Lincoln University
Te Whare Wānaka o Aoraki
CHRISTCHURCH • NEW ZEALAND

Lincoln University Digital Thesis

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- you will use the copy only for the purposes of research or private study
- you will recognise the author's right to be identified as the author of the thesis and due acknowledgement will be made to the author where appropriate
- you will obtain the author's permission before publishing any material from the thesis.

RURAL FINANCING IN THAILAND

A thesis
submitted in partial fulfilment
of the requirements for the Degree of
Doctor of Philosophy
in Economics and Finance

At

Lincoln University

By

Visit Limsombunchai

Lincoln University
2006

ABSTRACT

Abstract of a thesis submitted in partial fulfilment of the requirements for
The Ph.D. in Economics and Finance

RURAL FINANCING IN THAILAND

By Visit Limsombunchai

Rural financing in Thailand is heavily dependent on bank lending. Therefore, understanding the determinants of bank lending in the rural sector is an important element for promoting the development of credit accessibility to Thai farmers in the rural regions. Appropriate bank lending decisions would reduce lending costs and increase repayment rate and profits to the banks. Thus, a well-developed rural financial market would lead to sustainable development in the rural sector.

The purpose of this research is to identify critical factors in the bank lending decision and to investigate what factors affect the credit availability and loan price in rural lending in Thailand. This research also investigates the impact of the relationship lending (i.e., the relationship between the bank and the borrower) and the predictive power among the different estimation techniques in predicting the bank lending decision, amount of credit granted, and interest rate charged.

The data used in this research are obtained from the Bank for Agriculture and Agricultural Cooperative (BAAC). During the period of 2001 to 2003, a total of 18,798 credit files under the normal loan scheme are made available. The credit files are analyzed using the logistic regression (Logit), multiple linear regression (MLR), and four different types of the artificial neural networks (ANN), namely multi-layer feed-forward neural networks (MLFN), Ward networks (WD), general regression neural networks (GRNN), and probabilistic neural networks (PNN).

The results show that the total asset value (*Asset*), value of collateral (*Collateral*), and the length of the bank-borrower relationship (*Duration*) are crucial factors in determining bank lending decision, amount of credit granted, and interest rate charged. As expected, *Asset* has a positive impact on the bank lending decision and the amount of credit granted, while

Collateral has a positive and a negative influence on the amount of credit granted and the interest rate charged, respectively. However, *Collateral* has no significant impact on the bank lending decision, while *Asset* has a significant negative impact on the loan price in some specifications.

Duration has a significant negative impact on bank lending decision, amount of credit granted, and interest rate charged, which implies the importance of relationship lending in the Thailand rural financial market. However, the negative relationships between *Duration* and the bank lending decision, and between *Duration* and the amount of credit granted, contradict the postulated hypothesizes. The results imply that the bank uses information from the borrowers and monitors the lending risk via the lending decision and amount of credit granted. On the other hand, the relationship lending benefits the borrowers via loan pricing since the borrowers with a long term relationship with the bank receive a lower lending rate.

The predictive results of both in-sample and out-of-sample on bank lending decision, amount of credit granted, and interest rate charged show that in terms of predictive accuracy, most of the artificial neural networks models outperform the logistic and the multiple regression models. The empirical results also show the superiority of using the PNN model to classify and screen the loan applications, and the GRNN model to determine the amount of credit granted and interest rate charged.

Keywords: *Thailand, Rural financing, Lending decision model, Credit availability model, Loan pricing model, Logistic regression, Multiple linear regression, and Artificial neural networks.*

ACKNOWLEDGEMENTS

Writing this research is a difficult and long journey. The completion of this research is impossible without help and support of the following persons and organizations:

- Assoc. Prof. Dr. Christopher Gan and Assoc. Prof. Dr. Minsoo Lee, my principal supervisor and associate supervisor, respectively. My sincere thanks to both of them for their patience, dedication, support, encouragement, insightful comments, and invaluable suggestions for improvement on my work.

- To Mr. Thiraphong Tangthirasunan (The president of BAAC), I am indebted to you for the accessibility to the BAAC credit files. I am also thankful to BAAC staffs (Mr. Niluck Limlaemthong, Ms. Sunee Paiboon, Ms. Kanittha Kitsiwavast, and Ms. Sunisa Mingkwan) for their technical assistance and support in preparing the data set.

- I am thankful to Lincoln University in awarding me the Lincoln University Doctoral Scholarship to complete my PhD program.

- I would like to extend my sincere gratitude and appreciation to lecturers who allowed me to audit their subjects. These include: Dr. Christopher Gan (ECON 615: Applied Research Methods, and ECON 614: Commercial Banking), Dr. Minsoo Lee (QMET 609: Business Statistics and Econometrics, and ECON 621: Macroeconomic Analysis), Dr. Lana Friesen (ECON 605: Microeconomic Analysis), Mr. Bert Ward (ECON 609: Quantitative Economic Analysis), Dr. Sandhya Samarasinghe (COMP 627: Neural Networks Applications), Ms. Raylene Pierce (FINC 608: Foundations of Finance Theory), and Dr. Gerry Nartea (FINC 604: Finance Futures and Options).

- To Mr. Mike Clemes, I am thankful for your kindness and help on editing my thesis.

- To Ms. Annette Brixton, Ms. Dawn Dooling, Ms. Kristene Broad, and Ms. Deborah O'Connor, thank you very much for your kindness and assistance in many ways.

- To the Department of Agricultural and Resources Economics, Faculty of Economics, Kasetsart University, Thailand – my office. Thank you for granting me three years of study leave and providing some financial support to complete my PhD.

- To Assist. Prof. Orrachorn Sriswasdilek, I would like to acknowledge with much appreciation to your major role in helping me during the data collection period. Without your assistance, I would not have the data set for this research.

- To all my friends and colleagues at Lincoln and Thailand (my sincere apologies as there is not enough space to put all your names). Thank you all for the direct and indirect support in many ways. I appreciate them very much.

- Last, but not least, my family. I am grateful to my parents for their financial and moral support. Without their support and encouragement, I might not be able to complete my research and degree.

CONTENTS

Page

Page

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Thailand is an agricultural country, but industrialization has surpassed agriculture activities as evident in the past few decades. The country is made up of approximately 320 million rais (1 rai = 0.16 ha or 0.395 acre) and approximately 131 million rais (or about 40 percent) is considered as farmland and about 80 million rais (or about 25 percent) is forest land (see Table 1.1). During 1999-2000, 50 percent (about 65 million rais) of the total cultivated land were used to produce rice, 22 percent (about 28 million rais) for cash crops and 20 percent (about 26 million rais) for fruits (Office of Agricultural Economics, 2003).

The majority of the Thai population (about 55 percent or 35 million people, in 2002) resides in the rural areas and is involved in the agricultural sector. For example, in 1980, the employment in the agricultural sector accounted for 73 percent of the total employment, which is one of the highest proportions found among developing countries (see Table 1.2) (Krongkaew, 1995). The employment in this sector has dropped to 42 percent of the total employment in 2002 but it still absorbed more than two-fifth of the total labor force. Thus the agricultural sector still plays a major role in the Thai economy in terms of employment and production.

However, the agricultural sector in Thailand has shrunk over the past two decades, in line with the increased investments in the industrial, manufacturing and service sectors. For example, in 1988, the Gross Domestic Product (GDP) share of the agricultural sector (crops, livestock, forestry and fisheries) at current market prices was 16 percent, while the service and industrial sectors were 52 percent and 28 percent, respectively (see Table 1.3). The agriculture sector's share of the GDP was further reduced to 10 percent in 1993, while the service and industrial sectors' share of the GDP increased to 53 percent and 30 percent respectively. Furthermore its share of the total GDP declined to 10 percent in 2001, but the sector's total value increased from 252 million baht in 1988 to 320 and 532 million baht in 1993 and 2001, respectively.

Thailand agriculture sector generates foreign exchange earnings to the country. The share of agricultural exports in terms of the total value of exports is above 50 percent until 1988. For example, in 1987 more than half of the foreign exchange earnings came from agricultural products.

Table 1.1: Land utilization in Thailand

Year	Total area (Rai)	Forestry		Farmland		Unclassified land	
		Rai	%	Rai	%	Rai	%
1988	320,696,888	89,880,182	28.03	131,772,759	41.09	99,043,947	30.88
1989	320,696,888	89,635,625	27.95	131,831,185	41.11	99,230,078	30.94
1990	320,696,888	87,488,536	27.28	131,124,409	40.89	102,083,943	31.83
1991	320,696,888	85,436,284	26.64	133,076,188	41.50	102,184,416	31.86
1992	320,696,888	84,344,169	26.30	132,051,209	41.18	104,301,510	32.52
1993	320,696,888	83,450,623	26.02	131,207,893	40.91	106,038,372	33.06
1994	320,696,888	83,801,555	26.13	131,833,288	41.11	105,062,045	32.76
1995	320,696,888	82,178,161	25.62	132,478,570	41.31	106,040,157	33.07
1996	320,696,888	81,808,415	25.51	131,819,506	41.10	107,068,967	33.39
1997	320,696,888	81,441,164	25.40	131,107,608	40.88	108,148,116	33.72
1998	320,696,888	81,076,428	25.28	130,393,525	40.66	109,226,935	34.06
1999	320,696,888	80,610,219	25.14	131,341,384	40.95	108,745,285	33.91

Source: Department of Agricultural Economics (OAE.) (Various years)

Table 1.2: Population, employment and unemployment rate of Thailand

Year	Population (M. people)	Population		Employment		Unemployment (%)
		Agri.(%)	Non-agri.(%)	Agri.(%)	Non-agri.(%)	
1980	n.a.	n.a.	n.a.	72.50	27.50	n.a.
1990	54.55	66.80	33.20	66.50	33.50	n.a.
1996	60.12	n.a.	n.a.	45.29	54.71	3.55
1997	60.82	n.a.	n.a.	45.05	54.95	3.22
1998	61.47	n.a.	n.a.	44.68	55.32	7.27
1999	61.66	n.a.	n.a.	45.27	54.73	6.28
2000	61.88	56.20	43.80	44.39	55.61	5.81
2001	62.31	n.a.	n.a.	42.24	57.76	5.15
2002	62.80	54.85	45.15	41.64	58.36	3.65

Note: n.a. – not available

Source: National Statistic Office (NSO.) and Office of Agricultural Economics (OAE.) (Various years)

Table 1.3: Value of gross domestic product of Thailand by sector at current market prices (Million baht)

Year	Gross Domestic Product				
	Agriculture	Industry	Construction	Service	Total
1988	252.35	429.63	74.45	803.38	1,559.80
	16.18%	27.54%	4.77%	51.50%	100.00%
1989	279.95	528.60	102.12	946.32	1,856.99
	15.08%	28.47%	5.50%	50.96%	100.00%
1990	272.94	628.84	136.24	1,145.54	2,183.55
	12.50%	28.80%	6.24%	52.46%	100.00%
1991	317.09	747.27	168.28	1,274.00	2,506.64
	12.65%	29.81%	6.71%	50.83%	100.00%
1992	348.13	821.29	190.53	1,470.97	2,830.91
	12.30%	29.01%	6.73%	51.96%	100.00%
1993	320.05	936.62	220.77	1,687.78	3,165.22
	10.11%	29.59%	6.97%	53.32%	100.00%
1994	383.20	1,067.67	267.80	1,910.67	3,629.34
	10.56%	29.42%	7.38%	52.65%	100.00%
1995	458.98	1,240.78	302.64	2,183.82	4,186.21
	10.96%	29.64%	7.23%	52.17%	100.00%
1996	505.03	1,366.94	341.52	2,397.56	4,611.04
	10.95%	29.64%	7.41%	52.00%	100.00%
1997	513.99	1,443.24	271.82	2,503.55	4,732.61
	10.86%	30.50%	5.74%	52.90%	100.00%
1998	564.88	1,446.35	178.68	2,436.54	4,626.45
	12.21%	31.26%	3.86%	52.67%	100.00%
1999	502.83	1,534.07	166.25	2,433.93	4,637.08
	10.84%	33.08%	3.59%	52.49%	100.00%
2000	510.99	1,693.55	150.07	2,561.90	4,916.51
	10.39%	34.45%	3.05%	52.11%	100.00%
2001	532.08	1,764.95	152.36	2,674.02	5,123.42
	10.39%	34.45%	2.97%	52.19%	100.00%

Source: Bank of Thailand (BOT.) (Various years)

The proportion of agricultural exports has declined while the proportion of non-agricultural (industrial and service sectors) exports has increased as Thailand experienced an industrialization process since 1980s (see Table 1.4) (Krongkaew, 1995). The agricultural export share fell to 23 percent in 2001, but its export value experienced an increase from 194,198 million baht in 1988 to 677,893 million baht in 2001. This indicates that the values of agricultural export have been growing but its growth rates have been dominated by the growth rate in non-agriculture export. The agricultural trade balance has never been in deficit unlike the other sectors. The trade surplus in agricultural sector has tripled from 115,980 million baht in 1988 to 361,025 million baht in 2001.

The agriculture sector, associated with the rural sector, has contributed considerably to the growth of Thai economy in many ways. It has not only provided the supply of food and inputs, but has also been a source of employment and foreign exchange earnings. The important role of an agriculture sector in Thailand should be promoted by the policy makers and the industrialization process should not be promoted at the expense of the agriculture sector because the development of the agriculture sector complements the process of the industrialization.

Table 1.4: Value of exports, import and trade balance of Thailand (Millions baht)

Year	Export			Import			Trade balance	
	Total	Agriculture	%	Total	Agriculture	%	Total	Agriculture
1987	299,853	153,991	51.36	334,340	53,556	16.02	-34,487	100,435
1988	403,570	194,198	48.12	513,114	78,218	15.24	-109,544	115,980
1989	516,315	230,537	44.65	662,679	102,244	15.43	-146,364	128,293
1990	589,818	224,168	38.01	852,962	125,710	14.74	-263,144	98,458
1991	725,449	256,038	35.29	956,408	142,869	14.94	-230,959	113,169
1992	824,643	285,264	34.59	1,033,245	158,454	15.34	-208,602	126,810
1993	940,862	279,857	29.74	1,170,846	159,889	13.66	-229,984	119,968
1994	1,137,601	336,290	29.56	1,369,034	179,857	13.14	-231,433	156,433
1995	1,406,310	407,218	28.96	1,834,537	213,538	11.64	-428,227	193,680
1996	1,411,039	412,677	29.25	1,832,825	216,833	11.83	-421,786	195,844
1997	1,806,932	485,198	26.85	1,924,263	228,831	11.89	-117,331	256,367
1998	2,248,777	591,690	26.31	1,774,050	226,827	12.79	474,727	364,863
1999	2,214,249	556,498	25.13	1,907,391	228,097	11.96	306,858	328,401
2000	2,768,064	626,911	22.65	2,494,133	275,459	11.04	273,931	351,452
2001	2,893,176	677,893	23.43	2,756,656	316,868	11.49	136,520	361,025

Source: Department of Agricultural Economics (OAE.) (Various years)

1.2 Agricultural credit, rural finance, and debt burden

Agricultural credit is very important and necessary for the survival of the agricultural sector in Thailand. Farms in Thailand are generally under capitalized[1], where a majority of

[1] The capital-labour ratio of Thailand agricultural sector is lower than the other major Southeast Asian counties, such as Malaysia, Indonesia, and Philippine (Asian Development Bank, 2002).

the farmers are small-scale peasant farmers who are poor and lack of investment funds (Asian Development Bank, 2002). According to the National Statistic Office's 1999 socio-economic survey, poverty is mostly concentrated among farm households. For example, 5.3 million households or 54 percent are poor and about 4.2 million households or 72 percent of the country's ultra-poor (i.e. those below 80 percent of the poverty line[2]) are small farm owners or farm workers. Therefore, they borrow heavily to finance their farm productions, investment and private consumption. Table 1.5 shows the increasing trends of household outstanding debt and the average outstanding debt in Thailand agriculture sector.

The proportion of agricultural households with an outstanding debt has monotonically increased from 25 percent in 1988 to 34 percent in 1991 and to 60 percent in 1999 (see Table 1.5). The average outstanding debt has also risen from 6,047 baht per household in 1988 to 12,772 and 37,231 baht per household in 1991 and 1999, respectively. The average loan size in each year has increased from 3,831 baht per household in 1971 to 15,049 baht per household in 1995 and then 18,493 baht per household in 1999. The average loan size has increased nearly 500 percent over the last 28 years. The large portion of loan was primarily for agricultural productions (about 70 percent on the average) (see Table 1.6).

Table 1.5: Number of households, households with outstanding debt, and the average outstanding debt in agricultural sector of Thailand

Year	No. of household	No. of household with outstanding debt	%	Average outstanding debt (Baht/household)
1988	5,040,132	1,279,000	25.38	6,046.78
1990	5,073,471	1,408,000	27.75	7,828.94
1991	5,130,531	1,729,831	33.72	12,771.74
1995	5,502,782	2,857,993	51.94	24,672.13
1998	5,513,855	3,050,412	55.32	37,019.35
1999	5,642,890	3,379,163	59.88	37,231.00

Source: Office of Agricultural Economics (OAE.) (Various years)

The credit market in the rural areas is characterized by state-owned financial institutions, such as Bank for Agriculture and Agricultural Cooperative (BAAC), Government Saving Bank (GSB), and Government Housing Bank (GHB), private commercial banks,

[2] The poverty line in Thailand was equivalent to an average of 878 Baht/person/month in 1998.

cooperatives, informal institutions, such as production credit groups, credit unions, savings groups, village funds, and traditional informal sources such as relatives, neighbors, and money lenders (Poramacom, 2001).

Over the last 20 years, institutional credit has shown its growing role gradually replacing the informal credit market. The proportion of institutional credit has increased from 64 percent in 1971 to 91 percent in 1995, but declined to 84 percent in 1999 caused by the 1997 financial crisis (see Table 1.6). BAAC provides financial assistance to farmer, farmer association and/or agricultural co-operative and is the only formal credit source with a high significant share in Thailand agricultural credit market. In 1999, 62 percent of the total loan in agricultural sector came from BAAC, whereas, 13 percent came from agricultural co-operatives and 8 percent from commercial banks (see Table 1.7).

Table 1.6: Average loan size, sources of loan and borrowing purposes

Year	Average Loan size (Baht/household)	Sources of loan (%)		Borrowing Purposes (%)	
		Informal	Institution	Agri.	Non-agri.
1971	3,830.98	36.35	63.65	70.30	29.70
1976	2,187.07	36.97	63.03	79.44	20.56
1978	3,053.63	36.12	63.88	73.67	26.33
1980	4,360.63	42.10	57.90	76.53	23.47
1982	4,788.86	33.81	66.19	74.33	25.67
1986	3,206.43	29.55	70.45	76.49	23.51
1988	5,137.27	28.12	71.88	81.49	18.51
1990	6,759.27	17.95	82.05	77.99	22.01
1991	8,924.59	18.97	81.03	81.27	18.73
1995	15,048.83	9.01	90.99	70.39	29.61
1998	17,854.48	16.96	83.04	69.03	30.97
1999	18,493.14	16.10	83.90	67.30	32.70

Source: Office of Agricultural Economics (OAE.) (Various years)

BAAC's outstanding credit increased more than 9 times from 1987 to 1999. The outstanding credit in 1987 was 25 billion baht (35 percent) compared to 176 billion baht (53 percent) in 1996 and 229 billion baht (63 percent) in 1999. On the other hand, the proportion of outstanding credit relative to the total commercial banks debt has decreased from 64 percent in 1987 to 37 percent in 1999 (see Table 1.8). This demonstrates the important role of BAAC in Thailand's agricultural credit market relative to other formal credit institutions.

Table 1.7: Sources of loan in 1996 and 1999

Item	1996	1999
Informal:	***15.40***	***16.10***
Relatives	3.10	3.80
Neighbours	2.20	2.50
Landlord/money lenders	2.90	6.80
Rice mill owner	0.20	1.20
Farmer saving groups	1.40	0.90
Others	5.70	0.90
Institutional:	***84.60***	***83.90***
BAAC	59.70	61.80
Agricultural co-operatives	12.30	12.60
Commercial banks	12.30	8.30
Finance companies	0.30	1.20

Source: Office of Agricultural Economics (OAE) (1996, 1999)

Table 1.8: Outstanding credit classified by types of financial institutions

Year	Commercial banks[1]		BAAC		Finance companies		Total	
	B. Baht	%	B. Baht	%	B. Baht	%	B. Baht	%
1987	45.80	63.79	25.10	34.96	0.90	1.25	71.80	100.00
1992	131.00	65.43	65.20	32.57	4.00	2.00	200.20	100.00
1996	159.20	47.49	176.00	52.51	n.a.	n.a.	335.20	100.00
1997	156.60	44.92	192.00	55.08	n.a.	n.a.	348.60	100.00
1998	148.30	41.62	208.00	58.38	n.a.	n.a.	356.30	100.00
1999	133.80	36.93	228.50	63.07	n.a.	n.a.	362.30	100.00

Note: 1/ Including the Government Saving Bank (GSB) and Government Housing Bank (GHB).
n.a. – not available

Source: Bank of Thailand (BOT) (Various years)

BAAC has classified agricultural loans into the following categories (Natetayaluck, 2001a):

1. Short-term loans for agricultural production (0 – 12 months).

   This loan is generally used for annual inputs required in the production of farm crops and livestock such as purchasing seed, feed, fertilizer, pay for operating expenses, etc. In some cases, the loan period can be extended up to 18 months.

2. Loans for postponement of farm production sale (0 - 6 months).

   The purpose of this loan is to assist the farmers who want to store their products temporary and sell them at higher prices later, after the harvesting season.

3. Medium-term loans (1 – 3 years).

   This type of loan has to be repaid, between 1 to 3 years except for some special cases where the duration of loan may be increased to 5 years. The loans are mostly for land improvements, purchasing machinery and livestock.

4. Cash Credit loans.

   This type of credit allows the farmer to borrow extra cash up to a prespecified limit. Normally, the contract term of cash credit loans is not longer than 5 years.

5. Long-term loans for refinancing old debts.

   The main objective of the loan is for reimbursement of previous debts, redemption or repurchase of agricultural lands that belong to the farmer themselves or their parent, spouse or child. This type of loan prevents the farmers from losing their land to informal credit lenders. The pay back period of the loan is not longer than 10 years.

6. Long-term loans for agricultural investment.

   This type of loan is typically used for tree crops, purchasing a farm or additional lands, financing buildings and other permanent or long-life improvements. The loan contract term is between 15 – 20 years and may include a repaying grace period for the borrowers.

7. Loans for other occupations related to the agriculture.

   The loan may be utilized for operating expenses and/or investment in an agricultural business, e.g. procurement inputs, food processing, trading, etc.

The above 7 types of credit can be further classified into 3 categories according to the term of the loan:

1. Short-term credit: production credit with a repayment period of less than 1 year.
2. Medium-term credit or intermediate-term credit: repayment period is between 1 – 5 years.
3. Long-term credit: real-estate or long-term investment credit with a repayment period more than 5 years, but typically is not longer than 15 years.

The short-term loan in 1999 accounted for 60 percent of the total loan while the proportion of medium-term and long-term loans were about 22 percent and 18 percent, respectively (Agricultural Statistics of Thailand Crop Year 1999-2000). These figures indicated that most farmers placed a high demand on short-term credit rather than medium to long-term investment credits. The high demand for short-term credit was caused by the run-up in input prices, especially gasoline price during the gulf war, and the cumulative impact of low commodity prices problem.

The major problems facing Thailand's rural finance are similar to those experienced by most developing countries. These include (Davis et al., 1998; Thailand Development Research Institute, 1998):

1. Indebtedness. The outstanding debt of the agricultural household has continuously increased over time (see Table 1.5). This problem is related to the operational problems as farmers always have inadequate cash flows and profits to service loan repayments.
2. High interest rate charged. The interest rate charged on the agricultural loans is usually higher than traditional bank rates due to the higher risk and uncertainty on the agriculture productions.
3. Lack of collateral. Because of this problem, farmers may not have access to the credit, or may be charged a high interest rate to substitute for a high credit risk. This problem is partly due to slower progress made in reforming property rights and land title.
4. Limited sources of funds. Commercial banks and other financial institutions are generally wary of the involvement with rural lending, since they have to deal with higher risk. As a result, farmers and agricultural processors are mainly confined to dealing with BAAC.
5. Short of credit services. Financial institutions prefer to deal with large enterprises. They have been slow to adjust themselves to deal with small scale lending, such as small-scale processor, traders, and private farmers.
6. Lack of financial management knowledge and skills. About 60 percent of agricultural households incurred some form of debts (see Table 1.5), but it is unreasonable to expect farmers to have adequate financial management knowledge and skills in dealing with bankers or other financial institutions.

1.3 The impact of the 1997 financial crisis

In 1997, the Asian financial crisis which was caused by excessive short-term private borrowings and a series of misguided policies resulted in a severe deflation and contraction of the Thai economy. The deterioration triggered by a currency speculation, the fragility of financial institutions and capital outflow negatively affected both the financial and the real economic sectors. The baht devaluation and the sudden withdrawal of foreign funds led to a strict lending policy, financial distress in many enterprises and falling output, which resulted in a decrease demand for labor and massive layoffs. Consequently, the unemployment rate increased from 3 percent in the first quarter of 1997 to around 5 percent and 6 percent in the first quarter of 1998 and 1999, respectively. The sectors hit hardest included the financial services, property, and part of business services and manufacturing that served the domestic market (Bhipatboonthong, 2002). On the financial sector itself, a total of 56 financial firms were closed in late 1997. Furthermore, in 1998, six banks and 12 other financial companies were brought under the control of the Bank of Thailand.

The financial crisis resulted in an overall negative economic growth rate for Thailand. The bulk of the impact of the financial crisis was felt in 1998 when the overall economic growth rate dropped from –1 percent in 1997 to –11 percent in 1998. The preponderance of the decline in economic activity included the construction, industrial and service sectors, where their growth rates were –38 percent, –11 percent and –9 percent in 1998, respectively (see Table 1.9).

Table 1.9: Thailand's economic growth by sector at 1988 prices

Year	Growth rate				
	Agriculture	Industry	Construction	Service	Total
1994	4.95%	9.41%	14.15%	9.03%	8.99%
1995	3.45%	11.93%	6.72%	9.14%	9.24%
1996	4.14%	7.29%	7.05%	5.22%	5.90%
1997	-0.90%	2.20%	-25.64%	-0.78%	-1.37%
1998	-1.52%	-10.91%	-38.25%	-9.49%	-10.51%
1999	2.17%	12.09%	-6.84%	0.53%	4.45%
2000	6.44%	5.97%	-9.54%	4.08%	4.65%
2001	2.68%	1.48%	-0.95%	2.27%	1.94%

Source: Bank of Thailand (BOT) (Various years)

Despite favorable terms of trade in export arising from the baht devaluation and the improvement in prices for most agricultural products, productions and exports did not immediately improve, because of the El Nino effect in 1997. The growth rate of agricultural sector was –2 percent in 1998. However, the agricultural sector had the minimal impact from the financial crisis compared to the manufacturing and service sectors. As a result, the agricultural sector played a greater role in promoting economic growth during the financial crisis. The agriculture sector had a crisis carrying capacity and generated new employment for the unemployed urban.

An attempt of the Thai government to disentangle the effects from the economic and financial crisis was to strengthen the rural sector. The government strategic and policy frameworks for the rural development in Thailand include (Natetayaluck, 2001b):

I. Off-farm employment and rural enterprises must be expanded by:
    1. Creating the enabling environment for rural enterprise growth.
    2. Expanding rural credit for establishment of Small and Medium Enterprise (SMEs) and micro-enterprises.
    3. Increase private sector participation and development and Civil Society Organization (CSO) partnerships in development of off-farm employment.
    4. Strengthening vocational schooling and skills development in rural areas.
    5. Improving the incentive and regulatory framework for efficient intermediation.
    6. Reviewing the institutional framework for industrial development.

II. Rural finance markets must be strengthened by:
    1. Transforming of the Bank for Agriculture and Agricultural Cooperative (BAAC) into an independent commercial rural bank.
    2. Expanding micro finances for small farmers and poor households.
    3. Improving the policy framework for efficient intermediation.
    4. Reducing government intervention in rural banking.
    5. Improving the rural financial sectors training and supervision programmes.

However, the 1997 financial crisis has also changed the relative importance of loan sources in the agriculture sector. For example, the borrowing share of the informal sector has increased slightly in 1999 from 1996 (see Table 1.7). The proportion of loans from landlords and moneylenders has increased from 3 percent in 1996 to 7 percent in 1999. In contrast, the share of commercial banks loans has reduced from 12 percent in 1996 to 8 percent in 1999, with some of this gap partly filled by BAAC.

In addition, the data on repayment rate of BAAC's individual clients showed a drop in repayments of long term and medium-term loan after the crisis (see Table 1.10). The repayment rate was healthy during the early 1990s where short- and medium-term loans including cash credit line had an average repayment rate of about 85 percent while long-term loan averaged almost 80 percent. From 1995 to 1996, long term refinancing credit repayment rate depreciated by 23 percent points and dropped to 24 percent in 1999. While the repayment rate of short-term credit was 87 percent during the financial crisis year, it dropped by 6 percent and then increased to 83 percent in 1999. The medium-term credit repayment rate continued to decline significantly, and long-term investment credit also dropped significantly in 1998.

Default repayment took place before the financial crisis and the deterioration on the debt repayment was further impacted by the financial crisis. This gave rise to the question on bank rational lending decision-making.

Table 1.10: BAAC repayment rates by individual clients (%)

Loan category	Average 1992 - 1994	Average 1995 - 1996	1997	1998	1999
Short-term	90.70	88.95	86.80	81.30	83.10
Medium-term	86.63	84.75	72.30	67.90	64.40
Cash credit lines	88.03	86.75	82.50	74.90	76.30
Long-term(refinancing)	78.47	55.20	41.00	27.60	23.90
Long-term(investment)	81.07	78.55	70.20	52.20	59.20

Source: Bank for Agriculture and Agriculture Co-operatives (BAAC) (Various years)

1.4 Research objectives

The Thai rural sector is dependent heavily on bank lending. However, most commercial banks prefer to lend to large commercial farmers than to rural households. The commercial farmers normally acquire large loans with long repayment periods, which are more profitable for the banks, while rural households tend to acquire small loans that incur high transaction and administration costs with a high default risk. As a result, some rural households are denied access to formal credit and have to resort to informal borrowing with very high interest rates.

Understanding the capital market system and bank lending determinants in rural areas help to improve the development of credit accessibility to farmers in rural regions. The rationality of banks lending decisions includes reducing lending costs and increasing repayment rate and operating profits to the banks. As a consequence, a well-developed rural financial market would promote economic growth and development in the rural sector.

This research examines the determinants of bank lending decision, credit availability, and loan price in Thailand's rural financial market, and to provide policy makers recommendations to increase farmers' access to credit and to improve Thailand's rural financial system.

The objectives of this research include:

1. To provide an overview of the rural financial system in Thailand.
2. To identify critical factors in the lending decision process of the Bank for Agriculture and Agricultural Cooperative (BAAC) for the rural household (from the BAAC's perspective).
3. To investigate factors affecting the amount of credit granted and interest rate charged in rural lending.
4. To examine the impact of the relationship between lender and borrower on the lending decision, credit availability and credit price.

The research also examines the different bank lending behavior between agricultural lending and non-agricultural lending, and compares the results from different estimation techniques, such as multiple regression, logit regression, and artificial neural networks

1.5 Contribution of the research

This research is expected to contribute to the development of Thailand's rural financial market in addition to analyzing the rational behavior of bank lending processes. Since limited research has been conducted on rural finance regarding the lending process, an understanding of the rationale of the bank lending decision-making would improve the efficiency of Thailand rural financial markets and the rural sector development. Furthermore, the expansion of the commercial bank lending to the rural sector would be an effective way towards poverty alleviation and improving the quality life of the rural households.

In addition, this research would benefit both lenders and borrowers. Instead of using subjective evaluation decision-rules, which are bias and unreliable, lenders can apply an objective evaluation technique with a standard process and criteria to appraise their customer's credit risks and creditworthiness. A good credit risk assessment promotes a healthy credit market including competitive price of credit. Furthermore, it provides support to the lenders on credit risk management, reduction of default risk, increase in repayment rate and profits.

Borrowers can evaluate whether they qualify for new loans or an extension of existing loans. They will be able to estimate the credit availability and the price of credit corresponding to their risk level. This information would enhance the borrowers' decision-making process when they acquire loans. Self-assessment by the borrowers is also supportive to the credit suppliers as it may reduce the number of applications for credit from ineligible candidates and processing and administrative costs.

1.6 Data and methodology

This research uses the data and information from Thai bank credit files. All Thai banks have been approached to participate in this research. All participants who provided their customer credit files for this research were kept in strict confidence.

The data includes the loans granted in 1999 – 2003. As suggested by Heckman (1979), to avoid two types of estimation bias, choice and selection bias, which typically plague lending decision models, the credit files were be selected at random.

The analysis is divided into the following:

1. Descriptive analysis:

   Frequency table, graph, average, and percentage are used to answer Research Objective 1.

2. Quantitative analysis:

   To examine Research Objectives 2 and 4, a credit scoring model is developed via logistic regression and artificial neural networks (ANN) technique.

To examine Research Objectives 3 and 4, multiple regression and artificial neural networks (ANN) technique are employed to assess the credit availability and loan pricing models, and to verify the impact of relationship lending.

1.7 Outline of this thesis

The rest of this thesis is organized as follows. Chapter 2 presents an overview of the relevant literature and the theoretical background on credit scoring, the demand for credit and credit availability models, interest rate and loan pricing models, and relationship lending. Chapter 3 describes the empirical models, the estimation techniques, the data, and the data collection method. Chapter 4 presents the results, and discussions of the results generated by the analysis. The prediction capability among the different estimation techniques is also compared. Chapter 5 summarizes the major findings and implications, followed by the limitation of the research and recommendations for future research.

CHAPTER 2
LITERATURE REVIEW

This chapter provides an overview on bank lending decision, credit scoring model, demand for credit, and cost of credit. The chapter is organized as follows. Section 2.1 discusses the bank lending decision and creditworthiness evaluation techniques. Section 2.2 discusses the concept of credit scoring and explains how credit scoring works. The benefits and limitations of credit scoring are discussed in Section 2.3. The variables and modeling techniques that are commonly used in credit scoring models are summarized in Sections 2.4 and 2.5, respectively. Section 2.6 discusses the demand for credit and credit availability model, and Section 2.7 discusses about interest rate and loan pricing model. Section 2.8 reviews previous studies on relationship lending and the impact of relationship lending on credit availability and credit price.

2.1 Banks' lending decision: judgmental versus credit scoring

Generally, a bank lending decision depends upon the borrower credit risk, that is the probability of the borrower not repaying the loan. Credit analysis includes the valuation of the financial history and financial statements of the applicant credit background. It is the primary method used in appraising credit risks. The objectives of credit analysis are to determine the financial strength of the borrowers, to estimate the borrower's probability of repayment, and to reduce the risk of nonpayment to an acceptable level. There are two major problems in credit analysis: the assessment of all important factors about an applicant simultaneously and the evaluation of all applicants objectively (Plata and Nartea, 1998; Sinkey, 2002).

There are two main techniques that can be applied to evaluate a borrower creditworthiness (Crook, 1996):

1. Loan officer subjective assessment (judgmental technique) and
2. Credit scoring technique.

Creditworthiness is the characteristics of an individual that makes him or her qualify for a loan while someone who is not creditworthy is unqualified for the loan (Lewis, 1992). The subjective assessment of the borrower creditworthiness is normally based on the 6Cs' - Character, Capacity, Cash, Collateral, Conditions and Control (Rose, 1993) (see Table 2.1).

However, Lewis (1992), Crook (1996), and Glassman and Wilkins (1997) argue that the judgmental assessment technique seems to be inefficient, unexplainable, inconsistent and non-uniform. Thus, credit scoring models have become more preferable technique for creditworthiness and credit risk appraisal.

2.2 What is credit scoring?

Credit scoring, introduced in the 1950s, is a method in evaluating the credit risk of loan applications. Using historical data and statistical techniques, credit scoring tries to segregate the effect of various applicant characteristics on delinquencies and defaults (Turvey and Brown, 1990; Mester, 1997; Glassman and Wilkins, 1997; Frame et al., 2001). A surge in interest about credit scoring models has been motivated primarily by two requirements of major lending institutions over the last few years: emphasizing on increased efficiencies in processing of loan applications and prudent risk management (Glassman and Wilkins, 1997).

Credit scoring is broadly applied in consumer lending, especially in credit cards, and has been used in mortgage lending recently. Credit scoring has not been widely used in business lending because business loans differ substantially across borrowers, making it more difficult to build up an accurate method of scoring. However, this is changing. The complexity and flexibility of statistical models and advanced computing technology have made such scoring possible. Thus, many banks are using credit scoring to evaluate loan applications, which is a cost effective credit management tool (Mester, 1997).

Table 2.1: The six basic Cs' in lending

Character	Capacity	Cash	Collateral	Conditions	Control
• Customer past payment record. • Experience of other lender with current customer. • Purpose of loan Customer track record in forecasting. • Credit rating. • Presence of cosigners or guarantors of the proposed loan.	• Identity of customer and guarantors. • Copies of charters resolutions, agreements, and other documents bearing on the legal standing of the borrowing customer. • Description of history, legal structure, owners, natural of operations, productions, and principal customers and suppliers for a business borrower.	• Past earnings, dividends, and sales record. • Adequacy of projected cash flow. • Availability of liquid reserves. • Turnover of payables, receivables, and inventory. • Capital structure and leverage. • Expense controls. • Coverage ratios. • Recent performance of borrower stock and P/E ratio. • Management quality. • Content of auditor report and statement footnotes. • Recent accounting changes.	• Ownership of assets • Ages of assets. • Vulnerability to obsolescence. • Liquidation value. • Degree of specialization in assets. • Liens, encumbrances and restrictions. • Leases and mortgages issued. • Insurance covered. • Guarantees and warranties issued. • Bank relative position as creditor. • Lawsuits and tax situation. • Probable future financing needs.	• Customer position in industry and expected market share. • Customer performance vis-à-vis comparable firms in industry. • Competitive climate for customer product. • Sensitivity of customer and industry to business cycles and changes in technology. • Labor market conditions. • Impact of inflation on customer balance sheet and cash flow. • Long-run industry outlook. • Regulation; political and environmental factors.	• Applicable banking laws and regulations regarding the character and quality of acceptable loans. • Adequate documentation for examiners. • Signed acknowledgments and correctly prepared loan documents. • Consistency of loan request with bank written loan policy. • Inputs from noncredit personnel (such as economists or political experts) on external factors affecting loan repayment.

Source: Rose, 1993: p. 195

The overall idea of credit scoring models is quite straightforward. A large historical loans sample on the similar loan type is divided into those that paid and those that defaulted. Based on statistical probabilities, the combination of borrower characteristics that differentiate "good" from "bad" loans generate a score, which is an estimate of the riskiness of each new loan (Crook, 1996). Based on the score, banks or lenders can rank their loan applications or borrowers in term of risk, and then decide whether to make loans and how to price them. Crook (1996) argues that the aim of credit scoring is to predict risk, not to explain it. Therefore, it is not necessary that the predictive model also explains why some borrowers default on the loan repayment and others do not.

Credit history information and other data regarding repayment ability, which are generally provided by borrower, is analyzed (often electronically via a computer). The model then attempts to predict the applicant's likelihood of default based on previous experience with borrowers of similar loan profiles. A well-designed model should give high scores to borrowers whose loans will perform well and low scores to borrowers whose loans will not perform well. In some systems, the score is compared with a certain critical value (cut-off point) and the result is either accept or reject decision. To develop a good credit scoring model, sufficient historical data is needed to reflect loan performance in all economic conditions. However, there is no perfect scoring model. It may be possible that some bad borrowers may get a high score and receive the loans, and vise versa (McAllister and Mingo, 1994).

