

Rural Credit and Household Consumption: Evidence from Vietnam

by

Cuong Nguyen

Thesis submitted for the degree of Doctor of Philosophy
at Heriot-Watt University

Centre for Economic Reform and Transformation (CERT),
School of Management and Languages,
Heriot-Watt University,
Edinburgh, United Kingdom

February 2008

The copyright in this thesis is owned by the author. Any quotation from the thesis or use of any of the information contained in it must acknowledge this thesis as the source of the quotation or information

ABSTRACT

This thesis is a contribution to the research literature of the Vietnamese rural credit market during 1990s. The data used is from the Vietnam Living Standard Surveys carried out in 1992/93 and 1997/98, providing both cross-sectional and panel data.

The thesis consists of four main chapters. Chapter 2 provides an overview of the rural credit markets. We find a big expansion of the formal credit sector over years but financial system remains largely underdeveloped with a dual structure in which formal and informal sectors exist side by side. Chapter 3 is a study assessing determinant factors of household's credit participation formal and informal credit markets. We contribute to the literature as the first research attempts to estimate the credit demand and the credit supply functions separately. We are also the pioneer in applying the bivariate probit model with partial observability in empirical study using data from Vietnam. The empirical analysis in the Chapter 4 highlights the determinant factors of credit amount obtained. We find that poor rural households who seem to borrow most often are not those who borrow the largest loan size. The last chapter aims to assess the impact of formal credit on household consumption. To address problem of selection bias, we employ instrumental variable method. We show that on average each percent increased in household's formal credit per capita would lead to 0.26 percent increased in consumption per capita.

Our results draw some important policy implications. First, even though formal credit network continues to expand greatly to cover most rural areas, there is a question on the outreach and flexibility of credit services because rationing remains as a serious problem for rural households. Credit institutions should develop a better screening system, e.g. credit scoring system, to reduce collateral requirement which actually considered as one of the most obstacles of the formal credit access. Land regulations and fixed asset legal entitlement should also be reviewed together with improvement of administration procedure to enable households to use their property as collateral if required. In addition, increased demand for credit implies more investment opportunities. Hence, government should continue to improve the economic infrastructure to facilitate agricultural product trade market, provide better health and education system to maintain this demand.

To my family

Acknowledgement

I wish to thank my supervisors, Professor Mark Schaffer and Dr. Julian Fennema, for their invaluable guidance, advice, support and encouragement through out this research. I also thank Prof. Seth Armitage for his insightful discussion during my first year at Heriot-Watt University (HWU).

I wish to thank the Centre for Economic Reform and Transformation (CERT) and the School of Management and Languages for their financial support. This thesis would not have been possible without them.

I am grateful to all the comments and suggestions I have received from participants and discussants of the internal weekly lunch seminars at the HWU as well as conferences outside of the HWU.

I wish to thank the administrative staffs of the School for their efficient support.

Special thanks also to all my dear friends, brothers An, The, Phuc, Tung, little sisters Thanh, Nga and other Vietnamese gangs in Edinburgh and Glasgow, who made my years away from home much more enjoyable.

Finally, most of all, I thank my family: my parents, my wife Bao Ngoc, my daughter Ngoc Lan and one daughter to be named soon. I would not be able to complete this research without their love, understanding and support. This thesis is dedicated to them.

**ACADEMIC REGISTRY
Research Thesis Submission**



Name:	CUONG NGUYEN		
School/PGI:	MANAGEMENT AND LANGUAGES		
Version: <i>(i.e. First, Resubmission, Final)</i>	FINAL	Degree Sought:	PH.D.

Declaration

In accordance with the appropriate regulations I hereby submit my thesis and I declare that:

- 1) the thesis embodies the results of my own work and has been composed by myself
- 2) where appropriate, I have made acknowledgement of the work of others and have made reference to work carried out in collaboration with other persons
- 3) the thesis is the correct version of the thesis for submission*.
- 4) my thesis for the award referred to, deposited in the Heriot-Watt University Library, should be made available for loan or photocopying, subject to such conditions as the Librarian may require
- 5) I understand that as a student of the University I am required to abide by the Regulations of the University and to conform to its discipline.

* Please note that it is the responsibility of the candidate to ensure that the correct version of the thesis is submitted.

Signature of Candidate:		Date:	02/02/08
-------------------------	--	-------	----------

Submission

Submitted By <i>(name in capitals)</i> :	C Murray
Signature of Individual Submitting:	C Murray
Date Submitted:	11/03/08

For Completion in Academic Registry

Received in the Academic Registry by <i>(name in capitals)</i> :	CLAIRE RUSSELL		
Method of Submission <i>(Handed in to Academic Registry; posted through internal/external mail):</i>	HANDS IN TO ACADEMIC REGISTRY		
Signature:		Date:	11/3/08

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Contributions and Outline of the Dissertation	4
2	The Vietnamese Rural Credit Market in the 1990s	8
2.1	Introduction	8
2.2	The Vietnam Living Standard Survey	12
2.3	Credit Markets for Rural Households	14
2.3.1	Characteristics of Households and Communes in the Sample .	15
2.3.2	The Expansion of Formal Credit Sector	17
2.3.3	Features of Loan Contracts	19
2.3.4	Linkages between Formal and Informal Credit Sectors	23
2.4	Conclusion	26
3	Access to Credit in Rural Vietnam: A Partial Observability Model	27
3.1	Introduction	27
3.2	Credit Access and Credit Constraint	30
3.3	Estimating Determinants of Credit Participation by Univariate Probit Model	33
3.3.1	Econometric Specification	33
3.3.2	Variable Selection	35
3.3.3	Estimation Results	39
3.4	Bivariate Probit Model with Partial Observability	46
3.4.1	Econometric Specification	46
3.4.2	Identification Strategy	49

3.4.3	Estimation Results	53
3.5	Conclusion	56
4	The Determinants of Credit Amount Obtained by Rural House-	
	holds in Vietnam	57
4.1	Introduction	57
4.2	Econometric Framework	58
4.3	Model Identification	61
4.4	Results	63
4.5	Conclusion	69
5	The Impact of Formal Credit in Rural Vietnam	70
5.1	Introduction	70
5.2	Impact Assessments using Cross-Sectional Data	75
5.2.1	Econometric Specification	75
5.2.2	Results	78
5.3	Impact Assessments using Longitudinal Data	80
5.3.1	Addressing Contemporaneous Impacts	80
5.3.2	Results	83
5.3.3	Alternative Specifications	85
5.4	Conclusion	87
6	Conclusion	89
A	Tables and Figures Chapter2	104
B	Tables Chapter 3	120
C	Tables Chapter 4	126
D	Tables Chapter 5	132
E	VLSS Data Definition	137

List of Tables

A.1	Vietnam's Economic Indicators - 1990-2000	105
A.2	Mean characteristics of household and commune - VLSS1992/93 . . .	108
A.3	Mean characteristics of household and commune - VLSS1997/98 . . .	109
A.4	Characteristics of Borrowing Households by Credit Sectors	110
A.5	Distribution of Borrowing Households by Region	111
A.6	Distribution of Borrowing Households by Expenditure Quintiles and Credit Sectors	112
A.7	Distribution of Loan Number and Loan Amount by Expenditure per Capita Quintile	113
A.8	Distribution of Borrowing Households by Employment of Household Head	114
A.9	Loan Characteristics - VLSS1992/93	115
A.10	Loan Characteristics - VLSS1997/98	116
A.11	Collateral assets	117
A.12	Borrowing Purposes	118
A.13	Borrowing Lenders	119
B.1	Estimation Results: Univariate Probit Model (Model I)	121
B.2	Estimation Results: Bivariate Probit Model	122
B.3	Estimation Results: Bivariate Probit Model with Partial Observabil- ity (Model II) of Formal Sector in 1992/93	123
B.4	Estimation Results: Bivariate Probit Model with Partial Observabil- ity (Model II) of Formal Sector in 1997/98	124
B.5	Estimation Results: Bivariate Probit Model with Partial Observabil- ity (Model II) of Informal Sector in 1997/98	125
C.1	Estimation Results: OLS	128
C.2	Estimation Results: the Tobit model	129

C.3	Estimation Results: Heckman Selection Model for the Formal Sector Credit Amount Obtained	130
C.4	Estimation Results: Double Hurdle Selection Model for the Formal Sector Credit Amount Obtained	131
D.1	Estimation Results: Impacts of Formal Credit - OLS and OLS-FE estimates	133
D.2	Estimation Results: Impacts of Formal Credit - Instrumental Variable Estimates on Cross-Sectional Data	134
D.3	Estimation Results: Impacts of Formal Credit - Instrumental Variable Estimates on Longitudinal Data	135
D.4	Estimation Results: Impacts of Formal Credit - Household Fixed- Effect Estimates	136

List of Figures

A.1	Share of Agriculture in total GDP, Employment and Export	105
A.2	Poverty Distribution - Map of Vietnam.	106
C.1	Distribution of the credit amount obtained in level and in logarithm transformation	127

Chapter 1

Introduction

1.1 Motivation

The classical models of financial markets predict that capital should flow from businesses offering low returns to those offering high returns. The principle of diminishing marginal returns to capital states that enterprises with little capital should be able to earn higher returns on their investment than enterprises with more capital, for a given level of investment opportunity. In other words, poorer firms should be able to pay banks higher interest rates than richer firms. As a consequence, credit providers should target poorer areas to lend as they will be able to lend at a higher price. A study by Lucas (1990) found that borrowers in India should be willing to pay nearly sixty times as much for capital as borrowers in the United States. If investors follow what the theory predicts then capital should flow from the US to India. By the same logic, funds should not only move from rich to poor countries but also from richer areas to poorer areas within any given country. However, the reality is more complicated and sits at odds with the basic theory states. Funds and capital tend to stay in wealthier markets leaving poor markets on the periphery.

The rural credit markets in developing countries are interesting examples of these market failures where poor areas receive very little interest from credit organisations. It is common that segments of borrowers have different levels of access to certain

types of loans and certain types of credit institutions (Hoff & Stiglitz, 1990). Credit rationing theory states that the interest rates may not be equilibrium determined by credit supply and demand. And some would-be borrowers may find themselves excluded or dissuaded from the formal sector such that lending may be not available to them at any price. Possible reasons for these constraints have been mentioned in a vast number of studies emphasizing the limited availability of formal funding, the riskiness and uncertainty of rural credit market, the problem of asymmetric information between lenders and rural borrowers, the weak enforcement system in less developed countries, and the poorly developed economic infrastructure (see, for example, Hoff & Stiglitz (1993), Aghion & Morduch (2005), Bardhan (1999), Bell (1990), Ray (1998), Townsend (1995), among many others). Formal credit applicants may then adjust by turning to substitute but possibly more expensive financing sources, or may modify their first best allocation plans in other ways.

In contrast to the formal sector, the informal sector provides much easier conditions of access to loans. Borrowers and lenders often live in the same areas, reducing information asymmetry. Lenders usually do not require any collateral or documents, the relationship is governed by informal habits. The incentive to borrowers to repay is based on a relationship of trust, although sometimes repayment is enforced through violence (Siamwalla et al, 1990). The different levels of access to loans explains the fact that credit markets are often characterised by monopoly power of village moneylenders who charge borrowers a very high interest rate, such as 80 percent per year in the case of rural Pakistan (Aleem,1990), or from 20 to 120 percent per year in different rural areas in India (Timberg, 1984). By charging high interest rates and providing short term loans, informal lenders sometimes force borrowers into a state of default.

In 1960s, realising the failure of formal markets to provide financial services in poor rural areas, governments in low-income countries started intervening in the markets. Large government agricultural banks were given the responsibility for allocating funds. The interest rates were kept low, and banks would receive compensation for their loss, with the hope that providing subsidized credit would encourage

improvements in farm production and reduction of poverty. Philippines and India are two of the most cited examples for governments' cheap credit programmes (Meyer, 2002). However, the subsidized credit failed. The critics of the subsidized government banks, led by Rural Finance Programme at Ohio State University, argue that the rationing mechanism to screen risky borrowers failed due to the subsidized low interest rates, and funds were too cheap to act as an incentive for the poor to invest; they were more likely to consume. The poverty remained, while loan repayment rates plummeted. Braverman & Guasch (1986) find that, with very few exceptions, subsidized credit programmes often ended up with default rates between 40% to 95%. The poor household could have been better off without the subsidies (McKinnon (1973), Gonzalez-Vega (1984).

From the middle of the 1970s, microfinance started to become a popular credit strategy to replace the government-led subsidized programmes. Originating from the Grameen Bank in Bangladesh, the microfinance movement aims to solve problems of subsidized credit and find the way to reach the poor people in a more efficient way. The most striking idea is to lend to a group of borrowers rather than to an individual. This ultimately solves many problems: firstly, no collateral is required; secondly, the joint-liability condition of the group loan contract helps to overcome problems of moral hazard and adverse selection by enforcing the peer monitoring among group members. The success of Grameen had led to its replications in many other countries, including Vietnam.

In Vietnam, by the late 1980s, under the pressure of potential political and economic crisis, Vietnam started to transform its "central-planned" into "market-oriented" economy, the "*Doi Moi*" (reformation) programme as in Vietnamese. The reform has achieved remarkable successes, creating one of the fastest growing economies in the world. At the centre of "*Doi Moi*" is agricultural transformation: collectivist production was abandoned; irrigable land was reallocated to households. From the beginning of the transformation process, the Vietnamese government has focussed its economic policy toward balancing development between rural and urban areas. The rural credit markets in Vietnam, therefore, has received substantial

investment (World Bank, 1999). In 1999 as many as 5.9 million households had access to formal financial institutions, of which 2.7 million households were poor and low-income households (McCarty, 2001).

This thesis is a study of Vietnamese rural credit markets in 1990s. It sets out to answer three questions: what are the determinant factors of household's accessibility to the formal credit sector in rural Vietnam? What are the determinants of credit amount obtained by households? and what are the impacts of taking credit?

1.2 Contributions and Outline of the Dissertation

The thesis is organised in four main chapters. The first chapter aims to provide an overview of the Vietnamese rural credit markets in 1990s, using data from the Vietnam Living Standard Survey. Evidence from the VLSS data shows a big expansion of the formal credit sector in rural Vietnam. From a low 30% market share in 1992/93, the formal sector had grown to account for 50% of the rural credit market in 1997/98, although the development is not well-balanced across regions and households quintile by expenditure per capita. Formal funding is cheaper in term of interest rates but more often requires collateral in comparison with informal source. Average interest rates of both sectors had been cut by almost two third over time between the surveys, indicating a growing competition between lenders. However, the existence of informal lenders remains strong, presenting a complex linkages between agents in the credit markets.

The second chapter is a study of the determinants factors of credit accessibility of Vietnamese rural households, with concentration on access to the formal sector. We aims to estimate the demand and the supply channels separately. The demand for formal credit is identified by variable capturing health status of household members. The formal supply is identified by commune-level formal credit per capita, which reflects operation of formal lenders in the studied communes. We bases our analysis on the bivariate probit model with partial observability. This model is employed to

address the problem of limited information on the non-borrowing households. To the best of our knowledge, this is the first study of Vietnamese rural credit markets that allows supply of and demand for credit to be estimated distinctly.

Estimation results from this chapter present evidence of the impacts of economic development on credit behaviour of rural households. The role of household characteristics determining level of access to the formal sector is differed over the years between the two surveys, implying the changes of economic environment. The distinction between the demand for formal and informal credit arises from the observable variables, such as physical asset or human capital, reflecting household demand for investment capital. The supply of formal credit depends significantly on informational variables like distance from commune to the nearest bank branch. Prediction of formal credit market participation is high and increases over years, implying the expansion of the formal sector. The demand for formal credit is also inflated, suggesting household's need for investment capital grows with economic development.

The empirical analysis of chapter three highlights the determinant factors of credit amount obtained. The employed econometric models view probability of participating in credit markets and credit amount received as a joint determination of the function of household's demand for credit and the function of lender's decision on supply. The results of this chapter imply some important policy implications. First, even though formal credit network continues to expand greatly to cover most rural areas, there is question on the outreach and flexibility of credit services. Credit scoring system should be developed to reduce collateral requirement which actually considered as one of the most constraints for formal credit access. Land regulations and fixed asset legal entitlement should be reviewed together with improvement of administration efficiency to enable households to use their fixed property as collateral for credit.

The last chapter aims to assess the impact of formal credit on household consumption. Analyses using cross-section and longitudinal data have been used. In-

strumental variable method is employed to tease out the impact of credit on household consumption from other potential influences and biases. Results suggest that borrowing from the formal sector significantly improves consumption per capita of borrowing households. The challenges in impact assessments arise because credit is not provided to random households. On one hand, from the supply side, formal credit institutions carefully select areas to provide their financial services. The selected area is not necessary having "better" economic conditions compared to others, it may be poorer as result of credit targeted. However, by comparing the incidence of poverty in credit available and credit unavailable areas, which are different in characteristics, researchers may mistakenly report biased results. On the other hand, from the demand side, households who have taken credit may be systematically different from others who have not. When the characteristics that cause these differences are observable (e.g. age, education, assets) and are taken into account, impact evaluation would be less problematic. However, the differences often arise from self-selection process, which are driven by unobserved traits of borrowers, such as ability. A simple evaluation of credit programme without calculating these attributions will be biased.

We use instrumental variable to identify the credit amount that household obtained. The variables of credit per capita at commune-level, distance from commune committee to the nearest government bank branch (in form of log-transformation and dummy) are employed as the instruments. Our results are consistent with other studies of the Vietnamese rural credit markets (e.g. Pham et al (2002)) that also find positive and significant effects of credit on consumption. In particular, we find in our panel analysis that 1 percent increase in average credit per capita borrowed from the formal sector leads to 0.26 percent increase in average consumption per capita. Controlling for selection biases has been shown to be very important. In our study, estimations using simple OLS regression lead us to conclusion of no effects of credit. This result is clearly underestimated the more reliable (positive) results in which biases have been carefully addressed. However, a reservation should be pointed out. That is our core instrumental variables are at commune-level. These are useful to

control for intra-communal differences in formal credit accessibility but they do not allow controlling for inter-communal or household-level unobserved characteristics. Although this problem has been addressed by using household fixed-effects, and by cluster-robust test statistics and standard errors, results might be different if other instrumental variables had been used.

Chapter 2

The Vietnamese Rural Credit Market in the 1990s

2.1 Introduction

From independence in 1975 to mid-1980s, Vietnam remained as one of the poorest countries in the world. The gross domestic product (GDP) per capita in 1985 was US\$130 per year, placing Vietnam in the group of five poorest countries Glewwe (2004). Although school enrollment rates and life expectancy were relatively high compared to other poor countries in the same category in Africa, Latin America and South Asia, Vietnam's low income meant that the majority of the population was living in extreme poverty.

For more than a decade after the war, and similar to other communist countries, Vietnam's economy had been central planned where most production and resources were controlled by the government. Agricultural production had been organised in cooperatives, free trade was illegal, and goods and products could not move from one province to others without government's approval. In that framework, the economy did not grow and the inflation rate rose dramatically as government printed money to finance their deficits¹.

¹The annual inflation rate was recorded at 487 percent in 1986

In 1986, in response to poor economic performance that could push the country into the danger of economic and political crisis, Vietnam started its economic renovation ("*Doi Moi*") policy. A series of fundamental policies changes was gradually implemented, including legalisation of private economic activities, termination of price control, and reform of agricultural sector which is considered as one of the most important policy changes. By 1989, agricultural cooperatives were fully dismantled and farm households were allowed to cultivate on their land portions under lease contracts from the government. Agricultural reform rapidly turned Vietnam from a rice-importing country to one of the world's largest rice exporters.

There were regional differences that are likely to vary the pace of economic transition, especially the pace of agrarian transformation. After the re-unification in 1975, the collectivisation process of rural production for provinces in the South was far behind compared to those in the North. By the time Vietnam started its de-collectivisation in 1986, less than 10 percent of all of Southern farmers had joined collectives while almost all agricultural production in the North was collectivised (Pingali and Xuan, 1992). This is one of many reasons that market economy was more developed in the South at the beginning of the transition.

In 1988, two years after the economic reform started, the financial system reform was also implemented. For a long period before that, the Vietnamese banking system had a socialist mono-bank form in which the State Bank of Vietnam (SBV) was the only entity (Roman, 1997). The role of the SBV was limited to implementing government decisions on resource allocation or, more precisely, to financing state-owned enterprises. Demand for credit from private sector was low. People had little incentive and opportunity to expand private production due to its illegalised status. Moreover, there was literally no private agricultural production as most of production activities are collectivised. As a result, in this period the informal credit sector essentially satisfied all private demand (Do, Le, Nguyen and Dinh, 2001). However, the economic reform, the de-collectivisation and liberalisation of agricultural production, and the privatisation of small-scale trade and industry generated demand for investment capital. The financial system therefore needed to be restructured

to meet the requirement of economic development. In March 1988, the mono-bank system was broken into a two-tier system, with the SBV playing the role of central bank, and a number of independent commercial banks. Ten years later, by 1998, besides the SBV, the banking system included four large state-owned commercial banks ², 41 joint-stock commercial banks, 977 People's Credit Funds (PCFs), 50 cooperatives, and a small number of joint-venture banks and foreign bank branches (SBV, 1998). The stock market was opened in 2000 but remained small with limited channels for mobilizing capital and fewer than 30 companies listed in 2004.

Of the four state-owned commercial banks, the Vietnam Bank for Agricultural and Rural Development (VBARD) is the largest one in term of total capital and number of branches. It is also the key commercial bank serving rural areas and supporting development of rural production. In response to increasing demand of private credit and to enforce government policy of balancing economic growth and poverty reduction between rural and urban regions, the Government established in 1995 the Vietnam Bank for the Poor (VBP)³, a specialised government bank aimed at providing subsidised and directed credit for poor people, improving the out-reach and accessibility to formal credit for poorest households⁴. However, in rural areas informal credit continues to dominate the market in which relatives and money lenders are the most important sources of supply. The formal sector had expanded rapidly along with economic development to acquire 50 percent of market share in 1998 from about 30 percent in 1993 (Table A.4).

As seen in the figure A.1, the shares of agriculture in total GDP, export and employment have steadily decreased. Since the beginning to the end of 1990s, contribution of agriculture to GDP has been cut in half, from 40 percent to almost 20 percent. Although agriculture remains its important role in rural areas with the

²Namely the Vietnam Bank for Agriculture and Rural Development (VBARD), the Vietnam's Commercial and Industrial Bank (Vietcombank), the Vietnam Bank for Foreign Trade (Vietcombank) and the Vietnam Investment and Development Bank (Vietindbank)

³In 2004, VBP was restructured and renamed as the Vietnam Bank for Social Policies. However, through out this thesis we still use the name of VBP as it was during our study period.

⁴Because of the similarity in characteristics and operations, and of the fact that VBP operating via VBARD branch network, we should understand VBARD as referring to both VBARD and VBP except otherwise state.

share of four fifths of the country's population, its contribution to employment has declined progressively. By 2003, agriculture employs about 60 percent of labour force, down from 75 percent in 1990.

Agricultural land-use right was first privatised and then free exchange of land was legalised, formally in the first Land Law introduced in 1993. Questioning the efficiency of land reallocation during this period, a study of Ravallion et al (2006) concludes that the agrarian transition process favoured the "land-poor". In other words, households who possess low (high) amount of crop land under collective production would tend to see their land holding increased (decreased) over studying time. In another study, Do et al (2007) investigate the impact of land titling on rural production and find that households in provinces where the authorities have made better progress in issuing land-use right documents devote higher proportion of their land to multi-year crops and spend more time on non-farm work. They, however, do not find impact of land titling on household borrowing amount, even though farmers would have greater access to credit markets as they could pledge land as collateral.

Ten years after the economic reform, Vietnam's GDP growth rate had been maintained at about 8% annually. GDP per capita increased more than three times from USD130 in 1985 to almost USD400 in 2000 (see Table A.1). The proportion of people with per capita expenditures under the total poverty line has dropped dramatically from 58% in 1993 to 37% in 1998 (World Bank, 1999). Continuing through the 1990s, the Vietnamese economy transformed itself from a poor and closed economy into one of the most successful countries in the world in terms of economic growth, poverty reduction, and increased household welfare (Glewwe, 2004). This chapter plans to review a part of that remarkable transformation: the rural credit markets; using data from the Vietnam Living Standard Survey in 1992/93 and 1997/98.

The chapter is organised as follows. Section 2.2 will briefly discuss the VLSS and the sample. Section 2.3 provides a review of rural credit markets, focusing on the characteristics of rural households, the expansion of formal credit sector,

the features of loan contracts, and the linkages between formal and informal credit sector. Section 2.4 is conclusion.

2.2 The Vietnam Living Standard Survey

The primary data source used in this thesis is the Vietnam Living Standard Survey (VLSS) conducted in 1992/93 and 1997/98 by the General Statistical Office of the Government of Vietnam (GSO), funded by United Nation Development Programme (UNDP) and the World Bank. The surveys contain detailed information of 4,800 households from 150 communes in VLSS1992/93 and 6,000 households from 194 communes in VLSS1997/98. These surveys were constructed to create a panel data. As many households as possible that had already been interviewed in 1992/93 were sampled again in 1997/98. For some reasons, 495 households from VLSS1992/93 were not re-interviewed. To reach 6,000 households in the 1997/98 survey, an additional 1,695 households were selected from the total sample of 1995 Multi-Purpose Household Survey of the GSO.

Vietnam is geophysically divided into seven regions that internally are relatively homogeneous. Both 1992/93 and 1997/98 VLSS bases on this region category as basis for sampling, namely the Northern Uplands, the Red River Delta, the North Central Coast, the South Central Coast, the South East, the Highlands and the Mekong Delta. The Northern Uplands, Highlands and North Central Coast are the three poorest regions (see Figure A.2). For both surveys, samples were weighted basing on the statistics of Vietnam Population Census in 1989 with approximately 80% of Vietnamese households lived in rural areas. Commune samples were randomly selected from a total of proximately 10,000 communes in 646 districts, and 64 provinces and cities in Vietnam, and then an average of 32 households were randomly selected for interview in each commune. For the purpose of the thesis, we select only households who are residing in rural areas at the time of the surveys, which are including 3839 households in 120 communes for VLSS1992/93, and 4269 households in 150 communes for VLSS1997/98. Design of the VLSS allows us to

employ both cross-sectional and panel analyses. The panel data consisting 3,364 rural households will be used in Chapter 5.

The two surveys are similar in many respects which provide data in both household and commune levels. The household questionnaire was based on the format of the World Bank's Living Standards Measurement Surveys with adaptation to Vietnamese characteristics. The surveys collected information of household as a whole and also information of individuals within that household, such as personal profile, education, health status, employment, agriculture production, housing, consumption, credit and saving activities. For information on credit activities, households are asked to report whether they borrow any money from other individuals or organisations within the last 12 months, and details of loan they taken including source of loan, amount, duration, interest rate and collateral. The commune questionnaire was only applied to rural areas, collected basic data on commune geographical information, general economic conditions, transportation and infrastructure, and credit⁵.

The data processing are divided into two stages, the actual field interview and the final data process at the GSO office in Hanoi. In the first stage, the VLSS team calls for two rounds of interviews and data were entried in the field itself. Data collected in the first round is checked by the supervisor prior to the second round and necessary clarifications sought from the concerned household during the second round. All interviews were carefully supervised by senior staffs of the GSO and consultants of the World Bank. They observed some interviews and randomly revisited some of the interviewed households to check the quality of the work. In the second stage, at the GSO head office in Hanoi, the field data were intensively checked and corrected by senior experts using the questionnaires and the original recorded information. The data files were then processed to produce STATA data-sets, which are weighted and ready for use by researchers⁶. The VLSSs have been used by the GSO, the World Bank and the UN to calculate poverty lines for Vietnam in many reports (see, e.g., World Bank (1999), UNDP(2000)). No serious problems with the

⁵Credit questionnaire at commune level is only available in VLSS1997/98

⁶Calculation of sampling weights can be found in the GSO's Basic Information (GSO (2000) and GSO(2001))

data have been uncovered.

General statistics for households, household heads and communes for VLSS1992/93 and VLSS1997/98 are presented in Table A.2 and Table A.3, respectively. Definition and construction of variables used in the thesis can be found in Appendix E. As stated in the survey basic information documents from the GSO and the World Bank (i.e., GSO and World Bank (2000), GSO and World Bank (2001), Haughton et al (2001)), the VLSS data can be used directly without further weighting. All the results presented from hereon in this thesis come from the analysis of VLSS1992/93 and VLSS1997/98 unless otherwise stated.

Before moving to the next section which present stylised facts of Vietnamese rural credit market, it is appropriate at this point to define some key terms used extensively in the thesis. The formal credit sector consists all organised credit institutions (formal lenders), including Government banks (Bank for Agriculture and Rural Development, Bank for the Poor and other government banks), credit programmes (poverty alleviation programmes, job creation programmes), credit cooperatives, People's Credit Fund, and private banks. The informal sector includes individuals and non-organised credit providers (informal lenders), such as money lenders, relatives, friends, and rotated saving and credit associations (ROSCA). Formal borrowers are individuals or households who take loans from the formal lenders. Similarly, informal borrowers are those who borrow from the informal lenders. Other definitions and variable descriptions could be found in Appendix E.

2.3 Credit Markets for Rural Households

During 1990s, the financial sector in Vietnam was progressively developed in parallel with overall economic transformation and increased demand for credit from private sector. However, similar to other developing countries, the Vietnamese financial system remains largely underdeveloped with a dual structure in which formal and informal sectors exist side by side. At the corporate level, the financial system was

dominated by state-owned commercial banks with over 75 percent of outstanding bank credit (World Bank (2007, p.25)), leaving niche markets for joint-stock and private banks. In contrast to corporate credit market, the formal sector played a smaller role in providing credit for households, especially for poor households in rural areas. In 1993, the formal sector accounted for less than 30 percent of loans while the informal sector provided over 70 percent of loans made to rural households (see Table A.9).

This section will provide a background picture of Vietnam's rural credit markets in the 1990s. We first discuss the general characteristics of households and communes using the VLSS data. Comparison between borrowers and non-borrowers will be made. We then analyse the structure of rural credit markets emphasising on the expansion of formal sector as well as highlighting strong existence of informal lenders over five year between the two surveys. We finally discuss how the markets are divided by looking at the interaction of loan and household characteristics.

2.3.1 Characteristics of Households and Communes in the Sample

The characteristics of rural households are presented in the Tables A.2 and A.3 for the two surveys on 1992/93 and 1997/98, respectively. 1985 households (51.71%) in 1992/3 and 2334 households (54.69%) in 1997/8 are reported to have taken credit within the 12-months period prior to the survey date, regardless of the loan sources.

In 1992/93, the majority of heads of households are male (77%) and married (82%), with averaged age of 45 years. It appears that borrowers are younger than average by almost two years, more likely to be male (79%) and married (84%). The seven percent difference in gender of household head in both 1992/93 and 1997/98 between formal borrowers (82%) and non-borrowers (75%) shows that male-headed households seem to borrow more often than female-headed households. This suggests that men may be more active than women in credit participation. This is also true for married household head, indicating increase of credit activity in more stable

households. The figures also show differences in age, gender and marital status between households who took loans from formal sources and those who took loans from informal sources, although the variations are not substantial. Changes of these variables across the surveys are also small.

The average household size in 1992/93 is just less than five people while number of working adults in household is almost three⁷. Those who are not borrowing have smaller indicators in both terms which, in turn, suggests that smaller size households seem to borrow less. This feature is sensible in agricultural production where most households would use domestic labour in farm work and only hire external labour as the last alternative which is again often in labour exchange (Nguyen, 2007). Bigger households or, more importantly, households with higher number of working adults imply more labour input for production and therefore expansion may be more feasible. If saving is not enough to finance investment, borrowing is always an option, resulting in higher demand for credit. Using this reasoning to explain demand for consumption credit, it appears that large households consume more and again may be exposed to higher demand for consumption loans. In 1997/98, household size (5.17) and working-adult number (2.89) of the borrowing households are also larger than that of the non-borrowing households (4.51 and 2.76, respectively).

