and estimating among the observations with similar characteristics, which remarkably decreases the heterogeneity. Second, another advantage of PSM method is that, by applying different matching techniques, we can drop highly dissimilar observations. Thus, the estimates partly clear out the observations whose otherwise different characteristics might confuse outcomes. It is possible that the IV method overestimates treatment effects for this reason. Third, in case of continuous treatment, the approach investigates in detail of the impact through different level of treatment, where we can diagnose a non-linear relationship.

Overall, the study finds substantial effect of fertilizer adoption of paddy cultivation on productivity and welfare. On the one hand, the results show levels of effect are associated with the intensity of adoption across natural disaster conditions. This evidence suggests that the farmers could increase returns by not only deciding on whether to adopt, but also deciding to the extent of adoption. The findings also indirectly relate to the poor who are constrained by lack of credit access and who are vulnerable to shocks.

Around 90 percent of non-adopters are in the two provinces where the poverty rate is among the highest in the country. It is likely that they lack of capital to invest input for crop. Moreover, these low-income farmers also lack resources to cope with shocks. If farmers reside in the area that is frequently hit by disasters, they might to be unwilling to invest on risky technology adoption. Consequently, they are stuck in low productivity, low return, low risk activities, and trapping in poverty (Dercon & Christiaensen, 2011; Zimmerman & Carter, 2003).

References

Bia, M., & Mattei, A. . (2008). A Stata package for the estimation of the dose–response function through adjustment for the generalized propensity score. *Stata Journal, 8*(3), 354-373.

Blundell, R. , & Dias, M. C. . (2002). *Alternative Approaches to Evaluation in Empirical Microeconomics*. The Institute for Fiscal Studies and University College London. London.

Bound, J. , Jaeger, D. A. , & Baker, R. M. . (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak. *Journal of the American Statistical Association, 90*(430), 443-450.

Box, George E. P., & Cox, D. R. . (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological), 26*(2), 211-252.

Cochrane, W. (1993). *The development of American agriculture: A historical analysis*. University of Minnesota, Minneapolis.

Dercon, S., & Christiaensen, L. (2011). Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia. *Journal of Development Economics, 96*(2), 159-173.

Duflo, Esther, Kremer, Michael, & Robinson, Jonathan. (2008). How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya. *The American Economic Review, 98*(2), 482-488. doi: 10.2307/29730068

Gershon, F., Just, R. E., & Zilberman, D. . (1985). Adoption of agricultural innovations in developing countries: A survey. *Economic Development and Cultural Change, 33*(2), 255-298. doi: 10.2307/1153228

Harmeling, S. (2011). Global climate risk index 2012: Who suffers most from extreme weather events? Weather-related loss events in 2010 and 1991 to 2010.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica, 46*(6), 1251-1271.

Hazell, Peter B.R. . (2009). The Asian Green Revolution. In I. D. P. 00911 (Ed.): The International Food Policy Research Institute.

Heckman, J. (1997). Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations. *The Journal of Human Resources, 32*(3), 441-462.

Heckman, J., LaLonde, R., & Smith, J. (1999). The economics and econometrics of active labor market programs. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (1 ed., Vol. 3). Amsterdam: Elsevier.

Heckman, J., & Robb, R. (1985). Alternative Methods for Evaluating The Impact of Interventions: An Overview. *Journal of Econometrics, 30*(1985), 239-297.

Heckman, J., Urzua, S., & Vytlacil, E. (2006). Understanding Instrumental Variables In Models With Essential Heterogeneity. *Review of Economics and Statistics, v88*(3, Aug), 389-432.

Hirano, K., & Imbens, G. W. (2005). The Propensity Score with Continuous Treatments *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives* (pp. 73-84): John Wiley & Sons, Ltd.

Hoang, K. , Pham, V. B. , & Howeler, R. H. . (2000). *A review of cassava in Asia with country case studies on Thailand and Viet Nam.* Paper presented at the THE VALIDATION FORUM ON THE GLOBAL CASSAVA DEVELOPMENT STRATEGY, Rome.

Hossain, M., Tran, T. U., & Janaiah, A. (2003). Vietnam's experience with hybrid rice. *Economic and Political Weekly, 38*(25), 2523-2529.

Ivanic, M. , & Martin, W. (2010). *Promoting global agricultural growth and poverty reduction.* Paper presented at the 114th Seminar, Berlin, Germany.

Jeffrey, S. (2000). A critical survey of empirical methods for evaluating active labor market policies. *Schweizerische Zeitschrift fuer Volkswirtschaft und Statistik, 136*(3), 1-22.

Khandker, S. R., Koolwal, G. B., & Samad, H. (2010). *Handbook on impact evaluation: Quantitative methods and practice*: The World Bank. Washington, D.C.

Lechner, M. (1999). Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany after Unification. *Journal of Business & Economic Statistics, 17*(1), 74-90.

Leuven, E. , & Sianesi, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.

Nebitt, H. , Johnston, R., & Solieng, M. (2004). *Mekong River water: will river flows meet future agriculture needs in the Lower Mekong Basin.* Paper presented at the Water in Agriculture: ACIAR Proceedings No. 116.

Nguyen, V. B. , Mutert, E. , & Cong, D. S. (2003). *BALCROP: Balanced fertilization for better crops in Vietnam*.

Organization for Economic Cooperation and Development's Development Assistance Committee (OECD-DAC). (2002). Glossary of Key Terms in Evaluation and Results-Based Management Proposed Harmonized Terminology. OECD, Paris.

Pagan, A., & Ullah, A. (1999). *Nonparametric Econometrics*. Cambridge: Cambridge University Press.

Ravallion, M. . (2008). Evaluating anti-poverty programs. In E. E. Robert & P. T. Schultz (Eds.), *Handbook of Development Economics Volume 4*. Amsterdam, North-Holland.

Rosegrant, M. W., & Hazell, P. B. R. (2000). *Transforming the rural Asia economy: The unfinished revolution*. Hong Kong: Oxford University Press.

Rosenbaum, P., & Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika, 70*(1), 41-55.

Rosenbaum, P., & Rubin, D. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity *The American Statistician, 9*(1), 33-38.

Rosenzweig, M. R., & Binswanger, H. P. (1993). Wealth, weather risk and the composition and profitability of agricultural investments. *The Economic Journal, 103*(416), 56-78. doi: 10.2307/2234337

Rubin, D. (1974). Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology, 66*(5), 688-701.

Schultz, T. W. (1964). *Transforming traditional agriculture*. Yale University, Press, New Haven, CT.

Smith, J. A. , & Todd, P. E. . (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics, 125*(1–2), 305-353.

Sunding, D. , & Zilberman, D. . (1999). The agricultural innovation process: Research and technology adoption in a changing agricultural sector *Handbook of Agricultural Economcis*.

The World Bank. (2009). *Global facility for disaster reduction and recovery: Disaster risk management programs for priority countries*.

Thomas, T. , Christiaensen, L. , Do, Q. T. , & Le, D. T. . (2010). Natural disasters and household welfare: Evidence from Vietnam: The World Bank.

White, H. . (2006). Impact evaluation: The experience of the independent evaluation group of the World Bank. Washington D.C. : World Bank.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Wooldridge, J. M. (2005). *Introductory econometrics: A modern approach* (3th ed.).

World Bank. (2008). The development impact evaluation (DIME) initiative Washington D.C.

Young, B. K. , Wailes, J. E. , Cramer, L. G., & Nguyen, T. K. (2002). Vietnam's rice economy: Developments and prospects. Division of Agriculture, University of Arkansas.

Yu, Bingxin, Zhu, Tingju , Breisinger, Clemens , & Nguyen, Manh Hai. (2012). *Examine the agriculture, poverty, and climate change: Nexus in Vietnam*. Paper presented at the International Association of Agricultural Economists (IAAE) Triennial Conference, Foz do Iguaçu, Brazil.

Zimmerman, F. J., & Carter, M. R. (2003). Asset smoothing, consumption smoothing and the reproduction of inequality under risk and subsistence constraints. *Journal of Development Economics, 71*(2), 233-260.

Tables

Table 3.1: The major hazard events in Vietnam, 1999-2008

Year	Event	No. of people death	No. of people injured	No. of people missing	Economic loss (bil. VND)	Area affected
2008	Storm Kammuri	133	91	34	1,939.73	9 North and Central provinces
2007	Storm Lekima	88	180	8	3,215.51	17 North and Central provinces
2006	Storm Xangsane	72	532	4	10,401.62	15 Central and Southern provinces
2005	Storm No. 7	68	28		3,509.15	12 North and Central provinces
2004	Storm No. 2	23	22		298.199	5 Central provinces
2003	Rains and floods	65	33		432.471	9 Central provinces
2002	Floods	171			456.831	Mekong River Delta
2001	Floods	393			1,535.91	Mekong River Delta
2000	Flash floods	28	27	2	43.917	5 Northern provinces
1999	Floods	595	275	29	3,773.80	10 Central provinces

Source: The World Bank (2009) extracted from CCFSC's website.

Table 3.2: Land with problem of households (%)

Types of problem	Ratio over total land	Ratio over paddy land
Erosion	4.75	21.13
Dry land	29.03	92.01
Low-lying land	5.54	17.33
Sedimentation	2.42	2.42
Landslide	3.52	13.97
Stony soils/ clay	3.07	9.74
Others	0.78	3.26
No problem	41.88	97.26

Source: Calculation from the 2010 VARHS (N=2,280).
Note: The information of land problem is only available for crops land, forestry land, grass land, and garden; and excludes residential land, land for aquaculture, and uncategorized land, which take up around 9 percent of the total land area.

Table 3.3: Household's land plots suffered natural disasters in 2010

	Disaster	Flood/Storm/ Landslide		Drought		Plant diseases/ Insects	
		All	Exclusive	All	Exclusive	All	Exclusive
No. HH	696	274	217	189	151	314	251
Percent	30.53	12.02	11.89	8.29	6.62	13.77	11.01
Total obs.	2,280						

Source: Calculation from the 2010 VARHS.
Note: Around 1 percent of uncategorized disaster is excluded out of those groups.

Table 3.4: Chemical fertilizer adoption (1,000 VND/m^2).

Total sample (N=2,280)	Non-user (N=267)	User (N=2,013)				
		Mean	S.D.	Min	Med.	Max
Chemical fertilizer use	0	0.37	0.17	0.001	0.33	2.25

Source: Calculation from the 2010 VARHS.

Table 3.5: Outcomes by treatment status

	Non-user	User		Difference (%)	
		Under mean	Over mean	User Vs. Non-user	Under Vs. Over mean
No. of observations (N=2,280)	267	1132	881		
Percent	11.71	49.65	38.64		
Rice yield (kg/m^2)					
Total sample	0.276	0.397	0.477	57***	20***
Without disaster	0.274	0.408	0.486	61***	19***
Flood/Storm/Landslide	0.410	0.400	0.470	6	17***
Drought	0.228	0.401	0.439	82***	10
Plant disease/Insects	0.251	0.332	0.461	54***	39***
Rice revenue (thou. VND/m^2)					
Total sample	1.559	2.052	2.342	40***	14***
Without disaster	1.572	2.132	2.391	43***	12***
Flood/Storm/Landslide	2.378	1.937	2.288	-11	18***
Drought	0.907	2.064	2.268	136***	10
Plant disease/Insects	1.380	1.736	2.195	40***	26***
Agricultural income (thou. VND/capita)					
Total sample	280	376	458	132***	82***
Without disaster	275	381	424	125***	43**
Flood/Storm/Landslide	170	372	337	185*	-35
Drought	368	451	530	114	79
Plant disease/Insects	321	285	767	165	482***
Total income (thou. VND/capita)					
Total sample	772	1176	1484	70***	26***
Without disaster	783	1237	1503	73***	22
Flood/Storm/Landslide	562	971	988	74*	2
Drought	772	1292	1451	75	12
Plant disease/Insects	761	857	1916	71*	124***

Source: Calculation from the 2010 VARHS.

Note: (i) Significance at 10%, 5% and 1% are denoted as *, **, ***. Standard deviations are displayed in parentheses.

(ii) 2 observations of fertilizer user and 3 observations of rice yield are excluded out of the sample since their values are over ten times higher than means.

(iii) The average values of fertilizer users are not reported here for simplicity.

(iv) Incomes are reported at current prices.

Table 3.6: Summary of observable characteristics of the sample.

	Non-user	User		Difference (%)	
		Below mean	Over mean	User Vs. Non-user	Below Vs. Over mean
No. of observations (N=2,280)	267	1,132	881		
Percent of total sample	11.71	49.65	38.64		
Household characteristics					
HH size	5.58	5.00	4.61	-13***	-8***
No. of members < 10	1.14	0.78	0.61	-38***	-22***
No. of members 10-15	0.82	0.56	0.50	-35***	-11*
No. of members 16-25	1.23	1.17	1.09	-8	-7
No. of members 26-60	2.09	2.09	1.97	-2	-5***
Ethnic minority (dummy)	0.95	0.53	0.39	-51***	-25***
Education of head (0-12 grades)	3.49	5.45	6.24	66***	14***
Male headship (dummy)	0.92	0.85	0.85	-8***	-1
Age of head	44.38	49.04	50.29	12***	3**
No. of labors for wage/salary	0.70	0.95	1.04	42***	10*
No. of labors in HH agriculture	3.55	3.38	3.09	-8***	-9***
No. of labors in HH business	0.49	0.41	0.37	-20*	-9
Hired labor outside	0.22	0.44	0.46	107***	6
Loan value (thou. VND)	6,815	8,348	11,281	41*	35***
Assistance receipt (thou. VND)	245	2,614	2,761	993***	6
Distance to People's Committee (km)	5.99	3.00	2.53	-53***	-16***
Land and crop characteristics					
Land independent on public irrigation	0.45	0.25	0.25	-45***	1
Paddy land (acres)	1.96	1.69	1.92	-9	14
Ratio of land without problem (%)	0.44	0.47	0.46	5	-2
Hybrid seed (dummy)	0.32	0.48	0.53	57***	10**
Improved seed (dummy)	0.19	0.30	0.26	49***	-15**
Old local seed (dummy)	0.47	0.18	0.19	-61***	10
Instrumental variables					
Received info. of fertilizers	0.69	0.90	0.92	-32***	-3*
Crop insurance (dummy)	0.10	0.39	0.47	-305***	-20***
Crop insurance (ln)	0.54	1.76	2.14	-257***	-21***

Source: Calculation from the 2010 VARHS.

Note: Significance at 10%, 5% and 1% are denoted as *, **, ***

Table 3.7: Heckit model of the whole sample

	Model 1		Model 2	Model 3	Model 4
	Adoption	Level of adoption	Level of adoption	Level of adoption	Level of adoption
HH size	0.058	0.010			
	(0.080)	(0.011)			
No. of member <10	-0.098	-0.023*			
	(0.094)	(0.014)			
No. of members 10-15	-0.190**	0.005			
	(0.091)	(0.014)			
No. of members 15-25	-0.171**	-0.016			
	(0.082)	(0.011)			
No. of members 25-60	-0.137	-0.028**			
	(0.085)	(0.011)			
Education of head	0.261***	0.007	0.007***	0.008***	
	(0.038)	(0.006)	(0.002)	(0.002)	
Education of head (squared)	-0.011***	0.000			
	(0.003)	(0.000)			
Edu. of head * Ethnic	-0.145***	-0.000			
	(0.018)	(0.002)			
Gender of head (male)	-0.253*	0.013			
	(0.147)	(0.018)			
Age of head	-0.001	0.000			
	(0.004)	(0.001)			
No. of labors for salary	0.013	0.015***	0.010*		0.012**
	(0.043)	(0.006)	(0.006)		(0.006)
No. of labors in HH agriculture	0.214***	-0.003*	-0.007*		-0.011***
	(0.040)	(0.006)	(0.004)		(0.004)
No. of labors in HH business	-0.151***	-0.012**	-0.016**		-0.014**
	(0.047)	(0.007)	(0.007)		(0.007)
Hired labor outside	0.391***	-0.013			
	(0.092)	(0.012)			
Loan (ln)	-0.005	0.004***	0.003***	0.003***	0.004***
	(0.009)	(0.001)	(0.001)	(0.001)	(0.001)
Assistance receipt (ln)	0.051***	0.001			
	(0.016)	(0.002)			
Distance to People's Committee	-0.061***	-0.005**	-0.005**	-0.006***	-0.006***
	(0.009)	(0.002)	(0.002)	(0.002)	(0.002)
Land independent on public irrigation	-0.182***	0.040***	0.039***	0.041***	0.034**
	(0.084)	(0.014)	(0.014)	(0.013)	(0.014)
Paddy land	0.029	0.002			
	(0.088)	(0.006)			
Paddy land * Problematic land	-0.026	-0.004			

	(0.025)	(0.003)			
Paddy land * Hybrid seed	0.005	-0.003			
	(0.089)	(0.007)			
Paddy land * Improved seed	-0.014	0.000			
	(0.087)	(0.006)			
Paddy land * Old local seed	-0.074	-0.002			
	(0.086)	(0.006)			
N	2260	1995	1995	1995	1995
Log likelihood		-698	-710	-717	-719
LR test, chi2 (1)		0.35	0.91	1.52	3.42*

Note: ***, **, * denote the level of significance at 1%, 5%, and 10%.

Table 3.8: Heckit model of the whole sample including disaster variables

	Model 1		Model 2	Model 3	Model 4
	Adoption	Level of adoption	Level of adoption	Level of adoption	Level of adoption
HH size	0.058	0.011			
	(0.080)	(0.011)			
No. of member <10	-0.091	-0.023*			
	(0.095)	(0.014)			
No. of members 10-15	-0.192**	0.004			
	(0.092)	(0.014)			
No. of members 15-25	-0.167**	-0.017			
	(0.082)	(0.011)			
No. of members 25-60	-0.132	-0.028**			
	(0.086)	(0.011)			
Education of head	0.254***	0.007	0.007***	0.008***	
	(0.038)	(0.006)	(0.002)	(0.002)	
Education of head (squared)	-0.010***	0.000			
	(0.003)	(0.000)			
Edu. of head * Ethnic	-0.146***	-0.000			
	(0.019)	(0.002)			
Gender of head	-0.247*	0.014			
	(0.149)	(0.017)			
Age of head	-0.002	0.000			
	(0.004)	(0.001)			
No. of labors for salary	0.010	0.015**	0.010*		0.012**
	(0.043)	(0.006)	(0.006)		(0.006)
No. of labors in HH agriculture	0.212***	-0.004	-0.008**		-0.011***
	(0.041)	(0.006)	(0.004)		(0.004)
No. of labors in HH business	-0.155***	-0.011	-0.015**		-0.014*
	(0.048)	(0.007)	(0.007)		(0.007)

Hired labor outside	0.345***	-0.017			
	(0.094)	(0.012)			
Loan (ln)	-0.008	0.003***	0.003**	0.003**	0.003***
	(0.009)	(0.001)	(0.001)	(0.001)	(0.001)
Assistance receipt (ln)	0.052***	0.001			
	(0.016)	(0.002)			
Distance to People's Committee	-0.065***	-0.005**	-0.005**	-0.006***	-0.006***
	(0.010)	(0.002)	(0.002)	(0.002)	(0.002)
Land independent on public irrigation	-0.196**	0.040***	0.038***	0.040***	0.033**
	(0.085)	(0.014)	(0.014)	(0.014)	(0.014)
Paddy land	0.040	0.002			
	(0.102)	(0.006)			
Paddy land * Problematic land	-0.025	-0.003			
	(0.025)	(0.003)			
Paddy land * Hybrid seed	-0.004	-0.003			
	(0.104)	(0.007)			
Paddy land * Improved seed	-0.021	0.000			
	(0.103)	(0.006)			
Paddy land * Old local seed	-0.082	-0.002			
	(0.102)	(0.006)			
Disaster 1 (Flood/Storm)	0.434***	0.034**	0.029*	0.032*	0.032*
	(0.155)	(0.017)	(0.017)	(0.017)	(0.017)
Disaster 2 (Disease/Insect)	0.650***	0.003	0.004	0.005	0.005
	(0.185)	(0.020)	(0.020)	(0.020)	(0.020)
Disaster 3 (Drought)	-0.008	0.035**	0.040**	0.037**	0.039**
	(0.115)	(0.017)	(0.017)	(0.017)	(0.017)
N	2260	1995	1995	1995	1995
Log likelihood		-683	-694	-701	-703
LR test, chi2 (1)		1	1.83	2.61	4.82**

Note: ***, **, * denote the level of significance at 1%, 5%, and 10%.

Table 3.9: Estimated propensity score from logit model of the binary treatment

Dependant var. is chemical fertilizer user	All sample	Without disasters	With disasters
HH size	0.080	0.154	0.215
	(0.145)	(0.171)	(0.328)
No. of member <10	-0.141	-0.167	-0.228
	(0.171)	(0.202)	(0.381)
No. of members 10-15	-0.320*	-0.451**	-0.065
	(0.166)	(0.197)	(0.376)
No. of members 15-25	-0.295*	-0.416**	0.084
	(0.151)	(0.178)	(0.343)
No. of members 25-60	-0.226	-0.388**	0.216
	(0.158)	(0.184)	(0.387)
Education of head	0.528***	0.572***	0.161
	(0.075)	(0.086)	(0.195)
Education of head (squared)	-0.019***	-0.022***	-0.001
	(0.006)	(0.007)	(0.017)
Edu. of head * Ethnic	-0.325***	-0.341***	-0.172
	(0.045)	(0.049)	(0.112)
Gender of head	-0.462*	-0.590*	-0.263
	(0.280)	(0.331)	(0.715)
Age of head	-0.001	-0.006	0.020
	(0.007)	(0.008)	(0.019)
No. of labors for salary	0.016	-0.025	0.206
	(0.079)	(0.091)	(0.209)
No. of labors in HH agriculture	0.409***	0.503***	-0.164
	(0.075)	(0.088)	(0.203)
No. of labors in HH business	-0.286***	-0.281***	-0.338*
	(0.084)	(0.102)	(0.182)
Hired labor outside	0.696***	0.736***	0.082
	(0.174)	(0.216)	(0.381)
Loan (ln)	-0.012	-0.022	0.024
	(0.016)	(0.019)	(0.041)
Assistance receipt (ln)	0.106***	0.089**	0.202*
	(0.032)	(0.036)	(0.097)
Distance to People's Committee	-0.105***	-0.124***	-0.106***
	(0.016)	(0.021)	(0.034)
Land independent on public irrigation	-0.328**	-0.186	-0.929***
	(0.151)	(0.179)	(0.379)
Paddy land	0.039	0.051	44.461***
	(0.165)	(0.145)	(0.237)
Paddy land * Problematic land	-0.044	-0.059	-0.311*
	(0.046)	(0.056)	(0.175)
Paddy land * Hybrid seed	0.022	-0.003	-44.103

	(0.169)	(0.147)	(0.000)
Paddy land * Improved seed	-0.020	0.005	-44.467***
	(0.164)	(0.152)	(0.234)
Paddy land * Old local seed	-0.133	-0.138	-44.575***
	(0.162)	(0.142)	(0.232)
Constant	1.465***	1.593***	1.109
	(0.496)	(0.572)	(1.297)
N	2260	1570	474
Pseudo R2	0.24	0.27	0.25
Log likelihood	-618	-451	-112

Note: Significance at 10%, 5% and 1% are denoted as *, **, ***. Standard deviations are displayed in parentheses.

Table 3.10: Average treatment effect and balancing test for the whole sample.

Panel A: The whole sample					
	Before matching	After matching			
		Nearest neighbor	Nearest neighbor noreplacement	**Caliper norepalcement**	Kernel based matching
Yield	0.156***	0.056**	0.103***	**0.093*** **	0.067***
	(0.01)	(0.023)	(0.013)	**(0.014)**	(0.019)
Rice revenue	0.618***	0.189	0.394***	**0.33*** **	0.239**
	(0.051)	(0.125)	(0.069)	**(0.073)**	(0.102)
Agr. Income	0.232***	0.328**	0.165**	**0.273*** **	0.175
	(0.062)	(0.129)	(0.068)	**(0.075)**	(0.093)
Total Income	0.501***	0.16	0.067**	**0.169** **	0.133
	(0.048)	(0.123)	(0.059)	**(0.068)**	(0.089)
Treatment assignment					
Untreated	265	265	265	**265**	265
Treat	1995	1812	265	**233**	1812
Balancing test	27.95	8.06	8.36	**5.25**	8.00

Panel B: Without disasters					
	Before matching	After matching			
		Nearest neighbor	**Nearest neighbor noreplacement**	Caliper norepalcement	Kernel based matching
Yield	0.167***	0.067**	**0.111*** **	0.112***	0.073***
	(0.011)	(0.027)	**(0.014)**	(0.016)	(0.021)
Rice revenue	0.67***	0.194	**0.434*** **	0.421***	0.21*
	(0.059)	(0.145)	**(0.075)**	(0.081)	(0.114)
Agr. Income	0.181**	-0.09	**0.207** **	0.241***	0.081
	(0.07)	(0.128)	**(0.075)**	(0.086)	(0.102)
Total Income	0.505***	0.051	**0.125* **	0.141*	0.09
	(0.056)	(0.128)	**(0.069)**	(0.08)	(0.102)
Treatment assignment					
Untreated	210	210	**210**	210	210
Treat	1360	1169	**210**	174	1169
Balancing test	26.82	8.88	**5.64**	5.78	9.46

Panel C: With disasters					
	Before matching	After matching			
		Nearest neighbor	Nearest neighbor noreplacement	Caliper norepalcement	**Kernel based matching**
Yield	0.13***	0.038	0.097***	0.103**	**0.034**
	(0.02)	(0.072)	(0.032)	(0.044)	**(0.042)**
Rice revenue	0.531***	0.41	0.422**	0.401*	**0.23**
	(0.104)	(0.365)	(0.166)	(0.214)	**(0.212)**
Agr. Income	0.351**	-0.93	0.348**	0.588**	**0.463****
	(0.133)	(0.315)	(0.16)	(0.218)	**(0.206)**
Total Income	0.506***	0.285	0.18	0.503	**0.361****
	(0.1)	(0.247)	(0.131)	(0.16)	**(0.166)**
Treatment assignment					
Untreated	55	55	55	55	**55**
Treat	635	635	55	42	**479**
Balancing test	29.59	23.02	14.37	17.38	**10.72**

Note: (i) Significance at 10%, 5% and 1% are denoted as *, **, ***. Standard deviations are displayed in parentheses.
(ii) The log transformation of agricultural income and total income eliminates 20 observations with negative values that make the sample sizes of productivity and household welfare slightly different.
(iii) Common support region are imposed for all matching methods.