2.3 Benefits and limitations of credit scoring

Credit scoring techniques have a number of benefits compared to judgmental techniques for both lenders and borrowers. The perceived benefits of credit scoring that have led to its increasing use in loan assessment include (Chandler and Coffman, 1979; Lewis, 1992; Crook, 1996; Glassman and Wilkins, 1997; Mester, 1997):

1. Increase efficiencies and reduce costs.

   Traditional loan approval process usually takes up to at least two weeks. Credit scoring can reduce this process to days or hours, and fewer loan officers can handle a large number of applications. Credit scoring can increase efficiency by leaving loan officers to concentrate only on ambiguous cases. This means cost saving to the bank and benefits to the borrowers. The borrowers need to provide only the information used in the model and the application process becomes shorter.

2. More objective, reduce unfair lending practices and personal bias.

   Credit scoring is an objective process that can help lenders to ensure that the same decision criteria is applied to all borrowers, regardless of race, gender, or other factors prohibited by commercial law from being used in credit appraisal. The models are built on larger samples, and consider the characteristics of both good and bad borrowers, whereas judgmental methods are usually negatively biased towards perceptions of bad borrowers only.

3. Ability to control risk levels.

   For lenders, it is easier to control the number of new loans granted in high risk categories when credit scoring is used compared to judgmental method. Credit analyst can estimate credit risk of the borrowers with reasonable degree of confidence and can raise or drop the cut-off score to adjust the risk of the loan portfolio.

4. Continual learning system.

   Credit scoring models are based on statistical techniques. By repeatedly re-estimating the models with a larger data set, the systems can learn over time and can predict the next borrower behavior more accurately. With the recent development of neural networks – "self-learning" computer programs modeled on the human mind – the level that the credit scoring process can develop in terms of complexity and accuracy is unknown. Certainly, continual learning system will make the models more efficient in a short period.

According to Feldman (1997), credit scoring can alter small-business lending in three areas:

1. The interaction between borrowers and lenders.

   Credit scoring allows lenders to underwrite and monitor loans without actually meeting the borrowers. This development is in stark contrast to the perceived importance of a local bank-borrower relationship. In fact, because of the scoring systems, borrowers can obtain unsecured credit from distant lenders through direct marketing channels.

2. Loan pricing.

The price of small-business loan should decline, especially for high credit quality borrowers who will no longer have to bear the cost of extensive underwriting. In addition, increased competition, resulting from small businesses having access to more lenders should further lower borrowing costs.

3. Credit availability.

   Credit scoring should increase credit availability for small businesses. Better information about the repayment prospects of a small-business applicant makes it more likely that a lender will provide a larger loan amount and the loan price will be based on expected risk.

While the benefits of credit scoring are fairly well known, the limitations in using credit scoring should not be ignored. The weaknesses of credit scoring models include (Capon, 1982; Crook, 1996; Glassman and Wilkins, 1997; Mester, 1997):

1. Data and the accuracy of the models:

   Credit scoring models are extremely complex, and they are only good if data is available. Inaccurate and insufficient credit report information can invalidate the credit scoring results. The data on which the system is fed need an adequate sample of both well-performing and poorly performing loans. The data should be up to date and the models should be re-estimated frequently to ensure that changes in the relationship between potential factors and loan performance are captured.

2. Knowing the customers:

   The use of credit scoring models can not substitute the value gained by knowing the customers personally, although the models attempt to mimic actual behavior of borrowers. This is especially important when dealing with applications and customers who have not had a pristine credit history or where the ability to repay a loan may be the primary factor on which a credit decision is based. In such cases, credit scoring may not be able to capture the true repayment likelihood of the applicants.

3. Economic fluctuations:

   Because credit scoring models are based on historical data, the models are susceptible to biases due to the timeframe of the data and business cycle. A good credit scoring model needs to make accurate predictions in both good and bad economic situations. Therefore, the data on which the model is based should address both expansions and recessions. If the data on which the model is used does not include repayment behavior in an economic downturn, and if adjustments are not made, lenders may have to face with a higher risk than they have planned.

4. Selection bias:

   Generally, only accepted applicants information is used to estimate the coefficients in credit scoring models. Thus, the selection bias may lead to bias on estimated weights in credit scoring models.

2.4 Variables commonly used in credit scoring models

The pragmatism and empiricism of credit scoring implies that any characteristic and environment of the borrower that has obvious connections with default risk should be used in the scoring system (Lewis, 1992). Lewis (1992) suggests that there is no need to justify the case for any variable. If it helps the predictions, it should be used. However, some characteristics should not be used in the credit scoring models because they are discriminatory and legally banned, such as race, religion, and gender, and some of them are culturally unacceptable, such as health and conviction records.

However, the major factors commonly used in credit scoring models include the borrowers' liquidity, profitability, solvency (or leverage), efficiency and repayment capacity (Turvey and Brown, 1990; Turvey and Weersink, 1997; Novak and LaDue, 1999; Barney et al., 1999).

Liquidity reflects the capacity of borrower to generate cash to meet its short-term financial obligations. The variables commonly used to measure liquidity are (Lee et al. 1988; Rose, 1993):

1. $$\text{Current ratio} = \frac{\text{Current assets}}{\text{Current liabilities}}$$

2. $$\text{Acid-test ratio or quick ratio} = \frac{\text{Current assets} - \text{Inventories}}{\text{Current liabilities}}$$

3. Net Working capital = Current assets – Current liabilities

Low ratio on current and acid-test ratios and low net working capital indicate illiquidity and the borrower may not be able to generate sufficient funds to meet fixed financial obligations. Thus, default risk is expected to increase as liquidity decreases.

Profitability refers to efficiency of the borrower's activities and their ability to generate profit. Common measures of profitability include the following (Lee et. al., 1988; Barney et al., 1999):

1. $\text{Return on assets} = \dfrac{\text{Net return}}{\text{Total assets}}$

2. $\text{Return on equity} = \dfrac{\text{Net return}}{\text{Equity}}$

Low profitability ratios may imply misuse of resources, which may increase the probability of loan default. Thus, borrowers with high profitability ratios are relatively preferred.

Solvency (or leverage) indicates the amount of debt that a borrower has taken in addition to the loan being applied for. Key financial ratios used to analyze solvency include (Lee et al. 1988; Rose, 1993):

1. $\text{Leverage ratio} = \dfrac{\text{Total liabilities}}{\text{Total assets}}$

2. $\text{Debt-to-equity ratio} = \dfrac{\text{Total liabilities}}{\text{Equity}}$

The above ratios measure the overall financial position of the borrower because they reflect the likelihood that the sale of all assets should produce sufficient cash to cover all debt outstanding and indicate how much debt financing is used compared to equity financing. The higher ratios mean high debt financing, high financial risk and less likely additional loans will be granted.

The measure of efficiency includes gross ratio and capital turnover ratio. They are used to indicate the input-output efficiency of the business and the efficiency with which capital is being employed in the business, respectively (Lee et al. 1988; Turvey and Brown, 1990; Barney et al., 1999).

1. $$\text{Gross ratio} = \frac{\text{Total expenses}}{\text{Gross income}}$$

2. $$\text{Capital turnover ratio} = \frac{\text{Gross income}}{\text{Total assets}}$$

A high value of gross ratio (lower value of capital turnover ratio) implies inefficiencies in input use or production (capital utilization), and it is expected that the default risk increases as the gross ratio increases (capital turnover ratio decreases).

Repayment ability measures the borrower's ability to successfully meet principal and interest commitments from future cash flow. Repayment ability is measured by three variables (Rose, 1993; Barney et al., 1999):

1. $$\text{Interest expense ratio} = \frac{\text{Interest payment}}{\text{Total income}}$$

2. $$\text{Interest coverage ratio} = \frac{\text{Earning before tax and interest}}{\text{Interest payment}}$$

3. $$\text{Debt repayment ratio} = \frac{\text{Total debt and interest payment}}{\text{Total income}}$$

Higher ratios on both interest expense ratio and debt repayment ratio are expected to increase the probability of loan default.

The above ratios can be easily calculated from the borrower's financial statements. Thus, lenders always use these financial ratios in combination with other factors such as the borrower's personal attributes in the credit appraisal. However, the selection of financial ratios across studies is inconsistent. Lufburrow et al. (1984) used liquidity, solvency, collateral, repayment ability and repayment history when they estimated credit scoring for a production credit association in Illinois. Fischer and Moore (1986) utilized profitability, solvency, and efficiency with credit scoring function for the St. Paul Bank for Cooperative to assess the borrower's credit risk. Mortensen et. al. (1988) tried to predict the probability of the farmers' loan default in North Dakota with only two ratios, solvency and efficiency (see Table 2.2).

Miller and LaDue (1989) used combinations of profitability, solvency, and efficiency in their credit assessment models for a bank case focusing on dairy farms in New York. In addition, Turvey and Brown (1990) used liquidity, profitability, solvency, efficiency, and repayment capacity in estimating the credit scoring for Canada's farm credit corporation, while Barney et al. (1999) used them to estimate the farm debt failure prediction model for the farmers home administration in the United States. (see Table 2.2).

Table 2.2: Financial ratios commonly used in credit scoring models

	Liquidity	Solvency	Profitability	Efficiency	Repayment Capability	Others
Lufburrow et al. (1984)	✓	✓			✓	
Fischer and Moore (1986)		✓	✓	✓		
Mortensen et. al. (1988)		✓		✓		
Barry and Ellinger (1989)	✓	✓	✓	✓	✓	
Miller and LaDue (1989)		✓	✓	✓		
Turvey and Brown (1990)	✓	✓	✓	✓	✓	✓1/, 2/
Turvey and Weersink (1997)	✓	✓	✓		✓	✓1/
Novak and LaDue (1999)	✓	✓			✓	
Barney et al. (1999)	✓	✓	✓	✓	✓	✓2/
Wu and Wang (2000)	✓	✓	✓	✓	✓	✓

Note: 1/ Dummy variables on region and farm type.
2/ Dummy variables on loan for refinance or restructuring.

According to the literature, not all financial ratios have a significant impact on the probability of loan default. For example, Turvey and Brown (1990) reported that only current ratio and return on assets had a negative impact on the borrower's default risk, while leverage ratio had a positive impact on the default risk. Novak and LaDue (1999) and Turvey and Weersink (1997) found a positive relationship between leverage ratio and the probability of loan default. Furthermore, Wu and Wang (2000) indicated that debt-to-equity ratio was positively related to the borrower's default risk. The authors also found that capital turnover ratio had a negative relationship with the probability to default on the loan repayment, which was inconsistent with the financial theory. Table 2.3 summarizes the empirical relationships between the financial ratios and the probability of loan default.

Table 2.3: Empirical relationships between the financial ratios and the probability of loan default

Variable	Probability of loan default	Source
Liquidity		
- Current ratio	Negative	Turvey and Brown (1990)
- Quick ratio	Negative	Turvey and Weersink (1997)
- Net working capital	n.s.	n.a.
Profitability		
- Return on assets	Negative	Turvey and Brown (1990) Turvey and Weersink (1997)
- Return on equity	n.s.	n.a.
Solvency		
- Leverage ratio	Positive	Turvey and Brown (1990) Turvey and Weersink (1997) Novak and LaDue (1999)
- Debt-to-equity ratio	Positive	Wu and Wang (2000)
Efficiency		
- Gross ratio	n.s.	n.a.
- Capital turnover ratio	Positive	Wu and Wang (2000)
Repayment ability		
- Interest expense ratio	n.s.	n.a.
- Interest coverage ratio	n.s.	n.a.
- Debt repayment ratio	n.s.	n.a.

Note: n.s. = no significant impact found
n.a. = not available

2.5 Credit scoring and modeling techniques

Several statistical methods have been used to estimate credit scoring models in assessing agricultural credits, such as discriminant analysis (Dunn and Frey, 1976), linear probability models (Turvey, 1991), logit models (Mortensen et al., 1988), and probit models (Lufburrow et al., 1984). The last three methods are standard statistical techniques for estimating the probability of default based on historical data on loan performances and characteristics of the borrowers. The linear probability model assumes that the probability of default and the factors are linear in relationship. The logit model assumes that the probability of default is logistically distributed. The probit model assumes that the probability of default follows the standard cumulative normal distribution function and discriminant analysis divides borrowers into high and low default risk groups (Mester, 1997).

Historically, discriminant analysis has been popular. However, there are questions concerning the econometric properties since the technique is neither unbiased nor a consistent estimator, and most of the exogenous variables used in the model generally violate the normal distribution assumptions (Collins and Green, 1982; Ladue, 1989).

Collins and Green (1982) pointed out that the linear probability model could present reasonable prediction results compared to discriminant analysis and logit models. However, Judge et al. (1985), Greene (1997) and Pyndick and Rubinfeld (1998) indicated that the predictive value of linear probability models may not necessarily lie between zero and one, which violates the probability theory. Moreover, the variance of the models is generally heteroscedasticity, which lead to inconsistent estimation problem and invalid conventional measure of fit such as the R^2.

The logit and probit models are very similar to each other. Both of them provide asymptotically consistent, efficient and unbiased estimates. The logit model is generally preferred to the probit model because of its simplicity. Recently, the logit model has dominated the agricultural credit scoring literature (Barney et al., 1999; Novak and LaDue, 1999; Lee and Jung, 1999) (see Table 2.4).

Table 2.4: Credit scoring techniques

	DA.	LPM.	Logit	Probit	RPA.	ANN.
Dunn and Frey (1976)	✓					
Lufburrow et al. (1984)				✓		
Mortensen et. al. (1988)			✓			
Miller and LaDue (1989)			✓			
Turvey and Brown (1990)			✓			
Turvey (1991)	✓	✓	✓	✓		
Jensen (1992)						✓
Altman et al. (1994)	✓					✓
Turvey and Weersink (1997)			✓			
Novak and LaDue (1999)			✓		✓	
Lee and Jung (1999)			✓			✓
Barney et al. (1999)		✓	✓			✓
Wu and Wang (2000)	✓					✓

Note: DA. = Discriminant Analysis, LPM. = Linear Probability Model, Logit. = Logistic Model, Probit. = Probit Model, RPA. = Recursive Partitioning Algorithms, ANN. = Artificial Neural Networks.

Turvey (1991) empirically compared agriculture credit scoring models using four parametric methods with a single date set. The author recommended the logistic model over the probit model, linear probability model and discriminant analysis based on predictive power and ease of use, in addition to the consistent statistical property.

However, Novak and LaDue (1999) argued that the problems in a logit model include:

1. The need to pre select the exact explanatory variables without well-developed theory;
2. Inability to identify an individual variable's relative importance;
3. Reduction of the information space dimensionality; and
4. Limited ability to incorporate relative misclassification costs.

Thus, recursive partitioning algorithm (nonparametric classification method) was introduced as an alternative technique for credit scoring analysis, which allows direct incorporation of misclassification costs. However, recursive partitioning algorithm outperformed the logistic regression based on within-sample observation only, while the logistic regression is superior to recursive partitioning algorithm based on more appropriate out-of-sample observations.

Artificial neural network (ANN), a new classification technique, is artificial intelligent algorithms that allow for some learning process through experience to discern the relationship between the borrower characteristics and the probability of default. Since there are no assumptions about functional form, or about the distributions of the variables and errors of the model, ANN is more flexible than the standard statistical technique (Mester, 1997). It also allows for nonlinear relationship and complex classificatory equations. The user does not need to specify in detail about the functional form before estimating the classification equations, instead it lets the data determine the appropriate functional forms. The final equations may be very complex to explain or understand (Crook, 1996).

Altman et al. (1994) applied the linear discriminant analysis and the neural network techniques to analyze over 1,000 healthy, vulnerable and unsound Italian industrial firms from 1982 – 1992. The authors found that artificial neural network and discriminant analysis yielded the same degree of accuracy. Furthermore, Lee and Jung (1999) utilized artificial neural network and logistic regression with a total of 21,678 credit files from several credit unions in South Korea. The authors found that both models were very powerful for prediction delinquency. However, the logistic regression outperformed the neural network for the urban accounts, while the neural network outperformed the logistic regression for the rural accounts.

Barney et al. (1999) applied accounting data contained in the Farm and Home Plan (FHP) in predicting farm debt failure in the United States. The authors' predictive results indicated that the neural network model outperforms both the linear probability model and the logit model in predicting farm debt failure based on an error rate. In addition, Wu and Wang (2000) applied the neural network method to small business lending decisions in central New York. The authors found that the neural network has a stronger discriminating power for classifying the acceptance and rejection group than traditional parametric and non-parametric classifiers.

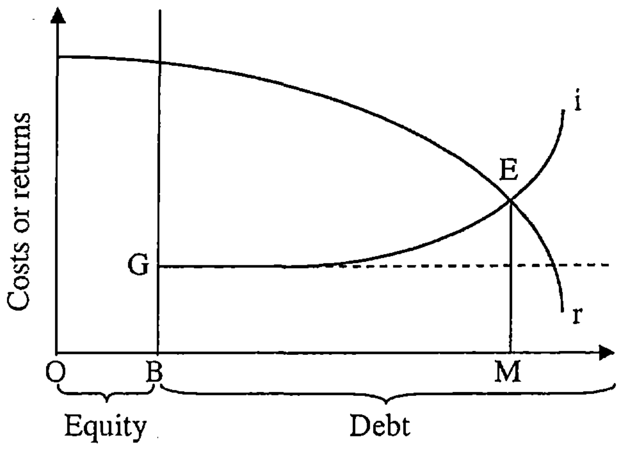
According to empirical studies, there is no unanimous agreement as to the best method for estimating credit scoring models and new methods continue to evolve. However, the logit models and artificial neural networks have been applied frequently in previous research.

2.6 Demand for credit and credit availability model

Demand for credit is a derived demand. Households or firms desire credit in order to make certain production and consumption expenditures as well as investments (Feder et al., 1993). To determine the optimal level of credit utilization, the costs and returns of credit must be considered.

The law of diminishing marginal return indicates that the return (i.e. marginal value products) from additional units of resources and resource services acquired with borrowed funds will decline at an accelerating rate, as indicated by curve r in Figure 2.1. Curve i represents the marginal cost of borrowing or interest rate and is upward sloping. This is because the borrowing cost increases when the leverage increases. The assumption that the borrowing cost is constant is unrealistic. The extremely high degree of credit use is disfavour for the lender, as it incurs cost in the form of increased financial risk. Thus, the lender might react by increasing the interest rate on the loan, or simply refusing to provide additional credit (Lee et al., 1988).

Figure 2.1: Equilibrium in credit utilization



Conceptually, the gain from the use of credit is maximized when the marginal rate of return r is equal to the marginal cost of borrowing i (see Figure 2.1). Thus, the optimal level of credit use is BM, where the curve r intersects the curve i at E. As long as the marginal rate of return exceeds the marginal cost of using a loan, the borrower will increase the level of credit use (Barry et al., 1995). The optimal credit use would increase the borrower's income. If some of this income is reinvested, saved, or used to repay debts, the wealth (asset) of the borrower will increase. Therefore, the credit utilization and the process of investing back a portion of the earnings result in the growth of the borrower's wealth.

However, it is possible that the financial institutions may limit the credit extended to the borrower, which is known as credit rationing. As a result, the amount of credit provided by the institutional lender might not meet with the borrower's demand for credit to maximize the borrower's gain. Thus, the borrower might decide to borrow from more expensive sources, such as non-financial institution lenders, as long as the return from credit use still exceed the cost of fund from those sources (see Figure 2.2). The borrowers who have an unlimited access to the institutional credit will not borrow from the more expensive source. Therefore, the amount of credit borrowing from a more expensive source can be used to measure the degree to which the borrowers are supply constrained by the financial institutions (Petersen and Rajan, 1994).

Figure 2.2: Sources and uses of funds

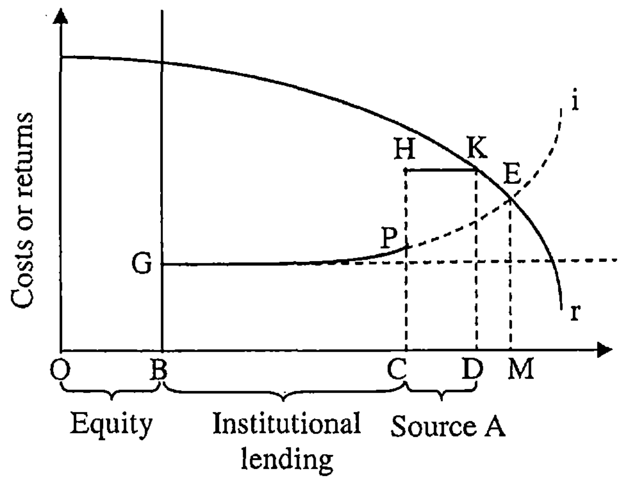


Figure 2.2 shows the sources and the uses of funds. Assuming that the financial institutions ration the amount of credit they offer to the borrower to the amount of BC. Since the borrower's demand for credit is BM, the borrower might want to borrow from the alternative source of funds, for instance Source A. If HK is the marginal cost of borrowing from Source A, the demand for credit of the borrower is only BD, where the marginal rate of return is equal to the marginal cost of borrowing. As a result, the borrower borrows only BC from the financial institution and borrows CD from Source A[3].

Since the availability of credit to the borrower is simultaneously determined by the borrower demand for credit and the supply of credit, it is difficult to measure credit availability directly. Changes in the level of credit use might be due to changes in demand for credit (assuming that the supply curve is observed), or by changes in supply of credit (assuming that the demand curve is observed). Although the amount of credit provided by the lender can be used as a proxy, it may underestimate the credit available to the borrower, since the borrower may receive a small amount of loan because the borrower is liquidity constrained, or because the borrower has little need for external funds (Petersen and Rajan, 1994). However, it can be expected that credit availability for a good borrower

[3] For CD to be an appropriate measure of institutional credit rationing, the marginal cost of borrowing from the alternative source must exceed the marginal cost of available institutional credit. If this is not true, the amount CD will be a function of the price financial institutions charge, as opposed to the volume of credit they are willing to offer (Petersen and Rajan, 1994).

should be greater than for a bad borrower. Consistent with this reasoning, the borrower who has a high asset value, high average earning, and low earning volatility would have greater access to credit. Furthermore, the borrower with a high education, high work experience, and high profitability would be granted a larger loan.

Bard et al. (2000) argue that borrower, loan, and lender characteristics, and financial market structure may influence the lenders' decision on the amount of the loan. Borrower characteristics that signal credit risk may affect the loan amount. Loan traits affecting the amount of loan include purpose of the loan and loan-to-value ratio. Bank attributes such as lending focus, lending polices, lending limits, reserve requirements, and available of funds are supply-side factors that could affect the availability of credit. Market structure features, such as market share and concentration, may also influence the lenders' lending polices. Therefore, the credit availability model could be expressed as follows:

$$A_i = f\left(B_i, L_i, C_i\right) \qquad (2.1)$$

where A_i is the loan amount for loan i;

B_i is a vector of borrower and loan characteristics believed to influence the loan amount;

L_i is a vector of bank characteristics that may affect the loan size;

C_i is the market structure variable hypothesised to influence loan supply.

2.7 Interest rate and loan pricing model

In general, interest rate is considered to be comprised of four components (Goodwin and Mishra, 2000):

1. a return to productive capital,
2. an adjustment reflecting a positive rate of time preference,
3. a premium for expected inflation, and
4. a default risk premium.

The first three components are expected to be identical across the borrowers and lenders for a given size and term of the loan. Thus, the interest rates differences across the identical loan contracts can only be explained by the differences in the default risk premiums. If the competition among the lenders in the lending market is sufficient, it will abolish the non-competitive loan pricing practices and the interest rates should equalize across the lenders. As a result, the differences in the interest rates are represented by the financial risk of the borrowers.

However, the competition among lenders may not be strong enough to fully eliminate the differences in interest rates across the lenders. Moreover, it is likely that different groups of lenders (for example, commercial banks and governmental banks) do not usually compete against one another for the same pool of loans. Different lenders appear to have different cost structures and different lending practices, which may be another important factor explaining the difference in the interest rates charged.

The market interest rate is determined by the interaction of the borrowers' demand for loans and the supply of loanable funds. The change in inflation and the financial market condition has an influence on the interest rate. Thus, the interest rate must be conditioned on the time of loan commencing. The supply of loanable funds depends on the current interest rate in the market, the cost of credit and other operating costs, the perceived riskiness of the loans relative to the alternative investment options available to the lenders, and the demand for loanable funds in the economy. The demand for loanable funds largely depends upon the current, and the future expectation, of the health of the economy. Since many loans have the contract term extended over several years, the expectations of future financial market conditions also play an important role in shaping the demand and supply of loanable funds.

Interest rate can be viewed as the loan price from the interaction of the demand and supply for loanable funds. When the individual loans are considered at a point in time, the factors related to the expectations of the future economy and the future financial market conditions are expected to be homogeneous across the competing lenders and borrowers. Therefore, the differences of the loan prices would be determined by the characteristics of the borrowers, the characteristics of the individual loans, and the likelihood that the loan will be repaid or default.

In general, the interest rate charged by a lender must cover the cost of funds, loan administration and service costs, a default risk premium and a profit for the lender (Lee et al., 1988). A simple loan pricing model can be expressed as follows (Rose, 1993):

$$\begin{array}{c}\text{Loan}\\\text{interest}\\\text{rate}\end{array} = \begin{array}{c}\text{Marginal cost}\\\text{of raising}\\\text{loanable}\\\text{funds to lend}\\\text{to the}\\\text{borrower}\end{array} + \begin{array}{c}\text{Non-funds}\\\text{operating}\\\text{costs}\end{array} + \begin{array}{c}\text{Estimated}\\\text{margin to}\\\text{protect the}\\\text{lender}\\\text{against}\\\text{default risk}\end{array} + \begin{array}{c}\text{Lender's}\\\text{desired}\\\text{profit}\\\text{margin}\end{array} \qquad (2.2)$$

The loan pricing model (equation 2.2) can be formally expressed in general form as follows (Bard et al., 2000):

$$R_i = f\left(B_i, L_i, C_i\right) \qquad (2.3)$$

where R_i is the interest rate for loan i;

B_i is a vector of borrower and loan characteristics that may influence credit risk (e.g. financial performance, production efficiency, risk-management ability, experience, and loan size) (Miller et al., 1993);

L_i is a vector of bank characteristics thought to influence the interest rate (e.g. lending costs, desired rate of return, and other lending policies);

C_i represents the market structure characteristic that may influence pricing behaviour (e.g. market concentration or market share).

2.8 Relationship lending: impact on credit availability and loan price

The issue of credit availability and the interest rate charged to borrowers, especially to small borrowers, have been widely discussed because many borrowers can not get enough credit at reasonable interest rates, even if they have an opportunity to invest in positive net present value projects. Stiglitz and Weiss (1981) suggest that the capital market frictions such as information asymmetries and agency costs may explain why credit does not always flow to borrowers with profitable investment opportunity.

Stiglitz and Weiss (1981) argue that the capital market differs from most markets because prices and interest rates do not always adjust to clear the market. The authors show that interest rate charged determines not only the demand for capital but also the riskiness of the borrowers. A higher interest rate either draws riskier applicants (the adverse selection effect) or influences borrowers to choose riskier investment projects (the incentive or moral hazard effect). Thus, lenders may decide to ration the quantity of loans rather than raise the interest rate to clear the market if an increase in the rate increases the average riskiness of borrowers.

To overcome asymmetric information problems, relationship lending (relationship building between the lender and the borrower) is one of the most reasonable strategies because it allows the lender to gather relevant information about the prospects and creditworthiness of the borrower over a considerable time period. The continuous contact between borrower and lender in the provision of various financial services can produce valuable input for the lender in making decisions. For example, banks (or lenders) can acquire information by monitoring the borrower performance over time under certain credit arrangements and/or through the provision of other services, such as deposit account or compensating balances. Banks may utilize this information in their decisions about the credit extension and the term of credit, such as interest rate charged, required collateral from the borrower, and attached other conditions to the loan. The borrower with close relation to the banks should have a lower cost of capital, lower required collateral, and greater availability of funds compared to a borrower without such relation (Petersen and Rajan, 1994; Berger and Udell, 2002).

Boot (2000) outlines five potential benefits in relationship lending. First, relationship lending facilitates information reuse through time, which encourages information production and monitoring by the lenders. Second, relationship lending facilitates flexible and implicit long-term contract. Third, loan contracts typically include covenants to mitigate agency costs that become suboptimal as new information arrives. Depending on the relative bargaining strengths, the development of long-term relationship facilitates low cost of renegotiation of the covenants. Fourth, relationship lending often involves collateral or personal guarantees, and long-term contractual relationship encourage monitoring and the efficient use of costly collateral. Fifth, relationship lending accommodates the inter-temporal smoothing of loan terms, which benefits young and informational opaque firms.

In recent years, there are a number of empirical studies that have investigated on relationship lending. For example, Berger and Udell (1995), Blackwell and Winters (1997), Athavale and Edmister (1999), and Bodenhorn (2003) found evidence of an inverse relationship between the bank-borrower relationship and the loan rate (see Table 2.5). On the other hand, Greenbaum et al. (1989), Sharpe (1990), Angelini et al. (1998), and Degryse and Cayseele (1998) suggested that borrowers with existing relationship were charged higher loan rates. Greenbaum et al. (1989) argued that the incumbent banks can charge higher loan rates due to their information advantage relative to their competitors. Degryse and Cayseele (2000) found offsetting relationship effects. The authors argued that on one hand, the loan rate increases with the length of the bank-borrower relationship. On the other hand, loan rates decline with the scope of the bank-borrower relationship, which is defined as the purchase of other financial services from the bank.

Table 2.5: The length of bank-borrower relationship and availability and cost of credit

	Availability of credit	Cost of credit
Greenbaum et al. (1989)		Positive
Sharpe (1990)		Positive
Petersen and Rajan (1994)	Positive	No
Berger and Udell (1995)	Positive	Negative
Blackwell and Winters (1997)		Negative
Cole (1998)	Positive	
Angelini et al. (1998)	Positive	Positive
Elsas and Krahnen (1998)		No
Harhoff and Korting (1998)	Positive	No
Athavale and Edmister (1999)		Negative
Degryse and Cayseele (2000)	No	Positive
Bodenhorn (2003)		Negative

Petersen and Rajan (1994) examined the effects of the banking relationship on the availability of credit to small firms and on the pricing of the credit. The authors found a positive relationship between the availability of credit and the length of time the firm associated with the financial institution. In contrast, they could not find any significant relationship between the duration of the lending and the loan price to support the hypothesis that relationship lending reduces loan rates. Thus, they conclude that the banking relationship affect the availability of credit more than the credit price.

Berger and Udell (1995) argued that relationship lending is less important in asset-based lending, such as mortgages and term loans. The authors contended that lines of credit, which tend to be cash-flow-based loans, are relationship loans. To avoid diluting the effect of the relationship lending, they excluded transaction-based loans and used only lines of credit data, unlike Petersen and Rajan's (1994) study. The authors found that the borrowers with longer banking relationships paid lower interest rate. However, the relationship was significant only in the sub sample of loans greater than $500,000 in total assets. The authors also found that the borrowers with longer banking relationships received larger loans and were less likely to pledge collateral.

Blackwell and Winters (1997) conducted a study on banking relationships and the effect of monitoring on loan pricing. The authors used 174 lines of credit data from 6 banks from two holding companies. Using regression analysis similar to that of Petersen and Rajan (1994) and Berger and Udell (1995), they observed a positive relationship between the loan's interest rate and the bank's monitoring effort. Their findings also showed that banks less frequently reviewed firms with whom they had longer lending relationships, and ultimately, charged lower interest rates (banks pass along the monitoring cost saving to the borrower in the form of lower interest rates).

Athavale and Edmister (1999) indicated that banks obtain private information about their customers, as well as monitoring borrowers, and have an information advantage in the production of other services. The authors results showed that relationship influenced the price of credit such that subsequent loans were priced significantly lower than prior loans to reflect the benefits of lower monitoring costs arising out of the relationship. Thus, they concluded that borrowers received some positive values from continuing the banking relationship.

Bodenhorn (2003) used the contract-specific loan records of a 19th-century U.S. bank to analyze the value of firm-bank relationships. The author determined that that small firms excluded from arm's-length markets found it advantageous to form extensive and durable relationships with banks. The author's results showed that repeat borrowing over long periods lead to lower interest costs, lower guarantee usage, and a greater likelihood of maintaining a banking relationship during financial panics and other macroeconomic downturns. Therefore, the author concluded that extended banking relationship was, indeed, valuable.

There are many studies that broadly confirm the importance of relationship lending on the credit availability (see Table 2.5). For example, Angelini et al. (1998), and Harhoff and Korting (1998) showed that the credit availability for small firms typically increased with the length of the bank-borrower relationship. In addition, Elsas and Krahnen (1998) found that banks continued to lend to the customers who had a good relationship with the banks in spite of a deterioration in the customer's credit rating.

The major factors included in the loan pricing and the credit availability models of a firm are (Petersen and Rajan, 1994; Berger and Udell, 1995; Blackwell and Winters, 1997; Keasey and Watson, 2000) (see Table 2.6):

1. Firm characteristics.

   These include firm size (total assets), and firm age.

2. Credit risk proxies.

   Key financial ratios, including current ratio, return on assets, leverage ratio, capital turnover ratio and interest coverage ratio, risk level, and collateral requirement dummies are conventionally utilized as credit risk proxies to control for the observable risk of the borrower.

3. Relationship indicators.

   Six variables have been used as relationship indicators in the previous literature. These include: duration of the relationship, borrowing concentration ratio, number of banks from which the firm borrows, housebank (major financial source) status, dummy variables on deposit accounts with and other services from current lender.

4. Dummy variables.

   These include bank, industry, region, loan type, loan size, and lending year dummies.

According to the literatures, the firms' size and age played an important role on both price and availability of credit. The results suggest that larger and older firms received more credit and pay lower interest than smaller and younger firms (Petersen and Rajan, 1994; Blackwell and Winters, 1997; Goodwin and Mishra, 2000). Furthermore, secured loans usually carried higher interest rates than unsecured loans, as riskier borrowers needed to provide collateral (Berger and Udell, 1995; Blackwell and Winters, 1997; Strahan, 1999).

Quick ratio, return on assets, capital turnover ratio, and interest coverage ratio were positively related to the credit availability, but negatively related to the loan price (Berger and Udell, 1995; Strahan, 1999), while leverage ratio was positively related to the loan price (Strahan, 1999; Keasey and Watson, 2000).

For relationship indicators, duration of the relationship, borrowing concentration ratio, housebank status, and the number of banks from which the firm borrows are considered as the key factors in determining the loan price and the credit availability. Long-term relationship and borrowing from only one or a few banks reduce the interest rate and increase the amount of credit (see Petersen and Rajan, 1994; Blackwell and Winters, 1997, Degryse and Cayseele, 2000). In addition, long-term loans and large loans are cheaper than the other loan types and sizes (see Keasey and Watson, 2000; Bodenhorn, 2003; Degryse and Ongena, 2005).

Table 2.6: Factors influencing the credit availability and cost of credit

Factors	Credit availability	Cost of credit
Firm characteristics		
- Firm size (total assets)	Positive • Petersen and Rajan (1994); Strahan (1999).	Negative • Petersen and Rajan (1994); Blackwell and Winters (1997); Strahan (1999); Goodwin and Mishra (2000).
- Firm age	Positive • Petersen and Rajan (1994); Angelini et al. (1998); Harhoff and Korting (1998); Akhavein et al. (2004).	Negative • Petersen and Rajan (1994); Blackwell and Winters (1997); Harhoff and Korting (1998); Strahan (1999); Degryse and Cayseele (2000); Goodwin and Mishra (2000); Keasey and Watson (2000).
Credit risk proxies		
- Quick ratio	Positive • Strahan (1999).	Negative • Berger and Udell (1995); Strahan (1999).
- Return on assets	Positive • Strahan (1999).	Negative • Strahan (1999).
- Leverage ratio		Positive • Keasey and Watson (2000).
- Capital turnover ratio	Positive • Strahan (1999).	Negative • Strahan (1999).
- Interest coverage ratio	Positive • Strahan (1999).	Negative • Strahan (1999).

Table 2.6: Factors influencing the credit availability and cost of credit (Cont)

Factors	Credit availability	Cost of credit
- Collateral requirement		Negative • Degryse and Cayseele (2000); Keasey and Watson (2000). Positive • Berger and Udell (1995); Blackwell and Winters (1997); Athavale and Edmister (1999); Strahan (1999); Bodenhorn (2003).
Relationship factors		
- Duration of the relationship	(See Table 2.5)	(See Table 2.5)
- Borrowing concentration ratio		Negative • Blackwell and Winters (1997).
- No. of borrowing bank	Negative • Petersen and Rajan (1994); Harhoff and Korting (1998); Angelini et al. (1998).	Positive • Petersen and Rajan (1994).
- Housebank status		Negative • Degryse and Cayseele (2000); Degryse and Ongena (2005).
Dummy variables		
- Loan type		Negative (Long-term loan) • Goodwin and Mishra (2000); Degryse and Cayseele (2000).
- Loan size		Negative • Strahan (1999); Degryse and Cayseele (2000); Keasey and Watson (2000); Bodenhorn (2003); Degryse and Ongena (2005). Positive • Goodwin and Mishra (2000).

CHAPTER 3
RESEARCH METHODOLOGY

This chapter discusses the research methodology. The research models and the variables used in the models are presented and discussed in the first section. The estimation techniques are discussed in Section 3.2. Section 3.3 provides the fundamental concept of the artificial neural networks and illustrates the networks' topologies. The data collection and the characteristics of borrowers and loans are discussed in Sections 3.4 and 3.5, respectively.

3.1 Research models

According to the banking literature discussed in Chapter 2, bank lending decision (credit scoring), credit availability, and loan pricing models are functions of borrower characteristics, credit risk proxies, relationship indicators, and dummy variables (see equation 3.1, 3.2, and 3.3).

Lending decision	= f (Borrower characteristics, Credit risk proxies, Relationship indicators, Dummy variables)	(3.1)
Credit availability	= f (Borrower characteristics, Credit risk proxies, Relationship indicators, Dummy variables)	(3.2)
Loan price	= f (Borrower characteristics, Credit risk proxies, Relationship indicators, Dummy variables)	(3.3)

where *Dependent variables* are:

- *Lending decision* = 1 if loan is paid (good loan or creditworthiness); 0 if loan is default (bad loan or not credit worthiness),
- *Credit availability* = Volume of loan granted (in Thai baht),

- *Loan price* = Interest rate charged, excluding application fee, borrowing fee, and other fees;

Borrower characteristics include:

- *Asset* (+, +, -) = Total asset value prior to credit decision (in Thai baht),
- *Age* (+, +, -) = Age of the borrower (in years),
- *Education* (+, +, -) = 0 if the highest qualification of the borrower is primary school or lower; 1 otherwise;

Credit risk proxies include:

- *Collateral* (+/-, +/-, +/-) = Total value of collateral pledged to the lending bank (in Thai baht),
- *Current ratio* (+, +, -) = Current assets divided by current liabilities (measures liquidity),
- *Return on asset* (+, +, -) = Net return divided by total assets (measures profitability),
- *Leverage ratio* (-, -, +) = Total liabilities divided by total assets (measures solvency),
- *Capital turnover ratio* (+, +, -) = Gross income divided by total assets (measures efficiency),
- *Debt repayment ratio* (-, -, +) = Total debt and interest payment divided by total income (measures repayment ability);

Relationship indicators include:

- *Borrowing from others* (-, -, +) = 1 if the borrower has an outstanding debt with other financial sources; 0 if the borrower borrows from only one bank,
- *Duration* (+, +, -) = The number of years of bank-borrower relationship prior to the credit decision;

Dummy variables include:

- *Sector* = 1 for agricultural loan; 0 for non-agricultural loan,
- *Province* = Dummy variables for province - Province j. (1 if the observation is in province j; 0 otherwise),
- *Major production* = Dummy variables for major production - Horticulture, Orchard/Vegetable, Livestock/Aquaculture, and Others (1 if the borrower's major production is in the identified major production group; 0 otherwise),

- *Loan type* = Dummy variables for loan type – Cash credit loan, Short-term loan (shorter than 12 months), Medium-term loan (between 1 to 5 years), and Long-term loan (longer than 5 years) (1 if the loan contract belongs to a specific loan type; 0 otherwise),
- *Loan size* = Dummy variables for loan size – Small loan (less than 0.1 million baht), Medium loan (between 0.1 to 1 million baht), and Large loan (more than 1 million baht) (1 if the amount of credit is categorized into a particular size; 0 otherwise)[4],
- *Lending bank* = Dummy variables for lending bank – Bank k. (1 if loan is granted by bank k; 0 otherwise),
- *Lending year* = Dummy variables for lending year (1 for the year of observation; 0 otherwise).