The average education level of household head is lower secondary school, or about 6 schooling years. Households who borrow from the formal sector does have higher education than average, but the difference is relatively small.

Approximately 80 percent of rural households work in agricultural production, as expected. A slightly lower number in 1997/98 suggests that off-farm employment becomes more popular and agricultural production plays less important role in the rural economy in parallel with progressing the industrialisation process. Average statistics of commune non-farm employment increase from 46 percent to 55 percent confirms this movement (also see Figure A.1). However, it is important to note that there is a big increase in household's annual-crop-land holding size from about 920

⁷We count any family member age between 16 and 60 as a working adult.

m² in 1992/93 to almost three times higher of 2,643 m² in 1997/98. This transformation was triggered by de-collectivisation of agricultural production started in 1988 which reallocated cultivated land previously farmed in cooperative to households.

Looking at general statistics in Table A.2 and Table A.3, formal borrowers actually acquire larger land holding than non-borrowers and informal borrowers⁸. Intuitively, smaller landholding size households may, on one hand, have less input for investment expansion, and therefore have less demand for credit. On the other hand, less land holding means the households may be poorer or may have less access to credit generally. Either of these will reduce borrowing propensity, as we will see in the next chapters.

In both surveys, statistics on variables of house ownership, total housing area and house value of formal borrowers are higher than those of sample average which in turn higher than those of informal borrowers, although the difference is not remarkable. About 97 percent of households reports legal ownership of their house. This is not a surprise as rural people are less mobile in comparison to urban people. It is typical to see a household with several generations living together and houses transferred between generations. As a house is one of the most popular assets and is often used as collateral for loans, it is in line with our expectation to see that these housing variables are larger for formal borrowers.

2.3.2 The Expansion of Formal Credit Sector

This section describes the picture of market's lenders and the distribution of lenders according to borrower's various sorting category. There are some terms and definitions we would like to make clear before going further. Firstly, borrowers are not individuals but households on debt at the time of the surveys. A borrower will belong to both types of formal and informal borrowers, if he borrows from both formal and informal sources. Secondly, due to the fact that a household could borrow more

⁸We do not have information on household's land-use documents. Annual crop land-holding size variable is the total of land size possessed by household regardless of land title.

than one loan at one time, to simplify we define formal borrowing as sum of all loans that a household borrow from formal lenders, and informal borrowing as sum of all loans that a household borrow from informal lenders. It is assumed that loan characteristics are homogenous for a household within a sector.

As mentioned in the previous section, nearly 55 percent of households in rural areas were borrowing in 1998, up slightly from 51 percent in 1993. Among borrowers, 82 percent of them borrow from one sector, either formal or informal, and 18 percent borrows from both sectors (see Table A.4). The notable point is the expansion of formal sector and the still strong existence of informal market.

According to household borrowing data in Table A.4, over 5 years the market share of informal lenders has been cut in half from about 60 percent to almost 30 percent in number of borrowers, excluding those who borrow from both sources. Formal borrowers have increased to nearly 50 percent from the quite low number of 23 percent. This expansion of formal sector was led by the VBARD and VBP, which lend to 80 percent of formal borrowers and account for 90 percent of total formal loan amount in 1998.

However, the development of formal sector and household borrowing activities are not uniformly distributed across regions. Table A.5 shows that while families from the North Central Coast were most likely to borrow in 1993, those from the South East was the least with about 42 percent. Because the Red River Delta is one of the two rice granaries of Vietnam⁹ and one of the most active economic areas of the economy, it is surprised that households here borrow less than most other regions (actually it came as the second lowest borrowing regions in 1993) especially in comparison with the Mekong Delta. From 1993 to 1998, borrowing became more popular in most regions, except the South Central Coast where the borrowing rate dropped dramatically from 50 percent to 38 percent of its population.

The regional variation also reflects in the sources of credit. In some regions, particularly in the Highlands in 1993, the role of formal institutions was relatively

⁹The other is Mekong Delta

weak with only 25 percent of market share. Formal sector borrowing was most popular in the North Central Coast at that time. The picture however changed dramatically by 1998. Formal institutions became almost three times more popular in the Highlands, lending to 70 percent of borrowers. The increasing role of formal credit in the Highlands could be interpreted as the result of government economic policy in an attempt to balance economic development among regions (World Bank, 1999). Market share of the formal sector also increased impressively in other regions. The smallest market share was in the Red River Delta of 53 percent. Informal lenders seem to be crowded out by the formal sector expansion although retaining its prominent role.

In addition to the fragmented picture of borrowing activities across regions, credit access also diverse among borrowing households categorised by quintile expenditure per capita (see Table A.6 and Table A.7). In the formal sector, the most often borrowing tends to come from poorer household and the biggest amount of borrowing comes from the middle quintile households. The former might result from higher demand for consumption credit of poor households, which is asserted by its outstanding rate of informal borrowing. The small share of formal credit for this quintile is no surprise. Conversely, households from upper quintile borrow less frequently but when they borrow they do borrow more often from formal banks and borrow larger amount (A.7). This result implies that richer households seem to have greater access to formal credit as they would have more valuable collateralisable assets.

2.3.3 Features of Loan Contracts

The discussion so far has presented the structure of rural financial system with the complex coexistence of formal and informal sector. In what follows, we inspect the differences between formal and informal lending attributions by looking at characteristics of loan contracts provided by each sector.

Interest rate

The characteristics of loan contracts including interest rates are presented in Table A.9 and Table A.10. In the formal sector, 75 percent of loans in 1993 and 95 percent of loans in 1998 are charged with positive interest rates. By contrast, in the informal sector, where friends and relatives are the main sources of credit, nearly 90 percent of loans from relatives in 1993 and 1998 are interest-free. However, as one could expect, 97 percent of loans from money lenders included in the two surveys is charged interest, sometimes at extremely high rates. In the study of Thailand's rural credit market in 1984-85, Siamwalla et al (1990) find that money lenders charge interest rates of up to 120 percent annually. In another case, Aleem (1990) also reports informal interest rates in rural areas of Pakistan of about 80 percent a year. It is no different in Vietnam.

In 1992/93, the average nominal annual interest rate charged by formal credit institutions was 46 percent. Within the formal sectors, VBARD and VBP provide loans with the lowest interest of 42 percent while private bank and other government organisations, e.g., job creation programmes impose 60 percent and 80 percent interest rates, respectively. It is no surprise to see that informal loans are charged at much higher interest rates. One could possibly borrow a non-interest loan from relatives or friends, but if one has to pay interest one would have to pay more with informal borrowing. The average annual interest rate for informal loans was 92 percent, within which money lenders charging 110 percent (Table A.9).

The expansion of formal credit sector, more availability and higher accessibility of formal funds may be part of the explanation for the cut in interest rates by almost two third of the 1993 rates. In 1997/98, formal interest rates was 15 percent on average and the informal interest rates was 38 percent. VBARD and VBP continues to be the cheapest sources for interest-bearing loans. Money lenders are of the most expensive, charging 50 percent annually on their credit (Table A.10). Relatives are the most frequent lenders providing interest-free loans, although there might be other hidden costs that we do not observe such as social and other obligations attached to such loans.

The VBARD and VBP's interest rates are determined within the interest rate framework specified by the State Bank of Vietnam (SBV). With the dominant role in the formal sector, the impact of VBARD's interest rates is crucial. A survey by McCarty (2001) finds that 50 percent of institutions referred to VBARD's rate as the guide interest rates while only 20 percent of them decided interest rates basing on operational cost and financial sustainability. These exogenously determined interest rates potentially have significant effect on the business of lenders. As it has been generally accepted in rural finance literature (see, e.g., Hoff et al (1990)), interest rates could be used as an indirect screening mechanism for lenders to select less risky borrowers. By setting interest rate without considering local market conditions and applicants' characteristics, lenders are unable to ration out risky borrowers without employing other screening methods. Consequently, formal lenders usually consider collateral assets as the effective tool against risk and to maintain the repayment rate. Credit guarantees from local government administration is also applied but far less often.

In terms of interest rates only, formal loans appear to be much cheaper than informal loans. So, why are some households willing to pay higher interest rates to borrow from informal lenders? Or put another way, why do people accept to pay 3.2 percent a month for a loan if another loan is available at 1.2 percent? One possible explanation is formal sector does not supply credit to satisfy all of the demand. There are households being refused to access formal credit services and therefore have to turn to informal credit sector.

Loan duration, loan size and collateral

In a loan contract, there are other important attributions beside interest rates, including the amount and length of the loan, or the requirement of collateral. Generally, the formal sector provides loans with larger amounts, longer durations but more often requiring collateral compared with the informal sector. In 1992/93, the average size for loans from the formal sector was VND906 thousands, in which the largest size of VND1,090 thousands came from VBARD loans¹⁰. Loans from the

¹⁰VND1,090 thousand equals USD97. VND//USD exchange rates in the 1990s are in Table

informal sector were smaller on average, VND820 thousand, but money lenders appear to provide the biggest average loan amount of VND1200 thousand, suggesting the limited availability of formal funding by that time. Additionally, the length of formal loans are observed to be almost 2 months shorter than the length of informal loans. Especially, VBARD loans are lent on average of 6 months duration which is a surprise as 87 percent of VBARD loans are reported as capital for investment (see Table A.12).

The pattern of loan contracts, notably the differences between the formal and the informal, were changed remarkably from 1992/93 to 1997/98. Formal sector provides longer loan duration while keep lending at larger loan size. VBARD became the largest credit provider with the average loan amount of VND4,100 thousand for the length of one and a half year. This suggests that formal lenders have aimed to provide long-term investment credit. For the loans obtained from informal lenders, the average term to maturity remained short at about 10 months, with the average amount of VND2,300 thousand. Due to its intrinsic characteristics of short-term, relatively small amount and high interest rate, more than 50 percent of informal loans are for consumption, especially loans from relatives with nearly 70 percent in 1998 (see Table A.12).

Although the formal loans appears to be cheaper, in term of interest rates, a large proportion of population still matches their demand with informal borrowing. One important barrier to access the formal sector is the requirement of physical collateral and/or guarantors. As presented in Table A.9 and Table A.10, 50 percent and 75 percent of formal loans, in 1993 and 1998 respectively, ask for collateral assets, such as house, land or durable goods. The collateralising rate is even higher with VBARD, at 76 percent and 87 percent in that two years. Land and house are the most common type of collateral, accounting for 90 percent of all cases (Table A.11). In contrast to the formal sector, the informal sector usually does not require collateral for their loans: only 5.4 percent of all loans in 1993 and 3.7 percent in 1998.

A.1.

Because of collateral requirement and other obstacles, e.g. bureaucracy and red tape, some rural households may not even try to apply for formal loans as they may preemptively assume that their applications would not be approved. In addition, even if some poor households may be able to provide collateral, many of them are simply afraid to risk using their land as it is their only means of subsistence UNDP (2000). This phenomenon is called credit self-rationing, where some people considered themselves out of formal credit market a priori. Therefore, although being considered as the cheaper source of credit, the formal sector would not be accessible to all households.

Thus, even though informal lenders provide credit that normally carries higher interest rates, shorter term and smaller size than credit from formal institutions, they still find demand for their loans. This reflects the segmented characteristic of Vietnamese rural credit market where particular lenders serve clients of their own market. It is widely reported in the literature that poor rural households are often less creditworthy. Lending in rural areas is arguably more expensive due to high transaction cost, small loan size and remoteness of clients. Lending to poor rural households is even suffered from more drawbacks as they are less likely to have collateralisable assets and lending without collateral is just riskier. As a result, formal lenders have less incentive to expand their operations in rural parts, specially in remote areas. Thus, a significant proportion of households may be left with no other option but finding their credit supply from the informal market.

2.3.4 Linkages between Formal and Informal Credit Sectors

The interactions and coexistence alongside each other of formal and informal credit sector in underdeveloped rural credit market have been well reported in the economic literature (see, e.g., Bell (1990), Bell et al (1997), Hoff & Stiglitz (1997), and Jain (1999)). As discussed in the previous section, the Vietnamese rural credit market has been divided into almost two halves in 1998: 49.7 percent of the number of loans was from the formal sector, and 50.3 percent from the informal sector (see Table

A.10). Although the two sectors typically compete with one another, they may also complement each other in a number of fundamental ways. For example, informal lenders could provide credit for households who are involuntarily excluded from formal credit services. Households who could not satisfy their demand by borrowing from institutional market, or in other words those who are partially rationed by formal lenders, could find their supplemental credit from informal sources. Jain (1999) suggests an explanation for this phenomenon using the informational differences between the two sectors that formal lenders may intend to partly finance some borrowers and therefore forcing them to seek informal lenders for the remainder of their credit demand. By doing this, formal institutions could indirectly screen borrowers over the information advantages of informal lenders. The evidence of the spillover of credit demand and the complement of formal and informal market in rural Vietnam is revealed in table A.4, represented in the stylised fact that 16.27 percent and 18.59 percent of indebted households in 1993 and 1998 borrowed from both credit sectors. Another explanation for this simultaneous two-sector borrowing is that formal borrowers may face unexpected consumption money demand such as unforeseen events of illness, accident or funeral. It is less likely that formal institutions are willing to supply consumption credit, if any, without time-consuming complex procedures. Hence the quick and simple way to obtain cash is from informal sources, including relatives and friends.

Additionally, formal institutions could finance excluded households indirectly by lending to local money lenders who then onlend to those formally rationed. This interaction between the two sectors has been studied theoretically (e.g. Varghese (2005), Floro & Ray (1997), Andersen et al (2006)) and empirically (e.g. Bell et al (1997), Gine (2005)) emphasizing the advantages of local lenders on screening and observing borrowers' behaviours. In the case of Vietnam, we find that among all household lenders, almost 49 percent of them are also borrowing from a third party in both year 1993 and 1998¹¹. This finding is striking. It implies a complex picture of credit activities in rural areas. How do we explain the existence of such credit

¹¹There are 654 household lenders in 1993 in which 315 of them (48.99%) also borrow. The numbers in 1998 are 655 lenders and 323 borrowing lenders (49.31%).

agents? And what are their roles in the rural credit market?

There are several possible explanations. In one case, borrowing lenders may have good access to formal sector, hence, have access to low interest credit. Then they can lend to other households or individuals using the credit they borrowed and make profit from this lending. Given the large differences between formal and informal interest rates¹², there are incentives for people to become the middlemen. Especially in markets where credit institutions fail to satisfy the whole demand for credit, the existence of credit intermediaries is an important channel to supply funds to creditworthy households. In other cases, borrowing lenders may have lent their money and later have to borrow from other sources in the unexpected event that they could not re-collect their lent money in time. This kind of consumption smoothing borrowing is also observed by Udry (1995) in his study of the Nigerian rural market.

Statistics on source of credit for borrowing lenders and other borrowers¹³ are presented in Table A.13. It reveals that about 42 percent and 61 percent of the borrowing lenders borrow from formal sector in 1993 and 1998 respectively. These numbers are not as high as expected. Specifically, borrowing lenders do borrow more often from VBARD (74.24 percent in 1993 and 83.82 percent in 1998) than other borrowers¹⁴ (64.26 percent and 81.71 percent) but in general, borrowing lenders even borrow less from formal source than normal borrowers in 1998. Borrowing from relatives and other individuals is surprisingly common. This evidence suggests the hypothesis that the financial intermediation may occur not only from the formal sector to the informal sector but also within the informal sector. Borrowing lenders may not be necessary to have better access to formal credit. However if they do, one may expect to see the monopoly rent on top of the loans they onlend to others. This is often called the profit of information advantages or in other cases, the premium for the risk that borrowing lenders have to bear while lending to less creditworthy

¹²The interest rate differences between formal and informal sector are about 46 percent and 23 percent in 1993 and 1998 respectively (see table A.2 and table A.3 for details)

¹³Borrowing lenders are households that borrow and lend simultaneously. Other borrowers include households that borrow but do not lend.

¹⁴Those borrowing but not lending

households.

2.4 Conclusion

This chapter aims to provide an overview of the Vietnamese rural credit markets in 1990s. It also presents description and definition of the Vietnam Living Standard Survey in 1992/93 and 1997/98, which will be the primary data for our analysis throughout the thesis.

The Vietnamese economic reform started by the end of 1980s had achieved a remarkable progress during 1990s, over the time of the VLSS surveys. Agricultural transformation and rural development were recognised as two of the most successful results of "*Doi Moi*". In just 5 years from 1992 to 1997, poverty rates were reduced from 58% to 37% of the population.

Evidence from the VLSS data shows a big expansion of the formal credit sector in rural Vietnam. From a low 30% market share in 1992/93, the formal sector had grown to account for 50% of the rural credit market in 1997/98, although the development is not well-balanced across regions and households quintile by expenditure per capita. Formal funding is cheaper in term of interest rates but more often requires collateral in comparison with informal source. Average Interest rates of both sectors had been cut by almost two third over time between the surveys, indicating a growing competition between lenders. However, the existence of informal lenders remains strong, presenting a complex linkages between agents in the credit markets.

Chapter 3

Access to Credit in Rural

Vietnam: A Partial Observability

Model

3.1 Introduction

As pointed out in the previous chapter, there was a considerable movement of borrowers in rural Vietnam from informal to formal sector during 1993 to 1998. The expansion of the formal credit sector is reported in both loan numbers and loan amount provided, and all sample communes are serviced by at least one formal credit institution. However, even when formal credit is available to people of a commune, not everybody can borrow, or at least there are people who cannot borrow as much as they would like to. If we assume that formal finance is more desirable and most people would choose formal finance as their first priority when they demand investment capital, then the unsuccessful applicants could be considered as riskier or weaker (in some aspects such as poorer or not targeted) than successful applicants. The credit rationing mechanism and selection process of formal institutions will push those failed applicants back to informal sources.

This chapter aims to study the determinant factors of credit participation in

the formal and informal credit markets. The decision to enter the credit markets well depends on both supply and demand factors. In turn, demand and supply are simultaneously affected by a number of variables. For example, large land possession may increase access to formal credit sector as land can be used as collateral for loans. But land will also affect total demand for credit, due to the potential higher investment. Considering both sides of the market and disentangling their effects is essential to understand how internal and external factors would behave in determining a household's credit decision.

Studies of credit markets have often been based on evaluation of borrower's characteristics to construct probabilistic model to estimate household's propensity to borrow. This approach has been applied widely in studies of developed countries' household debt markets (see, e.g., Cox and Jappelli (1993), Leece (2000)), and the developing credit markets, more recently (see, e.g., Diagne (1999), Ravi (2005), Swain (2002)). In this chapter, we also use this type of analysis together with instrumental variables to improve the identification distinction between determinants of supply and demand.

Of the earlier empirical literature of rationing in rural credit markets, Kochar (1997) and Bell et al (1997) are the closest to this thesis. Both papers study the determination of credit rationing in the formal sector in rural areas with consideration of the informal credit supply. Kochar estimates the extent to which rural households are constrained by government regulations which control access to production credit from government or formal sources. Bell et al estimate the extent of rationing in the regulated sector and the associated spillover of demand into the unregulated market. Similarly to us, the authors have to deal with problem of limited information or partially observed data on non-borrowing households. The econometric technique used by Kochar and Bell et al is learnt and applied in our work with appropriate adjustment, especially in the identification strategy.

Very few studies have analysed determination of credit participation by sectors in Vietnam. Majority are general reports from international organisations, such as the

World Bank or the Asian Development Bank, which have focused on discussion of data, and stylised facts of rural credit development (e.g., World Bank (1999), World Bank (2006)). Duong and Izumita (2002) examine rural household participation in the Vietnamese rural credit market using a small sample of 300 households collected by the authors. They found that the rural credit market in Vietnam is quite segmented, and that the formal sector specializes in lending for production purposes whereas the informal sector's lending is more diverse. Pham and Lensink (2007) use the VLSS data to assess lending policies of formal, informal and semiformal lenders for households in Vietnam. Urban households are included in their study. They found collateral, a guarantor and/or borrow for business-related activities will increase access to the formal sector, while household head being female increases the probability of participation in the informal market.

The contribution of this chapter is to fill the gap in literature by evaluating determinant factors of household participation in the rural credit markets using the VLSS data. The features of a transition economy like Vietnam make our analysis particularly interesting because of the changing in economic environment and legal regulations over the time of the survey, i.e., the Land Law that came to effect by December 1993, between the two survey dates. We expect these changes in the economic conditions will be reflected in the changes of credit behaviour of rural households.

The structure of this chapter is as follows. The next section will outline our definition of credit access and credit rationing used through out the thesis. Section 3.3 presents an univariate probit model to estimate probability of participation in each sector of credit, assuming that all rural households have positive demand. This helps to focus our analysis on the determinants of supply. Section 3.4 develops the bivariate probit model with partial observability to disentangle the supply and demand forces, which jointly determine probability of credit participation. The last section concludes.

3.2 Credit Access and Credit Constraint

Despite the frequent use of the terms "credit access" and "credit constraint" in the economics literature, it is not clear that these terms are always employed to refer to the same phenomenon. To avoid confusion, in this section we define meanings of these terms that we use in the thesis.

In the last three decades, following the lead of Hall (1978), a number of succeeding papers (e.g. Hall (1982), Hayashi (1985a, 1985b), Maki (1993) or Zeldes (1989)) have theoretically and empirically rejected the validity of the traditional Life Cycle-Permanent Income Hypothesis (LC-PIH) proposed by Friedman (1957) as an explanation of optimal consumption behaviour. The most important assumption of the LC-PIH is that consumer is constrained only by the lifetime budget constraint, so that consumption can be shielded from period-to-period fluctuation in income through borrowing and lending (Hayashi, 1985a). However, it has been becoming a consensus among economists that there exists a proportion of population facing credit constraints¹. In other words, far from the perfect world of fully functioning credit markets where households are insured from transitory income shocks by borrowing, credit markets are often imperfect especially in rural parts of less developed countries. In the imperfect markets, there are consumers who would like to borrow but are refused by lenders or are not able to borrow as much as they would like to at the market clearing interest rate.

In our study, the data does not allow us to identify directly households who practically rationed by lenders. In the credit section of both surveys, households were merely asked to report if they borrowed money from anyone within 12 months prior to the interview date. If the household interviewee answered "no", then there was no attempt to ask whether he had been refused by lenders or he did not borrow simply because of having no demand for credit. If this had been done, we would know directly from the surveys the level of credit constraints in the Vietnamese rural credit market.

¹Or, to use James Tobin's terminology, liquidity constraints.

Empirical models testing for the presence of credit constraints based on the LC-PIH has been among the most popular approaches. However, this approach is not feasible in our study because of the limitation of data. In general, empirical testing of the implications of the LC-PIH requires repeated observations on the same household over several time periods in order to estimate the level of household permanent income and its consumption behaviour. The two-period panel data from VLSS is short and the five years gap between the surveys will make our study inconclusive.

Thus, instead of pursuing the LC-PIH approach, we will estimate the level of credit constraints using a bivariate probit model which allows us separately estimate demand for and supply of credit for each household. And households whose estimated demand for credit is not satisfied by the estimated supply will be considered as the credit constrained households. In this framework, the proxying permanent income of households are incorporated into our model by a number of observable variables such as household head's education, employment and landholding size.

One good explanation for non-borrowing households is that lenders may not operate in the region in which the household resides, particularly in the case of the formal credit market. Data shows that 115 communes (out of 120) in 1992/93 and 135 communes (out of 135) in 1997/98 have formal institutions operate in their areas and providing financial services to people living in that commune. It appears that formal credit is available in, more or less, every commune sample in 1997/98, implying an uniform access at commune level across Vietnamese rural regions. This finding, however, does not suggest the same accessibility among households in one or different commune. Also, one could believe that living in a particular area may enable household to better access to formal institutions.

Access to credit is often confused or used interchangeably with participation in credit programmes in many studies. The crucial difference between the two concepts lies in the fact that access implies ability to borrow, i.e., availability of formal credit programmes and satisfying their eligibility criteria, while participation is something

that households choose to do freely after being able to access to formal credit. In other words, participation is more of a demand-side issue related to the potential borrower's choice of the optimal loan size while access is more of a supply-side issue related to the potential lender's choice of the maximum credit limit. In our study, a household would fall into one of four observable financing regimes as follows:

1. Doesn't borrow
2. Borrows from the formal sector only
3. Borrows from the informal sector only
4. Borrows from both credit sectors

Households who do not borrow could either (i) have access to credit but no demand; or (ii) have no access to any desirable source of credit. We cannot observe whether a non-borrowing household falls in the former or the latter case. Therefore, to avoid confusing the credit accessibility (supply side) and the credit participation (both demand and supply), we define household credit access as the actual credit participant who reported a positive amount of debt at the time of the survey.

In addition, it is noted that households who borrow from both formal and informal markets may be a source of disturbance in study of credit markets. However, controlling for this potential problem is complicated, due to the difficulty of finding good identification for borrowers's choice between the two sectors and the significant computational problems. Therefore, it is common in empirical studies of credit markets to ignore or assume away any noise caused by two-sector borrowers. Or even when the two sector choice model is estimated, confidence in the results are relatively low (see e.g. Kochar (1997), Bell et al (1997), Conning (2001)). This study is not exempted. Because households borrowing from both sectors of credit account for less than 20 percent of total borrowers (16.27% in 1992/93 and 18.59% in 1997/98, see Table A.4), we assume that these households do not have significant impact on our estimation, and they will be used as part of the studied sample for both sectors.

3.3 Estimating Determinants of Credit Participation by Univariate Probit Model

This section will develop an empirical model of rural credit market outcomes considering the joint determination of household's demand and lender's supply. The model concentrates on participation decisions, ignoring the information contained in borrowing volume which will be discussed in the next chapter. In what follow, we develop and estimate econometric models to understand the impacts of household and commune characteristics on credit participation.

The model incorporates the following structures and assumptions:

(i) The observed borrowing activities are jointly determined by the system of credit demand and supply which are, in turn, affected by household and commune characteristics

(ii) Households could only borrow from either formal, or informal, but not both credit sectors.

We start with a simple univariate model. We then extend our analysis to bivariate model to capture the complexity of the determinants of demand and supply.

3.3.1 Econometric Specification

Let D_{ij}^* and S_{ij}^{*2} , respectively, be latent variables of notional credit demand and credit supply of the household i in the commune j . Thus, the two variables will be expressed in a simplified form as

$$\begin{aligned} D_{ij}^* &= H_{ij}^d \beta + u_{ij} \\ S_{ij}^* &= H_{ij}^s \gamma + v_{ij} \end{aligned} \tag{3.1}$$

²Depending on the dependent variable of the equation to be analysed, we will understand D and S as demand and supply of formal credit, or informal credit, or simply credit regardless of source. In case of not being clearly stated, we should think of D and S as a general term of demand and supply of credit.

where H^d , H^s are the vectors of exogenous household and commune variables which respectively affect the household demand for credit and the supply which lenders is offered to the household; β and γ the vector of coefficients to be estimated; and u and v the normal distributed errors for demand and supply of household credit respectively.

Set D_{ij} and S_{ij} as dummy variables of demand and supply, where

$$\begin{aligned} D_{ij} &= 1 \text{ if } D_{ij}^* > 0 \\ &= 0 \text{ if otherwise} \end{aligned} \quad (3.2)$$

and

$$\begin{aligned} S_{ij} &= 1 \text{ if } S_{ij}^* > 0 \\ &= 0 \text{ if otherwise} \end{aligned} \quad (3.3)$$

In case of perfect or full observation, we will be able to observe values of both demand and supply, and therefore observe the equilibrium of borrowing. Loan amount that a household actually acquired is the intersection of demand and supply line i.e. $D = S$.

However, in our study, due to insufficient information provided by the survey data, we observe neither D or S . Instead we observe the value of B_{ij} , which is the product of D and S , the dummy variable of credit participation³

$$\begin{aligned} B_{ij} &= 1 \text{ if household } ij \text{ is a borrower } (D_{ij} = 1 \text{ and } S_{ij} = 1) \\ &= 0 \text{ if otherwise } (D_{ij} = 0 \text{ or } S_{ij} = 0) \end{aligned} \quad (3.4)$$

That is, a household will be observed as non-borrower even if his demand for credit is positive (but supply equals zero). In order to interpret the estimation results effectively, in a similar framework, Kochar (1997) makes an assumption that all

³Generally, we define B as household borrowing status. If we apply B by source of credit, then we should also interpret D and S by credit source accordingly.

households of the survey villages are having demand for credit. By assuming this, she drives the demand function away (i.e., D_{ij} always equals one) and hence the credit activities are solely determined by credit supply. However, this assumption appears to be strong and does not reflect the reality. Therefore, instead of following Kochar to make further assumption on demand and supply functions, we allow the model of probability of being credit participant to be estimated by a single equation which incorporates determinants of both demand and supply, as follows

$$P(\textit{participation}) = P(B_{ij} = 1|H_{ij}) = \beta_0 + \beta_1 H_{ij} + \varepsilon_{ij} \quad (3.5)$$

where β_0 is the constant; β_1 the conformable vector of coefficients to be estimated; H_{ij} vector of household and commune characteristics: $H_{ij} = (H_{ij}^s, H_{ij}^d)$; ε_{ij} the normal distributed estimation error.

3.3.2 Variable Selection

This section describes various dependent and explanatory variables that are used in estimating the econometric model mentioned above. The following chapters in this thesis will also refer to this section as a reference for definition and construction of dependent variables. Following the discussion in Section 3.2, we construct the dependent variables into two categories:

- (i) CRF is dummy variable for borrowing from the formal sector. This variable captured borrowing activities of households with lenders from the formal sector⁴. CRF takes the value one if household was borrowing from formal lenders at the time of the survey or within 12 months before the survey date. It equals zero otherwise. Settings are the same for both survey data in 1992/93 and 1997/98.
- (ii) CRI is dummy variable for borrowing from the informal sector. Similarly to CRF, CRI captures borrowing activities of households with lenders from the

⁴Definition and list of lenders of formal sector are in the Appendix E

informal sector. CRI equals one, zero otherwise, if household was borrowing from informal lenders at the time or within 12 months before the survey date. Again, settings are the same for both surveys.

The set of explanatory variables includes three categories:

- (i) Characteristics of head of household, i.e., age, gender, marital status, education level, and health condition;
- (ii) Characteristics of household, i.e., household size, ratio of number of working adults over household size as the proxy for labour power, working in agriculture production, house ownership, and irrigated land holding size;
- (iii) Characteristics of the commune that household residing in, i.e., commune population, formal credit per capita, distance from commune to the nearest formal bank branch.

Definition and construction of these variables can be found in the Appendix E. Summary statistics are in table A.2 and table A.3.

Both the formal and the informal credit supply are functions of these variables which also determine the household demand for loans. Variables describe characteristics of household head control for the household level human capital which are expected to have effects on both demand and supply for credit. The variables that identify supply-demand functions will be discussed later.

Married household head, a sign of household stability, is expected to have positive impacts on credit participation for both sectors. This positive impacts could come from both demand and supply functions, as a family household would be in higher propensity to demand for credit to keep the family income stable. Lenders, on the other hand, may consider married households a safer borrower hence may be more willing to fulfill the credit demand. Gender of head may also have impacts too. In some areas, households with male head may demand higher credit and lenders

may also tend to lend to males due to requirement of heavy labour in agriculture production in rural regions. However, it is possible that some formal credit programmes target female as their priority borrowers⁵. In that case, female household head may have higher participation probability in the formal sector.

While variables of age, gender and marital status of household head become standard variables to include in a household regression study, the education variable seems to have more to deliver. The coefficient for the variable education is expected to have a positive sign on the formal sector participation. More educated households may find more opportunities to invest and may be more confident to borrow. Also, being different from the informal sector where credit contracts are simple and sometime in oral form, the formal institutions normally require more complicated credit application and always in paper form. Better education may help household understand better the application procedures, thus the formal sector is more approachable. Similarly on the supply side, credit institutions may consider educated households as higher creditworthy customers and make credit better available to these people. The education variable of head's schooling years and schooling years squared are used to capture the educational level. The square form is to control for non-linearity of education.

Variables capturing household size and number of working adults are expected to generate positive signs. Large size households consume more but at the same time, they may have more domestic labour if the rate of working adults in the family is high, and hence have more input for production expansion. The expansion may lead to demand for credit. The dummy variable of agricultural employment (equal one if working in farm) reflects possibility of self-investment and therefore is expected to positively influence the credit participation.