Table 3.11: Propensity score distribution before and after matching for the whole sample

Summary of propensity score before matching

	Percentiles		
1%	4.045691		
5%	5.107901		
10%	5.254936	Obs.	23
25%	6.668897	Sum of Wgt.	23
50%	27.9512	Mean	29.29975
		Std. Dev.	21.46029
75%	43.51601		
90%	59.97123	Variance	460.5442
95%	68.60117	Skewness	0.427772
99%	71.3744	Kurtosis	2.054222

Panel A: Nearest-neighbor with replacement technique

Variable	Unmatched/ Matched	Mean		% Bias	% Reduct bias	t-test	
		Treated	Control			t	p>|t|
HH size	Unmatched	4.83	5.57	-36.6		-5.83	0.000
	Matched	4.92	5.05	-6.3	82.9	-1.92	0.055
No. of member <10	Unmatched	0.71	1.14	-40.5		-6.65	0.000
	Matched	0.75	0.93	-17.5	56.9	-5.38	0.000
No. of members 10-15	Unmatched	0.54	0.83	-34.9		-5.76	0.000
	Matched	0.57	0.62	-5.8	83.5	-1.86	0.063
No. of members 15-25	Unmatched	1.14	1.22	-6.7		-1.08	0.280
	Matched	1.17	1.00	13.8	-107.3	4.10	0.000
No. of members 25-60	Unmatched	2.04	2.08	-5.4		-0.85	0.394
	Matched	2.05	2.19	-16.9	-211.6	-4.35	0.000
Education of head	Unmatched	5.78	3.48	60.0		9.24	0.000
	Matched	5.38	5.56	-4.6	92.3	-1.37	0.171
Education of head (squared)	Unmatched	47.82	27.06	50.0		7.35	0.000
	Matched	42.82	47.23	-10.6	78.8	-3.07	0.002
Edu. of head * Ethnic	Unmatched	1.90	3.09	-33.7		-5.48	0.000
	Matched	2.10	1.94	4.4	87.1	1.42	0.157
Gender of head	Unmatched	0.85	0.92	-22.4		-3.12	0.002
	Matched	0.86	0.89	-10.4	53.4	-2.99	0.003
Age of head	Unmatched	49.59	44.36	40.7		6.16	0.000
	Matched	49.03	48.82	1.6	96.1	0.46	0.647
No. of labors for salary	Unmatched	0.98	0.70	28.0		4.04	0.000
	Matched	0.98	1.00	-2.5	91.0	-0.74	0.458
	Unmatched	3.25	3.53	-17.6		-2.72	0.007
No. of labors in HH agriculture	Matched	3.28	3.40	-7.0	60.2	-2.03	0.042
	Unmatched	0.39	0.49	-11.4		-1.84	0.066
No. of labors in HH business	Matched	0.41	0.43	-2.1	81.7	-0.67	0.500
Hired labor outside	Unmatched	0.45	0.22	51.1		7.32	0.000
	Matched	0.41	0.35	13.4	73.8	3.80	0.000
Loan (ln)	Unmatched	4.79	5.04	-5.3		-0.80	0.424
	Matched	4.86	4.55	6.7	-27.7	1.99	0.047
Assistance receipt (ln)	Unmatched	2.95	0.80	68.6		8.93	0.000
	Matched	2.47	2.29	5.8	91.6	1.53	0.126
Distance to People's Committee	Unmatched	2.80	6.00	-71.4		-14.04	0.000
	Matched	2.93	3.25	-7.0	90.2	-2.59	0.010
Land independent on public irrigation	Unmatched	0.25	0.45	-43.5		-7.05	0.000
	Matched	0.27	0.23	7.4	82.9	2.42	0.016
Paddy land	Unmatched	1.78	1.94	-5.1		-0.65	0.513
	Matched	1.79	1.36	13.7	-168.2	4.18	0.000
Paddy land * Problematic land	Unmatched	0.85	0.98	-5.8		-0.79	0.428
	Matched	0.86	0.65	9.2	-57.5	2.86	0.004
Paddy land * Hybrid seed	Unmatched	0.65	0.58	4.0		0.63	0.526

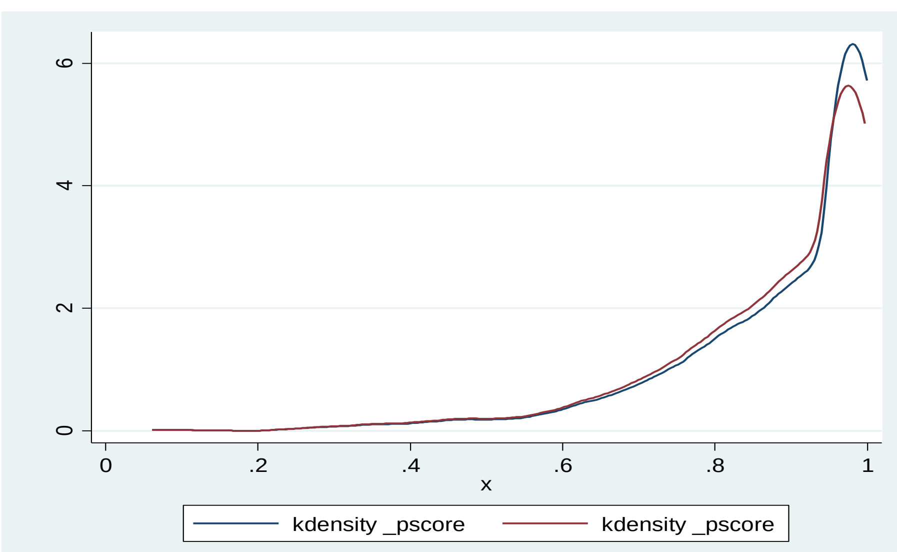
	Matched	0.64	0.63	0.6	85.2	0.17	0.867
Paddy land * Improved seed	Unmatched	0.55	0.35	8.9		1.06	0.288
	Matched	0.53	0.33	8.8	0.3	2.72	0.007
Paddy land * Old local seed	Unmatched	0.51	0.98	-22.2		-3.11	0.002
	Matched	0.55	0.36	9.3	58	3.20	0.001

Summary of propensity score after matching

	Percentiles		
1%	0.5981236		
5%	1.603028		
10%	2.091901	Obs	23
25%	4.603348	Sum of Wgt.	23
50%	7.003312	Mean	8.062615
		Std. Dev.	4.687349
75%	10.62472		
90%	13.82487	Variance	21.97124
95%	16.89578	Skewness	0.397035
99%	17.45477	Kurtosis	2.397067

Sample	Pseudo R2	p>chi2	Mean Bias	Med. Bias
Raw	0.239	0	29.3	28
Matched	0.038	0	8.1	7

Kernel distribution of propensity score before (red) and after (blue) matching



Panel B: Nearest-neighbor without replacement technique							
Variable	Unmatched/ Matched	Mean		% Bias	% Reduct bias	t-test	
		Treated	Control			t	p>|t|
HH size	Unmatched	4.83	5.57	-36.6		-5.83	0.000
	Matched	5.66	5.57	4.9	86.7	0.52	0.604
No. of member <10	Unmatched	0.71	1.14	-40.5		-6.65	0.000
	Matched	1.26	1.14	11.0	72.9	1.17	0.243
No. of members 10-15	Unmatched	0.54	0.83	-34.9		-5.76	0.000
	Matched	0.93	0.83	12.1	65.4	1.27	0.205
No. of members 15-25	Unmatched	1.14	1.22	-6.7		-1.08	0.280
	Matched	1.14	1.22	-6.4	4.0	-0.71	0.479
No. of members 25-60	Unmatched	2.04	2.08	-5.4		-0.85	0.394
	Matched	2.08	2.08	-0.9	83.9	-0.11	0.915
Education of head	Unmatched	5.78	3.48	60.0		9.24	0.000
	Matched	3.34	3.48	-3.5	94.1	-0.39	0.694
Education of head (squared)	Unmatched	47.82	27.06	50.0		7.35	0.000
	Matched	27.71	27.06	1.6	96.9	0.18	0.856
Edu. of head * Ethnic	Unmatched	1.90	3.09	-33.7		-5.48	0.000
	Matched	3.28	3.09	5.4	84.1	0.55	0.581
Gender of head	Unmatched	0.85	0.92	-22.4		-3.12	0.002
	Matched	0.94	0.92	7.1	68.2	1.03	0.301
Age of head	Unmatched	49.59	44.36	40.7		6.16	0.000
	Matched	42.65	44.36	-13.3	67.4	-1.57	0.118
No. of labors for salary	Unmatched	0.98	0.70	28.0		4.04	0.000
	Matched	0.57	0.70	-13.1	53.2	-1.70	0.090
No. of labors in HH agriculture	Unmatched	3.25	3.53	-17.6		-2.72	0.007
	Matched	3.46	3.53	-4.3	75.6	-0.50	0.618
No. of labors in HH business	Unmatched	0.39	0.49	-11.4		-1.84	0.066
	Matched	0.54	0.49	6.1	47.0	0.62	0.539
Hired labor outside	Unmatched	0.45	0.22	51.1		7.32	0.000
	Matched	0.12	0.22	-19.8	61.2	-2.79	0.005
Loan (ln)	Unmatched	4.79	5.04	-5.3		-0.80	0.424
	Matched	5.09	5.04	1.1	79.2	0.13	0.899
Assistance receipt (ln)	Unmatched	2.95	0.80	68.6		8.93	0.000
	Matched	0.36	0.80	-14.0	79.7	-2.67	0.008
Distance to People's Committee	Unmatched	2.80	6.00	-71.4		-14.04	0.000
	Matched	6.58	6.00	12.8	82.1	1.19	0.235
Land independent on public irrigation	Unmatched	0.25	0.45	-43.5		-7.05	0.000
	Matched	0.49	0.45	7.3	83.3	0.78	0.434
Paddy land	Unmatched	1.78	1.94	-5.1		-0.65	0.513
	Matched	1.50	1.94	-13.9	-172	-1.99	0.048
Paddy land * Problematic land	Unmatched	0.85	0.98	-5.8		-0.79	0.428
	Matched	0.68	0.98	-13.1	-125	-1.48	0.140

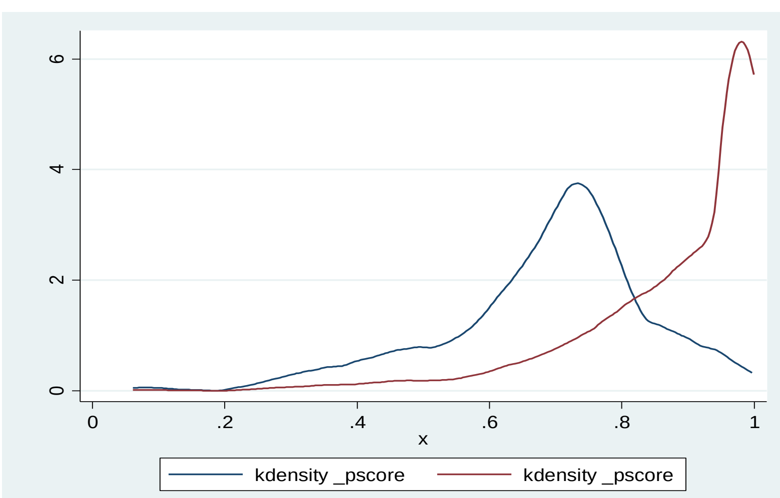
Paddy land * Hybrid seed	Unmatched	0.65	0.58	4.0		0.63	0.526
	Matched	0.53	0.58	-3.6	11.7	-0.51	0.611
Paddy land * Improved seed	Unmatched	0.55	0.35	8.9		1.06	0.288
	Matched	0.20	0.35	-7.2	19.2	-2.31	0.021
Paddy land * Old local seed	Unmatched	0.51	0.98	-22.2		-3.11	0.002
	Matched	0.77	0.98	-10.0	55.1	-0.98	0.329

Summary of propensity score after matching

	Percentiles		
1%	0.8719161		
5%	1.094557		
10%	1.564301	Obs	23
25%	4.280225	Sum of Wgt.	23
50%	7.156632	Mean	8.35642
		Std. Dev.	5.034791
75%	13.08763		
90%	13.89257	Variance	25.34912
95%	13.95326	Skewness	0.310976
99%	19.83522	Kurtosis	2.305237

Sample	Pseudo R2	p>chi2	Mean Bias	Med. Bias
Raw	0.239	0	29.3	28
Matched	0.056	0.012	8.4	7.2

Kernel distribution of propensity score before (red) and after (blue) matching



Panel C: Caliper without replacement technique							
Variable	Unmatched/ Matched	Mean		% Bias	% Reduct bias	t-test	
		Treated	Control			t	p>|t|
HH size	Unmatched	4.83	5.57	-36.6		-5.83	0.000
	Matched	5.68	5.45	11.1	69.6	1.10	0.273
No. of member <10	Unmatched	0.71	1.14	-40.5		-6.65	0.000
	Matched	1.23	1.11	11.3	72.1	1.12	0.264
No. of members 10-15	Unmatched	0.54	0.83	-34.9		-5.76	0.000
	Matched	0.84	0.77	7.6	78.1	0.78	0.437
No. of members 15-25	Unmatched	1.14	1.22	-6.7		-1.08	0.280
	Matched	1.27	1.20	5.5	16.8	0.57	0.567
No. of members 25-60	Unmatched	2.04	2.08	-5.4		-0.85	0.394
	Matched	2.04	2.06	-2.0	63.4	-0.23	0.822
Education of head	Unmatched	5.78	3.48	60.0		9.24	0.000
	Matched	3.56	3.73	-4.6	92.4	-0.49	0.624
Education of head (squared)	Unmatched	47.82	27.06	50.0		7.35	0.000
	Matched	27.58	28.84	-3.0	93.9	-0.34	0.732
Edu. of head * Ethnic	Unmatched	1.90	3.09	-33.7		-5.48	0.000
	Matched	3.20	3.29	-2.7	92.0	-0.27	0.790
Gender of head	Unmatched	0.85	0.92	-22.4		-3.12	0.002
	Matched	0.91	0.91	-1.4	94.0	-0.16	0.870
Age of head	Unmatched	49.59	44.36	40.7		6.16	0.000
	Matched	44.39	44.57	-1.4	96.6	-0.15	0.878
No. of labors for salary	Unmatched	0.98	0.70	28.0		4.04	0.000
	Matched	0.76	0.75	0.9	97.0	0.09	0.927
No. of labors in HH agriculture	Unmatched	3.25	3.53	-17.6		-2.72	0.007
	Matched	3.62	3.52	6.5	63.1	0.67	0.501
No. of labors in HH business	Unmatched	0.39	0.49	-11.4		-1.84	0.066
	Matched	0.45	0.44	0.5	95.7	0.05	0.958
Hired labor outside	Unmatched	0.45	0.22	51.1		7.32	0.000
	Matched	0.22	0.24	-4.7	90.8	-0.55	0.583
Loan (ln)	Unmatched	4.79	5.04	-5.3		-0.80	0.424
	Matched	4.60	5.07	-10.2	-93.7	-1.10	0.271
Assistance receipt (ln)	Unmatched	2.95	0.80	68.6		8.93	0.000
	Matched	0.71	0.90	-6.2	90.9	-0.94	0.349
Distance to People's Committee	Unmatched	2.80	6.00	-71.4		-14.04	0.000
	Matched	5.53	5.07	10.1	85.9	0.96	0.336
Land independent on public irrigation	Unmatched	0.25	0.45	-43.5		-7.05	0.000
	Matched	0.45	0.43	4.6	89.4	0.47	0.642
Paddy land	Unmatched	1.78	1.94	-5.1		-0.65	0.513
	Matched	1.65	1.75	-3.3	35.6	-0.50	0.618
Paddy land * Problematic land	Unmatched	0.85	0.98	-5.8		-0.79	0.428
	Matched	0.65	0.84	-8.5	-46.1	-1.16	0.245

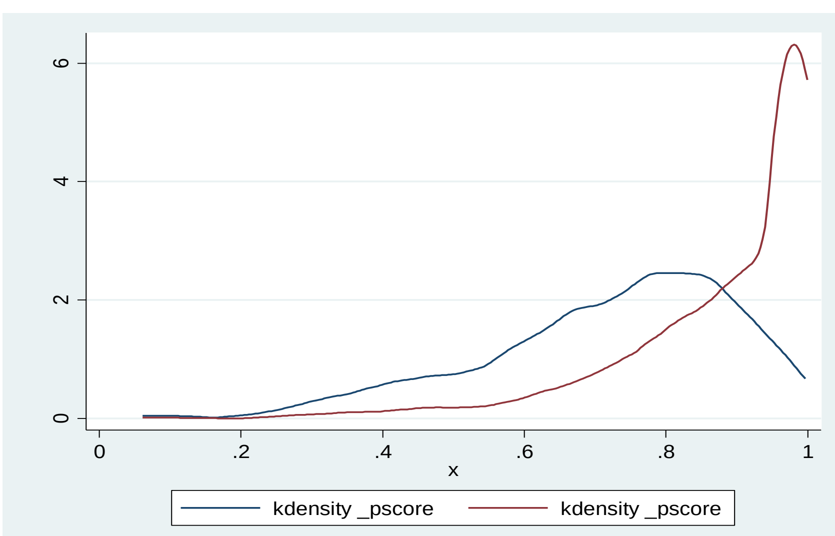
Paddy land * Hybrid seed	Unmatched	0.65	0.58	4.0		0.63	0.526
	Matched	0.67	0.58	5.9	-45.1	0.63	0.526
Paddy land * Improved seed	Unmatched	0.55	0.35	8.9		1.06	0.288
	Matched	0.30	0.35	-2.2	75.6	-0.43	0.670
Paddy land * Old local seed	Unmatched	0.51	0.98	-22.2		-3.11	0.002
	Matched	0.65	0.79	-6.7	70	-0.87	0.382

Summary of propensity score after matching

	Percentiles		
1%	0.4929801		
5%	0.8505756		
10%	1.353588	Obs	23
25%	2.160817	Sum of Wgt.	23
50%	4.699869	Mean	5.250992
		Std. Dev.	3.362003
75%	7.639369		
90%	10.17851	Variance	11.30307
95%	11.11758	Skewness	0.3661733
99%	11.27832	Kurtosis	2.018311

Sample	Pseudo R2	p>chi2	Mean Bias	Med. Bias
Raw	0.239	0	29.3	28
Matched	0.012	0.999	5.3	4.7

Kernel distribution of propensity score before (red) and after (blue) matching



Panel D: Kernel-based matching technique							
Variable	Unmatched/ Matched	Mean		% Bias	% Reduct bias	t-test	
		Treated	Control			t	p>|t|
HH size	Unmatched	4.83	5.57	-36.6		-5.83	0.000
	Matched	4.92	5.24	-15.7	57.2	-4.74	0.000
No. of member <10	Unmatched	0.71	1.14	-40.5		-6.65	0.000
	Matched	0.75	0.89	-13.4	66.8	-4.31	0.000
No. of members 10-15	Unmatched	0.54	0.83	-34.9		-5.76	0.000
	Matched	0.57	0.60	-2.9	91.6	-0.97	0.331
No. of members 15-25	Unmatched	1.14	1.22	-6.7		-1.08	0.280
	Matched	1.17	1.25	-6.6	0.3	-1.89	0.059
No. of members 25-60	Unmatched	2.04	2.08	-5.4		-0.85	0.394
	Matched	2.05	2.15	-11.2	-106.1	-3.10	0.002
Education of head	Unmatched	5.78	3.48	60.0		9.24	0.000
	Matched	5.38	5.08	7.8	86.9	2.41	0.016
Education of head (squared)	Unmatched	47.82	27.06	50.0		7.35	0.000
	Matched	42.82	40.08	6.6	86.8	2.04	0.042
Edu. of head * Ethnic	Unmatched	1.90	3.09	-33.7		-5.48	0.000
	Matched	2.10	2.36	-7.4	77.9	-2.36	0.019
Gender of head	Unmatched	0.85	0.92	-22.4		-3.12	0.002
	Matched	0.86	0.88	-8.9	60.3	-2.53	0.011
Age of head	Unmatched	49.59	44.36	40.7		6.16	0.000
	Matched	49.03	49.79	-5.9	85.5	-1.70	0.090
No. of labors for salary	Unmatched	0.98	0.70	28.0		4.04	0.000
	Matched	0.98	0.91	6.4	77.2	1.86	0.063
No. of labors in HH agriculture	Unmatched	3.25	3.53	-17.6		-2.72	0.007
	Matched	3.28	3.65	-23.2	-32.1	-6.45	0.000
No. of labors in HH business	Unmatched	0.39	0.49	-11.4		-1.84	0.066
	Matched	0.41	0.42	-0.7	94.0	-0.23	0.821
Hired labor outside	Unmatched	0.45	0.22	51.1		7.32	0.000
	Matched	0.41	0.32	19.4	62.1	5.55	0.000
Loan (ln)	Unmatched	4.79	5.04	-5.3		-0.80	0.424
	Matched	4.86	4.15	15.4	-193.4	4.61	0.000
Assistance receipt (ln)	Unmatched	2.95	0.80	68.6		8.93	0.000
	Matched	2.47	2.34	4.2	93.9	1.11	0.267
Distance to People's Committee	Unmatched	2.80	6.00	-71.4		-14.04	0.000
	Matched	2.93	3.11	-4.0	94.4	-1.56	0.118
Land independent on public irrigation	Unmatched	0.25	0.45	-43.5		-7.05	0.000
	Matched	0.27	0.29	-3.9	91	-1.23	0.221
Paddy land	Unmatched	1.78	1.94	-5.1		-0.65	0.513
	Matched	1.79	1.60	6.1	-19.9	1.83	0.067
Paddy land * Problematic land	Unmatched	0.85	0.98	-5.8		-0.79	0.428
	Matched	0.86	0.80	2.7	53.8	0.81	0.416

Paddy land * Hybrid seed	Unmatched	0.65	0.58	4.0		0.63	0.526
	Matched	0.64	0.68	-2.6	36.6	-0.67	0.502
Paddy land * Improved seed	Unmatched	0.55	0.35	8.9		1.06	0.288
	Matched	0.53	0.37	7.2	18.2	2.22	0.027
Paddy land * Old local seed	Unmatched	0.51	0.98	-22.2		-3.11	0.002
	Matched	0.55	0.52	1.5	93.1	0.52	0.606

Summary of propensity score after matching

	Percentiles		
1%	0.6920556		
5%	1.54248		
10%	2.566739	Obs	23
25%	3.899611	Sum of Wgt.	23
50%	6.620948	Mean	7.995325
		Std. Dev.	5.862708
75%	11.17304		
90%	15.65516	Variance	34.37134
95%	19.3929	Skewness	1.094392
99%	23.2113	Kurtosis	3.443817

Sample	Pseudo R2	p>chi2	Mean Bias	Med. Bias
Raw	0.239	0	29.3	28
Matched	0.03	0	8	6.6

Kernel distribution of propensity score before (red) and after (blue) matching

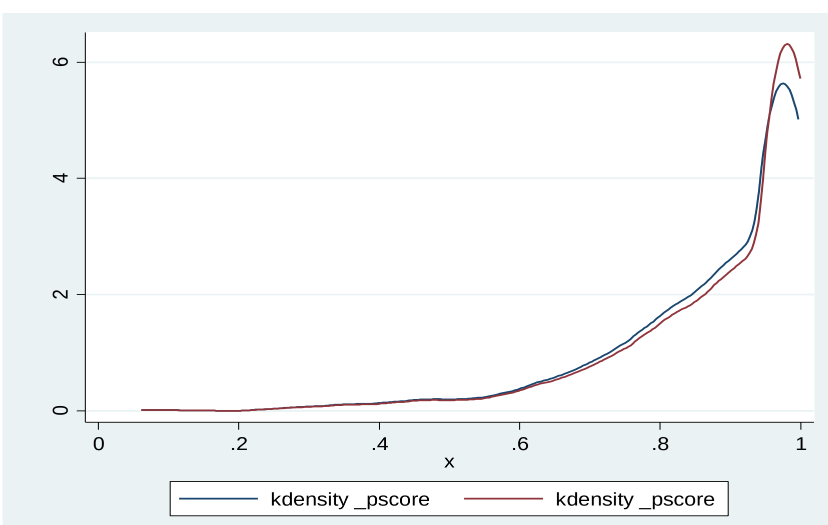


Table 3.12: Propensity score distribution before and after matching of households without disasters

Summary of propensity score before matching

	Percentiles		
1%	0.6761877		
5%	1.023039		
10%	2.20081	Obs	23
25%	10.98242	Sum of Wgt.	23
50%	26.81522	Mean	30.33446
		Std. Dev.	22.17862
75%	46.95355		
90%	62.04927	Variance	491.8913
95%	67.79373	Skewness	0.326898
99%	73.49481	Kurtosis	1.990973

Panel A: Nearest-neighbor technique							
Variable	Unmatched/ Matched	Mean		% Bias	% Reduct bias	t-test	
		Treated	Control			t	p>|t|
HH size	Unmatched	4.83	5.62	-38.3		-5.37	0.000
	Matched	4.97	5.11	-6.7	82.4	-1.62	0.106
No. of member <10	Unmatched	0.73	1.14	-38.2		-5.58	0.000
	Matched	0.79	0.76	2.1	94.4	0.55	0.583
No. of members 10-15	Unmatched	0.52	0.85	-38.7		-5.58	0.000
	Matched	0.57	0.62	-6.5	83.3	-1.71	0.087
No. of members 15-25	Unmatched	1.10	1.23	-11.0		-1.59	0.112
	Matched	1.15	1.23	-6.1	44.1	-1.43	0.153
No. of members 25-60	Unmatched	2.03	2.11	-9.0		-1.25	0.212
	Matched	2.06	2.19	-15.2	-68.6	-3.42	0.001
Education of head	Unmatched	5.76	3.37	62.0		8.42	0.000
	Matched	5.16	5.42	-6.8	89.0	-1.70	0.089
Education of head (squared)	Unmatched	47.86	26.41	50.9		6.60	0.000
	Matched	40.56	43.32	-6.5	87.1	-1.66	0.097
Edu. of head * Ethnic	Unmatched	1.62	2.94	-38.8		-5.68	0.000
	Matched	1.88	2.09	-6.2	83.9	-1.62	0.106
Gender of head	Unmatched	0.85	0.93	-26.4		-3.20	0.001
	Matched	0.86	0.90	-13.6	48.5	-3.18	0.001
Age of head	Unmatched	49.94	44.51	42.0		5.51	0.000
	Matched	49.25	49.95	-5.4	87.1	-1.27	0.203
No. of labors for salary	Unmatched	0.96	0.69	26.8		3.44	0.001
	Matched	0.96	0.80	15.3	43.0	3.75	0.000
No. of labors in HH agriculture	Unmatched	3.21	3.47	-16.4		-2.24	0.026
	Matched	3.26	3.52	-16.4	-0.2	-3.52	0.000
No. of labors in HH business	Unmatched	0.42	0.47	-6.3		-0.89	0.376
	Matched	0.44	0.48	-5.1	17.9	-1.32	0.185
Hired labor outside	Unmatched	0.43	0.19	54.4		6.77	0.000
	Matched	0.37	0.33	9.2	83.0	2.08	0.037
Loan (ln)	Unmatched	4.40	4.95	-11.8		-1.59	0.113
	Matched	4.54	3.97	12.2	-2.9	2.93	0.003
Assistance receipt (ln)	Unmatched	3.06	0.88	67.8		7.86	0.000
	Matched	2.45	2.70	-7.8	88.4	-1.66	0.097
Distance to People's Committee	Unmatched	2.64	5.84	-73.5		-12.99	0.000
	Matched	2.84	2.96	-2.8	96.2	-0.88	0.377
Land independent on public irrigation	Unmatched	0.23	0.45	-47.0		-6.76	0.000
	Matched	0.25	0.30	-9.8	79.1	-2.45	0.014
Paddy land	Unmatched	1.90	1.93	-1.0		-0.11	0.910
	Matched	1.92	1.52	12.1	-1083.5	2.97	0.003
Paddy land * Problematic land	Unmatched	0.99	1.04	-2.2		-0.25	0.799
	Matched	1.03	0.85	7.5	-239.7	1.80	0.073

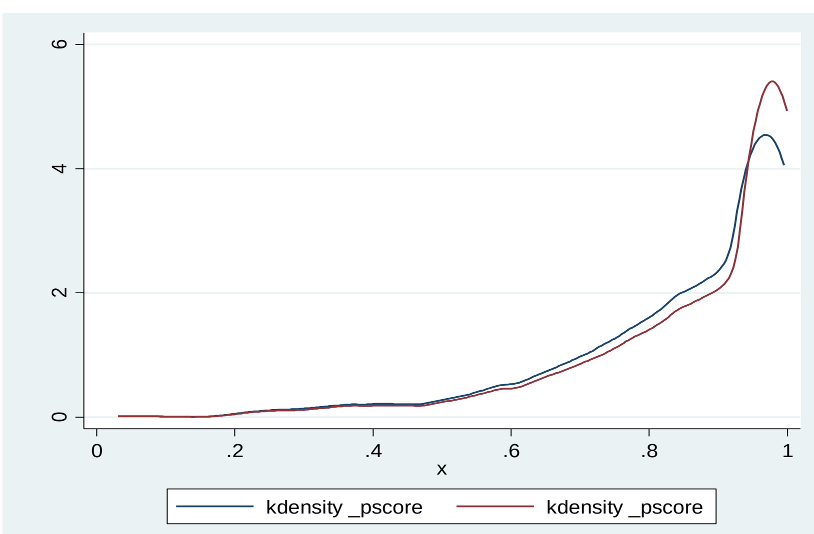
Paddy land * Hybrid seed	Unmatched	0.64	0.63	0.7		0.10	0.924
	Matched	0.64	0.46	10.6	-1472.8	2.70	0.007
Paddy land * Improved seed	Unmatched	0.61	0.29	13.6		1.44	0.150
	Matched	0.56	0.50	2.6	81.2	0.65	0.519
Paddy land * Old local seed	Unmatched	0.55	0.97	-20.9		-2.43	0.015
	Matched	0.62	0.53	4.8	77	1.17	0.242

Summary of propensity score after matching

	Percentiles		
1%	2.124431		
5%	2.564249		
10%	2.811932	Obs	23
25%	5.425393	Sum of Wgt.	23
50%	6.801488	Mean	8.33086
		Std. Dev.	4.162977
75%	12.10734		
90%	15.23701	Variance	17.33038
95%	15.28459	Skewness	0.467389
99%	16.40006	Kurtosis	2.216279

Sample	Pseudo R2	p>chi2	Mean Bias	Med. Bias
Raw	0.266	0	30.3	26.8
Matched	0.041	0	8.3	6.8

Kernel distribution of propensity score before (red) and after (blue) matching



Panel B: Nearest-neighbor without replacement technique							
Variable	Unmatched/ Matched	Mean		% Bias	% Reduct bias	t-test	
		Treated	Control			t	p>|t|
HH size	Unmatched	4.83	5.62	-38.3		-5.37	0.000
	Matched	5.51	5.62	-5.3	86.2	-0.51	0.613
No. of member <10	Unmatched	0.73	1.14	-38.2		-5.58	0.000
	Matched	1.20	1.14	5.3	86.3	0.51	0.612
No. of members 10-15	Unmatched	0.52	0.85	-38.7		-5.58	0.000
	Matched	0.85	0.85	0.6	98.5	0.05	0.957
No. of members 15-25	Unmatched	1.10	1.23	-11.0		-1.59	0.112
	Matched	1.18	1.23	-4.2	62.0	-0.40	0.686
No. of members 25-60	Unmatched	2.03	2.11	-9.0		-1.25	0.212
	Matched	2.08	2.11	-3.7	58.9	-0.41	0.680
Education of head	Unmatched	5.76	3.37	62.0		8.42	0.000
	Matched	3.37	3.37	0.0	100.0	0.00	1.000
Education of head (squared)	Unmatched	47.86	26.41	50.9		6.60	0.000
	Matched	28.04	26.41	3.9	92.4	0.40	0.692
Edu. of head * Ethnic	Unmatched	1.62	2.94	-38.8		-5.68	0.000
	Matched	3.30	2.94	10.4	73.1	0.93	0.354
Gender of head	Unmatched	0.85	0.93	-26.4		-3.20	0.001
	Matched	0.94	0.93	3.0	88.5	0.39	0.696
Age of head	Unmatched	49.94	44.51	42.0		5.51	0.000
	Matched	42.78	44.51	-13.3	68.3	-1.45	0.147
No. of labors for salary	Unmatched	0.96	0.69	26.8		3.44	0.001
	Matched	0.60	0.69	-8.0	70.2	-0.93	0.355
No. of labors in HH agriculture	Unmatched	3.21	3.47	-16.4		-2.24	0.026
	Matched	3.34	3.47	-7.8	52.1	-0.81	0.417
No. of labors in HH business	Unmatched	0.42	0.47	-6.3		-0.89	0.376
	Matched	0.48	0.47	1.1	82.8	0.11	0.914
Hired labor outside	Unmatched	0.43	0.19	54.4		6.77	0.000
	Matched	0.16	0.19	-6.4	88.2	-0.78	0.438
Loan (ln)	Unmatched	4.40	4.95	-11.8		-1.59	0.113
	Matched	5.31	4.95	7.5	36.5	0.78	0.436
Assistance receipt (ln)	Unmatched	3.06	0.88	67.8		7.86	0.000
	Matched	0.51	0.88	-11.3	83.3	-1.82	0.070
Distance to People's Committee	Unmatched	2.64	5.84	-73.5		-12.99	0.000
	Matched	5.74	5.84	-2.4	96.7	-0.21	0.834
Land independent on public irrigation	Unmatched	0.23	0.45	-47.0		-6.76	0.000
	Matched	0.47	0.45	5.2	89	0.49	0.625
Paddy land	Unmatched	1.90	1.93	-1.0		-0.11	0.910
	Matched	1.71	1.93	-6.9	-576.3	-0.91	0.365
Paddy land * Problematic land	Unmatched	0.99	1.04	-2.2		-0.25	0.799
	Matched	0.94	1.04	-4.4	-101.5	-0.44	0.662

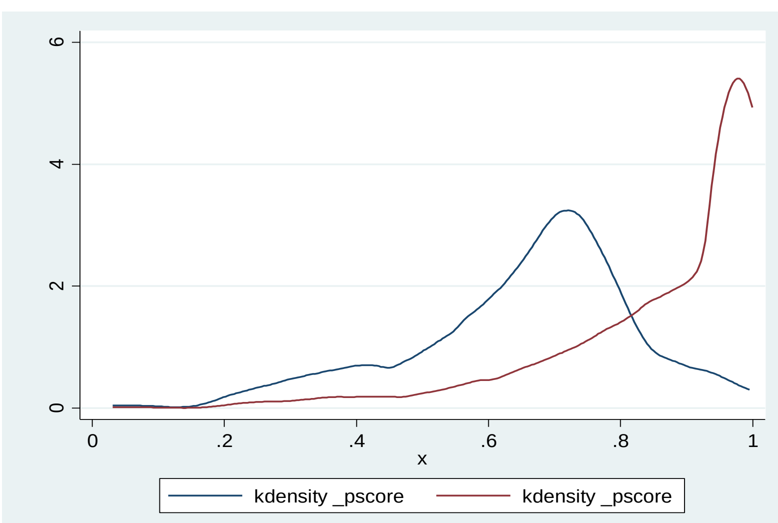
Paddy land * Hybrid seed	Unmatched	0.64	0.63	0.7		0.10	0.924
	Matched	0.72	0.63	4.9	-628.4	0.44	0.663
Paddy land * Improved seed	Unmatched	0.61	0.29	13.6		1.44	0.150
	Matched	0.19	0.29	-4.3	68.3	-1.41	0.160
Paddy land * Old local seed	Unmatched	0.55	0.97	-20.9		-2.43	0.015
	Matched	0.78	0.97	-9.7	53.8	-0.97	0.331