The sets of three positive and negative signs in the above specification indicate the hypothesized signs of the variable on bank lending decision, credit availability, and loan pricing models, respectively. For example, *Asset* (+, +, -) is positively related to the probability of a good loan, positively related to the loan amount, but negatively related to the loan price.

Bank lending decision (credit scoring), credit availability, and loan pricing models are based on quantifiable information, information from the financial statements and historical data, rather than intangible information, such as borrower characteristics and reliability of the borrower, which is difficult to verify and quantify. However, the literature has primarily focused on the bank-borrower relationship that assists the bank to obtain such information. The proximity relationship between the bank and the borrower has shown to facilitate monitoring and screening, and can overcome problems of asymmetric information (Boot, 2000). Thus, it is important to integrate all relevant information into the models.

The total asset value of the borrower (*Asset*) reflects the borrower's wealth. The lender generally prefers to lend to the borrower with higher level of wealth. According to Angelini et al. (1998), Cole (1998), and Degryse and Cayseele (2000), the variable *Age* is

[4] NZD 1 = 29.6791 baht (the exchange rate on 22 June 2005)

referred as the actual age of the borrowing firm. It represents the firm's experience in the business and also reflects the information revealed to the public as a whole. This research focuses on the rural lending and the borrowers typically are individual borrower. Thus, the *Age* variable in this research refers to the age of the borrower and reflects the borrower's reputation and the different investment opportunities between young and old borrowers. *Education* indicates the literacy level of the borrower.

Total value of collateral[5] (*Collateral*) indicates whether the loan is secured or not, and shows the ability and intention to repay the loan by the borrower. A relatively high collateral value compared to the loan amount will make the loan less risky from the lender's point of view. The key financial ratios used in *Credit risk proxies* include *Current ratio*, *Return on asset*, *Leverage ratio*, *Capital turnover ratio*, and *Debt repayment ratio*. They are employed to represent liquidity, profitability, solvency, efficiency, and repayment ability of the borrower, respectively. The purpose of using the financial variables in the models is to control for the observable risk of the borrower that determine the bank lending decision, amount of credit granted and the loan rate. Assuming that all else is equal, a riskier borrower is expected to have a lower probability of a good loan, receive a smaller amount of credit, and pay a higher loan rate.

Borrowing from others is a dummy variable indicating whether the borrower has an alternative source of funds or not. If the borrower has only a single source of funds, the borrower might also buy other products and services from the bank. Therefore, the bank could obtain more information about the borrower's financial status. These sources of information should reduce the monitoring costs of the bank, the expected cost of the loan, and the loan rate[6] (Degryse and Cayseele, 2000). The length of bank-borrower relationship (*Duration*) is another measurement of the relationship. The relationship between the bank and the borrower starts when the borrower buys a product from the bank. *Duration* indicates the proximity relationship between the bank and the borrower.

---

[5] A further distinction can be made between "inside collateral" (assets of the borrower) and "outside collateral" (assets belong to another party). However, the data limitations may prevent a clear distinction between inside and outside collateral. Therefore, in this research, it is assumed that there is no different between inside and outside collateral.

[6] Since the borrower also buys the other products and services from the bank, the bargaining power of the borrower would increase. In other words, cross-subsidization could negatively influence the loan rate (Degryse and Cayseele, 2000).

Dummy variables are included to describe the systematic effects relating to the type of borrower and the type of contract. These variables include *Sector*, *Province*, *Major production*, *Loan type*, *Loan size*, and *Lending bank*. Since it is generally assumed that an agricultural loan is more risky than non-agricultural loan, it is expected that an agricultural loan will be charged at a higher lending rate. *Province* dummy variables are included to account for the province effect, because different provinces have different risk levels, due to the different economic, social, and environmental conditions of each province. *Lending bank* and *Lending year* dummy variables are also included to control for the variation that might occur from the banks and to control for the business cycle effects, respectively.

Major production, *Loan type*, and *Loan size* dummy variables are hypothesized to influence bank lending decision, credit availability, and loan price. For example, the borrower who has a cash crop (*Horticulture*) as their major production would require a smaller amount of credit than the other farm types, and the contract term for the cash crop production is a short-term contract. Thus, this group of borrowers would have a higher probability to obtain a loan and should be charged a lower loan rate. This is because the *Short-term loan* is less risky than *Medium-term* or *Long-term loan*, and the lending risk is relatively low. In contrast, if the major production of the borrowers is either *Orchard* or *Livestock*, which may need a *Large* and *Long-term loan*, they would be expected to pay a higher loan rate.

With reference to the literature, it is expected that most of the variables in *Borrower characteristics*, *Credit risk proxies* and *Relationship indicators*, except *Collateral*, *Leverage ratio*, *Debt repayment ratio* and *Borrowing from others*, should have a positive relationship with the lending decision (probability of a good loan) and the availability of credit (volume of credit), but a negative relationship with the loan price (interest rate charge). Conversely, *Leverage ratio*, *Debt repayment ratio* and *Borrowing from others* would have a positive relationship with the interest rate charge and a negative relationship with the lending decision and the availability of credit (see Turvey and Brown, 1990; Petersen and Rajan, 1994; Berger and Udell, 1995; Degryse and Cayseele; 2000).

The literature shows that collateral can mitigate moral hazard and adverse selection problems in loan contracting (Stiglitz and Weiss, 1981; Chan and Thakor, 1987). The good borrowers signal their better prospect through their willingness to pledge collateral. Thus, *Collateral* should increase the availability of credit and decrease the price of credit. However, the collateralized loans are riskier and are charged higher loan rates than uncollateralized loans (see Berger and Udell, 1990; Blackwell and Winters, 1997; Athavale and Edmister, 1999; Bodenhorn, 2003). Therefore, its impact on lending decision, credit availability and credit price is ambiguous.

3.2 Estimation techniques

The bank lending decision (credit scoring) model will be analyzed using logistic regression and artificial neural networks, while both credit availability and loan pricing models will be estimated by multiple linear regression analysis and artificial neural networks technique. The bank lending decision model is given as follows (Gujarati, 1995):

$$P_i = E\left(Y_i = 1 \middle| X_{ij}\right) = \frac{1}{1+e^{-Z_i}} = \frac{1}{1+e^{-\left(\alpha+\sum_j \beta_j X_{ij} + \varepsilon_i\right)}} \tag{3.4}$$

where Y_i equals to 1 if loan is paid (good loan); 0 if loan is default (bad loan);

P_i is the estimated probability of a good loan (high value of P_i implies low default risk);

$Z_i = \alpha + \sum_j \beta_j X_{ij} + \varepsilon_i$

α and β_j are an intercept term and parameters, respectively.

X_{ij} are *Borrower characteristics*, *Credit risk proxies*, *Relationship indicators* and *Dummy variables*;

ε_i is the error term,

Equation 3.4 represents the cumulative logistic distribution function. If P_i is the probability of a good loan, then, the probability of a bad loan or (1-P_i) given as follows:

$$\left(1-P_i\right) = \frac{1}{1+e^{Z_i}} \tag{3.5}$$

Therefore, the odds ratio in favor of a good loan or $\frac{P_i}{\left(1-P_i\right)}$ can be written as follows:

$$\frac{P_i}{(1-P_i)} = \frac{1+e^{Z_i}}{1+e^{-Z_i}} = e^{Z_i} \tag{3.6}$$

Taking the natural log on equation 3.6 becomes:

$$Z_i = \ln\left(\frac{P_i}{1-P_i}\right) = \alpha + \sum_j \beta_j X_{ij} + \varepsilon_i \tag{3.7}$$

where Z_i is the natural logarithm of the odds ratio in favor of a good loan.

The model is a binary choice model and the use of ordinary least squares estimation technique is inappropriate (Maddala, 1983). Thus, to obtain efficient parameter estimates, the maximum likelihood estimation technique is applied to the logistic regression. The likelihood function L for the model is given as follows (Maddala, 2001):

$$L = \prod_{Y_i=1} P_i \prod_{Y_i=0} (1-P_i) \tag{3.8}$$

From equation 3.7, the probability of a good loan can be obtained by the following equation (Greene, 1997):

$$P_i = \Pr ob\left(Y_i = 1 \middle| X_{ij}\right) = \frac{e^{Z_i}}{1+e^{Z_i}} \tag{3.9}$$

The credit availability and loan pricing models are given as follow (see equation 3.10):

$$Y_i = \alpha + \sum_j \beta_j X_{ij} + \varepsilon_i \tag{3.10}$$

where Y_i is volume of credit grant or interest rate charge;

α and β_j are an intercept term and the parameters, respectively ;

X_{ij} are *Borrower characteristics*, *Credit risk proxies*, *Relationship indicators* and *Dummy variables*;

ε_i is the error term.

To estimate the credit availability and loan pricing models, the ordinary least squares (OLS) method is utilized.

Four different types of artificial neural networks will be applied in this research including multi-layer feed-forward neural network (MLFN) with one hidden layer, Ward network (WN), general regression neural network (GRNN) and probabilistic neural network (PNN). However, the PNN is only applicable to the lending decision model since it is a special neural network for choice modeling. Therefore, only the first three networks will be applied with credit availability and loan pricing models.

To develop the bank lending decision (credit scoring), credit availability (volume of credit) and loan pricing (interest rate charge) models via the artificial neural networks technique, the same set of dependent and independent variables used in the logistic and multiple regression models are utilized and the neural networks software package, NeuroShell2, is used to construct the models.

To evaluate and compare the forecast accuracy between different estimation techniques for all the models, the out-of-sample forecasting procedure will be employed. The sample size will be randomly divided into two different sets, namely the "estimation set" and the "forecasting set", or the "training set" and the "production set". The estimation and the forecasting sets contain 80% and 20% of the total sample, respectively. The forecast classification result, R-squared (R^2) and Root Mean Squared Error (RMSE) will be calculated and compared (see equation 3.11 and 3.12)[7]. The model with a higher percentage correct on the forecast classification, higher R^2 and lower RMSE is considered to be a relatively superior model.

[7] The forecast classification (at cut-off point = 0.5) will be applied only to the bank lending decision model since it is a binary choice model, while both R^2 and RMSE will be applied to both credit availability and loan pricing models.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2} \tag{3.11}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2} \tag{3.12}$$

where Y_i is actual value (volume of credit or interest rate charge);

$\overline{Y}$ is mean value of Y_i;

$\hat{Y}_i$ is estimated value of Y_i;

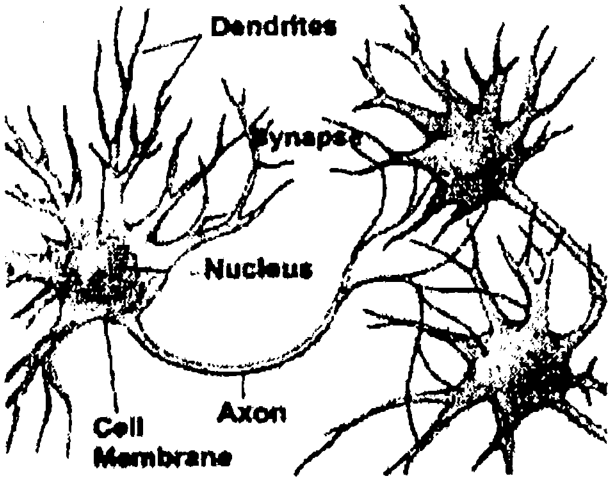
n is the number of observations on the forecasting set.

3.3 Artificial neural networks (ANN)

Statistical estimation techniques and artificial neural networks are closely related to each other. The major difference between them is that the statistical techniques have concentrated on linear problems that are relatively tractable, while neural network deals mainly with non-linear problems (Smith, 1996).

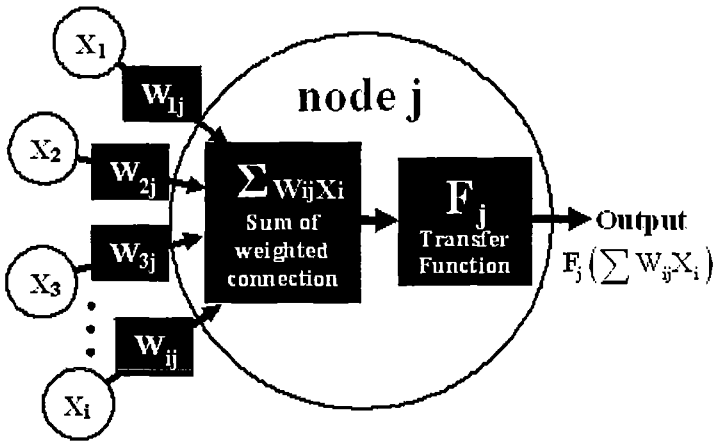
Neural computing inspired by the understanding of the biological nervous system and artificial neural networks (ANN) are computing models that imitate the human brain's working and learning processes (Martin and Jain, 1999). The basic functional element in biological and artificial systems are simple processing units called neurons, or nodes in artificial systems, which communicate by sending signals to each other through a large number of interconnections. These interconnections are the synapses in the biological systems (see Figure 3.1) and weighted connections in the artificial systems, W_{ij} in Figure 3.2. Each neuron in biological system and each node in artificial systems perform a relatively straightforward function. It obtains the inputs from other neurons (nodes) or external sources (X_i in Figure 3.2) and processes them into an output signal, which is transferred to other units (Pham and Liu, 1995).

Figure 3.1: Biological neuron



Source: Anonymous

Figure 3.2: Artificial neuron and structure of a computational unit (node j)



Source: Modified from Coakley and Brown (2000)

A biological neuron collects information via its dendrites. The information is processed by the soma and the neuron's axon transfers the signal to other neurons. An artificial neuron consists of two main components: a summation and a transfer function. The summation function sums up all the input signals from each connection (X_i) times the value of the connection weight between connection i and node j (W_{ij}) (see Figure 3.2). Following the summation of the inputs, the output for node j is determined by applying a transfer function F_j, also known as the activation function, to the summation value (see equation 3.13 and

3.14). The output is then transferred to the other artificial neurons in next layer of the network.

$$\text{Sum}_j = \sum\nolimits_j (W_{ij} X_i) \tag{3.13}$$

$$U_j = F_j(\text{Sum}_j) = F_j \sum\nolimits_j (W_{ij} X_i) \tag{3.14}$$

where U_j is output for node j;

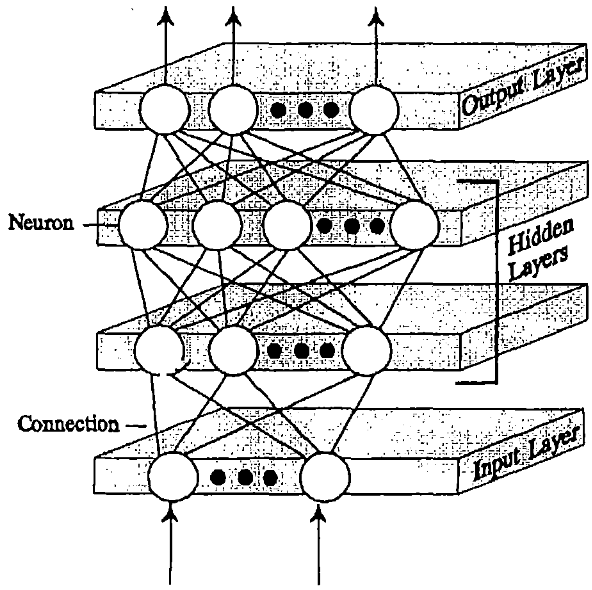
F_j is a transfer function in different functional forms: linear functions, linear threshold functions, step linear functions, sigmoid function, Gaussian functions and etc (Coakley and Brown, 2000).

The biological nervous system is systematized around a main nervous center, the brain. Information from outside or inside of the body arrives through specific sensory routes and are only perceptible to specific sensory regions located in definite area of the brain, where they are recognized, processed, and interpreted. These areas interrelate with many other regions of the brain and the decision is made to respond to the stimulus that has been received. The parallel processing of information by the brain makes it possible to analyze multiple pieces of data from various sources at the same time (Tarasenko, 1998).

Similarly, an artificial neural network consists of three main layers: input layer, hidden layer(s) and output layer. Generally, at one end of the network, the input layer, information or data is received. Thus, each neuron in the input layer brings the value of one independent variable into the network. Hidden layer(s), where the information is recognized, processed and interpreted, is always line in the middle of the network and may have more than one hidden layer. At the other end, the output layer, the network produces responses or outputs. Basically, one dependent variable is processed by one output node (see Figure 3.3).

The way neurons are organized and connected is known as the network's architecture that can vary from a single hidden layer to a complex web of neurons' clusters. Hidden layers in a neural network are known as feature detectors and different transfer functions applied to hidden layer groups detect different features in a pattern processed through a network. A network design may use, for example, linear function, logistic function and Gaussian function on three different hidden groups. The output layer will receive various views of the data and the mixture of the functions synchronizes in the output layer which may lead to better classification and prediction (Ward System Group Inc., 1993).

Figure 3.3: The artificial neural network structure with two hidden layers



Source: Coakley and Brown (2000)

The use of the neural network model is similar to the process utilized in building the statistical model. However, the neural network must first be trained from a set of data. Training begins with random values assigned to the weights on the interconnections and proceeds iteratively. Weights can be adjusted after the network has processed each example (example-by-example learning), or after all the examples have been processed (epoch-base training). The later technique is commonly used.

Thus, for a particular input, an output is produced from the network. The network then compares the network output to the actual output. The accuracy of this value is determined by the total mean square error (see equation 3.15) and back-propagation error correction method is used to reduce prediction errors; this is accomplished through the adjustment of the connection weights. As the iterative process of incremental adjustments continues, the weights gradually converge on the optimal set of values. Many epochs are required before the training is completed. This estimation technique, so called "back-propagation" technique, is commonly used in many types of artificial neural networks.

$$E = \frac{\frac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{K}\left(Y_{ik} - \hat{Y}_{ik}\right)^2}{NK} \tag{3.15}$$

where Y_{ik} is the k^{th} actual output for the n^{th} sample;

$\hat{Y}_{ik}$ is the k^{th} network output for the n^{th} sample;

N is the number of example in the data set;

K is the number of outputs of the network;

Several criterions can be used to decide when to stop the training process. These include: stopping after a certain number of learning epochs; stopping when the error (difference between the network output and the actual value) reaches a certain level; stopping after a certain number of examples since the minimum error has exceeded a specified number (Ward System Group Inc., 1993).

Generally, the number of dependent and independent variables determine the number of output and input neurons, respectively, while the number of hidden layers and nodes that are included in each hidden layer must be decided by the network developer. This decision determines the complexity of the mapping function and the performance of the network. If too many nodes are used, the network will tend to memorize the problem, and can not be generalized later. On the other hand, if too few nodes are used, the network will generalize well, but may not have enough power to learn the patterns adequately. Unfortunately, there is little theory to support the process for determining of the optimal number of hidden layers and nodes, and the optimal internal error threshold (Ward Systems Group Inc., 1993 and Lenk et al., 1997). Therefore, a trial-and-error process is usually applied to find the optimal artificial neural network model.

3.3.1 Multi-layer feed-forward neural network (MLFN)

The artificial neural network that is widely used is called the multi-layer feed-forward neural network (MLFN). The information in MLFN flows in the direction from the origin to the destination, one cannot return to the origin, and the computational units are grouped into 3 main layers (input, hidden, and output layers) (Hu et al., 1999). Figure 3.4 shows the structure of the multi-layer feed-forward neural network with one hidden layer and one network output, that is the most frequently used structure in many literatures[8]. Since the output of one layer is an input to the following layer, the output of the network can be exhibited algebraically as shown in equation 3.16.

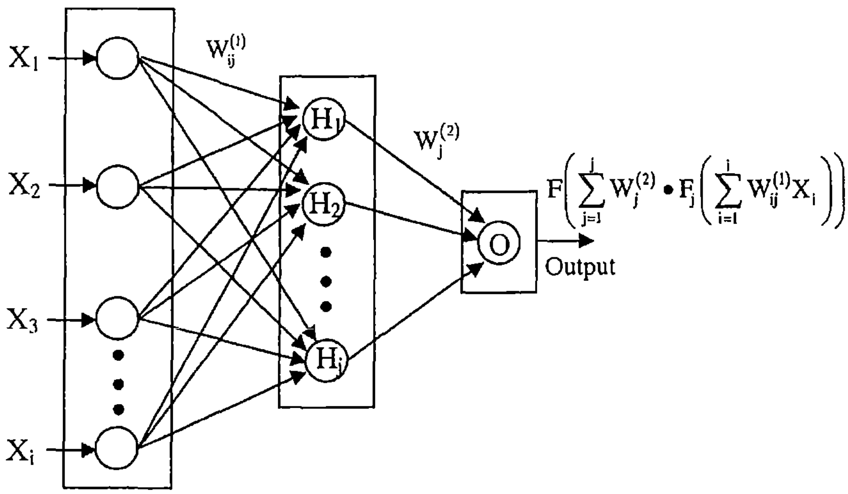
$$O = F\left(\sum_{j=1}^{J} W_j^{(2)}.U_j\right) = F\left(\sum_{j=1}^{J} W_j^{(2)}.F_j\left(\sum_{i=1}^{i} W_{ij}^{(1)} X_i\right)\right) \tag{3.16}$$

where O is the output of the network (i.e. bank's lending decision, volume of credit granted or loan price);

F is the transfer function in the output node;

$W_{ij}^{(1)}$ and $W_j^{(2)}$ are connection weights from input layer (node i) to hidden layer (node j) and from hidden layer (node j) to output layer, respectively.

Figure 3.4: MLFN structure with one hidden layer and one network output



Source: Modified from West et al. (1997), and Gradojevic and Yang (2000)

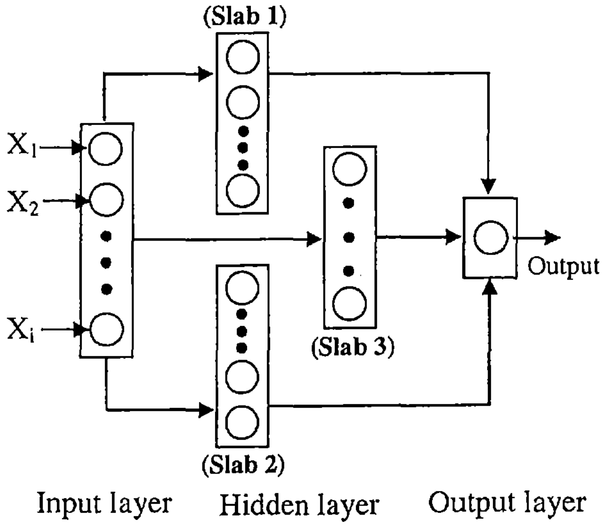
[8] Ward Systems Group Inc. (1993) argues that the three-layer Back-propagation network with standard connections is suitable for almost all problems.

3.3.2 Ward network (WN)

Since the hidden layer(s) and the activation function(s) in a neural network determine how well a problem can be learned, Ward Systems Group Inc. (1993) has invented the Back-propagation network architectures with multiple hidden slaps in the hidden layer, so called Ward Networks. The networks allow the different hidden slaps to use different activation functions. For example, a network design may use a Gaussian function on one hidden slap to detect features in the mid-range of the data and use a Gaussian complement in another hidden slap to detect features from the upper and lower extremes of the data. Thus, the network is expected to provide a better prediction result, since it offers two way of viewing the data.

The Ward network structure applied in this research includes the network with three hidden slabs (see Figure 3.5). The Ward network is a regular three-layer back-propagation network, but there are 3 slabs with 3 different activation functions in the hidden layer.

Figure 3.5: Ward network (3 hidden slaps with different activation functions)



Source: Modified from Ward Systems Group Inc. (1993)

3.3.3 Probabilistic neural network (PNN)

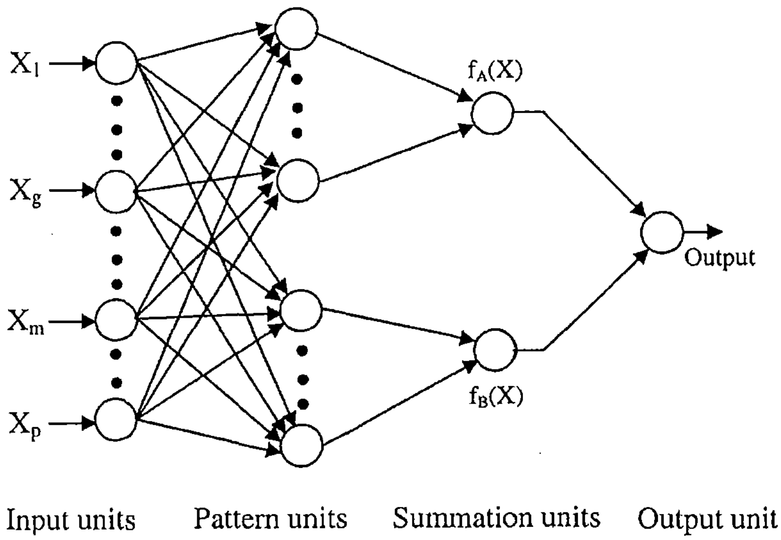
The PNN originally proposed by Specht (1990) is basically a classification network. Its general structure consists of 4 layers - an input layer, a pattern layer (the first hidden layer), a summation layer (the second hidden layer) and an output layer (see Figure 3.6)[9].

PNN is conceptually based on the Bayesian classifier statistical principle. According to the Bayesian classification theorem, X will be classified into class A, if the inequality in equation 3.17 holds (Albanis and Batchelor, 1999):

$$h_A c_A f_A(X) > h_B c_B f_B(X) \tag{3.17}$$

where X is the input vector to be classified;
h_A and h_B are prior probabilities for class A and B;
c_A and c_B are costs of misclassification for class A and B;
$f_A(X)$ and $f_B(X)$ are probabilities of X given the density function of class A and B, respectively.

Figure 3.6: The probabilistic neural network (PNN) architecture



Source: Modified from Specht (1990)

[9] Please note that PNN is not limited to the binary choice classification problem. It also can perform the multi level classifications. Since the bank lending decision is binary (accept or reject the loan application), the discussion in this research is based only on the binary-choice case.

To determine the class, the probability density function is estimated by a non-parametric estimation method developed by Parzen (1962) and extended by Cacoulos (1966). The joint probability density function for a set of p variables can be expressed as follows (Chen et al., 2003):

$$f_A(X) = \frac{1}{(2\pi)^{p/2}\sigma^p n_A} \sum_{j=1}^{n_A} e^{\frac{-(X-U_{Aj})'(X-U_{Aj})}{2\sigma^2}} \tag{3.18}$$

where p is the number of variables in the input vector X;

n_A is the number of training samples which belongs to class A;

U_{Aj} is the j^{th} training sample in class A (the centre of the j^{th} pattern unit of the A class);

σ is a smoothing parameter[10].

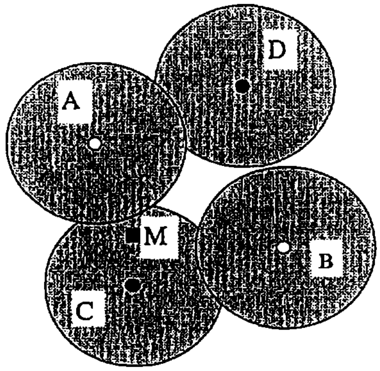
The working process of the PNN begins with the input layer, where the inputs are distributed to the pattern units. Then, the pattern unit, which is required for every training pattern, is used to memorize each training sample and estimate the contribution of a particular pattern to the probability density function (PDF). The summation layer comprises of a group of computational units with the number equal to the total number of classes. Each summation unit that delicate to a single class sums the pattern layer units corresponding to that summation unit's class. Finally, the output neuron(s), which is a threshold discriminator, chooses the class with the largest response to the inputs (Etheridge and Sriram, 1997; Albanis and Batchelor, 1999).

Figure 3.7 provides an illustrative example of the PNN working principle. In Figure 3.7A, there are four pattern units centred at four dots - A, B, C, and D. The black dots belong to class γ and the white dots belong to class α. The radii of the gray areas centred at these four dots are the square root of the variance, which is used to enhance the probability contribution from the samples of the same class and weaken the probability contribution from the sample of different classes. The square represents a testing sample, M. The PNN

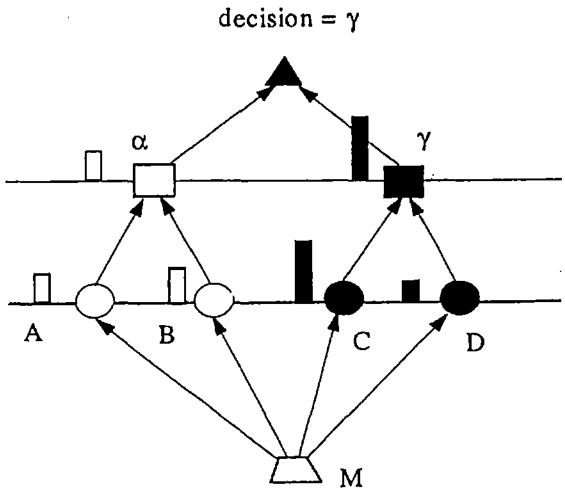
[10] In equation 3.18 the probability density function (PDF) is estimated with the simple covariance matrix $\sigma^2 I$, where I is the identity matrix. However, for the Adaptive PNN, the PDF is estimated with a full covariance matrix. Therefore, a separate smoothing parameter is adapted for each measurement dimension. Adaptive PNN usually outperforms the basic PNN in term of generalization accuracy. The price paid for these improvements is increased training time (Specht, 1996).

calculates the probability contributions of this sample to the four pattern units and sums the probability density functions (PDFs) according to different class labels of the pattern units. As shown in Figure 3.7B, the test sample, M, is fed into the four pattern units – A, B, C, and D, respectively. Each of unit is an independent and identical Gaussian distribution, and each of them calculates a contribution from the testing sample, M, to its PDF. The filled rectangles to the left of the pattern units in Figure 3.7B represent the contributions to the PDFs. The larger the contribution, the higher the rectangle. Then, the PDFs are summed to different class units α and γ, respectively. The results from the summation are called the conditional probability statistics. Because the testing sample has more contribution to class γ than that to class α, the filled rectangle aside the class unit γ is larger than the aside class α. Thus, the testing pattern M belongs to γ (Yang et al., 1999: p 69-70).

Figure 3.7: An illustration of the working principle of a simple PNN model



(a)



(b)

Source: Yang et al. (1999)

From the above example, it can be seen that the PNN is simpler and more intuitive than the back-propagation neural networks and should not be regarded as a "black box". Further, it should be noted that there are two important factors determining the performance of a PNN model: the positions of the pattern units and smoothing parameter. The pattern units are usually set by the training samples, but the function of the smoothing parameter should be optimally selected to minimize the misclassification rate (Yang et al., 1999). However, in practice, to choose an optimal smoothing parameter is not easy and straightforward, especially for the Adaptive PNN that needs to discover the best combination of σ's.

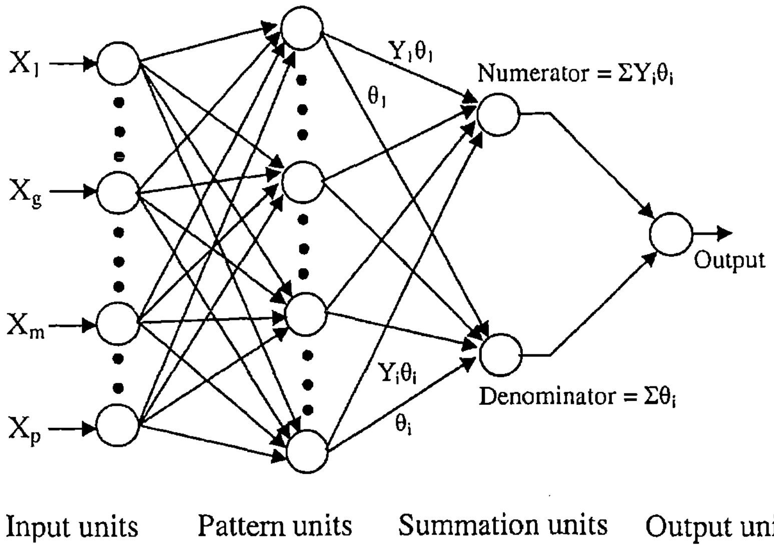
Therefore, in order to obtain the optimal value for the smoothing parameter, the PNN networks in this research are trained by the genetic adaptive learning (genetic algorithm)[11].

3.3.4 General regression neural network (GRNN)

The general regression neural network (GRNN) is a memory-based feed-forward network originally developed in the statistics literature by Nadaraya (1964) known as the Nadaraya-Watson kernel regression. The GRNN is based on nonlinear (kernel) regression theory. As its name implies, GRNN can estimate any arbitrary function between input and output vectors, drawing the function approximation directly from the historical data. Furthermore, it is consistent; that is, the estimation error approaches zero when the sample size becomes large (Wasserman, 1993).

The topology of the GRNN is shown in Figure 3.8. The network consists of four layers: the input layer, pattern layer (the first hidden layer), the summation layer (the second hidden layer), and the output layer.

Figure 3.8: Topology of the general regression neural network



Source: Modified from Specht (1991)

[11] Iterative algorithm can be used to train the PNN networks and generally the training is faster than using the genetic adaptive option. However, the iterative option should be applied when all of the input variables have the same impact on the output prediction. Since the input variables are of different types and some of them may have more of an impact on predicting the output than the others, using genetic adaptive learning to train the network should be more appropriate. This is because the genetic algorithm has the ability to search for the appropriate individual smoothing factors for each input as well as an overall smoothing factor (Ward Systems Group Inc., 1993).

In the GRNN model, the estimation of a dependent variable y (output) with respect to a given vector of independent variables X (inputs) can be regarded as finding the conditional mean of y given by X (also called the regression of y on X)[12]. Equation 19 summarizes this statistical concept (Wasserman, 1993):

$$E\left[y|X\right]=\frac{\int_{-\infty}^{\infty} yf(X,y)\,dy}{\int_{-\infty}^{\infty} f(X,y)\,dy} \tag{3.19}$$

where y is the output value estimated by GRNN;

X is the input vector for the estimation of y;

$E\left[y|X\right]$ is the conditional mean of the output y given by an input vector X;

f(X, y) is the joint probability density function (PDF) of X and y.

Specht (1991) shows that y can be estimated optimally as follows:

$$y=\frac{\sum_{i=1}^{n} Y_i\theta_i}{\sum_{i=1}^{n}\theta_i} \tag{3.20}$$

where $\theta_i = e^{\frac{-(X-U_i)'(X-U_i)}{2\sigma^2}}$ (or outputs of the pattern units) (3.21)

Y_i is the target (desired) output corresponding to input training vector X_i;

X and U_i are the input vector and the training vector i (or the centre of the pattern unit i), respectively;

σ is the smoothing parameter[13].

From Figure 3.8, the input layer is responsible for the reception of information. The number of input units is equal to the number of input variables. Then, the inputs are

---

[12] Although the GRNN can provide a multivariate vector of outputs without loss of generality, the case of univariate output is described here for simplicity.

[13] Similar to the Adaptive PNN, the Adaptive GRNN is a separate smoothing parameter adapted for each measurement dimension. The adaptation can greatly improve the estimation accuracy of the model.

presented to the second layer of processing neurons called pattern units. There is a unique pattern unit for each of training pattern. A pattern unit is used to combine and process the data in a systematic fashion such that the relationship between the input and the proper response is memorized. The outputs of the pattern units, θ_i (see equation 3.21), are subsequently forwarded to the summation layer. Theoretically, there are two different types of processing neurons in the summation layer. They are the weighted summation unit (or numerator neuron) and the simple arithmetic summation unit (or denominator neuron). After the sums are calculated, they are sent to the output unit. To obtain the GRNN regression output (y), the output unit performs a simple division of the signal coming from the weighted summation unit by the signal coming from the arithmetic summation unit (see equation 3.20) (Leung et al., 2000; Amaral et al, 2002; Chen and Leung, 2004)[14].

Similar to the PNN network, the GRNN requires supervised training. The network performs learning by examining the relationship between each pair of input vector X and the observed corresponding output y, and finally construes the underlying function by summarizing all of these relationships (Leung et al., 2000). According to equation 3.21, the smoothing parameter is regarded as the crucial factor to the performance of the GRNN. Therefore, to search for the optimal smoothing parameter, which minimizes the mean squared error, the genetic adaptive learning (genetic algorithm) is applied on the network training process.

3.3.5 Genetic algorithms (GAs)

Genetic algorithms are a means by which machines can emulate the mechanisms of biological genetics and natural selection. This involves searching high-dimensional spaces for superior solutions, if they are not yet optimal. The algorithms are simple, robust, and general; no knowledge of the search spaces is assumed (Wasserman, 1993). Therefore, they are widely used in many applications requiring the optimization of a certain multi-dimensional function, such as neural networks.

---

[14] In case of multivariate outputs, the number of weighted summation unit is always the same as the number of the GRNN output units. Each of the output units is connected only to its corresponding weighted summation unit and to the simple arithmetic summation unit. Thus, the output units divide its corresponding weighted summation unit by the arithmetic summation unit to obtain the network outputs ($y_j = \sum Y_{ij}\theta_i / \sum \theta_i$).

Genetic algorithms were first proposed by John Holland (1975) as an algorithmic concept based on a Darwinian-type survival-of-the fittest strategy with sexual reproduction, where stronger individuals in the population have a higher chance of creating offspring. A basic GA comprises three fundamental genetic operations found in natural genetics to guide their trek through the search space: selection, crossover, and mutation. These operations are used to modify the chosen solutions and select the most appropriate offspring to pass on to succeeding generations. GAs consider many points in the search space simultaneously and have been found to provide a rapid convergence to the global optimum solution in many types of problems; in other words, they usually exhibit a reduced chance of converging to local minima (Pham and Liu, 1995; Martin and Jain, 1999).

In computing terms, a GA maps a problem onto a set of binary strings (zeros and ones), and each string signifying a possible solution. Various portions of the bit-stings represent the parameters in the search problem. Then, the GA manipulates the most promising string in searching for improved solutions. A GA operates typically through a simple cycle of four stages (May, 1996):

1. Creation of population of strings.
2. Evaluation of each string.
3. Selection of best strings.
4. Genetic manipulation, to create the new population of strings.