In addition, facing the problem of limited information about loan applicants, one of the most popular solution for lenders is to require collateral. This crucial

⁵There is a line of literature stylised this gender targeting and its impact on credit access (see (Morduch 1999) for the review). Grameen Bank is the most famous microfinance institution that initially established to lend to female only.

condition is particularly high in the formal sector. As mentioned before, 75 percent of formal borrowers in 1997/98 have deposited fixed assets as collateral. The number is even higher if borrowing from VBARD or VBP where almost 90 percent of loans are collateralised. Of the collateral asset, land and house account for 90 percent. Therefore, we incorporate the variables of irrigated land holding size and house ownership into our model to proxy for availability of collateral and expect their coefficients to be positive.

Finally, we would like to discuss the distance variable, which measures distance from commune that household resides in to the nearest government bank branch. The distance variable is only available in VLSS1997/98 with value ranging from 0km (bank branch locates very closed to commune committee) to 45km. Average distance is 7.6km, median is 6km.

Recent economic research have demonstrated the importance of distance in explaining availability of formal credit. Sussman and Zeira (1995) shows that the cost of monitoring a borrower increase with borrower-lender distance. Consequently, if formal lenders are prevented from charging different prices to different borrowers to cover extra monitoring cost, they may choose not to lend to distant credit applicants. Formal lenders may also face severe problem of asymmetric information which comes together with adverse selection and higher risk. All increase with distance⁶. Following the literature, we argue that the distance hinders observability of formal lenders over households which eventually inflate both cost and risk of lending operation. As a result, formal institutions will be more cautious and less willing to lend to households in more distant communes. This clearly implies a shift in credit supply affecting household's participation.

In addition, from the borrower's side, the further the distance to formal banks, the higher the cost of capital borrower has to bear. In credit rationing theory, uninformed lenders are likely to charge higher interest rates to remote borrowers in order to compensate for the risk arising from adverse selection problem. Even

⁶See, for example, Dell'Ariccia (2001), Degryse and Ongena (2005) for further discussion

in the case that interest rates are unchanged, distant borrowers will have to bear the cost of transportation to bank branch (may be in term of time spent - cost of working hour wasted). It is also possible that distant households would become less informed of formal credit availability, hence would not have proper demand. In sum, distance variable is expected to reduce formal credit participation from both supply and demand sides.

The impact of distance on credit, however, may not be linear. As long as formal banks cannot observe borrower's locations, distance becomes indifferent (Degryse and Ongena, 2005). To capture this effect, we reclassify information on distance into three categories: DIST1[0-4km], DIST2[>4-10km] and DIST3[>10km]⁷.

Other variables, such as commune population, commune-level formal credit per capita, and health condition of household's key members are also included in our analysis. However, because these variables will be used as instrumental variables to identify supply and demand channels, and in order to make our argument focused, we will discuss these variables later in the separated section 3.4.2. Description of dependent and independent variables included in this chapter is in Appendix A.

3.3.3 Estimation Results

The univariate probit model is estimated as in equation 3.5. Despite limitations of this simple model, as will be discussed in the following section, results from this model will shed light on factors determining household's probability of participation to credit sectors, especially participation to the formal sector.

Estimation of Model I is applied to both databases, VLSS1992/93 and VLSS1997/98, with cluster robust estimator of variance. *Schooling years of household head* and *distance variables* are not available in 1992/93 data. The results are reported in Table B.1. Our interpretation of the results concentrates on participation in the formal sector.

⁷Do et al (2007) use dummy of distance <5km. However, in our case, distribution of households in each categories would be more equal. Mean values of DIST1, DIST2, DIST3 are, respectively, 0.39, 0.38, and 0.23.

Results are mixed between the two surveys. In 1992/93, the coefficients corresponding to *household size*, *ratio of working adults*, *health condition*, *land size*, and *commune credit per capita* are found to significantly affect household's participation to the formal credit sector. Most of these coefficients are also significant in 1997/98, except *health condition*. In addition, in 1997/98, the coefficients for *age of head* and *age squared*, *schooling years of head* and its squared, *household working in agriculture*, and *house ownership* are significant on the formal sector participation as well.

The coefficients for *gender of head* (equal one if head being male) have positive signs but are not significant at conventional level in both surveys, suggesting an indifferent effect of gender on formal borrowing. The signs of the coefficients for *household size* and *ratio of working adults* are positive at 1-percent significant level, as expected. Given the employment nature in rural Vietnam where agricultural production dominated, more labour available in a house is clearly an advantage making agriculture projects easier to form and implement. Without hiring extra labour, a small family may not have motivation and capacity to expand the family business, reducing demand for credit. Moreover, formal institutions may also be more willing to lend to large household size, with its potential of higher labour input. All these affect the formal sector participation.

The first difference between the 1992/93 and 1997/98 results is the signification of the coefficients for *irrigated land holding size*. In 1992/93, land size positively and significantly increase formal sector participation. This could be a result of better access to formal lenders due to availability of collateral assets, but it could also simultaneously stem from high demand for investment capital. Although the coefficient on this variable is still positive in 1997/98, it becomes less significant (10% level). Households with larger land holding may no longer have higher access to credit institutions. Again, it may be because of changes (tightening) in collateral policy which squeezes formal credit supply.

As regards to the use of fixed asset as collateral in less developed countries,

De Soto (2000) argues that one of the key reasons why poor people are excluded from main stream finance and then left out of economic development is that poor countries often lack of the legal entitlement to private property. As a result, without legal entitlement, assets such as land and houses that poor people may possess are difficult to trade, or more importantly, cannot be used as collateral for initial loan they demand. In such circumstance, value of property become less important as presented in table A.2. House value and house area of borrowing household are even smaller than that of all sample mean value. Unfortunately, we do not have data on land and house legal ownership certificate of sample households, we could not conclude on how the impact of legal entitlement on borrowers. However, in a study by Quy-Toan Do that uses land right registration number in different survey concludes that legal protection of private property indeed open many doors for rural households including accessibility to formal credit sector. Of the incentive and demand for investment credit, borrowers do seem to possess larger agricultural land size than mean of all sample. However, the single credit equation estimation does not allow us to clarify whether the less important role of land possession in determining formal credit access stemming from reduction in supply or demand. Further discussion, therefore, will be presented in the next section.

The second discrepancy is the significant level of the coefficients for *health condition* which are again positive and significant at 1-percent in the 1992/93 regression, but insignificant in the 1997/98 regression. As discussed previously, the higher value of the health variable reflects the worse health condition of the household members, and possibly higher demand for consumption credit. Positive sign of the coefficient means that households in 1992/93 seems to borrow more often from the formal sector if health condition of the household is not good. Although this result may sound peculiar, it does conform to what we have observed from the data. In 1992/93, almost 30 percent of all formal loans is for consumption purposes, which including loans to pay for health service expenses. Credit cooperatives, private banks and other formal credit institutions even provide more than 50 percent of their loans for consumption (see Table A.12). Together with the low collateral requirement of

the formal institutions⁸, the formal sector in 1992/93 seems to be approachable by households if they had consumption or urgent cash demand. This explains why bad health conditions lead to higher formal borrowing in 1992/93, although the marginal effect is small⁹. The situation in 1997/98 is different. Formal institutions now mostly provide credit for investment purposes¹⁰ and request collateral for their loans more frequently, which means borrowing for consumption becomes less popular and takes more time to process (e.g. more paper works related to collateral asset), even if possible. Indeed, bad health condition households may well find it more difficult to obtain consumption loans from the formal sector in 1997/98.

Other variables which significantly affect borrowing probability from the formal sector in 1997/98 are *age*, *education*, *employment in agriculture*, *house ownership*, and *commune-level credit per capita*. The coefficient of age squared in the regression is significant and negative which implies an inverse u-shape impact of age on formal borrowing activities. The result suggests that households having head at the age of 48 are most likely to obtain formal loan¹¹. Above that turning point, the older the household head, the lower the credit activity. This reflects that older people seems to be more settled, less productive, and are less likely to take new investment initiatives. Demand may play a more important role than supply here.

Similarly, to account for the non-linear effect of education, squared value of schooling years is included in the regression which also obtains negative and significant coefficient. This result casts an interesting implication. Unlike the conventional wisdom that higher education gives household better access and then higher probability of participation to the formal sector, in rural Vietnam, household heads with seven years in school or secondary school education appear to borrow more frequent

⁸The percentages of collateralised loans by credit cooperatives, private banks and other formal credit institutions are 17%, 5% and 8% in 1993, and 75%, 79% and 20% in 1997/98 respectively. See table A.2 and table A.10 for details.

⁹The marginal effect equals 0.0031

¹⁰Only 17 percent of all formal loans is for consumption, which is mainly from credit cooperatives.

¹¹Holding all other variables fixed, the relationship between the dependent variable *probability of formal credit sector participation* $Pr(B)$ and the variables *age* and *age_2* can be expressed in the function: $Pr(B_{ij}) = \beta_1 Age_{ij} + \beta_2 Age_{2ij} + v_{ij}$. This parabol function will achieve its maximum value of $Pr(B_{ij})$ at $Age = \frac{\beta_1}{2\beta_2}$ where the first derivative of $Pr(B_{ij})$ with respect to *Age* equals zero.

than household heads with college or university education, or households with very low or no education at all. A possible explanation is that it is easier for people with high education to find a paid job. These people may choose not to work in farm or small self-business, therefore even if they have better access to the formal sector, their low demand for investment credit decreases the borrowing rate. Or simply, people with higher education may be more likely wealthier, or at least less poor than low education people, hence demand for credit is not expected to linearly increased with education level. On the opposite direction, household heads having no or little education may find themselves excluded from the formal credit services. This phenomenon, however, has been discussed widely in the literature (see, e.g., Aghion (2005), Harper (1998)). We will come back to this latter when estimations of supply and demand equations are separately done.

Employment in agricultural production also increases borrowing from formal sources in 1997/98 while this coefficient in 1992/93 is positive but insignificant. This result may reflect the impacts of economic development, the decollectivisation of agricultural production, and the major change in the new land regulation enforced in late 1993, which eventually created better economic infrastructure and incentives for rural households to invest to increase productivity and production expansion. It also emerges from Table B.1 that *house ownership* and *commune credit per capita* are positive and statistically significant determinants of formal credit participation, as expected. There are many reasons for expecting these variables to be positive, in which limited information of potential borrowers and requirement of collateral assets play the central roles curbing the formal credit supply.

Dealing with the variables, which are not statistically significant in this specification, it is important to keep in mind that allowing demand and supply forces estimated simultaneously in a single equation may blur the picture of rural credit participation. It is likely that some variables drive credit demand and supply functions differently having a hidden impacts on participation. Take the distance variables as an example. As shown in Table B.1, living closed to formal bank branch will significantly increase propensity of formal borrowing. When distance increases,

the positive effect also disappears. We, however, are not able to tell whether this disappearance of significance comes from which sides of the market.

We now turn to the estimated probability of participation in the informal credit sector, which can also be found in Table B.1. Generally, there are fewer number of significant coefficients in both data sets. *Household size* positively and significantly increases informal borrowing, as expected. The coefficient for *commune credit per capita* shows a negative sign in 1997/98 data, which suggests that the formal sector expansion over the time between the two surveys has become a real and accessible source of credit for rural households. It is also a signal of the crowding out impact of the formal sector competition on informal lenders¹². The more availability of formal credit in a commune, the less likely households in that commune will borrow from informal sources. The more interesting results are on the *health variables*. Both regressions reveal that bad health condition of household members will push households to borrow more often from informal lenders, reassuring our hypothesis on informal credit demand in condition of negative income shock.

Until now, we have assumed that the participation decision to formal and informal credit sectors are independent and having no interaction. However, they may be related, especially when households have access to both credit sectors. To address this problem, we estimate a bivariate probit regression where errors from estimates of formal and informal credit sector participation are allowed to be correlated. Thus, similar to Model I specification, the sector participation equations are expressed as:

$$B_{ij}^s = 1 \text{ if } B_{ij}^{s*} = \beta_1 H_{ij}^s + \varepsilon_{ij}^s > 0, 0 \text{ otherwise}$$

where B_{ij}^s is the dummy variable of household participation in credit sector s , $s = (\text{formal}, \text{informal})$. All other variable specifications are as before, except the error terms ε_{ij}^s assumed to be distributed as a bivariate normal, i.i.d. across observations, with the correlation coefficient ρ to be estimated.

¹²See Jain (1999) for theoretical discussion of the crowding-out phenomenon in the rural credit market in developing countries.

The Wald test for the null hypothesis of $\rho = 0$, however, cannot be rejected ($p_{\chi^2}^{93} = 0.8871$, $p_{\chi^2}^{98} = 0.6583$) suggesting that the equations of credit sector participation could be independently estimated. Therefore, the results yielded from the bivariate model are almost the same to the results from the univariate model that we discussed before. For completeness, results of the bivariate probit regression are presented in Table B.2.

In summary, the results from Model I are fairly consistent over the two surveys, signs of coefficients mostly as expected. The differences in significant level of some coefficients between estimates from VLSS1992/93 and VLSS1997/98, e.g. *irrigated land holding size, health condition, farm, and house ownership*, imply the change of economic environment and the development of rural credit market. Moreover, the results suggest that a key underlying distinction between the formal and the informal credit sector participations is that the formal sector participation is particularly driven by observable variables, such as physical asset, e.g. *house ownership*, or human capital, e.g. *education of head, number of working adults in the household*. These variables reflect household demand for investment capital as well as its accessibility to the formal sector. The important roles of information variables like *credit per capita*, and *distance* are also highlighted. In contrast, the informal sector participation is significantly affected by *household size* and *health condition*.

Finally, we predict the probabilities of participation to credit sectors based on the estimated model. The results reveal a dramatic change of credit market over the five years gap between the two surveys. In 1993, only about 21% of households in rural area is predicted to borrow from the formal sector, implying that the remaining 79% is rationed out. This number is almost less than a half of the predicted 40% of the informal sector participation. However, in 1998, coinciding with the expansion of the formal sector, the predicted formal sector participation increases to 37%. This reflects a better accessibility to formal credit by 1997/98. The validity of these estimates, as well as the consistency of the Model I, however, is conditional on the validity of the assumption that credit participation is solely determined by a single credit supply function. We address this problem in the bivariate probit model below.

3.4 Bivariate Probit Model with Partial Observability

3.4.1 Econometric Specification

A possible reason for the popularity of this univariate probit model is its simplicity to implement and conformable results to theoretical predictions (see, for example, Iqbal (1981), Swain (2002)). However, the model suffers from substantial drawback due to its unrealistic assumptions on the single credit equation which incorporates both demand and supply forces. In practice, there will be a good proportion of non-borrowing households who have either no demand for or no supply of credit. There are several reasons. One will not borrow if one has been saving enough to self-finance the arise project. Conversely one does not borrow because he has no sufficiently profitable projects to invest or simply he is lacking of necessary skills to manage the potential investment and voluntarily withdraw himself from credit market. If a policy recommendation can be drawn from these observations, it may be that making credit available to poor households, as many countries did in 1970s by government credit subsidy, will not always work. Rather than that many poor households, especially those from the poorest segments, may be better off with government assistance in health, education and infrastructure (e.g. clean water, electricity, concrete road). These supports would help them to increase physical productivity and human capital, which empower them to new investment latter in time.

In order to relax the initial assumption and to address the problem of non-separable credit demand and supply, we employ the model of partial observability. This model is first introduced by Poirier (1980), and developed and applied in various empirical studies. Several possible bivariate probit models are available¹³ depending on level of observability of dependent variables in which Poirier's case is the least information one. In other studies (see e.g. Abowd (1982), Feinstein (1989), and

¹³ Meng & Schmidt (1985) provides a very good review on related literatures.

Heywood (1995)) at least one of the two dependent variable is observable and the model becomes a sequential decision model where the first step is known. For example, Abowd (1982) apply the partial observability model to examine job queues and union status which relax the assumption of other studies in the same topic that all workers who wish to work on union jobs are always able to find one¹⁴. Yet, union status is only observed if one has desire to join the union and be selected. Otherwise, we do not know a whether a non-union status person does not want to join the union or had been refused to join. The model that Abowd-Farber employed is in the same sense as the model introduced by Poirier, however there is a difference between the two. In Abowd-Farber study, more information has been observed which is the desire of non-union workers to join union. In other words, demand is observed for all samples and union status therefore is determined by selection function. In this case, one observes more than in Poirier's model but less than in the full observability case¹⁵.

In our study, observation of borrowing activities falls into the Poirier's type due to non-observability of both credit demand and supply. Yet we do not observe credit demand of households in the sample, hence it is not possible to separate whether a non-borrowing household is one that applied for credit but was rejected by lenders or simply having no credit demand. Following the structure specified in equations 3.1, 3.2, 3.3, and 3.4, credit participation status ($B = 1$) is only observed when there exist both demand and supply ($S = 1$ & $D = 1$). The last three combinations of S and D . ($S = 0$ and $D = 0$; or $S = 0$ and $D = 1$; or $S = 1$ and $D = 0$) are indistinguishable since we only learn that there is no credit activity (i.e. $B = 0$). According to Poirier (1980), the probability distribution of B is driven by a bivariate process representing binary choice of credit demander and supplier concerning the level of observability of

¹⁴See, for example, Lee (1978), Schmidt (1976)

¹⁵It is called censored probit model or partial partial observability model by

D and S

$$\Pr(B_{ij} = 1) = \Pr(D_{ij} = 1, S_{ij} = 1) \quad (3.6)$$

$$= \Pr(D_{ij}^* > 0, S_{ij}^* > 0) \quad (3.7)$$

$$= \Pr(H_{ij}^d \beta + u_{ij} > 0, H_{ij}^s \beta + v_{ij} > 0)$$

$$= \Pr(-H_{ij}^d \beta < u_{ij}, -H_{ij}^s \beta < v_{ij})$$

$$= 1 - F(-H_{ij}^d \beta, -H_{ij}^s \beta)$$

$$= F(H_{ij}^d \beta, H_{ij}^s \beta, \rho)$$

and

$$\Pr(B_{ij} = 0) = \Pr(D_{ij} = 0 \text{ or } S_{ij} = 0) \quad (3.8)$$

$$= 1 - \Pr(D_{ij} = 1 \ \& \ S_{ij} = 1) \quad (3.9)$$

$$= 1 - F(H_{ij}^d \beta, H_{ij}^s \beta, \rho) \quad (3.10)$$

We then obtain the log likelihood function¹⁶ of the sample as

$$\text{Log } L(\beta, \gamma; \rho) = \sum^n B_{ij} \ln(F(H_{ij}^d \beta, H_{ij}^s \beta; \rho)) + (1 - B_{ij}) \ln(1 - F(H_{ij}^d \beta, H_{ij}^s \beta; \rho)) \quad (3.11)$$

where ρ is the correlation between u_{ij} and v_{ij} ; $F(\cdot)$ denotes the bivariate standard normal distribution.¹⁷

The consequences of partial observability, as stated by Poirier himself, is that estimate of the model comes with cost i.e. the loss of efficiency. Poirier (1980, p.212) notes that "*estimators obtained from the model will be inefficient compared to those obtained in the case of fully observed choices*". This efficient lost is measured in a study by Meng & Schmidt (1985). Although there is no general conclusion on how

¹⁶The difference between full and partial observability likelihood function is in its information matrix where the former has four observable outcomes of S and D, and the latter has only two.

¹⁷The likelihood function for model with full observed information is originated from Maddala 1983, p. 123.

much efficiency has been foregone between full and partial observability, Meng & Schmidt do show that the loss does exist and there is evidence of efficiency improvement in moving from partial to censored and to full observability. The efficiency loss depends on the fraction data of the unobserved dependent variable in the dataset and the strength of identification especially near points of non-identification. In other word, sample split matters and the higher rate of observed "yes, yes" dependent variable the more efficiency we obtain. However, Meng & Schimidt concludes that as long as the model is identified we should not worry too much on how strong its identifying power is¹⁸.

3.4.2 Identification Strategy

While there is little we could do to improve the efficiency of model estimation due to limitation of available data, we can examine its identification. Under the general principle Rothenberg (1971), and as details in Poirier (1980), the partial observed bivariate probit models will be (locally) identified if and only if the information matrix corresponding to 3.11 is non-singular. In other word, for the model to be identified, we need at least one variable that is included in the explanatory variable sets of demand or supply equation (either H_{ij}^d or H_{ij}^s), but not in both. Indeed, finding valid identification is a crucial task in our study.

The identification variable sometimes simply stems from the method of sampling involved or from exogenous facts affecting one of the two equations. For example, Pitt and Khandker (1998) use the exogenous eligibility condition that only households own less than 0.5 acres of land are able to join any of the formal credit programmes in the area. This exogenous rule clearly defines the formal credit supply function, and hence identifies the model.

As already discussed in the previous section, most of variables capturing house-

¹⁸To find the effect of identification, Meng & Schimidt (1985) hold sample split and other features constant while increasing identification power. Generally, the cost of efficiency falls as identification power increases but the effect is only strong and predictable while near the point of nonidentification (p.80)

hold and commune characteristics will appear in both equations. The variables that are observable by econometrician will be reasonably expected to be observed by lenders, especially informal lenders. Therefore, it is exceptionally difficult to find "perfect" or "clean" instrumental variables, even in economic experiment. For instance, in the example by Pitt and Khandke cited above, in theory no household possessing larger than 0.5 acres of land should be found to borrow from formal lenders. But in reality, as pointed out by Morduch (1998, p.4), as many as 30 percent of formal borrowers actually own more than 0.5 acres. In another example, although principle interest rates of formal credit institutions could be a good instrument for predicting the demand for credit, the lack of variations in interest rates precludes such use.

One crucial argument we use in this dissertation to identify the demand for formal credit is that not all household characteristics can be observed by formal lenders due to limited information and lending policy. Therefore, the formal credit demand is primarily identified by a proxy variable measuring the health condition of the two most important persons in the household which is the number of illness days of household head and his/her spouse within four weeks before the time of the survey. This health proxy variable is assumed to reflect health condition of the family and hence the domestic labour supply. That is, the higher number of days of illness, the poorer health condition of family heads and the less labour input for production. As already pointed out, a great portion of formal loans is for investment (80 percent of all formal loans in 1998) and, in reality, formal loan applicants are often asked to submit their business plan explaining how to use the credit as a part of the application procedure (Pham & Izumida, 2002). In addition, due to its relatively large loan size, low interest rate with slow credit disbursement, the formal sector can be considered as the primary source for investment capital. Bad health problem and shortage of labour supply will consequently limit household's production expansion and therefore reduce household's demand for formal investment credit. However, this variable is not normally observable by institutions, hence it has no effect on supply decision and is excluded from the formal supply equation S . Because not-

healthy households are likely to voluntarily withdraw from the formal credit market, if there exists a negative relationship between the health variable and the formal credit participation, it is probably the result of little credit demand rather than of supply rationing.

More difficult, however, is the identification of the informal credit demand. Different from formal lenders, informal lenders, mostly friends and relatives, are able to closely observe household characteristics and local community environment. In other words, variables that influence credit demand will also be detected by informal lenders before making their own decision on supply. For instance, the bad health condition of household heads could increase demand for consumption credit, especially in urgent need of emergency cash. Friends and relatives are understandable to be the first option but they can also be well informed of the household situation and may not be willing to lend. As argued by Udry (1994), one should not expect to be able to find exclusion restrictions to identify the informal credit demand function. And we are not exempted. Therefore, to estimate determinants of informal sector participation, we have to solely rely on the identification from the informal supply equation which will be discussed below.

Identification of the formal sector supply is achieved by the variable measuring commune-level formal credit per capita, which equals total formal loan amount for all purposes lent to households in commune divided by commune's population. The use of this variable as instrument for formal credit supply is based on the extensive literature on credit rationing due to lending cost, adverse selection, and the importance of availability of borrower's information to formal lenders. The higher credit per capita a commune has, the more activities formal institutions undertake in the region. With more information and understanding about local environment, no doubt, formal institutions will be more confident to lend. Access to formal sector therefore will be higher for households residing in a commune that is better known by organised lenders than households from a commune where banks have little or no credit activity history.

A final issue which needs to be addressed is identification of the supply of informal loans. As widely discussed in literature on rural credit access, it is hard to define identification of informal supply without some relatively strong assumption on its availability according to the dynamic and the non-formality of lending conditions (see e.g. Aleem (1990), Bell (1997), Gine(2005), Conning (2001), and Kochar (1997)). For example, Kochar (1997) employs an indicator variable of "personal surety", a third party guarantee to secure the informal contract as the proxy for availability of informal credit supply. However, she also points out that the estimation basing on this variable is likely to yield biased results due to endogenous determination of informal borrowing. She then proposes a model to estimate probability of households having a security using joint distribution of the probability of being an informal borrower and of being the borrower providing some form of personal guarantee. This model, as she accepts, has a complicated structure without knowing the exact form of the joint distribution to be estimated. In our case, only about 20 percent of informal borrowers in 1997/98 needs some kind of guarantee for their loans¹⁹ which implies that the security variable as used by Kochar (1997) is not a promising identifier.

Originating from a widely observed facts that informal lenders are usually individuals living relatively close to borrowers (see e.g. Harper (1998)), normally in the same village or commune, we employ the variable of commune population as identification of availability of the informal credit. Indeed, as we have discussed in the earlier section that most of informal lenders in rural Vietnam are relatives and friends (about 80 percent of all informal loans in both surveys), higher population of a commune will increase number of potential informal lenders. This variable is assumed to be independent with household demand for the informal credit, hence serve to distinguish the supply from the demand equations of the informal credit market.

¹⁹In Kochar (1997) study, 52 percent of informal loans requires third party security

3.4.3 Estimation Results

The model of bivariate probit with partial observability (Model II) generalises Model I by relaxing the assumption of the singly determined equation of credit participation. The probability of credit sector participation is now jointly determined by household demand for credit and lender's decision on access. Estimates of the Model II are reported in Table B.3, B.4 and B.5.

We start by testing Model II against Model I to verify that allowing demand and supply equations to be estimated separately is necessary to strengthen the results. Because the models are non-nested, we cannot employ the conventional tests using the Likelihood ratio in selecting between the two. Instead, we apply the Akaike's information criterion (AIC) tests for the best-fit of each model estimation. The AIC critical values penalise the log-likelihood values for the addition of parameters, and thus select a model that fits well but has a minimum number of parameters²⁰. The results²¹ do support the selection of Model II over Model I.

In 1992/93, the coefficients for *household size*, *ratio of working adult in the household*, *irrigated land holding size*, and *health condition* are found to positively and significantly affect demand for formal credit, which is as expected and consistent with results from Model I. On the supply side, the coefficients for *household size* and *commune credit per capita*, the instrument for supply equation, are positive and statistically significant. The latter variable confirms the important role of local information and regional market familiarity for formal institutions in determining credit supply. However, the negative influence of *head's education* and *land holding size* on supply is somewhat surprising. This may indicate that formal institutions in 1992/93 do not consider household head's education as an important criteria for their decision on access. Also, the underdevelopment of legal regulation on private land title may hinder the possibility of using land as collateral asset. Banks therefore may consider land quality or agriculture productivity rather than household's

²⁰Discussion of using AIC and its derivations is provided in Koehler(1988). Stata command *icom* is employed to calculate the test.

²¹Reported in the results Tables B.3, B.4 and B.5.

land size as important factors determining the supply.

Turning to the VLSS1997/98 estimates, most of the parameters have the expected signs and significance level at 5% or better. On the demand equation, *household size*, *number of working adults in the household* and *working in farm* continue to highly associate with demand for formal credit, affirming the vital role of domestic labour input in agriculture production. The coefficients for *age of household head* is negative and significant as in the previous model. The signs of *schooling years of household head* and *its squared value* again imply an inverse u-shape effect on formal credit demand, which is shown to be the main mechanism to drive the result of formal sector participation in Model I.

The supply equation reveals that the sets of variables used by formal institutions in making their supply decisions differ considerably from those shaping the household's demand for credit. Thus, access to the formal sector is primarily determined by regional measures, i.e., *commune-level credit per capita*, and *distance to bank branch*. Indeed, households residing in a distant commune are likely to be neglected by formal credit institutions.

Estimates of the Model II for the informal sector participation are applied on both data sets, however the global convergence only obtained for VLSS1997/98 data. This is not uncommon in the literature, as noted in Abowd & Farber (1982), Meng & Schmidt (1985), and theoretically discussed in Maddala (1983). Instead of presenting results at local convergence²² for VLSS1992/93 data, which appears not to give us a consistent estimate of the true parameter value, we concentrate on discussing the 1997/98 results, presented in Table B.5.

Basing on estimates of access to formal sector, we use the predicted probability of having access to formal credit as an extra identifier for the informal demand. We expect that higher level of accessibility to the formal sector would reduce demand for informal loans.

As expected and relatively consistent with the results in Model I, the variables

²²Correlation between the residuals of demand and supply equation is either +1 or -1

associated with high consumption need, i.e. *household size* and *health condition*, positively affect informal credit demand. However, with supply equation separately estimated, we understand further that the positive effect of *number of adults in the household* on informal borrowing is driven by the better supply rather than by the higher demand. This reflects the importance of social connection for the informal sector accessibility. Also, the coefficient for availability of informal lenders, proxied by *commune population*, is found to be positive and significant in the supply equation. The coefficient on the predicted value of formal accessibility is negative and significant at 1% level, suggesting that households possibly prefer to borrow from formal lenders if they have access to both credit sectors.

Finally, we predict average probability of household having demand and access to each credit sectors based on the estimates from Model II. The probabilities of participation to the formal sector, i.e. $(\Pr(S^f = 1|D^f = 1))$, in 1992/93 and 1997/98, are 52.61% and 72.30%, respectively. These results are significantly higher than predicted values in the Model I (20.67% and 37.10%, respectively). Clearly, by estimating demand and supply equations independently, Model II provides more insides. First, the percentage of households having demand for formal credit ($D^f = 1$) is estimated at 39.63% in 1992/93 and 51.13% in 1997/98. The probabilities of having access to a formal source ($S^f = 1$) are, respectively, 76.00% and 86.02%. With regards to the informal credit sector participation, the Model II shows that almost 90% of rural households in 1997/98 has access to some source of informal credit (e.g. neighbours, friends, relatives or money lenders). Model I once again predicts a lower number at 27.91%.

The results imply that based on Model I solely will induce misleading conclusion on prediction of credit rationing. According to the Model I, nearly 80% of households in 1992/93 and 63% in 1997/98 having no access to formal credit. However, as the Model II reveals, the supply of formal credit in both years is adequately large, which significantly drives credit rationing rates down to 47% and 28%, respectively. The low rate of participation in 1992/93 appears to originated from the low demand rather than the short supply.

3.5 Conclusion

This chapter uses probability models to estimate the determinants factors of credit accessibility of Vietnamese rural households, with concentration on access to the formal sector. Unlike previous studies on Vietnam's rural credit markets (e.g., Pham et al (2002)), we aims to estimate the demand and the supply channels separately. The demand for formal credit is identified by variable capturing health status of household members. The formal supply is identified by commune-level formal credit per capita, which reflects operation of formal lenders in the studied communes. We bases our analysis on the bivariate probit model with partial observability. This model is employed to address the problem of limited information on the non-borrowing households. To the best of our knowledge, this is the first study of Vietnamese rural credit markets that allows supply of and demand for credit to be estimated distinctly.

Estimation results present evidence of the impacts of economic development on credit behaviour of rural households. The role of household characteristics determining level of access to the formal sector is differed over the years between the two surveys, implying the changes of economic environment. The distinction between the demand for formal and informal credit arises from the observable variables, such as physical asset, e.g. house ownership, or human capital, e.g. education of head, number of working adults in the household, reflecting household demand for investment capital. The supply of formal credit depends significantly on informational variables like *credit per capita*, and *distance*. Prediction of formal credit market participation is high and increases over years, implying the expansion of the formal sector. The demand for formal credit is also inflated, suggesting household's need for investment capital grows with economic development.