Summary of propensity score after matching

	Percentiles		
1%	0		
5%	0.5633319		
10%	1.079894	Obs	23
25%	3.710848	Sum of Wgt.	23
50%	5.158063	Mean	5.638401
		Std. Dev.	3.409725
75%	7.833782		
90%	10.434	Variance	11.62623
95%	11.33679	Skewness	0.438
99%	13.33279	Kurtosis	2.713205

Sample	Pseudo R2	p>chi2	Mean Bias	Med. Bias
Raw	0.266	0	30.3	26.8
Matched	0.036	0.576	5.6	5.2

Kernel distribution of propensity score before (red) and after (blue) matching



Panel C: Caliper without replacement technique							
Variable	Unmatched/ Matched	Mean		% Bias	% Reduct bias	t-test	
		Treated	Control			t	p>|t|
HH size	Unmatched	4.83	5.62	-38.3		-5.37	0.000
	Matched	5.34	5.43	-3.9	89.9	-0.36	0.716
No. of member <10	Unmatched	0.73	1.14	-38.2		-5.58	0.000
	Matched	1.05	1.07	-2.1	94.5	-0.19	0.850
No. of members 10-15	Unmatched	0.52	0.85	-38.7		-5.58	0.000
	Matched	0.78	0.76	1.4	96.5	0.12	0.903
No. of members 15-25	Unmatched	1.10	1.23	-11.0		-1.59	0.112
	Matched	1.19	1.20	-0.5	95.8	-0.04	0.967
No. of members 25-60	Unmatched	2.03	2.11	-9.0		-1.25	0.212
	Matched	2.05	2.09	-5.1	43.4	-0.50	0.614
Education of head	Unmatched	5.76	3.37	62.0		8.42	0.000
	Matched	3.18	3.69	-13.2	78.7	-1.21	0.226
Education of head (squared)	Unmatched	47.86	26.41	50.9		6.60	0.000
	Matched	25.88	28.60	-6.4	87.3	-0.62	0.539
Edu. of head * Ethnic	Unmatched	1.62	2.94	-38.8		-5.68	0.000
	Matched	2.75	3.18	-12.4	67.9	-1.03	0.305
Gender of head	Unmatched	0.85	0.93	-26.4		-3.20	0.001
	Matched	0.90	0.91	-3.7	86.1	-0.37	0.712
Age of head	Unmatched	49.94	44.51	42.0		5.51	0.000
	Matched	46.03	45.76	2.0	95.1	0.19	0.851
No. of labors for salary	Unmatched	0.96	0.69	26.8		3.44	0.001
	Matched	0.66	0.75	-9.6	64.0	-0.96	0.338
No. of labors in HH agriculture	Unmatched	3.21	3.47	-16.4		-2.24	0.026
	Matched	3.47	3.48	-0.4	97.8	-0.03	0.973
No. of labors in HH business	Unmatched	0.42	0.47	-6.3		-0.89	0.376
	Matched	0.43	0.43	0.0	100.0	0.00	1.000
Hired labor outside	Unmatched	0.43	0.19	54.4		6.77	0.000
	Matched	0.16	0.20	-10.3	81.0	-1.12	0.264
Loan (ln)	Unmatched	4.40	4.95	-11.8		-1.59	0.113
	Matched	4.39	4.74	-7.4	37.7	-0.69	0.490
Assistance receipt (ln)	Unmatched	3.06	0.88	67.8		7.86	0.000
	Matched	1.10	1.01	2.7	96.1	0.32	0.752
Distance to People's Committee	Unmatched	2.64	5.84	-73.5		-12.99	0.000
	Matched	4.82	4.66	3.8	94.8	0.34	0.734
Land independent on public irrigation	Unmatched	0.23	0.45	-47.0		-6.76	0.000
	Matched	0.42	0.38	8.7	81.4	0.76	0.445
Paddy land	Unmatched	1.90	1.93	-1.0		-0.11	0.910
	Matched	1.42	1.79	-11.1	-984.6	-1.64	0.101
Paddy land * Problematic land	Unmatched	0.99	1.04	-2.2		-0.25	0.799
	Matched	0.68	0.86	-7.5	-241.2	-0.89	0.375
Paddy land * Hybrid seed	Unmatched	0.64	0.63	0.7		0.10	0.924

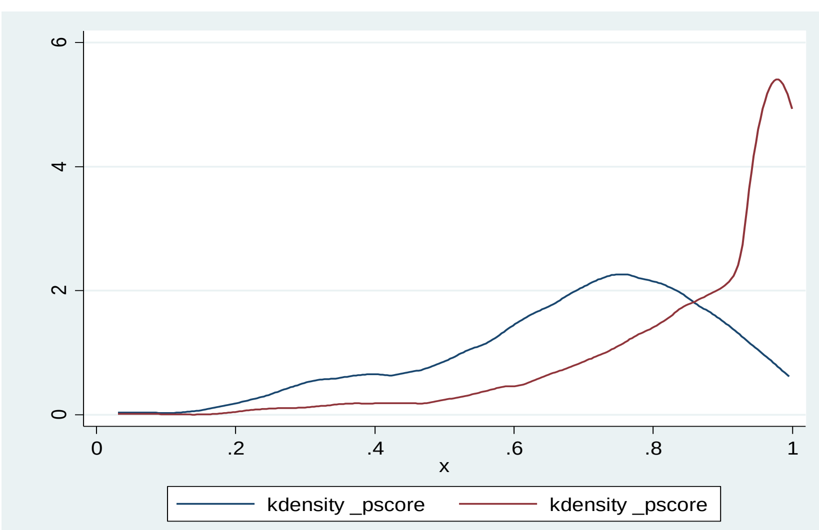
	Matched	0.68	0.66	1.2	-82.6	0.10	0.918
Paddy land * Improved seed	Unmatched	0.61	0.29	13.6		1.44	0.150
	Matched	0.26	0.27	-0.4	96.7	-0.13	0.900
Paddy land * Old local seed	Unmatched	0.55	0.97	-20.9		-2.43	0.015
	Matched	0.44	0.82	-18.8	10.3	-2.77	0.006

Summary of propensity score after matching

	Percentiles		
1%	0		
5%	0.3636371		
10%	0.4459245	Obs	23
25%	1.359767	Sum of Wgt.	23
50%	3.881846	Mean	5.771028
		Std. Dev.	5.029802
75%	9.644718		
90%	12.42486	Variance	25.29891
95%	13.2471	Skewness	0.829644
99%	18.78729	Kurtosis	2.985793

Sample	Pseudo R2	p>chi2	Mean Bias	Med. Bias
Raw	0.266	0	30.3	26.8
Matched	0.036	0.79	5.8	3.9

Kernel distribution of propensity score before (red) and after (blue) matching



Panel D: Kernel-based matching technique							
Variable	Unmatched/ Matched	Mean		% Bias	% Reduct bias	t-test	
		Treated	Control			t	p>|t|
HH size	Unmatched	4.83	5.62	-38.3		-5.37	0.000
	Matched	4.97	5.33	-17.0	55.6	-4.11	0.000
No. of member <10	Unmatched	0.73	1.14	-38.2		-5.58	0.000
	Matched	0.79	0.91	-11.7	69.4	-3.01	0.003
No. of members 10-15	Unmatched	0.52	0.85	-38.7		-5.58	0.000
	Matched	0.57	0.52	5.2	86.5	1.40	0.161
No. of members 15-25	Unmatched	1.10	1.23	-11.0		-1.59	0.112
	Matched	1.15	1.32	-13.4	-21.6	-3.06	0.002
No. of members 25-60	Unmatched	2.03	2.11	-9.0		-1.25	0.212
	Matched	2.06	2.22	-18.3	-102.1	-4.01	0.000
Education of head	Unmatched	5.76	3.37	62.0		8.42	0.000
	Matched	5.16	5.21	-1.4	97.8	-0.33	0.739
Education of head (squared)	Unmatched	47.86	26.41	50.9		6.60	0.000
	Matched	40.56	42.15	-3.8	92.6	-0.92	0.355
Edu. of head * Ethnic	Unmatched	1.62	2.94	-38.8		-5.68	0.000
	Matched	1.88	2.30	-12.2	68.5	-3.08	0.002
Gender of head	Unmatched	0.85	0.93	-26.4		-3.20	0.001
	Matched	0.86	0.90	-14.7	44.5	-3.45	0.001
Age of head	Unmatched	49.94	44.51	42.0		5.51	0.000
	Matched	49.25	50.33	-8.4	80.0	-1.93	0.053
No. of labors for salary	Unmatched	0.96	0.69	26.8		3.44	0.001
	Matched	0.96	0.86	9.3	65.5	2.18	0.029
No. of labors in HH agriculture	Unmatched	3.21	3.47	-16.4		-2.24	0.026
	Matched	3.26	3.63	-23.7	-45.0	-5.13	0.000
No. of labors in HH business	Unmatched	0.42	0.47	-6.3		-0.89	0.376
	Matched	0.44	0.47	-3.8	38.7	-1.00	0.318
Hired labor outside	Unmatched	0.43	0.19	54.4		6.77	0.000
	Matched	0.37	0.29	18.7	65.7	4.30	0.000
Loan (ln)	Unmatched	4.40	4.95	-11.8		-1.59	0.113
	Matched	4.54	4.05	10.4	12.3	2.49	0.013
Assistance receipt (ln)	Unmatched	3.06	0.88	67.8		7.86	0.000
	Matched	2.45	2.30	4.7	93.1	1.04	0.299
Distance to People's Committee	Unmatched	2.64	5.84	-73.5		-12.99	0.000
	Matched	2.84	2.93	-2.3	96.9	-0.71	0.475
Land independent on public irrigation	Unmatched	0.23	0.45	-47.0		-6.76	0.000
	Matched	0.25	0.29	-7.1	84.9	-1.78	0.075
Paddy land	Unmatched	1.90	1.93	-1.0		-0.11	0.910
	Matched	1.92	1.61	9.3	-813.8	2.22	0.027
Paddy land * Problematic land	Unmatched	0.99	1.04	-2.2		-0.25	0.799

	Matched	1.03	0.82	8.7	-293.9	2.01	0.044
Paddy land * Hybrid seed	Unmatched	0.64	0.63	0.7		0.10	0.924
	Matched	0.64	0.66	-1.4	-109.3	-0.31	0.754
Paddy land * Improved seed	Unmatched	0.61	0.29	13.6		1.44	0.150
	Matched	0.56	0.38	7.8	42.9	1.96	0.050
Paddy land * Old local seed	Unmatched	0.55	0.97	-20.9		-2.43	0.015
	Matched	0.62	0.53	4.5	78.5	1.10	0.273

Summary of propensity score after matching

	Percentiles		
1%	1.359805		
5%	1.414989		
10%	2.272079	Obs	23
25%	4.499206	Sum of Wgt.	23
50%	8.669534	Mean	9.460793
		Std. Dev.	6.029412
75%	13.35732		
90%	18.27087	Variance	36.35381
95%	18.66849	Skewness	0.6170245
99%	23.71946	Kurtosis	2.643281

Sample	Pseudo R2	p>chi2	Mean Bias	Med. Bias
Raw	0.266	0	30.3	26.8
Matched	0.04	0	9.5	8.7

Kernel distribution of propensity score before (red) and after (blue) matching

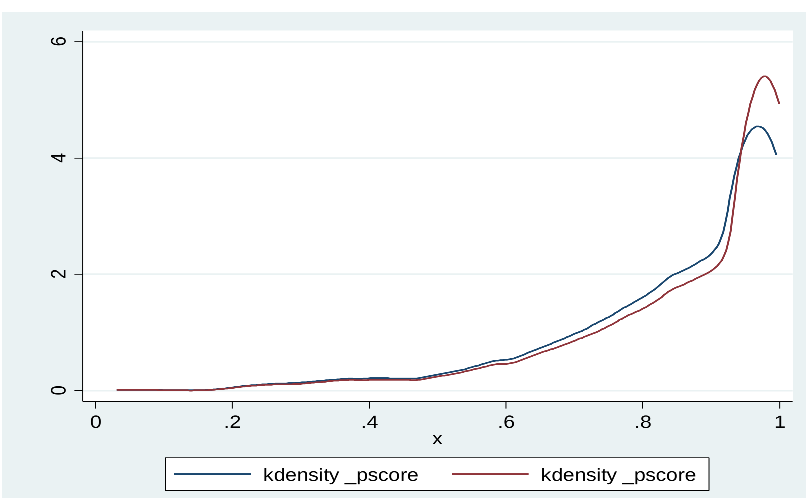


Table 3.13: Propensity score distribution before and after matching of households suffered from disasters

Summary of propensity score before matching

	Percentiles		
1%	5.145177		
5%	6.233551		
10%	8.628268	Obs	23
25%	10.37722	Sum of Wgt.	23
50%	27.12029	Mean	29.59041
		Std. Dev.	19.44298
75%	39.09868		
90%	50.6022	Variance	378.0293
95%	72.1411	Skewness	0.8914386
99%	77.0631	Kurtosis	3.3444

Panel A: Nearest-neighbor technique							
Variable	Unmatched/ Matched	Mean		% Bias	% Reduct bias	t-test	
		Treated	Control			t	p>|t|
HH size	Unmatched	4.84	5.36	-29.1		-2.09	0.037
	Matched	5.04	4.87	9.9	66.0	1.60	0.109
No. of member <10	Unmatched	0.67	1.13	-46.5		-3.40	0.001
	Matched	0.78	0.95	-17.9	61.5	-2.72	0.007
No. of members 10-15	Unmatched	0.57	0.76	-23.2		-1.820	0.070
	Matched	0.66	0.69	-3.5	85.1	-0.59	0.557
No. of members 15-25	Unmatched	1.22	1.16	5.1		0.37	0.714
	Matched	1.23	0.91	27.3	-431.4	4.46	0.000
No. of members 25-60	Unmatched	2.04	1.96	10.4		0.75	0.456
	Matched	2.03	2.04	-1.4	86.8	-0.25	0.802
Education of head	Unmatched	5.82	3.91	50.6		3.63	0.000
	Matched	4.88	4.87	0.4	99.2	0.06	0.950
Education of head (squared)	Unmatched	47.73	29.55	45.8		3.10	0.002
	Matched	36.43	37.17	-1.8	96.0	-0.32	0.751
Edu. of head * Ethnic	Unmatched	2.52	3.65	-30.7		-2.25	0.024
	Matched	3.24	2.52	19.7	35.9	3.25	0.001
Gender of head	Unmatched	0.86	0.89	-9.8		-0.67	0.503
	Matched	0.85	0.90	-15.0	-53.2	-2.38	0.017
Age of head	Unmatched	48.84	43.80	39.1		2.93	0.003
	Matched	47.87	44.90	23.0	41.1	3.80	0.000
No. of labors for salary	Unmatched	1.04	0.76	27.8		1.84	0.065
	Matched	0.99	0.70	29.3	-5.2	4.42	0.000
No. of labors in HH agriculture	Unmatched	3.35	3.78	-27.1		-1.94	0.053
	Matched	3.55	3.75	-12.9	52.3	-2.14	0.032
No. of labors in HH business	Unmatched	0.34	0.56	-27.1		-2.02	0.044
	Matched	0.36	0.29	7.7	71.5	1.29	0.198
Hired labor outside	Unmatched	0.49	0.33	34.0		2.36	0.018
	Matched	0.47	0.38	18.0	47.2	2.75	0.006
Loan (ln)	Unmatched	5.63	5.35	6.2		0.44	0.658
	Matched	5.41	3.84	35.0	-461.4	5.43	0.000
Assistance receipt (ln)	Unmatched	2.72	0.49	77.1		4.46	0.000
	Matched	1.75	1.25	17.6	77.2	2.79	0.005
Distance to People's Committee	Unmatched	3.14	6.63	-72.1		-6.48	0.000
	Matched	3.51	4.34	-17.2	76.2	-3.58	0.000
Land independent on public irrigation	Unmatched	0.29	0.47	-38.2		-2.84	0.005
	Matched	0.32	0.25	14.3	62.4	2.36	0.019
Paddy land	Unmatched	1.53	1.98	-14.9		-0.97	0.331
	Matched	1.50	1.48	0.6	95.9	0.11	0.914
Paddy land * Problematic land	Unmatched	0.54	0.74	-9.8		-0.76	0.446
	Matched	0.50	0.70	-10.3	-5.2	-2.16	0.031

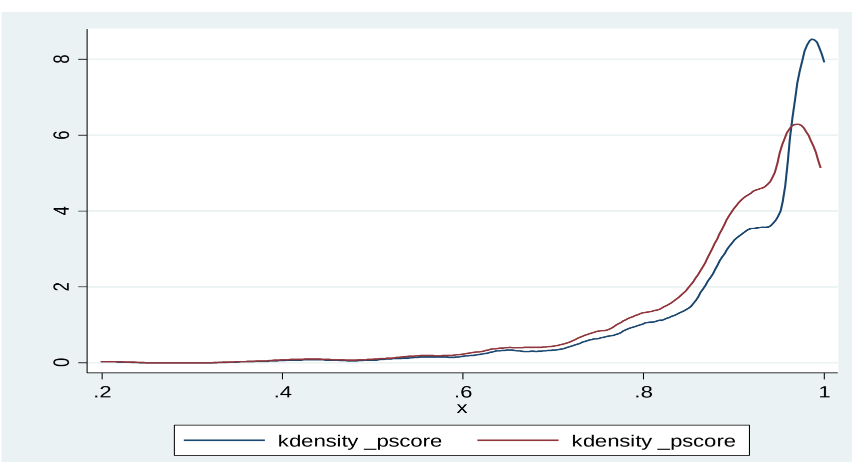
Paddy land * Hybrid seed	Unmatched	0.65	0.39	24.1		1.40	0.161
	Matched	0.52	0.91	-36.1	-49.8	-7.29	0.000
Paddy land * Improved seed	Unmatched	0.41	0.58	-8.6		-0.49	0.621
	Matched	0.48	0.19	14.5	-68	2.12	0.034
Paddy land * Old local seed	Unmatched	0.42	1.01	-23.2		-1.84	0.066
	Matched	0.51	0.38	5.0	78.2	1.02	0.308

Summary of propensity score after matching

	Percentiles		
1%	0.3867186		
5%	0.614317		
10%	1.370329	Obs	23
25%	5.046937	Sum of Wgt.	23
50%	14.49821	Mean	14.71951
		Std. Dev.	10.5128
75%	19.68475		
90%	29.27452	Variance	110.519
95%	34.99508	Skewness	0.4343568
99%	36.09293	Kurtosis	2.422963

Sample	Pseudo R2	p>chi2	Mean Bias	Med. Bias
Raw	0.229	0	29.6	27.1
Matched	0.135	0	14.7	14.5

Kernel distribution of propensity score before (red) and after (blue) matching



Panel B: Nearest-neighbor without replacement technique							
Variable	Unmatched/ Matched	Mean		% Bias	% Reduct bias	t-test	
		Treated	Control			t	p>|t|
HH size	Unmatched	4.84	5.36	-29.1		-2.09	0.037
	Matched	5.73	5.36	20.3	30.2	0.95	0.346
No. of member <10	Unmatched	0.67	1.13	-46.5		-3.4	0.001
	Matched	1.56	1.13	44.2	5	2.03	0.045
No. of members 10-15	Unmatched	0.57	0.76	-23.2		-1.82	0.070
	Matched	0.75	0.76	-2.2	90.7	-0.11	0.912
No. of members 15-25	Unmatched	1.22	1.16	5.1		0.37	0.714
	Matched	1.11	1.16	-4.7	9.1	-0.24	0.810
No. of members 25-60	Unmatched	2.04	1.96	10.4		0.75	0.456
	Matched	1.91	1.96	-7.2	30.9	-0.4	0.693
Education of head	Unmatched	5.82	3.91	50.6		3.63	0.000
	Matched	3.65	3.91	-6.7	86.7	-0.35	0.726
Education of head (squared)	Unmatched	47.73	29.55	45.8		3.1	0.002
	Matched	27.40	29.55	-5.4	88.2	-0.31	0.757
Edu. of head * Ethnic	Unmatched	2.52	3.65	-30.7		-2.25	0.024
	Matched	3.44	3.65	-5.9	80.7	-0.3	0.764
Gender of head	Unmatched	0.86	0.89	-9.8		-0.67	0.503
	Matched	0.89	0.89	0	100	0	1.000
Age of head	Unmatched	48.84	43.80	39.1		2.93	0.003
	Matched	40.29	43.80	-27.2	30.4	-1.4	0.164
No. of labors for salary	Unmatched	1.04	0.76	27.8		1.84	0.065
	Matched	0.45	0.76	-31	-11.5	-1.91	0.058
No. of labors in HH agriculture	Unmatched	3.35	3.78	-27.1		-1.94	0.053
	Matched	3.73	3.78	-3.4	87.4	-0.19	0.850
No. of labors in HH business	Unmatched	0.34	0.56	-27.1		-2.02	0.044
	Matched	0.84	0.56	32.6	-20.3	1.12	0.264
Hired labor outside	Unmatched	0.49	0.33	34		2.36	0.018
	Matched	0.22	0.33	-22.4	34.1	-1.28	0.202
Loan (ln)	Unmatched	5.63	5.35	6.2		0.44	0.658
	Matched	5.00	5.35	-7.9	-26.9	-0.41	0.685
Assistance receipt (ln)	Unmatched	2.72	0.49	77.1		4.46	0.000
	Matched	0.00	0.49	-16.8	78.2	-2	0.048
Distance to People's Committee	Unmatched	3.14	6.63	-72.1		-6.48	0.000
	Matched	8.13	6.63	31	57	1.23	0.220
Land independent on public irrigation	Unmatched	0.29	0.47	-38.2		-2.84	0.005
	Matched	0.51	0.47	7.6	80.1	0.38	0.706
Paddy land	Unmatched	1.53	1.98	-14.9		-0.97	0.331
	Matched	1.75	1.98	-7.6	48.8	-0.28	0.777
Paddy land * Problematic land	Unmatched	0.54	0.74	-9.8		-0.76	0.446
	Matched	0.94	0.74	10.5	-6.5	0.28	0.781
Paddy land * Hybrid seed	Unmatched	0.65	0.39	24.1		1.4	0.161

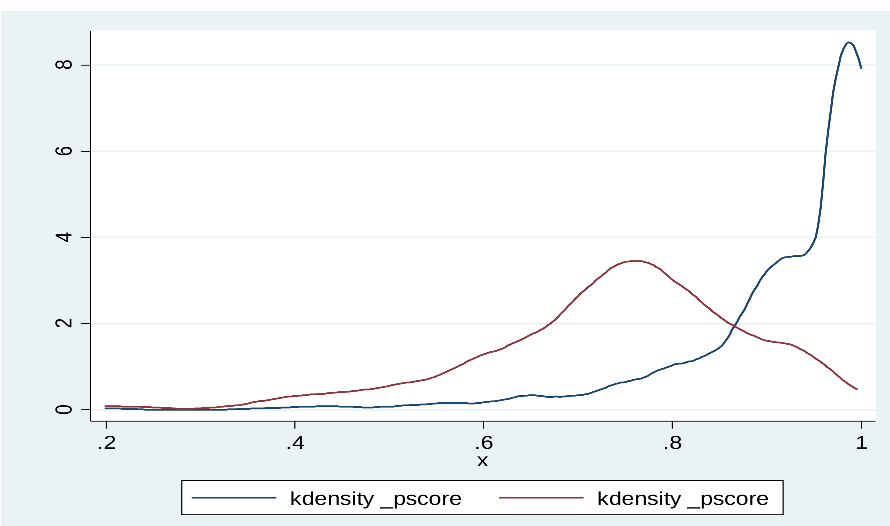
	Matched	0.36	0.39	-2.6	89.3	-0.25	0.800
Paddy land * Improved seed	Unmatched	0.41	0.58	-8.6		-0.49	0.621
	Matched	0.13	0.58	-23.5	-172.1	-2.74	0.007
Paddy land * Old local seed	Unmatched	0.42	1.01	-23.2		-1.84	0.066
	Matched	1.26	1.01	9.9	57.4	0.3	0.766

Summary of propensity score after matching

	Percentiles		
1%	0		
5%	2.164152		
10%	2.583019	Obs	23
25%	5.402595	Sum of Wgt.	23
50%	7.911068	Mean	14.37431
		Std. Dev.	12.19031
75%	23.47744		
90%	31.04984	Variance	148.6037
95%	32.63686	Skewness	0.8609254
99%	44.16815	Kurtosis	2.635159

Sample	Pseudo R2	p>chi2	Mean Bias	Med. Bias
Raw	0.229	0	29.6	27.1
Matched	0.321	0.001	14.4	7.9

Kernel distribution of propensity score before (red) and after (blue) matching



Panel C: Caliper without replacement technique							
Variable	Unmatched/ Matched	Mean		% Bias	% Reduct bias	t-test	
		Treated	Control			t	p>|t|
HH size	Unmatched	4.84	5.36	-29.1		-2.09	0.037
	Matched	4.81	5.40	-33.2	-14.2	-1.57	0.120
No. of member <10	Unmatched	0.67	1.13	-46.5		-3.40	0.001
	Matched	0.98	1.10	-12.0	74.1	-0.53	0.597
No. of members 10-15	Unmatched	0.57	0.76	-23.2		-1.82	0.070
	Matched	0.76	0.93	-19.8	14.6	-0.81	0.419
No. of members 15-25	Unmatched	1.22	1.16	5.1		0.37	0.714
	Matched	0.81	1.14	-28.6	-455.7	-1.43	0.157
No. of members 25-60	Unmatched	2.04	1.96	10.4		0.75	0.456
	Matched	2.00	1.93	9.4	9.5	0.48	0.635
Education of head	Unmatched	5.82	3.91	50.6		3.63	0.000
	Matched	4.17	3.83	8.8	82.5	0.40	0.691
Education of head (squared)	Unmatched	47.73	29.55	45.8		3.10	0.002
	Matched	31.64	29.02	6.6	85.6	0.33	0.741
Edu. of head * Ethnic	Unmatched	2.52	3.65	-30.7		-2.25	0.024
	Matched	3.64	3.50	3.9	87.4	0.17	0.866
Gender of head	Unmatched	0.86	0.89	-9.8		-0.67	0.503
	Matched	0.93	0.88	14.3	-45.9	0.74	0.463
Age of head	Unmatched	48.84	43.80	39.1		2.93	0.003
	Matched	41.52	44.36	-22.0	43.8	-1.17	0.244
No. of labors for salary	Unmatched	1.04	0.76	27.8		1.84	0.065
	Matched	0.64	0.79	-14.3	48.5	-0.70	0.483
No. of labors in HH agriculture	Unmatched	3.35	3.78	-27.1		-1.94	0.053
	Matched	3.45	3.88	-26.9	0.8	-1.25	0.215
No. of labors in HH business	Unmatched	0.34	0.56	-27.1		-2.02	0.044
	Matched	0.36	0.43	-8.5	68.5	-0.44	0.663
Hired labor outside	Unmatched	0.49	0.33	34.0		2.36	0.018
	Matched	0.24	0.36	-24.4	28.1	-1.19	0.238
Loan (ln)	Unmatched	5.63	5.35	6.2		0.44	0.658
	Matched	5.54	4.94	13.6	-117.9	0.60	0.548
Assistance receipt (ln)	Unmatched	2.72	0.49	77.1		4.46	0.000
	Matched	0.84	0.64	6.8	91.1	0.44	0.662
Distance to People's Committee	Unmatched	3.14	6.63	-72.1		-6.48	0.000
	Matched	6.41	4.56	38.3	46.9	1.57	0.120
Land independent on public irrigation	Unmatched	0.29	0.47	-38.2		-2.84	0.005
	Matched	0.31	0.45	-29.8	21.9	-1.35	0.182
Paddy land	Unmatched	1.53	1.98	-14.9		-0.97	0.331
	Matched	1.80	1.71	2.9	80.6	0.09	0.930
Paddy land * Problematic land	Unmatched	0.54	0.74	-9.8		-0.76	0.446
	Matched	1.25	0.42	42.2	-329.2	0.92	0.358

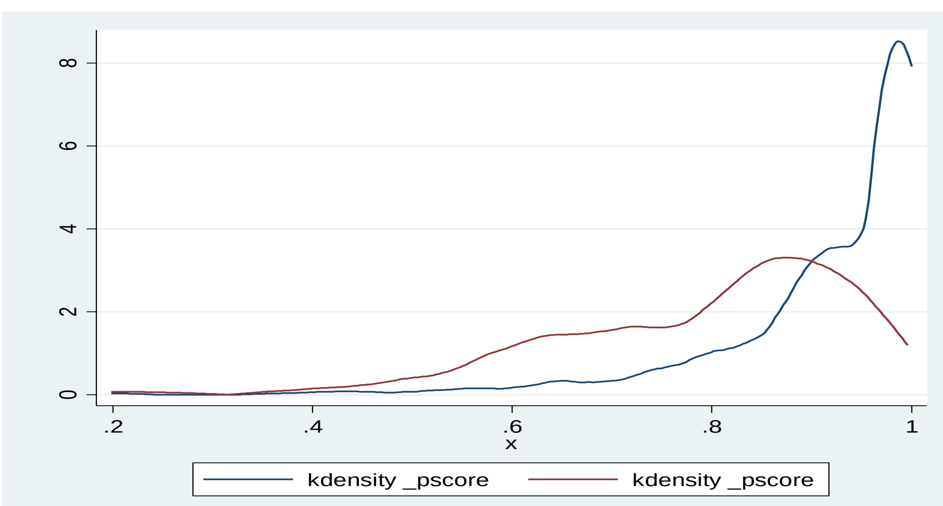
Paddy land * Hybrid seed	Unmatched	0.65	0.39	24.1		1.40	0.161
	Matched	0.40	0.42	-2.3	90.6	-0.19	0.849
Paddy land * Improved seed	Unmatched	0.41	0.58	-8.6		-0.49	0.621
	Matched	0.24	0.53	-15.1	-75.3	-1.45	0.151
Paddy land * Old local seed	Unmatched	0.42	1.01	-23.2		-1.84	0.066
	Matched	1.16	0.75	16.0	30.9	0.40	0.692

Summary of propensity score after matching

	Percentiles		
1%	2.274616		
5%	2.885342		
10%	3.878768	Obs	23
25%	8.547748	Sum of Wgt.	23
50%	14.33854	Mean	17.38485
		Std. Dev.	11.44456
75%	26.85716		
90%	33.18887	Variance	130.978
95%	38.29766	Skewness	0.6049059
99%	42.20342	Kurtosis	2.370706

Sample	Pseudo R2	p>chi2	Mean Bias	Med. Bias
Raw	0.229	0	29.6	27.1
Matched	0.252	0.134	17.4	14.3