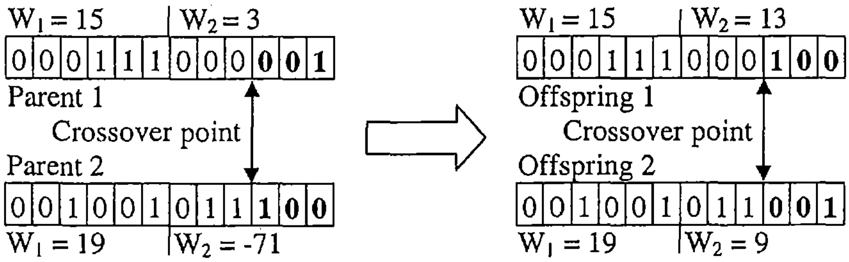
In the first stage, the algorithm starts with generating an initial population of potential solutions for being a starting point for the search process. Each element in the population is encoded into a string (or chromosome), to be manipulated by the genetic operators. In the second stage, the performance or fitness of each individual string from the current population is evaluated. Then, pairs of individuals are selected to mate with each other to form the offspring, which then form the next generation. Selection is based on the survival-of-the fittest strategy, where the strings with high fitness values (i.e. good solutions to the optimization problem under consideration) receive larger numbers of copies in the new population (May, 1996). The most commonly used strategy to select pairs of individuals is the method of roulette-wheel selection, in which every string is assigned a slot in a simulated wheel sized in proportionate to the string's relative fitness (P_i)[15]. This ensures

[15] $P_i = F_i / \Sigma F_i$ where F_i is the fitness value of string i.

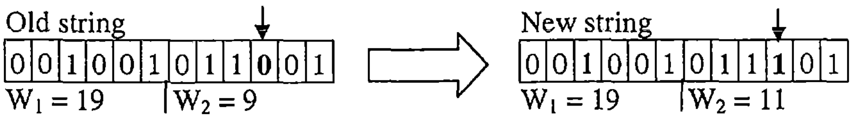
that highly fit strings have a greater probability to be selected to form the next generation through the genetic manipulation, crossover and mutation (Martin and Jain, 1999).

The crossover operator takes two parent strings and swaps part of their genetic information (bit-value) to produce two new strings (see Figure 3.9a). After the crossover point has been randomly chosen, portions of the parent strings are interchanged to produce the new offspring. The two individuals (children) resulting from each crossover operation will now be subjected to the mutation operator in the final step in forming the new generation. The mutation operator reverses one or more bit values at randomly selected locations in randomly selected strings according to a specified mutation probability (see Figure 3.9b). This operation is inspired by the possibility that the initially defined population might not contain all the information necessary to solve the problem. Thus, it forces the algorithm to search new areas, prevents the premature convergence and helps to find the global optimal solution (Pham and Liu, 1995).

Figure 3.9: (a) Crossover operation; (b) Mutation operation



(a)



(b)

In sum, GAs could surpass the other optimization techniques because they differ from the others in the following ways (Goldberg, 1989; Wasserman, 1993):

1. GAs work with a coding of the parameter set, not the parameters themselves.
2. GAs search from a population of the points, not a simple point.

3. Gas operate without any knowledge of the search space and mainly use payoff (objective function) information. No derivatives are calculated.
4. GAs use probability transition rules, not deterministic rules.

3.3.6 Applications of ANN in economics and finance

The artificial neural networks (ANNs) have been successfully employed in many different disciplines, including biology, psychology, statistics, mathematics, medical science, and computer science. However, the ANNs have been gaining popularity as a standard analytical tool in business, economics, and finance recent year. For example, in economics and finance disciplines, the ANNs have been applied to generalization problems including classification and prediction problems (West et al., 1997; Roa and Ali, 2002; DeTIENNE et al., 2003).

According to the literature, the ANNs have been successfully used for house price prediction (Tay and Ho, 1991; McCluskey, 1996; Limsombunchai et al., 2004), commodity price prediction (Khozadi et al., 1995), consumer choice prediction (West et al., 1997; Gan et al., 2005), bankruptcy prediction (Luther, 1998), credit evaluation (Altman et al. 1994; Barney et al., 2000; Lee and Jung, 1999; Wu and Wang 2000) (see section 2.5), predicting insurance losses (Kitchens et al., 2002), predicting bond ratings (Albanis and Batchelor, 1999), forecasting returns on the stock markets (Shachmurove and Witkowska, 2000; Yim, 2002), exchange rate forecasting (Gradojevic and Yang, 2000; Leung et al., 2000), and macroeconomic forecasting (Gonzalez, 2000).

Tay and Ho (1991) used a large sample data from the apartment sector in Singapore and found that a neural network model performs better than a multiple regression model in estimating the apartment value. The authors conclude that the neural network can generate valuation patterns for "true" open market sales in the presence of some "noise" as a way of establishing a robust estimator. Similar results can be found in MaCluskey (1996) and Limsombunchai et al. (2004) studies.

Kohzadi et al. (1995) compared the predictive power of the neural network model against the ARIMA model on the commodity price prediction. The weekly corn closing futures prices (US cents/bushel) from the Chicago Board of Trade from January 1974 to October

1993 were used in their analysis. The results showed the neural network model to be more accurate than the ARIMA.

West et al. (1997) examined the predictive relationship between retail store image variables and consumer patronage behaviour toward three nationwide mass-merchandise retailers. The authors directly compared a neural network model with discriminant analysis and logistic regression. The authors concluded that the neural network models offer superior predictive capabilities over traditional statistic methods in predicting consumer choice. Gan et al. (2005) compared the performance of logistic regression against the neural network models with respect to their ability to identify consumer choice on banking channels. The authors found that the probabilistic neural network (PNN) was the best model for consumers' choices prediction.

Luther (1998) developed a prediction model using artificial neural networks for predicting the outcome of bankruptcy, based on the firm's financial ratios at the time of filing for Chapter 11. The data set included 104 firms that filed for bankruptcy under Chapter 11 and had their case decided before December 1992. The authors argued that the ANNs model has significant higher prediction accuracy than the logit model in both training samples as well as the holdout samples at almost all cutoff points. In addition, the predictive accuracy of ANN was less sensitive to changes in the cutoff point in the model, thus making it a more robust technique than logit.

Kitchens et al. (2002) utilized the artificial neural networks to predict automobile insurance losses. The data consisted of over 174,000 records from private passenger automobile policies in the United States. The authors found that the neural network was more successful in accurately categorizing a loss than the multiple regression and logit models, but it was less successful in categorizing the policies with no-losses.

In Albanis and Batchelor (1999) study, the ability to classify long-term bond ratings of the linear discriminant analysis (LDA) and the probabilistic neural network (PNN) were compared and contrasted. The authors found that the PNN model classifies a number of bond issues into boundary rating groups significantly better than the LDA model while at the same time did not deviate from multivariate normality as opposed to the LDA model.

Shachmurove and Witkowska (2000) tested the dynamic interrelations among major world stock markets of Canada, France Germany, Japan, United Kingdom, and the United States using the artificial neural networks. Their data was derived from daily stock market indices and covered the period from January 3, 1987 through November 28, 1994, with a total of 2,064 observations per stock market. The authors concluded that the multi-layer feed-forward neural network (MLFN) models were better able to foresee the daily stock returns than the traditional forecasting models, in term of low mean squared errors.

Yim (2000) used the MLFN to predict Brazilian daily index returns. The predictive results from MLFN were compared with the structural time series models (STS) and the ARMA-GARCH models. The author found that the MLFN is superior to ARMA-GARCH and STS models, and the volatility derived from the ARMA-GARCH model is useful as input to a neural network.

In Gradojevic and Yang (2000) study, the ANNs were used to test for high-frequency Canada/U.S. dollar exchange rate forecasting. The authors reported that ANNs are consistently better in terms of the root mean squared error (RMSE) than random walk and linear models for the various out-of-sample set sizes. Moreover, the ANNs perform better than other models in terms of percentage of correctly predicted exchange rate changes. Leung et al. (2000) also used the ANNs to predict the monthly exchange rate of three currencies, British pound, Canadian dollar, and Japanese yen. The authors applied a specific neural network architecture called general regression neural network (GRNN) and compared its performance with the multi-layer feed-forward neural network and random walk models. The findings from their study showed that GRNN not only has a higher degree of forecasting accuracy but also performs statistically better than other evaluated models for different currencies.

Gonzalez (2000) developed a prediction model using neural networks to forecast the Canada's real GDP growth. For both the in-sample and out-of-sample periods, the forecasting accuracy of the neural networks was found to be superior to a well-established linear regression model developed by the department of finance, with the error reduction ranging from 13 to 40 percent. However, various tests showed that there was little evidence that the improvement in forecasting accuracy was statistically significant. Therefore, the authors argued that neural networks should complement standard econometric methods, rather than a substitute, since the method also presents various weaknesses.

3.4 Data collection

The data collection process turned out to be the greatest impediment because of the difficulty in accessing Thai banks data and information from their credit files. There are three major reasons why it is difficult to access the bank data completely. These are:

1. According to the legislation on data and information, banks cannot release their customers' private information to the third party for any reason that is not related to the bank's business.
2. Banks are afraid that if any data or information about their customers is made public, and someone could gain any private information advantage, they could be prosecuted by their customers.
3. Banks fear that the results published from research could affect practices negatively.

In December 2003, 8 banks that lend to the agricultural sector before and after the 1997 financial crisis were approached to participate in this research. Only the Bank for Agriculture and Agricultural Cooperatives (BAAC) agreed to cooperate[16] and allowed the researcher to disclose the bank's name. However, the researcher is not allowed to name the branches or provinces that are included in the sample and is not allowed to use the data set for any other research. Furthermore, only the credit files after the 1997 financial crisis are available and could be viewed. This is because the data files are retrieved from the "Credit BPR" (Credit Business Process Reengineering) database that was first implemented in 1996 and they are incomplete[17].

To avoid the impact of the 1997 financial crisis on the bank business and to incorporate with non-agricultural lending which started in 1999[18], the data set included loans granted in 2001 to 2003. About 3 to 5 provinces which had credit files available on Credit BPR

---

[16] 56.10 percent of the total loan in the agricultural sector came from BAAC, whereas 12.30, 10.80 and 9.40 percent came from the village and city fund, commercial banks and agricultural co-operatives, respectively (OAE, 2003).

[17] In 1996, there were only 40 branches that were connected to and could access the Credit BPR database. The number of the connected branches increased from time to time and there were about 250 branches linked to the database in 2003.

[18] Prior to the year 1999, BAAC lending was in agriculture. Under the 1999 BAAC Act, BAAC was allowed to provide non-agricultural lending.

database during the predetermined 3 years period were randomly selected from each region. The data set was retrieved from between 99 to 136 braches in 17 provinces (see Table 3.1).

Table 3.1: Number of provinces and branches included in the data set

Region	No. of provinces	No. of branches (Agriculture)			No. of branches (Non-Agriculture)		
		2001	2002	2003	2001	2002	2003
1. Northern	4	33	34	34	24	34	32
2. North-eastern	5	50	53	54	48	52	50
3. Central and Eastern	5	29	30	25	13	24	21
4. Southern and Western	3	18	19	16	14	17	14
Total	17	130	136	129	99	127	117

The total number of observations from the available data set was 242,168 samples, including 229,293 samples from agricultural lending and 12,875 samples from non-agricultural lending (see Table 3.2). However, the data set contained missing data on many variables due to the recent implementation of the database system. As a result, the samples that have no personal details (such as age, education, home address, etc.), financial details (such as asset, farm income, expense, net income, etc.) and debt repayment history were discarded from the data set. The usable data set consists of 18,798 loan contracts with 17,028 agricultural loan contracts and 1,770 non-agricultural loan contracts (see Table 3.3) All loans are credit services for individual borrowers and are under the normal loan scheme (excluding the government loans for specific projects).

Table 3.2: The number of observations on the available data set

Region	Agriculture			Non-Agriculture		
	2001	2002	2003	2001	2002	2003
1. Northern	41,896	24,772	15,271	403	2,311	2,749
2. North-eastern	63,402	20,540	15,675	725	1,456	2,481
3. Central and Eastern	12,863	8,051	5,640	57	258	320
4. Southern and Western	12,036	5,419	3,728	74	844	1,197
Sum	130,197	58,782	40,314	1,259	4,869	6,747
Total (N = 242,168)	N_1 = 229,293			N_2 = 12,875		

Table 3.3: The number of observations on the usable data set

Region	Agriculture			Non-Agriculture		
	2001	2002	2003	2001	2002	2003
1. Northern	-	3,073	2,301	-	272	206
2. North-eastern	325	1,836	1,918	-	168	337
3. Central and Eastern	173	1,989	1,767	-	49	68
4. Southern and Western	1,165	1,227	1,254	-	223	447
Sum	1,663	8,125	7,240	-	712	1,058
Total (N = 18,798)	N_1 = 17,028			N_2 = 1,770		

3.5 BAAC and its role in rural financing

The Bank for Agriculture and Agricultural Cooperative (BAAC) is a specialized bank. It was established by the 1966 BAAC Act, which has been amended several times, most recently in 1999, when lending mandate to former clients was expanded to cover non-farm activities. The BAAC's objective is to provide financial assistance to farmers, farmer associations, and agricultural cooperatives for the following purpose (ADB, 2002):

1. To undertake farm and other farm-related activities.
2. Other non-farm activities to increase income.
3. To develop agricultural knowledge to increase income or improve the quality of life of the farmers and their families.
4. To carry out the project intended to promote or support agricultural activities in joint venture with the entrepreneur to increase income or improve quality of life of the farmers and their families.

However, it should be noted that a limit of 20 percent is set on the total amount that BAAC may lend for items 3 and 4 above to ensure that sufficient funds are available for its main loan categories.

Since its establishment in 1996, the founding statement on credit lending to farmers stipulated that there was no need to offer credit to every farmer in Thailand. In the first 11 year of its operation, the bank's rate of new clients was very low due to the limited capital (BAAC, 2002). In 1971, the bank started providing credit by group guarantee without collateral. Since then, the bank has expanded the credit services (Poramacom, 2000). Its main products during that time period include short-term and medium-term loans for farmers and farmer institutions who used them as revolving credit for their production and agricultural investment.

During 1976-1986 (the 2nd era) BAAC's capital increased as mandated by state policy, enabling the bank to grow at a fast pace. For example, one policy aimed at accelerating government agencies involvement in setting up farming associations relied on BAAC to offer credit services to the newly established entities. The BAAC also took on the role in providing financial support services to projects set up under the government policy initiatives. These include the Crop Pledging Scheme, Long Term Loans for Agriculture for Farmer Project, and Long Term Loans for Investment in Fixed Assets for Farmer Institution Project (BAAC, 2002).

The 3rd era of the bank (from 1987 to 2000) was marked by a recession in Thailand. Before 1995, BAAC accelerated its effort to get more farmers as clients. The Entire Village Farmer Acceptance as Clients Scheme had been established and the bank increased the lending of long-term agriculture loans by cooperating with state agencies setting up special projects, which included the Beef Cow Raising Promotion and Dairy Cow Raising Promotion Project. In 1997, Thailand was hit by a severe financial and economic crisis. Most financial institutions faced liquidity problems and had a large volume of non-productive loans due the effects of a change in the foreign exchange regime, wrong management policy and loose lending policy. Financial institutions involved in the agricultural were less affected than other banks but as time passed, the income of farmers began to suffer. During this period, BAAC began to target more deposits. The main deposits included Thaweechoke Savings Accounts and Thaweesin Savings Certificates. After 1997, the farmers joined in special government-assisted projects promoted by BAAC and state agencies to solve their huge debt problems (BAAC, 2002).

In the 4th era (from 2001 to present), BAAC worked with state policy to help boost the grassroots economy and strengthen rural communities socially. These include setting up of the Debt Suspension and Debt Burden Reduction for Small-Scale Farmers Project, the Rural Village Fund Project, and the Community Enterprise Credit for One Tambon One Product Policy Support Project. During this period, BAAC changed its role to be a rural development bank as required by government policy (BAAC, 2002).

At the end of fiscal year 2002 (31 March 2003), the BAAC operating fund totalled 355.28 billion baht. There were 4 major sources of operating fund; deposit from public, borrowing from local and overseas sources, shareholder's equity, and other liabilities. The percentage of the sources of fund included 79.85, 7.88, 7.79, and 4.48 of deposit, borrowing, shareholder's equity, and other liabilities, respectively (BAAC, 2002). Therefore, currently BAAC has funded its loan outreach through the savings mobilization.

In the fiscal year 2002, BAAC provided credit services worth 120.62 billion baht to farmers and farmer institutions (including credit services for community enterprises but not loans extended under government policy project). Of this amount about 93.61 billion baht (77.61 percent) was lent directly to farmers, 26.97 billion baht (22.36 percent) to agricultural cooperatives, and 34 million baht (0.03 percent) to farmer associations. Credit services were classified by types of clients as follows (BAAC, 2002):

1. Credit services for individual farmers

   In the fiscal year 2002, BAAC disbursed loans directly to individual farmers worth 93.61 billion baht. During the year, repayments made to the bank amounted to 79.96 billion baht or 80.79 percent of the matured principal. By the end of fiscal year 2002, the principal outstanding was 241.72 billion baht (including credit services for community enterprises but not the special projects).

2. Credit services for community enterprises

   Credit services for community enterprises involved disbursing loans to BAAC client groups, groups or clubs of BAAC collective clients, farmers, members of farmer households or groups in general, which had at least 5 farmer members. These included collective groups or clubs for joint ventures, such as farm produce processing, activities involving agriculture, industry, commerce, services, artistry, handicrafts, knowledge and life quality development, vocation training and "micro-enterprises". The latter include BAAC clients who obtained loans for expenses either for agricultural-related or other activities to increase their income.

   BAAC promoted and supported operations of the community enterprises by the principle of *"Give more than Credit Lending"*, for example, giving advice, suggestions and support on marketing, production, finance accounting, and management, etc. As a result, the operations in fiscal year 2002, a total of 191,224 people had access to credit services from the bank worth some 11.426 billion baht.

3. Credit services for farmer institutions

The provision of credit services for farmer institutions is to make loans to agricultural cooperatives and farmer associations to use as a revolving fund so that farmers can more easily expand their operations. BAAC provided support in developing and strengthening farmer institutions in terms of skills, administration management and marketing. These include the following:

a) Loans to agricultural cooperatives

In fiscal year 2002, BAAC extended credit services to agricultural cooperatives worth 26.974 billion baht. During the year agricultural cooperatives repaid loans worth 26.88 billion baht to BAAC. By the end of 2002, the principal outstanding was 14.17 billion baht.

b) Loans to farmer associations

In fiscal year 2002, BAAC extended credit services to farmer associations worth 34 million baht. During the year, farmer associations repaid 37 million baht to BAAC. By the end of the year, the principal outstanding was 93 million baht.

4. Provision of credit services in the form of projects

In addition to normal credit services, BAAC extended credit services to farmers in the form of projects, which concentrated mainly on supervised credit. In cooperation with public and private agencies, BAAC provided technical, marketing and infrastructure construction support to help farmers improve their agriculture production. This assistance helped farmers increase production efficiency. The bank also participated in a feasibility study for special projects, taking in account technical, economic and social issues, administration management and marketing and finance. Once the operation was in progress, meetings were conducted with the farmers to explain the project and help them decide whether they should join it.

Credit operations in the form of projects are classified as projects under government policies and special projects of BAAC. These include the following:

a) Government projects initiated by "set policy or special cabinet resolutions" provided credit under special relaxed conditions to help farmers who are facing occupational hazards. For example, damages from disasters or natural calamities, which decrease output or quality level below normal standards or falling farm prices. Government assistance may be included in the provision of loans provided from BAAC on relaxed conditions. For example, low interest rates or interest compensation. Most of these credits were for production or marketing. Examples include the Loans for Postponement of the Sale of Farm Produce Scheme, and the Revolving Fund to Help Heavily Indebted Government Teachers Project.

By the end of 2002, BAAC extended credit services for the projects under government policies to some 2.72 million families of farmers who had joined the projects and the principal outstanding was 23.969 billion baht.

b) BAAC special projects extended credits for costs of investment and expenses of farm operations. This included providing credit services for the purchase of sugarcane fee discounted checks (fee checks), in cooperation between BAAC and public and private agencies. For this account, BAAC's Board of Directors had to approve and relax the bank's criteria of normal loan lending.

By the end of 2002, BAAC had implemented a total of 168 projects of this nature and disbursed a total of 10.142 billion baht in loans to serve 133,810 farm families. Of these projects, 166 were set up before 2002 and the repayment period is still outstanding, which includes a total of 3.708 billion baht in loans to serve 78,058 farm families. A total of 27 projects experienced operating problems, for example, lost output, marketing difficulties and low prices in dairy, para rubber, oil palm projects. BAAC restructured 17,664 farmers' debts worth 1.09 billion baht.

BAAC has generally maintained a timely loan repayment rate of about 85% with much of the arrears balance repaid later. Net income has increased steadily from 1989 to 1994 with a return on equity of 6.9-14.9 percent. In 1995-96, return on equity (ROE) was 3.3-4.0 percent; but during the 1997 financial crisis it dropped to 2.2 percent by end of June 1998. In 2003, the ROE of the bank was about 3.5 percent (Sakchuwong, 2004). Therefore, BAAC can be considered a unique rural credit institution since it has been operating profitably as a rural commercial bank (albeit lending only to farmers) and as an agency managing the implementation of a variety of directed credit rural development programs on behalf of government agencies (on an agency-fee basis) (ADB, 2002).

3.6 Characteristics of borrowers and loans

In this section, the information related to the borrowers and loans are discussed, including preliminary results and findings from the descriptive analysis (i.e. classification tables, compare means, correlation coefficients, etc.). There is no information about borrowers' current assets, current liabilities, and debt repayment available on the BAAC's Credit BPR database. As a result, *Current ratio* and *Debt repayment ratio* can not be calculated and are excluded from the research models.

Table 3.4 presents the descriptive statistics of the variables used in this research. The mean value of each variable is presented separately according to the lending sector. The t-test is used to test whether the mean values of two different sectors, agriculture and non-agriculture, are statistically different, whereas χ^2-test is for testing the relationship between the variables and lending sector. The statistical test results (both t-test and χ^2-test) are statistically significant at the 5 or 10 percent level, except for the age of the borrower and the value of collateral pledged to the bank.

Of the 18,798 new loans approved during 2001 to 2003, 16,202 (86.19 percent) can be considered as good loans. Since the loans are repaid, the borrowers can be regarded as good borrowers. The remaining, 2,596 (13.81 percent), are considered bad loans because the loans are default. The classification results indicate that non-agricultural loan has a higher default rate (20.40 percent) than agricultural loan (13.13 percent). Furthermore, on average, the amount of loan extended to the non-agricultural borrower is larger than the loan amount granted to the agricultural borrower. The non-agricultural loan is charged significantly lower than the agricultural loan (see Table 3.4).

Table 3.4: Characteristics of borrowers and loans classified by lending sector [1/]

Variable	Lending sector: Agriculture	Lending sector: Non-agriculture	Total	Statistical test[2/]
Dependent variables				
- Loan repayment				
Good loan	86.87	79.60	86.19	$\chi^2 = 71.19^{**}$
Bad loan	13.13	20.40	13.81	
Total	100.00	100.00	100.00	
- Volume of credit granted	133,866.54	206,856.00	140,739.15	$t = -10.39^{**}$
- Interest rate charged	0.0749	0.0676	0.0742	$t = 7.99^{**}$
Borrower characteristics				
- Asset	1,089,389.19	1,226,055.03	1,102,257.50	$t = -1.96^{*}$
- Age	50.04	50.23	50.06	$t = -0.72$
- Education				
≤ Primary school	89.52	87.06	89.29	$\chi^2 = 10.10^{**}$
> Primary school	10.48	12.94	10.71	
Total	100.00	100.00	100.00	
Credit risk proxies				
- Collateral	1,464,343.09	1,510,224.08	1,468,663.20	$t = -1.06$
- Return on asset[3/]	0.2709	0.3747	0.2808	$t = -3.59^{**}$
- Leverage ratio[3/]	0.0422	0.0290	0.0410	$t = 3.90^{**}$
- Capital turnover ratio[3/]	0.9000	1.1822	0.9269	$t = -4.36^{**}$
Relationship indicators				
- Borrowing from others				
Yes	17.91	13.39	17.48	$\chi^2 = 22.67^{**}$
No	82.09	86.61	82.52	
Total	100.00	100.00	100.00	
- Duration[4/]	2.17	3.03	2.26	$t = -10.82^{**}$
Dummy variables				
- Province				
Province 1	18.41	7.29	17.36	$\chi^2 = 1{,}226.87^{**}$
Province 2	8.09	17.85	9.01	
Province 3	2.91	0.90	2.72	
Province 4	2.14	0.96	2.03	
Province 5	6.82	3.67	6.53	
Province 6	3.46	3.50	3.47	
Province 7	2.29	2.88	2.35	
Province 8	6.24	11.19	6.71	
Province 9	5.13	7.29	5.34	
Province 10	10.71	3.45	10.02	
Province 11	1.82	0.00	1.65	
Province 12	4.09	1.98	3.89	
Province 13	6.05	0.45	5.52	
Province 14	0.41	0.73	0.44	
Province 15	2.48	0.56	2.30	
Province 16	13.76	36.05	15.86	
Province 17	5.17	1.24	4.80	
Total	100.00	100.00	100.00	

Table 3.4: Characteristics of borrowers and loans classified by lending sector [1/] (Cont)

Variable	Lending sector: Agriculture	Lending sector: Non-agriculture	Total	Statistical test[2/]
- Major production				
Horticulture	38.70	78.64	42.46	$\chi^2 = 1{,}105.69^{**}$
Orchard/Vegetable	27.57	16.27	26.50	
Livestock/Aquaculture	28.05	4.92	25.87	
Others	5.68	0.17	5.17	
Total	100.00	100.00	100.00	
- Loan type				
Cash credit loan	9.42	2.26	8.75	$\chi^2 = 396.99^{**}$
Short-term loan	8.69	0.11	7.88	
Medium-term loan	3.75	0.06	3.40	
Long-term loan	78.15	97.57	79.98	
Total	100.00	100.00	100.00	
- Loan size				
Small loan	62.53	45.65	60.94	$\chi^2 = 259.92^{**}$
Medium loan	36.97	51.86	38.37	
Large loan	0.50	2.49	0.69	
Total	100.00	100.00	100.00	
- Lending year				
2001	9.77	0.00	8.85	$\chi^2 = 300.15^{**}$
2002	47.72	40.23	47.01	
2003	42.52	59.77	44.14	
Total	100.00	100.00	100.00	
No. of observations	17,028	1,770	18,798	

Note: 1/ Classification results are in percentage and the results of the quantitative variables are mean values.
2/ T-test and χ^2-test are the equality test and test of independent, respectively.
3/ No. of observations for agricultural and non-agricultural loan are 16,560 and 1,750 samples, respectively. There are 18,310 observations in total.
4/ No. of observations for agricultural and non-agricultural loan are 3,974 and 479 samples, respectively. There are 4,453 observations in total.
* and ** represent 10% and 5% significant level, respectively.

The average total asset value and age of the BAAC borrowers are about 1,102,257 baht and 50 years old, respectively. Furthermore, approximately 89.29 percent of the borrowers received only primary education or less. There is no difference between the average age of the agricultural borrower and non-agricultural borrower. However, the education level and the average total asset value (wealth) of the non-agricultural borrower are slightly higher than the agricultural borrower.

All the loans in the usable data set are collateralized loans and the mean value of the collateral is about 1,468,663 baht. In addition, the mean value of collateral pledged to the bank is higher than the mean value of the borrower's total asset value. This is possible since the collateral include both "inside collateral" (assets of the borrower) and "outside collateral" (assets belong to another party). The mean value of the collateral is not significantly different according in the lending sector. However, on the average, the borrower who received non-agricultural loan is more profitable (0.37 versus 0.27 *Return on asset*), less leverage (0.03 versus 0.04 *Leverage ratio*), and high efficiency (1.18 versus 0.90 *Capital turnover ratio*).

Approximately 3,286 (17.48 percent) of the borrowers have outstanding debts with other lenders when they applied for a loan with BAAC. This implies that some borrowers have more than one debt. The mean value of *Duration* is 2.26 years. Thus, the average length of the bank-borrower relationship in the sample data is about 2 years[19]. The borrower who received a non-agricultural loan has a significantly longer relationship with the bank (about 3 years) and concentrated on one source of credit (only 13.39 percent has utilized more than one financial source).

The frequency distribution on the province in Table 3.4 shows that 18.41, 13.76 and 10.71 percent of the total observations on agricultural lending are in *Province 1*, *13*, and *10*, respectively, while 36.05, 17.85 and 11.19 percent of the total observations on non-agricultural lending are in *Province 16*, *2*, and *8*, respectively. This distribution indicates that the number and the type of loan (agricultural and non-agricultural loan) are dominated by some provinces, and therefore, the relationship between province and lending sector is not independent. The borrower who has *Horticulture* as the major production is considered

[19] Although some borrowers have a longer relationship with the bank, the maximum length of the bank-borrower relationship is only 7 years.

as a major borrower on both agricultural and non-agricultural lending. Furthermore, the samples on both lending sectors are dominated by *Long-term loan*. Overall, *Small loan*, *Medium loan*, and *Large loan* account for 60.94, 38.37, and 0.69 percent of the total number of loans, respectively. However, the majority of the non-agricultural loans is *Medium loan* (about 51.86 percent), whereas 62.53 percent of the agricultural loans is *Small loan*.

Table 3.5 shows the characteristics of borrowers and loans classified by loan performance (good/bad loan). The mean values on the volume of credit granted and the interest rate charged show that the bad loan (default borrower) receives more credit and is charged slightly higher than the good loan. Furthermore, on average, the bad loan has a higher *Return on asset* and *Capital turnover ratio* (more profitable and efficiency) than the good loan. The average length of the bank-borrower relationship of the bad loan is longer than the good loan. However, the good loan has a higher asset value (wealth) than the bad loan and the collateral value pledged by the good loan is lower than the bad loan. In addition, the mean value of *Leverage ratio* for the good loan is lower than for the bad loan.

The classification results and the χ^2-tests shown in Table 3.5 also indicate that *Borrowing from others* and the loan performance (good/bad loan) are independent. In contrast, the loan performance is not independent with *Education*, *Province*, *Major production*, *Loan type*, *Loan size* and *Lending year*, since the χ^2-tests on these variables are all significant at the 5 percent level. The borrower with primary school or below has a higher probability to default on the loan repayment than the borrower who has a high school qualification and above (18.92 percent versus 13.20 percent). The borrower with *Long-term loan* and *Large loan* is more likely to default on the loan repayment.

The mean values of the volume of credit granted (loan amount) and interest rate charged (loan price) are calculated and compared according to the groups of each variable presented in equation 3.1 (see Table 3.6). The independent-sample t-test and one-way ANOVA (F-test) are used to test the significant mean differences between two different groups and among multiple groups, respectively. The test results in Table 3.6 show that the loan amount and loan price vary with *Education, Borrowing from others*, *Province*, *Major production*, *Loan type*, *Loan size*, and *Lending year*.

Table 3.5: Characteristics of borrowers and loans classified by loan performance [1/]

Variable	Loan performance		Total	Statistical test[2/]
	Good loan	Bad loan		
Dependent variables				
Volume of credit granted	134,015.52	182,702.28	140,739.15	t = -9.68**
Interest rate charged	0.0740	0.0755	0.0742	t = -1.66*
Borrower characteristics				
Asset	1,112,507.60	1,038,285.20	1,102,257.50	t = 1.66*
Age	50.20	49.15	50.06	t = 4.91**
Education				
≤ Primary school	81.08	18.92	100.00	χ^2 = 49.44**
> Primary school	86.80	13.20	100.00	
Total	86.19	13.81	100.00	
Credit risk proxies				
Collateral	1,458,804.85	1,530,190.54	1,468,663.20	t = -1.72*
Return on asset[3/]	0.2595	0.4132	0.2808	t = -4.92**
Leverage ratio[3/]	0.0374	0.0630	0.0410	t = -4.61**
Capital turnover ratio[3/]	0.8240	1.5672	0.9269	t = -9.16**
Relationship indicators				
Borrowing from others				
Yes	86.24	13.76	100.00	χ^2 = 0.01
No	86.18	13.82	100.00	
Total	86.19	13.81	100.00	
Duration[4/]	2.17	2.69	2.26	t = -8.38**
Dummy variables				
Province				
Province 1	95.19	4.81	100.00	χ^2 = 912.09**
Province 2	96.22	3.78	100.00	
Province 3	82.23	17.77	100.00	
Province 4	75.13	24.87	100.00	
Province 5	87.94	12.06	100.00	
Province 6	76.07	23.93	100.00	
Province 7	84.58	15.42	100.00	
Province 8	85.25	14.75	100.00	
Province 9	86.54	13.46	100.00	
Province 10	90.50	9.50	100.00	
Province 11	70.00	30.00	100.00	
Province 12	75.00	25.00	100.00	
Province 13	87.86	12.14	100.00	
Province 14	91.46	8.54	100.00	
Province 15	66.97	33.03	100.00	
Province 16	79.20	20.80	100.00	
Province 17	83.92	16.08	100.00	
Total	86.19	13.81	100.00	

Table 3.5: Characteristics of borrowers and loans classified by loan performance [1/] (Cont)

Variable	Loan performance: Good loan	Loan performance: Bad loan	Total	Statistical test[2/]
Major production				
Horticulture	87.89	12.11	100.00	$\chi^2 = 264.94^{**}$
Orchard/Vegetable	88.26	11.74	100.00	
Livestock/Aquaculture	84.54	15.46	100.00	
Others	69.93	30.07	100.00	
Total	86.19	13.81	100.00	
Loan type				
Cash credit loan	92.58	7.42	100.00	$\chi^2 = 171.05^{**}$
Short-term loan	93.99	6.01	100.00	
Medium-term loan	89.05	10.95	100.00	
Long-term loan	84.60	15.40	100.00	
Total	86.19	13.81	100.00	
Loan size				
Small loan	88.87	11.13	100.00	$\chi^2 = 184.77^{**}$
Medium loan	82.16	17.84	100.00	
Large loan	73.64	26.36	100.00	
Total	86.19	13.81	100.00	
Lending year				
2001	88.27	11.73	100.00	$\chi^2 = 27.38^{**}$
2002	84.83	15.17	100.00	
2003	87.23	12.77	100.00	
Total	86.19	13.81	100.00	
No. of observations			18,798	

Note: 1/ Classification results are in percentage and the results of the quantitative variables are mean values.
2/ The t-test and χ^2-test are the equality test and test of independent, respectively.
3/ No. of observations is 18,310.
4/ No. of observations is 4,453.
* and ** represent 10% and 5% significant level, respectively.

Table 3.6: The mean values of volume of credit granted and interest rate charged

Variable	Volume of credit granted	Interest rate charged
Education		
≤ Primary school	133,738.19	0.0733
> Primary school	199,082.82	0.0816
t-test	**-11.27****	**-10.67****
Borrowing from others		
Yes	117,864.11	0.0712
No	145,584.91	0.0748
t-test	**-7.79****	**-5.75****
Province		
Province 1	71,408.85	0.0693
Province 2	93,614.83	0.0734
Province 3	104,992.36	0.0687
Province 4	120,947.58	0.0715
Province 5	86,574.12	0.0709
Province 6	111,596.37	0.0683
Province 7	73,997.71	0.0770
Province 8	78,148.86	0.0678
Province 9	81,296.34	0.0670
Province 10	195,807.71	0.0817
Province 11	276,429.14	0.0848
Province 12	171,975.73	0.0648
Province 13	217,487.16	0.0801
Province 14	243,182.21	0.0796
Province 15	196,100.24	0.0828
Province 16	238,878.20	0.0791
Province 17	154,221.02	0.0829
F-test	**165.64****	**43.97****
Major production		
Horticulture	113,211.55	0.0692
Orchard/Vegetable	173,244.77	0.0805
Livestock/Aquaculture	140,051.90	0.0780
Others	202,846.99	0.0638
F-test	**107.53****	**153.71****
Loan type		
Cash credit loan	97,594.80	0.0835
Short-term loan	32,781.40	0.0743
Medium-term loan	114,820.19	0.0841
Long-term loan	156,803.42	0.0727
F-test	**177.46****	**64.66****
Loan size		
Small	52,962.23	0.0685
Medium	253,888.21	0.0831
Large	1,609,175.49	0.0888
F-test	**19,433.04****	**529.71****
Lending year		
2001	170,225.63	0.0837
2002	128,966.19	0.0765
2003	147,367.45	0.0699
F-test	**42.10****	**184.51****

Note: * and ** represent 10% and 5% significant level, respectively.

The borrower who borrows only from BAAC and who has a higher education could obtain a larger amount of loan from the bank (see Table 3.6). However, they are charged a higher rate. Thus, the result on the loan price contradicts the hypothesis. This is because the average values of the interest rate charged are calculated without controlling for the impacts of the other factors that might influence the loan rate.

The correlation matrix in Table 3.7 illustrates the volume of credit granted has a positive relationship with *Asset*, *Collateral*, *Capital turnover ratio* and *Duration*. In contrast, it is negatively related to *Age*. The low and insignificant correlation coefficients between the volume of credit granted and *Return on asset* (r = 0.00), and between the volume of credit granted and *Leverage ratio* (r = -0.01) imply that both variables (*Return on asset* and *Leverage ratio*) do not have a linear association with the amount of credit granted. The results do not mean that they are independent, since they might have a non-linear relationship with the loan amount.

The correlation coefficients also show that *Age*, *Return on asset*, *Capital turnover ratio* and *Duration* are negatively related to the loan price (see Table 3.7). However, it is found that *Asset* and *Collateral* are positively related to the interest rate charged, whereas the *Leverage ratio* shows a negative relationship. The correlation coefficient does not control for the other factors' influences, and further investigation should be conducted to examine these relationships.

Table 3.7: Correlation matrix

	Volume	Interest	Asset	Age	Collateral	ROA	LR	CTR	Duration
Volume	1.00								
Interest	0.19**	1.00							
Asset	0.36**	0.04**	1.00						
Age	-0.01**	-0.09**	0.11**	1.00					
Collateral	0.49**	0.08**	0.50**	0.07**	1.00				
ROA	0.00	-0.04**	-0.07**	-0.06**	-0.02**	1.00			
LR	-0.01	-0.03**	-0.06**	-0.03**	-0.01	0.03**	1.00		
CTR	0.02**	-0.04**	-0.09**	-0.10**	0.01	0.72**	0.03**	1.00	
Duration	0.06**	-0.39**	0.12**	0.10**	0.08**	0.11**	0.02	0.17**	1.00

Note: Volume = Volume of credit granted.
Interest = Interest rate charged.
ROA = Return on asset.
LR = Leverage ratio.
CTR = Capital turnover ratio.
* and ** represent 10% and 5% significant level, respectively.

CHAPTER 4

RESEARCH RESULTS AND FINDINGS

This chapter discusses the empirical results on the bank lending decision model, credit availability model and loan pricing model, and the critical factors affecting the lending decision, amount of credit and loan price. The estimated results are expected to reflect the borrower's risk level, and the bank-borrower relationship should not dominate the bank lending decision, the amount of credit availability, and the lending rate.

The chapter is organized as follows: Section 4.1 discusses the estimated results on the bank lending decision model. In addition, both the logistic regression and the artificial neural networks classification outputs are computed and compared. Section 4.2 and 4.3 discuss the estimated results on credit availability models and loan pricing models, respectively. The predictive power of the multiple regression analysis technique and the artificial neural networks on both the credit availability and the loan pricing model are also evaluated. Section 4.4 summarizes the findings.

4.1 Bank lending decision model

The bank lending decision models are estimated using the logistic regression via the maximum likelihood estimation technique. The estimated results of the bank lending decision models are presented in Table 4.1. In general, both models (with and without *Duration* variable) fit the data quite well. The coefficients in both models are significantly different from zero at the 10 percent level. The chi-square statistic using a likelihood ratio test on both models fail to accept the null hypothesis that the parameter estimates for the models are equal to zero. Model 1 (without *Duration*) and Model 2 (with *Duration*) correctly predict bank lending decision 86.51 and 84.29 percent, respectively. However, it should be noted that Model 1 and 2 have produced a 93.61 and 89.52 percent Type I error (wrongly reject H_0 or accepting a bad loan as a good loan), and 0.61 and 1.08 percent Type II error (wrongly accept H_0 or rejecting a good loan as a bad loan), respectively. Although Model 1 has a higher overall percentage correct and a lower percentage of Type II error, Model 2 can predict the bad loan group better than Model 1 and has a lower percentage of Type I error[20].