Chapter 4

The Determinants of Credit Amount Obtained by Rural Households in Vietnam

4.1 Introduction

Thus far, the focus has been on the determinant factors of overall probability of credit market participation of rural households. In what follows, the focus will be on the determinant factors of the amount of credit obtained. The VLSSs provide data on every separate loans that a household held in the reference years including loans with different lenders and different loans with the same lender. Hence the amount of loan that a household received from the same source of credit could be derived. Assuming that a lender's decisions on supply of different loans to the same household are homogenous, then household characteristics will be judged identically on every loan application and sum of loan amounts obtained from the same lender could be considered as one loan amount. Therefore, the unit of observation here is a loan amount obtained from each source of credit. Every household will only appear once in the dataset whether it does not borrow, or borrow only one loan from one source, or borrow more than one loan from more than one source of credit.

Depending on source of credit, a household may appear as non-formal borrower even though it does take credit from formal source, and vice versa. Until now, all types of lenders from the same sector are considered as homogenous and hence they will act in the same way in considering characteristics of credit applicants.

The reason for studying the determinants of loan amount obtained is the fact that not every household could be able to borrow as much as they want to or in other words there is a high possibility of quantitative credit rationing (Petrick, 2005). In the previous chapter, the analysis has been on credit participation of households which is referring to complete rationing where some households could not borrow even the smallest loan amount. Those models do not aim to and also are not able to explain what factor causing the differences in loan size that households had. The variation of average loan size by lenders can be found in table A.9 and A.10. For both reference years, loans from government banks and relatives are respectively larger than the average size of those from the formal and the informal sector. Also, the loan size has actually increased more than three times for the 5 years gap between the two surveys. These increases, however, are not likely equally distributed among borrowers. Hence, the following analysis contributes to build up a bigger picture of relationship between lenders and borrowers in rural credit market.

The chapter is structured as follows. The next section provides econometric framework for estimation, followed by the section on identification strategy. Results are discussed in the sequent section. Conclusion is the last.

4.2 Econometric Framework

The empirical model is as follows. Let L_{ij} be observed value of the dependent variable, the loan amount that household i in the commune j receives by the time of the survey, which is sum of all loans from the same source of credit. One will only observe that $L_{ij} > 0$ if household is a borrower and $L_{ij} = 0$ if not. The loan amount L_{ij} is determined by households and regional characteristics which can be

expressed as

$$L_{ij} = L_{ij}^* = H_{ij}\beta + \varepsilon_{ij} \text{ if } H_{ij}\beta + \varepsilon_{ij} > 0 \quad (4.1)$$

$$= 0 \text{ otherwise} \quad (4.2)$$

where L_{ij}^* is the latent dependent variable of loan amount, H_{ij} is the vector of household and regional characteristics, β the vector of coefficients to be estimated and ε_{ij} the residual, assumed to be normally and independently distributed.

The censoring feature of L_{ij} destroys the assumption of linear relationship between the loan amount obtained and household characteristics. Hence estimation of the model following the conventional linear original least square regression will result in bias. In economics, this type of censored model is called the model of limited dependent variable or the Tobit model, which first introduced by the pioneering work of Tobin (1958)¹. Following the literature, the model considered in equation (4.1) can be consistently estimated by maximizing its likelihood function².

However, it has been argued that the standard Tobit model faces limitations. As noted by Cragg (1971), among others (see, for example, Gronau (1973), Heckman (1976), Fin & Schmidt (1984), and Blundell & Meghir (1987)), the standard Tobit specification is quite restrictive because it imposes that both participation decision and decision concerning the size of the dependent variable depend on the same factors and the same mechanism. In other words, the Tobit model assumes that the probability of the observations being censored and the value of continuous observations on the dependent variable are governed by the same stochastic process. Thus,

¹The Tobit model has become popularly applied in various areas of empirical economics, especially recently due to the advance of computational capacity and the increased availability of microdata. An excellent literature review of theoretical development and empirical application of the Tobit model and its generations is provided in Amamiya (1984).

²The standard likelihood function for the Tobit model is: $L = \prod_0 [1 - \Phi(\hat{H}_{ij}\beta/\sigma)] \prod_1 \sigma^{-1} \phi[(L_{ij} - H_{ij}\beta)/\sigma]$ where \prod_0 and \prod_1 refer to product of the likelihood over zero and positive observations of L_{ij} , Φ and ϕ refer to the standard normal distribution and density function, respectively.

any variable that increases the probability of a non-zero value must also increase the conditional mean of the positive value. Therefore, in the case that the zero value of dependent variable is generated through a different process i.e. self-selection or selection on unobservable characteristics, the estimate of equation (4.1) may be potentially biased. For example, some households may be less likely to borrow due to their low demand, but once they decide to borrow, these households may borrow more than others.

By incorporating the participation process, the model is then rewritten with an additional selection equation, as

$$P_{ij}^* = H_{ij}^p \gamma + u_{ij} \quad (4.3)$$

and

$$\begin{aligned} L_{ij} &= L_{ij}^* = H_{ij}^L \beta + \varepsilon_{ij} \text{ if } P_{ij}^* > 0 \\ &= 0 \text{ if otherwise} \end{aligned} \quad (4.4)$$

where P_{ij}^* is the latent variable capturing the selection process, H^p and H^L the vectors of characteristics, γ and β the coefficients to be estimated, and u_{ij} and ε_{ij} the error terms for the equations (4.3) and (4.4), respectively. $(u_{ij}, \varepsilon_{ij})$ is assumed to be independent of H^p and H^L , and normally distributed. The correlation coefficient between u_{ij} and ε_{ij} is $\rho_{u\varepsilon}$.

Given the above structure, there are two distinguished cases regarding the assumption on distribution of error terms. First, if the errors are independent i.e. $\rho_{u\varepsilon} = 0$ then the selection process do not have effect on the outcome equation, or in other words, there is no sample selection problem. In this case, the model becomes a *two-part model* which was presented in detail by Cragg (1971) and applied by, for example, Duan et al (1983) in a study of individual annual medical expenses. This

model can be estimated in two-step procedure, using probit regression to obtain consistent estimates of γ for the participation equation and then carrying out OLS regression for the outcome on the selected sample to get consistent estimates for β (Wooldridge, 2002, p.563).

However, any correlation between the two errors, i.e. $\rho_{u\varepsilon} \neq 0$, will require an alternative estimation procedure to address potential bias. In other words, we need to take account of the sample selection problem and the different mechanism of participation and outcome³. This selection model can be estimated straightforward using MLE method, as long as $(u_{ij}, \varepsilon_{ij})$ are bivariate normal distributed⁴. Alternatively, Heckman (1979) suggests a two-step procedure, which is less efficient compared to MLE but consistent and simple to implement. In the first step, the probit selection equation is estimated by MLE and sample selection correction term is constructed. For each observation in the selected sample, we compute $\hat{\lambda}_{ij} = \frac{\phi(H_{ij}^p \hat{\gamma})}{\Phi(H_{ij}^p \hat{\gamma})}$ (the inverse Mills ratio). In the second step, the outcome equation (4.4) is then estimated by OLS, including the correction term $\hat{\lambda}_{ij}$ as an additional regressor, as below

$$E(L_{ij}|H_{ij}^L, P_{ij}^* > 0) = H_{ij}^L \beta + \sigma_u \rho_{u\varepsilon} \hat{\lambda}_{ij} \quad (4.5)$$

4.3 Model Identification

Selection in the standard Tobit model specified in equation (4.1) does not require any condition in order to be identified. In theory, the bivariate selection model is also identified without restrictions on the regressors. Even if exactly the same regressors appear in the equation (4.3) and (4.4), i.e. $H^p = H^L$, the model still formally identified due to the non-linearity of the inverse Mills ratio. However, this

³(Amamiya 1984) calls the model a type 2 Tobit model, (Wooldridge 2002) calls it a probit selection model. Here, in our particular situation, we prefer to call it the selection model.

⁴The likelihood function of the model is $L = \prod_0 P(P_{ij}^* \leq 0) \prod_1 f(L_{ij}|H_{ij}, P_{ij}^* > 0) P(P_{ij}^* > 0)$. The where assumptions on the error terms are $(u_{ij}, \varepsilon_{ij}) \sim N(0, 0, \sigma_u^2, \sigma_\varepsilon^2, \rho_{u\varepsilon})$ and $\sigma_u^2 = 1$. Derivation and exact form of the likelihood function provided in (Amemiya 1985, p.386)

case ($H^p = H^L$) often leads to multicollinearity problem between the inverse Mills ratio term $\hat{\lambda}$ and the regressors H^L included in outcome equation, as discussed in many research (see, for example, Nawata & Nagase (1996), Leung & Yu (1996)). Therefore, as a rule of thumb in empirical study, estimation of the selection model requires at least one regressor in the participation equation (P^*) be excluded from the outcome equation (L^*). In other words, H^L should be a subset of H^p . Hence the centre point in estimating the selection model is to identify the participation and outcome equations.

Recalling discussion from the previous Chapter 3, a household is observed to participate in credit market if and only if he has demand for loan and that demand is satisfied (either fully or partly) by lender's supply. In other words, $P^* > 0$ if and only if ($D^* > 0$ & $S^* > 0$). However, as also noted, demand and supply channels are not observed in the data. We therefore have to estimate these equations using instrumental variables which are argued to affect either demand or supply but not both functions. That is, in addition to household and commune characteristics that generally included, *health condition of household members* is used to identify the demand, while *commune credit per capita*, and *commune population* are instruments for the supply⁵.

The picture is, however, more blurred in estimating the selection model, to find determinants of credit amount. As discussed, the loan value is only observed positive if household pass the participation or selection hurdle and the main problem is this hurdle is not clearly defined. One cannot say whether a household's participation to credit market solely depend on its demand (and that means there is no credit rationing), or on its supply (i.e. all household demand for credit), or both. Moreover, it is even harder to find a variable influencing the participation but not the loan size, using what the data provides. Indeed, identification of participation equation has always been at the centre of debates in empirical research using this model.

In order to take account of the complications and to make our discussion more

⁵Commune population is used for informal sector supply only.

consistent, we further assume that:

- (i) Household borrows if and only if he has demand for credit, i.e., participation or selection equation always incorporates specifications affecting the demand
- (ii) The loan amount that household received are determined by lender decision on credit supply, i.e., outcome equation always includes variables influencing the supply

With these assumptions imposed and the model specifications as in equations (4.3) and (4.4), there are four possible specifications as follows

$$(1) P_{ij}^* = D_{ij}^*; L_{ij}^* = S_{ij}^*$$

$$(2) P_{ij}^* = D_{ij}^*; L_{ij}^* = (D_{ij}^* \& S_{ij}^*)$$

$$(3) P_{ij}^* = (D_{ij}^* \& S_{ij}^*); L_{ij}^* = S_{ij}^*$$

$$(4) P^* = (D_{ij}^* \& S_{ij}^*); L_{ij}^* = (D_{ij}^* \& S_{ij}^*)$$

Given the exclusion restrictions discussed above, only specification (3) satisfies. That is, a household must have demand for credit and the demand is satisfied by a supply source then the household can become a borrower, i.e. $P_{ij}^* = (D_{ij}^* \& S_{ij}^*)$. In other words, household's participation decision depends on both demand and supply channels. However, once a household borrows, loan amount that household can obtain is solely determined by supply decision from lender, i.e. $L_{ij}^* = S_{ij}^*$.

4.4 Results

General specification for variables included in our models is similar to what already presented in Section 3.3.2. In short, the following covariates are included in the vector H_{ij} : household characteristics (i.e. *age, gender, marital status, education level of household head, health condition of household members, household size,*

farm, irrigated land holding size, house ownership) and commune characteristics (i.e. *commune population, formal credit per capita, and distance from commune to the nearest bank branch*). Identification of demand for and supply of credit has also been discussed in the previous section and as detail in Section 3.4.2. Further discussion of variable construction are given in the Appendix E. Estimations are implemented on VLSS1992/93 and VLSS1997/98, for both formal and informal sectors.

The dependent variable L is the borrowing amount obtained by a household. It is measured in thousand VND, and calculated for formal and informal source separately. The empirical estimation needs to take account of two complications: (1) Borrowing amounts are zero for almost 50 percent of both survey samples,⁶ and (2) Positive borrowing amount are highly skewed, especially for the formal loans⁷. In addition, estimators of the limited dependent variable models has been showed not robust against specification errors, i.e., in the existence of heteroskedasticity or non-normal distribution of the residuals (see, e.g., Maddala and Nelson (1975), Arabmazaar and Schmidt (1981)). One common practice to solve these problems is to make logarithmic transformation of the interest variables (Maddala, 1983). However, without careful consideration, taking log-transformation of the loan amount variable will destroy its censoring characteristics because $\ln(L = 0)$ is not defined. To get over this problem, following Fische et al (1979) and others, we assume household's borrowing amount to be one thousand VND for those with no borrowing in the sample⁸. Thus, the log-transformation of L will be $\ln(1 + L)$ rather than $\ln(L)$ maintaining the censoring point at zero. Distribution of loan size in level and of loan size after taking log-transform is presented in Figure C.1. Similarly, as we conjecture that the marginal impact on loan size to decrease with these variables, we also take log-transformations of (one plus) these variables are specified, including *irrigated land holding size, commune population, and commune-level formal credit per capita*.

⁶That is non-borrowing households are observed having zero loan amount

⁷In the VLSS93, mean of formal loan amount is 1,029, median 500, skewness 9.35 and kurtosis 136.17. In the VLSS98, the numbers respectively are 4,143 (mean); 2000 (median), 8.74 (skewness), 132.25 (kurtosis)

⁸1,000 VND = 1 unit measuring loan size

We start with a simple OLS regression of $\ln(1 + L)$ on characteristics of household and commune, to address the linear relationship, if any. Regression results are presented in Table C.1 with robust standard error, adjusted for clustering to control for potential correlation of the error terms corresponding to households living in the same commune. In all estimates, the coefficients on *household size*, *working adult rate*, *log commune population*, and *log formal credit per capita* are positive and significant. The coefficient on *health condition* suggests positive effects on loan size from both sectors in VLSS1992/93, while it is only significant for informal borrowing in VLSS1997/98. The coefficients for *irrigated land holding* are positive in both years. Estimated coefficients of *education* and *distance* are also as expected. The OLS estimate, however, is potentially biased because the sample it selects (households with positive loan size) is not representative of the population, as already discussed.

Table C.2 presents results from Tobit model for pooled VLSS1992/93 and VLSS1997/98 data, as specified in equation (4.1), to take account of the censored value of loan size. Nevertheless, estimated results do not change much from the OLS estimations, on both sign of coefficients and significant level. In VLSS1992/93 regression, the coefficients on *household size* and *working adult rate* remain significant and positive for the formal sector, those on *age and gender of household head* become statistically insignificant. The positive effect of bad health status may sound suspicious but it is consistent with the stylised facts that consumption loans account for large proportion of formal loans in 1993. This kind of direct competition between formal and informal lenders on consumption credit is confirmed by the significantly negative impact of *commune formal credit per capita* on informal loan size.

Turning to the VLSS1997/98 results, again the differences between Tobit and OLS estimates are not many, even though the Tobit model does show changes in significant level for some variables. Similar to the OLS, Tobit results pick up the impact of *age and education of head* on the formal loan size. Households having its head being 44 years of age, or spending 7 years in school are found to borrow the largest loan amount. The signs and significant levels of the coefficients on *house*

ownership, land holding size, commune population and *credit per capita* remain the same to the OLS results, reassuring the importance of collateralisable assets and commune popularity in determining household's formal borrowing. The positive and significant coefficient on DIST1 variable implies that households residing close to bank branch would borrow larger loans. Estimated results for the informal sector are as expected. *Bad health condition of household members* and *large commune population* are positively associated with informal loan size. The negative and significant of the coefficient on *credit per capita* suggest the competition between formal and informal lenders.

Nevertheless, the Tobit estimates are also suffered from some drawbacks, as discussed in Section 4.2. In what follows, we present results from estimates of the selection models specified in equation (4.3) and (4.4), which incorporate a participation process to take account of the selection bias that may arise from the treatment of zero value of the outcome. There are two cases regarding the relationship between selection and outcome processes, measuring in correlation of the estimate errors. As already noted, if the correlation coefficient $\rho_{u\varepsilon} = 0$, the system can be estimated by a probit regression for participation and OLS regression for the loan size. The latter results are exactly the same to what we obtained in Table C.1. Thus, it is not necessary to repeat that result here.

In the case that $\rho_{u\varepsilon} \neq 0$ or in other words, participation process has influence on the outcome, a Heckman two-step estimation is applied. Results are in Table C.3. We however only estimates the model for the formal sector because of our inability to find defensible identification for the informal sector participation process. The participation equation estimates are obtained from probit regression of discrete borrowing variable ($CRF = 1$ if borrow) on regressors of household characteristics. In 1993, the results show that participation to the formal sector is significantly influenced by *household size, land holding size, and commune population*. However, turning to the credit amount, household size seems not to be a significant factor. Instead, number of working adult in the household significantly and positively increase the amount. The interesting result is the negative sign of the coefficient on

farm which indicates that peasant households borrow less than average. That is, poor rural households who seem to borrow most often are not those who borrow the largest loan amount.

Generally, the results in 1997/98 are consistent with the estimates using 1992/93 data but with more significant coefficients. The coefficients on *farm* remains statistically insignificant. The *DIST1* significantly positively affect the participation but not the loan size. This result is different from what we obtain from the Tobit model, suggesting selection may cause noise in the Tobit results.

Further more, it is important to note that the coefficients on the inverse Mills ratio are statistically significant at 10% level. Though the Heckman second-step estimates of the outcome equation with inclusion of the inverse Mills ratio are still consistent, the 10% significancy implies that selection bias is relatively weak. In addition, in the VLSS1997/98 regression the identification variable of *health condition* is not statistically significant within conventional level. The estimation of the Heckman model therefore is identified solely by the non-linearity of the inverse Mills ratio. As pointed out by Sartori (2003), the insignificant of identification variable would make the model results somehow inconclusive.

The Heckman selection model can also be estimated using a one-stage procedure using maximum likelihood method⁹ which is less consistent but more efficient compared to the two-stage procedure. However, the regression cannot converge using the VLSS1997/98 data due to the weak identification as mentioned above. Hence, we are not going further with the Heckman one-stage procedure.

To check the robustness of our results, we employ a double hurdle model to estimate the equation system specified in (4.3) and (4.4). Instead of observing loan amount given that the participation condition $P_{ij}^* > 0$ satisfied, the double hurdle model requires further that L_{ij}^* must also be greater than zero (i.e. $L_{ij}^* > 0$). Thus the probability of observation of L_{ij} is driven by a bivariate process rather than only

⁹The likelihood function of the model is $L = \prod_0 P(P_{ij}^* \leq 0) \prod_1 f(L_{ij}|H_{ij}, P_{ij}^* > 0) P(P_{ij}^* > 0)$. The where assumptions on the error terms are $(u_{ij}, \epsilon_{ij}) \sim N(0, 0, \sigma_u^2, \sigma_\epsilon^2, \rho_{u\epsilon})$ and $\sigma_u^2 = 1$.

by the selection process as in the Heckman model (Fennema 2005, p.109)¹⁰. Results from the double hurdle estimates are presented in Table C.4¹¹. We only run the regression the formal sector data as the regression on informal sector cannot converge¹². In both years, the results confirm the positive impacts of *education*, *number of working adult*, *commune population*, and *commune credit per capita* on borrowing amount. The coefficient on *farm* are negative and significant in 1992/93. However, the Wald test for the independent errors between selection and outcome equations cannot be rejected ($p = 0.1089$). In 1998, the regression reports highly significant coefficients on *commune population*, and *commune credit per capita*. The *distance dummy* is again affecting participation but not the loan size. The coefficients on *age of head*, *education*, and *house ownership* are significant with expected signs. Tests for hypothesis of independent errors are rejected ($p = 0.0000$).

The VLSS panel data is also available, as described in Section 2.2. However, we will not employ the panel data analysis for several reasons. Firstly, the time span between the two surveys is relatively long (five years). During this period, the structure of the Vietnamese economy had changed quickly and so did the rural credit market. Assessing determinants of household's loan size would become a challenge to incorporate this development of economic infrastructure. Secondly, our main aim is to understand how borrowing and lending behaviours had change from 1992/93 to 1997/98. To satisfy this objective a comparison between results of two cross-sectional data covering two studied period will be the best strategy. Because of these given points, we will only employ panel data model in the next chapter to evaluate impacts of credit on household consumption.

¹⁰The likelihood for the double hurdle model, as described in (Jones 1992), is $L = \prod_0 [1 - \Phi(\gamma H_{ij}^p, \beta H_{ij}/\sigma_\epsilon, \rho_{u\epsilon})] \prod_+ \left\{ \Phi\left(\frac{(\gamma H_{ij}^p + \frac{\rho_{u\epsilon}}{\sigma_\epsilon}(L - \beta H_{ij}))}{\sqrt{1 - \rho_{u\epsilon}}}\right) \frac{1}{\rho_{u\epsilon}} \phi((L - \beta H_{ij})/\sigma_\epsilon) \right\}$. Where Φ is the density

of the bivariate normal distribution; $(u_{ij}, \epsilon_{ij}) \sim N(0, 0, \sigma_u^2, \sigma_\epsilon^2, \rho_{u\epsilon})$ and $\sigma_u^2 = 1$.

¹¹Estimation is done by the *dhurdle* command in Stata, which is developed by (Fennema 2005)

¹²The problem of inability to get programme converged is not uncommon in maximum likelihood estimation, especially in the case of the double hurdle model. As noted in Jones (1992) and Camaron et. al. (2005, p. 551) this convergence difficulty has been a great limitation for the double hurdle model and it continues to discourage researchers from applying this model.

4.5 Conclusion

This chapter analyse the determinants of credit amount obtained by rural households. Tobit model is used to account for the censoring characteristics of borrowing amount, where its value from non-borrowing household observation is zero. However, it is shown that Tobit model is not appropriate model due to its limitation in explaining the selection process. Borrowing household first need to have access to credit source before the loan amount can be decided.

Heckman model is employed to address the selection bias. The results show that fixed assets variables such as house, land holding, or human capital variables such as education, and number of working adult in the household are among the most important factors affecting household's credit participation. However, the credit amount seems to be driven by commune level formal credit per capita, and population, rather than any other variables.

The statistics of the inverse Mill-ratio, which is used to correct for the selection bias in the Heckman model, is significant at 10% level, suggesting the selection may not be consistent. To strengthen the results, we employ double-hurdle model which allow us to estimate the selection and outcome process in a bivariate procedure. Results from double-hurdle confirms the importance of correcting for selection bias. House, land and commune credit per capita continues to have positive and significant impacts on credit amount. This suggests that wealthier households may enjoy both of better access and higher loan amounts from the formal sector.

Chapter 5

The Impact of Formal Credit in Rural Vietnam

5.1 Introduction

Lack of access to productive capital is generally seen as one cause of poverty in rural areas of developing countries. Reasons for this shortage of financial services are many. On one hand, commercial banks have little incentive to expand their operations in poor areas because of high cost and tremendous risk, as analysed in Hoff & Stiglitz (1990). On the other hand, the poor often find themselves excluded from the formal credit market because they cannot either meet the eligible conditions or be able to put up acceptable collateral as usually required by formal institutions. Even though poor people may still be able to obtain credit from alternative informal sources, such as moneylenders or relatives, such lenders with limited resources often charge extremely high interest rates on short term and small size loans making them unreliable providers of investment capital.

Since the 1970s, the poor in developing world have started to gain better access to formal credit thanks to the development of microfinance programmes, inspired by the success of Grameen Bank in Bangladesh. The number of people received credit from microfinance institutions worldwide rose dramatically from 13 million

by 1997 to 113 millions by 2006 (Daley-Harris 2006). Microfinance has become an important instrument to alleviate poverty in political agenda, capturing much attention from policy makers. In the academic world, economists have also shown increased interest in microfinance. A number of works have studied to explain how and why the microfinance works, particularly to understand how it solves the problem of asymmetric information that discourages commercial banks from lending to the poor. Other issues attracting lots of economic research are mechanism of joint-liability group lending, financial sustainability and outreach, and above all, the evaluation of impact of credit programmes on poverty. That is, regardless of how microcredit programmes work, how many clients these programmes serve, the crucial question that economists and policy makers ultimately want to answer is how credit accessibility affects the life of borrowers.

In the theoretical economic literature, the role of capital in production and economic development is well established, but it is harder to estimate the respective contribution of capital especially at household or individual level. As in rural agricultural production, capital would serve to finance farming activities, to insure harvest against production risk, and to help people smooth consumption in the face of income shock. In practice, however, simply making cheap credit available to the poor has been proved to not always work by the evidence of the failure of many subsidised credit programmes in developing countries, for example in India and Philippines, from 1950s through out the beginning of 1970s. The new wave of microfinance development urges researchers to produce careful evaluation of credit programmes to find the impacts of credit accessibility, or in other words, to answer how borrowers would have done without taking loans.

Unlike natural science where effects could mostly be measured through experimentation in a strictly controlled environment to judge the causal direction of the relationship being studied and minimise potential biases, it is usually hard and costly to set up experiments in economic research. Indeed, most empirical evaluations of credit programme have been using data collected from designed surveys (non-experimental) and employing econometric techniques to control for unwanted

effects from other factors. That is, an impact evaluation must estimate the counterfactual to find how the results would change had the programme never taken place.

However, econometric estimations based on non-experimental data may suffer from potential biases. In the situation of measuring credit programme impacts, generally, the biases will arise from two main sources: the non-randomness of credit programme placements and the endogeneity of programme participation or self-selection into credit programme by households¹. In the former problem, the availability of formal credit is unlikely to be randomly allocated across geographical regions but endogenously determined by institutions. In other words, formal institutions carefully select areas of operation and that decision on where/whether to provide financial services causes potential biases for impact evaluation. Clearly, the selected commune is not necessarily a "better" commune compared to others; it may be poorer as result of credit targeted. By comparing the incidence of poverty in "credit-available" and "credit-unavailable" areas, which are different in characteristics, researchers may mistakenly report biased results.

Additionally, households who have taken credit may be systematically different from others who have not. When the characteristics that cause these differences are observable (e.g. age, education, assets) and are taken into account, the impact evaluation should be less problematic. Often, though, the differences arise from the so-called self-selection process, driven by unobserved traits of borrowers such as entrepreneurial skills, hardworking or better focus. Moreover, formal lenders are unlikely provide credit at random but usually select clients based on screening applicant's characteristics. In many cases, borrowing eligibility conditions are even predetermined. The nonrandom credit could (self-)selectively allocate to either richer or poorer borrowers. A simple evaluation of credit programme without calculating these attributions, will be very likely biased. For example, measuring impacts of credit by comparing the increase of consumption between the poorer borrowing households who received credit through a poverty reduction programme

¹Especially in the case of group lending

with the not-that-poor non-borrowing households who were not eligible to join the programme, one may overestimate the programme effects².

The econometric literature suggests several methods to address the selection bias problems, including matching, difference-in-difference and instrumental variables. The general idea of these methods is to incorporate the selection process in statistical models to arrive at an unbiased estimate i.e. matching method using exogenous eligible condition of lender to separate borrowers from non-borrowers for comparison; instrumental variable method using exogenous instrument to separate borrowing activities from household outcomes. As pointed out in Cameron (2005), the validity of those models depends on how well the models are specified. Unfortunately, the majority of existing empirical studies on credit programme so far are of descriptive nature or fail to adequately address the selection bias in the estimate model³. The collection of few serious empirical studies, however, presents mixed results. Pitt and Khandker (1998), by using a quasi-experiment data of the Grameen Bank in Bangladesh to construct instrumental variables for individual credit participation and applying village fixed effects estimation to control for unobserved village-level heterogeneity, find that joining credit programme increases annual consumption of borrowers, especially of female borrowers. Khandker (2005) employs the analysis of panel data from Bangladesh to eliminate the differences in characteristics of credit programme location and borrower's unobservable attributes that do not change over time, reports the positive results. Also using data from Bangladesh to measure total and noncredit impacts of credit programmes on farm productivity, McKernan (2002), treating credit activities as endogenous and correcting for selection bias, finds evidence of large positive effects.

Nevertheless, there are other studies reporting the opposite results. For example, Coleman (1999) and (2002) take the advantage of data on prospective borrowers in Northeast Thailand, where households were self-selected to join the village bank

²As in e.g. Hossain (1988) who incorrectly presents evidence of the positive impact of the Grameen Bank without control for selection biases.

³See Morduch & Haley (2002) for a comprehensive review of recent literature on impact assessment.

programme but some of them were forced to wait one year before getting their first loans. This provides an unique way to address the selection biases. Coleman finds the programme loans having little impact on household consumption. Morduch (1998) uses the same dataset in Pitt and Khandker (1998) but estimates the impact by an alternative method, the difference-in-difference estimation with careful consideration of the de facto member selection into credit programme, finds no evidence of positive effects.

Thus, impact assessment exercise is surprisingly difficult to implement cleanly. At the heart of most empirical argument is the problem of identification, that is to separate out the causality relationship of credit and variables of interest outcomes. Aghion (2005) suggests the use of interest rates, lender characteristics or distance to/from the closest main city as instrument variables for loan impact but each has limitations. The interest rate hardly varies within a lender and also among lenders⁴. This is a key constraint for the interest rate playing its role because variation is necessary for identification. So do characteristics of rural lenders. The variable of distance is an option that we have adopted and will continue to be used in this chapter to identify credit participation.

Following the literature on impact evaluation discussed above and the results presented in the previous chapters, we consider self-selection as the main source of potential bias that we need to address in our estimation models. Unlike other microfinance studies where credit programmes may not operate in all studied villages (and hence non-randomness of programme placement requires attention), programme placement across communes in Vietnam is not a problem using VLSS data set. According to results in Chapter 3, credit institutions are found to almost uniformly distributed among communes in Vietnam⁵. However, it is possible that lending institutions do not homogeneously operate among areas. Households from some particular communes may find easier to obtain credit compared to households

⁴In Vietnam, study by McCarty (2001) shows that as many as 70 out of 78 formal institutions included in the survey set interest rates with reference to the government banks, such as VBARD.

⁵Formal credit institutions are reported to operate in all 135 communes in VLSS1997/98 and in 115 out of 120 communes in VLSS1992/93

living elsewhere due to targeted credit programmes or being close to a formal institution's branch office. This problem will be addressed by employing commune-level instrumental variables such as distance from commune committee to the nearest bank branch or commune credit per capita. As analysed above, these variables will capture variation of credit among households residing in different communes.

Similar to the previous chapters, this chapter is going to use data from the VLSS1992/93 and the VLSS1997/98. One development is a panel data set will be created based on these two surveys. We estimate impacts using both cross-sectional and panel data. In all regressions, the dependent variable (household's outcome) is real consumption per capita which is adjusted to commune price level to make it comparable between households from different communes. The credit variable which we would like to measure the impact of is household's formal credit per household member⁶. That is only impact of borrowing from the formal sector will be studied.

The chapter is organised as follows. Section 5.2 first sets up the econometric model to estimate the impacts using cross-sectional data. Instrumental variables are employed to overcome potential biases of OLS estimations. Results are presented and discussed. Section 5.3 explains how longitudinal data can solve problem of simultaneity biases and describes estimation framework followed by results. Section 5.4 is conclusion.

5.2 Impact Assessments using Cross-Sectional Data

5.2.1 Econometric Specification

Consider the following model linking credit and household outcome

$$Y_{ij} = \beta_0 + \beta_1 L_{ij} + \beta_2 H_{ij}^Y + \mu_j^Y + u_{ij} \quad (5.1)$$

⁶See Appendix A for more details

where Y_{ij} is the real per capita expenditure⁷ of household i in commune j ; L_{ij} the household per capita borrowing amount from the formal sector and β_1 the coefficient to be estimated which measuring the impact of borrowing; μ_j the unmeasured commune j fixed effects (time invariant) on commune credit availability and household outcomes; H_{ij} the vector of household and commune characteristics; u_{ij} the error term representing unmeasured household and commune characteristics that affect the outcomes.