Kernel distribution of propensity score before (red) and after (blue) matching



Panel D: Kernel-based matching technique							
Variable	Unmatched/ Matched	Mean		% Bias	% Reduct bias	t-test	
		Treated	Control			t	p>|t|
HH size	Unmatched	4.84	5.36	-29.1		-2.09	0.037
	Matched	5.04	4.93	5.9	79.8	0.92	0.356
No. of member <10	Unmatched	0.67	1.13	-46.5		-3.40	0.001
	Matched	0.77	0.85	-7.8	83.2	-1.20	0.232
No. of members 10-15	Unmatched	0.57	0.76	-23.2		-1.82	0.070
	Matched	0.66	0.85	-22.9	1.6	-3.78	0.000
No. of members 15-25	Unmatched	1.22	1.16	5.1		0.37	0.714
	Matched	1.22	0.91	26.9	-422.1	4.29	0.000
No. of members 25-60	Unmatched	2.04	1.96	10.4		0.75	0.456
	Matched	2.03	1.91	15.7	-51.5	2.62	0.009
Education of head	Unmatched	5.82	3.91	50.6		3.63	0.000
	Matched	4.89	4.49	10.8	78.7	1.81	0.070
Education of head (squared)	Unmatched	47.73	29.55	45.8		3.10	0.002
	Matched	36.51	31.50	12.6	72.4	2.27	0.023
Edu. of head * Ethnic	Unmatched	2.52	3.65	-30.7		-2.25	0.024
	Matched	3.25	3.06	5.1	83.5	0.81	0.418
Gender of head	Unmatched	0.86	0.89	-9.8		-0.67	0.503
	Matched	0.85	0.83	5.8	41.2	0.82	0.413
Age of head	Unmatched	48.84	43.80	39.1		2.93	0.003
	Matched	47.87	47.24	4.9	87.5	0.72	0.474
No. of labors for salary	Unmatched	1.04	0.76	27.8		1.84	0.065
	Matched	1.00	0.75	25.1	9.7	3.85	0.000
No. of labors in HH agriculture	Unmatched	3.35	3.78	-27.1		-1.94	0.053
	Matched	3.54	3.62	-5.0	81.7	-0.81	0.415
No. of labors in HH business	Unmatched	0.34	0.56	-27.1		-2.02	0.044
	Matched	0.35	0.30	6.0	77.9	1.04	0.299
Hired labor outside	Unmatched	0.49	0.33	34.0		2.36	0.018
	Matched	0.47	0.37	18.7	45.1	2.86	0.004
Loan (ln)	Unmatched	5.63	5.35	6.2		0.44	0.658
	Matched	5.40	4.50	20.1	-222.8	3.10	0.002
Assistance receipt (ln)	Unmatched	2.72	0.49	77.1		4.46	0.000
	Matched	1.76	1.89	-4.7	93.9	-0.65	0.514
Distance to People's Committee	Unmatched	3.14	6.63	-72.1		-6.48	0.000
	Matched	3.50	3.64	-3.0	95.9	-0.60	0.550
Land independent on public irrigation	Unmatched	0.29	0.47	-38.2		-2.84	0.005
	Matched	0.32	0.29	7.3	80.9	1.18	0.240
Paddy land	Unmatched	1.53	1.98	-14.9		-0.97	0.331
	Matched	1.48	1.22	8.6	42.4	1.52	0.128
Paddy land * Problematic land	Unmatched	0.54	0.74	-9.8		-0.76	0.446
	Matched	0.50	0.56	-2.9	70.1	-0.61	0.542

Paddy land * Hybrid seed	Unmatched	0.65	0.39	24.1		1.40	0.161
	Matched	0.52	0.62	-9.9	59.1	-2.16	0.031
Paddy land * Improved seed	Unmatched	0.41	0.58	-8.6		-0.49	0.621
	Matched	0.48	0.27	10.9	-25.9	1.57	0.116
Paddy land * Old local seed	Unmatched	0.42	1.01	-23.2		-1.84	0.066
	Matched	0.49	0.33	6.2	73.1	1.32	0.186

Summary of propensity score after matching

	Percentiles		
1%	2.939704		
5%	2.955918		
10%	4.68328	Obs	23
25%	5.058248	Sum of Wgt.	23
50%	7.803426	Mean	10.71517
		Std. Dev.	7.288932
75%	15.72312		
90%	22.85522	Variance	53.12853
95%	25.1435	Skewness	0.9884122
99%	26.8647	Kurtosis	2.688389

Sample	Pseudo R2	p>chi2	Mean Bias	Med. Bias
Raw	0.229	0	29.6	27.1
Matched	0.075	0	10.7	7.8

Kernel distribution of propensity score before (red) and after (blue) matching

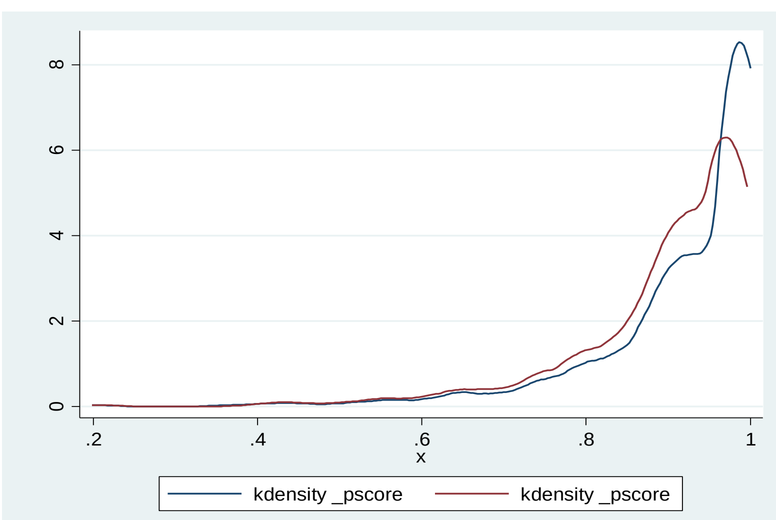


Table 3.14: Estimated treatment effect on outcomes in the case of continuous treatment

Panel A: All sample									
Treatment level		Yield		Revenue		Agricultural income		Total income	
		Average response	Treatment effect	Average response	Treatment effect	Average response	Treatment effect	Average response	Treatment effect
Average	0.37	0.416		2.064		5.612		6.866	
Intervals									
t0	0.002	0.254	0.049	1.196	0.261	5.595	0.060	6.779	-0.066
t1	0.222	0.396	0.047	2.062	0.173	5.508	0.119	6.660	0.260
t2	0.443	0.462	0.011	2.263	0.050	5.654	-0.054	7.082	-0.037
t3	0.663	0.490	0.016	2.397	0.077	5.594	0.043	6.947	-0.070
t4	0.883	0.527	0.015	2.568	0.073	5.701	0.037	6.819	-0.026
t5	1.104	0.557	0.009	2.716	0.048	5.753	-0.013	6.785	0.011
t6	1.324	0.572	0.001	2.802	0.014	5.695	-0.059	6.828	0.040
t7	1.545	0.570	-0.006	2.809	-0.024	5.542	-0.096	6.932	0.065
t8	1.765	0.551	-0.014	2.733	-0.062	5.312	-0.128	7.088	0.086
t9	1.985	0.515	-0.022	2.575	-0.099	5.012	-0.157	7.291	0.106
t10	2.206	0.462	-0.030	2.335	-0.136	4.648	-0.186	7.536	0.125

Panel B: Without disaster									
Treatment level		Yield		Revenue		Agricultural income		Total income	
		Average response	Treatment effect	Average response	Treatment effect	Average response	Treatment effect	Average response	Treatment effect
Average	0.359	0.417		2.091		5.565		6.877	
Intervals									
t0	0.002	0.228	0.054	1.143	0.269	5.529	0.097	6.851	-0.096
t1	0.222	0.398	0.055	2.084	0.214	5.493	0.119	6.641	0.295
t2	0.443	0.471	0.014	2.301	0.067	5.650	-0.090	7.155	-0.076
t3	0.663	0.507	0.020	2.494	0.115	5.506	-0.003	6.899	-0.137
t4	0.883	0.551	0.019	2.751	0.112	5.491	-0.043	6.636	-0.072
t5	1.104	0.590	0.013	2.984	0.085	5.348	-0.123	6.512	-0.019
t6	1.324	0.614	0.005	3.151	0.049	5.034	-0.191	6.495	0.019
t7	1.545	0.621	-0.003	3.235	0.011	4.577	-0.250	6.556	0.049
t8	1.765	0.610	-0.011	3.236	-0.026	3.993	-0.305	6.679	0.074
t9	1.985	0.581	-0.019	3.157	-0.062	3.289	-0.358	6.856	0.096
t10	2.206	0.535	-0.027	2.999	-0.097	2.469	-0.410	7.081	0.118

Panel C: Flood/ Storm/ Landslide									
Treatment level		Yield		Revenue		Agricultural income		Total income	
		Average response	Treatment effect	Average response	Treatment effect	Average response	Treatment effect	Average response	Treatment effect
Average	0.393	0.452		2.168		5.428		6.644	
Intervals									
t0	0.008	0.489	0.001	2.161	0.051	5.407	-0.133	6.476	-0.062
t1	0.151	0.476	-0.059	2.191	-0.151	5.264	0.116	6.435	0.188
t2	0.294	0.411	0.021	2.020	0.073	5.418	0.058	6.677	0.065
t3	0.438	0.444	0.019	2.140	0.098	5.503	0.042	6.767	0.035
t4	0.581	0.468	0.011	2.274	0.071	5.546	-0.041	6.798	-0.053
t5	0.724	0.482	0.007	2.371	0.046	5.477	-0.064	6.709	-0.089
t6	0.867	0.492	0.003	2.432	0.021	5.391	-0.024	6.585	-0.066
t7	1.010	0.495	-0.001	2.457	0.001	5.370	0.037	6.499	-0.025
t8	1.154	0.493	-0.004	2.455	-0.015	5.437	0.097	6.473	0.014
t9	1.297	0.487	-0.005	2.431	-0.027	5.587	0.149	6.501	0.047
t10	1.440	0.479	-0.007	2.389	-0.038	5.811	0.195	6.573	0.073

Panel D: Drought									
Treatment level		Yield		Revenue		Agricultural income		Total income	
		Average response	Treatment effect	Average response	Treatment effect	Average response	Treatment effect	Average response	Treatment effect
Average	0.352	0.401		2.093		5.737		6.887	
Intervals									
t0	0.013	0.358	0.000	1.868	0.021	5.027	0.361	6.302	0.287
t1	0.137	0.364	0.038	1.910	0.194	5.551	0.337	6.710	0.281
t2	0.262	0.412	0.028	2.157	0.149	5.899	-0.055	7.017	0.049
t3	0.386	0.443	0.001	2.323	0.005	5.826	-0.053	7.070	-0.002
t4	0.511	0.443	-0.015	2.316	-0.080	5.767	0.004	7.063	-0.039
t5	0.635	0.423	-0.021	2.211	-0.116	5.779	0.057	7.012	-0.057
t6	0.759	0.397	-0.020	2.065	-0.130	5.856	0.106	6.940	-0.059
t7	0.884	0.372	-0.017	1.902	-0.137	5.993	0.147	6.868	-0.053
t8	1.008	0.352	-0.013	1.730	-0.143	6.180	0.183	6.802	-0.047
t9	1.132	0.337	-0.009	1.552	-0.149	6.411	0.213	6.744	-0.044
t10	1.257	0.326	-0.006	1.365	-0.158	6.679	0.239	6.689	-0.044

Panel E: Plant disease/ Insect									
Treatment level		Yield		Revenue		Agricultural income		Total income	
		Average response	Treatment effect	Average response	Treatment effect	Average response	Treatment effect	Average response	Treatment effect
Average	0.396	0.391		1.836		5.746		6.786	
Intervals									
t0	0.013	0.324	-0.011	1.086	0.329	5.647	-0.110	6.309	0.140
t1	0.222	0.339	0.047	1.783	0.183	5.518	0.115	6.616	0.164
t2	0.430	0.426	0.025	2.075	0.062	5.772	0.107	6.953	0.123
t3	0.639	0.470	0.013	2.187	0.037	5.972	0.069	7.162	0.045
t4	0.848	0.493	0.007	2.259	0.028	6.105	0.053	7.230	0.007
t5	1.057	0.505	0.003	2.314	0.020	6.213	0.050	7.236	-0.004
t6	1.266	0.510	0.000	2.348	0.008	6.319	0.053	7.227	-0.004
t7	1.474	0.508	-0.003	2.357	-0.007	6.434	0.059	7.222	-0.001
t8	1.683	0.499	-0.007	2.333	-0.024	6.560	0.065	7.221	0.001
t9	1.892	0.484	-0.010	2.274	-0.042	6.700	0.072	7.223	0.002
t10	2.101	0.462	-0.013	2.176	-0.061	6.854	0.079	7.227	0.001

Note: The marginal effect is computed as: $(E[Y(t+0.1)\ ]\ ) - \mathrm{E}[\mathrm{Y(t)}]$, where t is treatment level; t0 and t10 correspond to the min and max of chemical fertilizer adoption.

Table 3.15: Estimated generalized propensity scores of the continuous treatment

Dependant var. is amount of fertilizer used	All sample	No disaster	Flood/Storm/ Landslide	Drought	Plant disease/ Insect
HH size	0.007	0.005	0.097	0.013	0.028
	(0.021)	0.023	(0.073)	(0.088)	(0.091)
No. of member <10	-0.039	-0.031	-0.08	-0.056	-0.163
	(0.026)	0.029	(0.085)	(0.102)	(0.108)
No. of members 10-15	0.005	-0.002	-0.117	0.112	-0.051
	(0.026)	0.029	(0.081)	(0.112)	(0.108)
No. of members 15-25	-0.029	-0.028	-0.129*	0.016	-0.058
	(0.022)	0.024	(0.074)	(0.096)	(0.088)
No. of members 25-60	-0.049**	-0.05**	-0.031	0.068	-0.069
	(0.022)	0.023	(0.07)	(0.091)	(0.093)
Education of head	0.026**	0.022*	-0.009	0.001	0.073*
	(0.011)	0.012	(0.033)	(0.044)	(0.042)
Education of head (squared)	-0.001	-0.001	0.002	0.003	0.0004
	(0.001)	0.001	(0.003)	(0.004)	(0.003)
Edu. of head * Ethnic	-0.007*	-0.016***	0.023*	-0.024	-0.003
	(0.004)	0.005	(0.012)	(0.016)	(0.017)
Gender of head	0.02	0.039	0.125	-0.203	-0.106
	(0.034)	0.036	(0.099)	(0.144)	(0.155)
Age of head	0.001	0.001	-0.0002	0.011**	-0.011**
	(0.001)	0.001	(0.004)	(0.005)	(0.005)
No. of labors for salary	0.033***	0.036***	-0.011	-0.054	0.069
	(0.011)	0.012	(0.035)	(0.049)	(0.045)
No. of labors in HH agriculture	0.001	-0.001	-0.014	-0.006	0.051
	(0.011)	0.012	(0.035)	(0.061)	(0.048)
No. of labors in HH business	-0.023	-0.017	0.017	-0.072	-0.104*
	(0.014)	0.016	(0.053)	(0.07)	(0.061)
Hired labor outside	-0.01	-0.053**	0.015	0.005	0.024
	(0.023)	0.026	(0.068)	(0.095)	(0.098)
Loan (ln)	0.007***	0.006**	-0.004	0.005	0.006
	(0.002)	0.003	(0.007)	(0.01)	(0.01)
Assistance receipt (ln)	0.006***	0.008**	0.012	-0.021	0.018
	(0.003)	0.003	(0.011)	(0.014)	(0.014)
Distance to People's Committee	-0.011***	-0.006	-0.012	-0.016	-0.032**
	(0.004)	0.005	(0.012)	(0.011)	(0.013)
Land independent on public irrigation	0.075***	0.042	0.115	0.105	0.072
	(0.026)	0.03	(0.071)	(0.098)	(0.111)
Paddy land	0.009	0.005	-0.125	-0.358	-0.084
	(0.012)	0.012	(0.176)	(0.386)	(0.169)
Paddy land * Problematic land	-0.012*	-0.005	0.019	0.005	0.075*
	(0.006)	0.007	(0.051)	(0.078)	(0.045)

Paddy land * Hybrid seed	-0.011	-0.004	0.08	0.162	0.048
	(0.013)	0.012	(0.176)	(0.384)	(0.17)
Paddy land * Improved seed	-0.009	-0.011	0.071	0.453	0.111
	(0.012)	0.011	(0.174)	(0.385)	(0.169)
Paddy land * Old local seed	-0.01	-0.003	0.075	0.219	0.089
	(0.012)	0.011	(0.176)	(0.391)	(0.168)
Constant	-1.059***	-0.995***	-1.069***	-1.575***	-0.949***
	(0.079)	0.086	(0.242)	0.342	(0.325)
Observations	1995	1360	207	142	215
Log likelihood	-1386	-810	-116	-104	-202

Note: Significance at 10%, 5% and 1% are denoted as *, **, ***. Standard deviations are displayed in parentheses. This is regression of Box-Cox transformation.

Table 3.16: Balancing test and generalized propensity score distribution in the case of continuous treatment

Panel A: All sample						
	Treatment Interval 1 [.0017, .3694]			Treatment Interval 2 [.3701, 2.2059]		
	Mean Difference	S. D.	t-value	Mean Difference	S. D.	t-value
HH size	-0.136	0.079	-1.729	0.140	0.078	1.792
No. of member <10	-0.046	0.040	-1.138	0.057	0.041	1.390
No. of members 10-15	-0.028	0.035	-0.809	0.033	0.035	0.959
No. of members 15-25	-0.026	0.052	-0.498	0.017	0.052	0.319
No. of members 25-60	-0.061	0.037	-1.620	0.057	0.037	1.543
Education of head	0.079	0.133	0.594	-0.127	0.136	-0.934
Education of head (squared)	0.439	1.580	0.278	-0.685	1.591	-0.431
Edu. of head * Ethnic	-0.099	0.145	-0.682	0.043	0.146	0.296
Gender of head	0.000	0.016	-0.016	0.001	0.016	0.037
Age of head	0.450	0.589	0.764	-0.622	0.589	-1.056
No. of labors for salary	-0.007	0.047	-0.159	-0.008	0.047	-0.176
No. of labors in HH agriculture	-0.126	0.066	-1.907	0.121	0.065	1.853
No. of labors in HH business	0.003	0.038	0.090	-0.002	0.037	-0.052
Hired labor outside	0.005	0.023	0.242	-0.007	0.022	-0.304
Loan (ln)	0.361	0.210	1.722	-0.294	0.208	-1.409
Assistance receipt (ln)	0.234	0.160	1.468	-0.234	0.159	-1.470
Distance to People's Committee	-0.070	0.134	-0.524	0.147	0.134	1.090
Land independent on public irrigation	-0.004	0.020	-0.185	0.001	0.020	0.051
Paddy land	0.401	0.181	2.214	-0.420	0.184	-2.289
Paddy land * Problematic land	0.281	0.122	2.314	-0.277	0.122	-2.264
Paddy land * Hybrid seed	0.098	0.071	1.375	-0.080	0.070	-1.142
Paddy land * Improved seed	0.093	0.136	0.685	-0.106	0.140	-0.760
Paddy land * Old local seed	0.170	0.109	1.568	-0.203	0.109	-1.855

Note: Test that the conditional mean of the pre-treatment variables given the generalized propensity score is not different between units who belong to a particular treatment interval and units who belong to all other treatment intervals.

GPS evaluated at the representative point of each treatment interval

	Obs	Mean	Std. Dev.	Min	Max
GPS of interval 1	1995	0.688	0.095	0.381	0.823
GPS of interval 2	1995	0.603	0.127	0.120	0.820
Balancing test	0.01				

Panel B: Households without disasters						
	Treatment Interval 1 [.0017, .3694]			Treatment Interval 2 [.3701, 2.2059]		
	Mean Difference	S. D.	t-value	Mean Difference	S. D.	t-value
HH size	-0.065	0.065	-0.991	0.085	0.065	1.303
No. of member <10	-0.009	0.034	-0.264	0.018	0.034	0.530
No. of members 10-15	-0.005	0.029	-0.182	0.003	0.029	0.105
No. of members 15-25	-0.009	0.043	-0.207	0.013	0.043	0.313
No. of members 25-60	-0.047	0.032	-1.480	0.058	0.032	1.822
Education of head	-0.054	0.128	-0.425	0.034	0.127	0.268
Education of head (squared)	-1.164	1.498	-0.778	0.857	1.487	0.576
Edu. of head * Ethnic	0.035	0.107	0.326	-0.023	0.107	-0.212
Gender of head	0.005	0.014	0.400	-0.007	0.013	-0.495
Age of head	0.066	0.492	0.134	-0.047	0.488	-0.096
No. of labors for salary	-0.012	0.039	-0.320	0.005	0.039	0.125
No. of labors in HH agriculture	-0.083	0.053	-1.569	0.094	0.053	1.783
No. of labors in HH business	0.016	0.032	0.511	-0.004	0.031	-0.126
Hired labor outside	-0.005	0.019	-0.245	0.010	0.019	0.542
Loan (ln)	0.278	0.176	1.576	-0.315	0.176	-1.784
Assistance receipt (ln)	0.142	0.129	1.104	-0.164	0.127	-1.295
Distance to People's Committee	0.073	0.103	0.705	-0.030	0.104	-0.288
Land independent on public irrigation	-0.013	0.016	-0.793	0.012	0.016	0.747
Paddy land	0.247	0.159	1.560	-0.283	0.161	-1.755
Paddy land * Problematic land	0.180	0.110	1.641	-0.208	0.112	-1.855
Paddy land * Hybrid seed	0.066	0.061	1.095	-0.078	0.061	-1.287
Paddy land * Improved seed	0.039	0.121	0.323	-0.068	0.125	-0.542
Paddy land * Old local seed	0.121	0.091	1.329	-0.114	0.090	-1.262

Note: Test that the conditional mean of the pre-treatment variables given the generalized propensity score is not different between units who belong to a particular treatment interval and units who belong to all other treatment intervals.

GPS evaluated at the representative point of each treatment interval

	Obs	Mean	Std. Dev.	Min	Max
GPS of interval 1	1360	0.764	0.110	0.403	0.909
GPS of interval 2	1360	0.650	0.142	0.038	0.907
Balancing test	0.05				

Panel C: Flood/ Storm/ Landslide						
	Treatment Interval 1 [.0081, .3684]			Treatment Interval 2 [.3714, 1.4400]		
	Mean Difference	S. D.	t-value	Mean Difference	S. D.	t-value
HH size	-0.023	0.025	-0.913	0.020	0.025	0.815
No. of member <10	0.002	0.014	0.157	0.000	0.014	0.023
No. of members 10-15	-0.015	0.011	-1.381	0.013	0.011	1.219
No. of members 15-25	-0.017	0.015	-1.152	0.017	0.015	1.150
No. of members 25-60	0.003	0.011	0.250	-0.003	0.011	-0.321
Education of head	0.037	0.049	0.755	-0.033	0.049	-0.679
Education of head (squared)	0.364	0.571	0.639	-0.274	0.574	-0.478
Edu. of head * Ethnic	0.036	0.046	0.780	-0.019	0.047	-0.415
Gender of head	-0.003	0.005	-0.596	0.003	0.005	0.586
Age of head	0.126	0.187	0.674	-0.176	0.189	-0.934
No. of labors for salary	-0.007	0.015	-0.481	0.006	0.015	0.421
No. of labors in HH agriculture	-0.024	0.022	-1.112	0.023	0.021	1.056
No. of labors in HH business	-0.003	0.009	-0.330	0.004	0.009	0.438
Hired labor outside	0.001	0.007	0.122	-0.001	0.007	-0.212
Loan (ln)	0.029	0.065	0.441	-0.061	0.066	-0.927
Assistance receipt (ln)	-0.039	0.045	-0.872	0.021	0.046	0.470
Distance to People's Committee	0.007	0.041	0.179	-0.009	0.041	-0.218
Land independent on public irrigation	-0.003	0.007	-0.435	0.003	0.007	0.499
Paddy land	-0.053	0.044	-1.224	0.034	0.043	0.787
Paddy land * Problematic land	-0.057	0.040	-1.441	0.039	0.040	0.986
Paddy land * Hybrid seed	0.005	0.013	0.402	-0.006	0.013	-0.492
Paddy land * Improved seed	-0.004	0.014	-0.295	0.005	0.014	0.341
Paddy land * Old local seed	-0.052	0.042	-1.240	0.033	0.042	0.799

Note: Test that the conditional mean of the pre-treatment variables given the generalized propensity score is not different between units who belong to a particular treatment interval and units who belong to all other treatment intervals.

GPS evaluated at the representative point of each treatment interval

	Obs	Mean	Std. Dev.	Min	Max
GPS of interval 1	207	0.743	0.173	0.055	0.940
GPS of interval 2	207	0.752	0.159	0.002	0.940
Balancing test	0.1				

Panel D: Drought						
	Treatment Interval 1 [.0132, .3632]			Treatment Interval 2 [.3852, 1.257]		
	Mean Difference	S. D.	t-value	Mean Difference	S. D.	t-value
HH size	0.003	0.025	0.116	-0.009	0.024	-0.381
No. of member <10	-0.009	0.015	-0.596	0.009	0.014	0.621
No. of members 10-15	0.008	0.010	0.848	-0.009	0.010	-0.952
No. of members 15-25	0.005	0.015	0.312	-0.003	0.015	-0.195
No. of members 25-60	-0.001	0.010	-0.078	-0.005	0.010	-0.450
Education of head	0.014	0.051	0.263	0.005	0.051	0.089
Education of head (squared)	0.179	0.571	0.313	0.078	0.562	0.138
Edu. of head * Ethnic	-0.042	0.048	-0.867	0.027	0.047	0.585
Gender of head	-0.006	0.004	-1.454	0.005	0.004	1.304
Age of head	0.106	0.149	0.713	-0.051	0.148	-0.349
No. of labors for salary	0.009	0.015	0.584	-0.010	0.015	-0.653
No. of labors in HH agriculture	0.001	0.017	0.079	-0.011	0.018	-0.617
No. of labors in HH business	0.003	0.009	0.305	0.001	0.010	0.081
Hired labor outside	-0.006	0.007	-0.963	0.004	0.007	0.631
Loan (ln)	-0.045	0.060	-0.754	0.008	0.060	0.137
Assistance receipt (ln)	-0.020	0.052	-0.385	0.028	0.053	0.529
Distance to People's Committee	-0.040	0.067	-0.599	0.056	0.066	0.850
Land independent on public irrigation	0.003	0.006	0.527	-0.001	0.006	-0.097
Paddy land	-0.020	0.018	-1.123	0.000	0.020	-0.008
Paddy land * Problematic land	0.001	0.010	0.132	0.001	0.010	0.134
Paddy land * Hybrid seed	-0.028	0.016	-1.770	0.016	0.017	0.926
Paddy land * Improved seed	0.014	0.011	1.285	-0.018	0.012	-1.456
Paddy land * Old local seed	-0.005	0.009	-0.572	0.001	0.010	0.148

Note: Test that the conditional mean of the pre-treatment variables given the generalized propensity score is not different between units who belong to a particular treatment interval and units who belong to all other treatment intervals.

GPS evaluated at the representative point of each treatment interval

	Obs	Mean	Std. Dev.	Min	Max
GPS of interval 1	142	0.565	0.203	0.003	0.791
GPS of interval 2	142	0.483	0.232	0.000	0.791
Balancing test	0.05				

Panel E: Plant diseases/ Insects						
	Treatment Interval 1 [.0129, .3650]			Treatment Interval 2 [.3712, 2.1008]		
	Mean Difference	S. D.	t-value	Mean Difference	S. D.	t-value
HH size	-0.003	0.026	-0.099	-0.002	0.026	-0.071
No. of member <10	-0.012	0.014	-0.834	0.010	0.014	0.690
No. of members 10-15	0.000	0.012	0.037	-0.001	0.012	-0.127
No. of members 15-25	0.006	0.019	0.285	-0.008	0.020	-0.422
No. of members 25-60	0.000	0.012	-0.031	0.000	0.012	0.025
Education of head	0.055	0.050	1.097	-0.038	0.047	-0.808
Education of head (squared)	0.412	0.542	0.761	-0.276	0.497	-0.555
Edu. of head * Ethnic	0.017	0.055	0.303	-0.019	0.056	-0.344
Gender of head	0.005	0.005	0.989	-0.005	0.005	-0.947
Age of head	-0.040	0.190	-0.212	0.051	0.192	0.267
No. of labors for salary	0.020	0.017	1.142	-0.013	0.017	-0.764
No. of labors in HH agriculture	-0.001	0.026	-0.030	-0.002	0.026	-0.078
No. of labors in HH business	-0.017	0.013	-1.373	0.015	0.013	1.158
Hired labor outside	0.012	0.007	1.596	-0.012	0.007	-1.591
Loan (ln)	-0.034	0.071	-0.480	0.028	0.072	0.385
Assistance receipt (ln)	0.069	0.056	1.233	-0.058	0.056	-1.042
Distance to People's Committee	-0.105	0.056	-1.882	0.087	0.056	1.559
Land independent on public irrigation	0.010	0.007	1.472	-0.009	0.007	-1.268
Paddy land	0.099	0.073	1.356	-0.099	0.074	-1.330
Paddy land * Problematic land	0.024	0.018	1.310	-0.027	0.018	-1.441
Paddy land * Hybrid seed	0.002	0.024	0.078	-0.009	0.025	-0.364
Paddy land * Improved seed	0.042	0.061	0.688	-0.039	0.063	-0.614
Paddy land * Old local seed	0.058	0.039	1.476	-0.054	0.040	-1.349

Note: Test that the conditional mean of the pre-treatment variables given the generalized propensity score is not different between units who belong to a particular treatment interval and units who belong to all other treatment intervals.

GPS evaluated at the representative point of each treatment interval

	Obs	Mean	Std. Dev.	Min	Max
GPS of interval 1	215	0.445	0.158	0.031	0.644
GPS of interval 2	215	0.393	0.189	0.008	0.644
Balancing test	0.05				

Table 3.17: Estimates of conditional distribution of treatment given covariates.