The significant positive sign on *Sector* dummy variable indicates that the agricultural loan has a higher probability of a good loan than the non-agricultural loan, holding other factors constant, and this could lead BAAC to have a higher preference to lend to the agricultural sector than the non-agricultural sector[21] (see Table 4.1). In contrast, the significant negative coefficients on *Medium-term loan*, *Long-term loan*, *Medium loan* and *Large loan* dummy variables show that default risk on the loan increases with the size and the period of loan, since a higher negative value indicates a higher default risk on the loan or a lower probability of a good loan. Furthermore, the dummy variables on the major production show that the borrowers who have a cash crop (*Horticulture*) as a major production crop have a higher probability of a good loan. This is because horticultural productions involve a small amount of credit and mainly are short-term loans, unlike livestock and aquaculture productions, which are medium-term or long-term investments requiring a relatively high investment funding.

The estimated coefficients of the province dummy variables are significant different from zero. A negative (or a positive) value implies of a decrease (or an increase) in the probability of a good loan relative to the other provinces. For example, in Table 4.1, Model 2, *Province 4* (coefficient value = -3.0082) has a lower probability of a good loan than *Province 5* (coefficient value = -1.2678) on the average, assuming other factors are held constant. The results show the effect of provincial differences on the loan performance and the probability of a good loan. Therefore, it is necessary for the bank to use the different lending criteria on the different provinces, and this should not be claimed as a provincial discrimination. This is because different provinces have different economic, social, environmental, physical, and biological conditions, and there are also differences in the production system and risk, which all influence the loan performance and the default risk.

[20] It is generally accepted that the misclassification cost of Type I error is more costly than Type II error. For Type I error, the lending bank may not lose only the principal but also the interest on the principal. On the other hand, for Type II error, the lending bank loses only the interest and expected profit that may be obtained from lending to the borrower.

[21] It should be noted here that BAAC is considered as a government bank specifically for agriculture borrowing. BAAC is also involved in non-agricultural loan since 2001.

Table 4.1: Bank lending decision models

Independent variables[1/]	Coefficients[2/]	
	Model 1	Model 2
Borrower Characteristics		
Log(asset)	0.3005**	0.3568**
Age	-0.0019	-0.0021
Education	0.1372**	0.1268
Credit risk proxies		
Log(collateral)	-0.0278	-0.0741
Return on asset	0.0510*	0.0516
Leverage ratio	-0.8957**	-0.3424
Capital turnover ratio	-0.0734**	-0.0802**
Relationship indicators		
Borrowing from others	0.0928	-0.0784
Duration		-0.1555**
Dummy variables[3/]		
Sector	0.4620**	0.7978**
(Province)		
Province 2	0.3612**	-0.4844
Province 3	-1.2037**	-0.8204*
Province 4	-1.8634**	-3.0082**
Province 5	-1.2186**	-1.2678**
Province 6	-1.5655**	-2.2047**
Province 7	-0.9760**	-0.7551**
Province 8	-1.1367**	-1.1496**
Province 9	-1.1047**	-1.4100**
Province 10	-1.1636**	-1.8059**
Province 11	-2.9225**	-2.7329**
Province 12	-1.8796**	-0.7683
Province 13	-1.1986**	-1.4572**
Province 14	-0.9118**	-2.0319**
Province 15	-2.2798**	-2.3291**
Province 16	-1.8834**	-1.9453**
Province 17	-1.5190**	-1.7594**

Table 4.1: Bank lending decision models (Cont)

Independent variables[1/]	Coefficients[2/] Model 1	Model 2
(Major production)		
Horticulture	0.8285**	0.8033**
Orchard/Vegetable	0.8682**	0.5804**
Livestock/Aquaculture	0.7163**	0.3850**
(Loan type)		
Short-term loan	-0.2112	-0.3968
Medium-term loan	-0.2018	-0.8016**
Long-term loan	-0.7279**	-1.1433**
(Loan size)		
Medium loan	-0.3695**	-0.2855**
Large loan	-0.4836**	-0.7328*
(Lending year)		
2001	0.3243**	0.0117
2002	-0.4195**	-0.3729**
Constant	-0.6233	0.1901
No. of Observations	18,310	4,444
Log likelihood	-6,510.05	-1,755.37
LR statistic	1,710.00**	475.59**
df	35	36
McFadden R^2	0.1161	0.1193

Classification table[4/]

	Model 1			Model 2		
	BL	GL	Overall	BL	GL	Overall
Correct	162	15,678	15,840	77	3,669	3,746
	6.39%	99.39%	86.51%	10.48%	98.92%	84.29%
Incorrect	2,374	96	2,470	658	40	698
	93.61%	0.61%	13.49%	89.52%	1.08%	15.71%

Note: 1/ Dependent variable is bank lending decision (good/bad loan).
2/ Maximize using logistic likelihood function and quasi-maximum likelihood (QML) standard errors and covariance.
3/ To avoid singularity problem, a dummy variable is dropped from each group.
4/ BL and GL stand for bad loan and good loan, respectively.
*, ** represent 10% and 5% significant level, respectively.

In 2001, the Thai government under Prime Minister Dr. Thaksin Shinawatra and his Thai Rak Thai Party (TRT) implemented a three-year debt suspension programme. Under the program, farmers with debt of up to 100,000 baht did not need to repay their BAAC debts and interest for three years. They were also given training in marketing, diversifying and finding supplemental income sources. However, they could not borrow more from the BAAC during this period (Bangkok Post, 2000 and 2002). There were about 2.19 million farmers (holding about 74 percent of the BAAC's debts) with debts totalling 87.70 billion baht participated in the programme. About 1.14 million farmers with debts totalling 51.11 billion baht were in the debt suspension programme and another 1.05 million farmers with 36.59 billion baht debt in the debt-reduction scheme. Therefore, it is not surprising that the dummy variable for the 2002 lending year on both Model 1 and 2 are negative, since some previous good debtors were encouraged to default on debt repayments in anticipation of the three-year debt moratorium programme (see Table 4.1).

In Model 1 (without *Duration*), only *Asset*, *Education*, *Return on asset*, *Leverage ratio* and *Capital turnover ratio* are significant at the 10 percent level (see Table 4.1). As expected, the probability of a good loan increases with increased *Asset*, *Education* and *Return on asset*. However, the probability of a good loan does not decrease with only increased *Leverage ratio* but also *Capital turnover ratio*. The findings contradict the hypothesis on *Capital turnover ratio* showing that the borrower who has a higher gross income to total assets has a higher probability of defaulting on debt repayment. It implies that when the borrower has earned more, they prefer to spend their money for other activities or purposes rather than repaying their debt. The estimated signs on *Age* and *Borrowing from others* contradict the hypothesised signs, but they are not significantly different from zero. In addition, the results from Model 1 show that *Collateral* has no significant impact on the bank lending decision.

When *Duration* is included in the Model 2, the estimated results showed *Asset* and *Capital turnover ratio*, are significantly different from zero at the 10 percent level, while *Education*, *Return on asset* and *Leverage ratio* are insignificant (see Table 4.1). Furthermore, the estimated coefficient on *Capital turnover ratio* is negative which is consistent with the estimated result in Model 1. However, the relationship between *Duration* and the bank lending decision contradict the postulated hypothesis. The estimated coefficient is negative and significant at the 95 percent confident level. This suggests that the borrower who has a longer relationship with the bank has a higher probability of defaulting on debt repayment and the bank should cautiously deal with this group of borrowers. The coefficient of *Borrowing from others* is negative in Model 2 and not significant. Consequently, the results imply that multiple financial source have no influence on the bank lending decision.

In this research, *Duration* is restricted to a maximum of 7 years. Some borrowers may have a longer relationship with the bank, but there is no available information to estimate the length of the relationship. Therefore, *Duration* is a censoring variable and it may induce inconsistent estimates of the length of the bank-borrower relationship (Ongena and Smith, 2000).

The results in Table 4.1 are consistent with Cole (1998) findings who argued that the bank is more likely to extend credit to large borrowers than small borrowers. The estimated coefficients on *Borrowing from others* showed that borrowing from other sources of funds does not have a significant impact on the bank lending decision. This finding contradicts Cole (1998) and Harhoff and Korting (1998) findings, who concluded that the multiple source borrowings are less likely to get credit. The *Duration* has a negative impact on the bank lending decision, which contradicts the postulated hypothesis. Cole (1998) and Bodenhorn (2003) found a positive relationship between the length of relationship and probability of loan approval. However, the empirical result supports the theory that the bank-borrower relationship generates private information useful for the lender in assessing the borrower's creditworthiness.

4.1.1 Bank lending decision models for agricultural and non-agricultural lending

Agricultural lending and non-agricultural lending differ in their lending behaviours. Most of the agricultural products are seasonal, perishable, and difficult to store. Moreover, the production depends on the natural environment condition. Thus, agricultural lending can be considered risky, compared to non-agricultural lending, and the banks may apply different lending decision criterion on agricultural lending versus non-agricultural lending. The *Sector* coefficient is found to be positive and significant at the 5 percent level (see Table 4.1). This implies that the bank lending decision is impacted by the lending sectors.

The estimated results on bank lending decision models for agricultural lending are shown in Table 4.2. *Sector* is discarded from the model because the model is segregated according to the lending sectors. The chi-square tests on both Model 1 and 2 strongly reject the hypothesis of no explanatory power and the models accurately predict 87.19 and 85.30 percent of the total observations, respectively. Model 1 (without *Duration*) has a higher overall percentage correct classification, but it has a larger Type I error than Model 2 (with *Duration*). In addition, the overall percentage correct classification of both models (with and without *Duration*) improved slightly when compare to the estimated results of the aggregate lending (both agricultural and non-agricultural lending information) (see Table 4.1).

The estimated results on the bank lending decision models for agricultural lending (see Table 4.2) yield similar outcomes and conclusions as the models for aggregate lending (see Table 4.1). The results indicate that horticultural production, short-term loan, and small borrowing are more stable and have a lower default risk. Furthermore, the estimated coefficients of the provinces show that the probability of a good loan differ according to the residential province. The significant negative sign on the year 2002 dummy variable signals an abnormal default rate on debt repayment from the farmers, since the introduction of the debt moratorium programme in 2002.

In Model 1, only 4 variables from *Borrower characteristics*, *Credit risk proxies* and *Relationship indicators* are significant at the 5 percent level of significance (see Table 4.2). They are *Asset*, *Education*, *Leverage ratio*, and *Capital turnover ratio*. The estimated coefficients exhibit that both *Asset* and *Education* have a positive relationship with the bank lending decision, while *Leverage ratio* and *Capital turnover ratio* are negatively correlated with the bank lending decision. When *Duration* is added to the model (see Model 2), the estimated outputs reinforce the imperative of *Asset*, *Capital turnover ratio* and *Duration*. The estimated sign is positive on *Asset* but negative on *Capital turnover ratio* and *Duration*, and all of them are significantly different from zero at the 5 percent level of significance.

Table 4.2: Bank lending decision models for agricultural lending

Independent variables[1/]	Coefficients[2/]	
	Model 1	Model 2
Borrower Characteristics		
Log(asset)	0.3197**	0.3719**
Age	-0.0009	-0.0016
Education	0.1686**	0.1769
Credit risk proxies		
Log(collateral)	-0.0339	-0.0689
Return on asset	0.0383	0.0050
Leverage ratio	-0.9629**	-0.8326
Capital turnover ratio	-0.0634**	-0.0596**
Relationship indicators		
Borrowing from others	0.1081	0.0329
Duration		-0.1915**
Dummy variables[3/]		
(Province)		
Province 2	0.3397*	-0.1766
Province 3	-1.2561**	-0.8480*
Province 4	-1.8665**	-2.6726**
Province 5	-1.2657**	-1.2191**
Province 6	-1.5949**	-2.1101**
Province 7	-1.0039**	-0.7339*
Province 8	-1.2343**	-1.1030**
Province 9	-1.1283**	-1.4990**
Province 10	-1.1496**	-1.6164**
Province 11	-2.7958**	-2.4990**
Province 12	-1.9438**	-0.6575
Province 13	-1.1482**	-1.2083**
Province 14	-0.8085*	-2.4747**
Province 15	-2.2483**	-2.1955**
Province 16	-1.6257**	-1.7089**
Province 17	-1.5205**	-1.6943**

Table 4.2: Bank lending decision models for agricultural lending (Cont)

Independent variables[1/]	Coefficients[2/]	
	Model 1	Model 2
(Major production)		
Horticulture	0.8767**	0.8959**
Orchard/Vegetable	0.8425**	0.4711**
Livestock/Aquaculture	0.6909**	0.3031
(Loan type)		
Short-term loan	-0.2789*	-0.4525
Medium-term loan	-0.2832	-0.8265**
Long-term loan	-0.7631**	-1.1379**
(Loan size)		
Medium loan	-0.4016**	-0.3559**
Large loan	-0.6262**	-1.7228**
(Lending year)		
2001	0.1507	-0.0585
2002	-0.3757**	-0.3479**
Constant	-0.4007	0.6708

No. of Observations	16,560	3,965
Log likelihood	-5,720.45	-1,489.09
LR statistic	1,446.85**	398.97**
df	34	35
McFadden R^2	0.1123	0.1182

Classification table[4/]

	Model 1			Model 2		
	BL	GL	Overall	BL	GL	Overall
Correct	131	14,308	14,439	56	3,326	3,382
	6.02%	99.48%	87.19%	9.30%	98.90%	85.30%
Incorrect	2,046	75	2,121	546	37	583
	93.98%	0.52%	12.81%	90.70%	1.10%	14.70%

Note: 1/ Dependent variable is bank lending decision (good/bad loan).
2/ Maximize using logistic likelihood function and quasi-maximum likelihood (QML) standard errors and covariance.
3/ To avoid singularity problem, a dummy variable is dropped from each group.
4/ BL and GL stand for bad loan and good loan, respectively.
*, ** represent 10% and 5% significant level, respectively.

The results in Model 2 (see Table 4.2) suggest that *Age*, *Education*, *Collateral*, *Return on asset*, *Leverage ratio* and *Borrowing from others* have no significant impact on the bank lending decision for agricultural lending. Since agricultural lending accounts for 90.44 and 89.22 percent of the total observations in Model 1 and 2 of the aggregate lending, respectively, the estimated outputs on the bank lending decision models for the aggregate and agricultural lending yield almost identical conclusions.

Table 4.3 presents the estimated results on the bank lending decision models for non-agricultural lending. *Sector*, *Province 11* and the *2001* lending year are excluded from the model because the model is segregated according to the lending sectors, and there is no information available either on *Province 11* or the *2001* lending year. The dummy variable for long-term loan is included in the loan type group, since the majority of non-agricultural lending is long-term loan. There are only 3 different types of the borrower major production on non-agricultural lending. They are *Horticulture*, *Orchard/Vegetable*, and *Livestock/Aquaculture*[22]. The *Horticulture* dummy variable is removed from the major production group to avoid the dummy trap problem.

The overall percentage correct prediction of Model 1 and 2 are 81.14 and 76.62 percent, respectively (see Table 4.3). However, the predictive powers are low when compared to the lending decision models for aggregate lending (see Table 4.1) and agricultural lending (see Table 4.2). However, the lending decision model for non-agricultural lending can detect a Type I error relatively better than the models for aggregate lending and agricultural lending.

The likelihood ratio (LR) on both Model 1 and 2 fail to accept the null hypothesis at the 5 percent level of significance (see Table 4.3). It can be concluded that both models have significant explanatory power and they can be used to explain the bank lending decision for non-agricultural loans.

[22] There were only 3 observations in "Others" category or about 0.17 percent of the total observations on non-agricultural lending.

Table 4.3: Bank lending decision models for non-agricultural lending

Independent variables[1/]	Coefficients[2/]	
	Model 1	Model 2
Borrower Characteristics		
Log(asset)	0.2308**	0.2009
Age	-0.0027	0.0021
Education	-0.0605	-0.1084
Credit risk proxies		
Log(collateral)	-0.0635	0.0076
Return on asset	0.1657*	0.3395**
Leverage ratio	-0.1484	0.1706
Capital turnover ratio	-0.1537**	-0.2333**
Relationship indicators		
Borrowing from others	-0.1003	-0.1489
Duration		0.0661
Dummy variables[3/]		
(Province)		
Province 1	2.5775**	2.8699**
Province 2	2.7023**	1.2465**
Province 3	1.9517**	34.0266
Province 4	0.2169	-33.7642
Province 5	1.8097**	0.6444
Province 6	1.2126**	0.3365
Province 7	2.1108**	2.2376**
Province 8	1.8403**	1.3142**
Province 9	1.7263**	1.9697**
Province 10	2.1065**	0.2664
Province 12	2.1261**	1.5666
Province 13	0.6784	-0.2011
Province 14	1.3843	33.4604
Province 15	0.7552	0.9524
Province 17	1.8678**	1.0726

Table 4.3: Bank lending decision models for non-agricultural lending (Cont)

Independent variables[1/]	Coefficients[2/] Model 1	Model 2
(Major production)		
Orchard/Vegetable	-0.5669**	0.1145
Livestock/Aquaculture	0.3045	0.4481
(Loan type)		
Long-term loan	0.8353*	0.3219
(Loan size)		
Medium loan	-0.1650	-0.0233
Large loan	0.1225	0.1664
(Lending year)		
2002	-0.7444**	-0.6047**
Constant	-1.8880	-2.5336

	Model 1	Model 2
No. of Observations	1,750	479
Log likelihood	-735.61	-240.85
LR statistic	306.10**	84.28**
df	29	30
McFadden R^2	0.1716	0.1488

Classification table[4/]

	Model 1			Model 2		
	BL	GL	Overall	BL	GL	Overall
Correct	66	1,354	1,420	31	336	367
	18.38%	97.34%	81.14%	23.31%	97.11%	76.62%
Incorrect	293	37	330	102	10	112
	81.62%	2.66%	18.86%	76.69%	2.89%	23.38%

Note: 1/ Dependent variable is bank lending decision (good/bad loan).
2/ Maximize using logistic likelihood function and quasi-maximum likelihood (QML) standard errors and covariance.
3/ To avoid singularity problem, a dummy variable is dropped from each group.
4/ BL and GL stand for bad loan and good loan, respectively.
*, ** represent 10% and 5% significant level, respectively.

The estimated coefficients in Model 1 are significantly different from zero at the 5 percent level, while only some variables are significant in Model 2 (see Table 4.3). The insignificant variables in Model 2 may be caused by the small sample size, since there are only 479 observations on non-agriculture lending. The results in Model 1 show the significant impacts of *Asset*, *Return on asset* and *Capital turnover ratio* on non-agricultural lending decision.

When *Duration* is included in the model (see Model 2 in Table 4.3), *Asset* becomes insignificant, and only *Return on asset* and *Capital turnover ratio* have a significant influence on the bank lending decision. However, it is found that *Duration* is not significantly different from zero. The positive sign on *Return on asset* indicates that the probability of a good loan for non-agricultural loan increases with increased *Return on asset*. *Capital turnover ratio*, in contrast, has a negative sign. The result implies that the borrowers with a higher *Capital turnover ratio* are more likely to default on debt repayment, and therefore, the probability that their loan contact will be approved would be lower or decreased. Although the estimated coefficient on *Capital turnover ratio* contradicts the priori hypothesized sign, the result are consistent with the lending decision models for aggregate lending (see Table 4.1) and agricultural lending (see Table 4.2).

Table 4.4 presents the marginal effects of the bank lending decision models. The marginal effect of *Sector* in Model 2 shows the probability of a good loan on agricultural loan increases by 11.33 percent. The result indicates that the borrower can access agricultural credit from BAAC easier than non-agricultural credit. Furthermore, it also implies that the default probability of the agricultural loan contract is lower than non-agricultural loan contract by 11.33 percent on average.

The marginal effects for agricultural lending show that a 1 percent increases in *Asset* and *Capital turnover ratio* would change the probability of a good loan by 0.0387 and -0.0062 percent, respectively (see Table 4.4, Model 2). Furthermore, the marginal effect on *Duration* shows that, when the relationship between the bank and the borrower increase by 1 year, the probability of a good loan (or default) would decrease (or increase) by 1.99 percent on the average.

Table 4.4: Marginal effects of bank lending decision models

Independent Variables[1/]	Marginal Effects[2/]					
	Aggregate		Agriculture		Non-agriculture	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Borrower Characteristics						
Log(asset)	***0.0285***	***0.0406***	***0.0289***	***0.0387***	***0.0313***	0.0302
Age	-0.0002	-0.0002	-0.0001	-0.0002	-0.0004	0.0003
Education	***0.0136***	0.0147	***0.0161***	0.0190	-0.0081	-0.0161
Credit risk proxies						
Log(collateral)	-0.0026	-0.0084	-0.0030	-0.0072	-0.0086	0.0011
Return on asset	***0.0045***	0.0059	0.0030	0.0005	***0.0226***	***0.0511***
Leverage ratio	***-0.0851***	-0.0389	***-0.0874***	-0.0868	-0.0193	0.0257
Capital turnover ratio	***-0.0073***	***-0.0091***	***-0.0060***	***-0.0062***	***-0.0208***	***-0.0351***
Relationship indicators						
Borrowing from others	0.0086	-0.0091	0.0095	0.0034	-0.0138	-0.0232
Duration		***-0.0177***		***-0.0199***		0.0099
Dummy variables[3/]						
Sector	***0.0512***	***0.1133***				
(Province)	Yes	Yes	Yes	Yes	Yes	Yes
(Major production)	Yes	Yes	Yes	Yes	Yes	Yes
(Loan type)	Yes	Yes	Yes	Yes	Yes	Yes
(Loan size)	Yes	Yes	Yes	Yes	Yes	Yes
(Lending year)						
2001	***0.0279***	0.0013	0.0132	-0.0062		
2002	***-0.0413***	***-0.0443***	***-0.0352***	***-0.0377***	***-0.1054***	***-0.0998***
No. of observations	18,310	4,444	16,560	3,965	1,750	479

Note: 1/ Dependent variable is bank lending decision (good/bad loan).
2/ Marginal effect is at the mean value. For dummy variable, marginal effect is P|1 - P|0.
3/ See Appendix 1 for the marginal effects on province, major production, loan type, and loan size dummy variables.
Bold and italic represent 10% significant level or below.

The marginal effects for non-agricultural lending show in Table 4.4 (see Model 2) indicate that if *Return on asset* increases by 1 percent, the probability of a good loan would increase by 0.0511 percent. On the other hand, the probability of a good loan would decrease by 0.0351 percent if the *Capital turnover ratio* increases by 1 percent.

In summary, the results from the logistic regression models reveal that the factors determining bank lending decision are influenced by lending sectors (agriculture or non-agriculture). Borrowing for agricultural production or borrowing related to agriculture activities would have a higher probability of a good loan. Both *Asset* and *Capital turnover ratio* play a very crucial role on the bank lending decision models for both agricultural and non-agricultural lending. Furthermore, *Duration* has a significant impact on the bank lending decision only on agricultural lending, while *Return on asset* is significant only on non-agricultural lending decision. This suggests that the bank places heavy emphasis on *Return on asset* only on non-agricultural lending, but not on agricultural lending. However, both *Capital turnover ratio* and *Duration* are negatively correlated with the bank lending decision, which contradict the hypotheses. The estimated parameter on the 2002 lending year dummy variable indicates that the default risk in 2002 is higher than other years. The higher default risk in 2002 on both agricultural and non-agricultural lending might be partly instigated by the debt suspension programme of the Thai government.

4.1.2 Artificial neural networks and bank lending decision models

In this research, four different types of the artificial neural networks (ANN) are applied to model the bank lending decision. This includes multiple-layer feed-forward neural network (MLFN), Ward network (WN), general regression neural network (GRNN), and probabilistic neural network (PNN). These networks are considered supervised networks.

The first network, MLFN with one hidden layer is the fundamental network in the family of the ANN and has been widely used in many disciplines, including biology, psychology, statistics, mathematics, medical science, and computer science. Recently, networks have been applied to a variety of business areas such as accounting and auditing, finance, management, marketing, and production. The second network, WD (multiple hidden slabs with different activation functions) was invented by Ward Systems Group Inc (1993). The logistic function is used as the activation functions for all MLFN models and WN uses the 3 hidden slabs structure that has Gaussian, Tanh, and Gaussian Complement as the activation functions for 3 hidden slabs. The last two networks, GRNN and PNN, are special class of the ANN invented by Specht (1990 and 1991). The first two networks, MLFN and WD can handle both classification and prediction problems. The PNN is normally applied to classification problem. However, instead of producing continuous valued outputs, GRNN can categorise data like PNN. The networks can be used to model the bank lending decision.

Since the neural network model is nonlinear and their training process is always regarded as a black-box, it is very difficult to write out the algebraic relationship between a dependent variable and an independent variable. Furthermore, the learned outputs, connection weights or coefficients, can not be interpreted and tested. Therefore, only the classification results of the models, and the relative contribution factors of some selected methods, are presented in this research.

The ANN models will use the same data set and the same set of independent variables used in the logistic regression models. The number of hidden neurons on both MLFN and WD is set up to the default value given by the NeuroShell2 package, which can be computed with the following formula[23]:

$$\text{No. of hidden neurons} = \frac{1}{2}(\text{Inputs} + \text{Outputs}) + (\text{No. of training patterns})^{1/2} \quad (4.1)$$

In case of WN, which has 3 hidden slabs, the number of hidden neurons in each slab is equal to the number given by equation 4.1 divided by the number of hidden slabs (Ward System Group Inc., 1993). For GRNN and PNN, the networks require that the number of pattern units (the first hidden layer) must be at least equivalent to the number of training patterns, while the number of summation units (the second hidden layer) must equal to 2 for GRNN[24] and equal to the number of classes for PNN[25].

Table 4.5 shows the classification results of the neural networks on the bank lending decision models for the aggregate lending. The classification results show the neural networks, PNN and GRNN, exhibit a superior ability to learn and memorize the patterns corresponding to the borrower default risk (see Model 1 and 2). The overall percentage

[23] Since there is no science to determine the number of hidden neurons, getting the right number of hidden neurons is a matter or trial and error. However, the defaults are usually reliable (Ward Systems Group Inc., 1993).
[24] Theoretically, there are two different types of summation units. They are simple arithmetic summations and weighted summations.
[25] Since there are only two types of borrowers, good and bad borrower, the number of summation unit is two for PNN.

correct of PNN and GRNN, for both Model 1 and 2, are about 96 and 92 percent, respectively. Both networks offer a better classification results than the logistic regression. Similarly the WN and MLFN yield almost the same level of correctness as the logistic regression.

Between the PNN and the GRNN classification, the PNN has a lower Type I error than GRNN on both Model 1 and 2, but the GRNN Type II error on both models is slightly lower than the PNN (see Table 4.5). The results indicate that the PNN can predict the bad loan group better than the GRNN, but the GRNN is more accurate in detecting the good loan than the PNN, since it has 99 percent predictive power. However, the PNN is considered to be the best classification technique since it has the highest overall percentage correct in both Model 1 and 2.

When *Duration* is introduced into the models, the results in Table 4.5 also show that both the PNN and the GRNN Type I error reduced from 11.65 to 10.20 percent and 54.18 to 46.26 percent, respectively,. Therefore, the bank-borrower relationship can be considered as an important factor that would help to improve the classification capability of the bank lending decision models.

The classification results of the neural networks on the bank lending decision models for agricultural lending and non-agricultural lending are shown in Table 4.6 and Table 4.7, respectively. The classification results provide similar findings and conclusions as the models for aggregate lending (see Table 4.5). Table 4.6 and Table 4.7 suggest that the PNN and the GRNN are superior classification techniques, since both networks have a higher overall percentage correct prediction than the logistic regression models. In contrast, the WN and the MFLN produce almost identical results found in the logistic regression models. Therefore, the WN and the MFLN do not outperform the logistic regression.

Table 4.5: Classification tables of the neural networks on the bank lending decision models

	Model 1[1/]			Model 2[2/]		
	BL	GL	Overall	BL	GL	Overall
Logistic Regression (Logit)						
Correct	162	15,678	15,840	77	3,669	3,746
	6.39%	99.39%	86.51%	10.48%	98.92%	84.29%
Incorrect	2,374	96	2,470	658	40	698
	93.61%	0.61%	13.49%	89.52%	1.08%	15.71%
Probabilistic Neural Network (PNN)						
Correct	2,238	15,369	17,607	660	3,632	4,292
	88.35%	97.48%	96.16%	89.80%	97.92%	96.58%
Incorrect	295	397	692	75	77	152
	11.65%	2.52%	3.84%	10.20%	2.08%	3.42%
General Regression Neural Network (GRNN)						
Correct	1,162	15,723	16,885	395	3,698	4,093
	45.82%	99.68%	92.22%	53.74%	99.70%	92.10%
Incorrect	1,374	51	1,425	340	11	351
	54.18%	0.32%	7.78%	46.26%	0.30%	7.90%
Ward Network (WN)						
Correct	413	15,524	15,937	88	3,662	3,750
	16.29%	98.42%	87.04%	11.97%	98.73%	84.38%
Incorrect	2,123	250	2,373	647	47	694
	83.71%	1.58%	12.96%	88.03%	1.27%	15.62%
Multiple-layer Feed-forward Neural Network (MLFN)						
Correct	289	15,546	15,835	83	3,670	3,753
	11.40%	98.55%	86.48%	11.29%	98.95%	84.45%
Incorrect	2,247	228	2,475	652	39	691
	88.60%	1.45%	13.52%	88.71%	1.05%	15.55%
No. of Inputs			35			36
No. of Observations			18,310			4,444

Note: 1/ without Duration.
2/ with Duration.
BL and GL stand for bad loan and good loan, respectively.

Among the four different networks, the PNN is the best classification technique in terms of overall accuracy. Although the GRNN can be regarded as a superior technique when compared to the logistic regression, it does not outperform the PNN. Furthermore, not all the artificial neural networks are able to deal with the classification problem better than the logistic regression.

Table 4.6: Classification tables of the neural networks on the bank lending decision models for agricultural lending

	Model 1[1/]			Model 2[2/]		
	BL	GL	Overall	BL	GL	Overall
Logistic Regression (Logit)						
Correct	131	14,308	14,439	56	3,326	3,382
	6.02%	99.48%	87.19%	9.30%	98.90%	85.30%
Incorrect	2,046	75	2,121	546	37	583
	93.98%	0.52%	12.81%	90.70%	1.10%	14.70%
Probabilistic Neural Network (PNN)						
Correct	1,905	14,227	16,132	532	3,295	3,827
	87.51%	98.92%	97.42%	88.37%	97.98%	96.52%
Incorrect	272	156	428	70	68	138
	12.49%	1.08%	2.58%	11.63%	2.02%	3.48%
General Regression Neural Network (GRNN)						
Correct	807	14,357	15,164	297	3,353	3,650
	37.07%	99.82%	91.57%	49.34%	99.70%	92.06%
Incorrect	1,370	26	1,396	305	10	315
	62.93%	0.18%	8.43%	50.66%	0.30%	7.94%
Ward Network (WN)						
Correct	322	14,229	14,551	15	3,355	3,370
	14.79%	98.93%	87.87%	2.49%	99.76%	84.99%
Incorrect	1,855	154	2,009	587	8	595
	85.21%	1.07%	12.13%	97.51%	0.24%	15.01%
Multiple-layer Feed-forward Neural Network (MLFN)						
Correct	315	14,224	14,539	51	3,339	3,390
	14.47%	98.89%	87.80%	8.47%	99.29%	85.50%
Incorrect	1,862	159	2,021	551	24	575
	85.53%	1.11%	12.20%	91.53%	0.71%	14.50%
No. of Inputs			34			35
No. of Observations			16,560			3,965

Note: 1/ without Duration.
2/ with Duration.
BL and GL stand for bad loan and good loan, respectively.

Table 4.8 shows the relative contribution factors of the PNN models. The mean value of the contribution factor is presented at the end of Table 4.8. It shows on the average how much an input can contribute to the model. For example, if the model consists of 35 inputs, an input can contribute about 0.0286 (1/35 = 0.0286) or 2.86 percent to the model. Thus, a higher (or a lower) input contribution value which is above (or below) the mean value of 0.0286 is considered as an (or less) important input.

Table 4.7: Classification tables of the neural networks on the bank lending decision models for non-agricultural lending

	Model 1[1/]			Model 2[2/]		
	BL	GL	Overall	BL	GL	Overall
Logistic Regression (Logit)						
Correct	66	1,354	1,420	31	336	367
	18.38%	97.34%	81.14%	23.31%	97.11%	76.62%
Incorrect	293	37	330	102	10	112
	81.62%	2.66%	18.86%	76.69%	2.89%	23.38%
Probabilistic Neural Network (PNN)						
Correct	325	1,317	1,642	123	339	462
	90.53%	94.68%	93.83%	92.48%	97.98%	96.45%
Incorrect	34	74	108	10	7	17
	9.47%	5.32%	6.17%	7.52%	2.02%	3.55%
General Regression Neural Network (GRNN)						
Correct	103	1,383	1,486	105	340	445
	28.69%	99.42%	84.91%	78.95%	98.27%	92.90%
Incorrect	256	8	264	28	6	34
	71.31%	0.58%	15.09%	21.05%	1.73%	7.10%
Ward Network (WN)						
Correct	113	1,323	1,436	22	341	363
	31.48%	95.11%	82.06%	16.54%	98.55%	75.78%
Incorrect	246	68	314	111	5	116
	68.52%	4.89%	17.94%	83.46%	1.45%	24.22%
Multiple-layer Feed-forward Neural Network (MLFN)						
Correct	65	1,357	1,422	23	340	363
	18.11%	97.56%	81.26%	17.29%	98.27%	75.78%
Incorrect	294	34	328	110	6	116
	81.89%	2.44%	18.74%	82.71%	1.73%	24.22%
No. of Inputs			29			30
No. of Observations			1,750			479

Note: 1/ without Duration.
2/ with Duration.
BL and GL stand for bad loan and good loan, respectively.

The relative contribution factors in Table 4.8 show that *Asset*, *Age*, *Collateral* and *Capital turnover ratio* are important variables in explaining the aggregate lending decision models. *Education* and *Return on asset* are important only in Model 1 and the relative contribution of *Duration* in Model 2 verifies the essential of the bank-borrower relationship in the bank lending decision model. In contrast, *Sector* contributes to the models less than 1 percent and much lower than the mean values of the contribution factor, which implies that it is not an important factor in the PNN models.

Table 4.8: Relative contribution factors of PNN on bank lending decision models

Independent variables[1/]	Relative Contribution					
	Aggregate		Agriculture		Non-agriculture	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Borrower Characteristics						
Log(asset)	***0.0639***	***0.0567***	***0.0567***	***0.0571***	***0.0703***	0.0293
Age	***0.0526***	***0.0377***	***0.0364***	***0.0544***	0.0126	0.0053
Education	***0.0360***	0.0110	0.0034	***0.0418***	0.0249	0.0139
Credit risk proxies						
Log(collateral)	***0.0506***	***0.0570***	***0.0476***	***0.0376***	***0.0880***	***0.0549***
Return on asset	***0.0548***	0.0108	0.0115	0.0036	0.0058	0.0253
Leverage ratio	0.0158	0.0087	***0.0567***	***0.0409***	0.0072	***0.0562***
Capital turnover ratio	***0.0465***	***0.0583***	0.0115	0.0155	0.0188	0.0301
Relationship indicators						
Borrowing from others	0.0174	0.0268	0.0009	***0.0342***	***0.0725***	0.0029
Duration		***0.0325***		***0.0355***		0.0091
Dummy variables[2/]						
Sector	0.0013	0.0018				
(Province)	Yes	Yes	Yes	Yes	Yes	Yes
(Major production)	Yes	Yes	Yes	Yes	Yes	Yes
(Loan type)	Yes	Yes	Yes	Yes	Yes	Yes
(Loan size)	Yes	Yes	Yes	Yes	Yes	Yes
(Lending year)	Yes	Yes	Yes	Yes	Yes	Yes
No. of inputs	35	36	34	35	29	30
No. of observations	18,310	4,444	16,560	3,965	1,750	479
Mean value of the contribution factor[3/]	0.0286	0.0278	0.0294	0.0286	0.0345	0.0333

Note: 1/ Dependent variable is bank lending decision (good/bad loan).
2/ See Appendix 2 for the relative contribution on province, major production, loan type, loan size and lending year dummy variables.
3/ Equal to 1 over the number of inputs used in the model.
Bold and italic indicate that the relative contribution is above the mean value of contribution factors.

For the agricultural lending decision models, the relative contribution factors of Model 1 and 2 show *Asset*, *Age*, *Collateral* and *Leverage ratio* as important factors in determining the bank lending decision (see Table 4.8). In Model 2, *Education* and *Borrowing from others* also have a relatively high contribution to the model. The relative contribution of *Duration* illustrates a significant role of the bank-borrower relationship on the bank lending decision.

In case of non-agricultural lending decision models, only *Collateral* has a high contribution above the relative contribution mean value and can be regarded as a crucial factor in explaining the bank lending decision in both Model 1 and 2. *Asset* and *Borrowing from others* have a high contribution to the bank lending decision only in Model 1, while *Leverage ratio* has a high contribution to the bank lending decision only in Model 2. The relative contribution of *Duration* in Model 2 is 0.0091 which is much lower than the relative contribution mean value of 0.0345. Therefore, it can be concluded that the *Duration* does not have a significant influence on the bank lending decision for non-agricultural lending.

Some of the results from the relative contribution factors of the PNN models are similar to those in the logistic regressions. These include: 1) *Asset* as a key factor for the bank lending decision models in both the PNN and the logistic regressions, and 2) *Duration* has a significant role only in the models for aggregate lending and agricultural lending, but not for non-agricultural lending.

In contrast, some of the findings are inconsistence across the estimation techniques. For example, 1) *Sector* is not an important factor in the PNN model but is highly significant in the logistic regressions; 2) The relative contribution factors of the PNN models show that *Collateral* is an imperative input for all the models, particularly for the non-agricultural lending decision models, but it is not significant in the logistic regressions; 3) *Age* is insignificant in all the logistic models, however, it is an important factor on the PNN aggregate and agricultural lending decision models; 4) the PNN emphasized *Leverage ratio* rather than *Capital turnover ratio* in the agricultural lending decision models, but *Leverage ratio* is significant only in the Model 1 where *Capital turnover ratio* is significant in both Model 1 and 2 of the logistic regression models; and 5) The logistic regression results show that *Return on asset* has a significant impact on non-agricultural lending decision models. Conversely, its contributions on the PNN models are lower than the mean values on both Model 1 and 2.

The comparative results between the logistic regression and the PNN suggest that both estimation techniques pay attention to the different sets of independent variables. As a result, they yield different in-sample classification results. The PNN does not only pay attention to the value of *Asset*, but also to the value of *Collateral*, compared to the logistic regression. Furthermore, *Age* is also taken into account in the PNN for the aggregate lending and the agricultural lending decision models. Unlike the logistic regression, the PNN puts more weight on the *Leverage ratio* than *Capital turnover ratio* or *Return on asset* on the agricultural lending decision models.

There are no assumptions about the functional form and the distributions of the variables on the PNN model. Thus the PNN allows for nonlinear relationship and complex classification equations. In addition, the model is estimated via the genetic adaptive learning process, not the one-shot estimation process. Therefore, it has a higher ability to learn and memorize a complex non-linear relationship than the logistic regression[26]. As a result, the PNN generates an excellent in-sample classification results and becomes the superior classification model for the bank lending decision model.

4.1.3 Out-of-sample forecast

It is generally accepted that the within-sample forecast always yields an upward bias, since the model predicts well only in the in-sample forecast but its performance may be relatively poor on the out-of-sample forecast. Hence, to examine the future classification power of the model, the out-of-sample forecasting technique is more appropriate.