The impact of credit on household consumption β_1 can be obtained by estimating equation 5.1 with ordinary least squares. This result will consistently estimate the causal effect of credit only if credit obtained by household is randomly assigned in the population. However, as discussed previously, credit has been long considered endogenous that associates with self-selected participation and thus \hat{B}_1^{OLS} cannot be interpreted as the coefficient capturing impact of credit.

A standard way to approach problems of endogeneity and simultaneity bias is to use instrumental variables that explain the credit amount obtained by household, but that have no direct influence on household consumption. Then the estimated credit variable basing on instrumental variables can be used to correct for the selection bias in assessing impact of credit in the outcome equation. Thus, the borrowing amount by the i^{th} household in the j^{th} commune can be expressed as

$$L_{ij} = \alpha_0 + \alpha_1 Z_{ij} + \alpha_2 H_{ij}^L + \mu_j^L + v_{ij} \quad (5.2)$$

where Z_{ij} is the vector of instrumental variables that will not appear in the outcome equation; H_{ij}^L the vector of commune and household characteristics determining credit amount; μ_j the unmeasured commune j fixed effects influencing household's credit activities; v_{ij} the normally distributed error terms, satisfying $E(v_{ij}|H_{ij}^L, Z_{ij}) = 0$. All other variable specifications are as before.

The idea of using instrumental variable approach may be simple but economic

⁷Henceforth called in short as household consumption/expenditure

research shows that finding convincing instrumental variables has been frustrating. Generally, a valid instrumental variable must satisfy two requirements: (i) it must be correlated with the endogenous variable (i.e., $\text{corr}(Z_{ij}, L_{ij}) \neq 0$) and (ii) it must have no direct association with the error term (i.e., $\text{corr}(Z_{ij}, u_{ij}) = 0$). The former requirement of instrumental variable's relevance has attracted much attention in the literature in recent years, largely motivated by the Bound, Jaeger and Baker (1995) paper. Staiger and Stock (1997) provide a complete analysis of this issue highlighting the problem of weak correlation between instruments and the endogenous variables. The key result is that using weak instruments can still persist substantial bias even in large finite samples. F-statistics for the test of joint null hypothesis that the coefficients on excluded instruments in the first stage equal zero has emerged as an indicator to avoid the weak instrument problem⁸.

The latter requirement of orthogonality of a valid instrument is harder to test. When the equation is exactly identified, i.e., the number of endogenous variables equal the number of instruments, then test is not possible. Only when the equation is over-identified, i.e., the number of instruments is more than the number of endogenous variables, an over-identifying restriction test can be applied to inspect whether the suspect instrument is valid, i.e., being independent with the error term of the outcome equation.

Finding at least one valid instrument for the loan amount is indeed a crucial task in our study. Following previous discussion on determinant of credit participation (see Chapters 3 and 4, and particularly section 3.4.2), there are two variables that we could consider to use as instruments for household borrowing. The first is the *distance from commune to the nearest formal bank branch* and the second is the *formal credit per capita in each commune*. It has been shown in Chapter 3 and Chapter 4 that both of these variables work well to explain the supply of formal credit while having little impact on household's credit demand. In Chapter 4 we find that *the distance variable* significantly reduce credit amount that a household would

⁸Also mentioned in Staiger and Stock (1997) paper, among others such as Baum, Schaffer and Stillman (2003), Stock and Yogo (2005), there is a rule of thumb that as long as the first stage F-statistics >10 , the weak instrument problem is substantially avoided.

receive while *commune credit per capita* would positively correlated with loan size. Similar argument can be imposed here. We argue that *the distance* and *the commune per capita credit* variables have no direct relationship with household consumption. They affect household consumption only through the channel of credit.

5.2.2 Results

The OLS regression results are presented in Table D.1. Commune fixed-effects are applied to capture the commune's time-invariant characteristics that may affect the commune's credit availability and household outcomes. Standard errors are robust and adjusted for clustering on commune. Regressions are run on both VLSS1992/93 and VLSS1997/98 cross-sectional data to estimate impacts of borrowing from the formal sector. The dependent variable (Y_{ij}) is log of household real per capita expenditure which is adjusted by commune price level. The credit variable is the log of (one plus) household formal borrowing amount per capita⁹. As before, other household and commune characteristics are included as control variables.

In 1992/93, the coefficients of interest are small and positive in both OLS and OLS-FE regressions, but only significant in the latter case. Other regressors including *age of household head*, *marital status*, *number of working adult in household*, *education level*, *log of commune population*, and *log of commune credit per capita* are all found to have positive and significant effects on household consumption. Similarly, regressions (3 & 4) of 1997/98 data find these household characteristics statistically improve consumption. Coefficients on household size are significant and negative in all regressions, which is, however, not a surprise. The negative sign of coefficients on *farm* and *agriculture land holding size* suggests self-consumption. i.e., the outcome variable of consumption does not incorporate household's self-consumption value. Thus, being a farmer with large land holding will significantly reduce household's expenditure recorded in the data due to self-supply of both food and non-foodstuffs.

⁹"One plus" is an ad hoc but common technique used to avoid problem of undefined $\log(0)$. The zero is the borrowing amount of non-borrowing households.

While general household characteristics behave as expected and consistent with different specifications, the coefficients reflecting impacts of credit are small (smaller than 0.01) and insignificant (excepted in regression (2)). Based on the OLS FE results only, one may conclude that borrowing does not have any effects on borrower's consumption. However, we have discussed previously that using simple regressions of OLS FE without careful consideration for potential biases arise from non-random selection of borrowers will generate misleading results. For instance, if borrowing decision and borrowing amount are not randomly provided among household sample (which is very much unlikely), then borrowing becomes endogenously determined and OLS estimates are inconsistent. The tests for hypothesis that credit variable can be treated as exogenous have been rejected at 1% level in all regressions (see Table D.2).

Instrumental variables are employed to solve problem of endogeneity. The log (one plus) of formal borrowing amount will be estimated as in equation 5.2 using three instruments LNCAPI; DIST1; and DIST2 for 1997/98, and one instrument LNCAPI for 1992/93. The predicted value of credit is then used to estimate the outcome equation 5.1. All instruments are excluded from outcome equation. Standard errors are again robust and adjust for clustering within commune. Results are shown in Table D.2.

In Table D.2, regression (1) presents results of IV estimation using 1992/93 data. The coefficient on credit becomes large (0.142) and statistically significant at 1% level. High value of the first-stage F-statistics (50.44) confirms that LNCAPI is indeed a powerful instrument explaining variation of credit amount. Other variables of household characteristics remain their signs as in the OLS FE regressions with the same or stronger significant level.

Regressions (2) to (5) are from 1997/98 data. Again, the coefficients of interest in all regressions turns consistently positive (0.11) and significant at 1% level. Regression (2) is estimated using three instruments LNCAPI; DIST1 and DIST2, with first-stage F-statistics equals 119.29 implying weak instrument problem is clearly

avoided. However, the LR test for redundancy of DIST1 and DIST2 cannot be rejected ($P(\chi^2(2)) = 0.16$), suggesting efficiency of the estimation is not improved by using one (or both) of these instruments. To check the robustness, regressions (3), (4) and (5) drop DIST2, DIST1, and both of the instruments out of estimation, respectively¹⁰. The coefficient on credit are unchanged, remaining value at 0.11 and significant at 1% level across regressions. First-stage F-statistics are large in all cases (>100). Hansen J-statistics of over-identification are overall small indicating the validity of the instruments used.

In summary, the OLS FE estimates are shown to be inconsistent in case of credit endogeneity. Under simple OLS estimations, credit is found having no significant effects on household consumption. Estimations are improved by employing instrumental variables to identify variation of borrowing amount. We find that, on average, taking credit does improve household consumption in both surveys. Because both consumption and credit are in logarithm form, the relationship is an elasticity. For instance, in 1992/93, the IV results (see Table D.2) suggest that one percent increase in average formal borrowing amount per capita would yield 0.14 percent increase in the average consumption per capita. Similarly, in 1997/98, 0.11 percent increase in average consumption per capita would be expected if formal borrowing per capita increases by one percent.

5.3 Impact Assessments using Longitudinal Data

5.3.1 Addressing Contemporaneous Impacts

Despite the promising results that have been presented, the instrumental variable estimations on cross-sectional data faces several fundamental limitations. Firstly, cross-section results may not be robust as it often depends significantly on statistical methods that used to address problem of endogeneity (i.e., credit participation)

¹⁰When LNCAAP is dropped, DIST1 and DIST2 becomes insignificant in the first-stage regression and F-stat is small (0.94). This suggests LNCAAP should be kept as the main instrument in all regressions.

and relatively strong assumptions that statistical framework bases on. Some studies show that sophisticated econometric model does not necessarily mean that the methodology could overcome limitation of cross-section data and provide reliable results (see, for example, Lalonde (1986)).

Secondly, using cross-sectional data cannot solve the problem of contemporaneous impact. Due to the fungible characteristics of credit, by borrowing household could temporarily increase consumption by spending all or part of loans they obtained on food or consumer goods rather than use the borrowing money to buy essential inputs for production. Report of consumption during or short-time after household's borrowing will be potentially upward biased. These drawbacks of cross-sectional data is generally recognised and a clear solution to improve the results is to employ longitudinal data which enable us to observe borrower's consumption in over time.

While longitudinal data analysis is widely applied in economic research, it is less popular in credit impact studies due to unavailability of data. To date, the most comprehensive work using panel data is from Khandker (2005). In the paper, Khandker estimates the impact of credit received on borrower's consumption by employing household-level fixed-effect method using full panel data collected from Bangladesh. To reduce estimation biases due to measurement error in data collection and time-varying heterogeneity of household characteristics, Khandker also employs eligible condition of land holding size as instrumental variable for borrowing activities. His results from panel data show that cross-section analysis does overestimate impacts of borrowing¹¹.

The VLSS data set has two periods of data on each respondent household, providing us a solution for contemporaneous impact problem. With the record of household borrowing activities in 1992/93, we could assess credit impacts on the same household's consumption in 1997/98. There are two clear reasons to do so. First,

¹¹Earlier, Pitt and Khandker (1998) use cross-section data from the first round survey and find that borrowing would increase annual consumption of female borrowers by 18 taka. Employing panel data, Khanker (2005) finds smaller increase of 8 taka.

investments take time to pay off and we would like to see how much help credit could be in enhancing borrower's consumption in long term. Second, using lags variable would reduce problem of simultaneity effects between credit and consumption. If we estimate impacts of preceding credit on current consumption, then it is reasonable to suppose that household's current consumption and other characteristics cannot affect lagged borrowing directly. However, if there is a variable that affects consumption today as well as borrowing activities yesterday such as household's land holding size, then today consumption may be indirectly correlated with lagged credit amount. Hence, problem of simultaneity cannot be avoided completely.

To account for this still a possible problem, we employ instrumental variables for credit amount in 1992/93. As such, we estimate the following system:

$$Y_{ij}^{98} = \beta_0 + \beta_1 L_{ij}^{93} + \beta_2 H_{ij}^{98} + \mu_j^Y + u_{ij} \quad (5.3)$$

$$L_{ij}^{93} = \alpha_0 + \alpha_1 Z_{ij} + \mu_j^L + v_{ij} \quad (5.4)$$

where Y_{ij}^{98} is the real per capita expenditure in 1997/98 of the household i in the commune j ; H_{ij}^{98} the corresponding household characteristics in 1997/98; μ_j the commune-level fixed-effects; and Z_{ij} the vector of instrumental variables that is correlated with credit amount obtained by household in 1992/93, but that would not directly influence household consumption in 1997/98.

Selection of variables to include in Z_{ij} is basing on the results from Chapter 4. These variables must significantly correlated with borrowing amount in 1992/93, therefore the instruments are also in lagged time. The first variable to include is the 1992/93 log of (one plus) commune credit per capita which has been proved as a strong instrument for formal credit variation. Others 1992/93 variables employed as instruments consist *household size, number of working adult in the household, health status of household head and his/her spouse, log of (one plus) irrigated land holding size, and log of (one plus) commune population*, which are all significant at

1% level in determining credit amount (see Table C.2). Further specification tests will be carried out to check the validity of these instruments.

5.3.2 Results

The longitudinal sample is constructed basing on households who appear in both 1992/93 and 1997/98 surveys. Of the 3,839 rural households included in the 1992/93 sample, 3,364 households are re-interviewed in 1997/98, which constitute the number of household in the longitudinal data.

The dependent variable for us to measure the impacts of credit is the log of (one plus) real per capita expenditure in 1997/98. As discussed before, we expect this variable to be positively influenced by household variables such as number of working adult or education of household head, while it may negatively correlated with others such as working in farm or household size. The credit variable is measured by log of (one plus) per capita credit amount, but in the lagged value. That is we measure impact of borrowing in 1992/93 on the consumption in 1997/98. Some control variables like *land holding size*, *commune credit per capita*, and *commune population* also appear in logarithm form, as before. Estimate results of the equation 5.3 are shown in Table D.3.

We start with OLS regression (1) which treats borrowing in 1992/93 as exogenous variable. No effect of credit has been found. The coefficient of interest is small (0.006) and not significant within the conventional level. Regressions (2) to (6) employ instrumental variables for the 1992/93 credit variable which is suspected to behave endogenously. Regression (2) is estimated by two-stage methods using 6 instruments for the credit variable: HHSIZE93, HHADULT93, HEALTH93, LNLAND93, LNPOP93, and LNCAP93. The coefficient on log credit becomes significant at 5% level. First-stage F-statistics (9.77) is close to the threshold (10), indicating a possible weak instrument problem. The Hansen J-statistics (J=26.12; p=0.00) strongly rejects the null hypothesis that the instruments are independent of the second-stage errors, suggesting existence of invalid instruments. GMM estimate

is applied in regression (3) using the same regressors and instruments as in (2) to generate efficient estimates of the coefficients as well as consistent estimates of the standard errors (Baum et al, 2003). The coefficient of interest remains significant at 10% level; other coefficients mostly unchanged.

To cope with the invalid instrument problem that picked up in regression (2) and (3), regression (4) only keeps 3 strongest instruments basing on the t-statistics in the first-stage regression (2). The remaining ones are HEALTH93, LNLAND93, LNCA93. GMM estimates is done in regression (5). Results are not much improved, even though F-statistics (18.32) suggesting weak instruments are now avoided. The coefficients on credit are almost unchanged (0.09) at 5% significant level. However, statistics of over-identification test remain large.

One possible explanation for correlation between the lagged instruments and the current consumption is little variation overtime (5 years) of some household characteristics, such as *size of households, commune population, or land holding size*. While these variables certainly have their influence on credit activities in 1992/93, they may also persist as important factors determining consumption in 1997/98. The lagged household characteristics appear not to be good instrument overall.

Regression (6) holds LNCA93 as the only instrument. The results gain strength. The coefficient of interested variable is positive, higher in value (0.113), and significant at 1% level. Weak instrument problem is eliminated (F-stat=44.32). Specification test for the null hypothesis that credit variable can be treated as exogenous is rejected at 1% level ($p=0.00$) suggesting instrumental variable is more appropriate method for estimating the consumption effects of borrowing¹². The finding is in line with previous results using cross-sectional data. Thus, borrowing from the formal sector has a significant positive effect on per capita consumption.

¹²Because the specification is exactly identified, we cannot test for the validity of this only instrument.

5.3.3 Alternative Specifications

The final step to assess the robustness of our results is to exploit full potential of the panel data. Different from the study of lagged credit which quietly simplifies the impacts of recent borrowing and weights more on far-lagged borrowing, in this section we assess impacts by investigating how changes in credit affect changes in outcomes.

Recalling the outcome equation (5.1) that we modify to incorporate time period t :

$$Y_{ijt} = \beta_0 + \beta_1 L_{ijt} + \beta_2 H_{ijt} + \mu_j + \nu_{ij} + u_{ijt} \quad (5.5)$$

where Y_{ijt} is the real per capita expenditure of household i in commune j at the period t (1992/93, 1997/98); μ_j and ν_{ij} are unmeasured determinants of household's credit activities that are time invariant and fixed within a commune j and a household i , respectively; L_{ijt} , H_{ijt} are credit and characteristic variables specified as in the previous models; u_{ijt} is the error term.

As before, we are interested in estimating the coefficient β_1 , which reflects the average impacts of credit. Unlike prior models where only cross-section commune fixed-effect is possible, panel data allows us to estimate a model with household fixed effects specified. That is ν_{ij} captures nonrandom household selection into the formal credit programme, such as household's management ability, energy level or talents. Similar to commune fixed effects μ_j , ν_{ij} is also assumed not changing over time. Because the household fixed effects method also resolves any commune-level endogeneity (Khandker, 2005), we can simplify our model by omitting commune level unmeasured fixed effects (μ_j) from equation (5.5)

An alternative to fixed effects is to apply random effects regression for the panel model. The fundamental difference between the two methods is that while fixed effects regression remove the unobserved effect which can be any time-constant ex-

planatory variables determining credit activities of a household, the random effects estimator is attractive when the unobserved effect is uncorrelated with all the explanatory variables (Wooldridge, 2002). In our case, household's unobserved characteristics, which are thought to affect its credit activities such as management skills, may be well correlated with included household variables. Hence, fixed effects estimation appears to be the appropriate selection¹³.

With a panel data set containing two period of observations like the VLSSs, household fixed-effect estimation is the same as differencing the data by household and running a cross-section regression on the differenced data. By taking the difference between 1992/93 and 1997/98, the time invariant commune and household characteristics are eliminated. Selection bias problem due to omitted unobservable variables is controlled. A consistent estimate of the impact β_1 is more likely to be obtained.

Results of the fixed-effect estimate are presented in Table D.4. The coefficient on credit variable remains positive (0.03) and significant at 1% level. Several variables at commune and household level have been dropped because of the differencing, such as the distance, number of years in school of household head. These variables are either unchanged over time or not available in both surveys.

Finally, because the household fixed-effect method may still not yield consistent estimates of the credit effects using panel data if our assumption on the time invariant of ν_{ij} is violated, i.e., the unmeasured household specific characteristics may change over time. For example, experience could be learnt differently between households, and hence production and consumption growth may vary, leading to inconsistent estimation. A standard solution to correct this time-varying heterogeneity is to use instrumental variables. As before, we employ credit per capita at commune-level as the instrument for household credit. Test result (see Table D.4) shows that household-level fixed-effect instrumental variable (FE-IV) method

¹³To be certain, we carry out a Hausman test which compares the fixed effects and random effects estimates by testing the hypothesis that there is no correlation between unobserved fixed effects and included explanatory variables. The test result (Chi2(12)=449.78, p>chi2=0.0000) rejects the hypothesis. In other words, fixed effects is preferred to random effects.

is more appropriate in estimating effects of credit on household consumption.

Column (2) in Table D.4 presents FE-IV results. The coefficient of interest is 0.263 and significant at 1% level. This value is greater than any other estimated coefficients of credit. The OLS results in Section D.1 reveal no significant impacts of borrowing in both cross-section data sets. The OLS-IV results, also on cross-section data, show significant effects of about 0.11 (elasticity) after controlling for potential bias by employing instrumental variable. Using similar method on longitudinal data shows the same result (coefficient of credit = 0.11, significant at 1% level). And the results from this section, applying full panel data analysis, display the statistically significant coefficients of 0.025 (without instrument) and of 0.263 (with instrument). These finding suggests that the estimated coefficients capturing positive impacts of credit are quite robust - they are not sensitive to changes in econometric specification. In addition, selection bias is clearly a major problem and not addressing them carefully will underestimate the effects of credit, as in our case.

5.4 Conclusion

This chapter aims to assess the impact of formal credit on household consumption using data from VLSS1992/93 and VLSS1997/98. Analyses using cross-section and longitudinal data have been used. Instrumental variable method is employed to tease out the impact of credit on household consumption from other potential influences and biases. Results suggest that borrowing from the formal sector significantly improves consumption per capita of borrowing households.

The challenges in impact assessments arise because credit is not provided to random households. On one hand, from the supply side, formal credit institutions carefully select areas to provide their financial services. The selected area is not necessary having "better" economic conditions compared to others, it may be poorer as result of credit targeted. However, by comparing the incidence of poverty in credit available and credit unavailable areas, which are different in characteristics,

researchers may mistakenly report biased results. On the other hand, from the demand side, households who have taken credit may be systematically different from others who have not. When the characteristics that cause these differences are observable (e.g. age, education, assets) and are taken into account, impact evaluation would be less problematic. However, the differences often arise from self-selection process, which are driven by unobserved traits of borrowers, such as ability. A simple evaluation of credit programme without calculating these attributions will be biased.

We use instrumental variable to identify the credit amount that household obtained. The variables of credit per capita at commune-level, distance from commune committee to the nearest government bank branch (in form of log-transformation and dummy) are employed as the instruments. Our results are consistent with other studies of the Vietnamese rural credit markets (e.g. Pham et al (2002)) that also find positive and significant effects of credit on consumption. In particular, we find in our panel analysis that 1 percent increase in average credit per capita borrowed from the formal sector leads to 0.26 percent increase in average consumption per capita.

Controlling for selection biases has been shown to be very important. In our study, estimations using simple OLS regression lead us to conclusion of no effects of credit. This result is clearly underestimated the more reliable (positive) results in which biases have been carefully addressed. However, a reservation should be pointed out. That is our core instrumental variables are at commune-level. These are useful to control for intra-communal differences in formal credit accessibility but they do not allow controlling for inter-commune or household-level unobserved characteristics. Although this problem has been addressed by using household fixed-effects, and by cluster-robust test statistics and standard errors, results might be different if other instrumental-variables had been used.

Chapter 6

Conclusion

The thesis is a contribution to the research literature of the Vietnamese rural credit market during 1990s. The data is from the Vietnam Living Standard Surveys carried out in 1992/93 and 1997/98, providing both cross-sectional and panel data. We aim to answer three main research questions: What are the determinant factors of household's accessibility to the formal credit sector in rural Vietnam? What are the determinants of credit amount obtained by households? And what are the impacts of taking credit on household consumption?

Chapter 2 provides an overview of the Vietnamese rural credit markets. We emphasize on some interesting findings from the stylised facts. First, there was a big expansion of the formal credit sector. From a very low level of 30% market share in 1992/93, the formal sector had grown to account for 50% of the rural credit markets in 1997/98. However, similar to other developing countries, the Vietnamese financial system remains largely underdeveloped with a dual structure in which formal and informal sectors exist side by side.

Second, although the development of formal sector and household borrowing activities are not equally distributed across regions, there exists an uniform access to formal credit at commune-level, i.e. every commune included in the data had access to formal institutions in 1997/98. However, credit access at household-level really diversifies. The most frequent borrowers are poor households but borrowers

who borrow biggest amounts are rich households.

Third, interest rates from informal lenders were really high, at about 90% annual in 1992/93. Together with the expansion of the formal sector, the interest rates of the informal sector had been driven down to approximately 40% annual in 1997/98, indicating a growing competition between the two sectors. In term of interest rates, formal loans appear to be cheaper than informal loans, however it often requires collateral.

Finally, the existence of informal lenders remains strong in both studied periods, presenting a complex linkages between agents in the credit markets. In addition, almost fifty percent of lending households are also borrowers at the same time. The existence of borrowing-lenders suggests that financial intermediation are active in rural areas. This indicates the different level of credit access among population. Some households may find themselves being excluded from the formal credit market and they have to seek alternative credit from the informal source, accepting higher interest rate and smaller loan amount.

Chapter 3 is a study assessing determinant factors of household's credit participation in the Vietnamese formal and informal credit markets. The literature on this research area is very little. To the best of our knowledge, we can only name two articles, Pham et al (2002) and Pham et al (2007), which have a similar research question to ours. However, Pham et al (2002) uses a small sample of 300 households collected by the authors. Pham et al (2007) uses the VLSS data but they includes urban households in their study. Our work is different from the two research mentioned above not only in term of data but also in term of methodology. We contribute to the literature as the first research attempts to estimate the credit demand and the credit supply functions separately to find determinant factors of credit access in rural Vietnam. We are also the pioneer in applying the bivariate probit model with partial observability in empirical study using data from Vietnam. The contributions and findings from this chapter are briefly summarised as follows.

Firstly, we find that education of household head has an inverse u-shape impact

on demand for credit. This result implies that low and high education household heads seems to have less demand for credit compared to household heads with secondary school degree. A possible explanation is that high educated people may be easier to find a paid job and they may choose not to work in farm or small self-business, therefore even if they have better access to the formal sector, their low demand for investment credit reduces the borrowing rate. On the opposite direction, there is high propensity that low educated people are poor and less skillful, and hence have lower demand for credit. In other words, while educated households may be self-rationed out of the formal credit market, low or non-educated households are more likely to be rationed out.

Secondly, the results show that land holding size significantly affects demand for and supply of formal credit. This suggests the importance of input capital and incentives for investment. On the supply, holding of fixed asset appears to be crucial for households to have access to the formal sector due to its requirement of collateral. Also demonstrated in our results, variables having impacts on supply decisions appears to differ considerably from those shaping household's demand. That is, access to the formal sector is primarily determined by regional measures, i.e., commune-level credit per capita and distance to bank branch. Indeed, the further distance from commune to a government bank branch, the less likely the bank will lend to commune's residents.

Further more, household characteristics determining level of access to the formal sector are different over the years between the two surveys, implying changes in the financial sector and economic environment. Prediction of formal credit market participation is high and increases over years, implying the expansion of the formal sector. Credit constraint is cut in half from 34 percent in 1992/93 to just about 15 percent in 1997/98. Over these years, the predicted demand for formal credit is also inflated from 40 percent to 52 percent suggesting household's need for investment capital growing with economic development.

The empirical analysis in the Chapter 4 highlights the determinant factors of

credit amount obtained. We started by employing simple OLS and Tobit model to cast an overview results. Then the Heckman two-stage model is applied to address problem of selection bias. A double-hurdle model is also used to check the robustness of the Heckman model's results. This piece of work again contributes as an unique study of the determinants of loan size to the literature of Vietnamese rural credit market.

We find that participation to the formal sector is significantly influenced by variables of household size, land holding size, and commune population. However, turning to the credit amount, household size seems not to be a significant factor. Instead, number of working adult in the household significantly and positively increase the amount. The interesting result is the negative sign of the coefficient on farm which indicates that peasant households borrow less than average. That is, poor rural households who seem to borrow most often are not those who borrow the largest loan size. Wealthier households are those who enjoy both better access and higher loan amounts from the formal sector.

The last chapter aims to assess the impact of formal credit on household consumption. Analyses using cross-section and longitudinal data have been used. The challenges in impact assessments arise because credit is not provided to random households. On one hand, from the supply side, formal credit institutions carefully select areas to provide their financial services. The selected area is not necessary having "better" economic conditions compared to others, it may be poorer as result of credit targeted. However, by comparing the incidence of poverty in credit available and credit unavailable areas, which are different in characteristics, researchers may mistakenly report biased results. On the other hand, from the demand side, households who have taken credit may be systematically different from others who have not. When the characteristics that cause these differences are observable (e.g., age, education, assets) and are taken into account, impact evaluation would be less problematic. However, the differences often arise from self-selection process, which are driven by unobserved traits of borrowers, such as ability. A simple evaluation of credit programme without calculating these attributions will be biased.

We use instrumental variable to identify the credit amount that household obtained. The variables of credit per capita at commune-level, distance from commune committee to the nearest government bank branch (in form of log-transformation and dummy) are employed as the instruments. Our results are consistent with other studies of the Vietnamese rural credit markets (e.g., Pham et al (2002)) that also find positive and significant effects of credit on borrower's consumption. In particular, we find in our panel analysis that 1 percent increase in average credit per capita borrowed from the formal sector leads to 0.26 percent increase in average consumption per capita. Controlling for selection biases has been shown to be very important. In our study, estimations using simple OLS regression lead us to conclusion of no effects of credit. This result is clearly underestimated the more reliable (positive) results in which biases have been carefully addressed. However, a reservation should be pointed out. That is our core instrumental variables are at commune-level. These are useful to control for intra-communal differences in formal credit accessibility but they do not allow controlling for inter-communal or household-level unobserved characteristics. Although this problem has been addressed by using household fixed-effects, and by cluster-robust test statistics and standard errors, results might be different if other instrumental variables had been used.

The results from this thesis imply some important policy implications. First, even though formal credit network continues to expand greatly to cover most rural areas, there is a question on the outreach and flexibility of credit services because rationing remains as a serious problem for rural households. Credit institutions should develop a better screening system, e.g. credit scoring system, to reduce collateral requirement which actually considered as one of the most obstacles of the formal credit access. Land regulations and fixed asset legal entitlement should be reviewed together with improvement of administration procedure to enable households to use their property as collateral if required.

Finally, the increased demand for credit implies more investment opportunities. Hence, on one hand, government should continue to improve the economic infrastructure to facilitate agricultural product trade market, provide better health

and education system to maintain this demand. On the other hand, credit institutions should be encouraged to expand their services to distant and mountain areas. Formal lenders should not limit themselves to lending only for production purposes but should diversify to include consumption.

In the future research, we should explore further the different development level across provinces in Vietnam which currently receives very little attention. In addition, a study of the interaction between formal and informal sectors will make great contribution to the literature. Further work also required to fully explain credit behaviours of households who borrow from both sectors and credit behaviours of households who come from different consumption quintiles. Currently, we still assume that loans from formal or informal lenders are homogeneous and lenders behave indifferently within its sector. This assumption is strong. We therefore may want to relax this assumption by studying household's credit from lenders rather than from sectors as we currently do.

Bibliography

- Abowd, J. & Farber, H. (1982), 'Job queues and the union status of workers', *Industrial and Labour Relation Review* 35(3), 354–367.
- Aghion, B. & Morduch, J. (2005), *The Economics of Microfinance*, The MIT Press.
- Aleem, I. (1990), 'Imperfect information, screening, and the costs of informal lending: A study of a rural credit market in pakistan', *The World Bank Economic Review* 4(3), 329–49.
- Amamiya, T. (1984), 'Tobit models: a survey', *Journal of Econometrics* 24, 3–61.
- Amemiya, T. (1985), *Advanced Econometrics*, Harvard University Press.
- Andersen, T. B. & Malchow-Moller, N. (2006), 'Strategic interaction in undeveloped credit markets', *Journal of Development Economics* 80, 275–298.
- Arabmazaar, A. & Schmidt, P. (1981), 'An investigation of the robustness of the tobit estimator to non-normality', *Econometrica* 50, 1055–1063.
- Baum, C., Schaffer, M. & Stillman, S. (2003), 'Instrumental variables and gmm: Estimation and testing', *Stata Journal* Vol. 3(1), 1–31.
- Bell, C. (1990), 'Interactions between institutional and informal credit agencies in rural india', *The World Bank Economic Review* 4(3), 297–327.
- Bell, C., Srinivasan, T. & Udry, C. (1997), 'Rationing, spillover, and interlinking in credit markets: the case of rural punjab', *Oxford Economic Papers* 49, 557–585.
- Blundell, R. & Meghir, C. (1987), 'Bivariate alternatives to the probit model', *Journal of Econometrics* 34, 179–200.

- Bound, J., Jaeger, D. & Baker, R. (1995), 'Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak', *Journal of the American Statistical Association* 90, 443–450.
- Cameron, C. & Trivedi, P. (2005), *Microeconometrics-methods and applications*, Cambridge University Press.
- Coleman, B. (1999), 'The impact of group lending in northeast thailand', *Journal of Development Economics* Vol. 60, 105–141.
- Coleman, B. (2002), Microfinance in northeast thailand: Who benefits and how much?, Technical report, Asia Development Bank, Working Paper Series No.9.
- Conning, J. (2001), Mixing and matching loans: Complementarity and competition amongst lenders in a rural credit market in chile, Technical report, Department of Economics, Hunter College, City University of New York.
- Cox, D. & Jappelli, T. (1993), 'The effects of borrowing constraints on consumer liabilities', *Journal of Money, Credit and Banking* 25(2), 197–213.
- Cragg, J. (1971), 'Some statistical models for limited dependant variables with applications to the demand for durable goods', *Econometrica* 39(5), 829–844.
- Daley-Harris, S. (2006), State of the microfinance summit campaign report 2006, Technical report, Washington, DC: Microcredit Summit Campaign.
- de Soto, H. (2000), *The Mystery of capital*, Basic Book.
- Degryse, H. & Ongena, S. (2005), 'Distance, lending relationships, and competition', *The Journal of Finance* 60(1), 231–66.
- Dell'Aricca, G. (2001), 'Asymmetric information and the market structure of the banking industry', *European Economic Review* 45, 1957–80.
- Diagne, A. (1999), Determinants of household access to and participation in formal and informal credit market in malawi, Technical report, Food Consumption and Nutrition Division, DISCUSSION PAPER NO. 67.