Panel A: Rice yield					
	All sample	Without disaster	Flood/Storm/ Landslide	Drought	Plant disease/ Insects
Fertilizer	0.47***	0.53***	0.045	-0.11	0.232**
	(0.056)	(0.069)	(0.193)	(0.186)	(0.104)
Fertilizer squared	-0.17***	-0.177***	-0.037	0.045	-0.081
	(0.037)	(0.048)	(0.142)	(0.131)	(0.055)
GPS	0.17**	0.261***	-0.338	0.029	-0.343
	(0.076)	(0.085)	(0.209)	(0.241)	(0.252)
GPS squared	-0.026	-0.084	0.166	-0.053	0.364
	(0.07)	(0.07)	(0.165)	(0.248)	(0.292)
Fertilizer*GPS	-0.237***	-0.31***	0.285*	0.51**	0.442***
	(0.071)	(0.080)	(0.166)	(0.213)	(0.163)
Cons.	0.253***	0.227***	0.489***	0.359***	0.321***
	(0.019)	(0.023)	(0.056)	(0.061)	(0.053)
Obs.	1995	1360	207	142	215
R-squared	0.16	0.19	0.07	0.129	0.347
Balancing test	0.01	0.05	0.1	0.05	0.05

Panel B: Rice revenue					
	All sample	Without disaster	Flood/Storm/ Landslide	Drought	Plant disease/ Insects
Fertilizer	2.324***	2.59***	0.655	0.292	1.56*
	(0.302)	(0.376)	(0.941)	(1.025)	(0.603)
Fertilizer squared	-0.818***	-0.791***	-0.346*	-0.624	-0.496
	(0.199)	(0.262)	(0.692)	(0.724)	(0.316)
GPS	1.658***	1.869***	-1.01	-0.352	1.068
	(0.41)	(0.462)	(1.028)	(1.33)	(1.458)
GPS squared	-711*	-0.765**	0.415	0.309	-0.443
	(0.378)	(0.381)	(0.805)	(1.371)	(1.687)
Fertilizer*GPS	-1.91***	-2.292***	0.848	2.172*	0.064
	(0.382)	(0.433)	(0.809)	(1.174)	(0.941)
Cons.	1.192***	1.138***	2.156***	1.864***	1.066***
	(0.1)	(0.126)	(0.272)	(0.365)	(0.308)
Obs.	1995	1360	207	142	215
R-squared	0.11	0.127	0.069	0.108	0.183
Balancing test	0.01	0.05	0.1	0.05	0.05

Panel C: Agricultural income					
	All sample	Without disaster	Flood/Storm/ Landslide	Drought	Plant disease/ Insects
Fertilizer	0.952**	-0.935**	-0.779	0.683	0.637
	(0.401)	(0.375)	(0.879)	(0.921)	(0.625)
Fertilizer squared	-0.625**	0.47*	0.559	-0.347	-0.117
	(0.625)	(0.262)	(0.646)	(0.651)	(0.328)
GPS	-1.151**	-0.915**	0.805	2.049*	0.657
	(0.545)	(0.462)	(0.951)	(1.195)	(1.512)
GPS squared	1.356***	0.697*	-0.716	-1.669	-1.481
	(0.502)	(0.381)	(0.752)	(1.232)	(1.749)
Fertilizer*GPS	-0.462	2.402***	0.966	-0.015	2.212**
	(0.507)	(0.433)	(0.755)	(1.054)	(0.976)
Cons.	5.594***	6.852***	6.482***	6.292	6.301***
	(0.133)	(0.125)	(0.255)	(0.302)	(0.319)
Obs.	1995	1372	208	142	215
R-squared	0.009	0.069	0.014	0.058	0.215
Balancing test	0.01	0.05	0.1	0.05	0.05

Panel D: Total income					
	All sample	Without disaster	Flood/Storm/ Landslide	Drought	Plant disease/ Insects
Fertilizer	-0.544*	1.215**	-1.544	0.285	0.278
	(0.301)	(0.494)	(1.44)	(1.229)	(0.797)
Fertilizer squared	0.401**	-1.179***	1.232	0.826	0.13
	(0.199)	(0.345)	(1.059)	(0.868)	(0.418)
GPS	-0.856**	-1.063*	0.636	3.006*	-1.136
	(0.409)	(0.607)	(1.557)	(1.594)	(1.93)
GPS squared	0.795**	1.142**	-0.665	-1.797	0.654
	(0.377)	(0.501)	(1.232)	(1.643)	(2.231)
Fertilizer*GPS	1.934***	-0.285	1.262	-2.247	1.728
	(0.381)	(0.57)	(1.237)	(1.406)	(1.245)
Cons.	6.78***	5.527***	5.419***	5.022***	5.643***
	(0.1)	(0.165)	(0.417)	(0.403)	(0.407)
Obs.	1995	1360	207	142	215
R-squared	0.05	0.022	0.011	0.066	0.108
Balancing test	0.01	0.05	0.1	0.05	0.05

Note: Significant at 10%, 5% and 1% are denoted as *, **, ***. Standard deviations are displayed in parentheses.

Table 3.18: OLS and instrumental variable estimates of the binary treatment

Panel A: OLS	Yield			Revenue			Agricultural income			Total income		
	All sample	Without disaster	With disasters	All sample	Without disaster	With disasters	All sample	Without disaster	With disasters	All sample	Without disaster	With disasters
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Chemical fertilizer	0.108***	0.114***	0.097***	0.407***	0.426***	0.403***	0.247***	0.21***	0.387***	0.194***	0.167***	0.328***
	(0.01)	(0.012)	(0.02)	(0.053)	(0.061)	(0.107)	(0.058)	(0.066)	(0.125)	(0.044)	(0.049)	(0.092)
Obs.	2,260	1,570	690	2,260	1,570	690	2,260	1,570	690	2,260	1,570	690
R-squared	0.19	0.24	0.18	0.16	0.21	0.13	0.27	0.29	0.25	0.36	0.41	0.3

Note: Significant at 10%, 5% and 1% are denoted as *, **, ***. Standard deviations are displayed in parentheses.

Panel B: 2 SLS												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Chemical fertilizer	0.218***	0.217***	0.281***	1.039***	1.134***	0.982***	0.957***	0.715***	1.276***	0.487***	0.219***	0.281***
	(0.037)	(0.039)	(0.084)	(0.193)	(0.209)	(0.421)	(0.213)	(0.222)	(0.5)	(0.155)	(0.163)	(0.084)
Excluded IV												
Fertilizer information (dummy)	0.192***	0.238***	0.075***	0.192***	0.238***	0.075***	0.192***	0.238***	0.075***	0.192***	0.238***	0.075***
	(0.019)	(0.023)	(0.031)	(0.019)	(0.023)	(0.031)	(0.019)	(0.023)	(0.031)	(0.019)	(0.023)	(0.031)
Crop insurance (dummy)	0.125***	0.114***	0.136***	0.125***	0.114***	0.136***	0.125***	0.114***	0.136***	0.125***	0.114***	0.136***
	(0.013)	(0.016)	(0.031)	(0.013)	(0.016)	(0.031)	(0.013)	(0.016)	(0.031)	(0.013)	(0.016)	(0.031)
Included IV	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Endogeneity test												
Hausman t-test	-0.12***	-0.113***	-0.197***	0.353***	0.354***	0.362***	0.185***	0.159**	0.325***	0.169***	0.162***	0.25***
Weak IV test												
Relevance condition, F-test	96.13***	77.45***	23.06***	96.13***	77.45***	23.06***	96.13***	77.45***	23.06***	96.13***	77.45***	23.06***
Overidentifying, Sargan test	7.75***	6.61***	0.81	1.75	2.68	0.01	21.33***	27.7***	0.05	13.9***	13.59***	0.22
Obs.	2,260	1,570	690	2,260	1,570	690	2,260	1,570	690	2,260	1,570	690
Centered R-squared	0.15	0.27	0.18	0.1	0.14	0.18	0.22	0.27	0.18	0.35	0.41	0.15

Note: (i) Significant at 10%, 5% and 1% are denoted as *, **, ***.
(ii) The null hypothesis of the Hausman test is that the IV is exogenous, i.e., an OLS estimator of the same equation would yield consistent estimates, hence the IV process is unnecessary.
(iii) The null hypothesis for relevance condition is that the parameters of exclued IVs are jointly equal to zero, or the excluded IVs are jointly uncorrelated with treatment variable.
(iv) The null hypothesis of Sagan test for exclusion restriction or overidentifying restrictions is that the IVs are valid instruments, i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. The statistic is chi-squared distributed.

Table 3.19: OLS and instrumental variable estimates of the continuous treatment

Panel A: OLS	Yield		Revenue		Agricultural income		Total income	
	All sample	Without disaster	All sample	Without disaster	All sample	Without disaster	All sample	Without disaster
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Chemical fertilizer	0.192***	0.216***	0.796***	0.909***	0.186***	0.002	0.219***	0.035
	(0.012)	(0.015)	(0.063)	(0.082)	(0.074)	(0.095)	(0.053)	(0.069)
Obs.	1,955	1,360	1,955	1,360	1,955	1,360	1,955	1,360
R-squared	0.21	0.25	0.17	0.21	0.27	0.30	0.35	0.40

Panel B: 2SLS	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Chemical fertilizer	0.313***	0.319***	1.412***	1.559***	3.033***	3.148***	1.713***	1.44***
	(0.074)	(0.081)	(0.4)	(0.438)	(0.608)	(0.673)	(0.393)	(0.413)
Excluded IV								
Fertilizer information (dummy)	0.063***	0.052**	0.063***	0.052**	0.063***	0.052**	0.063***	0.052**
	(0.02)	(0.023)	(0.02)	(0.023)	(0.02)	(0.023)	(0.02)	(0.023)
Crop insurance (ln)	0.016***	0.018***	0.016***	0.018***	0.016***	0.018***	0.016***	0.018***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Included IV	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Endogeneity test								
Hausman t-stat.	-0.124*	-0.107	0.78***	0.885***	0.111	-0.117	0.18***	-0.017
Weak IV test								
Relavance condition, F test	25.81***	24.76***	25.81***	24.76***	25.81***	24.76***	25.81***	24.76***
Overidentifying, Sargan test	2.18	1.27	0.39	8.53***	8.54***	6.15***	7.01***	5.38**
Obs.	1,955	1,360	1,955	1,360	1,955	1,360	1,955	1,360
Centered R-squared	0.17	0.23	0.13	0.18	-0.27	-0.26	0.1	0.22

Note: (i) Significant at 10%, 5% and 1% are denoted as *, **, ***.
(ii) The null hypothesis of the Hausman test is that the IV is exogenous, i.e., an OLS estimator of the same equation would yield consistent estimates, hence the IV process is unnecessary.
(iii) The null hypothesis for relevance condition is that the parameters of exclued IVs are jointly equal to zero, or the excluded IVs are jointly uncorrelated with treatment variable.
(iv) The null hypothesis of Sagan test for exclusion restriction or overidentifying restrictions is that the IVs are valid instruments, i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. The statistic is chi-squared distributed.

Panel C: OLS												
	Yield			Revenue			Agricultural income			Total income		
	Flood/ Storm	Drought	Plant diseases	Flood/ Storm	Drought	Plant diseases	Flood/ Storm	Drought	Plant diseases	Flood/ Storm	Drought	Plant diseases
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Chemical fertilizer	0.092***	0.046	0.167***	0.577***	0.052	0.673***	0.051	0.508	0.836***	0.034	0.25	0.714***
	(0.036)	(0.048)	(0.027)	(0.179)	(0.262)	(0.159)	(0.241)	(0.307)	(0.195)	(0.158)	(0.223)	(0.144)
Obs.	207	142	215	207	142	215	207	142	215	207	142	215
R-squared	0.25	0.37	0.45	0.24	0.37	0.28	0.37	0.36	0.32	0.29	0.4	0.47

Panel D: 2SLS												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Chemical fertilizer	-0.119	1.663	0.477***	-0.385	13.524	2.122***	-2.808	-16.232	1.462**	-1.611	-15.648	1.914***
	(0.355)	(4.122)	(0.121)	(1.727)	(33.429)	(0.656)	(2.878)	(41.442)	(0.69)	(1.793)	(39.05)	(0.58)
Excluded IV												
Fertilizer information (dummy)	0.049	0.015	0.107	0.049	0.015	0.107	0.049	0.015	0.107	0.049	0.015	0.107
	(0.06)	(0.058)	(0.09)	(0.06)	(0.058)	(0.09)	(0.06)	(0.058)	(0.09)	(0.06)	(0.058)	(0.09)
Crop insurance (ln)	-0.009	-0.002	0.036***	-0.009	-0.002	0.036***	-0.009	-0.002	0.036***	-0.009	-0.002	0.036***
	(0.09)	(0.009)	(0.01)	(0.09)	(0.009)	(0.01)	(0.09)	(0.009)	(0.01)	(0.09)	(0.009)	(0.01)
Included IV	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Endogeneity test												
Hausman t-test	0.214	-1.619	-0.335***	0.588***	0.036	0.557***	0.083	0.528*	0.786***	0.052	0.269	0.618***
Weak IV test												
Relavance condition, F-test	1.00	0.07	7.6***	1.00	0.07	7.6***	1.00	0.07	7.6***	1.00	0.07	7.6***
Overidentifying, Sargan test	0.3	0.16	0.00	0.05	0.06	0.46	0.49	0.08	0.97	0.8	0.04	0.09
Obs.	207	142	215	207	142	215	207	142	215	207	142	215
Centered R-squared	0.11	-13.94	0.06	0.12	-15.74	-0.03	-0.11	-15.74	0.29	-0.14	-25.66	0.28

Table 3.20: Distribution of non-users and natural disaster shocks by provinces in the sample

		Total sample	Non-user (fertilizers)	Flood/Storm/ Landslide	Drought	Plant diseases/ Insects
	Obs.	2,280	267	217	151	251
1	Ha Tay	15.91	3	2.3	3.31	8.37
2	Phu Tho	9.9	0	3.23	35.76	1.99
3	Lao Cai	11.5	0	6.91	22.52	1.2
4	Lai Chau	12.15	40.82	3.23	2.65	23.11
5	Dien Bien	12.11	48.69	3.69	8.61	14.34
6	Nghe An	5.49	0.75	0.92	7.28	11.95
7	Quang Nam	8.82	0.37	44.7	0	0.4
8	Khanh Hoa	1.25	0.37	0.92	0.66	1.2
9	Dak Lak	9.81	2.62	30.88	9.93	10.36
10	Dak Nong	5.23	2.62	0	3.97	19.92
11	Lam Dong	0.73	0.37	2.76	5.3	1.2
12	Long An	7.09	0.37	0.46	0	5.98
	Total (%)	100	100	100	100	100

Source: Calculation from the 2010 VARHS.

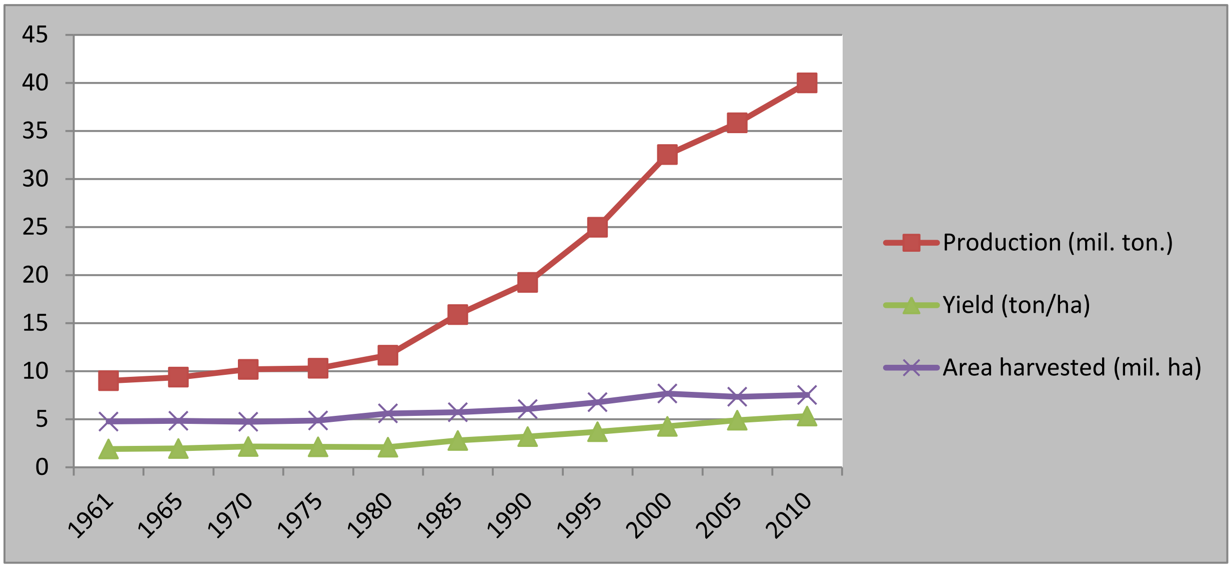
Table 3.21: The treatment effect in two mountainous provinces

	Before matching	After matching			
		Nearest neighbor	Nearest neighbor noreplacement	Caliper norepalcement	Kernel based matching
Yield	0.264***	0.02	0.026**	0.019	0.014
	(0.01)	(0.023)	(0.011)	(0.014)	(0.016)
Rice revenue	0.224***	0.038	0.14**	0.041	0.007
	(0.062)	(0.13)	(0.065)	(0.084)	(0.092)
Agr. income	0.195***	0.178*	0.138**	0.119*	0.15**
	(0.054)	(0.092)	(0.058)	(0.072)	(0.076)
Total Income	0.081	0.114	0.028	0.069	0.049
	(0.053)	(0.096)	(0.057)	(0.072)	(0.077)
Treatment assignment					
Untreated	237	237	237	237	237
Treated	318	308	237	154	308
Balancing test	16.74	9.08	10.7	3.32	6.42

Note: Significant at 10%, 5% and 1% are denoted as *, **, ***.

Figures

Figure 3.1: Development of paddy cultivation in Vietnam, 1961-2010

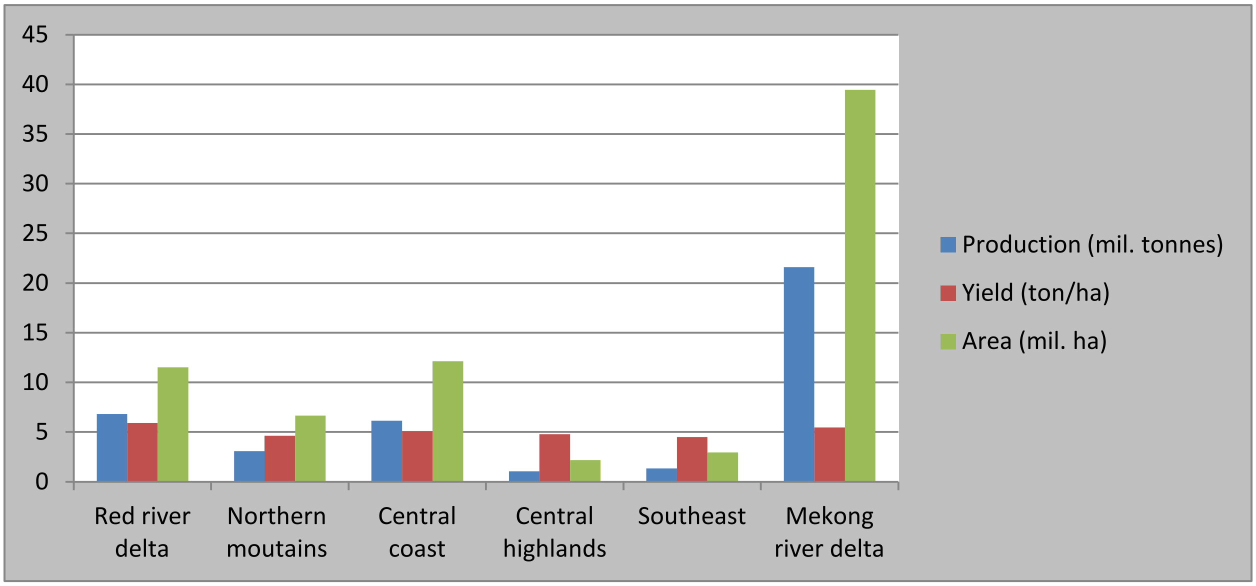


Source: Calculation from FAOSTAT database.

Figure 3.2: Ratio between imported and consumed nutrient fertilizers in Vietnam, 2002-10

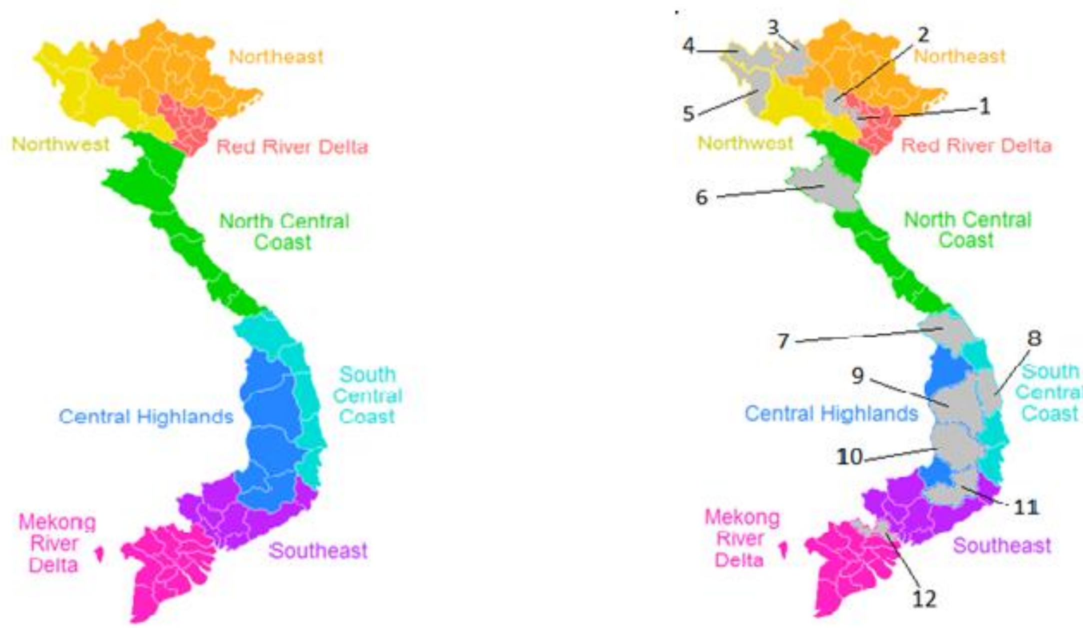
Source: Calculation from FAOSTAT database.

Figure 3.3: Paddy cultivation by regions of Vietnam in 2010



Source: Calculation from GSO database.

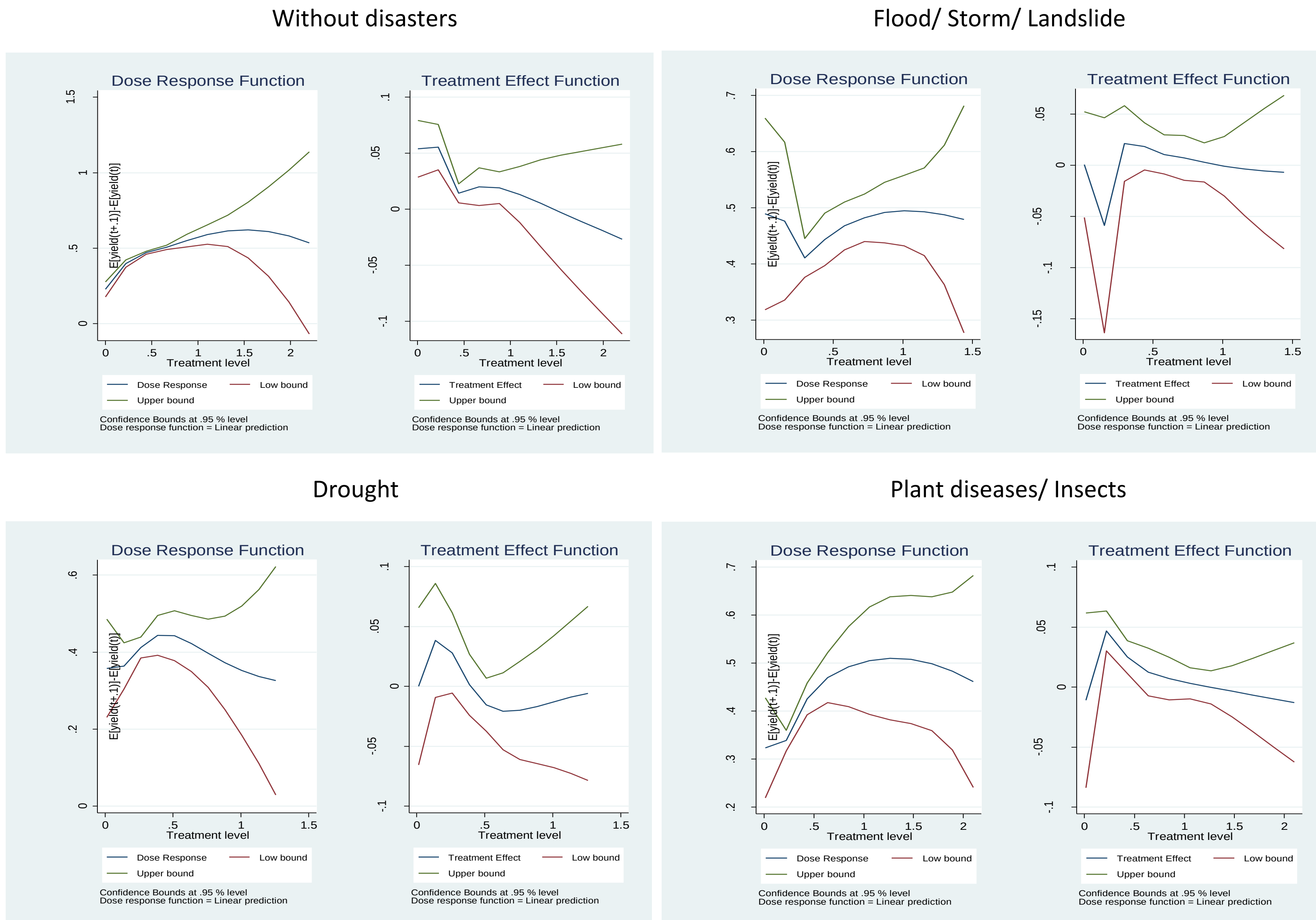
Figure 3.4: Eight regions of Vietnam and twelve provinces in the sample



Source: The Vietnamese regions map from wikimedia.org

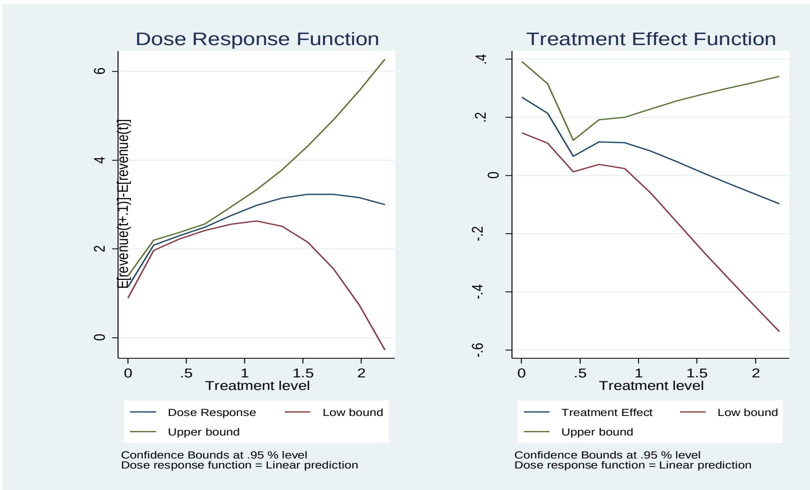
Figure 3.5: Dose-response and marginal effect of continuous treatment