To perform the out-of-sample forecast, the data set is randomly divided into 2 sub-samples: a training (or estimation) sample and a forecast sample. The training sample and the forecast sample contain 80 and 20 percent of the total observations, respectively. To evaluate the forecast accuracy of the model, the classification rates (percentage correct and percentage incorrect classifications) are computed and compared.

The classification results of out-of-sample prediction for the logistic regression and the artificial neural networks on the aggregate lending decision models are presented in Table 4.9. The results show that the PNN and the GRNN models are overfitting to the training sample. The overall percentage correct on out-of-sample forecasting is relatively lower than in-sample forecasting in most cases. In contrast, the logistic regression, the WN, and

[26] The genetic adaptive learning process usually provides the global optimum solution in many types of problems. However, the estimation process is very time consuming.

the MLFN are not overfitting to the training sample, as the percentage correct of the models on both in-sample and out-of-sample forecasting are more likely the same[27]. However, the overfitting model can still be regarded as superior for future prediction, if it can perform better than the other models on out-of-sample forecasting. Further investigations need to be carried out to verify the best prediction model for the bank lending decision.

The results in Table 4.9 show that the GRNN has the highest overall percentage correct on both Model 1 and 2, while the overall percentage correct of the PNN model is better than the logistic regression only in Model 2. Furthermore, both the WN and the MLFN models yield almost the same overall percentage correct as the logistic regression models. Therefore, the accuracy of the 3 models is similar to each other. In term of the correctness, only the GRNN can predict the bank lending decision more accurately than the logistic regression (see Table 4.9). The major finding on out-of-sample forecasting is quite different from the in-sample forecasting, where the PNN is considered the best classification model.

A closer examination of the logistic regression performance, however, indicates that the logistic regression can predict well only on the good loan group. An examination of the Type I error rate shows that the logistic regression is unable to predict the bad loan group, as it has more than 90 percent of Type I error. The Type I error of the PNN is smaller than both the GRNN and the logistic regression, especially when *Duration* is introduced into the lending decision model (Model 2). The results show the PNN has the highest percentage correct when predicting the bad loan group and its performance is fairly well on predicting the good loan group. This suggests that the PNN can be used to predict both good and bad loan groups, because it strikes a balance between Type I and Type II error.

[27] Model "overfitting" can be quantified by contrasting the within-sample accuracy (where overfitting to the training sample may be present) with the out-of-sample accuracy (which is an unbiased assessment of the model's performance). In the absence of overfitting, the within-sample and out-of-sample accuracy would coincide (West and et al., 1997: p387).

Table 4.9: Classification results of out-of-sample forecasting on the bank lending decision models

	Model 1 1/			Model 2 2/		
	BL	GL	Overall	BL	GL	Overall
Logistic Regression (Logit)						
Correct	33	3,117	3,150	14	731	745
	6.30%	99.33%	86.02%	9.46%	98.78%	83.90%
Incorrect	491	21	512	134	9	143
	93.70%	0.67%	13.98%	90.54%	1.22%	16.10%
Probabilistic Neural Network (PNN)						
Correct	226	2,817	3,043	73	675	748
	43.38%	90.00%	83.10%	49.32%	91.22%	84.23%
Incorrect	295	313	608	75	65	140
	56.62%	10.00%	16.90%	50.68%	8.78%	15.77%
General Regression Neural Network (GRNN)						
Correct	105	3,097	3,202	34	731	765
	20.04%	98.69%	87.44%	22.97%	98.78%	86.15%
Incorrect	419	41	460	114	9	123
	79.96%	1.31%	12.56%	77.03%	1.22%	13.85%
Ward Network (WN)						
Correct	75	3,085	3,160	13	728	741
	14.31%	98.31%	86.29%	8.78%	98.38%	83.45%
Incorrect	449	53	502	135	12	147
	85.69%	1.69%	13.71%	91.22%	1.62%	16.55%
Multiple-layer Feed-forward Neural Network (MLFN)						
Correct	55	3,090	3,145	12	730	742
	10.50%	98.47%	85.88%	8.11%	98.65%	83.56%
Incorrect	469	48	517	136	10	146
	89.50%	1.53%	14.12%	91.89%	1.35%	16.44%
No. of Observations			3,662			888

Note: 1/ without Duration.
2/ with Duration.
BL and GL stand for bad loan and good loan, respectively.

Although the GRNN has the highest overall percentage correct and a very low Type II error on both Model 1 and 2, it is not reasonable to conclude that the GRNN is a superior model for the bank lending decision model. This is because the GRNN Type I error is higher than the PNN, and the overall percentage correct is calculated under the assumption that the misclassification costs of both Type I and Type II errors are identical. Since it is generally accepted that Type I error (accepted a bad loan as a good loan) has a higher misclassification cost, the overall percentage correct may be misleading in this case, as it ignores the relative cost difference between Type I and Type II error[28].

Table 4.10 presents the out-of-sample forecast results on the bank lending decision models for agricultural lending. The results show that the logistic regression, the WN, and the MLFN have the same level of prediction accuracy. The GRNN has the highest overall prediction accuracy with a relatively small Type II error on both Model 1 and 2. The PNN is more accurate than the logistic regression only in Model 1. However, the PNN Type I error is lower than the logistic regression on both models, particularly Model 2.

The results in Table 4.10 also show that the logistic regression can predict the good loan group slightly better than the GRNN and the PNN, but it can correctly predict the bad loan group only about 4 and 5 percent on Model 1 and 2, respectively. Although the PNN is not the best model in term of overall prediction accuracy, it has the lowest Type I error (59.83 percent) in Model 2 and its Type I error is marginally higher than the GRNN in Model 1. The results in Table 10 are similar to the findings and conclusions in the aggregate lending decision models.

The out-of-sample forecast classification results of the non-agricultural lending decision models are shown in Table 4.11. The results indicate that both the PNN and the GRNN perform better than the logistic regression on out-of-sample forecasting. Furthermore, the PNN has the highest overall prediction accuracy and the lowest Type I error on both Model 1 and 2. Even though the PNN has a relatively high Type II error when compare to the other models, the percentage correct prediction on the good loan group is higher than 90 percent. Therefore, the PNN can be considered as the superior model in predicting the bank lending decision for non-agricultural lending.

[28] If the misclassification cost of Type I error is significantly greater than Type II error, then a large overall accuracy rate with a large Type I error rate would be costlier than a small overall accuracy rate with a small Type I error rate (Etheridge and Sriram, 1996).

Table 4.10: Classification results of out-of-sample forecasting on the bank lending decision models for agricultural lending

	Model 1 1/			Model 2 2/		
	BL	GL	Overall	BL	GL	Overall
Logistic Regression (Logit)						
Correct	18	2,851	2,869	6	670	676
	4.05%	99.41%	86.62%	5.13%	99.11%	85.25%
Incorrect	426	17	443	111	6	117
	95.95%	0.59%	13.38%	94.87%	0.89%	14.75%
Probabilistic Neural Network (PNN)						
Correct	49	2,846	2,895	47	619	666
	11.04%	99.23%	87.41%	40.17%	91.57%	83.98%
Incorrect	395	22	417	70	57	127
	88.96%	0.77%	12.59%	59.83%	8.43%	16.02%
General Regression Neural Network (GRNN)						
Correct	69	2,849	2,918	18	667	685
	15.54%	99.34%	88.10%	15.38%	98.67%	86.38%
Incorrect	375	19	394	99	9	108
	84.46%	0.66%	11.90%	84.62%	1.33%	13.62%
Ward Network (WN)						
Correct	47	2,838	2,885	1	674	675
	10.59%	98.95%	87.11%	0.85%	99.70%	85.12%
Incorrect	397	30	427	116	2	118
	89.41%	1.05%	12.89%	99.15%	0.30%	14.88%
Multiple-layer Feed-forward Neural Network (MLFN)						
Correct	47	2,843	2,890	5	671	676
	10.59%	99.13%	87.26%	4.27%	99.26%	85.25%
Incorrect	397	25	422	112	5	117
	89.41%	0.87%	12.74%	95.73%	0.74%	14.75%
No. of Observations			3,312			793

Note: 1/ without Duration.
2/ with Duration.
BL and GL stand for bad loan and good loan, respectively.

Table 4.11: Classification results of out-of-sample forecasting on the bank lending decision models for non-agricultural lending

	Model 1[1/]			Model 2[2/]		
	BL	GL	Overall	BL	GL	Overall
Logistic Regression (Logit)						
Correct	6	271	277	5	57	62
	8.22%	97.83%	79.14%	17.24%	86.36%	65.26%
Incorrect	67	6	73	24	9	33
	91.78%	2.17%	20.86%	82.76%	13.64%	34.74%
Probabilistic Neural Network (PNN)						
Correct	39	253	292	19	62	81
	53.42%	91.34%	83.43%	65.52%	93.94%	85.26%
Incorrect	34	24	58	10	4	14
	46.58%	8.66%	16.57%	34.48%	6.06%	14.74%
General Regression Neural Network (GRNN)						
Correct	10	270	280	13	64	77
	13.70%	97.47%	80.00%	44.83%	96.97%	81.05%
Incorrect	63	7	70	16	2	18
	86.30%	2.53%	20.00%	55.17%	3.03%	18.95%
Ward Network (WN)						
Correct	18	261	279	3	66	69
	24.66%	94.22%	79.71%	10.34%	100.00%	72.63%
Incorrect	55	16	71	26	-	26
	75.34%	5.78%	20.29%	89.66%	0.00%	27.37%
Multiple-layer Feed-forward Neural Network (MLFN)						
Correct	8	269	277	3	64	67
	10.96%	97.11%	79.14%	10.34%	96.97%	70.53%
Incorrect	65	8	73	26	2	28
	89.04%	2.89%	20.86%	89.66%	3.03%	29.47%
No. of Observations			350			95

Note: 1/ without Duration.
2/ with Duration.
BL and GL stand for bad loan and good loan, respectively.

The results in Model 1 (without *Duration*) in Table 4.11 show that the overall percentage correct of the GRNN, the WN, and the MLFN are very close to the overall percentage correct of the logistic regression, but the GRNN, the WN, and the MLFN Type I errors are lower than the logistic regression. In contrast, all 3 models perform much better than the logistic regression when the *Duration* is included (see Model 2). The 3 models have higher overall prediction accuracy and lower Type II error than the logistic regression. However, the WN and the MLFN Type I errors are larger than the logistic regression. Therefore, the

WN and the MLFN might not outperform the logistic regression on the out-of-sample forecasting.

4.1.4 Expected misclassification loss of the models

The misclassification costs of Type I and Type II errors must be differentiated and accounted for when interpreting the results (Anandarajan et al., 2001, and Etheridge and Sriram, 1997). The expected loss of misclassification on out-of-sample forecasting must be estimated. The bank lending decision model that offers the smallest expected loss is considered the most preferable model.

According to Koh (1992), the expected misclassification loss (EL) of the model can be calculated by using the following equation:

$$EL = (PB)\,(PI)\,(CI) + (PG)\,(PII)\,(CII) \tag{4.2}$$

where PB = prior probability of being bad loan,

PG = prior probability of being good loan,

PI = conditional probability of Type I error (= No. of Type I errors given by the model / No. of bad loans),

PII = conditional probability of Type II error (= No. of Type II errors given by the model / No. of good loans),

CI = misclassification costs of Type I error,

CII = misclassification costs of Type II error.

It is important to differentiate between Type I and Type II error costs, because the misclassification loss to the bank from Type I error is generally greater than Type II error. The consequences of incorrect classification are intangible and immeasurable, such as loss of existing and potential clients, loss of depositor's trustworthy, etc. Thus, CI and CII are not quantified in this research. However, to overcome the misclassification costs dilemma, the relative misclassification costs of Type I and Type II errors are used. The relative cost ratios are assumed to vary accordingly from 1:1, 2:1, 3:1, 4:1 and 5:1, with the relatively higher misclassification cost on Type I error where a bad loan is classified as a good loan[29] (Koh, 1992; Novak and LaDue, 1999).

Table 4.12 summarizes the model expected misclassification loss on out-of-sample forecasting at different relative cost ratios. For aggregate lending, the GRNN with *Duration* (Model 2) has the lowest expected loss of misclassification when the relative cost ratio of Type I and Type II errors is 1:1. Although the PNN with *Duration* (Model 2) has lower overall percentage correct than the GRNN (Model 2) on out-of-sample forecasting, when the cost ratio is 2:1 or higher, the PNN becomes the top performer since it has the lowest expected loss.

The PNN and the GRNN give lower expected cost than the logistic regression at all levels of relative cost ratio, except the PNN at 1:1 cost ratio. The logistic regression with *Duration* (Model 2) has lower expected cost of misclassification than the Model 2 of WN and MLFN at all relative cost ratios (see Table 4.12). However, when the cost ratio is 3:1 or higher, the logistic regression without *Duration* should not be used to predict the bank lending decision for the aggregate lending because it has the highest expected cost.

The expected loss results on out-of-sample prediction for agricultural lending in Table 4.12 show that the GRNN (Model 1 without *Duration*) offers the lowest misclassification cost at 1:1 relative cost ratio. When the relative cost ratio increases, the PNN (Model 2 with *Duration*) generates the minimum relative error costs. Thus, the PNN and the GRNN can perform better than the logistic regression, excluding the Model 2 PNN at 1:1 relative cost ratio. However, the Model 2 logistic regression has lower expected loss than the WN and the MLFN in most cases, but not Model 1 where the logistic regression is relatively more costly than the WN and the MLFN.

[29] According to BAAC annual report (2003), doubtful accounts are estimated at 40, 70 and 100 percent of the loan principal overdue for less than 1 year, 1 year or more but not exceeding 2 years, and more than 2 years, respectively. Furthermore, the uncollected due interest payments are recorded in full amount as doubtful accounts. Thus, the relative cost ratio between Type I and Type II errors should be much greater than 1:1.

Table 4.12: Expected loss of misclassification on out-of-sample forecasting

Model	Cost Ratio (CI : CII)				
	1:1	2:1	3:1	4:1	5:1
Aggregate lending					
Logistic Regression (Logit)					
Model 1	0.1355	0.2653	0.3951	0.5249	0.6547
Model 2	0.1359	0.2613	0.3867	0.5121	0.6375
Probabilistic Neural Network (PNN)					
Model 1	0.1646	0.2430	0.3214	0.3998	0.4783
Model 2	0.1459	***0.2160***	***0.2862***	***0.3564***	***0.4266***
General Regression Neural Network (GRNN)					
Model 1	0.1220	0.2328	0.3435	0.4543	0.5650
Model 2	***0.1172***	0.2238	0.3305	0.4372	0.5439
Ward Network (WN)					
Model 1	0.1332	0.2519	0.3706	0.4893	0.6079
Model 2	0.1403	0.2666	0.3930	0.5193	0.6457
Multiple-layer Feed-forward Neural Network (MLFN)					
Model 1	0.1371	0.2611	0.3851	0.5090	0.6330
Model 2	0.1389	0.2662	0.3935	0.5207	0.6480
Agricultural lending					
Logistic Regression (Logit)					
Model 1	0.1313	0.2574	0.3835	0.5097	0.6358
Model 2	0.1324	0.2571	0.3819	0.5066	0.6313
Probabilistic Neural Network (PNN)					
Model 1	0.1236	0.2406	0.3575	0.4745	0.5914
Model 2	0.1519	***0.2305***	***0.3092***	***0.3878***	***0.4665***
General Regression Neural Network (GRNN)					
Model 1	***0.1168***	0.2278	0.3388	0.4499	0.5609
Model 2	0.1228	0.2340	0.3453	0.4565	0.5677
Ward Network (WN)					
Model 1	0.1266	0.2442	0.3617	0.4793	0.5968
Model 2	0.1329	0.2632	0.3936	0.5239	0.6543
Multiple-layer Feed-forward Neural Network (MLFN)					
Model 1	0.1251	0.2427	0.3602	0.4778	0.5953
Model 2	0.1323	0.2581	0.3840	0.5098	0.6356
Non-agricultural lending					
Logistic Regression (Logit)					
Model 1	0.2055	0.3938	0.5821	0.7703	0.9586
Model 2	0.2782	0.4479	0.6177	0.7875	0.9573
Probabilistic Neural Network (PNN)					
Model 1	0.1644	0.2600	0.3555	0.4511	0.5466
Model 2	***0.1189***	***0.1897***	***0.2604***	***0.3311***	***0.4019***
General Regression Neural Network (GRNN)					
Model 1	0.1971	0.3742	0.5512	0.7283	0.9053
Model 2	0.1373	0.2505	0.3636	0.4768	0.5900
Ward Network (WN)					
Model 1	0.2005	0.3550	0.5096	0.6642	0.8187
Model 2	0.1839	0.3678	0.5518	0.7357	0.9196
Multiple-layer Feed-forward Neural Network (MLFN)					
Model 1	0.2056	0.3883	0.5709	0.7536	0.9363
Model 2	0.2080	0.3919	0.5759	0.7598	0.9437

Note: Model 1 and 2 are without and with Duration, respectively.
Bold and italic indicate the minimum expected loss.

In case of non-agricultural lending, the PNN with *Duration* (Model 2) yields the lowest expected misclassification loss consistently for all the levels of relative cost ratios (see Table 4.12). In addition, all the artificial neural networks models (PNN, GRNN, WN and MLFN) perform better than the logistic regression models when compared in terms of relative error costs, except for the MLFN (Model 1) when the relative cost ratio is 1:1. The Model 2 logistic regression grants the highest misclassification cost for all relative cost ratios, except at a 5:1 cost ratio, where the Model 1 logistic regression has the highest expected cost of incorrect classification.

In summary, in terms of prediction accuracy, the PNN can be considered as the best prediction model for in-sample forecasting (see Table 4.5, 4.6, and 4.7). However, for out-of-sample forecasting, the overall percentage correct of the GRNN is the highest on aggregate and agricultural lending, while the PNN is still the best prediction model for non-agricultural lending (see Table 4.9, 4.10 and 4.11). The GRNN and the PNN perform better than the logistic regression on out-of-sample forecasting. In most cases, the logistic regression performances are quite similar to the WN and the MLFN on both in-sample and out-of-sample forecast. Therefore, the results suggest that in term of precision, only some types of neural networks can predict the bank lending decision better than the logistic regression.

To account for misclassification costs, the expected loss is calculated at various cost ratios. The results in Table 4.12 show that when the relative cost between Type I and Type II errors is 2:1 or higher, the PNN with *Duration* (Model 2) proposes the lowest misclassification costs. Therefore, it can be concluded that the PNN (Model 2) is the superior model in predicting the bank lending decision, since the Type I error (classifying a bad loan as a good loan) is more costly than the Type II error (classifying a good loan as a bad loan).

4.2 Credit availability model

The estimated results of the ordinary least squares (OLS) regressions for the credit availability models are shown in Table 4.13. The estimated results in Model 1 and 2 (without and with *Duration*) show that all explanatory variables can explain for 74.10 and 75.76 percent of the total variation in credit availability, respectively. Furthermore, the null hypothesis that all explanatory variable coefficients are jointly zero can be rejected, since the F-statistics on both models are highly significant. Therefore, the results indicate that both models fit the data quite well. However, the coefficient of determination (R^2) and Root Mean Squares Error (RMSE) of Model 2 are slightly higher and lower than Model 1, respectively. Therefore, the result implies that the model that includes *Duration* might explain the variation of credit availability slightly better than the model without *Duration*.

The dummy variables for Sector, Provinces, Major productions, Loan types, Loan sizes and Lending years in both models are highly significant in explaining credit availability. The estimated results are consistent with the financial theory. For example, the positive coefficient of *Long-term loan*, compared to the other dummy variables in the same group, suggests that long-term loan is normally granted a larger amount of credit than short-term loan and medium-term loan. Furthermore, the borrowers who have livestock or aquaculture as the major production received a bigger amount of credit from the bank when compared to the other types of production.

The negative sign on *Sector* implies that agricultural loans generally obtained a lower amount of credit than non-agricultural loans. This is because an agricultural loan is more risky in nature and has a higher operating cost than non-agricultural loan. The result is consistent with the standard practice of the commercial bank.

The positive and negative coefficients of the province dummy variables are indicative of an increase or decrease in the amount of credit granted on the loan contract in a specific province relative to the other provinces. For instance, the volume of credit for the loan agreement in *Province 15*, which is in the irrigation area and a major tourist province, is higher than *Province 9*, which is in the rainy area and has paddy rice as a major crop, by 24.41 percent on the average (the coefficient of *Province 9* is not significantly different from zero) (see Table 4.13, Model 2). Since the credit risk and the credit worthiness of the borrower might differ from province to province, the availability of credit varies in accordance with the province risk.

Table 4.13: Credit availability models

Independent variables[1/]	Coefficients[2/]	
	Model 1	Model 2
Borrower Characteristics		
Log(asset)	0.0515**	0.0651**
Age	-0.0026**	-0.0014
Education	-0.0351**	-0.0633**
Credit risk proxies		
Log(collateral)	0.1587**	0.1711**
Return on asset	-0.0053	0.0010
Leverage ratio	0.0145	0.0747
Capital turnover ratio	0.0081**	0.0092**
Relationship indicators		
Borrowing from others	-0.0393**	-0.0593**
Duration		-0.0229**
Dummy variables[3/]		
Sector	-0.1211**	-0.0800**
(Province)		
Province 2	0.0580**	0.1412**
Province 3	0.0576**	0.1195**
Province 4	0.0616**	0.1143
Province 5	0.0116	0.1048*
Province 6	-0.0115	0.1224**
Province 7	0.1191**	0.1459**
Province 8	-0.0551**	0.0452
Province 9	-0.0166	0.0660
Province 10	0.2224**	0.1948**
Province 11	0.1612**	0.1796**
Province 12	0.0665**	-0.1128
Province 13	0.2513**	0.1878**
Province 14	0.2277**	0.1710*
Province 15	0.2797**	0.2441**
Province 16	0.1463**	0.1308**
Province 17	0.1714**	0.1248**

Table 4.13: Credit availability models (Cont)

Independent variables1/	Coefficients2/	
	Model 1	Model 2
(Major production)		
Horticulture	0.0180	-0.0455
Orchard/Vegetable	0.0529**	0.0224
Livestock/Aquaculture	0.1396**	0.0831*
(Loan type)		
Short-term loan	-0.1666**	-0.1877**
Medium-term loan	0.1725**	0.1132**
Long-term loan	0.3684**	0.2678**
(Loan size)		
Medium loan	1.2116**	1.1832**
Large loan	2.7777**	2.7331**
(Lending year)		
2001	0.0387*	0.0222
2002	-0.0087	0.0341**
Constant	7.8563**	7.6453**
R-squared	0.7410	0.7576
Adjusted R-squared	0.7405	0.7556
RMSE	0.5007	0.4774
F-statistic	1493.69**	382.58**
No. of Observations	18,310	4,444

Note: 1/ Dependent variable is log(volume of credit).
2/ White adjustment for estimation a heteroscedasticity consistent covariance matrix.
3/ To avoid singularity problem, a dummy variable is dropped from each group.
*, ** represent 10% and 5% significant level, respectively.

The results in Model 1 and 2 also demonstrate that *Asset*, *Collateral* and *Capital turnover ratio* have a positive impact on the availability of credit as hypothesised. From Model 2, the estimated coefficients indicate that a 1 percent increased in *Asset*, *Collateral* and *Capital turnover ratio* would increase the amount of credit granted by 0.0651, 0.1711 and 0.0092 percent, respectively (see Table 4.13). Thus, the results show that bank places a high emphasis on the value of collateral pledged when determining the volume of credit.

The results in Table 4.13 show that *Education*, *Borrowing from others*, and *Duration* have a negative impact on the credit availability. Only *Borrowing from others* is found as hypothesised and is consistent with Petersen and Rajan (1994) and Angelini et al. (1998) findings[30]. The estimated coefficient on *Borrowing from other* in Model 2 suggests that if the borrowers have an outstanding debt with other financial sources, the amount of credit on the loan agreement would be reduced by 5.93 percent on the average.

Although the estimated coefficient of *Duration* is negative and contradicts with the hypothesised sign[31], the effect of the bank-borrower relationship on credit availability could not be rejected, because the coefficient is significantly different from zero at the 5 percent level (see Table 4.13, Model 2). The result shows that an additional year of the bank-borrower relationship would decrease the amount of credit granted to the borrower by 2.29 percent. This implies that a longer relationship between the bank and the borrower enables the bank to be a more efficient lender. The bank would utilize the information and monitors the lending risk by the use of credit availability channel (controlling the loan amount).

Education and *Age* are expected to have a positive relationship with the amount of credit granted. However, the estimated results in Table 4.13 provide evidences of negative relationships on both variables. The negative sign on *Age* is consistent with Petersen and Rajan (1994), Angelini et al. (1998), and Harhoff and Korting (1998) findings. However, *Age* is eliminated when the *Duration* is introduced into the model. Therefore, the result confirms the influence of the bank-borrower relationship on the credit availability. The negative sign on *Education* implies that the borrower whose education level is higher than primary school is more likely to receive less credit from the bank. This implies that an educated borrower is a rational borrower and borrows only the required amount from the bank to minimize the interest expense.

In Table 4.13, Model 1, the effect of *Return on asset* and *Leverage ratio* on the amount of credit granted is found to be negative and positive, respectively. The results are not

---

[30] Petersen and Rajan (1994) and Angelini et al. (1998) concluded that those firms that maintain multiple-bank relationships are more credit constrained than single-bank firms.

[31] Petersen and Rajan (1994), Burger and Udell (1995) and Akhavein et al. (2004) found that the duration of the relationship has a positive relationship to the credit available to borrowers. Therefore, the result found on this research is not consistent with them.

consistent with the financial theory. The estimated sign on *Return on asset* is positive when *Duration* is introduced into the model (see Model 2) but *Leverage ratio* is positive and inconsistence with the hypothesised sign. However, the t-statistic shows that both *Return on asset* and *Leverage ratio* do not have a significant impact on the volume of credit granted.

The results in Table 4.13 show that *Asset*, *Education*, *Collateral*, *Capital turnover ratio* are the key factors affecting the amount of credit granted to the borrower in rural lending. *Age*, *Return on asset* and *Leverage ratio*, on the other hand, do not have a significant impact on the amount of loan granted. Furthermore, the significant of *Borrowing from others* and *Duration* confirms the important of relationship lending in the rural financial market. The negative relationship between *Duration* and the availability of credit, which contradict the hypothesis, signifies that the bank uses the information obtained from the relationship to monitor the lending risk through credit availability.

4.2.1 Credit availability models for agricultural and non-agricultural lending

The estimated coefficients of *Sector* are significant at the 5 percent level, implying that the amount of credit provided to the borrower would depend upon the *Sector* (see Table 4.13). Thus, the factors affecting the volume of credit may not have the same influence on the different lending sectors.

To estimate the credit availability models for agricultural lending and non-agricultural lending, *Sector* is excluded from the models because the model is segregated according to the lending sectors. For non-agricultural lending, *Province 11* and the year *2001* dummy variables are removed from the models, due to insufficient information. Furthermore, to avoid the singularity problem when the non-agricultural lending models are estimated, *Horticulture*, *Short-term loan*, and *Medium-term loan* dummy variables are also excluded from the models. This is because most of the non-agricultural loans are *Long-term loan* and there are only three different groups found on the borrower's major production.

The results in Table 4.14 (Model 1) confirm the important of *Asset*, *Age*, *Education*, *Collateral*, *Capital turnover ratio* and *Borrowing from others* as important factors in determining the credit availability for agricultural lending. The effects of *Age* and

Education to the volume of credit are negative and consistent with the results in the aggregate model discussed previously (see Table 4.13).

When *Duration* is included into the credit availability model, *Age* and *Capital turnover ratio* become insignificance (see Table 4.14, Model 2). Only *Asset*, *Education*, *Collateral*, *Borrowing from others* and *Duration* are significantly different from zero at the 5 percent level. The effects of *Asset* and *Collateral* are both uniformly positive, while the effects of *Education*, *Borrowing from other* and *Duration* are negative. Furthermore, the bank relies on the borrower characteristics (*Asset* and *Education*) rather than financial performance (*Return on asset*, *Leverage ratio*, and *Capital turnover ratio*) when determining the amount of credit for an agricultural loan.

The province, major production, loan type, loan size and lending year dummy coefficients on both models are significant at the 5 percent level. The estimated results show that the volume of credit granted on the agricultural loan is partially determined by province, loan type, loan size and lending year. However, the amount of credit funding might not necessarily be determined by farm type or the borrower major production. This is because the coefficients of major production dummy variables in Model 2 are not significantly different from zero (see Table 4.14). The borrower may borrow the money from the bank for investing in other agricultural production in addition to their major production.

Table 4.15 shows the estimated results of the credit availability models for non-agricultural lending. Overall, the results are quite different from the models for aggregate lending and agricultural lending. In Model 1, apart from the dummy variables which are used as control variables, only *Asset* and *Collateral* have significant influence on the availability of credit on non-agricultural loan contract. The variables such as *Age*, *Education*, *Capital turnover ratio* and *Borrowing from others* are significant in the credit availability models for aggregate lending and agricultural lending but insignificant in the non-agricultural lending models. Both *Return on asset* and *Leverage ratio* do not impact the volume of credit granted, and the results are similar to the aggregate lending and agricultural lending models.

Table 4.14: Credit availability models for agricultural lending

Independent variables[1/]	Coefficients[2/]	
	Model 1	Model 2
Borrower Characteristics		
Log(asset)	0.0533**	0.0651**
Age	-0.0027**	-0.0011
Education	-0.0401**	-0.0735**
Credit risk proxies		
Log(collateral)	0.1574**	0.1718**
Return on asset	-0.0064	0.0067
Leverage ratio	0.0154	0.0529
Capital turnover ratio	0.0082**	0.0068
Relationship indicators		
Borrowing from others	-0.0396**	-0.0541*
Duration		-0.0251**
Dummy variables[3/]		
(Province)		
Province 2	0.0622**	0.1372**
Province 3	0.0543**	0.1013*
Province 4	0.0679**	0.1924
Province 5	0.0131	0.1132*
Province 6	-0.0172	0.1276**
Province 7	0.0929**	0.0922**
Province 8	-0.0771**	0.0185
Province 9	-0.0216	0.0395
Province 10	0.2339**	0.2061**
Province 11	0.1663**	0.1842**
Province 12	0.0921**	-0.0099
Province 13	0.2588**	0.2036**
Province 14	0.2556**	0.1934*
Province 15	0.2873**	0.2501**
Province 16	0.1565**	0.1348**
Province 17	0.1810**	0.1325**

Table 4.14: Credit availability models for agricultural lending (Cont)

Independent variables[1/]	Coefficients[2/]	
	Model 1	Model 2
(Major production)		
Horticulture	0.0127	-0.0543
Orchard/Vegetable	0.0461*	0.0096
Livestock/Aquaculture	0.1408**	0.0795
(Loan type)		
Short-term loan	-0.1525**	-0.1859**
Medium-term loan	0.1806**	0.1221**
Long-term loan	0.3831**	0.2887**
(Loan size)		
Medium loan	1.2032**	1.1430**
Large loan	2.7930**	2.6529**
(Lending year)		
2001	0.0342*	0.0248
2002	-0.0094	0.0335*
Constant	7.7337**	7.5617**
R-squared	0.7384	0.7564
Adjusted R-squared	0.7379	0.7542
RMSE	0.4992	0.4645
F-statistic	1371.85**	348.54**
No. of Observations	16,560	3,965

Note: 1/ Dependent variable is log(volume of credit).
2/ White adjustment for estimation a heteroscedasticity consistent covariance matrix.
3/ To avoid singularity problem, a dummy variable is dropped from each group.
*, ** represent 10% and 5% significant level, respectively.

Table 4.15: Credit availability models for non-agricultural lending

Independent variables[1/]	Coefficients[2/]	
	Model 1	Model 2
Borrower Characteristics		
Log(asset)	0.0433**	0.0604*
Age	-0.0012	-0.0029
Education	0.0030	0.0067
Credit risk proxies		
Log(collateral)	0.1679**	0.1803**
Return on asset	0.0026	-0.0214
Leverage ratio	0.0304	0.0840
Capital turnover ratio	0.0063	0.0177**
Relationship indicators		
Borrowing from others	-0.0497	-0.0539
Duration		0.0009
Dummy variables[3/]		
(Province)		
Province 1	-0.0550	-0.1539
Province 2	-0.0782*	0.0124
Province 3	0.1839*	0.5132**
Province 4	-0.0993	-0.5705**
Province 5	-0.0738	-0.1415
Province 6	-0.0809	-0.1434
Province 7	0.1704**	0.2820**
Province 8	-0.0850*	0.0091
Province 9	-0.1398**	0.1295
Province 10	0.0014	0.1524
Province 12	-0.5053**	-0.7273**
Province 13	-0.0501	-0.3462
Province 14	0.0059	0.0085
Province 15	-0.0349	0.0940
Province 17	-0.0992	-0.1501

Table 4.15: Credit availability models for non-agricultural lending (Cont)

Independent variables[1/]	Coefficients[2/]	
	Model 1	Model 2
(Major production)		
Orchard/Vegetable	0.0104	0.0984
Livestock/Aquaculture	-0.0476	0.2277**
(Loan type)		
Long-term loan	0.0591	0.2180
(Loan size)		
Medium loan	1.2336**	1.4383**
Large loan	2.7624**	2.9618**
(Lending year)		
2002	-0.0144	-0.0201
Constant	8.1851**	7.4837**
R-squared	0.7416	0.7715
Adjusted R-squared	0.7372	0.7562
RMSE	0.5017	0.5379
F-statistic	170.22**	50.43**
No. of Observations	1,750	479

Note: 1/ Dependent variable is log(volume of credit).
2/ White adjustment for estimation a heteroscedasticity consistent covariance matrix.
3/ To avoid singularity problem, a dummy variable is dropped from each group.
*, ** represent 10% and 5% significant level, respectively.

The estimated results of Model 2 in Table 4.15 show that *Capital turnover ratio* is significant after *Duration* is included in the model. There are 3 major variables (*Asset*, *Collateral* and *Capital turnover ratio*) which can be used to resolve the volume of credit supply to the non-agricultural loan contract. The significant of *Capital turnover ratio* in Model 2 implies that the bank does not consider only the borrower characteristic (*Asset*), but also concern about the borrower's financial performance (efficiency) on the volume of credit granted to non-agricultural lending. The result is quite different from agricultural lending where only the borrower characteristics are considered when the volume of credit is determined.

The estimated signs of *Borrowing from others* and *Duration* are negative and positive, respectively, and consistent with the expected signs (see Table 4.15). However, they are not significantly different from zero. The results indicate that the bank-borrower relationship does not have an impact on the availability of credit for non-agricultural lending. As a result, it can be concluded that the relationship lending is not a crucial factor in determining the amount of credit granted for a non-agricultural loan.

In Model 1, there are only 5 out of the 15 provinces that are significant at the 10 percent level, whilst only 4 provinces are significant in Model 2 (see Table 4.15). The estimated results show that on average, some provinces could receive a larger volume of credit than the others. Therefore, the impact of province on the availability of credit still exists. There is no evidence that the amount of credit granted is influenced by the types of loans. This is because most of the non-agricultural loans are long term loans and the coefficients of *Long-term loan* are insignificant in both models. Alternatively, the significance of *Livestock/Aquaculture* dummy variable in Model 2 implies that the borrower who has an animal farm would be able to receive more financial support from the bank on their non-agricultural loan contract. The insignificance of many variables on both models could be partly caused by the small number of observations in non-agricultural lending.

In summary, Table 4.13, 4.14, and 4.15 suggest that the factors determining the credit availability are influenced by lending sectors. In Table 4.13, the negative signs on *Sector* demonstrate that on the average, borrowing for investing in agricultural productions would receive a smaller amount of credit from the bank than borrowing for non-agricultural productions. Furthermore, the results from Table 4.14 and 4.15 indicate that the bank pays more attention to the borrower characteristics (*Asset* and *Education*) when determining the amount of credit for an agricultural loan. For the non-agricultural loan, the bank does not consider only the borrower characteristic (Asset), but they are also concerned about the borrower's financial performance (*Capital turnover ratio*) when determining that amount of credit. *Collateral* may be considered a very important factor influencing the availability of credit for both agricultural and non-agricultural lending, because the bank could make a claim over it in case of default. There is also no evidence of relationship lending found on the credit availability models for non-agricultural lending. In addition, for agricultural lending, the negative sign shown on *Borrowing from others* and *Duration* suggest that the bank manages the lending risk by controlling the amount of credit granted.

4.2.2 Artificial neural networks and credit availability models

The artificial neural networks technique is employed to model the bank's decision on the quantity of loan to the borrower in the rural area. Three different forms of the artificial neural networks, including general regression neural network (GRNN), Ward network (WN) and multi-layer feed-forward neural network (MLFN), are conducted. Because the probabilistic neural network (PNN) is a specific network for a classification problem, it is inappropriate to apply the PNN on the prediction problem where the output is a continuous value. The data set, input, and output variables used for the neural networks models are identical to the data set and variables used for the multiple regression models in the previous subsection. The number of hidden neurons used on each network is set as the default value provided by Neuroshell2 (see Subsection 5.1.2 for details).

The results show that both the GRNN and the WN, can estimate the amount of credit granted more precisely than the regression model (see Table 4.16). However, it is inconclusive whether that the artificial neural networks technique outperforms the regression technique, since the results from the MLFN models are better than the regression models in some cases only. In addition, the results imply that the information gain from the bank-borrower relationship would help to improve the model performance. Furthermore, the results in Table 4.16 also show that the GRNN is the best model in assessing the volume of credit granted, since it yields the highest R^2 and the lowest RMSE on both aggregate and segregated models.

However, the results do not provide a strong and conclusive evidence of superiority in term of prediction capability as shown by the sample results. The out-of-sample forecast and forecast evaluation are conducted in the next section to explore the predictive power of the artificial neural networks models and the regression models.

The GRNN is considered as the best network for in-sample forecast and only the relative contribution factor from the GRNN on the credit availability models are presented in Table 4.17. The mean value of the contribution factor is used as an indicator to specify the relative important of the factors because the connection weights of the neural networks can not be tested directly.

Table 4.16: Neural Networks R-squared (R^2) and Root Mean Squared Error (RMSE) on credit availability models

Type of networks	Model 1[1/]		Model 2[2/]	
	R^2	RMSE	R^2	RMSE
Aggregate lending				
Multiple Linear Regression (MLR)	0.7410	0.5007	0.7576	0.4774
General Regression Neural Network (GRNN)	***0.7911***	***0.4497***	***0.8325***	***0.3967***
Ward Network (WN)	0.7430	0.4990	0.7897	0.4450
Multi-layer Feed-forward Network (MLFN)	0.7358	0.5060	0.7872	0.4472
No. of Observations		18,310		4,444
Agricultural lending				
Multiple Linear Regression (MLR)	0.7384	0.4992	0.7564	0.4645
General Regression Neural Network (GRNN)	***0.7924***	***0.4447***	***0.8386***	***0.3781***
Ward Network (WN)	0.7472	0.4909	0.7855	0.4359
Multi-layer Feed-forward Network (MLFN)	0.7318	0.5060	0.7608	0.4604
No. of Observations		16,560		3,965
Non-agricultural lending				
Multiple Linear Regression (MLR)	0.7416	0.5017	0.7715	0.5379
General Regression Neural Network (GRNN)	***0.8572***	***0.3728***	***0.8696***	***0.4062***
Ward Network (WN)	0.8004	0.4405	0.7897	0.5158
Multi-layer Feed-forward Network (MLFN)	0.7545	0.4889	0.7676	0.5422
No. of Observations		1,750		479

Note: 1/ without Duration.
2/ with Duration.
Bold and italic indicate (alternatively) the highest R^2 or the lowest RMSE.