- Do, Q.-T. & Iyer, L. (2007), 'Land titling and rural transition in vietnam', *Economic Development and Cultural Change* (forth coming) .
- Do, Q.-T., Le, T.-T., Nguyen, V.-N. & Dinh, L.-T. (2001), *Credit*, UNDP and Vietnam Statistical Publishing House, chapter 12, pp. 201–219.
- Duan, N., Manning, W., Morris, C. & Newhouse, J. (1983), 'A comparison of alternative models for the demand for medical care', *Journal of Business and Economic Statistics* 1, 535–543.
- Feinstein, J. (1989), 'The safety regulations of u.s. nuclear power plants: Violations, inspections, and abnormal occurrences', *Journal of Political Economy* 97(1), 115–154.
- Fennema, J. (2005), Transition, Investment and Capacity Utilisation, PhD thesis, Heriot-Watt University. Ph.D. thesis.
- Fin, T. & Schmidt, P. (1984), 'A test of the tobit specification against an alternative suggested by cragg', *Review of Economics and Statistics* 66, 174–177.
- Fishe, R., Maddala, G. & Trost, P. (1979), Estimation of a heteroscedastic tobit model, Technical report, Manuscript, University of Florida.
- Floro, M. & Ray, D. (1997), 'Vertical links between formal and informal financial institutions', *Review of Development Economics* 1, 34–56.
- Friedman, M. (1957), *A Theory of Consumption Function*, Princeton University Press.
- Gine, X. (2005), Access to capital in rural thailand: an estimated model of formal vs. informal credit, Technical report, World Bank Policy Research Working Paper No. 3502, The World Bank.
- Glewwe, P., Agrawal, N. & Dollar, D. (2004), *Economic growth, poverty, and household welfare in Vietnam*, The World Bank.
- Gronau, R. (1973), 'The effects of children on the housewife's value of time', *Journal of Political Economy* 81, 168–199.

- GSO & WorldBank (2000), Basic information - vietnam living standards survey 1992-93, Technical report, Vietnam General Statistics Office and the World Bank.
- GSO & WorldBank (2001), Basic information - vietnam living standards survey (vlss) 1997-98, Technical report, Vietnam General Statistics Office and the World Bank.
- Hall, R. (1978), 'Stochastic implication of the life cycle-permanent income hypothesis: Theory and evidence', *The Journal of Political Economy* 86(6), 971-987.
- Hall, R. & Mishkin, F. (1982), 'The sensitivity of consumption to transitory income: Estimates from panel data on households', *Econometrica* 50, 461-81.
- Harper, M. (1998), *Profit for the Poor: Cases in Micro-finance*, Intermediate Technology Publisher.
- Haughton, D., Haughton, J. & Phong, N., eds (2001), *Living Standards During an Economic Boom*, UNDP and Vietnam Statistical Publishing House.
- Hayashi, F. (1985a), 'The effects of liquidity constraints on consumption: A cross-sectional analysis', *The Quarterly Journal of Economics* 100(1), 183-206.
- Hayashi, F. (1985b), Test for liquidity constraints: a critical survey, Technical report, NBER.
- Heckman, J. (1976), 'The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models', *Annals of Economic and Social Measurement* 5, 475-492.
- Heywood, J. & Mohanty, M. (1995), 'Estimation of the us federal job queue in presence of an endogenous union queue', *Economica* 62, 479-493.
- Hoff, K. & Stiglitz, J. (1990), 'Imperfect information and rural credit markets-puzzles and policy perspectives', *The World Bank Economic Review* 4(3), 235-250.

- Hoff, K. & Stiglitz, J. (1997), 'Moneylenders and bankers: price-increasing subsidies in a monopolistically competitive market', *Journal of Development Economics* 52, 429–462.
- Hossain, M. (1988), Credit for alleviation of rural poverty: the grameen bank in bangladesh, Technical report, International Food Policy Research Institute, Research Report No. 65.
- Iqbal, F. (1981), Rural savings, investment and interest rates in developing countries: Evidence from india, Technical report, The Rand Corporation, Santa Monica, Rand Note No. N-1763-AID.
- Jain, S. (1999), 'Symbiosis vs. crowding-out: the interaction of formal and informal credit markets in developing countries', *Journal of Development Economics* 59, 419–444.
- Jones, A. (1992), 'A note on computation of the double-hurdle model with dependence with application to tobacco expenditure', *Bulletin of Economic Research* 44(1), 67–74.
- Khandker, S. (2005), 'Microfinance and poverty: Evidence using panel data from bangladesh', *The World Bank Economic Review* 19, 263–286.
- Kochar, A. (1997), 'An empirical investigation of rationing constraints in rural markets in india', *Journal of Development Economics* 53, 339–371.
- Koehler, A. & Murphree, E. (1988), 'A comparison of the akaike and schwarz criteria for selecting model order', *Applied Statistics* 37(2), 187–195.
- Lalonde, R. (1986), 'Evaluating the econometric evaluations of training programs with experimental data', *American Economic Review* 76(4), 604–620.
- Lee, L. F. (1978); 'Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables', *International Economic Review* 19(2), 415–33.

- Lecce, D. (2000), 'Choice of mortgage instrument, liquidity constraints and the demand for housing debt in the uk', *Applied Economics* **32**, 1121–1132.
- Leung, S. & Yu, S. (1996), 'On the choice between sample selection and two-part models', *Journal of Econometrics* **72**, 197–229.
- Lucas, R. (1990), 'Why doesn't capital flow from rich to poor countries?', *The American Economic Review* **80**(2), 92–96.
- Maddala, G. (1983), *Limited-dependant and Qualitative Variables in Econometrics*, Cambridge University Press.
- Maddala, G. & Nelson, F. (1975), Specification errors in limited dependent variable models, Technical report, NBER working paper No. 96.
- Maki, A. (1993), 'Liquidity constraints: a cross-section analysis of the housing purchase behaviour of japanese households', *The Review of Economics and Statistics* **75**, 429–437.
- McCarty, A. (2001), Microfinance in vietnam: A survey of schemes and issues, Technical report, The State Bank of Vietnam and DFID.
- McKernan, S.-M. (2002), 'The impact of microcredit programs on self-employment profits: Do noncredit program aspects matter?', *The Review of Economics and Statistics* **84**(1), 93–115.
- Meng, C.-L. & Schmidt, P. (1985), 'On the cost of partial observability in the bivariate probit model', *International Economic Review* **26**(1), 71–85.
- Meyer, R. (2002), *Microfinance, Poverty Alleviation, and Improving Food Security: Implications for India*, Boca Raton Press.
- Morduch, J. (1998), Does microfinance really help the poor? new evidence from flagship programs in bangladesh, Technical report, Department of Economics, Harvard University.
- Morduch, J. (1999), 'The microfinance promise', *Journal of Economic Literature* **37**, 1569–1614.

- Morduch, J. & Haley, B. (2002), Analysis of the effects of microfinance on poverty reduction, Technical report, NYU Wagner Working Paper No. 1014.
- Nawata, K. & Nagase, N. (1996), 'Estimation of sample-selection bias models', *Econometric Reviews* 15, 387–400.
- Nguyen, S. (2007), 'Contending views and conflicts over land in vietnam's red river delta', *Journal of Southeast Asian Studies* 38(2), 309–334.
- Petrick, M. (2005), 'Empirical measurement of credit rationing in agriculture: a methodological survey', *Agricultural Economics* 33(2), 191–203.
- Pham, D. & Izumida, Y. (2002), 'Rural development finance in vietnam: a micro-econometric analysis of household surveys', *World Development* 30 (2), 319–335.
- Pham, T. & Lensink, R. (2007), 'Lending policies of informal, formal and semiformal lenders evidence from vietnam', *Economics of Transition* 15(2), 181–209.
- Pingali, P. & Xuan, V.-T. (1992), 'Vietnam: Decollectivisation and rice productivity growth', *Economic Development and Cultural Change* 40(4), 697–718.
- Pitt, M. & Khandker, S. (1998), 'The impact of group-based credit programs on poor households in bangladesh: Does the gender of participants matter?', *Journal of Political Economy* 106(5), 958–996.
- Poirier, D. (1980), 'Partial observability in bivariate probit models', *Journal of Econometrics* 12(2), 209–217.
- Ravallion, M. & Walle, D. V. D. (2006), 'Land reallocation in an agrarian transition', *The Economic Journal* 116(514), 924–943.
- Ravi, S. (2005), Access to credit and response to risk: Benefits of institutionalizing cooperation, Technical report, Department of Economics, New York University.
- Roman, L. (1997), *An Institutional Challenge: Transition in Vietnamese Banking Sector*, University of Stockholm, chapter 5, pp. 121–139.

- Rothenberg, T. (1971), 'Identification in parametric models', *Econometrica* **39**, 577–91.
- Sartori, A. (2003), 'An estimator for some binary-outcome selection models without exclusion restrictions', *Political Analysis* **11**, 111–138.
- SBV (1998), Annual report, Technical report, State Bank of Vietnam.
- Schmidt, P. & Strauss, R. (1976), 'The effects of unions on earnings and earnings on unions', *International Economic Review* **17(1)**, 204–12.
- Siamwalla, A., Pinthong, C., Poapongsakorn, N., Satsanguan, P., Nettayarak, P., Mingmancenakin, W. & Tubpun, Y. (1990), 'The thai rural credit system: Public subsidies, private information, and segmented markets', *The World Bank Economic Review* **4(3)**, 271–95.
- Staiger, D. & Stock, J. (1997), 'Instrumental variables regression with weak instruments', *Econometrica* **65(3)**, 557–586.
- Stock, J. & Yogo, M. (2005), *Testing for Weak Instruments in Linear IV Regression*, Cambridge University Press, chapter 5.
- Sussman, O. & Zeira, J. (1995), Banking and development, Technical report, CERP Discussion paper 1127.
- Swain, R. (2002), 'Credit rationing in rural india', *Journal of Economic Development* **27(2)**, 1–20.
- Timberg, T. & Aiyar, C. (1984), 'Informal credit markets in india', *Economic Development and Cultural Change* **33(1)**, 43–59.
- Udry, C. (1994), 'Risk and insurance in a rural credit market: An empirical investigation in northern nigeria', *The Review of Economic Studies* **61(3)**, 495–526.
- UNDP (2000), Compendium of rural development assistance in viet nam, Technical report, United Nations Development Programme.

- Varghese, A. (2005), 'Bank-moneylender linkage as an alternative to bank competition in rural credit markets', *Oxford Economic Papers* 57, 315–335.
- WB (1999), Vietnam development report 2000: Attacking poverty, Technical report, World Bank Report No. 19914-VN.
- WB (2007), Vietnam: Developing a comprehensive strategy to expand access to microfinance services. volume i: The microfinance landscape in vietnam, Technical report, World Bank.
- Wooldridge, J. (2002), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press.
- Zeldes, S. (1989), 'Consumption and liquidity constraints: An empirical investigation', *The Journal of Political Economy* 97(2), 305–46.

Appendix A

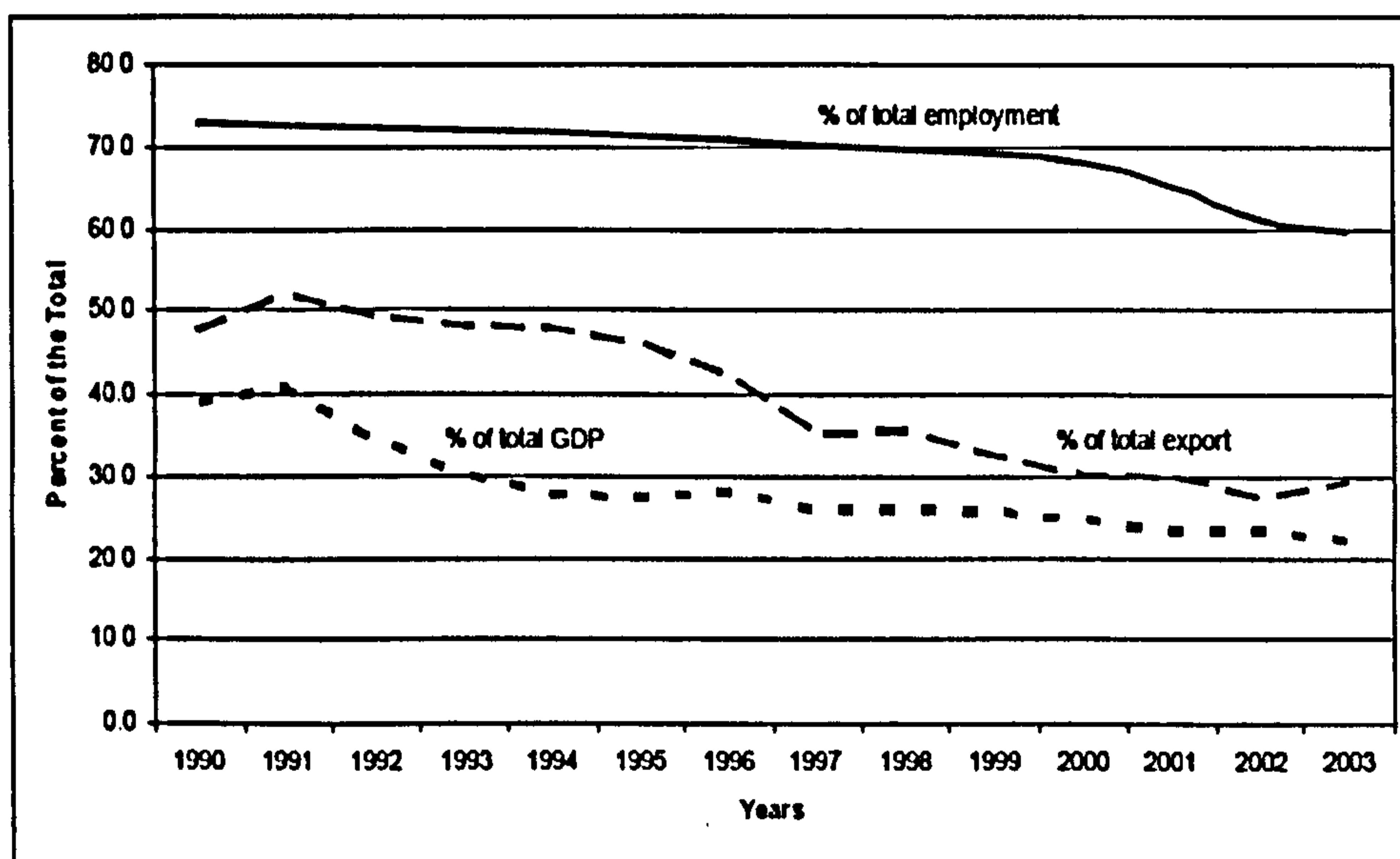
Tables and Figures Chapter2

Table A.1: Vietnam's Economic Indicators - 1990-2000

Figure A.1: Share of Agriculture in total GDP, Employment and Export

Year	GDP growth (%)	GDP per capita (USD)	Exchange rate (VND/USD)
1990	5.10	98	6,483
1991	5.96	113	10,037
1992	8.65	143	11,202
1993	8.07	187	10,641
1994	8.84	227	10,966
1995	9.54	284	11,038
1996	9.34	332	11,033
1997	8.15	356	11,683
1998	5.76	356	13,268
1999	4.77	370	13,943
2000	6.79	397	14,168

Source: APEC Energy Statistics (2003)



Source: World Bank (2006, Vol.1, p.18)

Figure A.2: Poverty Distribution - Map of Vietnam.

Table A.2: Mean characteristics of household and commune - VLSS1992/93

Variable	Pooled	Non- borrowing household	Borrowing household		
			Total	Formal borrower	Informal borrower
Male head of household (Male=1)	0.77 (0.42)	0.75 (0.43)	0.79 (0.4)	0.82 (0.39)	0.79 (0.41)
Age of household head	44.85 (14.79)	46.76 (15.49)	43.07 (13.87)	43.83 (13.54)	42.75 (13.89)
Marital status of head (Married = 1)	0.82 (0.38)	0.8 (0.4)	0.84 (0.37)	0.85 (0.36)	0.83 (0.37)
Household size	4.97 (2.12)	4.7 (2.08)	5.23 (2.12)	5.35 (2.00)	5.22 (2.16)
Working adult number	2.68 (1.26)	2.64 (1.26)	2.71 (1.27)	2.83 (1.29)	2.67 (1.26)
Education level of household head	1.57 (1.47)	1.62 (1.56)	1.52 (1.38)	1.57 (1.33)	1.5 (1.39)
Work in agriculture (Yes=1)	0.83 (0.38)	0.82 (0.39)	0.84 (0.37)	0.85 (0.35)	0.84 (0.37)
House ownership (Yes=1)	0.97 (0.18)	0.96 (0.19)	0.97 (0.16)	0.98 (0.15)	0.97 (0.16)
House area (m2)	61.85 (32.76)	64.29 (33.92)	59.6 (31.48)	63.14 (31.03)	57.68 (30.93)
House value (VND mil.)	7.83 (11.2)	8.45 (11.92)	7.25 (10.44)	8.18 (12.21)	6.98 (10.37)
Annual-crop-land holding size (m2)	918.78 (1346.88)	882.09 (1338.33)	953.06 (1354.26)	1145.32 (1516.63)	898.42 (1260.8)
Household's annual expenditure (VND mil.)	5.35 (3.87)	5.21 (3.94)	5.48 (3.81)	5.79 (4.04)	5.34 (3.82)
Expenditure per capita (VND mil.)	1.12 (0.71)	1.16 (0.76)	1.07 (0.65)	1.1 (0.59)	1.05 (0.65)
Commune's nonfarm employment (Yes=1)	0.46 (0.5)				
Commune population ('000)	8.69 (4.95)				
N	3839	1854 ^a (48.29%)	1985 ^a (51.71%)	798 ^b (40.20%)	1510 ^b (76.07%)

Note: ^a) As percentage of the whole sample; ^b) As percentage of borrowing households only. Total percentage larger than 100% due to some households borrow from both formal and informal sectors; Education level ranging from 1-6 including pre-school, primary school, lower secondary school, upper secondary school, vocational training, university; The land holding size is calculated basing on areas of land for agricultural production purpose only. Standard deviations are in parentheses.

Table A.3: Mean characteristics of household and commune - VLSS1997/98

Variable	Pooled	Non-borrowing household	Borrowing household		
			Total	Formal borrower	Informal borrower
Male head of household (Male=1)	0.78 (0.41)	0.75 (0.44)	0.81 (0.39)	0.82 (0.38)	0.81 (0.4)
Age of household head	47.29 (13.84)	49.97 (14.93)	45.07 (12.43)	45.54 (12.09)	44.64 (12.81)
Marrital status of head (Married = 1)	0.82 (0.38)	0.78 (0.42)	0.86 (0.35)	0.87 (0.34)	0.84 (0.36)
Household size	4.87 (1.94)	4.51 (2.04)	5.17 (1.8)	5.3 (1.83)	5.06 (1.74)
Working adult number	2.84 (1.27)	2.76 (1.26)	2.89 (1.27)	2.98 (1.29)	2.83 (1.25)
Education level of household head	2.51 (1.23)	2.37 (1.32)	2.62 (1.14)	2.63 (1.11)	2.6 (1.17)
Years in school of household head	6.44 (4.07)	6.08 (4.28)	6.74 (3.87)	6.73 (3.81)	6.75 (3.96)
Work in agriculture (Yes=1)	0.76 (0.43)	0.76 (0.43)	0.76 (0.42)	0.79 (0.41)	0.73 (0.44)
House ownership (Yes=1)	0.97 (0.17)	0.96 (0.2)	0.98 (0.14)	0.98 (0.12)	0.97 (0.17)
House area (m2)	67.7 (34.45)	67.02 (37.02)	68.25 (32.17)	70.04 (32.98)	66.28 (30.51)
House value (VND mil.)	23.84 (31.33)	24.18 (33.28)	23.57 (29.62)	23.29 (29.4)	23.58 (28.53)
Annual-crop-land holding size (m2)	2643.67 (4773.54)	2449.81 (4592.43)	2802.87 (4914.29)	3079.61 (5335.11)	2587.05 (4371.64)
Household's annual expenditure (VND mil.)	11.09 (7.38)	10.69 (8.06)	11.42 (6.75)	11.82 (7.01)	10.93 (6.32)
Expenditure per capita (VND mil.)	2.4 (1.43)	2.53 (1.67)	2.29 (1.19)	2.31 (1.19)	2.24 (1.17)
Commune's nonfarm employment (Yes=1)	0.55 (0.5)				
Commune population ('000)	9.02 (5.12)				
N	4269	1935 ^a (45.31%)	2334 ^a (54.69%)	1576 ^b (67.52%)	1192 ^b (53.35%)

Note: ^a) As percentage of the whole sample; ^b) As percentage of borrowing households only, the total percentage larger than 100% due to some households borrow from both formal and informal sectors; Education level ranging from 1-6 including pre-school, primary school, lower secondary school, upper secondary school, vocational training, university; The land holding size is calculated basing on areas of land for agricultural production purpose only. Standard deviations are in parentheses.

Table A.4: Characteristics of Borrowing Households by Credit Sectors

Source of loan	VLSS1992/93	VLSS1997/98
Formal sector¹	23.93	48.93
- VBARD & VBP	65.91	81.98
- Private banks	13.28	0.19
- Cooperative	24.69	4.12
- Other credit organisations	4.26	17.70
Informal sector¹	59.80	32.48
- Money lenders	25.89	23.91
- Relatives	57.22	50.25
- ROSCA	-	2.01
- Other individuals	29.21	35.07
Both sectors¹	16.27	18.59
Overall in-debt households²	51.71	54.69

Note: Numbers are in percentage. Numbers in bold are: ¹) percentages of total borrowing households; ²) percentages of total sample households. Numbers in sub-category are percentages of total borrowing households from the appropriate sector.

Table A.5: Distribution of Borrowing Households by Region

Region	VLSS93			VLSS98		
	Formal Borrower	Informal Borrower	All Borrower	Formal Borrower	Informal Borrower	All Borrower
Northern Uplands	38.73	78.1	51.81	68.48	47.29	57.68
Red River Delta	38.11	79.16	46.39	53.13	69.01	49.04
North Central Coast	52.92	71.35	59.48	67.7	54.21	59.33
South Central Coast	33.16	74.74	49.48	69.84	38.62	37.65
Highlands	25.37	80.6	52.34	71.68	45.58	61.41
South East	40.88	66.42	42.81	72.43	44.52	58.56
Mekong Delta	38.78	77.78	57.38	72.1	49.08	59.16
Full sample	40.20	76.07	51.71	67.52	51.07	54.69
N	798	1,510	1,985	1,576	1,192	2,334

Note: Numbers are in percentages. Numbers in All Borrower column are percentages of borrowing household in each region regardless of credit sectors. Results are calculated within region, and finance source are calculated within borrowing household. Result for all region as contribution toward total borrowing households in the sample is not necessary as number of sample among region are also vary depending on their weight of population.

Table A.6: Distribution of Borrowing Households by Expenditure Quintiles and Credit Sectors

	Expenditure per capita quintile					% of full sample	Per capita exp.
	Low	Low-mid	Middle	Up-mid	Upper		
VLSS93							
- Formal borrower	18.36	21.48	22.27	21.48	20.34	20.79	<i>1096.53</i>
- Informal borrower	46.22	42.58	38.93	36.59	32.33	39.33	<i>1052.93</i>
- All borrower	55.86	54.3	51.56	49.87	46.94	51.71	<i>1077.72</i>
- <i>Per capita exp.</i>	<i>514.77</i>	<i>740.57</i>	<i>945.92</i>	<i>1231.47</i>	<i>2160.01</i>		
VLSS98							
- Formal borrower	38.29	37.47	40.05	33.72	35.05	36.92	<i>2313.60</i>
- Informal borrower	31.62	30.56	29.27	24.59	23.56	27.92	<i>2244.45</i>
- All borrower	58.55	56.51	58.67	50.12	49.59	54.69	<i>2292.54</i>
- <i>Per capita exp.</i>	<i>1126.22</i>	<i>1652.12</i>	<i>2077.45</i>	<i>2671.15</i>	<i>4467.59</i>		

Note: Household per capita expenditure (number in *italic*) is measured in '000 VND. Other numbers are in percentages (borrowers).

Table A.7: Distribution of Loan Number and Loan Amount by Expenditure per Capita Quintile

	Expenditure per capita quintile					Total
	Low	Low-mid	Middle	Up-mid	Upper	
VLSS1992/93						
- Formal sector	17.67 (8.26)	20.68 (10.95)	21.43 (17.08)	20.68 (23.19)	19.55 (40.49)	100% (100%)
- Informal sector	23.51 (10.68)	21.66 (12.94)	19.80 (17.07)	18.61 (17.47)	16.42 (41.82)	100% (100%)
VLSS1997/98						
- Formal sector	20.75 (9.75)	20.30 (13.62)	21.70 (17.97)	18.27 (19.09)	18.97 (39.58)	100% (100%)
- Informal sector	22.65 (9.76)	21.90 (12.62)	20.97 (21.81)	17.62 (21.81)	16.86 (33.98)	100% (100%)

Note: Numbers are in percentages of loans and loan amount (in parentheses), by household per capital quintiles.

Table A.8: Distribution of Borrowing Households by Employment of Household Head

Employment	VLSS93				VLSS98			
	Pooled	Borrowing household			Pooled	Borrowing households		
		All	Formal	Informal		All	Formal	Informal
Agriculture	82.65	83.63	85.34	83.51	69.15	70.99	73.54	68.14
Self-emp	10.97	10.73	10.15	10.86	10.52	10.93	9.71	12.25
Paid-emp	5.05	4.69	3.51	4.64	11.78	12.13	11.10	13.09
Unemp	1.33	0.96	1.00	0.99	8.55	5.96	5.65	6.63
Total	100%	100%	100%	100%	100%	100%	100%	100%
N	3,839	1,985	798	1,510	4,269	2,334	1,576	1,192

Note: Numbers are in percentages.

Table A.9: Loan Characteristics - VLSS1992/93

Variable	Loan number (%)	Loan size (‘000 VND)	Loan duration (month)	Interest rate (%)	Collateral (%)
Formal sector	29.05	906.52	8.81	45.98	50.60 (31.33)
- VBARD & VBP	60.04	1092.63 (2224.47)	6.64 (6.97)	42.73	75.99
- Private banks	13.25	919.78 (1485.75)	9.84 (10.07)	59.41	5.61
- Cooperative	22.63	418.42 (973.9)	14.37 (25.62)	41.01	17.23
- Other credit organisations	4.08	831.73 (1394.91)	30.3 (31.54)	77.73	8.12
Informal sector	70.95	820.04	10.67	91.93	5.37
- Money lenders	22.91	1211.11 (3756.2)	8.23 (9.47)	109.39	11.21
- Relatives	51.33	726.59 (1385.5)	13.43 (16.31)	74.93	0.67
- Other individuals	25.76	658.46 (1520.92)	8.77 (8.89)	110.28	9.54

Note: Loan number (in bold) is percentage of total loans. Loan size is in ‘000 VND. Loan duration is in month. Interest rate is mean of annual interest rate (%), weighted by loan amount from the same lenders: $\bar{r} = \frac{\sum L_i r_i}{\sum L_i}$. Collateral is the percentage of collateralised loans amount. Numbers in sub-category are percentage within the sector. Standard deviation in parentheses

Table A.10: Loan Characteristics - VLSS1997/98

Variable	Loan number (%)	Loan value (‘000 VND)	Loan duration (month)	Interest rate (% annual)	Collateral (%)
Formal source	49.70	3638.13	18.40	14.97	74.61
		(5273.02)	(17.42)	(1.52)	(25.61)
- VBARD & VBP	78.16	4149.69	17.75	15.29	87.05
		(5741.94)	(15.40)		
- Private banks	0.17	1900.00	14.55	18.4	78.95
		(964.37)	(13.43)		
- Cooperative	3.96	3095.38	17.93	20.08	74.94
		(3392.72)	(47.05)		
- Other credit organisations	17.72	1518.7	24.85	12.37	19.65
		(1798.02)	(17.82)		
Informal source	50.30	2348.27	9.58	37.91	3.74
		(5942.11)	(14.12)	(9.15)	(5.15)
- Money lenders	20.97	2761.23	9.53	49.67	13.51
		(4798.27)	(19.44)		
- Relatives	46.06	2164.33	11.11	30.66	0.13
		(3471.82)	(8.86)		
- ROSCA	1.32	1577.04	20.27	85.89	0.00
		(1528.66)	(22.51)		
- Other individuals	31.65	2374.59	7.64	38.67	2.67
		(8867.31)	(6.79)		

Note: Loan number (in bold) is percentage of total loans. Loan size is in ‘000 VND. Loan duration is in month. Interest rate

is mean of annual interest rate (%), weighted by loan amount from the same lenders: $\bar{r} = \frac{\sum L_i r_i}{\sum L_i}$. Collateral is the

percentage of collateralised loans amount. Numbers in sub-category are percentage within the sector. Standard deviation in parentheses

Table A.11: Collateral assets

Variable	VLSS1997/98				Total
	Land	House	Furniture/ Fixed-assets	Others	
Formal source	58.59	31.64	4.77	5.00	100%
- Bank for the poor	44.81	42.21	10.39	2.60	
- Other government banks	62.20	27.87	4.05	5.88	
- Private banks	100.00	-	-	-	
- Other credit organisations	45.31	50.00	3.91	0.78	
Informal source	8.02	90.38	1.60	-	100%
- Money lenders	29.17	68.75	2.08	-	
- Relatives	3.77	95.60	0.63	-	
- ROSCA	-	-	-	-	
- Other individuals	4.95	92.08	2.97	-	

Note: Numbers are in percentage of collateralised loan number. In sub-category is the percentage within credit sectors.

Table A.12: Borrowing Purposes

Variable	VLSS93			VLSS98		
	Investment	Consumption	Repay/ Relend	Investment	Consumption	Repay/ Relend
Formal Sector	71.63	28.37	-	80.45	17.44	2.12
- VBARD & VBP	87.13	12.87	-	82.68	14.94	2.53
- Private banks	49.17	50.83	-	100	-	-
- Cooperative	47.80	52.20	-	60.56	39.44	-
- Other credit organisations	48.65	51.35	-	74.84	23.58	1.57
Informal Sector	43.45	56.32	.23	41.50	54.38	4.13
- Money lenders	59.83	39.49	.68	62.20	33.07	4.72
- Relatives	37.57	62.36	.07	27.84	68.10	4.06
- ROSCA	-	-	-	37.50	50.00	12.50
- Other individuals	41.09	58.77	.14	47.83	48.70	3.48

Note: Loan for consumption includes loan to build or buy house, consumer durables, general consumption on food and non-schooling and others. Investment consists loan for working capital and basic investment. All numbers are in percentage of numbers from the same source of credit.

Table A.13: Borrowing Lenders

Source of loan	VLSS1992/93		VLSS1997/98	
	Borrowing lenders	Other borrowers	Borrowing lenders	Other borrowers
Formal sector	41.90	39.88	63.16	68.22
- VBARD & VBP	74.24	64.26	83.82	81.71
- Private banks	9.09	14.11	0.00	0.22
- Cooperative	18.18	25.98	3.43	4.23
- Other credit organisations	7.58	3.60	18.14	17.64
Informal sector	58.10	60.12	36.84	31.78
- Money lenders	19.23	27.12	15.68	25.42
- Relatives	52.99	57.99	56.76	49.06
- ROSCA	-	-	5.95	1.29
- Other individuals	39.32	27.35	34.59	35.15

Notes: Borrowing lenders are households that borrow and lend simultaneously. Other borrowers include households that borrow but do not lend money. Numbers are in percentages.