5A: Rice yield

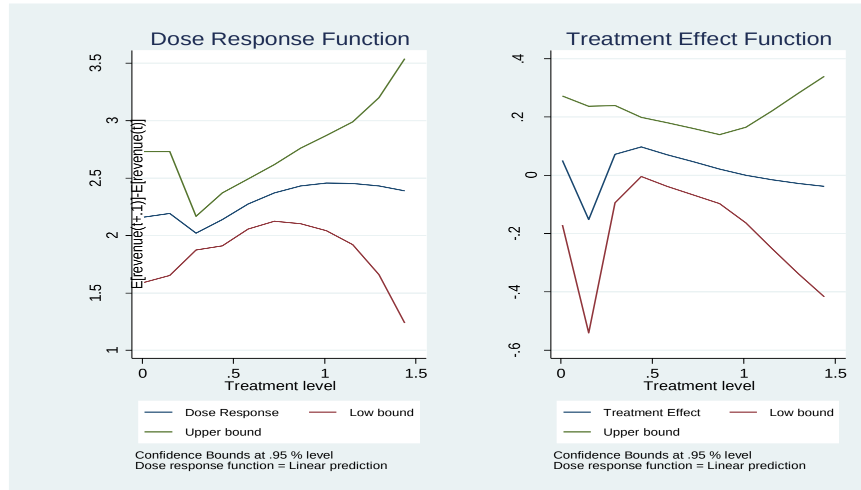


5B: Rice revenue

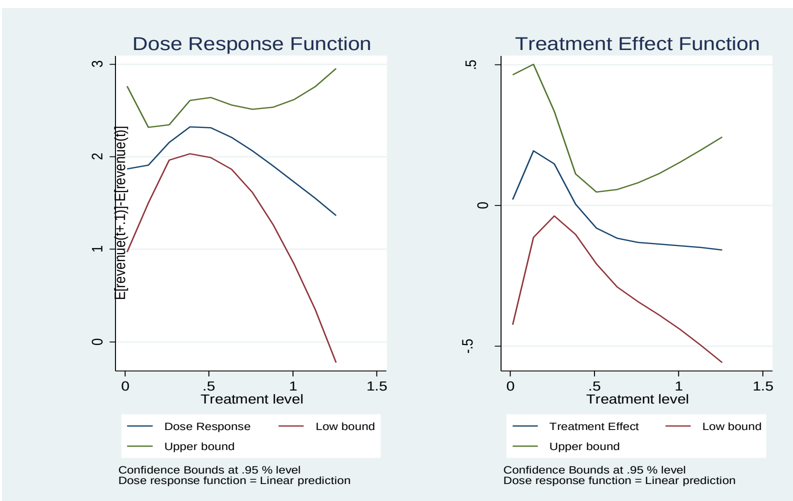
Without disasters



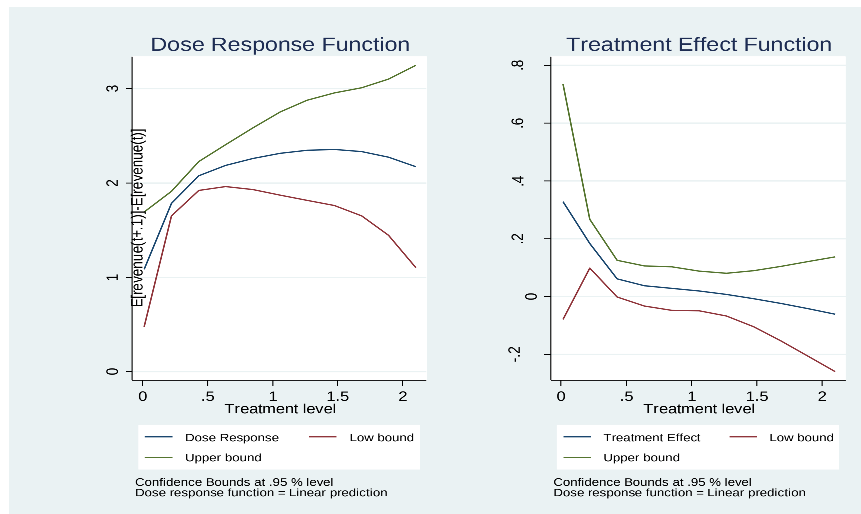
Flood/ Storm/ Landslide



Drought

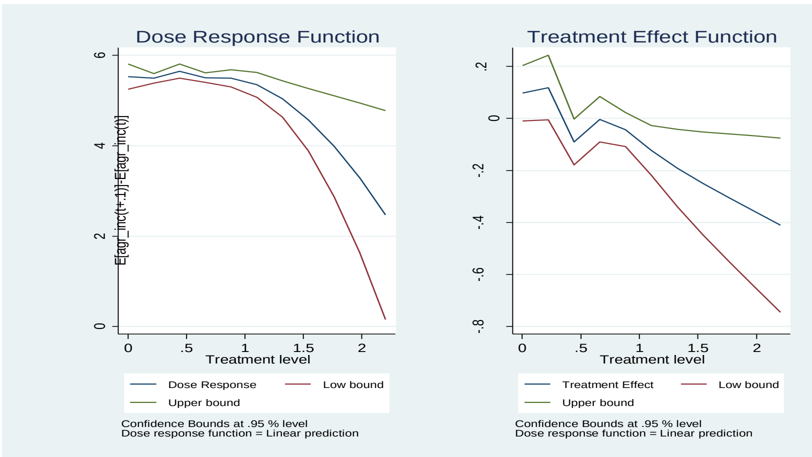


Plant diseases/ Insects

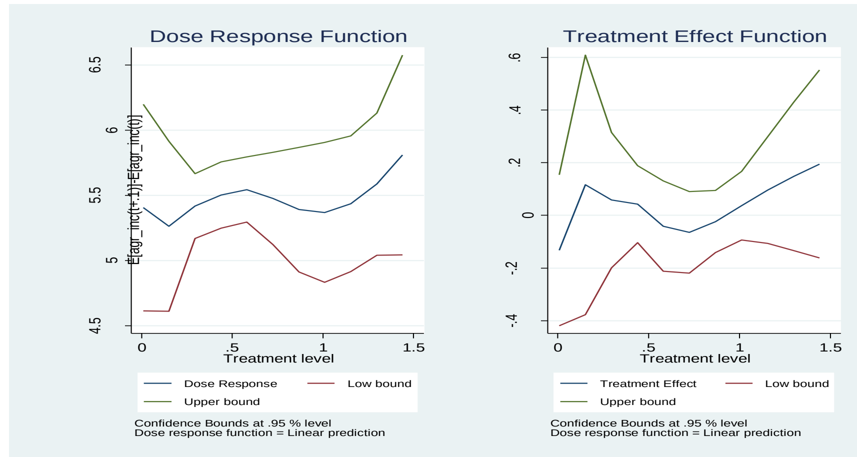


5C: Agricultural income

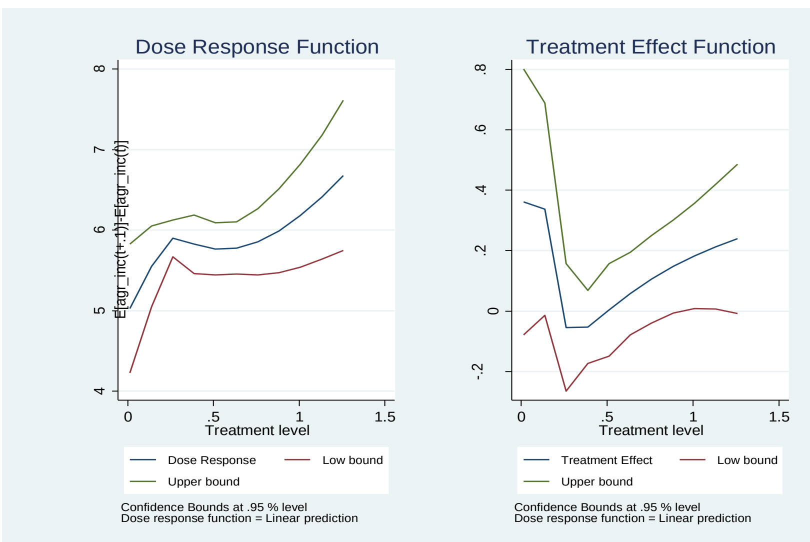
Without disasters



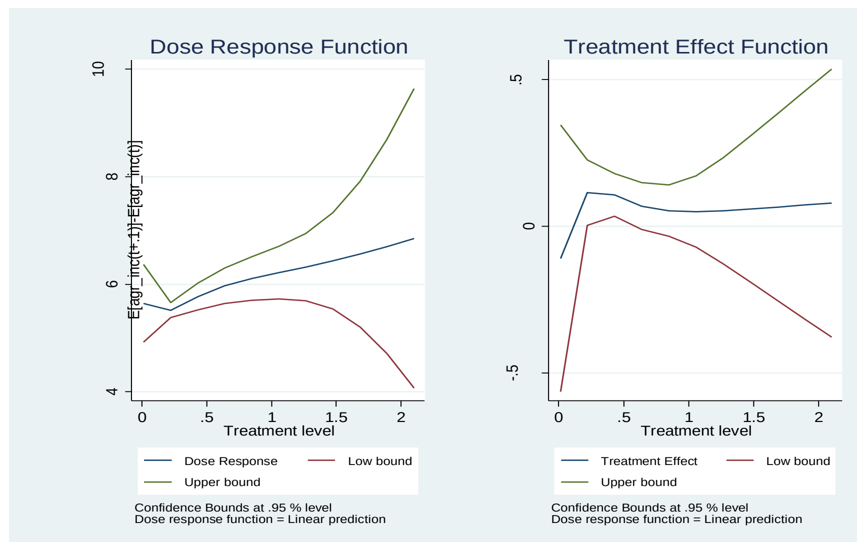
Flood/ Storm/ Landslide



Drought

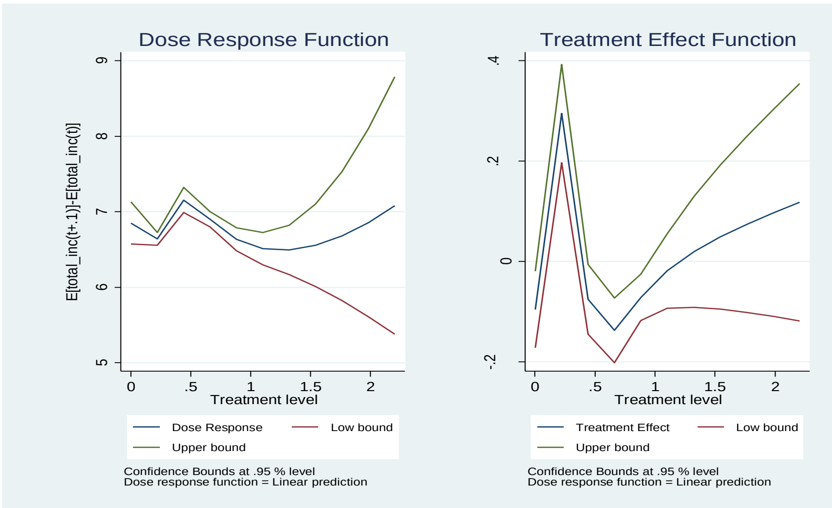


Plant diseases/ Insects

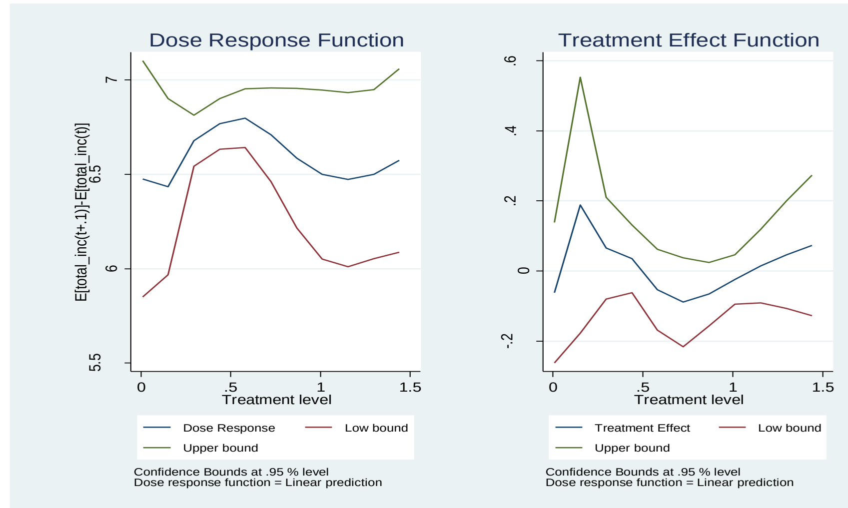


5D: Total income

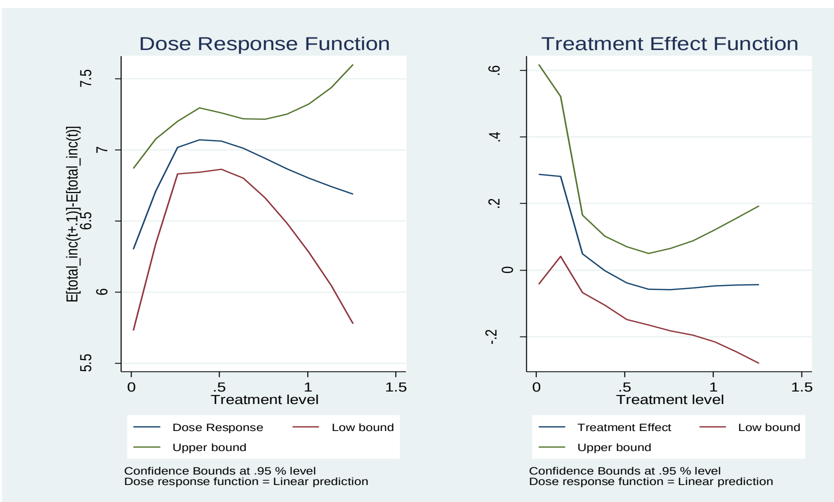
Without disasters



Flood/ Storm/ Landslide



Drought



Plant diseases/ Insects

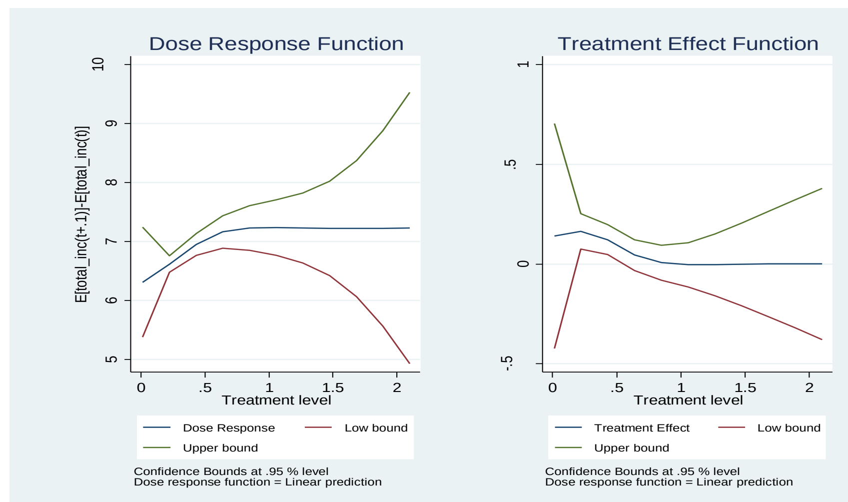
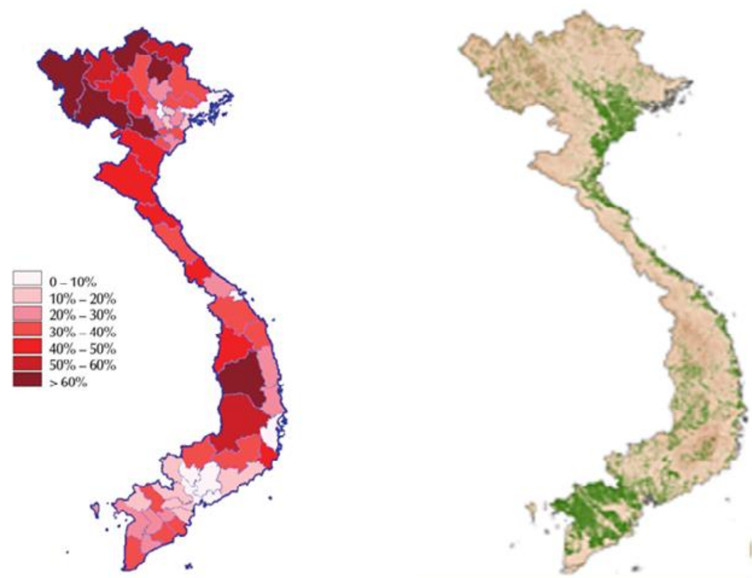
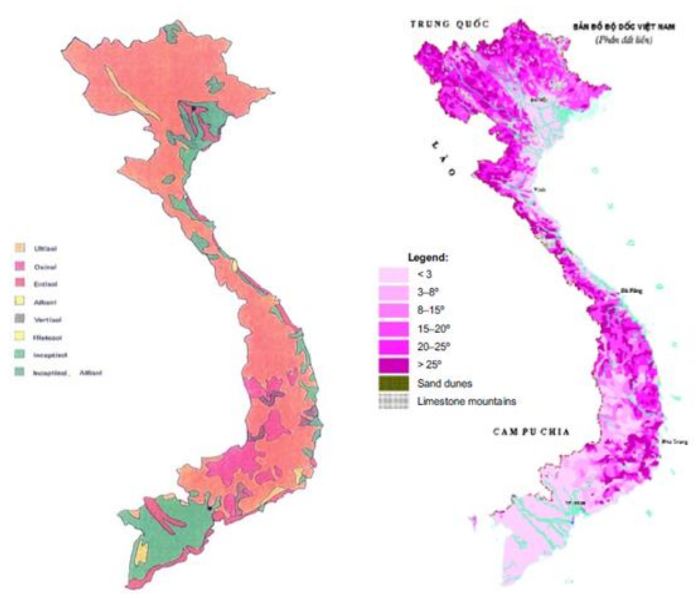


Figure 3.6: Provincial poverty incidence in 2002 (left) and area of paddy cultivation in 2009 (right) of Vietnam



Source: GSO estimates base on the VHLSS 2002 data (left) and Nel Garcia & Andy Nelson, IRRI (right)

Figure 3.7: Map of soils (left) and slopes (right) in Vietnam



Source: K. Hoang, Pham, and Howeler (2000) (left) and V. B. Nguyen, Mutert, and Cong (2003) (right)

Figure 3.8: Maps of temperature and precipitation in Vietnam

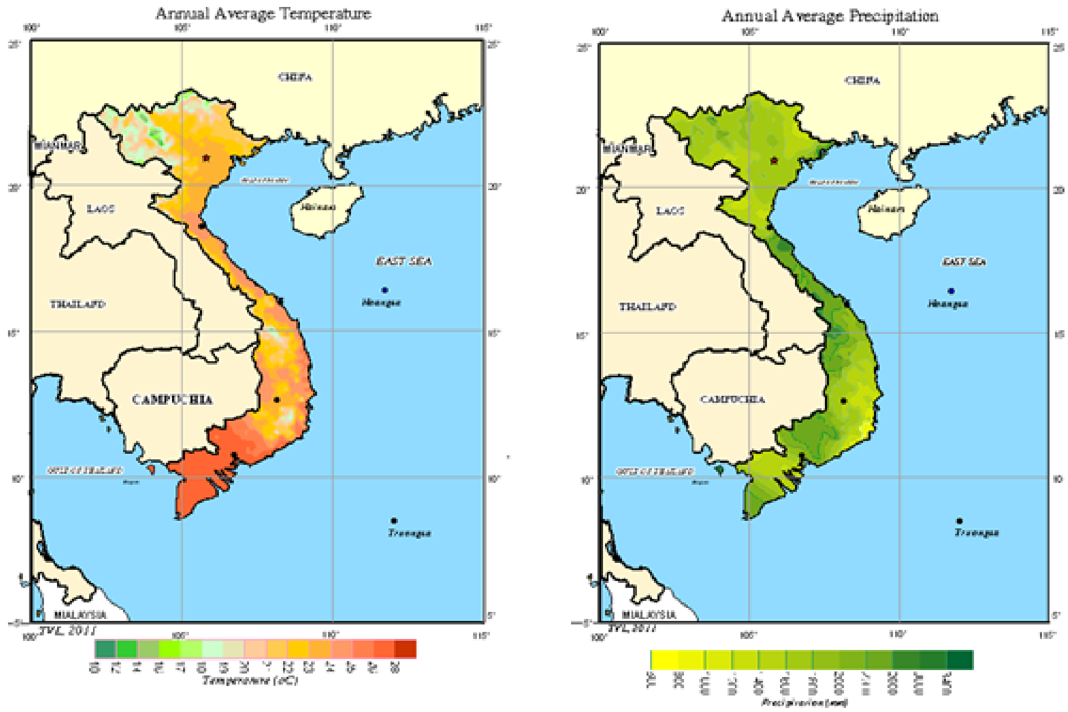
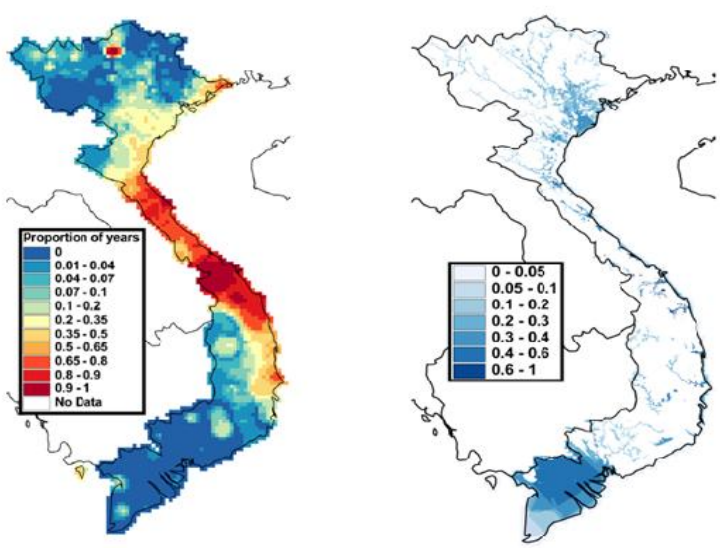


Figure 3.9: Proportion of years in which 300 millimeters of rain fell in a consecutive 5-day period (left) and proportion of riverine and coastal flooding 1985-2007 (right) in Vietnam



Source: Thomas et al. (2010)

Chapter 4

Farmland and child labour

4.1. Introduction

According to the International Labour Organization (ILO), child labour refers to children engaging in any work that is mentally, physically, socially or morally dangerous and harmful to children, and interferes with their schooling. The definition of child labour varies in the literature. Based on the ILO classification, working children generally consist of three groups: children in employment, child labourers and children in hazardous work. The categories of children in employment are the broadest range, containing children in all types of paid and non-paid activities. The categories of child labourers are narrower, which exclude children over 14 years old in hazardous work and include children over 11 engaging in light work (i.e. work that does not influence children's ability to attend school and vocational training). Children in hazardous work include those who are involved in any work that may harm their safety, moral development, and physical and mental health (Diallo, Hagemann, Etienne, Gurbuzer, & Mehran, 2010). The paper uses child labourer and child worker interchangeably.

The ILO estimates that around 176 million children between the ages of 5 and 14 were engaged in economic activities in 2008, which accounts for 14.5 percent (Diallo et al., 2010). Categorised into four groups of economic activity – agriculture, industry, service and other uncategorised activities – the data show that child labour in agriculture accounts for 60 percent of children 5 to 17 years old, and is the highest rate among other sectors. The service and industrial sectors and the uncategorised group, respectively, follow with 26 percent, 7 percent and 7.5 percent. Two-thirds of child labourers are engaged in unpaid family activities, with 64 percent for boys and 73 percent for girls. Paid employment and self-employment account for 21 percent and 5 percent in the same age group. Thus, the majority of child labourers are predominantly employed in agricultural activities on family farms.

Over the past two decades, studies on child labour have increasingly gained attention in both theoretical and empirical directions. The main issues concentrate on addressing three relevant areas, including the causes, the consequences and the policies of child labour. Understanding the causes of child labour plays an important role in designing appropriate policies for governing and reducing child labour. In terms of causes, there are five major groups of factors that have been studied and found to be associated with child labour: poverty, credit market imperfections, land and labour market imperfections, parental characteristics, and macroeconomic factors (Fors, 2010). Among these factors, poverty has been analysed widely and found as a dominant cause of child labour. The effects of poverty that are correlated with imperfect markets, household characteristics and other external factors generate a complicated mechanism affecting child labour. Most empirical findings show a positive relationship between poverty and child labour. However, the results are diverse depending on differences in not only the sample characteristics but also the definitions and methodologies.

Child labour prevalently occurs in rural and agricultural settings where land is the primary source of wealth in agriculture and for farm households. Farmland has been taken into consideration as another factor that may affect child labour. Additionally, in developing countries, the economies are commonly characterised by the presence of imperfections in the goods, financial and labour markets. The decision of employing children may also be affected by the situation of the labour market and credit market. This study therefore focuses on investigating the impacts of farmland on child labour in the context of an agrarian economy.

Taking the case study of Vietnam, child labour has been studied in several papers using data from the 1990s. Figure 4.1 graphs the evolution of poverty and child labour in Vietnam from 1993 to 2008. Both poverty and child labour rates remarkably decreased over this period. In particular, poverty declined from 58 percent to 14.5 percent, while child labour declined from 45 percent to 10.5 percent. However, the progress of child labour reduction was less significant over the period of 2002-2008, reaching only around 2 percent. Intuitively, a question arises whether farm households had any incentives or constraints to employ their children, which resulted in poorer progress in child labour reduction and better progress in poverty reduction.

Moreover, around 70 percent of the population in Vietnam live in rural areas. Associated with the characteristics of the agricultural economy, rural households mainly depend on income sources from agriculture. According to a report by the World Bank in 2006, about 80 percent of rural households are directly or indirectly dependent on agriculture; about one-third of rural households are wage earners with only 4 percent being full-time labourers; and about 34 percent of households engage in non-agricultural business, with only 3 percent in full-time work (B. T. Hoang, 2009). We observe that, although the economic activities are diversifying, the rural households still prevalently work on their farms. The low rate of full-time workers reflects the immobility in the labour market.

Given this context, the paper aims to examine whether the farm size of rural households has an effect on child labour. This paper is organised into seven sections. The following section, section 2, reviews the major trends of child labour in literature. Section 3 develops the theoretical model of farm households. Section 4 describes survey and descriptive statistics. Section 5 introduces econometric specifications. Section 6 reports the empirical results. Finally, section 7 discusses the concluding remarks.

4.2. Literature review

This section briefly introduces the relevant directions in the literature about child labour. We divide it into three directions, i.e. the impacts of poverty, farmland, and both farmland and market failures on child labour. Then we present a literature review on child labour in Vietnam.

In general

First, considered as the primary cause of child labour, poverty has been substantially focused on in the majority of the literature. Basu and Van (1998) contribute a fundamental framework on child labour. The authors proposed an assumption that child labour is caused by subsistence poverty. Basu (2000) extends the model to analyse the effect of the minimum adult wage on child labour. The study demonstrated that child labour incidence may fall or rise as the adult wage is raised by the minimum law. A number of empirical studies show the relationship between living standards and child labour. Krueger (1996) presents an obvious trend in a cross-country sample that employment of young children is common in low-income countries and uncommon in high-income countries. Kambhampati and Rajan (2005) find a relationship in India that an increase in parents' wage reduces child labour, but the effects of the mother's wage and father's wage are different. Beegle, Dehejia and Gatti (2006) show that crop shocks are significantly related to household wealth and found a strong effect, that an unexpectedly poor harvest led to an increase in the level of child labour. Duryea, Lam, and Levison (2007) find that an unemployment shock of male household heads increases the probability of children entering the labour force in urban Brazil.

Second, related to farmland, Bar and Basu (2009) use an overlapping generation model to investigate the effect of landholding on child labour. Under the assumption of an imperfect labour market, the authors revealed that landholding has a positive effect on child labour in the short run but negative in the long run. However, in the short run, there exists a critical level below which landholding will cause a decrease in child labour. Basu, Das, and Dutta (2010) propose a similar model that exhibits an inverted-U curve relationship between land size and child labour. The framework is supported by the significant relationship of the data set in Northern India. The existence of a turning point shows that land size increases child hours worked if it is below such a level of land size, and then decreases the hours worked if the land size continues to increase over that level.

Third, regarding farmland and market failures, a paper by Bhalotra and Heady (2003) has been widely referred to in the literature recently. This study is motivated by the remarkable observation that children of land-rich households are often more likely to be employed than the children of land-poor households. The authors call it a “wealth paradox” and challenge the assumption that poverty is the primary cause of child labour. The paper suggests that this paradox can be explained by imperfect land and labour markets. An imperfect credit market is likely to weaken the effect of this paradox. Based on an analysis of these effects, a model was constructed and applied to a survey in rural Pakistan and Ghana. The findings show that a wealth paradox exists for girls in both countries, whereas it is not statistically significant for boys after conditioning on other covariates. Dumas (2013) looks into the relationship between market imperfections and child labour. The study developed a rural household model and found that, on average, market imperfections increase child labour but the effects are heterogeneous by land ownership using a sample in Madagascar. Households with medium-sized plots may increase child labour when the labour market improves, contrary to other sizes (small or large size). Similar evidence was also found in rural India (Congdon Fors, 2007) and Burkina Faso (Dumas, 2007).

In Vietnam

In Vietnam, child labour has been investigated in relation to household characteristics and external factors after its economic reforms. During the 1990s, child labour in Vietnam declined sharply. Several studies used the household living standard surveys (VLSS) in 1992-93 and 1997-98 to analyse the determinants of this decline.

Firstly, child labour is found to be correlated with living standards. Using nonparametric decomposition on panel data in the period of 1993-1998, Edmonds (2001) shows that an increase in per capita expenditure can explain the 59 percent decline in the incidence of child labour. This rate is highest in the group of households moving out of poverty, at 80 percent. By applying the multinomial logit equation and altruistic model, Rosati and Tzannatos (2006) find that household income had a negative impact on child labour in both the 1993 and 1998 surveys and the relationship is non-linear. The study examines for groups of child labour across schooling status (with and without schooling). The extent of the effect tends to diminish for higher income households. This impact was also revealed in the Vietnam household living standards survey (VHLSS) in 2006 by using the bivariate probit model (UCW Project, 2009).

Secondly, other household characteristics are also examined. Edmonds and Turk (2002) further analyze impacts on the decline of child labour using parametric and non-parametric methods. Child labour in rural areas and ethnic minority groups declined at a slower pace than in urban areas. Children in households owning businesses are likely to work more than children in households without a business. In terms of migration, child labour in households where the head had moved decreased less than in households where the head had never moved. In addition, the education of the parents significantly influences the probability of working, especially the education of the mother (Rosati and Tzannatos, 2006; UCW Project, 2009). Children of highly educated parents are less likely to work.

Therefore, previous studies on child labour in Vietnam have left room for investigating the impact of land and market failures. Especially over the previous decade, child labour reduction has been going at a slower pace compared to the 1990s. We will attempt to develop a rural household model and apply appropriate econometric techniques to test the impact of land on child labour using a survey carried out in 2008. Although the model focuses on the land factor and does not take imperfect markets into account, we still show evidence of an incomplete labour market in this sample.

4.3. Theoretical framework

We suppose that farm household has one adult and one child. The utility function of household is:

$$u = u(x, e) = u(x - c(e)) \qquad (1)$$

where x is total household consumption, $e \in [0,1]$ is the amount of children's work (i.e. take value 0 if children do not work and value 1 if children work with all effort). And $c(e)$ is cost function of child labour with $c_e > 0$ and $c_{ee} > 0$ (and where subscripts indicate derivatives), because the cost of child labour rises in the quantity of supplied. Let assume adult always work regardless of the wage or leisure and his labour input is denoted as ℓ . First consider the special case in which the two types of labor are identical the production function becomes $f(k, \ell + e)$ where k is the land holding. We assume $f_{\ell\ell} < 0$ and $f_{k\ell} > 0$. The household maximizes

$$u = f(k, \ell + e) - w(1 - \ell) - c(e) \qquad (2)$$

subject to $e \geq 0$, i.e. the child cannot take off-farm work.

The first order conditions for an internal solution are

$$f_\ell = f_e = w \qquad (3)$$

An internal solution with respect to ℓ in ensured by the Inada condition $\lim_{\ell \to 0} f_\ell(k, \ell) = \infty$. An internal solution with respect to the child's labor input is ensured by the condition $c_e(0) < w$, i.e. the cost of the child undertaking a small amount of work is lower than the wage the parent can obtain (or pay) in the labor market. If this is not satisfied and $c_e(0) \geq w$ the child will not work.

Differentiation with respect to the land holding gives

$$f_{k\ell} + f_{\ell\ell} \frac{\partial e}{\partial k} = c_{ee} \qquad (4)$$

implying

$$\frac{\partial e}{\partial k} = \frac{c_{ee} - f_{k\ell}}{f_{\ell\ell}} \qquad (5)$$

The sign is ambiguous and depends on whether increased land holding increases the marginal product of labor (and hence of the child's labor) by more or less than it increases the marginal cost associated with the child's labor.

In the more general case, the farm production function is $f(k,\ell,e)$. We assume $f_k, f_\ell, f_e > 0$, $f_{kk}, f_{\ell\ell}, f_{ee} < 0$, $f_{k\ell}, f_{ke} > 0$ (the marginal products of each labor type is increasing in the land holding) and $f_{\ell e} > 0$ (the child and the parent are complementary factors). We also assume $f_{\ell\ell} f_{ee} - f_{\ell e}^2 > 0$ to guarantee concavity of the production function.

We hold land fixed. The first order conditions for the two labor inputs give

$$\begin{aligned} f_\ell &= w \\ f_e &= c_e \end{aligned} \qquad (6)$$

Now vary the land holding. We get

$$\begin{aligned} f_{k\ell} + f_{\ell\ell}\frac{\partial \ell}{\partial k} + f_{\ell e}\frac{\partial e}{\partial k} &= 0 \\ f_{ke} + f_{\ell e}\frac{\partial \ell}{\partial k} + f_{ee}\frac{\partial e}{\partial k} &= c_{ee} \end{aligned} \qquad (7)$$

which we may rewrite as

$$\begin{pmatrix} f_{\ell\ell} & f_{\ell e} \\ f_{\ell e} & f_{ee} \end{pmatrix} \begin{pmatrix} \dfrac{\partial \ell}{\partial k} \\ \dfrac{\partial e}{\partial k} \end{pmatrix} = \begin{pmatrix} -f_{k\ell} \\ c_{ee} - f_{ke} \end{pmatrix} \qquad (8)$$

It follows that

$$\begin{pmatrix} \dfrac{\partial \ell}{\partial k} \\ \dfrac{\partial e}{\partial k} \end{pmatrix} = \frac{1}{f_{\ell\ell} f_{ee} - f_{\ell e}^2} \begin{pmatrix} f_{ee} & -f_{\ell e} \\ -f_{\ell e} & f_{\ell\ell} \end{pmatrix} \begin{pmatrix} -f_{k\ell} \\ c_{ee} - f_{ke} \end{pmatrix} \qquad (9)$$

The impact of the increased land holding on the child's labor input is therefore

$$\frac{\partial e}{\partial k} = -\frac{f_{ee} f_{k\ell} + (c_{ee} - f_{ke}) f_{\ell e}}{f_{\ell\ell} f_{ee} - f_{\ell e}^2} \qquad (10)$$

The denominator of this expression is positive by assumption. In the numerator, the first term $-f_{ee}f_{k\ell} > 0$. The second term $-(c_{ee} - f_{ke})\, f_{\ell e}$ may be either positive or negative. Child labor may therefore either increase or decrease with land-holding. It will be negative if c_{ee} is sufficiently high, i.e. if the costs associated with child labor increase sufficiently fast. The result is similar to that obtained in the simple case.