The relative contribution factor of the GRNN on credit availability models in Table 4.17 shows that *Asset* and *Collateral* are important factors in granting credit. Their contributions to the model are higher than the mean value of the contribution factor in all models. The borrower characteristics such as *Age* and *Education* are considered imperative factors only in Model 1 (non-agricultural lending). Therefore, the results show that the bank mainly focuses on the borrower's *Asset* and *Collateral* when approving the amount of credit.

Table 4.17: Relative contribution factor of GRNN on credit availability models

Independent variables[1/]	Relative Contribution					
	Aggregate		Agriculture		Non-agriculture	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Borrower Characteristics						
Log(asset)	***0.0482***	***0.0421***	***0.0514***	***0.0447***	***0.0469***	***0.0412***
Age	0.0086	0.0117	0.0081	0.0035	***0.0356***	0.0005
Education	0.0076	0.0081	0.0021	0.0232	***0.0477***	0.0000
Credit risk proxies						
Log(collateral)	***0.0639***	***0.0546***	***0.0541***	***0.0555***	***0.0671***	***0.0593***
Return on asset	0.0134	***0.0458***	0.0131	***0.0356***	0.0040	0.0070
Leverage ratio	0.0129	0.0137	***0.0500***	0.0103	***0.0369***	0.0090
Capital turnover ratio	***0.0330***	0.0214	0.0023	0.0152	0.0108	0.0007
Relationship indicators						
Borrowing from others	0.0231	0.0222	0.0032	0.0110	0.0137	0.0169
Duration		0.0123		***0.0333***		0.0245
Dummy variables[2/]						
Sector	***0.0348***	0.0161				
(Province)	Yes	Yes	Yes	Yes	Yes	Yes
(Major production)	Yes	Yes	Yes	Yes	Yes	Yes
(Loan type)	Yes	Yes	Yes	Yes	Yes	Yes
(Loan size)	Yes	Yes	Yes	Yes	Yes	Yes
(Lending year)	Yes	Yes	Yes	Yes	Yes	Yes
No. of inputs	35	36	34	35	29	30
No. of observations	18,310	4,444	16,560	3,965	1,750	479
Mean value of the contribution factor[3/]	0.0286	0.0278	0.0294	0.0286	0.0345	0.0333

Note: 1/ Dependent variable is log(volume of credit).
2/ See Appendix 3 for the relative contribution of province, major production, loan type, loan size and lending year dummy variables.
3/ Equal to 1 over the number of inputs used in the model.
Bold and italic indicate that the relative contribution is above the mean value of contribution factors.

In addition, the relative contribution of relationship indicators (*Borrowing from others* and *Duration*) are quite low in all the models, except in Model 2 (agricultural lending), while the relative contribution of *Duration* is higher than the mean value of the contribution factor (see Table 4.17). Therefore, the results signify that *Duration* has a significant role in determining the volume of credit for agricultural lending. There is no evidence of relationship lending in the credit availability models for non-agricultural lending.

In the credit availability models for aggregate lending (Model 1), the relative contribution of *Sector* is about 3.48 percent and is relatively higher compared to the contribution factors mean value of 2.86 percent (see Table 4.17). However, in Model 2, its contribution is only about 1.61 percent. The results show that *Sector* might not be a key factor affecting the amount of credit granted in the rural lending. This is because its contribution to the model is eliminated when *Duration* is included in the model.

The results from both the regression and the GRNN models reveal the robust influence of *Asset* and *Collateral* on the amount of credit granted. Furthermore, the results indicate that *Duration* has a significant impact on the credit availability only on agricultural lending. However, the effect of *Sector* on the volume of credit granted is significant only in the regression models. The regression results indicate that the bank concentrates more on the borrower characteristics when making decisions on the quantity of loan for an agricultural loan. The relative contribution factors of the GRNN models show a relatively low contribution from the borrower characteristics, except on the *Asset*. Furthermore, the regression results suggest that the borrower characteristics (*Age* and *Education*) do not affect the volume of credit for a non-agricultural loan, but the contribution factors of the GRNN model provide evidence that they might be important factors. Therefore, only *Asset* and *Collateral* are found significant in all the models. For the remaining variables their effects on credit availability are ambiguous depending on the estimation techniques.

4.2.3 Out-of-sample forecast

The predictive power of the regression and the artificial neural networks models (GRNN, WN, and MLFN) are examined via the out-of-sample forecasting technique. The sample is randomly separated into two different sets. The estimation set consists of 80 percent of the total sample and 20 percent is the forecasting set. To evaluate the forecasting accuracy of the models, only the sample data in the estimation set is used to estimate the models, and the out-of-sample forecasting is conducted over the forecasting set.

Table 4.18 presents the out-of-sample forecast results of the regression and the artificial neural networks models on the credit availability for aggregate lending, agricultural lending, and non-agricultural lending. The out-of-sample forecast R^2 and RMSE reported on Table 4.18 are not significantly different from the R^2 and RMSE of the in-sample forecast results shown in Table 4.16. Thus, there is no evidence of overfitting problem.

The predictive results of the models for aggregate lending indicate that the GRNN can predict the amount of credit granted better than the regression and the other two neural networks models (WN and MLFN). Furthermore, the results also show that, in Model 1, both the WN and the MLFN yield the same level of accuracy as the regression model. In Model 2, however, all the neural networks models can predict the amount of credit granted better than the regression model (see Table 4.18).

For agricultural lending, the forecast results illustrate that the GRNN and the WN are superior predictive model for Model 1 (without *Duration*) and Model 2 (with Duration), respectively (see Table 4.18). Furthermore, Model 1 results indicate that the regression model can predict the volume of credit slightly better than the MLFN. Although the predicted results demonstrate that the WN is more precise than GRNN in Model 2, the R^2 and RMSE of both the GRNN and the WN are not significantly different from each other. In term of prediction accuracy, all the neural networks models in Model 2 outperform the regression model.

The out-of-sample forecast results of the credit availability models for non-agricultural lending are presented in the last section of Table 4.18. The results clearly demonstrate that the R^2 and RMSE from the neural networks models are better than the R^2 and RMSE of the regression models. The results suggest that the GRNN should be considered as the outstanding prediction model, since its R^2 and RMSE are, correspondingly, higher and lower than the other models.

Table 4.18: Out-of-sample forecast results on credit availability models

Type of networks	Model 1[1/]		Model 2[2/]	
	R^2	RMSE	R^2	RMSE
Aggregate lending				
Multiple Linear Regression (MLR)	0.7396	0.5099	0.7467	0.5022
General Regression Neural Network (GRNN)	***0.7581***	***0.4915***	***0.7707***	***0.4777***
Ward Network (WN)	0.7392	0.5099	0.7666	0.4817
Multi-layer Feed-forward Network (MLFN)	0.7322	0.5167	0.7704	0.4785
No. of Observations		3,662		888
Agricultural lending				
Multiple Linear Regression (MLR)	0.7358	0.5096	0.7675	0.4542
General Regression Neural Network (GRNN)	***0.7574***	***0.4884***	0.7825	0.4393
Ward Network (WN)	0.7444	0.5010	***0.7869***	***0.4347***
Multi-layer Feed-forward Network (MLFN)	0.7296	0.5158	0.7743	0.4472
No. of Observations		3,312		793
Non-agricultural lending				
Multiple Linear Regression (MLR)	0.7509	0.4993	0.6956	0.5689
General Regression Neural Network (GRNN)	***0.8042***	***0.4427***	***0.8345***	***0.4195***
Ward Network (WN)	0.7972	0.4506	0.7992	0.4626
Multi-layer Feed-forward Network (MLFN)	0.7628	0.4868	0.7813	0.4827
No. of Observations		350		95

Note: 1/ without Duration.
2/ with Duration.
Bold and italic indicate (alternatively) the highest R^2 or the lowest RMSE.

The out-of-sample forecast results reveal the prediction capability of the neural networks models, especially GRNN. The results show that the neural networks models provide a better predictive result than the regression model, or are at least equivalent to the regression result. Furthermore, the results also show the important of the bank-borrower relationship, since Model 2, which includes *Duration*, can predict the credit availability more accurately than Model 1. Finally, to predict the amount of credit granted, the models for agricultural and non-agricultural lending should be employed instead of utilizing the aggregate model. This is because the R^2 and RMSE of the aggregate models are, correspondingly, lower and higher than the R^2 and RMSE of the segregated models (agricultural and non-agricultural models).

4.3 Loan pricing model

Multiple linear regression and artificial neural networks are used to estimate the loan pricing models. The results of the four regression loan pricing models are presented in Table 4.19. Model 1 and 2 are the models without and with *Duration*, respectively. The volume of credit is not included in the loan pricing models, since it is assumed to be an endogenous variable. In Model 3 and 4, however, the volume of credit is assumed to be determined before the loan price (or interest rate charge). As a result, it is considered as an exogenous variable. In Model 3 and 4 (without and with Duration, respectively) the dummy variables on loan size (*Small Loan*, *Medium Loan*, and *Large Loan*) are excluded and replaced by the volume of credit in natural logarithmic form[32].

The results in Table 4.19 show that the models can explain only 13 to 31 percent of the total variation in loan price. This indicates that there are a number of unobservable factors that may be correlated with the loan price. This includes the borrower's credit history and the use of BAAC as the main channel for governmental interference in the rural financial market.

The estimated coefficients of the loan pricing models in Table 4.19 show that the value of asset and collateral are negatively correlated with the loan price. Both variables are negative and statistically significant in three of the four models. The effects of *Asset* and *Collateral* on the loan price are consistent with Strahan (1999), Degryse and Cayseele (2000), Goodwin and Mishra (2000), and Keasey and Watson (2000) findings. From Model 4, a 1 percent increased in value of asset and collateral would reduce the interest rate charge (or loan price) by 0.0012 and 0.0055 percent, respectively. The higher the value of asset and collateral, generally, would represent a safer loan contract and the borrower would be charged a lower loan rate.

[32] Bodenhorn (2003) argues that the loan amount might account for the economy of scale in lending. That is whether lower rates were paid on the larger loan. Furthermore, it may also be correlated with unobserved credit risks, if the bank allowed safer borrowers larger lending limits.

Table 4.19: Loan pricing models

Independent variables[1/]	Coefficients[2/]			
	Model 1	Model 2	Model 3	Model 4
Borrower Characteristics				
Log(asset)	-0.0011**	-0.0002	-0.0017**	-0.0012**
Age	-0.0002**	0.0000	-0.0002**	0.0000
Education	-0.0044**	0.0008	-0.0040**	0.0017
Credit risk proxies				
Log(collateral)	-0.0003	-0.0031**	-0.0022**	-0.0055**
Return on asset	-0.0005	0.0002	-0.0004	0.0002
Leverage ratio	-0.0039	-0.0102	-0.0042	-0.0120
Capital turnover ratio	-0.0002	0.0001	-0.0003**	0.0000
Log(volume of credit)			0.0129**	0.0132**
Relationship indicators				
Borrowing from others	0.0017**	0.0018	0.0022**	0.0027
Duration		-0.0070**		-0.0067**
Dummy variables[3/]				
Sector	0.0060**	0.0110**	0.0075**	0.0123**
(Province)				
Province 2	0.0073**	0.0137**	0.0066**	0.0119**
Province 3	0.0007	0.0023	0.0001	0.0007
Province 4	0.0017	0.0004	0.0011	-0.0011
Province 5	0.0010	0.0080**	0.0009	0.0068**
Province 6	0.0021	0.0106**	0.0023*	0.0091**
Province 7	0.0092**	0.0037	0.0076**	0.0017
Province 8	0.0026**	0.0166**	0.0031**	0.0162**
Province 9	0.0003	0.0077**	0.0004	0.0070**
Province 10	0.0051**	0.0097**	0.0024**	0.0070**
Province 11	0.0044**	0.0114**	0.0021	0.0087**
Province 12	-0.0072**	0.0104**	-0.0078**	0.0118**
Province 13	0.0015*	0.0048*	-0.0014	0.0022
Province 14	0.0022*	0.0093**	-0.0007	0.0069**
Province 15	0.0058**	0.0094**	0.0025	0.0062**
Province 16	0.0030**	0.0078**	0.0012	0.0060**
Province 17	0.0094**	0.0100**	0.0075**	0.0083**

Table 4.19: Loan pricing models (Cont)

Independent variables[1/]	Coefficients[2/]			
	Model 1	Model 2	Model 3	Model 4
(Major production)				
Horticulture	0.0094**	0.0135**	0.0091**	0.0142**
Orchard/Vegetable	0.0158**	0.0205**	0.0151**	0.0204**
Livestock/Aquaculture	0.0160**	0.0198**	0.0142**	0.0188**
(Loan type)				
Short-term loan	-0.0028**	-0.0070**	-0.0007	-0.0046*
Medium-term loan	-0.0017*	-0.0008	-0.0037**	-0.0024
Long-term loan	-0.0109**	-0.0066**	-0.0153**	-0.0103**
(Loan size)				
Medium loan	0.0170**	0.0149**		
Large loan	0.0290**	0.0295**		
(Lending year)				
2001	0.0072**	0.0029*	0.0067**	0.0026*
2002	0.0080**	0.0058**	0.0080**	0.0053**
Constant	0.0857**	0.1090**	-0.0172**	0.0111
R-squared	0.1380	0.2669	0.1756	0.3147
Adjusted R-squared	0.1363	0.2609	0.1740	0.3092
RMSE	0.0289	0.0261	0.0283	0.0252
F-statistic	83.55**	44.56**	114.45**	57.83**
No. of Observations	18,310	4,444	18,310	4,444

Note: 1/ Dependent variable is loan price (interest rate).
2/ White adjustment for estimation a heteroscedasticity consistent covariance matrix.
3/ To avoid singularity problem, a dummy variable is dropped from each group.
*, ** represent 10% and 5% significant level, respectively.

In Model 1 and 3, both *Age* and *Education* are negatively significant at the 5 percent level, but when *Duration* is introduced into the models (see Model 2 and 4), both variables become positive but insignificant. The results imply that the influences of some of the borrower characteristics could be eliminated by the length of the bank-borrower relationship[33].

[33] Angelini et al. (1998) found no relationship between age of the borrowing firm and the borrowing cost. In contrast, both Petersen and Rajan (1994) and Degryse and Cayseele (2000) studies show the negative relationship between the age of firm and the loan price.

Return on asset, *Leverage ratio* and *Capital turnover ratio* coefficients are not significantly different from zero at the 10 percent level in all the models, except *Capital turnover ratio* in Model 3 (see Table 4.19). Thus, the borrower financial performance appears to have no significant impact on the loan price. Furthermore, the estimated results show that *Volume of credit* has a positive significant influence on the interest rate charge, where a 1 percent increased in the volume of credit would increase the loan price by 0.0132 percent (see Model 4). This result supports the study of Goodwin and Mishra (2000) who found a positive relationship between loan size and loan price.

Borrowing from others has positive signs in all the models. However, the significant impact of *Borrowing from others* on the loan price is eliminated when *Duration* is included into the model (see Table 4.19). The estimated coefficients of *Duration* on both Model 2 and 4 are negative and statistically significant at the 5 percent level. This shows that the borrower who maintains a longer relationship with the bank would have a lower borrowing cost[34]. The results in this research are consistent with the findings of Berger and Udell (1990), Berger and Udell (1995), and Bodenhorn (2003).

The results show *Sector* is positive and significantly influence the loan rate in all four models (see Table 4.19). The result suggests that agricultural loan is charged higher than non-agricultural loan. Furthermore, the estimated coefficients of the dummy variables on major production and loan type reveal that the borrowing cost for horticultural production is lower than the other production types and the *Long-term loan* is fairly cheaper than the other types of loans. The results of *Medium loan* and *Large loan* in Model 1 and 2 also confirm that loan size has a positive influence on the loan price, as *Large loan* is charged more than 1 percent higher than *Medium loan*. In addition, the coefficients of provinces exhibit the province effect on the loan price, where some provinces are charged higher and some provinces are charged lower.

The results confirm that the volume of credit (loan size) is positively related to the credit risk and loan price. The borrower with higher level of wealth, as reflected in the value of asset, appears to have lower lending rate. Furthermore, a higher value of collateral pledged can command a lower loan rate. The loan price varies significantly across the provinces,

[34] Angelini et al. (1998) argue that close relation helps the bank to overcome information asymmetry problem, consequently, the bank shares this gain with the borrower by lower lending rate.

major productions, loan types and lending years. Since *Duration* is found to be negative and highly significant, it can be concluded that the relationship lending exists and is important to the borrower.

4.3.1 Loan pricing models for agricultural and non-agricultural lending

The results of loan pricing models for agricultural lending are shown in Table 4.20. The results are quite similar to the loan pricing models for aggregate lending in Table 4.19. The results show that *Asset*, *Age* and *Education* have a negative relationship with the loan price (see Table 4.20, Models 1 and 3). However, their influences are eliminated and become insignificant when *Duration* is included into the model (see Model 2 and 4). All the financial performance indicators, *Return on asset*, *Leverage ratio* and *Capital turnover ratio*, are not significant in all the models, except *Capital turnover ratio* in Model 3. Although the estimated coefficients of *Borrowing from others* have the hypothesised signs, only one of the four models is significant at the 5 percent level. Therefore, the loan price for an agricultural loan is primarily determined by *Collateral*, *Volume of credit* and *Duration*, regardless of the impacts of control variables such as province, loan type, etc.

As hypothesised, the results show that both *Collateral* and *Duration* are negatively related to the loan price, while the *Volume of credit* is positively related to the loan price. The impact of *Collateral* on the loan price is much smaller than the *Volume of credit*, as shown in Model 4 (see Table 4.20), where a 1 percent increased in *Volume of credit* would increase the interest rate by 0.0095 percent, but a 1 percent increased in value of *Collateral* would decrease the loan price by only 0.0042 percent. The results show that *Duration* is negatively significant at the 5 percent level in Model 2 and 4, and the effect of the relationship lending on the loan pricing process is validated.

The results of the control variables in Table 4.20 disclose that the bank offers different lending rates to different provinces, as the credit risk of each province is different. Furthermore, the results also suggest that horticultural production is less risky and is priced slightly lower than the other agricultural productions. In addition, long-term loan is relatively cheaper than other types of loans. In addition, the results from the control variable groups (province, major production, loan type, and loan size dummy variables) are similar to the findings found on the loan pricing models for aggregate lending in Table 4.19.

Table 4.21 presents the results of loan pricing models for non-agricultural lending. The results are quite similar to the finding in Table 4.19 and 4.20. However, only *Asset* and *Age* are negative and significant in Model 1 and 3, but not *Education*. Therefore, the *Education* level of the borrower has no impact on the interest rate charged for non-agricultural lending, compared to the agricultural lending. All variables (*Asset*, *Age* and *Education*) become insignificant in Model 2 and 4 when *Duration* is included into the models. The results suggest that both *Asset* and *Age* are important factors and have a significant influence on the loan price, but both variables are dominated by *Duration*.

In addition, *Collateral*, *Volume of credit* and *Duration* have the hypothesised signs and are highly significant in all the models (see Table 4.21). The coefficients of *Borrowing from others* are positively significant only in Model 1 and 3, and its impact on the loan price is deleted with the addition of *Duration* (see Models 2 and 4).

The result from cross comparison between agricultural and non-agricultural lending clearly indicates that the interest rate charge on non-agricultural lending is more responsive to the loan size (amount of credit) than agricultural lending. The estimated coefficients of *Volume of credit* on Model 4 in Table 4.20 and 4.21 show that a 1 percent increased in the loan amount would increase the lending rate for non-agricultural and agricultural loan by 0.0312 and 0.0095 percent, respectively. Furthermore, *Collateral* indicates that a 1 percent increased in value of collateral would decrease the borrowing cost for non-agricultural loan and agricultural loan by 0.0141 and 0.0042 percent, respectively. It can be concluded that *Collateral* has a stronger impact on the loan price for non-agricultural lending. In terms of the borrowing cost, the influence of the relationship lending is more important for agricultural lending than non-agricultural lending (see Table 4.20 and 4.21).

Table 4.20: Loan pricing models for agricultural lending

Independent variables[1/]	Coefficients[2/]			
	Model 1	Model 2	Model 3	Model 4
Borrower Characteristics				
Log(asset)	-0.0007**	0.0001	-0.0013**	-0.0007
Age	-0.0002**	0.0000	-0.0002**	0.0000
Education	-0.0051**	0.0000	-0.0047**	0.0007
Credit risk proxies				
Log(collateral)	-0.0002	-0.0023**	-0.0019**	-0.0042**
Return on asset	-0.0004	0.0003	-0.0003	0.0002
Leverage ratio	-0.0024	-0.0068	-0.0026	-0.0082
Capital turnover ratio	-0.0002	0.0001	-0.0003**	0.0000
Log(volume of credit)			0.0114**	0.0095**
Relationship indicators				
Borrowing from others	0.0013	0.0015	0.0017**	0.0020
Duration		-0.0072**		-0.0069**
Dummy variables[3/]				
(Province)				
Province 2	0.0063**	0.0126**	0.0057**	0.0114**
Province 3	0.0000	0.0028	-0.0005	0.0018
Province 4	0.0008	-0.0001	0.0002	-0.0018
Province 5	0.0012	0.0055	0.0011	0.0047
Province 6	0.0012	0.0114**	0.0014	0.0103**
Province 7	0.0094**	0.0031	0.0083**	0.0023
Province 8	0.0011	0.0158**	0.0018*	0.0159**
Province 9	-0.0013	0.0071**	-0.0011	0.0069**
Province 10	0.0074**	0.0142**	0.0049**	0.0121**
Province 11	0.0088**	0.0166**	0.0066**	0.0142**
Province 12	-0.0082**	0.0108**	-0.0091**	0.0109**
Province 13	0.0041**	0.0097**	0.0015	0.0075**
Province 14	0.0015	0.0126**	-0.0016	0.0106**
Province 15	0.0077**	0.0132**	0.0046**	0.0107**
Province 16	0.0107**	0.0139**	0.0090**	0.0124**
Province 17	0.0105**	0.0125**	0.0086**	0.0112**

Table 4.20: Loan pricing models for agricultural lending (Cont)

Independent variables[1/]	Coefficients[2/]			
	Model 1	Model 2	Model 3	Model 4
(Major production)				
Horticulture	0.0106**	0.0154**	0.0105**	0.0162**
Orchard/Vegetable	0.0140**	0.0179**	0.0136**	0.0181**
Livestock/Aquaculture	0.0154**	0.0183**	0.0139**	0.0179**
(Loan type)				
Short-term loan	-0.0022**	-0.0067**	-0.0006	-0.0049*
Medium-term loan	-0.0019*	0.0001	-0.0039**	-0.0011
Long-term loan	-0.0094**	-0.0041**	-0.0136**	-0.0072**
(Loan size)				
Medium loan	0.0145**	0.0097**		
Large loan	0.0219**	0.0213**		
(Lending year)				
2001	0.0023*	0.0005	0.0019	0.0004
2002	0.0081**	0.0052**	0.0082**	0.0049**
Constant	0.0847**	0.1052**	-0.0037	0.0381**
R-squared	0.1382	0.2522	0.1706	0.2835
Adjusted R-squared	0.1364	0.2456	0.1689	0.2773
RMSE	0.0282	0.0239	0.0277	0.0233
F-statistic	77.91**	37.86**	102.99**	45.74**
No. of Observations	16,560	3,965	16,560	3,965

Note: 1/ Dependent variable is loan price (interest rate).
2/ White adjustment for estimation a heteroscedasticity consistent covariance matrix.
3/ To avoid singularity problem, a dummy variable is dropped from each group.
*, ** represent 10% and 5% significant level, respectively.

In summary, the regression results confirm the significant impacts of *Collateral, Volume of credit* and *Duration* on the loan price. *Asset*, *Age* and *Education* also have an impact on the loan rate, but their influences are overlooked when *Duration* is included in the model. Thus, the effect of relationship lending on the loan price is imperative, especially in agricultural lending. In addition, the lending rate for non-agricultural lending is quite sensitive to the value of *Collateral* and *Volume of credit.*

Table 4.21: Loan pricing models for non-agricultural lending

Independent variables[1/]	Coefficients[2/]			
	Model 1	Model 2	Model 3	Model 4
Borrower Characteristics				
Log(asset)	-0.0024**	-0.0011	-0.0034**	-0.0033
Age	-0.0002**	0.0002	-0.0001**	0.0002
Education	0.0014	0.0049	0.0012	0.0051
Credit risk proxies				
Log(collateral)	-0.0033**	-0.0089**	-0.0071**	-0.0141**
Return on asset	-0.0005	-0.0012	-0.0005	-0.0006
Leverage ratio	-0.0133	-0.0176	-0.0144	-0.0224*
Capital turnover ratio	0.0000	0.0006	-0.0002	0.0000
Log(volume of credit)			0.0254**	0.0312**
Relationship indicators				
Borrowing from others	0.0057**	0.0059	0.0069**	0.0075
Duration		-0.0038**		-0.0038**
Dummy variables[3/]				
(Province)				
Province 1	0.0185**	0.0227**	0.0191**	0.0267**
Province 2	0.0216**	0.0208**	0.0234**	0.0212**
Province 3	0.0221**	0.0070	0.0173**	-0.0106
Province 4	0.0330**	0.0051	0.0351**	0.0252**
Province 5	0.0190**	0.0226**	0.0201**	0.0273**
Province 6	0.0233**	0.0182**	0.0250**	0.0232**
Province 7	0.0224**	0.0124**	0.0182**	0.0032
Province 8	0.0238**	0.0320**	0.0254**	0.0309**
Province 9	0.0238**	0.0256**	0.0268**	0.0202**
Province 10	0.0198**	0.0241**	0.0193**	0.0199**
Province 12	0.0339**	0.0373**	0.0460**	0.0572**
Province 13	0.0260**	0.0136	0.0278**	0.0258*
Province 14	0.0249**	0.0088	0.0251**	0.0096
Province 15	0.0266**	0.0133*	0.0270**	0.0099*
Province 17	0.0285**	0.0145*	0.0309**	0.0197**

Table 4.21: Loan pricing models for non-agricultural lending (Cont)

Independent variables[1/]	Coefficients[2/]			
	Model 1	Model 2	Model 3	Model 4
(Major production)				
Orchard/Vegetable	0.0054**	0.0146**	0.0057**	0.0120**
Livestock/Aquaculture	0.0062**	0.0155**	0.0080**	0.0081**
(Loan type)				
Long-term loan	-0.0088**	-0.0087	-0.0104**	-0.0152**
(Loan size)				
Medium loan	0.0348**	0.0510**		
Large loan	0.0682**	0.0837**		
(Lending year)				
2002	0.0059**	0.0048	0.0061**	0.0051
Constant	0.1221**	0.1689**	-0.0901**	-0.0617**
R-squared	0.2703	0.3961	0.3645	0.5101
Adjusted R-squared	0.2580	0.3557	0.3542	0.4785
RMSE	0.0316	0.0347	0.0295	0.0313
F-statistic	21.98**	9.80**	35.25**	16.12**
No. of Observations	1,750	479	1,750	479

Note: 1/ Dependent variable is loan price (interest rate).
2/ White adjustment for estimation a heteroscedasticity consistent covariance matrix.
3/ To avoid singularity problem, a dummy variable is dropped from each group.
*, ** represent 10% and 5% significant level, respectively.

4.3.2 Artificial neural networks and loan pricing models

Table 4.22 presents the neural networks R^2 and RMSE on loan pricing models. The in-sample estimation results indicate that the GRNN is the best network for loan pricing because it has the highest R^2 and lowest RMSE on both Model 1 and 2 (with and without *Duration*) of aggregate lending, agricultural lending, and non-agricultural lending. Furthermore, the R^2 and RMSE of the neural networks models are, correspondingly, higher and lower than the R^2 and RMSE of the regression models (see Table 4.22).

Table 4.22: Neural Networks R squared (R^2) and Root Mean Squared Error (RMSE) on loan pricing models

Type of networks	Model 1[1/]		Model 2[2/]	
	R^2	RMSE	R^2	RMSE
Aggregate lending				
Multiple Linear Regression (MLR)	0.1756	0.0283	0.3147	0.0252
General Regression Neural Network (GRNN)	***0.3752***	***0.0246***	***0.5297***	***0.0209***
Ward Network (WN)	0.3066	0.0259	0.5100	0.0213
Multi-layer Feed-forward Network (MLFN)	0.2649	0.0267	0.4807	0.0219
No. of Observations		18,310		4,444
Agricultural lending				
Multiple Linear Regression (MLR)	0.1706	0.0277	0.2835	0.0233
General Regression Neural Network (GRNN)	***0.5300***	***0.0208***	***0.6069***	***0.0173***
Ward Network (WN)	0.2920	0.0256	0.4249	0.0209
Multi-layer Feed-forward Network (MLFN)	0.2542	0.0262	0.3903	0.0215
No. of Observations		16,560		3,965
Non-agricultural lending				
Multiple Linear Regression (MLR)	0.3645	0.0295	0.5101	0.0313
General Regression Neural Network (GRNN)	***0.6035***	***0.0233***	***0.9698***	***0.0078***
Ward Network (WN)	0.4704	0.0270	0.6338	0.0270
Multi-layer Feed-forward Network (MLFN)	0.4790	0.0267	0.5871	0.0287
No. of Observations		1,750		479

Note: 1/ and 2/ are the models without and with Duration, respectively, assuming that volume of credit is determined before the loan price. Therefore, dummy variables on the loan size are excluded and replaced by log(volume of credit).
Bold and italic indicate (alternatively) the highest R^2 or the lowest RMSE.

Since the performance of Model 2 (with *Duration*) is better than the Model 1 (without *Duration*), it can be concluded that the length of the bank-borrower relationship is an important factor in determining the lending rate. Furthermore, a cross comparison between the loan pricing models for agricultural lending and non-agricultural lending illustrates the neural networks models could recognize the pattern and the variation of the loan rate in non-agricultural lending much better than in agricultural lending. The result implies that the interest rate charge on non-agricultural loan contract is more uniformity.

Table 4.23: Relative contribution factor of GRNN on loan pricing models

Independent variables[1/,2/]	Relative Contribution					
	Aggregated		Agriculture		Non-agriculture	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Borrower Characteristics						
Log(asset)	***0.0466***	0.0100	0.0136	0.0101	***0.0495***	***0.0459***
Age	0.0278	0.0135	0.0187	0.0060	0.0108	***0.0364***
Education	0.0027	0.0046	0.0088	***0.0345***	0.0210	0.0045
Credit risk proxies						
Log(collateral)	***0.0369***	0.0254	***0.0324***	***0.0467***	***0.0361***	***0.0379***
Return on asset	0.0082	0.0086	***0.0445***	0.0009	0.0077	0.0058
Leverage ratio	0.0103	0.0073	0.0171	0.0170	0.0011	0.0184
Capital turnover ratio	0.0203	0.0168	0.0196	***0.0400***	0.0011	0.0324
Log(volume of credit)	***0.0483***	***0.0549***	***0.0667***	***0.0582***	***0.0720***	***0.0571***
Relationship indicators						
Borrowing from others	0.0228	0.0071	0.0139	0.0203	0.0003	0.0232
Duration		***0.0409***		***0.0555***		0.0297
Dummy variables[3/]						
Sector	***0.0411***	***0.0303***				
(Province)	Yes	Yes	Yes	Yes	Yes	Yes
(Major production)	Yes	Yes	Yes	Yes	Yes	Yes
(Loan type)	Yes	Yes	Yes	Yes	Yes	Yes
(Loan size)	No	No	No	No	No	No
(Lending year)	Yes	Yes	Yes	Yes	Yes	Yes
No. of inputs	34	35	33	34	28	29
No. of observations	18,310	4,444	16,560	3,965	1,750	479
Mean value of the contribution factor[4/]	0.0294	0.0286	0.0303	0.0294	0.0357	0.0345

Note: 1/ Dependent variable is loan price (interest rate).
2/ Assume that volume of credit is determined before the loan price. Therefore, dummy variables on the loan size are excluded and replaced by log(volume of credit).
3/ See Appendix 4 for the relative contribution of province, major production, loan type and lending year dummy variables.
4/ Equal to 1 over the number of inputs used in the model.
Bold and italic indicate that the relative contribution is above the mean value of the contribution factors.

Table 4.23 shows the relative contribution factor of the neural networks on loan pricing models. However, only the results from the GRNN are shown because it is considered as the superior network. To verify which factor has a significant impact on the loan price, the relative contribution of each factor is compared with the model's relative contribution mean value.

The relative contribution factors shown in Table 4.23 confirm the important role of the *Volume of credit* and *Collateral* in determining the loan price, since both factors contribute to the model more than the mean value of the relative contribution factors. Furthermore, the contribution values of *Duration* on the loan pricing models for agricultural lending and non-agricultural lending suggest that the length of the bank-borrower relationship has a stronger influence on the agricultural lending rate. The value of *Asset* has a significant influence only on the non-agricultural lending. In addition, the high relative contribution of *Sector* indicates that *Sector* has a significant impact in determining the loan price.

The impact of *Age* and *Education* on the loan price is not obvious in the GRNN models, compared to the regression models (see Table 4.23). The considerable effect of *Return on asset*, *Leverage ratio* and *Capital turnover ratio* to the borrowing cost can not be detected on the loan pricing models for aggregate lending and non-agricultural lending. Nevertheless, a strong influence of *Return on asset* and C*apital turnover ratio* on the loan price is noticeable only in Model 1 and 2 of the agricultural lending, respectively.

The results from the regression models and the GRNN models reveal that there are three major factors influencing the loan pricing. They are *Collateral*, *Volume of credit* and *Duration* (see Table 4.19, 4.20, 4.21 and 4.23). Despite *Age* and *Education* having a significant impact on the lending rate in the regression models (but their influences can be purged by *Duration*), the relative contributions of both factors are lower than the contribution factors mean values in most of the GRNN models (see Table 4.23). The effect of *Return on asset*, *Leverage ratio* and *Capital turnover ratio* on the interest rate charged are not significant in the regression and the GRNN models. It can be concluded that these variables have no impact on the loan pricing. The relative contributions of *Borrowing from others* are lower than the mean value of the relative contribution factors in all the GRNN models and its estimated coefficients are significant only in 3 regression models (see Table 4.19 and 4.20), but the relationship lending is still important because the robust influence of the *Duration* on the loan price could be detected on both the regression and the GRNN models.

4.3.3 Out-of-sample forecast

The out-of-sample forecast evaluation results of the loan pricing models are shown in Table 4.24. The R^2 and RMSE of all neural networks models are, respectively, higher and lower than the R^2 and RMSE of the regression models. Therefore, the results suggest that the artificial neural networks models can predict the loan price more accurately than the regression models. Furthermore, the results also indicate that Model 2 which includes *Duration* yields a higher R^2 and lower RMSE than Model 1. Hence, it can be concluded that the length of the bank-borrower relationship plays an important role in the loan pricing model as it can improve the models' predictive power significantly.

Table 4.24: Out-of-sample forecast results of loan pricing models

Type of networks	Model 1[1/]		Model 2[2/]	
	R^2	RMSE	R^2	RMSE
Aggregate lending				
Multiple Linear Regression (MLR)	0.1602	0.0282	0.2942	0.0264
General Regression Neural Network (GRNN)	***0.2617***	***0.0265***	0.3475	0.0254
Ward Network (WN)	0.2606	0.0265	0.4231	0.0239
Multi-layer Feed-forward Network (MLFN)	0.2416	0.0268	***0.4233***	***0.0239***
No. of Observations		3,662		888
Agricultural lending				
Multiple Linear Regression (MLR)	0.1726	0.0282	0.2355	0.0236
General Regression Neural Network (GRNN)	***0.2616***	***0.0266***	***0.3289***	***0.0221***
Ward Network (WN)	0.2614	0.0266	0.3062	0.0225
Multi-layer Feed-forward Network (MLFN)	0.2379	0.0270	0.3021	0.0226
No. of Observations		3,312		793
Non-agricultural lending				
Multiple Linear Regression (MLR)	0.3218	0.0302	0.4950	0.0314
General Regression Neural Network (GRNN)	***0.4234***	***0.0278***	***0.8448***	***0.0174***
Ward Network (WN)	0.3994	0.0284	0.6668	0.0255
Multi-layer Feed-forward Network (MLFN)	0.4015	0.0284	0.6584	0.0258
No. of Observations		350		95

Note: 1/ and 2/ are the models without and with Duration variable, respectively, assuming that volume of credit is determined before the loan price. Therefore, dummy variables on the loan size are excluded and replaced by log(volume of credit).
Bold and italic indicate (alternatively) the highest R^2 or the lowest RMSE.

However, it should be noted that, in most cases, the out-of-sample forecast R^2 is lower than the in-sample R^2 (see Table 4.21 and 4.23). The results highlight the overfitting problem on both the regression and the artificial neural networks models. In addition, it appears that the overfitting problem is enormous using the GRNN model, which is regarded as the superior model for in-sample forecast. Despite the overfitting problem on the GRNN models, the models have the highest R^2 and lowest RMSE in most cases of the out-of-sample forecasting, except on the Model 2 of aggregate lending (see Table 4.24). Therefore, it is reasonable to conclude that the GRNN is the superior model for loan price prediction.

4.4 Summary of findings

In summary, the results show that *Sector* has a significant impact on the bank lending decision, the volume of credit granted, and the interest rate charged. The agricultural borrowers have a higher probability of a good loan compared to non-agricultural borrowers. This implies that they have a higher probability in obtaining a loan from BAAC. However, they are charged at a higher loan rate and receive a smaller amount of loan than the non-agricultural borrowers. The influence of the relationship lending on the bank lending decision, the credit availability and the loan price is imperative, especially on agricultural lending. The results of the agricultural lending show that BAAC utilizes the information obtained from the relationship to monitor the borrower's credit risk through the bank lending decision and the credit availability. However, a long term relationship with the bank benefits the borrower via loan pricing.

The total asset value and the value of collateral are also important in explaining the bank lending decision, credit availability, and loan price. The results indicate that the probability of a good loan increases with increased in the total asset value. Furthermore, the amount of credit granted is increased with increased in the total asset value and the value of collateral, while the loan price is decreased with increased value of collateral. The results also show that loan price is positively related to the volume of credit. For the rest variables (*Age*, *Education*, *Return on asset*, *Leverage ratio,* and *Capital turnover ratio*) their effects on the bank lending decision, credit availability, and loan price are ambiguous.

The forecast results of both the in-sample and the out-of-sample on the bank lending decision, credit availability, and loan price show that most of the artificial neural networks models outperform the logistic and multiple regression analysis techniques. In addition, the results indicate that the PNN models can be successfully implemented to screen the borrowers, and the GRNN models can be employed to determine the volume of credit granted and the interest rate charged.

CHAPTER 5

SUMMARY AND CONCULSIONS

This chapter summarizes the research. Section 5.1 presents a summary of the research objectives, data and methodology, and major findings. The implications of the research findings are discussed in Section 5.2. Section 5.3 discuses the research limitations and section 5.4 provides recommendations for further research.

5.1 Summary and major findings

Rural financing in Thailand is heavily dependent on bank lending. However, banks prefer to lend to large commercial farmers rather than to small-scale peasant farmers who are poor and lack investment funds. Thus understanding the bank lending process for the rural sector is an important element for promoting the development of credit accessibility to Thai farmers in rural regions. Thus, appropriate bank lending decisions would reduce lending costs and increase repayment rate and operating profits to the banks. As a consequence, a well-developed rural financial market would lead to a sustainable development in the rural sector.

The purpose of this research is to identify critical factors in the bank lending decision process and to investigate factors affecting the amount of credit granted and interest rate charged for rural lending in Thailand. This research also investigates the impact of the relationship between the lender and the borrower on the lending decision, credit availability and credit price. Furthermore, it examines the predictive power among the different estimation techniques in predicting the bank lending decision, amount of credit granted, and interest rate charged.