Appendix B

Tables Chapter 3

Table B.1: Estimation Results: Univariate Probit Model (Model I)

Independent variables	Dependent Variable: Dummy for Borrowing by Credit Sector			
	Formal Sector		Informal Sector	
	1992/93	1997/98	1992/93	1997/98
GENDER	0.129 (0.081)	0.02 (0.078)	0.022 (0.072)	0.095 (0.081)
AGE	0.016 (0.011)	0.029 (0.013)**	-0.012 (0.011)	-0.017 (0.015)
AGESQR	-0.0002 (0.0001)*	-0.0003 (0.000)***	-0.000 (0.000)	0.000 (0.000)
MARITS	-0.078 (0.093)	0.069 (0.089)	-0.15 (0.075)*	-0.134 (0.084)
HHSIZE	0.052 (0.013)***	0.087 (0.012)***	0.061 (0.013)***	0.049 (0.013)***
HHADULT	0.196 (0.104)*	0.334 (0.089)***	0.071 (0.088)	0.194 (0.095)**
EDULEVEL	0.01 (0.02)	0.031 (0.027)	-0.022 (0.019)	-0.042 (0.035)
EDUYEAR		0.042 (0.023)*		0.008 (0.027)
EDUYEARSQR		-0.003 (0.001)***		0.001 (0.001)
FARM	0.073 (0.071)	0.158 (0.059)***	0.048 (0.063)	-0.062 (0.06)
HOUSE	0.16 (0.182)	0.34 (0.148)**	0.215 (0.124)*	0.048 (0.133)
LAND	0.098 (0.027)***	0.011 (0.007)*	-0.019 (0.027)	-0.007 (0.007)
POP	0.012 (0.01)	0.005 (0.007)	-0.001 (0.01)	0.014 (0.011)
HEALTH	0.01 (0.003)***	0.001 (0.004)	0.008 (0.003)**	0.015 (0.003)***
CAP	0.125 (0.051)**	0.029 (0.006)***	-0.013 (0.03)	-0.012 (0.007)*
DIST1		0.192 (0.088)**		-0.001 (0.112)
DIST2		0.069 (0.087)		-0.016 (0.11)
_cons	-2.093 (0.289)***	-2.53 0.355	-0.165 (0.271)	-0.223 (0.356)
Pr(participation)	0.2067	0.3710	0.3931	0.2791
Log Likelihood	-1875.96	-2534.30	-2518.66	-2384.60
Akaike Criterion	3779.92	5104.60	5065.331	4805.20
N	3839	4143	3839	4143

Note: Standard errors in parentheses are robust and adjusted for commune clustering. *, **, and *** indicate significance at the 10%, 5%, and 1% level

Table B.2: Estimation Results: Bivariate Probit Model

Independent variables	Dependent Variable: Dummy for Borrowing by Credit Sector			
	VLSS1992/93		VLSS1997/98	
	Formal Sector	Informal Sector	Formal Sector	Informal Sector
GENDER	0.129 (0.081)	0.022 (0.072)	0.02 (0.078)	0.095 (0.081)
AGE	0.016 (0.011)	-0.012 (0.011)	0.029 (0.013)**	-0.017 (0.015)
AGESQR	-0.0002 (0.0001)*	-0.000 (0.000)	0.000 (0.000)***	0.000 (0.000)
MARITS	-0.078 (0.093)	-0.15 (0.075)*	0.069 (0.089)	-0.134 (0.084)
HHSIZE	0.052 (0.013)***	0.061 (0.013)***	0.087 (0.012)***	0.049 (0.013)***
HHADULT	0.196 (0.104)*	0.071 (0.088)	0.334 (0.089)***	0.194 (0.095)**
EDULEVEL	0.01 (0.02)	-0.022 (0.019)	0.031 (0.027)	-0.042 (0.035)
EDUYEAR			0.042 (0.023)*	0.008 (0.027)
EDUYEARSQR			-0.003 (0.001)***	0.001 (0.001)
FARM	0.073 (0.071)	0.048 (0.063)	0.158 (0.059)***	-0.062 (0.06)
HOUSE	0.16 (0.182)	0.215 (0.124)*	0.34 (0.148)**	0.048 (0.133)
LAND	0.098 (0.027)***	-0.019 (0.027)	0.011 (0.007)*	-0.007 (0.007)
POP	0.012 (0.01)	-0.001 (0.01)	0.005 (0.007)	0.014 (0.011)
HEALTH	0.01 (0.003)***	0.008 (0.003)**	0.001 (0.004)	0.015 (0.003)***
CAP	0.125 (0.051)**	-0.013 (0.03)	0.029 (0.006)***	-0.012 (0.007)*
DIST1			0.192 (0.088)**	-0.001 (0.112)
DIST2			0.069 (0.087)	-0.016 (0.11)
_cons	-2.093 (0.289)***	-0.165 (0.271)	-2.53 0.355	-0.223 (0.356)
Log Likelihood	-4394.61		-4918.73	
rho	.0053		-.0159	
Wald test (H ₀ :rho=0) p-val	0.8871		0.6583	
N	3839		4143	

Note: Standard errors in parentheses are robust and adjusted for commune clustering. *, **, and *** indicate significance at the 10%, 5%, and 1% level

Table B.3: Estimation Results: Bivariate Probit Model with Partial Observability (Model II) of Formal Sector in 1992/93

Variable	Demand		Supply	
	Coefficient	Std. Err.	Coefficient	Std. Err.
GENDER	0.1285	0.1283	-0.0912	0.1925
AGE	0.0083	0.0201	0.0187	0.0309
AGESQR	-0.0001	0.0002	-0.0002	0.0003
MARITS	-0.1437	0.1442	0.2182	0.2234
HHSIZE	0.088	0.0238***	-0.0794	0.033*
HHADULT	0.3643	0.2016*	-0.3529	0.3191
EDULEVEL	0.0571	0.0347	-0.0779	0.0504
FARM	0.1091	0.1141	-0.1708	0.1778
HOUSE	0.0535	0.267	0.0204	0.3908
LAND	0.1627	0.0294***	-0.0921	0.0372**
POP	0.0143	0.0098	0.018	0.0184
HEALTH	0.0121	0.0035***		
CAP			1.6978	0.3849***
_cons	-1.5168	0.4906***	0.1039	0.971
Pr(Demand)	.3963	.1230		
Pr(Supply)	.7600	.1930		
Pr(Supply Demand)	.5261	.3288		
rho	-.8392			
LR test (H0: rho=0) p-val	0.0170			
Log likelihood	-1712.36			
Akaike criterion	3478.72			
N	3839			

Note: Standard errors in parentheses are robust and adjusted for commune clustering. *, **, and *** indicate significance at the 10%, 5%, and 1% level

Table B.4: Estimation Results: Bivariate Probit Model with Partial Observability (Model II) of Formal Sector in 1997/98

Variable	Demand		Supply	
	Coefficient	Std. Err.	Coefficient	Std. Err.
GENDER	0.0735	0.116	-0.0268	0.2038
AGE	0.0444	0.0183**	-0.0244	0.0324
AGESQR	-0.0005	0.0002***	0.0003	0.0003
MARITS	-0.0126	0.1371	0.0819	0.2485
HHSIZE	0.1139	0.0217***	-0.0849	0.035**
HHADULT	0.4355	0.1975**	-0.2892	0.345
EDUYEAR	0.0915	0.0282***	-0.0667	0.0469
EDUYEARSQR	-0.0052	0.0019***	0.0037	0.0031
FARM	0.3059	0.0796***	-0.2615	0.1284*
LAND	-0.0041	0.0067	0.0399	0.0235**
POP	0.0063	0.007	0.0124	0.0114
HEALTH	0.0005	0.0034		
CAP			0.2426	0.0356***
DIST1	0.2023	0.0857**	0.0233	0.1411
DIST2	0.0345	0.0858	-0.0697	0.145**
_cons	-2.3247	0.4209***	1.326	0.7156*
Pr(Demand)	0.5113	0.1494		
Pr(Supply)	0.8602	0.1420		
Pr(Supply Demand)	0.7230	0.2651		
rho	-0.8965			
LR test (H0: rho=0) p-val	0.0022			
Log likelihood	-2450.53			
Akaike criterion	4963.07			
N	4143			

Note: Standard errors in parentheses are robust and adjusted for commune clustering. *, **, and *** indicate significance at the 10%, 5%, and 1% level

Table B.5: Estimation Results: Bivariate Probit Model with Partial Observability (Model II) of Informal Sector in 1997/98

Variable	Demand		Supply	
	Coefficient	Std. Err.	Coefficient	Std. Err.
GENDER	0.0096	0.0982	0.4021	0.3222
AGE	-0.0313	0.0149*	0.0279	0.0544
AGESQR	0.0001	0.0001	0.0001	0.0006
MARITS	-0.0966	0.105	-0.1714	0.3772
HHSIZE	0.0846	0.0209***	-0.2246	0.0777***
HHADULT	0.0675	0.1269	0.0947	0.6047*
EDUYEAR	-0.0109	0.0082	0.0348	0.0277
FARM	-0.0857	0.0689	-0.016	0.2837
LAND	0.0041	0.0069	-0.0302	0.0185
POP			0.0953	0.0489**
HEALTH	0.0078	0.0047*	0.1366	0.0977
Predicted Formal Access	-0.6509	0.1551***		
_cons	0.9173	0.3888	0.3984	1.355
Pr(Demand)	.3209	.0993		
Pr(Supply)	.9565	.0695		
Pr(Supply Demand)	.8872	.1565		
rho	-.9346			
LR test (H0: rho=0) p-val	0.2327			
Log likelihood	-2365.90			
Akaike criterion	4781.80			
N	4143			

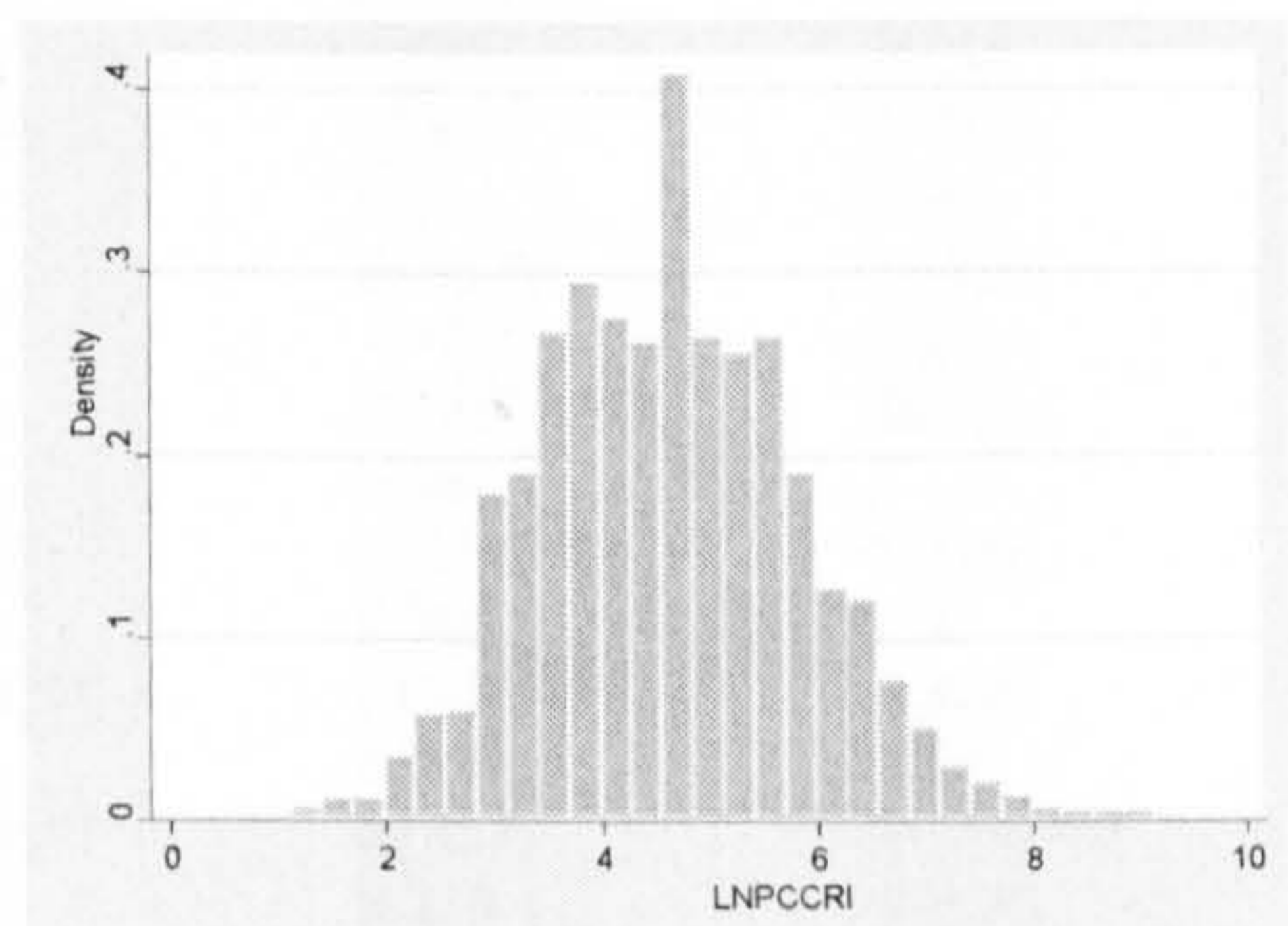
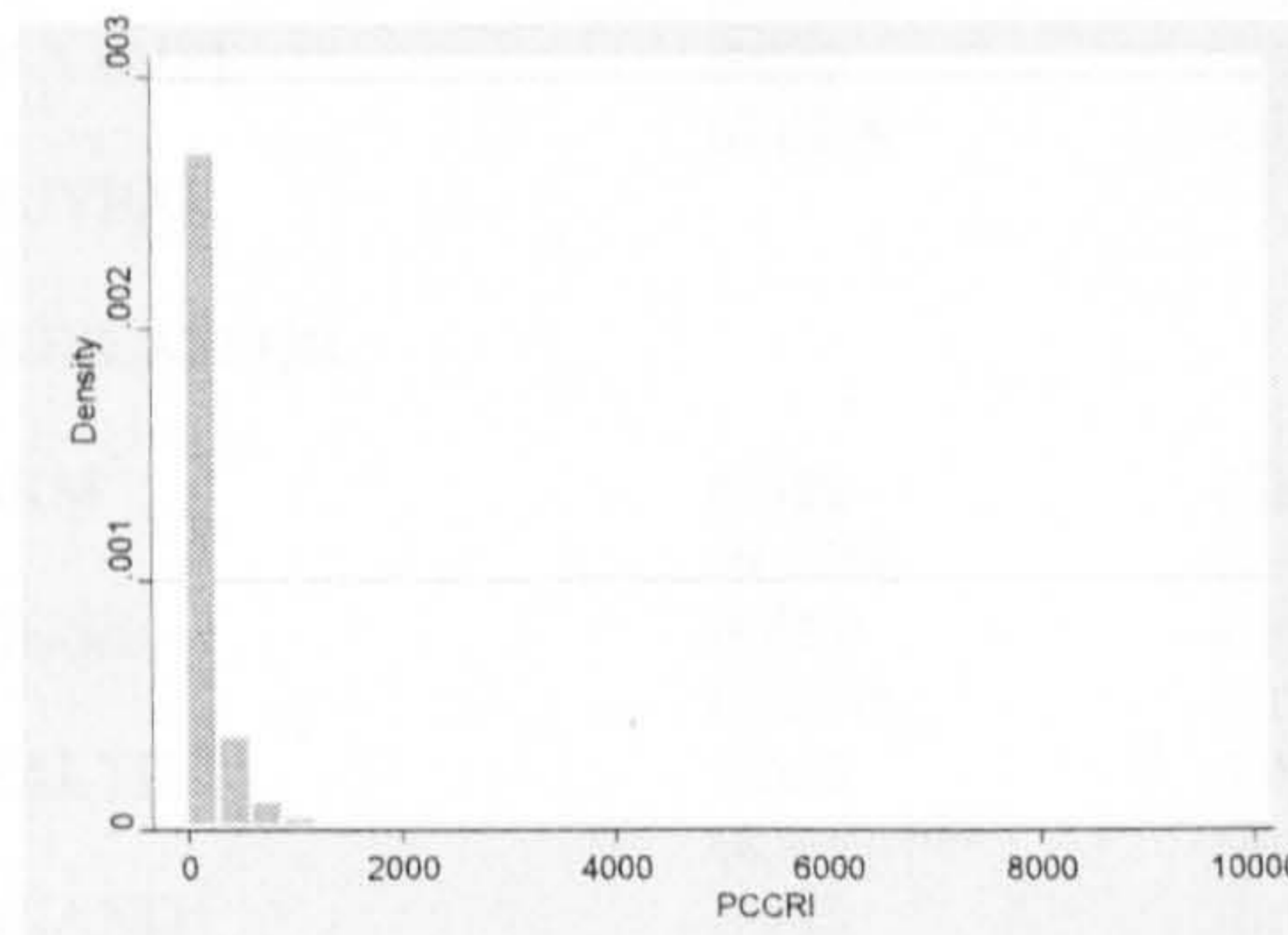
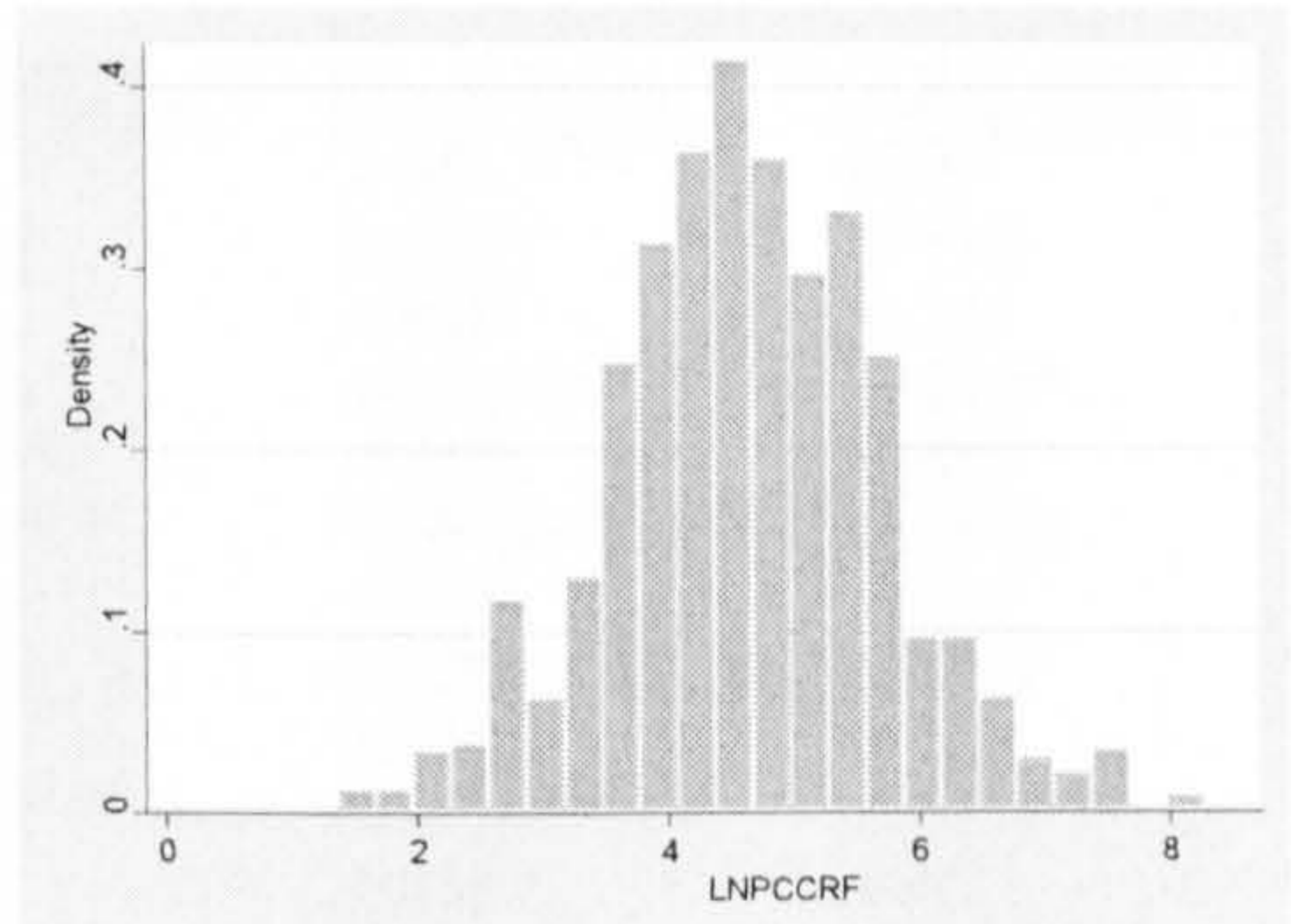
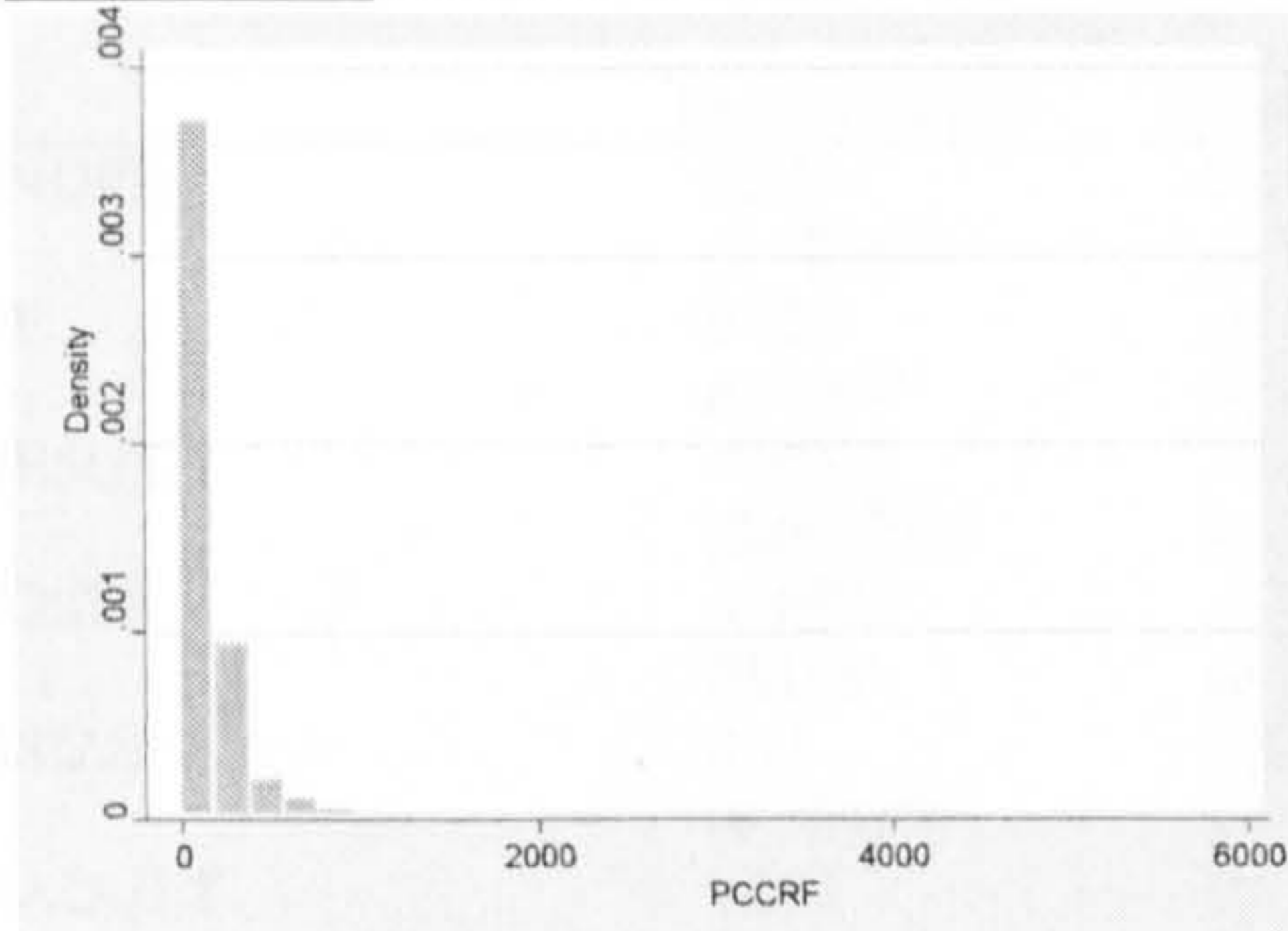
Note: Standard errors in parentheses are robust and adjusted for commune clustering. *, **, and *** indicate significance at the 10%, 5%, and 1% level

Appendix C

Tables Chapter 4

Figure C.1: Distribution of the credit amount obtained in level and in logarithm transformation