4.4. Survey and descriptive statistics

There are two important household surveys: Multi-Purpose Household Survey (MPHS) and Vietnam Living Standard Survey (VLSS) conducted in the 1990s. Since the two surveys are partly overlapped, General Statistics Office of Vietnam decided to merge MPHS and VLSS to become a new Vietnam Household Living Standard Survey (VHLSS). The VHLSS was implemented biennially during the period 2000-2010 with technical assistance from UNDP and the World Bank. This paper uses the 2008 VHLSS which contains more detailed information on farm land. There are two modules of questionnaire: short household questionnaire (including income information) and long household questionnaire (including income and consumption expenditure information) in all waves of the VHLSS. The study employs the long questionnaire.

Based on the relevant documents and manual instruction of Phung & Nguyen (2006) and GSO (2008), the survey could be briefly summarized as the following. The interviews in each survey were conducted from May to November in each year. Sampling was at three levels: communes/wards at the first stage, census enumerate on areas (EA) at the second stage and households at the third stage. At the first stage, the sample was selected from the master frame designed for four waves of the VHLSS in this period which included 3,063 communes/wards from 1999 Population Census. At the second stage, wards and communes were partitioned into EAs and three EAs in communes/wards selected. Only one EA constitutes for each wave of survey and the two others are used for the sequential rotated waves. At the third stage, a sample of households was selected systematically with twenty households in each rural EA and ten households in each urban EA. This is technically a three-stage design (including the selection of households), but it is operationally equivalent to a two-stage design since only one EA is selected within each commune for each wave of survey. The sample is rotated 50 percent from one wave to the successive wave of the VHLSS based on the master sample. More specifically, the current survey keeps 50 percent of households in the previous survey, and randomly selects another 50 percent of households from EAs which are different to the ones used in the previous survey, as mentioned at the second stage.

4.4.1. Basic information in the survey

Table 4.1 reports the basic information in the survey. The communes/wards and EAs are randomly selected with the probability proportional to size. Based on the ratio of the urban population to the rural population, the sample size in rural areas is around three times larger than that in urban areas. Out of a total of 9,189 households, the urban and rural households account for 26 percent and 74 percent. The number of households in the sample accounts for only 0.05 percent of the country. The ratio of households between the two areas is almost equal to the ratio of the population distributed in the 1999 Population Census, with around 23.7 percent in urban areas and 76.3 percent in rural areas. The household size in the survey is 4.12 and the household size in rural areas is slightly larger than that in urban areas'. The education levels of the household heads are divided into five groups, including without education, primary, lower-secondary, upper-secondary, and college and beyond. Around 41 percent of the household heads had attended a lower-secondary school, making up the largest fraction. This is followed by primary, upper-secondary, without education, and college and beyond, at around 28 percent, 19 percent, 7 percent and 5 percent, respectively. The data show a noticeable gap in the education of the household heads between the two areas, as the proportion with upper-secondary level and college and above level are higher in urban areas. Female headship, which could affect household expenditure, and children's investment should also be tracked (Joshi, 2004). The female headship is higher in urban households than in rural households, approximately 38 percent and 21 percent.

The per capita total expenditure and living expenditure in the urban areas are nearly twice as high as that in the rural areas. The poverty rate is measured by the percentage of households having an average income or expenditure per capita under the poverty line. There are two methods to compute poverty in Vietnam. The first method is based on international standards that were developed by the General Statistics Office (GSO) with support from the World Bank.[21] The second method relies on the poverty line of the Ministry of Labour, Invalids and Social Affairs. We compute the poverty rate based on the poverty line of GSO. The poverty rate of the whole survey is 17 percent, while the rate of rural areas is 22 percent and of urban areas is only 3 percent.

As in other developing countries, a large share of the labour force is engaged in agricultural activities, especially in rural areas. In the survey, 69 percent of households are involved in

[21] The poverty line developed by GSO and the World Bank is VND 280.000 (2008 Statistical Yearbook of Vietnam, 2009).

agricultural, forestry, aquaculture and fisheries activities.[22] This rate is 87 percent in rural areas and 28 percent in urban areas. These farming activities heavily rely on land, which is a primary factor of production. The ratio of households managing land in farming activities is nearly the same as the ratio of farming households. Almost all farm households report that they use their land for farming activities.[23]

Moreover, Table 4.2 shows more details about land used by households for farming activities. Approximately 23 percent of households use one piece of land, and 26 percent of households use more than five pieces of land.[24] Households report all the pieces of land depending on the features of their management in production at the time of interviewing. These pieces are assigned one of eight categories: annual crops, perennial crops, forestry, water surface, grass field, resident, shifting cultivation and others.[25] Annual crop land is used for growing plants that have a time period from cultivating to harvesting not exceeding one year. Perennial crop land is used for the cultivation of plants with a growing cycle of more than one year, from planting to harvesting. Forestry land is land with natural forest or planted forest that has forest standards. Water surface land is used for growing aquaculture products. Grassland is used to grow grass for cattle. A pond or garden adjacent to the residential land is the area that lies in or surrounds the residential land area of the household. Shifting cultivation land refers to an area that is managed by mountainous households by clearing trees or creating forest fires for cultivation of some crops. Other agricultural land includes areas to build glasshouses and other kinds of development purposes for farming activities. Table 4.3 shows that the majority of land used by households is annual crop land, which accounts for 59 percent on average. Perennial crop land and residential land follows with around 9 percent and 7 percent, respectively.

In Vietnam, the land belongs to all the people, with the State as the representative owner. The State grants land use rights to land users via the form of allocation of land, lease of land and recognition of land use rights for persons currently using the land stably; it regulates the rights and obligations of land users (Vietnam National Assembly, 2003). According to the law on land amended in 2001 by government, the quota on the allocation of agricultural

---

[22] In this paper, we mention agricultural or farming activities, which include agricultural (crops, livestock), forestry, aquaculture and fisheries activities.

[23] The differences may be due to missing data, since there are households with positive outputs of production but with zero land area.

[24] The different pieces of land used by households are attributed by different characteristics, such as type of land (annual crops, forestry, grass field, water surface, etc.), type of irrigation (gravity, pump, etc.), type of ownership (long-term contract, rent, borrow, etc.) and others.

[25] Residential land is still reported here because I find that there are households who use it and generate a low revenue of output, for instance planting vegetables, or households rent it out to others.

land varies with the type of land.[26] The quota on the allocation of land to each household or individual for planting annual crops, for aquaculture and for salt production is no more than 3 hectares (or 7.4 acres) of each type of land. The quota on the allocation of land for perennial crops is no more than 10 hectares (or 24.7 acres) in the plains and no more than 30 hectares (or 74 acres) in the midlands and in mountain regions. However, in this survey, the average area of agricultural land per household is 2.35 acres, which reflects that the production activities of the households are small-scale (Table 4.1).

In the survey, since child labourers under 10 years of age account for a small fraction, the analysis focuses on children between the ages of 10 and 14.[27] In Table 4.1, households having child labourers, out of all the households having children from 10 to 14 years old, are estimated at around 7 percent in urban areas and 15 percent in rural areas. As expected, child labour in rural areas is remarkably higher than in urban areas. When child labour is estimated in terms of the individual (not in terms of households employing child), the rate is lower. Out of the total 11 percent of working children, as high as 82 percent of the children participate in agricultural activities (including 83 percent working for their households), followed by 14 percent in the industrial sector and 4 percent in the service sector.

4.4.2. Land, labour and product markets

In 1986, after the transition, the government decentralised agricultural land from a collective system and assigned it to individuals and households for a period of up to 15 years. In 1993, government implemented a new land law that allowed land-use rights to be transferred, exchanged, leased, mortgaged and inherited. This important reform was expected to bring benefits for landholders, particularly in the agricultural sector. First, the security of land usage is enhanced, which may affect agricultural investment decisions. Second, access to credit could be facilitated if the land served as collateral. Based on the two rounds of the survey, VLSS 1992-93 and 1997-98, the additional land rights were found to increase long-term investment but were irrelevant for credit access (Do & Iyer, 2004). In this survey, the data also reveal similar results. In Table 4.4, the long-term land-use certificate possession of households with loans and those without loans is equal at around 77 percent.

[26] Land allocation by the State means granting land use rights by the State by way of an administrative decision to an entity that has requirements for land use (Vietnam National Assembly, 2003).

[27] According to the labour laws of Vietnam, the minimum age for full-time employment is 18. Workers between 15 and 18 years of age have special provisions. Many children have reportedly worked in violation of this law. The law established working hours as eight hours per day with a mandatory 24-hour rest period per week.

In Vietnam, the procedure for issuing land-use certificates to farmers is time-consuming and proceeded at an unequal pace across the regions. For instance, after three years of the new law taking effect, 72 percent of the farmers in the Mekong delta sample reported that they possessed a certificate, while the Red River delta had only 8 percent (Hare, 2008). Empirical evidence suggests that the certificate contributes rather small in the absence of the appropriate conditions and constitutions. Hence, the land-use certificate seems to be inadequate to represent the actual land property rights and could cause bias among regions. The possession of land could be alternatively distinguished by the information which was provided in the questionnaire: how the household got their plot of land. Depending on the security level of the land, we divide it into two groups: long-term and short-term land. Long-term land includes long-term use allocation, signing a contract, gifted or inherited land, purchased land, proclaimed land and bartered land. The land users in this group are the landowners. Short-term land includes auctioned, rented and borrowed land. This makes sense because the ratio of land-use certificates in the long-term group is more than twice as that in the short-term group, at around 79 percent and 30 percent, as shown in Table 4.4.

Panel A of Table 4.5 shows properties of land use. The majority of land operated in farming activities is long-term land, consisting of 93 percent. Short-term land is around 6 percent, and the rest is uncategorised land. As the assumption in a study by Basu et al. (2010), land is considered immobile. We are aware that farmers still sell land and move, but this is rare and harmless. The high ratio of long-term land in the sample could also make it appropriate to use that assumption.[28] Furthermore, the land is investigated to determine whether each plot is managed by the households or for other purposes, such as renting, lending without payment, in exchange for other plots, fallowing and others. Nearly 91 percent of the land area is managed by the households themselves.

Over the past two decades, rapid urbanisation and industrialisation have been witnessed as the primary attributes of the economic development progress. Consequently, labourers have moved out of the agricultural sector to participate in the industrial and service sectors. Despite the structural changes in the labour force, labour in agriculture still constitutes the largest fraction. In this survey, 55 percent of labourers work in agriculture, forestry and aquaculture, and 65 percent of labourers engage in at least one agricultural activity for their household. As in other developing countries, the labour market is underdeveloped due to a shortage of labour incentives, labour mobility and information. In particular, the labour market in rural areas may be characterised by seasonal labour demand and high transaction

[28] In this survey, around 7 percent of land was attained through purchase.

costs. That salient feature leads to the situation that workers in family enterprises or unpaid family labourers are pervasive in agriculture as well as non-agriculture.

The labour market in Vietnam is not an exception. Panel B displays the labour usage of households involved in farm production or farm service. The work information was retrieved from the questions that asked whether the individuals work for a salary, whether they work for their household's farm or whether they do non-farm work for their household. The household members could work in more than one of these types of jobs. A household belongs to the working category in panel B if there is at least one member that participates in household production or service in agriculture, forestry or fishery. The information on hiring outside labour is computed indirectly from the expenditure of the household's production, which includes an item for hiring outside labour. As predicted, around 84 percent of households only employed their household members, whereas merely 0.2 percent of households hired outside labourers exclusively. Households using both household members and hired labour make up around 12 percent. There were around 4 percent of farm households without family workers or hiring expenses. It could be the case that households rent out land, share land and production, or exchange products for the labour cost. The distribution is similar between the rural and urban areas. The imperfect labour market may contribute to the effects that cause child labour.

Regarding the final use of agricultural output, the households could sell either all or some of their products. The remaining products could be consumed by the household or left for other purposes. We measure this indicator as the percentage of revenue obtained from sold products out of the total revenue. To check the product market and household output use, Panel C shows that the majority of farm households utilise farm output for both purposes: sell in the market and household consumption. This rate is higher in rural areas than in urban areas: at 91 percent and 72 percent, respectively. The ratio of households retaining or selling all of the output is lower in rural areas than in urban areas. As an essential feature of the agrarian economy of developing countries, in the situation of imperfect markets, farm households are typically dependent on agriculture as both consumers and producers. There is no separability between consumption and production.

4.4.3. Child labour situation

As the basic information of the survey in Table 4.1 shows, child labour occurs mostly in rural areas. This analysis therefore focus on child labour in rural areas. Table 4.6 shows child labour in rural areas across industries. Children are involved prevalently in agricultural (crops and livestock) activities, at 90 percent out of the total, followed by aquaculture and silviculture, at only 4 percent and 2 percent, respectively. In the industrial sector, children are engaged in the processing industry and mining industry at around 5 percent and 3 percent, followed by construction and utility production and distribution. A small proportion of child labour appears in the service sector at merely 2 percent, including hotels and restaurants, finance and credit, and public and personal service. These figures reflect the typical traits of the economy in rural areas, in which agriculture is a dominant source of livelihood and self-employed households.

Table 4.7 summarises the distribution of child labour in relation with other main factors that may influence child labour in rural areas. Child labourers are observed in two categories of work: farm work and non-farm work. Child labourers of 10-14 years old engaging in farm work is around 11 percent, whereas those engaging in non-farm work is more than 2 percent.

Children may occasionally work, for example, during periods of harvest, summer holidays, etc. The questionnaire records the number of hours worked per day, days worked per week and weeks worked during the last 12 months. Based on these numbers, the working hours are calculated using the average per week from the previous 12 months. The data show that children who are employed in farm work work less hours than those in non-farm work, around 11 hours and 21 hours correspondingly. In terms of days per year, children are involved in farm work for 118 days and non-farm work 180 days. Thus, the probability of being employed in farm work is higher than in non-farm work, but with less hours worked.

Although the analysis does not focus on the relationship between schooling and child labour, we will still look at that common phenomenon of whether schooling has a negative effect on child labour. In Table 4.7, child labour in farm work and non-farm work is substantially higher in the group of children without schooling. So children are more likely to be employed and work longer if they do not go to school.

To check the distribution of child labour over household living expenditure, the sample is divided into three groups at two cut-off points: 25th and 50th percentile. Child labour in farm work clearly decreases in richer households in the three groups defined by living expenditure. However, child labour in non-farm work in the richest group is slightly higher than that in the middle group. In terms of hours worked, children in the richest households (both farm work and non-farm work) work more than those in the poorest households but less than those in the households in the middle group.

To observe the relationship between child labour and land size, we split the per-capita managed land by percentiles in the group of households who have children aged 10 to 14 years old. The distribution of land by percentiles partly reflects that the agricultural sector in Vietnam is characterised by small-scale farm households. We divide this roughly into three groups at two breaking points at the 50^{th} percentile and 80^{th} percentile, which correspond to an average land area of 0.23 and 0.68 acres. In the farm work group, the figures suggest that child labour increases if the farmland is larger. Children also work longer in larger farmlands. This trend is opposite to that for the non-farm work group. The child labour rate for doing non-farm work is highest for the group of households with the smallest size. Apparently, this is because farm work is dependent on the land.

Another important asset of agricultural activities is cattle. This factor may affect child labour in rural areas as well. The cattle information in the survey includes three types: drafting, ploughing and breeding cattle; breeding pigs; and basic herds of poultry and cattle. These types are normalised to be the same unit as the number of cattle. The cattle per capita is also divided into two groups at the 50^{th} percentile point. Child labour in the group over the 50^{th} percentile is higher than the group below the 50^{th} percentile, at around 20 percent and 9 percent. Nonetheless, the hours worked by child labourers in the higher group is one hour and a half less than in the group below the point.

The data suggest the potential impact of land use for farming activities on child labour. Child labour in the group of households employing only their household members is higher than child labour in the group of households who hired outside labour. In contrast, the hours worked is slightly lower in the group of households who employed only their household members. Households with outside labour account for around 10 percent of the total. As shown in Table 4.7, the majority of households use their products both for household consumption and to sell in the market. In Table 4.8, child labour incidence appears to decrease when the ratio of sold output increases. However, the data appear to be in the opposite direction for the hours worked. Children in households who sell more are likely to work longer.

Capturing an overview of these main factors, the specifications of the model could be appropriately identified to investigate the effects on child labour. The descriptive statistics initially reflect a quite clear impact of farmland on the possibility of being employed but still an ambiguous impact on the hours worked.

4.4.4. Variable summary

To examine the impact on child labour in rural areas, the model includes the variables of the child, household characteristics and regional indicator, which are shown in Table 4.8. The gender of the child labourers is fairly equal, although the empirical evidence reports mixed results on gender. The existence of gender bias may be derived from parental gender bias, labour market discrimination, the type of work and differences in the return of human capital. Edmonds and Pavcnik (2005) examined data for 36 countries and found that child labour for girls is higher than for boys when including market work and domestic work; but when the data on domestic work is excluded, child labour for girls is less than for boys. According to a report by Diallo et al. (2010) in 2008, there were 176 million boys in economic activity compared to 130 million girls. It is expected that child work is affected by the availability of labour and the household size. Larger households and more adult labourers are expected to decrease child labour. The education level of the parents or head affects the decision of child labour. An increase in parental education has a direct positive effect on their children's education and job opportunities, resulting in a decrease in child work. The gender of the head represents the bargaining power in the household. Table 4.8 shows the gap of child work, especially in the ratio of participation across the regions. Therefore, the regional factor is included to capture different demographic characteristics, conditions and economic levels. There are eight socio-economic regions in Vietnam, including Red River Delta, North East, North West, North Central Coast, South Central Coast, Central Highlands, South East and Mekong River Delta.

4.5. Econometric specifications

This study's main objective is to examine the effects of farm land on child labour. Since the unemployed children account for a larger share of the sample, almost all values of the dependent variable are equal to zero. The linear regression OLS produces biased and inconsistent estimates when a significant ratio of dependent variable's value is equal to zero. The Tobit model is a traditional approach to deal with data with many zeros, which estimates the relationship between non-negative dependent variable and an independent variable (Tobin, 1958). The marginal effect of an independent variable is treated conditional on limited fraction of dependent variable which takes positive values as in this case.

There is an argument pointing out an inadequacy of the Tobit model in solving that problem. In the Tobit model, the choice of being censored (participation) and expected value conditional on un-censored (level of participation) are determined by the same factors. The model considers dependent variable to be censored at zero but ignores the source of zeros, in which could be caused by deliberate household's decision or certain circumstances (e.g. financial conditions, characteristics of demographic (Newman et al, 2003; Martinez-Espineira, 2006).

Heckman (1979) proposes the two-stage estimation procedure to deal with zero observation. The author pointed out that estimation on selected subsample results in selection bias. The first stage is Probit estimation and the second stage is censoring estimation on selected subsample. In other words, the first stage estimates the probability of observed positive outcome or participated decision. The second stage estimates the level of participation conditional on observed positive values. The model assumes that these two stages are affected by different sets of independent variables, which is contrast to Tobit model. Another extended point in the Heckit model is that all zero observations are assumed to be derived from respondent's deliberate choices.

Cragg (1971) proposes the double-hurdle model, which is developed from Tobit model and Heckit model. The double-hurdle and Heckit models are similar in building two stages of decision. The first hurdle refers to the participation decision and the second hurdle refers to the level of participation decision. Both models are allowed to be affected by different sets of explanatory variables. However, the Heckit model assumes there is no zero observation in second stage of decision, whereas the double-hurdle model permits potential zero values in the second hurdle which appear from deliberate choice or random circumstances.

In this paper, the first stage or the first hurdle is a decision on whether a child is employed, and the second stage is decision on how many hours that child engages in. According to the assumption of the Heckit model, all the observed observations are positive in second stages However, in the double-hurdle model, there are zero observations which have potential positive hours worked. The zero working hours reported may be due to imperfect labour market.

Both Heckit and double-hurdle models have been used widely in previous empirical studies, mainly in consumption decision and labour supply. The two-stage Heckman model was used in demand analysis of fish in Cheng and Capps (1988), habit of consuming seventeen goods in Heien and Durham (1991), etc. The Tobit model has been applied in labour supply by Blundell and Meghir (1987), Blundell et al (1987, 1998); in household consumption by Deaton and Irish (1984), Yen and Jones (1997), Burton et al (1996); or in loan default analysis by Moffatt (2005), land investment by Bekele and Mekenon (2010), etc.

4.5.1. The standard Tobit model

The model is defined as:

$$h_i^* = x_{1i}\,\beta_1 + v_i \;; \quad v_i \sim N(0,\sigma^2) \qquad (11)$$

$$h_i = \begin{cases} h_i^* \; if \; h_i^* > 0 \\ 0 \; otherwise \end{cases} \qquad (12)$$

where h_i^* is a latent unobserved endogenous variable which presents the optimal working hours; h_i is the corresponding observed variable which measures actual hours worked; x_{1i} and β_1 are vectors of independent variables and their parameters, respectively; v_i is a homoskedastic and normally distributed error term. The condition (12) implies that the observed number of hours are positive continuous if the positive number of hours are desired, and zero otherwise. Due to the non-negative values of hours worked, dependent variable h_i is censored at zero. This means that the observed zero on the dependent variable can be either "true" zero (i.e. individual deliberate choice) or censored zero (i.e. data collection methods, certain circumstances). Using maximum likelihood method, the likelihood function of standard Tobit is:

$$lnL = \sum_{0} \ln\,[1 - \emptyset(\frac{x_{1i}\,\beta_1}{\sigma})] + \sum_{+} \ln\,[\frac{1}{\sigma}\,\varphi(\frac{h_i - x_{1i}\,\beta_1}{\sigma})] \qquad (13)$$

Where "0" denotes the zero observations (hours worked h_i is zero) in the sample and "+" indicates the positive observations (hours worked h_i is positive); $\emptyset(.)$ and $\varphi(.)$ denotes standard normal cumulative distribution function and standard normal probability density function, respectively.

4.5.2. The generalized Tobit or Heckit

As the argument above, Heckman (1979) propose the two-stage estimation method to correct for the selection bias. The first step estimates the participation decision and the second step estimates for level of participation. According to Heckman (1979) and Flood and Gråsjö (1998), the Tobit model is modified as:

The participation decision:

$$d_i^* = x_{2i}\,\beta_2 + u_i\ ; \qquad u_i \sim N(0,1) \qquad (14)$$

$$d_i = \begin{cases} 1\ if\quad d_i^* > 0 \\ 0\ otherwise \end{cases} \qquad (15)$$

The level of participation decision:

$$h_i^* = x_{1i}\,\beta_1 + v_i\ ; \qquad v_i \sim N(0,\sigma^2) \qquad (16)$$

$$h_i = \begin{cases} h_i^*\ if\ d_i = 1 \\ 0\ \ if\ d_i = 0 \end{cases} \qquad (17)$$

In this model, x_{1i} and x_{2i} are vectors of explanatory variables in two stages of decision. Hence, the model assumes that the decisions of participation and level of participation are affected by separated sets of factors. As in Tobit model, β_1 and β_2 are corresponding vectors of parameters; d_i^* is a latent variable that denotes binary censoring; d_i is the observed value representing the participation decision. The observed number of hours worked equals to unobserved latent value when a positive number of hours worked is reported; otherwise it takes the value zero. The error terms u_i and v_i are assumed to be independently distributed. This assumption implies that there is no relationship between the two stages of decision.

However, Heckman (1979) assumes that the two error terms are correlated and the first stage dominates the second one. Thus the error terms follow the bivariate normal distribution:

$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix}\right] \qquad (18)$$

Where ρ is correlation coefficient of the error terms. The domination assumption means that if the child reports positive hours worked, this is the intentional purpose of their parents. In other words, the participation is a deliberate choice. Then the model is estimated by Probit for the decision on participation and standard OLS for the positive hours worked. The log-likelihood function for this approach is:

$$lnL = \sum_0 \ln\left[1 - \emptyset\left(\frac{x_{2i}\beta_2}{\sigma}\right)\right] + \sum_+ \ln\left[\emptyset\left(\frac{x_{2i}\beta_2 + \frac{\rho}{\sigma}(h_i - x_{1i}\beta_1)}{\sqrt{1-\rho^2}}\right) \frac{1}{\sigma}\ \varphi\left(\frac{h_i - x_{1i}\beta_1}{\sigma}\right)\right] \quad (19)$$

If the error terms are independent, ρ=0, the equation (23) is simplified as:

$$lnL = \sum_0 \ln\left[1 - \emptyset\left(\frac{x_{2i}\beta_2}{\sigma}\right)\right] + \sum_+ \ln\left[\emptyset(x_{2i}\beta_2)\ \frac{1}{\sigma}\ \varphi\left(\frac{h_i - x_{1i}\beta_1}{\sigma}\right)\right] \qquad (20)$$

4.5.3. The double-hurdle model

The double-hurdle extends the standard Tobit and Heckit models to overcome the zero hours worked. This model is similar as the Heckit model, but there is a slight modification in the equation (17) as following:

$$h_i = \begin{cases} h_i^*, & if\ d_i = 1\ and\ h_i^* > 0 \\ 0, & if\ d_i = 0 \end{cases} \qquad (21)$$

This equation implies that the observed hours worked can be either censored at zero or by data processing, and other circumstances. Household may be willing to send children to work in the market but there is not any opportunities. In other words, household would send their children to work if there exits the perfect labour market. Assuming the error terms are independent, the log-likelihood function of double-hurdle is expressed as:

$$lnL = \sum_{0} \ln [1 - \emptyset(x_{2i}\,\beta_2)\varphi(\frac{x_{1i}\,\beta_1}{\sigma})] + \sum_{+} \ln [\emptyset(x_{2i}\,\beta_2)\,\frac{1}{\sigma}\varphi(\frac{h_i - x_{1i}\,\beta_1}{\sigma})] \qquad (22)$$

The first term demonstrates for the observations with zero values. It implies that the zero observations are affected by both participation and level of participation decisions. It is contrast to Heckit model which indicates that all zero observations are only from participation decision. The different point is shown by the additional term in equation (22), $\varphi\left(\frac{x_{1i}\,\beta_1}{\sigma}\right)$, which contributes for the effect of possible zero values in second stage decision. The second term in equation (22) expresses the conditional distribution and density function of censoring rule and observed positive values.

Under the assumption of independence between two error terms, the log-likelihood function of the double-hurdle is the summation log-likelihood of Probit model and truncated regression model (McDowell, 2003; Aristei et al, 2007). We estimate the model by maximizing two components separately (John, 1989; McDowell, 2003). Although this can be estimated by the user-written program of Burke (2009) in Stata, but the package cannot compute the marginal effect after each stages.

4.5.4. Test for model appropriateness

To select the more appropriate model, we can employ the likelihood ratio (LR) test. The test uses the log likelihood of the models being compared to test whose parameter estimates are a more proper fit for the data. The LR statistic is defined as two times the log of the ratio of the likelihood functions of two models evaluated at their log likelihood (Greene, 2000). We test the Heckman model against the Tobit model, and the double-hurdle model against Tobit model by comparing each pair of log likelihood values. The log likelihood of double-hurdle model is the sum of the log likelihoods of the truncated regression model and probit model. The LR statistic can be expressed as:

$$LR = 2(log_{Tobit} - log_{Heckman}) \qquad (23)$$

$$LR = 2\left[log_{Tobit} - \left(log_{probit} + log_{truncated}\right)\right] \qquad (24)$$

LR statistic is chi-square distributed with degrees of freedom equal to the difference between the numbers of parameters estimated in the two models. The null hypothesis in the equation (23) is that the Heckman model provides a better fit than the Tobit model, and in the equation (28) is that the double-hurdle model provides a better fit than the Tobit model. The LR statistic is compare with the critical value at a specific significance level.

4.6. Empirical results

This analysis emphasises on evaluating the impact of farmland on child labour. We also track child labour in different types of work and in relation to education. We test the models for three groups of children: those who work on family farms, those that do all kinds of work and those that do on-farm work but without school enrolment. If the results vary across these groups, we could further observe how households allocate differently their children's time endowment, conditional on alternative characteristics of the children and households as well.

The marginal effects of the explanatory variables are evaluated at the sample means based on the three models. The marginal effects are estimated for participation (probability of working) and the level of participation (hours worked). Since the assumption in the Heckman and double-hurdle models allows different explanatory variables in the two stages, several insignificant ones are dropped from the second stage. [29]

Model selection

Table 4.9 shows the result of the likelihood ratio test for three categories of child labour. In the groups of child labourers engaging in all kinds of work, and on-farm work and without schooling, the log likelihoods are fairly equal in the test with the Heckman model and with the double-hurdle model, which are 240 and 214 for the former group and 60 for latter group. For the group of child labourers on-farm work, the log likelihoods is 236 in the test with the Heckman model and is 588 in the test with the double-hurdle model.

Since the critical value of $\chi(19)$ is equal to 36 at significance level of 1 percent, the LR statistics show that the null hypothesis is strongly rejected in all three groups. The test confirms a better fit of the Tobit model over the Heckman model and the double-hurdle model. This indicates that the decision whether to employ child labourers and to what extent to employ child labourers are affected by the same set of factors.

[29] The selection of explanatory variables in the second stage of the Heckman model and the double-hurdle bases on the descriptive statistic of this sample and tries of running regression in an attempt to choose a more appropriate set of factors.

Main variables

As an overview, the results of these models show the identical impacts in terms of the sign of the parameters. The parameters have almost a similar level of significance in the first step of the three models, but less significance in the second step of the Heckman and the double-hurdle models. However, the magnitude of the parameters is less different between the Tobit and the double-hurdle models. The changes in the impact of the three models across the three categories of child labour are relatively consistent.

First of all, we check the effect of farmland. Table 4.10 reports the estimated parameters for child labour in all sorts of work, which includes family farm work, non-family farm work and work for wage. Farmland used in agricultural activities by households has a positive effect on the possibility of being employed but the effect is not statistically significant. The marginal effect of land on child labour decreases, which is shown by the negative parameters of land squared, but the effect is still not statistically significant.