VLSS1992/93 Formal & Informal Sector



VLSS1997/98: Formal & Informal Sector

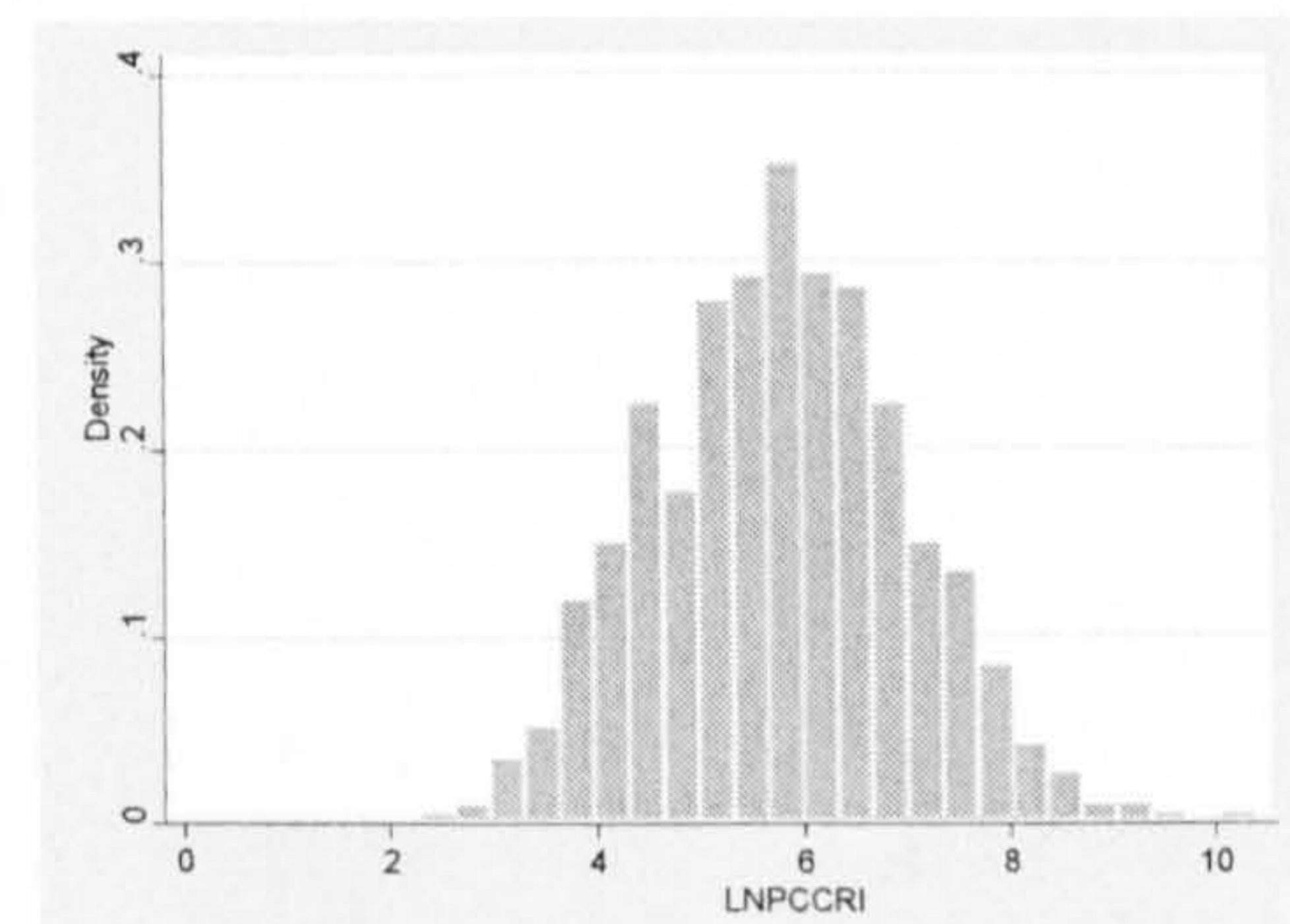
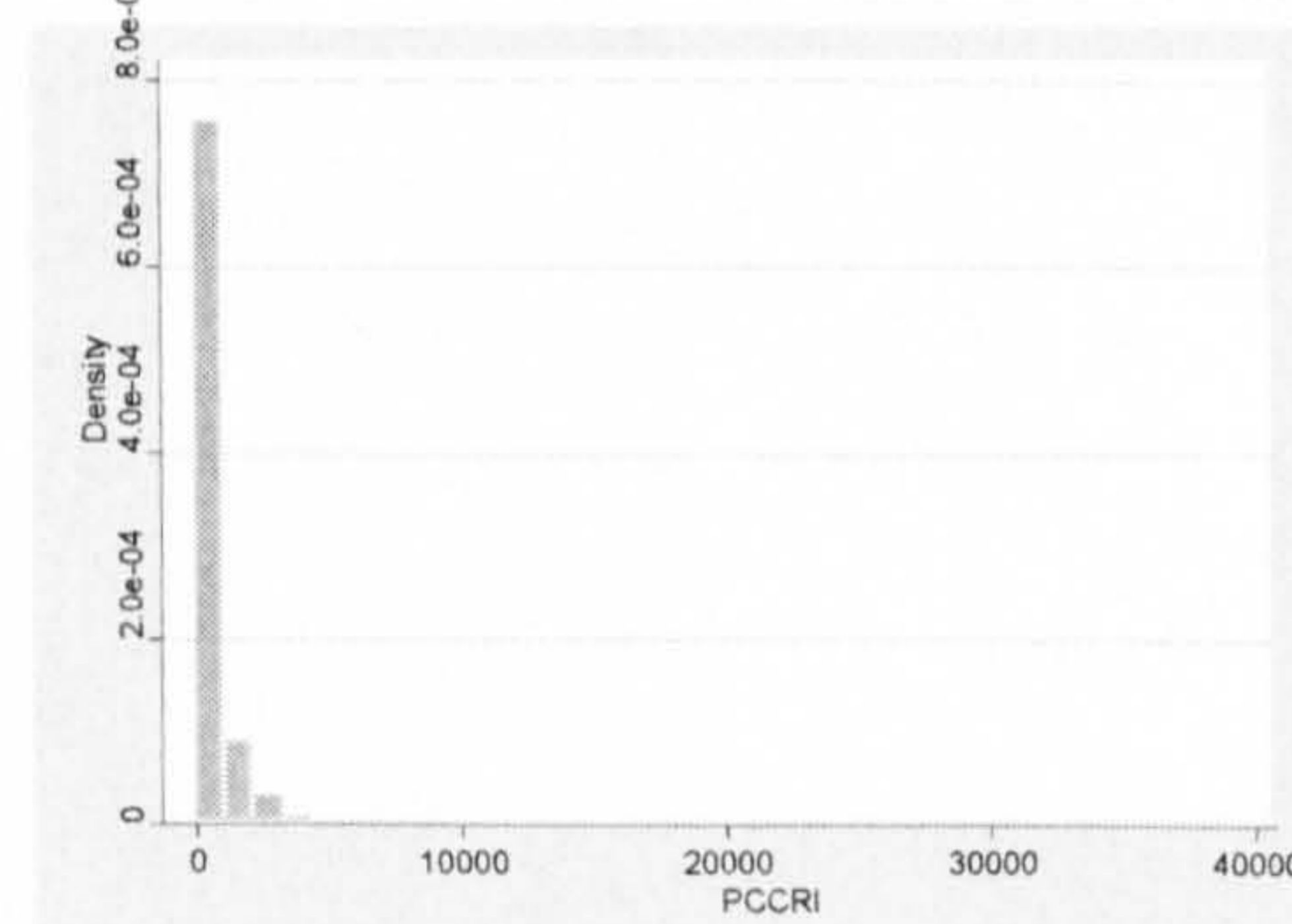
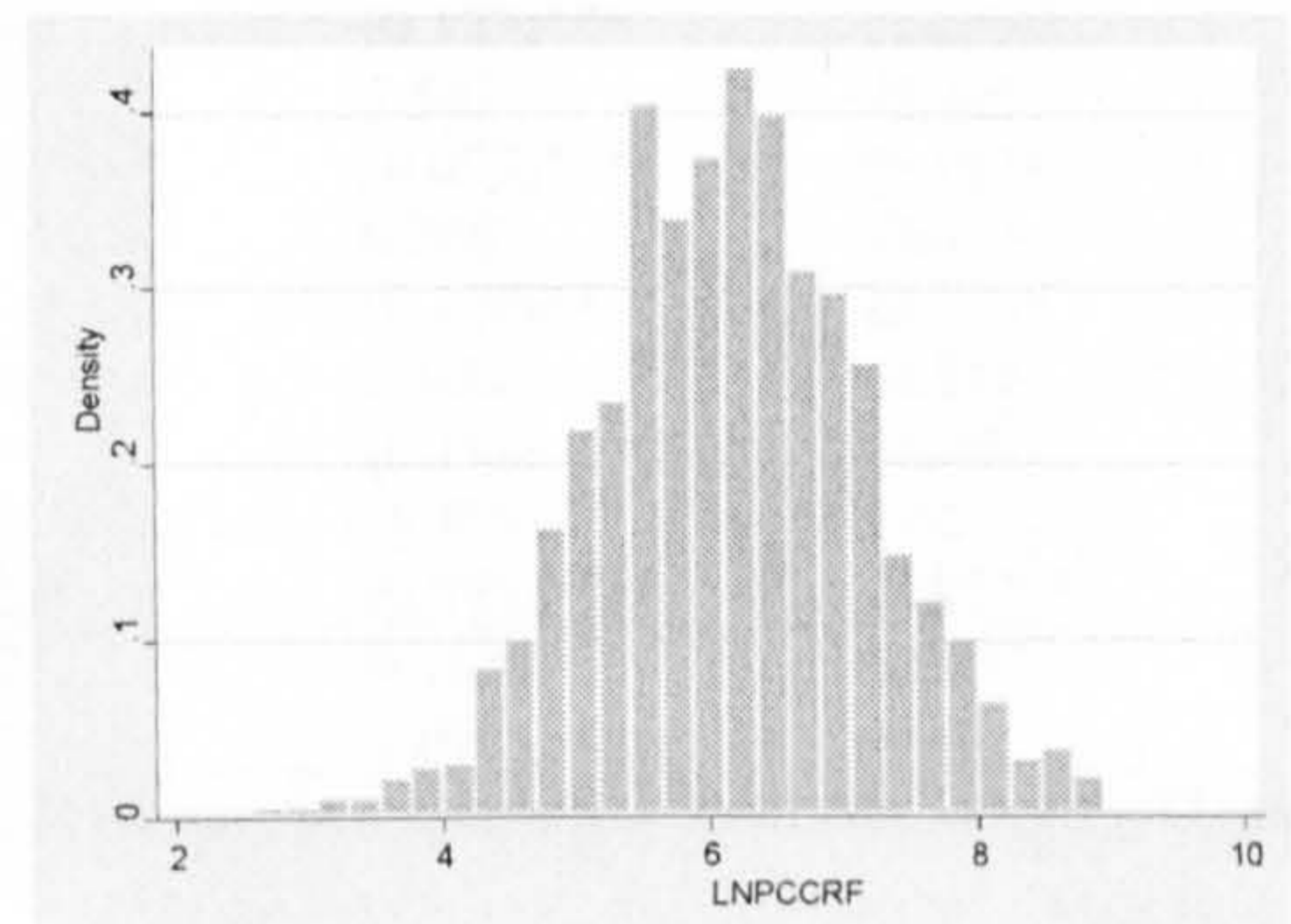
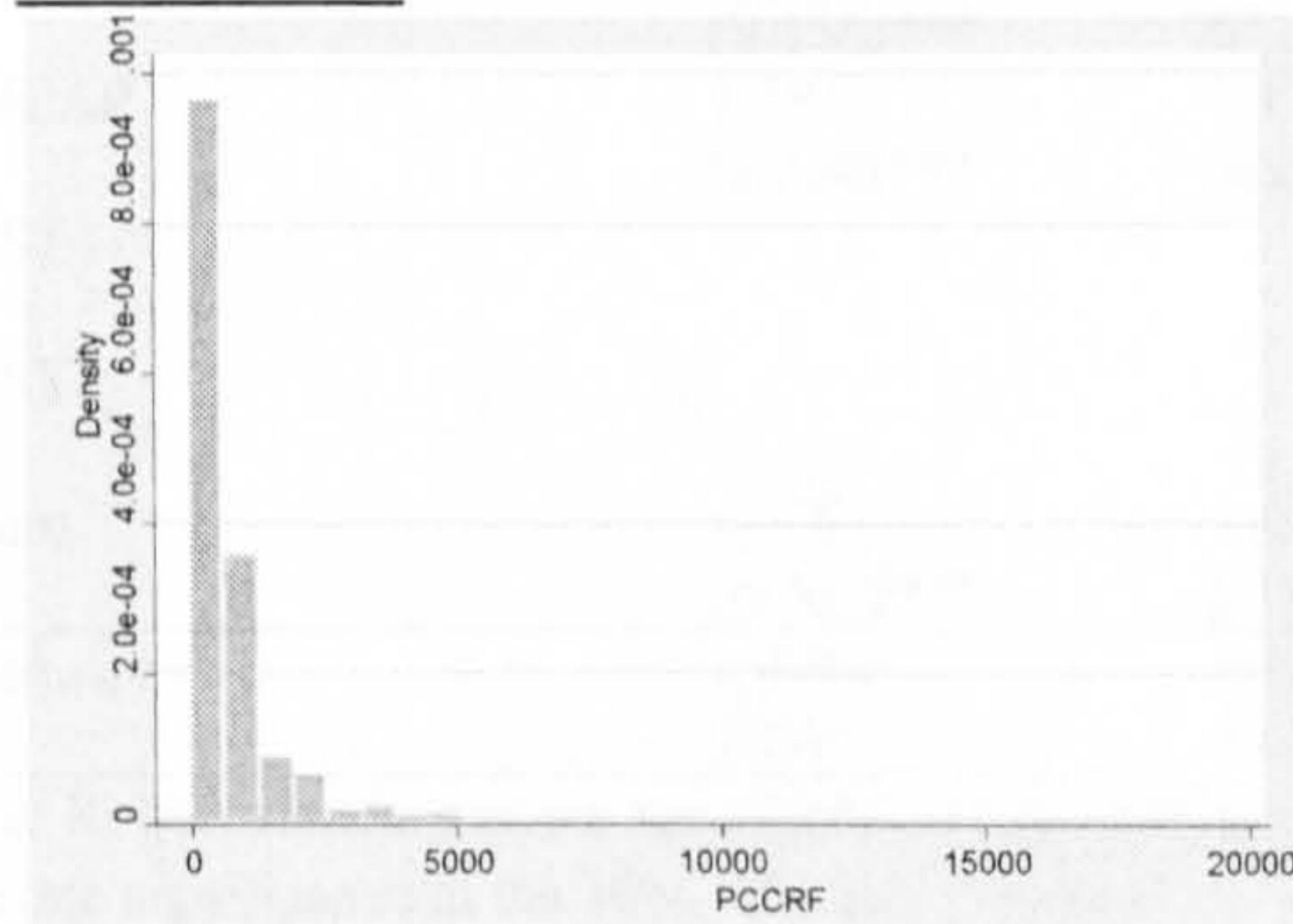


Table C.1: Estimation Results: OLS

Independent variables	Dependent Variable: Ln(1+Credit Amount)			
	VLSS1992/93		VLSS1997/98	
	Formal Sector	Informal Sector	Formal Sector	Informal Sector
GENDER	0.245 (0.123)**	0.02 (0.175)	0.081 (0.194)	0.217 (0.196)
AGE	0.029 (0.017)*	-0.023 (0.023)	0.087 (0.03)***	-0.055 (0.033)*
AGESQR	-0.0003 (0.0002)**	-0.00004 (0.0002)	-0.001 (0.0002)***	0.0002 (0.0003)
MARITS	-0.097 (0.143)	-0.226 (0.173)	0.187 (0.222)	-0.307 (0.207)
HHSIZE	0.11 (0.024)***	0.189 (0.028)***	0.203 (0.031)***	0.136 (0.031)***
HHADULT	0.291 (0.167)*	0.135 (0.201)	0.628 (0.21)***	0.485 (0.209)**
EDULEVEL	0.035 (0.029)	-0.012 (0.043)	0.057 (0.071)	-0.067 (0.086)
EDUYEAR			0.15 (0.057)***	0.016 (0.062)
EDUYEARSQR			-0.009 (0.003)***	0.002 (0.003)
FARM	0.028 (0.106)	0.043 (0.154)	0.463 (0.159)***	-0.194 (0.168)
HOUSE	0.225 (0.234)	0.563 (0.242)**	0.902 (0.332)***	0.08 (0.301)
HEALTH	0.016 (0.006)***	0.013 (0.008)***	0.000 (0.011)	0.036 (0.008)***
LNLAND	0.413 (0.109)***	-0.114 (0.147)	0.259 (0.104)**	-0.008 (0.117)
LNPOP	0.535 (0.156)***	0.311 (0.236)	0.812 (0.173)***	0.536 (0.27)**
LNCAP	1.551 (0.228)***	0.128 (0.25)	1.541 (0.093)***	-0.145 (0.164)
DIST1			0.268 (0.129)**	-0.027 (0.123)
DIST2			0.106 (0.112)	-0.074 (0.204)
_cons	-2.538 (0.582)***	1.436 (0.828)*	-5.855 (0.791)***	1.825 (0.853)**
R-square	0.0984	0.0342	0.1429	0.0337
N	3839	3839	4143	4143

Note: Robust standard errors are reported in parentheses and adjusted for clustering at commune level. *, **, and *** indicate significance at the 10%, 5%, and 1% level

Table C.2: Estimation Results: the Tobit model

Independent variables	Dependent Variable: Ln(1+Credit Amount)			
	VLSS1992/93		VLSS1997/98	
	Formal Sector	Informal Sector	Formal Sector	Informal Sector
GENDER	0.938 (0.624)	0.089 (0.412)	0.137 (0.515)	0.74 (0.653)
AGE	0.145 (0.095)	-0.058 (0.062)	0.266 (0.089)***	-0.161 (0.106)
AGESQR	-0.002 (0.001)*	-0.0002 (0.001)	-0.003 (0.001)***	0.000 (0.001)
MARITS	-0.352 (0.703)	-0.768 (0.467)*	0.62 (0.59)	-1.137 (0.735)
HHSIZE	0.461 (0.103)***	0.432 (0.069)***	0.537 (0.087)***	0.483 (0.112)***
HHADULT	1.613 (0.918)*	0.544 (0.603)	2.342 (0.712)***	1.954 (0.879)**
EDULEVEL	0.101 (0.141)	-0.09 (0.092)	0.167 (0.198)	-0.312 (0.252)
EDUYEAR			0.416 (0.146)***	0.046 (0.182)
EDUYEARSQR			-0.027 (0.008)***	0.007 (0.01)
FARM	0.482 (0.529)	0.171 (0.345)	1.087 (0.37)***	-0.651 (0.454)
HOUSE	1.314 (1.184)	1.463 (0.751)*	3.042 (0.993)***	0.449 (1.091)
HEALTH	0.086 (0.026)***	0.043 (0.018)**	0.008 (0.026)	0.136 (0.031)***
LNLAND	2.268 (0.33)***	-0.145 (0.224)	0.749 (0.19)***	-0.035 (0.25)
LNPOP	2.546 (0.428)***	0.413 (0.271)**	1.824 (0.343)***	1.663 (0.43)***
LNCAP	6.538 (0.429)***	0.146 (0.275)	3.753 (0.207)***	-0.763 (0.25)***
DIST1			0.777 (0.317)**	0.008 (0.399)
DIST2			0.358 (0.317)	-0.163 (0.394)
_cons	-18.742 (2.626)***	-2.454 (1.62)	-27.954 (2.514)***	-3.145 (2.132)
Log likelihood	-3695.64	-6449.92	-6452.41	-5382.18
N	3839	3839	4143	4143

Note: Robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level

Table C.3: Estimation Results: Heckman Selection Model for the Formal Sector Credit Amount Obtained

Independent variables	VLSS1992/93		VLSS1997/98	
	Selection	Credit Amount	Selection	Credit Amount
GENDER	0.118 (0.079)	0.16 (0.14)	0.032 (0.073)	0.076 (0.229)
AGE	0.018 (0.012)	0.028 (0.02)	0.035 (0.013)***	0.107 (0.058)*
AGESQR	-0.0002 (0.000)*	-0.0003 (0.000)	-0.0003 (0.000)***	-0.001 (0.001)*
MARITS	-0.061 (0.09)	0.085 (0.145)	0.06 (0.083)	0.265 (0.276)
HHSIZE	0.053 (0.013)***	-0.016 (0.032)	0.073 (0.013)***	0.016 (0.096)
HHADULT	0.178 (0.116)	0.67 (0.215)***	0.297 (0.101)***	0.838 (0.501)*
EDULEVEL	0.01 (0.018)	0.081 (0.031)***	0.029 (0.028)	0.046 (0.093)
EDUYEAR			0.054 (0.021)**	0.113 (0.092)
EDUYEARSQR			-0.004 (0.001)***	-0.006 (0.006)
FARM	0.07 (0.068)	-0.248 (0.116)**	0.164 (0.053)	0.24 (0.253)
HOUSE	0.153 (0.149)	0.338 (0.27)	0.359 (0.137)***	1.257 (0.669)*
LNLAND	0.299 (0.042)***	0.045 (0.152)	0.106 (0.028)***	0.257 (0.151)*
LNPOP	0.279 (0.054)***	0.945 (0.152)***	0.168 (0.049)***	1.265 (0.249)***
LNCAP	0.785 (0.051)***	1.472 (0.35)***	0.484 (0.029)***	1.703 (0.578)***
HEALTH	0.011 (0.003)***		0.001 (0.004)	
DIST1			0.161 (0.057)***	0.123 (0.255)
DIST2			0.088 (0.056)	-0.017 (0.204)
_cons	-2.986 (0.321)***	-0.973 (1.821)	-3.703 (0.339)***	-8.21 (6.049)
Lamda	0.963	s.e. 0.598*	3.12	s.e. 1.79*
rho	0.757		1.00	
N	3839		4143	

Note: Dependent variable is LNPCCRF. Standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level

Table C.4: Estimation Results: Double Hurdle Selection Model for the Formal Sector
Credit Amount Obtained

Independent variables	VLSS1992/93		VLSS1997/98	
	Selection	Credit Amount	Selection	Credit Amount
GENDER	0.118 (0.079)	0.109 (0.121)	0.036 (0.073)	0.023 (0.079)
AGE	0.018 (0.012)	0.023 (0.018)	0.035 (0.013)***	0.044 (0.014)***
AGESQR	-0.0002 (0.0001)*	-0.0003 (0.0002)	-0.0004 (0.0001)***	-0.0003 (0.0001)***
MARITS	-0.063 (0.09)	0.107 (0.131)	0.056 (0.083)	0.135 (0.093)
HHSIZE	0.053 (0.013)***	-0.037 (0.023)	0.075 (0.013)***	-0.114 (0.015)***
HHADULT	0.178 (0.116)	0.589 (0.186)***	0.29 (0.101)***	0.27 (0.121)**
EDULEVEL	0.009 (0.018)	0.079 (0.029)***	0.028 (0.028)	-0.001 (0.03)
EDUYEAR			0.055 (0.021)***	0.017 (0.023)
EDUYEARSQR			-0.004 (0.001)***	0.0003 (0.001)
FARM	0.068 (0.068)	-0.277 (0.103)***	0.162 (0.053)***	-0.051 (0.059)
HOUSE	0.154 (0.149)	0.258 (0.241)	0.365 (0.137)***	0.511 (0.177)***
LNLAND	0.3 (0.042)***	-0.085 (0.084)	0.104 (0.028)***	0.069 (0.029)**
LNPOP	0.281 (0.054)***	0.832 (0.098)***	0.178 (0.049)***	0.964 (0.052)***
LNCAP	0.789 (0.052)***	1.154 (0.162)***	0.49 (0.03)***	0.852 (0.052)***
HEALTH	0.012 (0.003)***		0.003 (0.004)	
DIST1			0.163 (0.057)***	-0.18 (0.06)
DIST2			0.09 (0.056)	0.728 (0.577)
_cons	-2.992 (0.321)***	0.648 (0.886)	-3.758 (0.34)***	-0.158 (0.06)
Likelihood ratio	-2896.90		-4346.56	
rho	0.3946		0.4443	
Wald test (H ₀ :rho=0) p-val	0.1089		0.0049	
N	3839		4143	

Note: Dependent variable is LNPPCRF. Standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level

Appendix D

Tables Chapter 5

Table D.1: Estimation Results: Impacts of Formal Credit - OLS and OLS-FE estimates

Log per capita expenditure	VLSS1992/93		VLSS1997/98	
	OLS (1)	OLS FE (2)	OLS (3)	OLS FE (4)
LNPCCRF	0.007 (0.005)	0.01 (0.004)***	-0.002 (0.002)	0.001 (0.002)
GENDER	0.012 (0.027)	0.011 (0.023)	-0.05 (0.028)**	-0.048 (0.022)**
AGE	0.013 (0.004)***	0.014 (0.003)***	0.018 (0.004)***	0.013 (0.003)***
MARITS	0.106 (0.031)***	0.114 (0.026)***	0.074 (0.025)***	0.076 (0.024)**
IHSIZE	-0.044 (0.005)***	-0.057 (0.004)***	-0.081 (0.006)***	-0.079 (0.004)***
HHADULT	0.308 (0.037)***	0.236 (0.032)***	0.219 (0.035)***	0.18 (0.028)***
EDULEVEL	0.049 (0.008)***	0.054 (0.005)***	0.037 (0.009)***	0.038 (0.008)***
EDUYEAR			0.041 (0.01)***	0.03 (0.006)***
EDUYEARSQR			-0.001 (0.001)***	-0.001 (0.0003)**
FARM	-0.238 (0.028)***	-0.165 (0.021)***	-0.166 (0.024)***	-0.07 (0.017)***
HOUSE	0.113 (0.041)***	0.103 (0.038)***	-0.041 (0.055)	-0.005 (0.038)
HEALTH	0.001 (0.001)	0.001 (0.001)	0.0004 (0.001)	0.001 (0.001)
LNLAND	-0.057 (0.03)*	0.082 (0.022)***	0.052 (0.018)***	0.065 (0.011)***
LNPOP	0.311 (0.042)***		0.312 (0.038)***	
LNCAP	0.161 (0.043)***		0.144 (0.024)***	
DIST1			-0.014 (0.032)	
DIST2			0.045 (0.057)	
_cons	5.742 (0.145)***	6.46 (0.077)***	6.165 (0.153)***	7.154 (0.078)***
R-squared	0.2499	0.1043	0.3214	0.2052
N	3839	3839	4143	4143

Note: Dependent variable is the log of real expenditure per capita adjusted by price indexes of regions and months; Variable of interest (LNPCCRF) is the log of (one plus) household per capita formal credit amount obtained; Regression (1,2) & (3, 4) are estimated using full sample data from the VLSS1992/93 and VLSS1997/98, respectively. Regression 2 & 4 introduce commune fixed-effect. Figures in brackets are standard errors which are robust and adjusted for clustering on communes (in regression 1 & 3). R-squared in the FE regressions is the overall R-squared. *, **, and *** indicate significance at the 10%, 5%, and 1% level

Table D.2: Estimation Results: Impacts of Formal Credit - Instrumental Variable Estimates on Cross-Sectional Data

Log per capita expenditure	VLSS1992/93		VLSS1997/98			
	IV (1)	IV (2)	IV (3)	IV (4)	IV (5)	
LNPCCRF	0.142 (0.041)***	0.11 (0.022)***	0.112 (0.022)***	0.11 (0.022)***	0.11 (0.021)***	
GENDER	-0.017 (0.03)	-0.063 (0.034)*	-0.063 (0.034)*	-0.063 (0.034)*	-0.058 (0.033)*	
AGE	0.006 (0.001)***	0.009 (0.001)***	0.009 (0.001)***	0.009 (0.001)***	0.009 (0.001)***	
MARITS	0.118 (0.033)***	0.057 (0.033)*	0.056 (0.033)*	0.057 (0.033)*	0.059 (0.032)*	
HHSIZE	-0.049 (0.006)***	-0.091 (0.007)***	-0.091 (0.007)***	-0.091 (0.007)***	-0.091 (0.007)***	
HHADULT	0.284 (0.04)***	0.162 (0.043)***	0.16 (0.043)***	0.161 (0.043)***	0.154 (0.043)***	
EDULEVEL	0.047 (0.008)***	0.035 (0.011)***	0.035 (0.011)***	0.035 (0.011)***	0.033 (0.011)***	
EDUYEAR		0.031 (0.011)***	0.031 (0.011)***	0.031 (0.011)***	0.026 (0.011)**	
EDUYEARSQR		-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.000 (0.001)	
FARM	-0.238 (0.031)***	-0.209 (0.029)***	-0.209 (0.029)***	-0.209 (0.029)***	-0.206 (0.028)***	
HOUSE	0.094 (0.048)**	-0.132 (0.077)*	-0.133 (0.077)*	-0.132 (0.077)*	-0.123 (0.076)	
HEALTH	0.000 (0.001)	0.000 (0.002)	0.000 (0.002)	0.000 (0.002)	0.000 (0.002)	
LNLAND	-0.097 (0.031)***	0.025 (0.021)	0.025 (0.021)	0.025 (0.021)	0.027 (0.021)	
LNPOP	0.255 (0.042)***	0.232 (0.039)***	0.232 (0.039)***	0.232 (0.039)***	0.23 (0.038)***	
_cons	6.069 (0.124)***	6.782 (0.14)***	6.781 (0.14)***	6.782 (0.14)***	6.818 (0.139)***	
Robust first-stage F-stat	50.44	119.29	160.07	148.76	300.10	
Hansen J-stat (P-val)	0.00	1.506 (0.47)	0.054(0.81)	1.194(0.27)	0.00	
Endog test Chi_2 (P-val)	8.513(0.00)	20.06(0.00)	21.20(0.00)	20.41(0.00)	21.13(0.00)	
LR IV redundancy test		3.62(0.16)	0.435(0.50)	3.142(0.07)		
N	3839	4143	4143	4143	4143	

Note: Instrumental variable for regression (1) is LNCAP using VLSS1992/93; Regression 2-5 using VLSS1997/98. Instrumental variables in these regression are: IV(2): LNCAP, DIST1, and DIST2; IV(3): LNCAP, DIST1; IV(4): LNCAP, DIST2; IV(5): LNCAP. LNCAP instrument is significant at 1 percent level in all first-stage regressions, DIST1 & DIST1 become insignificant at 10 percent level and F-stat is really small (0.94) in the first-stage regression when LNCAP is dropped. Hence result from that regression is not reported here. Redundancy test in IV(2) is for DIST1 and DIST2; IV(3) is for DIST1; IV(4) for DIST2. Standard errors are reported in parentheses which are heteroskedastic-robust and adjusted for clustering on commune. *, **, and *** indicate significance at the 10%, 5%, and 1% level.

Table D.3: Estimation Results: Impacts of Formal Credit - Instrumental Variable Estimates on Longitudinal Data

Log per capita expenditure (1997/98)	OLS (1)	IV (2)	IV-GMM (3)	IV (4)	IV-GMM (5)	IV-GMM (6)
LNPCCRF93	0.006 (0.006)	0.095 (0.04)**	0.07 (0.037)*	0.089 (0.04)**	0.054 (0.039)*	0.113 (0.042)***
GENDER98	-0.055 (0.029)*	-0.073 (0.03)**	-0.064 (0.029)**	-0.072 (0.03)**	-0.065 (0.03)**	-0.077 (0.031)**
AGE98	0.008 (0.001)***	0.007 (0.001)***	0.008 (0.001)***	0.008 (0.001)***	0.007 (0.001)***	0.007 (0.001)***
MARITS98	0.064 (0.026)**	0.087 (0.028)***	0.092 (0.027)***	0.085 (0.029)***	0.086 (0.029)***	0.092 (0.03)***
HHISIZE98	-0.067 (0.007)***	-0.071 (0.007)***	-0.073 (0.006)***	-0.071 (0.007)***	-0.075 (0.007)***	-0.072 (0.007)***
HHADULT98	0.26 (0.038)***	0.22 (0.044)***	0.23 (0.043)***	0.223 (0.044)***	0.223 (0.044)***	0.212 (0.045)***
EDULEVEL98	0.048 (0.011)***	0.046 (0.011)***	0.042 (0.011)***	0.046 (0.011)***	0.044 (0.011)***	0.046 (0.012)***
EDUYEAR98	0.027 (0.011)**	0.021 (0.011)*	0.02 (0.011)*	0.022 (0.011)*	0.016 (0.011)	0.02 (0.012)*
EDUYEARSQR98	-0.001 (0.001)	-0.001 (0.001)	0.000 (0.001)	-0.001 (0.001)	0.000 (0.001)	-0.001 (0.001)
FARM98	-0.163 (0.033)***	-0.168 (0.033)***	-0.155 (0.031)***	-0.168 (0.033)***	-0.134 (0.032)***	-0.17 (0.034)***
HOUSE98	-0.003 (0.049)	-0.044 (0.055)	-0.045 (0.053)	-0.042 (0.055)	-0.019 (0.055)	-0.052 (0.057)
HEALTHI98	0.002 (0.001)	0.002 (0.002)	0.001 (0.001)	0.002 (0.002)	0.001 (0.001)	0.002 (0.002)
LNLAND98	0.045 (0.023)*	0.024 (0.025)	0.034 (0.023)	0.025 (0.025)	0.042 (0.024)*	0.02 (0.025)
LNPOP98	0.207 (0.04)***	0.201 (0.043)***	0.206 (0.042)***	0.201 (0.042)***	0.184 (0.042)***	0.2 (0.044)***
_cons	6.78 (0.129)***	6.843 (0.131)***	6.829 (0.129)***	6.838 (0.131)***	6.878 (0.131)***	6.855 (0.133)***
Robust first-stage F-stat		9.77	9.77	18.32	18.32	44.32
Hansen J-stat (P-val)		26.12(0.00)	26.12(0.00)	17.70(0.00)	17.70(0.00)	0.00
Endog test Chi_2 (P-val)		1.34(0.24)	1.34(0.24)	0.52(0.47)	0.52(0.47)	6.96(0.00)
N		3364	3364	3364	3364	3364

Note: Dependent variable is the log of real per capita consumption in 1997/98. The variable of interest is log of (one plus) per capita formal credit amount in 1992/93. Regression (1) is estimated by OLS. Regression (2) is estimated by two-stage methods using 6 instruments for the credit variable: HHISIZE93, HHADULT93, HEALTHI93, LNLAND93, LNPOP93, and LNCAP93; Regression (4) keeps 3 strongest instruments used in (2), basing on the t-stat in the first-stage regression of (2), including HEALTHI93, LNLAND93, LNCAP93. Regression (6) keeps only LNCAP93 as the instrument. Regression (3), (5) employ GMM estimation to improve efficiency using the same regressors and instruments as in (2) & (4), respectively. Regression (6) is exactly identified hence GMM and IV 2SLS estimators coincide.

As before, standard errors are reported in parentheses which are heteroskedastic-robust and adjusted for clustering on commune. *, **, and *** indicate significance at the 10%, 5%, and 1% level.

Table D.4: Estimation Results: Impacts of Formal Credit - Household Fixed-Effect Estimates

Log per capita expenditure	FE (1)	IV-FE (2)
LNPCCRF	0.025 (0.003)***	0.263 (0.015)***
GENDER	0.000 (0.052)	0.14 (0.077)*
AGE	0.015 (0.001)***	0.012 (0.002)***
MARITS	-0.032 (0.051)	-0.148 (0.076)*
HHSIZE	-0.093 (0.008)***	-0.094 (0.012)***
HHADULT	0.034 (0.051)	-0.054 (0.072)
EDULEVEL	0.171 (0.007)***	0.088 (0.01)***
FARM	-0.189 (0.025)***	-0.093 (0.041)**
HOUSE	-0.082 (0.069)	-0.095 (0.091)
HEALTH	0.002 (0.001)*	0.001 (0.002)
LNLAND	0.114 (0.01)***	-0.003 (0.02)
LNPOP	0.322 (0.055)***	0.354 (0.083)***
R-squared	0.4841	0.3839
Robust first-stage F-stat		439.49
Hansen J-stat (P-val)		0.00
Endog test Chi_2 (P-val)		445.56(0.00)
N	3364	3364

Note: Dependent variable is difference in log of real per capita consumption between 1997/98 and 1992/93. The variable of interest is difference in log of (one plus) per capita formal credit amount. Instrumental variable is difference in log of (one plus) commune credit per capita. Other regressors are also in difference. Standard errors in parentheses are heteroskedastic-robust. *, **, and *** indicate significance at the 10%, 5%, and 1% level.

Appendix E

VLSS Data Definition

Dependent Variables

Credit

CR	Dummy of credit participation regardless of source of credit. CR equals 1 if household borrows
CRF	Dummy of formal credit participation. CRF equals 1 if household borrows from a formal source of credit
CRI	Dummy of informal credit participation. CRI equals 1 if household borrows from an informal source of credit
ACR; ACRF; ACRI	Variables of credit amount obtained by borrowing household from any source; formal source; and informal source, respectively. Unit: '000 VND
PCCRF; PCCRI	Variables of credit amount per household member from formal source; and informal source, respectively. Unit: '000 VND $PCCRF=ACRF/HHSIZE$; $PCCRI=ACRI/HHSIZE$
LNPCCRF	$\log(1+PCCRF)$
LNPCCRI	$\log(1+PCCRI)$

Expenditure

PCEXP	Expenditure per capita of survey household readjusted by price indexes of regions and months (VLSS1992/93: January 1993=1; VLSS1997/98: January 1998=1). This variable equals total expenditure including consumption of food and non-foodstuffs divided to household size. It is comparable between the two surveys.
LNPCEXP	$\log(1+PCEXP)$

Independent Variables

GENDER	Gender of household head
AGE	Age of household head
AGESQR	Square value of age of household head
MARITS	Dummy indicating whether household head is married
HHSIZE	Household size – number of household members
HHADULT	Ratio of number of working adults (age 16-60) and household size
EDULEVEL	Education level of household head; Value ranges from 1-6 including pre-school, primary school, lower secondary school, upper secondary school, vocational training, university, respectively.
EDUYEAR	Number of schooling year of household head
EDUYEARSQR	EDUYEAR squared
FARM	Dummy indicating whether household work in agricultural production
HOUSE	Dummy indicating whether household owns the house
HEALTH	Total number of illness days of household head and his/her spouse
LAND	Area of irrigated land holding size; Unit: m ²
LNLAND	$\ln(1+Land)$
POP	Commune's population; Unit: 1000 people
LNPOP	$\ln(1+POP)$
CAP	Total amount of formal credit borrowed by commune's households divided by commune's population
LNCAP	$\ln(1+CAP)$
DISTANCE	Distance from commune's committee to the nearest government's bank branch
DIST1	Dummy for DISTANCE [0-<4km]
DIST2	Dummy for DISTANCE [4-<10km]
DIST3	Dummy for DISTANCE [>10km]