To diminish the endogenity problem of household welfare, we include the interaction term between farmland and poverty, instead of poverty or income variable alone. This variable has statistically significant positive effect in the first stage of decision in three models. The effect is significant in the second stage of decision in the Tobit model only. The result can be understood as the poor household is more likely to employ their children keeping farmland constant; or among the poor households, the land-rich household is more likely to employ their children than the land-poor household. In the Tobit and the double-hurdle models, if households is living in poverty, each additional unit (acres per capita) of farmland increases the probability of child labour by about 4 percent points. This impact is higher in the Heckman model, and the probability of child labour increased by approximately 25 percent points.

Moreover, cattle also appear to have a positive effect on the participation decision of child labour in the Heckman and double-hurdle models. An increase of one unit (number of cattle per capita) of cattle leads to an increase of 4 percent points in the Heckman model and around 0.8 percent points in the double-hurdle model. The interaction term of cattle and poverty is statistically insignificant in the Heckman model and the double-hurdle model, but significant positive in both stages in the Tobit model. So similar to the interaction term of land, if households are poor, a household with more cattle is more likely to employ their children and employ longer than a household with less cattle,

Table 4.11 examines the impact on child labour in the groups of children working only on their household farm. As expected, farmland appears to have a positive and statistically significant on the possibility of child labour. This is understandable that children working on farm are more influenced by their family's farmland than children working on non-farm activities. The effect is around 2 percent points in the Tobit model and the double-hurdle model, and 15 percent points in the Heckman model. Land size also has a statistically significant positive effect on the hours worked in the Tobit model and the double-hurdle model, but not in the Heckman model, which are 44 percent and 219 percent points, correspondingly. The marginal effect of land on child labour declines when the land increases in the second stage of only the double-hurdle model. Again, the interaction term of land and poverty cause positive impact on participation decision of three models, but a positive impact on hours worked of the Tobit model. Similarly, cattle also has a positive impact on the probability of child labour in the Heckman and the double-hurdle models, and the impact of the interaction term with poverty is statistically significant in both stages in only the Tobit model.

Table 4.12 shows the results of the group of children without school enrolment. The effect of farmland on the probability of being employed is positive in the Heckman model and the double-hurdle model, and nearly greater than in the other two categories of the sample in Table 4.10 and Table 4.11, which are around 125 percent points in the Heckman model, and 52 percent points in the double-hurdle model. The impact of land size is statistically insignificant in the Tobit model. Interestingly, the outside labour variable has a negative effect on child labour, in both stages of the Tobit model and in the first stage of the Heckman model and the double-hurdle model. The result indicates that hiring outside labour is associated with a decline in the probability of children working on farms by around 23 percent points in the Tobit model, 84 percent points in the Heckman model and 32 percent points in the double-hurdle model. The impact suggests that there is a trade-off between child labour and outside labour. Hence, this provides informative clues to imply the effect of the labour market, which was introduced in the literature review. Although we do not analyse imperfect market in the theoretical model, this empirical evidence demonstrate a potential relationship of labour market and child labour. A household may have an incentive to use their children's effort on farm rather than hiring outside labour. This could caused by imperfect labour market, the shortage of seasonal labour in rural market, or the moral hazard problem (Bhalotra & Heady, 2003).

Other variables

Along with the main variables of interest in this analysis, we explored other determinants that may also be connected with child labour. Tracking children's characteristics, the results show that significantly more older children are employed than younger children in both stages of the decision in the Tobit model, except in the second stage of the Heckman model the double-hurdle model.[30] The possibility of being employed between boys and girls is statistically indifferent, as the descriptive data shows in the previous section.

In addition, the structure of the household provides clues on the allocation of labour sources within the household that influences child labour. Children in households are categorised by age in order to investigate whether there is a supplement or substitute effect on child labour within the household. The number of children under 10 years old has a positive impact in the three models in the group of children working on farms, except the second stage of the Heckman and the double-hurdle models. Household members from 15 to 19 years of age have statistically insignificant effect on child labour. Therefore, children are likely to be employed and work more hours on farms if their households have younger children. A possible reason is that households that have younger children probably need to employ the older ones to trade-off with the cost of the younger ones.

Regarding the head of household, we test for the educational level and female headship. The education of the household head is tested on four levels: without education, primary, secondary, high school and beyond. The expected negative impact appears to be the highest in the group of household heads with higher education. The effects are strongly and consistently significant for the three groups and in the three models in the first stage, and less significant in the second stage. This implies that children are more likely to be employed if their household head has lower education. In addition, the female headship variable has no effect on child labour.

Furthermore, the regional differences are controlled by the dummy variables of the eight socio-economic regions. The estimated parameters show that children from the North East, North West and North Central Coast are likely to work more than those from other regions. The evidence is in line with the fact that these regions have lower incomes on average compared to the other regions. These regions also have the highest poverty rates over the country. Therefore, we are aware that poverty is still a significant factor affecting child labour, as shown in the descriptive statistics. We, however, exclude the income due to endogeneity. So the regional dummies employed in the model, instead of the income variable, partly capture the household living standard.

[30] I test these excluded variables in the second stage of the Heckman model and the double-hurdle, then decide to drop due to their insignificance.

4.7. Conclusion

This paper used the sub-sample of rural areas and for children from 10 to 14 years old from the 2008 Vietnam household living standard survey. Although the preliminary data showed child labour occurs in low-income households, the study aimed to test another factor, farmland, which may potentially affect child labour. The results support the hypothesis that child labour on family farm increases in land-rich households and decreases in land-poor households. Although the theoretical model does not take into account imperfect markets, the evidence shows a substitute between outside labour and domestic child labour in the group of children without school enrolment.

The hypothesis was investigated by comparing across three models, where the dependent variables were examined for the two stages of decision-making, including the possibility of participation and the extent of participation. Then the effects were decomposed into two stages of decision: probability to participate in economic activities and the level of participation. Livestock also plays a similar role in causing child labour. This reflects that productive assets are crucial determinants affecting child labour, together with the poverty factor.

In short, these findings demonstrated the effects of land as another important indicator that causes child labour, which supports the theoretical framework. Furthermore, the findings suggest the implication that a mechanism should be operated for both poverty and child labour reduction. In the case of the credit subsidy policy or poverty alleviation programme, poor households receive money and invest in the agrarian sector. On the one hand, the programme may increase living standards and decrease poverty, but on the other hand, it may also increase child labour. With that concern, policies should be implemented along with other programmes that eliminate the constraints of imperfect markets and improve the knowledge of parents about the negative effects of child labour.

By investigating children in different groups, the findings showed a critical point that the effect of labour market on children without schooling. Hence, child labour also has a negative effect on their education. To improve children’s education, it is worth paying attention to the child and adult labour markets problem as well.

References

Abbott, C. P. , Hurt, C. , & Tyner, W. E. . (2009). What's driving food price? *Farm Foundation Issue Report*. Oak Brook, IL, USA: Farm Foundation.

Aksoy, M. A., & Dikmelik, A. I. (2007). The Role of Services in Rural Income: The Case of Vietnam. *The World Bank Research Working Paper 4180*.

Aksoy, M. A., & Isik-Dikmelik, A. (2008). Are low food prices pro-poor? Net food buyers and sellers in low-income countries *Policy Research Working Paper 4642*: The World Bank.

Arndt, C., Garcia, A. , Pham, H.H. , & Thurlow, J. . (2009). *A Provisional 2007 Social Accounting Matrix for Vietnam*. Department of Economics, University of Copenhagen, and Central Institute of Economic Management: Hanoi.

Ball, R., & Pounder, L. (1996). "Efficient but Poor" Revisited. *Economic Development and Cultural Change, 14*(4), 735-760.

Bar, T., & Basu, K. . (2009). Children, education, labor, and land: in the long run and short run. *Journal of the European Economic Association, 7*(2-3), 487-497.

Basu, Kaushik. (2000). The intriguing relation between adult minimum wage and child labour. *The Economic Journal, 110*(462), C50-61.

Basu, Kaushik, Das, S. , & Dutta, B. . (2010). Child labor and household wealth: theory and empirical evidence of an inverted-U. *Journal of Development Economics, 91*(8-14).

Basu, Kaushik, & Van, Pham. (1998). The economics of child labor. *American Economic Review, 88*(3), 412-427.

Becker, G. S. (1965). A Theory of Allocation of Time. *Economic Journal, 75*, 493–517.

Beegle, K., Dehejia, R. , & Gatti, R. (2006). Child labor and agricultural shocks. *Journal of Development Economics 80*(1), 80-96.

Benson, T. , Minot, N., Pender, T. , Robles, M., & Braun, J. v. . (2008). Global food crisis: Monitoring and Assessing Impact to Inform Policy Responses: IFPRI Issue Brief 55, Washington, D.C. .

Bhalotra, S. , & Heady, C. (2003). Child farm labour: the wealth paradox. *World Bank Economic Review, 17*(2), 197-227.

Bia, M., & Mattei, A. . (2008). A Stata package for the estimation of the dose–response function through adjustment for the generalized propensity score. *Stata Journal, 8*(3), 354-373.

Blundell, R. , & Dias, M. C. . (2002). *Alternative Approaches to Evaluation in Empirical Microeconomics*. The Institute for Fiscal Studies and University College London. London.

Bound, J. , Jaeger, D. A. , & Baker, R. M. . (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak. *Journal of the American Statistical Association, 90*(430), 443-450.

Box, George E. P., & Cox, D. R. . (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological), 26*(2), 211-252.

Busjeet, G., Demombynes, G., & Sobrado, C. (2008). Food prices and the poor in Latin America *LAC Poverty and Gender Group, The World Bank, Washington, D.C.*

Castel, P. . (2009). *Vietnam Health Insurance: Use of Health Care Services by the Poor Efficiency and Equity Issues in the Province of Kon Tum*. Poverty Assessment. Vietnamese Academy of Social Sciences - Centre for Analysis and Forecasting.

Centre for International Economics, Canberra and Sydney. (2002). Vietnam Poverty Analysis. Australian Agency for International Development.

Cervantes-Godoy, D., & Dewbre, J. (2010, 29-30 November 2010). *Economic Importance of Agriculture for Sustainable Development and Poverty Reduction: The Case Study of Vietnam.* Paper presented at the Global Forum on Agriculture: Policies for Agricultural Development, Poverty Reduction and Food Security, OECD Headquarters, Paris.

Charlebois, P. , & Hamann, N. . (2010). *The Consequences of a Strong Depreciation of the U.S. Dollar on Agricultural Markets*. Agriculture and Agri-Food Canada.

CIEM, (Central Institute for Economic Management). (2007). Bao Cao Cua Chinh Phu Ve Co Phan Hoa Cac Doanh Nghiep Nha Nuoc (in English: Report of Goverment on Privatization of SOEs).

Cochrane, W. (1993). *The development of American agriculture: A historical analysis*. University of Minnesota, Minneapolis.

Congdon Fors, H. . (2007). The determinants of rural child labor: an application to India. *Working Papers in Economics 256, University of Gothenburg*.

Conway, T. , Turk, C. H., & Blomquist, J. (2001). Vulnerability, Risk and Social Protection in Vietnam *first draft, prepared for the Poverty Task Force Vietnam*.

Coxhead, I. , Linh, Vu Hoang, & Tam, Le Dong. (2011). *Global Market Shocks and Poverty in Vietnam: The Case of Rice*. Staff Paper No. 559. University of Wisconsin-Madison, Department of Agricultural & Applied Economics.

Dawe, D. , & Slayton, T. (2010). The World Rice Market Crisis of 2007–2008. In D. Dawe (Ed.), *The Rice Crisis*. London: Earthscan.

de Janvry, A., Fafchamps, M., & Sadoulet, E. (1991). Peasant Household Behaviour With Missing Markets: Some Paradoxes Explained. *The Economic Journal, 101*(409), 1400-1417.

Deaton, A. (1990). *Saving in Developing Countries: Theory and Review.* Paper presented at the Proceedings of The World Bank Annual Conference on Development Economics.

Dercon, S., & Christiaensen, L. (2011). Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia. *Journal of Development Economics, 96*(2), 159-173.

Diallo, Y., Hagemann, F., Etienne, A., Gurbuzer, Y., & Mehran, F. (2010). Global child cabour developments: measuring trends from 2004 to 2008. Geneva, Switzerland: International Labour Office, Statistical Information and Monitoring Programme on Child Labour.

Dillon, J. L. (1971). An Expository Review of Bernoullian Decision Theory in Agriculture: Is Utility Futility? *Review of Marketing and Agricultural Economics, 39*(1), 3-80.

Do, Quy-Toan, & Iyer, Lakshmi. (2004). Land Rights and Economic Development: Evidence from Vietnam. *World Bank Policy Working Research Paper No. 3120*. Washington, DC.

Dufhues, T., & Buchenrieder, G. (2005). Outreach of credit institutes and households' access constraints to formal credit in Northern Vietnam: Universitaet Hohenheim, Department of Agricultural Economics and Social Sciences in the Tropics and Subtropics. Discussion Paper No. 01/2005.

Duflo, E. . (2006). Poor but Rational? . In R. B. A. Banerjee, and D. Mookherjee (Ed.), *In Understanding Poverty* (pp. 367–378). Oxford: Oxford University Press.

Duflo, Esther, Kremer, Michael, & Robinson, Jonathan. (2008). How High Are Rates of Return to Fertilizer? Evidence from Field Experiments in Kenya. *The American Economic Review, 98*(2), 482-488. doi: 10.2307/29730068

Dumas, C. . (2007). Why do parents make their children work? A test of the poverty hypothesis in rural areas of Burkina Faso. *Oxford Economic Papers 59*, 301-329.

Dumas, C. . (2013). Market Imperfections and Child Labor. *World Development, 42*, 127-142.

Duryea, S., Lam, D. , & Levison, D. . (2007). Effects of economic shocks on children's employment and schooling in Brazil. *Journal of Development Economics 84*(1), 188-214.

Edmonds, Eric. (2001). Will child labor decline with improvements in living standards? A case study for Vietnam.

Edmonds, Eric, & Turk, Carrie. (2002). Child labor in transition in Vietnam.

Fahey, S. , McCarty, A., Scheding, J., & Huong, T. (2000). Strategic Framework for AusAID Support for the Education and Training Sector.

FAO. (2008). *Growing demand on agriculture and rising prices of commodities: An opportunity for smallholders in low-income, agricultural-based countries?* Paper prepared by the Trade, Markets, and Agricultural Development Economics Division of FAO for the Round Table, 31st session of IFAD's Governing Council.

Ferreira, F. H. G., Fruttero, A., Leite, P., & Lucchetti, L. . (2011). Rising Food Prices and Household Welfare: Evidence from Brazil in 2008 *Policy Research Working Paper 5652*: The World Bank.

Fiona, W. , & Newman, C. (2011). *Income Shocks and Household Risk-Coping Strategies: Evidence from rural Vietnam*. Department of Economics, Trinity College Dublin.

Fors, Heather Congdon. (2010). Child labour: A review of recent theory and evidence with policy implications *Journal of Economics Surveys, December*.

Frydman, R., C. , Gray, M. Hessel, & Rapaczynski, A. (1999). When Does Privatization Work? The Impact of Private Ownership on Corporate Performance in the Transition Economies. *Quarterly Journal of Economics 114*(4), 1153-1191.

Fujii, T. (2011). Impact of food inflation on poverty in the Philippines *Research Collection School of Economics. Paper 1315*.

Gershon, F., Just, R. E., & Zilberman, D. . (1985). Adoption of agricultural innovations in developing countries: A survey. *Economic Development and Cultural Change, 33*(2), 255-298. doi: 10.2307/1153228

Gilbert, C. L. . (2010). How to understand high food prices? *Journal of Agricultural Economics, 61*(2), 398–425.

Glewwe, P. (1998). An Overview of Economic Growth and Household Welfare in Vietnam in the 1990s. In N. A. Paul Glewwe, David Dollar (Ed.), *Economic Growth, Poverty, and Household Welfare in Vietnam*. Washington, D.C., USA: The World Bank

Greene, W. H. . (2000). *Econometric Analysis* (4th ed.). Englewood Cliffs, NJ, USA: Prentice Hall.

GSO. (2008). *Operational Handbook: Vietnam Household Living Standard Survey 2008*. Hanoi, Vietnam

Hare, Denise. (2008). The origins and influence of land property rights in Vietnam. *Development Policy Review, 26*(3), 339-363.

Harmeling, S. (2011). Global climate risk index 2012: Who suffers most from extreme weather events? Weather-related loss events in 2010 and 1991 to 2010.

Harvard Vietnam Program. (2008). Choosing success : the lessons of East and Southeast Asia and Vietnam's future : a policy framework for Vietnam's socioeconomic development, 2011-2020. *Asia Programs, Harvard University*.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica, 46*(6), 1251-1271.

Hazell, Peter B.R. . (2009). The Asian Green Revolution. In I. D. P. 00911 (Ed.): The International Food Policy Research Institute.

Headey, D. (2011). Rethinking the global food crisis: The role of trade shocks. *Food Policy, 36*(2), 136-146.

Headey, D., & Fan, S. (2010). *Refections on the Global Food Crisis*. IFPRI. Washington D. C., U.S.A.

Heckman, J. (1997). Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations. *The Journal of Human Resources, 32*(3), 441-462.

Heckman, J., LaLonde, R., & Smith, J. (1999). The economics and econometrics of active labor market programs. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (1 ed., Vol. 3). Amsterdam: Elsevier.

Heckman, J., & Robb, R. (1985). Alternative Methods for Evaluating The Impact of Interventions: An Overview. *Journal of Econometrics, 30*(1985), 239-297.

Heckman, J., Urzua, S., & Vytlacil, E. (2006). Understanding Instrumental Variables In Models With Essential Heterogeneity. *Review of Economics and Statistics, v88*(3, Aug), 389-432.

Hirano, K., & Imbens, G. W. (2005). The Propensity Score with Continuous Treatments *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives* (pp. 73-84): John Wiley & Sons, Ltd.

Hoang, B. T. (2009). *Rural employment and life: Challenges to gender roles in Vietnam's agriculture at present*. Paper presented at the FAO-IFAD-ILO Workshop on Gaps, Trends and Current Research in Gender Dimensions of Agricultural and Rural Employment Differentiated Pathways out of Poverty, Rome.

Hoang, K. , Pham, V. B. , & Howeler, R. H. . (2000). *A review of cassava in Asia with country case studies on Thailand and Viet Nam.* Paper presented at the THE VALIDATION FORUM ON THE GLOBAL CASSAVA DEVELOPMENT STRATEGY, Rome.

Hossain, M., Tran, T. U., & Janaiah, A. (2003). Vietnam's experience with hybrid rice. *Economic and Political Weekly, 38*(25), 2523-2529.

IMF. (2008). Food and Fuel Prices - Recent Developments, Macroeconomic Impact, and Policy Responses. In International Monetary Fund Report (Ed.), *Prepared by the Fiscal Affairs, Policy Development and Review, and Research Departments*.

Ivanic, M. , & Martin, W. (2010). *Promoting global agricultural growth and poverty reduction.* Paper presented at the 114th Seminar, Berlin, Germany.

Ivanic, M., & Martin, W. (2008). Implications of higher global food prices for poverty in low-income countries *Policy Research Working Paper Series 4594, The World Bank.*

Jeffrey, S. (2000). A critical survey of empirical methods for evaluating active labor market policies. *Schweizerische Zeitschrift fuer Volkswirtschaft und Statistik, 136*(3), 1-22.

Joshi, Shareen. (2004). Female Household-headship In Rural Bangladesh: Incidence, determinants and impact on childrens' schooling: Economic Growth Center, Yale University. Center Discussion Paper No. 894.

Kambhampati, U. S., & Rajan, R. . (2005). Does child work decrease with parental income? The luxury axiom revisited in India. *European Journal of Development Research, 17*(4), 649-680.

Khandker, S. R., Koolwal, G. B., & Samad, H. (2010). *Handbook on impact evaluation: Quantitative methods and practice*: The World Bank. Washington, D.C.

Krueger, Alan. (1996). Observation on international labor standards and trade. *NBER Working Papers 5632*: National Bureau of Economic Research.

Lanjouw, P. , Marra, M. , & Nguyen, C. . (2013). *Vietnam's Evolving Poverty Map Patterns and Implications for Policy*. Policy Research Working Paper 6355. The World Bank.

Le, D. D. (1996). Legal Consequences of State-Owened Enterprises Reform In N. J. F. Chee Yuen Ng, Frank Cong Hiep Huynh (Ed.), *State-Owned Enterprise Reform in Vietnam: Lessons from Asia*.

Lechner, M. (1999). Earnings and Employment Effects of Continuous Off-the-Job Training in East Germany after Unification. *Journal of Business & Economic Statistics, 17*(1), 74-90.

Leuven, E. , & Sianesi, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.

Lustig, N. (2009). Coping with rising food prices: Policy dilemmas in the developing world *Working Paper Number 164*. Washington, D.C.: The Center for Global Development.

Mac Millan, J., & Woodruff, C. . (1999). Interfirm Relationships and Informal Credit in Vietnam. *Quarterly Journal of Economics, 114*(4), 1285-1320.

McCarty, A. (2001). Microfinance in Vietnam: A Survey of Schemes and Issues: DFID and State Bank of Vietnam.

Mendola, M. (2007). Farm Household Production Theories: A Review of "Institutional" and "Behavioral" Responses. *Asian Development Review, 24*(1), 49-68.

Millar, J. R. . (1970). A Reformulation of A. V. Chayanov's Theory of the Peasant Economy. *Economic Development and Cultural Change, 18*(2), 219-229.

Mitchell, D. (2008). A note on rising food prices. Washington, D.C.: World Bank.

Myrdal, G. (1968). *Asian Drama: An Inquiry into the Poverty of Nations*. New York: Pantheon.

Nebitt, H. , Johnston, R., & Solieng, M. (2004). *Mekong River water: will river flows meet future agriculture needs in the Lower Mekong Basin.* Paper presented at the Water in Agriculture: ACIAR Proceedings No. 116.

Newbery, B. M., & Stiglitz, J. E. (1981). *The Theory of Commodity Price Stabilization: A Study in the Economics of Risk*. Oxford: Clarendon Press.

Nguyen, K. T. , Khuat, O. T. H., Ma, S., Pham, D. C., Khuat, G. T. H., & Ruger, J. P. (2012). Effect of Health Expenses on Household Capabilities and Resource Allocation in a Rural Commune in Vietnam. *PLOS ONE, 7*(10).

Nguyen, V. B. , Mutert, E. , & Cong, D. S. (2003). *BALCROP: Balanced fertilization for better crops in Vietnam*.

Nielsen, C. P. . (2003). Vietnam's Rice Policy: Recent Reforms and Future Opportunities. *Asian Economic Journal, 17*(1), 1-26.

Nose, M. , & Yamauchi, F. . (2012). Long-term Impacts of Global Food Crisis on Production Decisions: Evidence from Farm Investments in Indonesia *Policy Research Working Paper 6065*: The World Bank.

Nuryartono, N. . (2006). *Impact of Smallholders' Access to Land and Credit Markets on Technology and Land Use Decisions: The Case of Tropical Forest Margins in Central Sulaewsi-Indonesia*: Cuvillier Verlag Gottingen.

Organization for Economic Cooperation and Development's Development Assistance Committee (OECD-DAC). (2002). Glossary of Key Terms in Evaluation and Results-Based Management Proposed Harmonized Terminology. OECD, Paris.

Pagan, A., & Ullah, A. (1999). *Nonparametric Econometrics*. Cambridge: Cambridge University Press.

Paxson, C. (1991). Using weather variability to estimate the response of savings to transitory income in Thailand. *American Economic Review, 8*(1).

Pham, B. D. , & Izumida, Y. (2002). Rural Development Finance in Vietnam: A Microeconometric Analysis of Household Surveys. *World Development, 30*(2), 319–335.

Pham, H. N. (2010). The Vietnamese Rice Industry During the Global Food Crisis. In D. Dawe (Ed.), *The Rice Crisis* (pp. 219-232). London: Earthscan.

Phung, D. T., & Nguyen, P. . (2006). Vietnam Household Living Standard Survey (VHLSS) 2002 and 2004: Basic Information.

Phung, D.T., & Waibel, H. (2010). Food Price Crisis, Poverty and Welfare in Vietnam: An Ex-post Decomposition Analysis *Working Paper. Leibniz Universität Hannover, Institute of Development and Agricultural Economics, Germany*.

Poverty Working Group. (2000). Attacking Poverty *Vietnam Development Report 2000*.

Ranjula Bali, S., Nguyen, V. S. , & Vo, V. T. . (2008). Microfinance and Poverty Reduction in the Mekong Delta in Vietnam. *African and Asian Studies, 7*(2-3), 191-215.

Ravallion, M. . (2008). Evaluating anti-poverty programs. In E. E. Robert & P. T. Schultz (Eds.), *Handbook of Development Economics Volume 4*. Amsterdam, North-Holland.

Ravallion, M. , & van de Walle, D. . (2008). *Land in Transition: Reform and Poverty in Rural Vietnam*. Washington, D.C., USA: Palgrave Macmillan and the World Bank.

Ray, D. (2006). Aspirations, Poverty, and Economic Change. In R. B. A. Banerjee, and D. Mookherjee (Ed.), *In Understanding Poverty* (pp. 409–421). Oxford: Oxford University Press.

Rosati, Furio Camillo, & Tzannatos, Zafiris. (2006). Child labor in Vietnam. *Pacific Economic Review, 11*(1), 1-31.

Rosegrant, M. W., & Hazell, P. B. R. (2000). *Transforming the rural Asia economy: The unfinished revolution*. Hong Kong: Oxford University Press.

Rosenbaum, P., & Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika, 70*(1), 41-55.

Rosenbaum, P., & Rubin, D. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity *The American Statistician, 9*(1), 33-38.

Rosenzweig, M. R., & Binswanger, H. P. (1993). Wealth, weather risk and the composition and profitability of agricultural investments. *The Economic Journal, 103*(416), 56-78. doi: 10.2307/2234337

Rubin, D. (1974). Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology, 66*(5), 688-701.

Schultz, T. W. (1964). *Transforming traditional agriculture*. Yale University, Press, New Haven, CT.

Sjöholm, F. . (2006). *State owned enterprises and equitization in Vietnam.* Working Paper 228. Stockholm School of Economics, Sweden.

Smith, J. A. , & Todd, P. E. . (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics, 125*(1–2), 305-353.

Sunding, D. , & Zilberman, D. . (1999). The agricultural innovation process: Research and technology adoption in a changing agricultural sector *Handbook of Agricultural Economcis*.

The World Bank. (2009). *Global facility for disaster reduction and recovery: Disaster risk management programs for priority countries*.

Thomas, T. , Christiaensen, L. , Do, Q. T. , & Le, D. T. . (2010). Natural disasters and household welfare: Evidence from Vietnam: The World Bank.

Thurlow, J., Tarp, F., McCoy, S. , Nguyen, M. H. , Breisinger, C. , & Arndt, C. . (2010). The Impact of the Global Commodity and Financial Crises on Poverty in Vietnam *New Directions in Development Economics. Working Paper No. 2010/98*: UNU-WIDER, Helsinki.

Timmer, P. (2009). *Did Speculation Affect World Rice Prices?* Agricultural and Development Economics Division of the Food and Agriculture Organization of the United Nations.

Timmer, P. . (2010). Did Speculation Affect World Rice Prices? In D. Dawe (Ed.), *The rice Crisis*: Earthscan.

Tobin, J. . (1958). Estimation of relationships for limited dependent variables. *Econometrica, 26*(1), 24-36.

Tran, T. U. , Hossain, M. , & Janaiah, A. (2000). Modern Farm Technology and Infrastructure in Vietnam: Impact on Income Distribution and Poverty. *Economic and Political Weekly, 35*(52/53), 4638-4643.

Trostle, R. (2008). Global agricultural supply and demand: Factors contributing to the recent increase in food commodity prices. Washington, D.C.: United States Department of Agriculture.

UCW Project. (2009). Understanding children's work in Vietnam *Understanding Children's Work*: University of Rome "Tor Vergata".

Udry, C. (1995). Risk and saving in northern Nigeria. *The American Economic Review, 85*(5), 1287-1300.

Valero-Gil, J. N., & Valero, M. (2008). The effects of rising food prices on poverty in Mexico. *Agricultural Economics, 39*(2008), 485–496.

Van de Walle, D. (1996). *Infrastructure and Poverty in Viet Nam*. Washington, D.C. : The World Bank

Vietnam National Assembly. (2003). *Law on Land*. (No. 13-2003-QH11).

Vo, T. T. (2008). *Impacts of global financial crisis on Vietnam's economy and recommended policy responses.* Institue of Development Studies.

von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Vu, L. , & Glewwe, P. . (2011). Impacts of rising food prices on poverty and welfare in Vietnam. *Journal of Agricultural and Resource Economics, 36*(1), 14-27.

Vuong, Quan Hoang. (2010). *Financial Markets in Vietnam's Transition Economy: Facts, Insights, Implications*. Saarbrücken, Germany.

White, H. . (2006). Impact evaluation: The experience of the independent evaluation group of the World Bank. Washington D.C. : World Bank.

Wodon, Q., Tsimpo, C., & Coulombe, H. . (2008). Assessing the impact on poverty of rising cereal prices: The case of Ghana *Policy Research Working Paper 4740*: The World Bank.

Wodon, Q., & Zaman, H. (2008). Rising food prices in Sub-Saharan Africa: Poverty impact and policy responses *Policy Research Working Paper 4738*: The World Bank.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Wooldridge, J. M. (2005). *Introductory econometrics: A modern approach* (3th ed.).

World Bank. (2000). Vietnam 2010 Pillars of Development *Vietnam Development Report 2001, Joint Report Asian Development Bank and UNDP*.

World Bank. (2001). Vietnam Growing Healthy: A review of Vietnam's Health Sector *Report No.22210-VN, June 2001*.

World Bank. (2008). The development impact evaluation (DIME) initiative Washington D.C.

World Bank. (2009). Global economic prospects 2009: Commodities at a crossroads. Washington, D.C.

World Bank. (2011). Boom, Bust and Up Again? Evolution, Drivers and impact of Commodity Prices: Implications for Indonesia *A Trade Development Report.*

Young, B. K. , Wailes, J. E. , Cramer, L. G., & Nguyen, T. K. (2002). Vietnam's rice economy: Developments and prospects. Division of Agriculture, University of Arkansas.

Yu, Bingxin, Zhu, Tingju , Breisinger, Clemens , & Nguyen, Manh Hai. (2012). *Examine the agriculture, poverty, and climate change: Nexus in Vietnam*. Paper presented at the International Association of Agricultural Economists (IAAE) Triennial Conference, Foz do Iguaçu, Brazil.

Zezza, A., Davis, B. , Azzarri, C. , Covarrubias, K. , Tasciotti, L. , & Anriquez, G. . (2008). The impact of rising food prices on the poor *ESA Working Paper No 08-07*: FAO (Food and Agriculture Organization of the United Nations).

Zimmerman, F. J., & Carter, M. R. (2003). Asset smoothing, consumption smoothing and the reproduction of inequality under risk and subsistence constraints. *Journal of Development Economics, 71*(2), 233-260.