The data used in this research are obtained from the Bank for Agriculture and Agricultural Cooperative (BAAC), a major lender in Thailand agricultural sector. During the period of 2001 to 2003, a total of 18,798 credit files under the normal loan scheme (excluding the government loans for specific projects) from between 99 to 136 branches in 17 provinces are made available. The credit scoring model is analyzed using the logistic regression (Logit) and four different types of the artificial neural network, namely multi-layer feed-forward neural network (MLFN), Ward network (WD), general regression neural network (GRNN), and probabilistic neural network (PNN). The credit availability and loan pricing models are estimated by multiple linear regression (MLR) and three artificial neural network models, including MLFN, WD, and GRNN.

Table 5.1 summarizes the estimated results of all the models and Table 5.2 presents the models' ranking in term of predictive accuracy. In summary, the results show that:

- Relationship lending in agricultural lending is important in explaining the bank lending decision, credit availability, and loan price.
- On agricultural lending, the length of the bank-borrower relationship (*Duration*) has a negative impact on the bank lending decision, volume of credit, and loan price. This implies that the bank utilizes information of the borrowers and monitors the lending risk via lending decision and credit availability. However, banking relationships benefit the borrowers via loan pricing since the borrowers with a long term relationship with the bank commands a lower borrowing rate. In contrast, for non-agricultural lending, *Duration* influences only loan price and has no impact on lending decision or volume of credit.
- If the borrower has an outstanding debt with other creditors, it would reduce the amount of credit granted (particularly for the agriculture borrowers) and increase the lending rate, but it has no significant impact on the bank lending decision.
- Total asset value (*Asset*) has a positive impact on lending decision and credit availability. On the other hand, it has a negative impact on credit price, but the impact is significant only in the models without *Duration*. The relative contribution factors of the GRNN models show that *Asset* has a strong influence on loan pricing.
- Age of borrower (*Age*) and borrower's education level (*Education*) have no significant effect on all the models including *Duration*, except on the credit availability model for agricultural lending, where *Education* shows a significant negative relationship. The results imply that the influence on some of the borrower characteristics could be eliminated by the bank-borrower relationship. Furthermore, the results shows that the borrower who has higher education tend to borrow smaller amount.

Table 5.1: Factors affecting the bank lending decision, credit availability, and loan price

Factors	Lending decision	Volume of credit	Loan price
Borrower characteristics			
Asset	(+)	(+)	• (-) only in the model without Duration. • High contribution on GRNN model
Age	(0) • High contribution in PNN model	• (-) only in the model without Duration. • (0) on non-ag. lending	• (-) only in the model without Duration
Education	• (+) only in the models without Duration • (0) on non-ag. lending	(-) • (0) on non-ag. lending	• (-) only in the model without Duration • (0) on non-ag. lending
Credit risk proxies			
Collateral	• Negative in most cases but not sig. • High contribution in PNN models	(+)	(-)
Return on asset	• (+) only on non-ag. lending	(0)	(0)
Leverage ratio	• (-) only in the model without Duration	(0)	(0)
Capital turnover ratio	(-)	(+) • But not sig. in some models	(0)
Volume of credit			(+)
Relationship indicators			
Borrowing from others	(0)	(-) • (0) on non-ag. lending	• (+) only in the model without Duration
Duration	(-) • (0) on non-ag. lending	(-) • (0) on non-ag. lending	(-)
Dummy variables			
Sector	(+)	(-)	(+)
Province	✓ Province effect	✓ Province effect	✓ Province effect
Major production	• Horticulture > Orchard > Livestock (ag. lending)	• (0) (ag. lending) • Livestock > Hort., Orchard (non-ag.)	• Orchard > Livestock > Horticulture
Loan type	• Short > Medium > Long (ag. lending)	• Long > Medium > Short (ag. lending)	• Medium, Short > Long
Loan size	• Small > Medium > Large (ag. lending)	• Large > Medium > Small	• Large > Medium > Small
Lending year	• High default risk in 2002		• 2002 > 2001, 2003 (ag. lending)

Note: (+), (-), and (0) represent positive, negative, and no significant impact, respectively.

Table 5.2: Ranking of the prediction models based on prediction accuracy of the model

	Lending decision	Volume of credit	Loan price
In-sample forecast	1. PNN 2. GRNN 3. WN, MLFN, and Logit	1. GRNN 2. WN 3. MLFN 4. MLR[2/]	1. GRNN 2. WN 3. MLFN 4. MLR
Out-of-sample forecast	Aggregate and Agricultural lending 1. GRNN 2. PNN[1/], WN, MLFN, and Logit Non-agricultural lending 1. PNN[1/] 2. GRNN 3. WN, MLFN, and Logit	1. GRNN 2. WN 3. MLFN 4. MLR[2/]	1. GRNN 2. WN 3. MLFN 4. MLR

Note: 1/ PNN (with Duration) yields the lowest misclassification cost in most cases and can be considered as the superior model in term of relative error cost.
2/ WN and MLFN provide the same accuracy level as MLR in some models.

- Value of collateral (*Collateral*) has a positive and a negative influence on credit availability and loan pricing models, respectively, but it has no impact on the bank lending decision model. However, it should be noted that most of the *Collateral* coefficients on the bank lending decision models are negative, but they are not significantly different from 0. This indicates that higher collateral pledged means a risky loan with a high default risk. In addition, the relative contribution factors of the PNN models show that *Collateral* has a strong influence on the bank's lending decision.
- There is a reverse relationship between *Capital turnover ratio* and lending decision. This result suggests that the borrower who has a higher income to assets tend to default on debt repayment compared to the borrower who has a lower income to assets. Generally, the borrower who has a relatively high income may prefer to spend his or her money on other activities rather than repaying the debt.

- *Return on assets* has significant impact only on the lending decision model for non-agricultural lending.
- Agricultural lending has a lower default risk (a higher probability of good loan) compare to non-agricultural lending. However, the borrower in the agricultural sector is charged higher interest rates and receives a smaller amount of credit due to higher risk in the agricultural sector.
- Bank lending decision, credit availability, and loan price differ according to the provinces. The results suggest that there is a province effect on the bank lending decision, the volume of credit granted, and the interest rate charged. The results can not explain the discrimination on the bank lending policy, since different provinces have different levels of risk.
- Horticulture farm in general has a lower default risk and receives a lower interest rate compared to other types of farm and farming activities. Nevertheless, the amount of credit granted is not significantly different according to the farm type and the major production of the farm, since the borrower may need a credit for other farming activities apart from the major production.
- For non-agricultural lending, if the borrower's major production is livestock or aquaculture, the borrower receives a larger amount of loan in terms of non-agricultural borrowing.
- *Long-term loan* and *Large loan* on agricultural lending have a higher default risk. *Large loan* is charged a higher interest rate. However, *Long-term loan* is charged a lower interest rate. This is because most of the long-term loans are for purchasing long-term asset that can be pledged as collateral to secure the loan.
- The default risk in 2002 is higher than in other years, due to the implementation of the debt suspension programme of the Thai government in 2001.
- According to the in-sample classification results, most of the artificial neural networks yield almost the same classification results as the logistic regression on the bank lending decision models. However, the probabilistic neural networks (PNN) and the general regression neural networks (GRNN) predict better than the logistic regression model and they can also detect a Type I error (wrongly reject H_0 or accept a bad loan as a good loan) much better than the logistic regression. The PNN is considered the best prediction model since it has the highest overall percentage collect classification of all the models.

- In the out-of-sample forecast, the PNN is no longer the best classification model in term of classification accuracy. However, it can be regarded as the superior model for predicting the bank lending decision. The PNN with *Duration* can identify Type I error much better than the other models and offers the lowest expected misclassification loss on out-of-sample forecast.
- For credit availability models, most of the artificial neural networks models can predict slightly better than the multiple regression technique in the in-sample forecast. However, on loan pricing models, all the artificial neural networks models perform better than the multiple regression model in terms of in-sample forecast accuracy. Both the in-sample and the out-of-sample forecast results point out that the GRNN is the superior prediction model to both credit availability and loan pricing in term of prediction accuracy.
- The forecast results of in-sample and out-of-sample on the bank lending decision, credit availability, and credit price show that most of the artificial neural networks models outperform the logistic regression and the multiple regression models, especially on the loan pricing models. In addition, the results indicate the superiority of using the PNN model to classify and screen the borrowers, and the GRNN model to determine the volume of credit granted and loan price.

5.2 Implications of the research findings

The findings of this research have important implications for academics, borrowers, banks, and policy makers. For academics, the significant effects of the relationship indicators in the bank lending decision model, credit availability model, and loan pricing model imply that the relationship lending (relationship building between bank and borrower) can assist the bank (or lender) to overcome the incomplete and asymmetric information problems in the bank lending decision process. The research results also show that the PNN and the GRNN can be successfully implemented to screen the loan applications, and to predict the amount of credit granted and loan price, respectively. Thus, this research exhibits the potential of the neural methodology, especially the PNN and the GRNN, as an analysis tool for generalization problems (classification and prediction). However, the results in this research also show that the neural networks models might not necessarily perform better than the logistic and the multiple regression models. This is because some of the artificial neural network models yield almost the same level of accuracy as the logistic and the multiple regression models on both in-sample and out-of-sample forecasting.

For the borrowers, the research findings show that borrowing from many sources of funds can deteriorate the credit availability and raise the loan price. Therefore, the borrowers should use or rely on a single financial source rather than dealing with many financial sources. In addition, the research findings show that the borrowers who have a longer relationship with the bank should receive a lower loan rate. Hence, it is a good idea for the borrowers to maintain the relationship with the bank, since it can benefit them via the loan price.

For the banks, the results found in this research show that a good credit scoring model which has the ability to detect a bad loan could help the bank to reduce the loan losses from bad borrowers. Consequently, it would improve the financial stability and the profitability of the bank. Therefore, the credit scoring model should be developed and used to support credit officers in screening loan applications.

The results from the bank lending decision model, credit availability model, and loan pricing model show that the bank pays much attention to two major factors when making the lending decisions. These factors include the total asset value and the value of collateral. Focusing on these two factors alone might cause over lending and under pricing problems when there is a financial crisis or a high depreciation on the land or asset price. Hence, the banks should consider the potential of the borrower's capability to repay the loan.

For policy makers, to promote the development of credit accessibility to farmers in rural regions, the property rights reformation and land titling programme must be accelerated. This is because the programme has progressed slowly during the past decades, but both asset and collateral play a very important role in determining the bank lending decision, the amount of credit granted, and the interest rate charge.

Since the default risk in 2002 is higher than the other years, the research findings show that the three-year debt moratorium programme introduced by the Thai government under Prime Minister Dr. Thaksin Shinawatra and his Thai Rak Thai Party (TRT) distorts the country's credit culture and encourages the good debtors to default on the debt repayments.

An inappropriate rural financial policy is not only inefficient in achieving the key goals of improving the farmers' welfare, but also distorts the country's monetary and financial culture. Therefore, the policy makers should be more concerned about this issue when implementing or introducing new rural financial policy.

5.3 Research limitations

There are a number of limitations related to the data set, the estimation techniques, and the variables used in this research. These include:

- The data set used in this research is from BAAC only and it covers only 17 out of 76 provinces. Therefore, the results in this research may not be applicable to the whole country (even if the data set covers all four different geographic regions). Moreover, they can not represent the lending behaviour of all the commercial banks in the rural financial market in Thailand.
- Only the data and information of the applicants who have been granted a credit in the past are observed. There is no data and information on the applicants who were rejected. Thus, the models are parameterised using a sample of accepted applicants only. This may lead to biased estimates of the parameters.
- The length of the bank-borrower relationship (*Duration*) in the data set is restricted to a maximum of seven years due to the lack of available information. Thus, *Duration* is a censoring variable which may cause inconsistent estimates of *Duration*.
- The models ignore the potential exposure to the future credit risk. This is because credit scoring model, credit availability model, and loan pricing model are typically static models in nature. In addition, the information about financial distress of the borrower in the past is not included in the models.
- In the case where there is a natural disaster, such as flooding or drought, BAAC may allow the borrower to postpone debt repayment or reschedule the repayment period. However, in some cases, the borrower and the credit officer may decide to use the "ever green" strategy (borrowing a new debt to repay an old debt) to resolve the overdue problem. As a result, a bad loan can become a good loan. Unfortunately, this research can not differentiate the use of this tactic.
- In this research, the cut-off point of the bank lending decision model is set equal to 0.50. Since the classification results can be improved by adjusting the cut-off point,

fixing the cut-off point of the model at 0.50 may limit the classification ability of the model.

- There are some drawbacks on using the neural networks. They are (Limsombunchai et al., 2004, and Gan et al., 2005):
  - Firstly, the neural networks lack theoretical background concerning the explanatory capabilities. The connection weights in the networks can not be interpreted or used to identify the relationships between dependent and independent variables. This means the neural networks are regarded as a "black box".
  - Secondly, there are no formal techniques for non-linear methods to test the relative relevance of the independent variables and to carry out the variable selection process. This is because the traditional statistical tests are either impossible or meaningless on the neural networks models.
  - Lastly, the neural networks learning process can be very time consuming, especially when the data set is large, and the model consists of many independent variables.

5.4 Recommendations for future research

To improve the research results and to increase the generalizability of the research findings, the credit files from other commercial banks should be collected and included in the data set. Furthermore, the numbers of the sample province should be increased and the information of the applicants who did not qualify for the credit should be taken into account.

In addition, there are a number of variables that can be added to the models to enhance the performance of the models. These include geographic distance between the lending bank and the borrower (see Degryse and Ongena, 2005), farm diversification index (see Goodwin and Mishra, 2000), the borrower's credit rating or the borrower's credit bureau score (see Wu and Wang, 2000; Athavale and Edmister, 1999; Elsas and Krahnen, 1998), and the management ability of the borrower (see Angelini et al., 1998; Ellinger et al., 1992).

To identify the suitable cut-off point for the lending decision model, the Kolmogorov-Smirnov statistic (K-S test) can be used. The K-S test measures the distance between the distribution functions of the two classifications (good loan and bad loan). The score that generates the greatest separability between the functions is considered the suitable cut-off value for good or bad loan. However, the significant weakness of the K-S test is that it assumes the relative costs of misclassification errors are equal. As a result, it does not incorporate relevant information regarding the misclassification rates and their respective costs (Nargundkar and Priestley, 2003).

Since the PNN, GRNN, WN, and MLFN models are all supervised models, it might be useful to apply unsupervised neural network model such as the self-organizing map (SOM) with the credit scoring model. The SOM (Kohonen, 1982) is an artificial neural network method based on unsupervised algorithm technology that has been successfully applied to data mining (Shanmuganathan and Sallis, 2001). According to Ripley (1996), the SOM is mainly a clustering algorithm. It is argued that the SOM offers a powerful tool to visualise and analyse multi-dimensional data sets, and the distribution of variables in a data set can be make visible and easily understood. Despite the fact that this unsupervised model would suit clustering strategies, it has commonly been used in conjunction with supervised approaches in providing some explanation of the classification process (Vellido et al., 1999).

Collateral requirement is considered as another key issue in the lending decision process. However, it has not been addressed in this research. Therefore, further research can be extended to include the determinants of collateral (what factors determine the extent of total debt collateralization) and the influence of the bank-borrower relationship on the collateral requirement. These could be achieved by using the multiple regression analysis and the artificial neural networks techniques.

REFERENCES

Akhavein, J., L.G. Goldberg, and L. White, 2004, "Small banks, small business, and relationships: an empirical study of lending to small farms", ***Journal of Financial Services Research***, 26(3): 245-261.

Albanis, G.T., and R.A. Batchelor, 1999, Using probabilistic neural networks and rule induction techniques to predict long-term bond ratings, in M. Torres (Ed.), Proceeding of the 5th Annual Conference on Information Systems, Analysis and Synthesis, Orlando: IIIS.

Altman, E.I., M. Glancario, and F. Varetto, 1994, "Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks (the Italian experience)", ***Journal of Banking and Finance***, 18: 505–529.

Amaral, T.G.B., M.M. Crisostomo, and V.F. Pires, 2002, "Helicopter motion control using a general regression neural network", in M. Masoud (Ed.), ***Computational Intelligence in Control***, Idea Group Inc.

Anandarajan, M., P. Lee, and A. Anandarajan, 2001, "Bankruptcy prediction of financially stressed firms: An examination of the predictive accuracy of artificial neural networks", ***International Journal of Intelligent Systems in Accounting***, Finance & Management, 10: 69-81.

Angelini, P., R. Di Salvo, and G. Ferri, 1998, "Availability and cost of credit for small business: Customer relationships and credit cooperatives", ***Journal of Banking and Finance***, 22(6-8): 925-954.

Asian Development Bank (ADB), 2002, "Policy and program priorities for balancing the role of agriculture in rural and environmental development", ***Development of Agricultural Sector and Policy***, Thailand (Draft final report).

Athavale, M., and R.O. Edmister, 1999, "Borrowing relationships, monitoring, and the infulence on loan rates", ***Journal of Financial research***, 22(3): 341–352.

Bangkok Post, 2000, "Big break for farmers", 8th August 2000.

Bangkok Post, 2002, "119,000 farmers exit programme", 12th July 2002.

Bank for Agriculture and Agriculture Co-operative (BAAC), (Various years), ***BAAC Annual Reports***, Bangkok, Thailand.

Bank of Thailand (BOT), (Various years), ***Bank of Thailand Monthly Bulletin***, Retrieved August 5, 2003 from www.bot.or.th

Bard, S.K., P.J. Barry, and P.N. Ellinger, 2000, "Effects of commercial bank structure and other characteristics on agricultural lending", ***Agricultural Finance Review***, 60: 17-31.

Barney, D.K., O.F. Graves, and J.D. Johnson, 1999, "The farmers home administration and farm debt failure prediction", ***Journal of Accounting and Public Policy***, 18: 99–139.

Barry, P.J., and P.N. Ellinger, 1989, "Credit scoring, loan pricing, and farm business performance", ***Western Journal of Agricultural Economics***, 14(1): 45-55.

Barry, P.J., P.N. Ellinger, J.A. Hopkin, and C.B. Baker, 1995, ***Financial Management in Agriculture***, (5th ed.), Interstate Publishers.

Berger, A.N., and G.F. Udell, 1990, "Collateral, loan quality, and bank risk", ***Journal of Monetary Economics***, 25(1): 21-42.

Berger, A.N., and G.F. Udell, 1995, "Relationship lending and lines of credit in small firm finance", ***Journal of Business***, 68: 351–381.

Berger, A.N., and G.F. Udell, 2002, "Small business credit availability and relationship lending: the importance of bank organizational structure", ***Economic Journal***, 112: 32–53.

Bhipatboonthong, S., 2002, ***Micro-credit in Thailand***, Seminar on Micro-credit Experience in Taiwan, Department of Finance, National Chen-Chi University and International Cooperation and Development Fund, Taiwan.

Blackwell, D.W., and D.B. Winters, 1997, "Banking relationships and the effect of monitoring on loan prices, ***Journal of Financial Research***, 20: 275–289.

Bodenhorn, H., 2003, "Short-term loans and long-term relationships: Relationship lending in early America", ***Journal of Money, Credit, and Banking***, 35(4): 485-505.

Boot, A.W.A., 2000, "Relationship banking: what do we know?", ***Journal of Financial Intermediation***, 9(1): 7-25.

Cacoullos, T., 1966, "Estimation of a multivariate density", ***Annals of the Institute of Statistical Mathematics (Tokyo)***, 18(2): 179-189.

Capon, N., 1982, "Credit scoring systems: a critical analysis", ***Journal of Marketing***, 40: 82–91.

Chan, Y., and A.V. Thakor, 1987, "Collateral and competitive equilibria with moral hazard and private information", ***Journal of Finance***, 42(4): 345-363.

Chandler, G.C., and J.Y. Coffman, 1979, "A comparative analysis of empirical vs. judgmental credit evaluation", ***Journal of Retail Banking***, 1: 15 – 26.

Chen, A., M.T. Leung, and H. Daouk, 2003, "Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index", ***Computers & Operations Research***, 30(6): 901-923.

Chen, A.S., and M.T. Leung, 2004, "Regression neural network for error correction in foreign exchange forecasting and trading", ***Computers & Operations Research***, 31(7): 1049-1068.

Coakley, J.R., and C.E. Brown, 2000, "Artificial neural networks in accounting and finance : modeling issues", ***International Journal of Intelligent Systems in Accounting, Finance & Management***, 9: 119–144.

Cole, R.A., 1998, "The importance of relationships to the availability of credit", ***Journal of Banking and Finance***, 22(6-8): 959-977.

Collins, R.A., and R.D. Green, 1982, "Statistical methods for bankruptcy forecasting", ***Journal of Economics and Business***, 34: 349–354.

Crook, J.N., 1996, "Credit scoring: an overview", Working paper no.96/13, Department of Business studies, The University of Edinburgh.

Davis, J.R., A. Gaburici, and P.G. Hare, 1998, "What's wrong with Romanian rural finance? Understanding the determinants of private farmers' access to credit", Center for Economic Reform and Transformation (CERT), Discussion Paper, Heriot-Watt University, Edinburgh, UK.

Degryse, H., and P.V. Cayseele, 2000, "Relationship lending within a bank-based system: evidence from European small business data", ***Journal of Financial Intermediation***, 9(1): 90-109.

Degryse, H., and, S. Ongena, 2005, "Distance, lending relationships, and competition", ***Journal of Finance***, 60(1): 231-266.

DeTIENNE, K. B., D.H. DeTIENNE, and S.A. Joshi, 2003, "Neural networks as statistical tools for business researchers", ***Organizational Research Methods***, 6(2): 236-265.

Dunn, D.J., and T.L. Frey, 1976, "Discriminant analysis of loans for cash grain farms", ***Agricultural Finance Review***, 36: 60–66.

Ellinger, P.N., N.S. Splett, and P.J. Barry, 1992, "Consistency of credit evaluations at agricultural banks", ***Agribusiness***, 8(6): 517-536.

Elsas, R., and J.P. Krahnen, 1998, "Is relationship lending special? Evidence from credit-file data in Germany", ***Journal of banking and Finance***, 22(10-11): 1283-1316.

Etheridge, H.L., and R.S. Sriram, 1997, "A comparison of the relative costs of financial distress models: artificial neural networks, logit and multivariate discriminant analysis", ***International Journal of Intelligent Systems in Accounting, Finance & Management***, 6(3): 235-248.

Feder, G., L.J. Lau, J.Y. Lin, and X. Luo, 1993, "The nascent rural credit market in China", in Hoff, K., A. Braverman, and J.E. Stiglitz (Eds.), ***The economics of rural organization: theory, practice, and policy***, Oxford University Press.

Feldman, R., 1997, "Small business loans, small banks and a big change in technology called credit scoring", Federal Reserve Bank of Minneapolis, ***The Region***, September, 19-25.

Fischer, M.L., and K. Moore, 1986, "An improved credit scoring function for the St. Paul Bank for cooperatives", ***Journal of Agricultural Cooperative***, 1: 11–21.

Frame, W.S., A. Srinivasan, and L. Woosley, 2001, "The effect of credit scoring on small-Business lending", ***Journal of Money, Credit and Banking***, 33(3): 813–825.

Gan, C., V. Limsombunchai, M. Clemes, and A. Weng, 2005, "Consumer choice prediction: artificial neural networks versus logistic models", Working paper no. 104, Commerce Division, Lincoln University.

Glassman, C.A., and H.M. Wilkins, 1997, "Credit scoring: probabilities and pitfalls", ***Journal of Retail Banking Service***, 19(2): 53–56.

Goldberg, D.E., 1989, ***Genetic Algorithms in Search, Optimization, and Machine Learning***, Addison-Wesley.

Gonzalez, S., 2000, "Neural networks for macroeconomic forecasting: a complementary approach to linear regression models", Working paper no. 2000-07, Department of Finance Canada.

Goodwin, B.K., and A.K. Mishra, 2000, "An analysis of risk premia in U.S. farm-level interest rates", ***Agricultural Finance Review***, 60: 1-16.

Gradojevic, N., and J. Yang, 2000, "The application of artificial neural networks to exchange rate forecasting: the role of market microstructure variables", Working Paper 2000–23, Bank of Canada.

Greenbaum, S.I., G. Kanatas, and I. Venezia, 1989, "Equilibrium loan pricing under the bank-client relationship", ***Journal of Banking and Finance***, 13: 221–235.

Greene, W.H., 1997, ***Econometric Analysis***, (3rd ed.), Prentice Hall.

Gujarati, D.N., 1995, ***Basic Econometrics***, (3rd ed.), McGraw-Hill.

Harhoff, D., and T. Korting, 1998, "Lending relationships in Germany – empirical evidence from survey data", ***Journal of Banking and Finance***, 22 (10-11): 1317-1353.

Heckman, J.J., 1979, "Sample selection bias as a specification error", ***Econometrica***, 47: 153–162.

Holland, J.H., 1975, ***Adaptation in Natural and Artificial Systems***, MIT Press.

Hu, M.Y., M. Shanker, and M. S. Hung, 1999, "Estimation of posterior probabilities of consumer situational choices with neural network classifiers", ***International Journal of Research in Marketing***, 16(4): 307-317.

Jensen, H.L., 1992, "Using neural networks for credit scoring", ***Journal of Managerial Finance***, 18(6): 15–16.

Judge, G.G., W. E. Griffiths, R.C. Hill, H. Lutkepohl, and T. Lee, 1985, ***The Theory and Practice of Econometrics***, (2nd ed.), John Wiley and Sons.

Keasey, K., and R. Watson, 2000, "Interest rate premia on UK small firm bank borrowings: a research note", ***Journal of Business Finance and Accounting***, 27: 247–259.

Kitchens, F.L., J.D. Johnson, and J.N. Gupta, 2002, "Predicting automobile insurance losses using artificial neural networks", in Smith, Kate (Ed), ***Neural Networks in Business: Techniques and Applications***, Idea Group Publishing.

Koh, H.C., 1992, "The sensitivity of optimal cutoff points to misclassification costs of Type I and Type II errors in the going-concern prediction context", ***Journal of Business Finance and Accounting***, 19(2): 187-197.

Kohonen, T., 1982, "Self-organized formation of topologically correct feature maps", ***Biological Cybernetics***, 43(1): 59-69.

Kohzadi, N., M.S. Boyd, I. Kaastra, B.S. Kermanshahi, and D. Scuse, 1995, "Neural networks for forecasting: an introduction", ***Canadian Journal of Agricultural Economics***, 43: 463-474.

Krongkaew, M., 1995, "Contributions of agriculture to industrialization", in Medhi Krongkaew (Ed.), ***Thailand's Industrialization and its Consequences***, Macmillan.

Ladue, E., 1989, "The role of credit evaluation in agricultural finance: discussion", ***American Journal of Agricultural Economics***, 71: 1152–54.

Lee, T.H., and S.C. Jung, 1999, "Forecasting creditworthiness: logistic vs. artificial neural net", ***Journal of Business Forecasting***, Winter: 28–30.

Lee, W.F., M.D. Boehlje, A.G. Nelson, and W. G. Murry, 1988, ***Agricultural Finance***, (8th ed.), Iowa State University.

Lenk, M.M., E.M. Worzala, and A. Silva, 1997, "High-tech valuation : should artificial neural networks bypass the human valuer?", ***Journal of Property Valuation & Investment***, 15(1): 8-26.

Leung, M.T., A.S. Chen, and H. Daouk, 2000, "Forecasting exchange rates using general regression neural networks", ***Computers & Operations Research***, 27(11-12): 1093-1110.

Lewis, E.M., 1992, ***An Introduction to Credit Scoring***, (2nd ed.), Fair, Isaac and Co.

Limsombunchai, V., C. Gan, and M. Lee, 2004, "House price prediction: hedonic price model vs. artificial neural network", ***American Journal of Applied Sciences***, 1(3): 193-201.

Lufburrow, J., P.J. Barry, and B.L. Dixon, 1984, "Credit scoring for farm loan pricing", ***Agricultural Finance Review***, 44: 8-14.

Luther, R., 1998, "An artificial neural network approach to predicting the outcome of Chapter 11 bankruptcy", ***The Journal of Business and Economic Studies***, 4(1): 57-73.

Maddala, G.S., 1983, ***Limited-dependent and qualitative variables in econometrics***, Cambridge University Press.

Maddala, G.S., 2001, ***Introduction to econometrics***, (3rd ed.), John Wiley.

Martin, N.M., and L.C. Jain, 1999, "Introduction to neural networks, fuzzy systems, genetic algorithms and their fusion", in Jain, L. C. and N. M., Martin (Eds.), ***Fusion of Neural Networks, Fuzzy Sets and Genetic Algorithms: Industrial Application***, CRC Press LLC.

May, G.R., 1996, "Applications of neural networks in semiconductor manufacturing process", in C. H. Chen (Ed.), ***Fuzzy Logic and Neural Network Handbook***, McGraw-Hill.

McAllister, P.H., and J.J. Mingo, 1994, “Commercial loan risk management, credit scoring and pricing: the need for a new shared database”, ***Journal of Commercial Lending***, May 1994: 6–22.

McCluskey, W., 1996, “Prediction accuracy of machine learning models for the mass appraisal of residential property”, ***New Zealand Valuers' Journal***, July: 41-47.

Mester, L.J., 1997, “What's the point of credit scoring?”, ***Business Review***, Federal Reserve Bank of Philadelphia, Sep./Oct. 1997: 3–16.

Miller, L.H., and E.L. LaDue, 1989, “Credit assessment models for farm borrowers: a logit analysis”, ***Agricultural Finance Review***, 49: 22–36.

Miller, L.H., P.N. Ellinger, P.J. Barry, and K. Lajili, 1993, “Price and nonprice management of agricultural credit risk”, ***Agricultural Finance Review***, 53: 28-41.

Mortensen, T., D.L. Watt, and F.L. Leistritz, 1988, “Predicting probability of loan default”, ***Agricultural Finance Review***, 48: 60–67.

Nadaraya, E.A., 1964, “On estimating regression”, ***Theory Probab. Application***, 10: 186-190.

Nargundkar, S., and J.L. Priestley, 2003, “Assessment of evaluation methods for prediction and classifications of consumer risk in the credit industry”, in Zhang, G.P. (Ed), ***Neural Networks in Business Forecasting***, Idea Group Inc.

Natetayaluck, P., 2001a, “The analysis of debt moratorium project”, ***Journal of Economic Monitoring***, 9(1): 27-33.

Natetayaluck, P., 2001b, “Rural credit policy and accessibility of the credit in Thailand”, ***Journal of Economic Monitoring***, 9(6): 14–23.

National Statistic Office (NSO.), (Various year), ***Socio-economic Survey Report***, Retrieved August 5, 2003 from www.nso.go.th

Novak, M.P., and E. LaDue, 1999, “Application of recursive partitioning to agricultural credit scoring”, ***Journal of Agricultural and Applied Economics***, 31(1): 109–122.

Office of Agricultural Economics (OAE.), (Various year), ***Agricultural Statistics of Thailand***, Ministry of Agriculture and Cooperatives, Retrieved August 5, 2003 from www.oae.go.th

Ongena, S., and D.C. Smith, 2000, "Bank relationships: a reviews", In Harker, P. T. and S. A., Zenics (Ed.), ***Performance of Financial Institutions: Efficiency, Innovation, Regulation***, Cambridge University Press.

Parzen, E., 1962, "On estimation of a probability density function and mode", ***Annals of Mathematical Statistics***, 33: 1065-1076.

Petersen, M.A., and R.G. Rajan, 1994, "The benefit of lending relationships: evidence from small business data", ***Journal of Finance***, 49: 3–38.

Pham, D.T., and X. Liu, 1995, ***Neural Networks for Identification, Prediction, and Control***, Springer.

Plata, V., and G.N. Nartea, 1998, "Credit analysis procedures of rural lenders in Canterbury", Paper presented at the Fifth Annual Conference of the New Zealand Agricultural and Resource Economics Society (Inc.).

Poramacom, N., 2000, "Agricultural credit provided by Bank for Agriculture and Agricultural Cooperatives (BAAC) and assets and indebtedness of farm households", ***Kasetsart Journal (Social Science)***, 21: 191-198.

Poramacom, N., 2001, ***Agricultural Credit in Asia and the Pacific***, Country report, Report of the APO Symposium, 8 – 14 September 1999, Asian Productivity Organization (APO), Tokyo, Japan.

Pyndick, R.S., and D.L. Rubinfeld, 1998, ***Econometric Models and Economic Forecasts***, (4th ed.), McGraw Hill.

Rao, C.P., and J. Ali, 2002, "Neural network model for database marketing in the new global economy", ***Marketing Intelligence & Planning***, 20(1): 35-43.

Ripley, B., 1996, ***Pattern recognition and neural networks***, Cambridge University Press.

Rose, P.S., 1993, ***Commercial Bank Management***, (2nd ed.), IRWIN.

Sakchuwong, S., 2004, "A performance analysis of the Bank for Agriculture and Agricultural Cooperatives in providing agricultural credit", Master Thesis, Faculty of Economics, Kasetsart University, Thailand.

Shachmurove, Y., and D. Witkowska, 2000, "Utilizing artificial neural network model to predict stock markets", CARESS working paper no. 00-11, University of Pennsylvania.

Shanmuganathan, S., and P. Sallis, 2001, "Self organising maps in biological process modelling", in proceedings of The Fifth Biannual Conference on Artificial Neural Networks and Expert Systems, 126-131.

Sharpe, S.A., 1990, "Asymmetric information, bank lending and implicit contracts: a stylized model of customer relationships", ***Journal of Finance***, 45: 1069–1087.

Sinkey, J.F., 2002, ***Commercial Bank Financial Management***, (6th ed.), Prentice Hall.

Smith, M., 1996, ***Neural Networks for Statistical Modeling***, International Thomson Computer Press.

Specht, D., 1990, "Probabilistic neural networks", ***Neural Networks***, 3(1): 109-118.

Specht, D., 1991, "A general regression neural network", ***IEEE Transactions on Neural Networks***, 2(6): 568-576.

Specht, D.F., 1996, "Probability neural networks and general regression neural networks", in C. H. Chen (Ed.), ***Fuzzy Logic and Neural Network Handbook***, McGraw-Hill.

Stiglitz, J.E., and A. Weiss, 1981, "Credit rationing in markets with imperfect competition", ***American Economic Review***, 71: 393–410.

Strahan, P.E., 1999, "Borrower risk and the price and non-price terms of bank loans", Staff Reports, no. 90 (October 1999), Federal Reserve Bank of New York.

Tarasenko, L., 1998, ***A Guide to Neural Computing Applications***, Arnold.

Tay, D.P.H., and D.K.H. Ho, 1991, "Artificial intelligence and the mass Appraisal of residential apartments", ***Journal of Property Valuation & Investment***, 10(2): 525-539.

Thailand Development Research Institute (TDRI), 1998, ***Project of Financial Market in Thai Rural Sector in 1996***, TDRI, Thailand.

Turvey, C.G., 1991, "Credit scoring for agricultural loans: a review with application", ***Agricultural Finance Review***, 51: 43–54.

Turvey, C.G., and A. Weersink, 1997, "Credit risk and the demand for agricultural loans", ***Canadian Journal of Agricultural Economics***, 4: 201–217.

Turvey, C.G., and R. Brown, 1990, "Credit scoring for a federal lending institution: the case of Canada's farm credit corporation", ***Agricultural Finance Review***, 50: 47–57.

Vellido, A., P.J.G. Lisboa, and K. Meehan, 1999, "Segmentation of the on-line shopping market using neural networks", ***Expert Systems with Applications***, 17: 303-314.

Ward System Group Inc., 1993, NeuroShell 2, release 4.0.

Wasserman, P.D., 1993, ***Advanced methods in neural computing***, Van Nostrand Reinhold, New York.

West, P.M., P.L. Brockett, and L.L. Golden, 1997, "A comparative analysis of neural networks and statistical methods for predicting consumer choice", ***Marketing Science***, 16(4): 370-391.

Wu, C., and X.M. Wang (2000), "A neural network approach for analyzing small business lending decisions", ***Review of Quantitative Finance and Accounting***, 15(3): 259-276.

Yang, Z.R., M.B. Platt, and H.D. Platt, 1999, "Probabilistic neural networks in bankruptcy prediction", ***Journal of Business Research***, 44(2): 67-74.

Yim, J., 2002, "A comparison of neural networks with time series models for forecasting returns on a stock market", Working paper no. 07/2002, Business School, RMIT University.

APPENDICES

Appendix 1: Marginal effects of bank lending decision models

Independent variables[1/]	Marginal Effects[2/]					
	Aggregate		Agriculture		Non-agriculture	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Borrower Characteristics						
Log(asset)	***0.0285***	***0.0406***	***0.0289***	***0.0387***	***0.0313***	0.0302
Age	-0.0002	-0.0002	-0.0001	-0.0002	-0.0004	0.0003
Education	***0.0136***	0.0147	***0.0161***	0.0190	-0.0081	-0.0161
Credit risk proxies						
Log(collateral)	-0.0026	-0.0084	-0.0030	-0.0072	-0.0086	0.0011
Return on asset	***0.0045***	0.0059	0.0030	0.0005	***0.0226***	***0.0511***
Leverage ratio	***-0.0851***	-0.0389	***-0.0874***	-0.0868	-0.0193	0.0257
Capital turnover ratio	***-0.0073***	***-0.0091***	***-0.0060***	***-0.0062***	***-0.0208***	***-0.0351***
Relationship indicators						
Borrowing from others	0.0086	-0.0091	0.0095	0.0034	-0.0138	-0.0232
Duration		***-0.0177***		***-0.0199***		0.0099
Dummy variables[3/]						
Sector	***0.0512***	***0.1133***				
(Province)						
Province 1					***0.1681***	***0.1891***
Province 2	***0.0308***	-0.0646	***0.0278***	-0.0196	***0.2143***	***0.1297***
Province 3	***-0.1731***	***-0.1223***	***-0.1767***	***-0.1184***	***0.1341***	0.2262
Province 4	***-0.3229***	***-0.6218***	***-0.3143***	***-0.5412***	0.0277	-0.8250
Province 5	***-0.1747***	***-0.2160***	***-0.1771***	***-0.1924***	***0.1347***	0.0789
Province 6	***-0.2496***	***-0.4353***	***-0.2481***	***-0.3951***	***0.1091***	0.0457
Province 7	***-0.1317***	***-0.1086***	***-0.1316***	***-0.0973***	***0.1429***	***0.1835***
Province 8	***-0.1573***	***-0.1086***	***-0.1699***	***-0.1670***	***0.1514***	***0.1340***
Province 9	***-0.1524***	***-0.2448***	***-0.1511***	***-0.2506***	***0.1393***	***0.1642***
Province 10	***-0.1580***	***-0.3255***	***-0.1486***	***-0.2637***	***0.1440***	***0.0369***
Province 11	***-0.5767***	***-0.5615***	***-0.5394***	***-0.4937***		
Province 12	***-0.3210***	-0.1134	***-0.3263***	-0.0868	***0.1409***	0.1423
Province 13	***-0.1698***	***-0.2470***	***-0.1535***	***-0.1799***	0.0712	-0.0322
Province 14	***-0.1221***	***-0.4023***	***-0.1001***	***-0.4946***	0.1141	0.2254
Province 15	***-0.4242***	***-0.4611***	***-0.4062***	***-0.4107***	0.0777	0.1055
Province 16	***-0.2857***	***-0.2696***	***-0.2302***	***-0.2214***		
Province 17	***-0.2365***	***-0.3195***	***-0.2280***	***-0.2865***	***0.1323***	0.1150

Appendix 1: Marginal effects of bank lending decision models (Cont)

Independent variables[1/]	Marginal Effects[2/]					
	Aggregate		Agriculture		Non-agriculture	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
(Major production)						
Horticulture	***0.0761***	***0.0805***	***0.0748***	***0.0775***		
Orchard/Vegetable	***0.0719***	***0.0619***	***0.0674***	***0.0469***	***-0.0857***	0.0168
Livestock/Aquaculture	***0.0601***	***0.0417***	***0.0562***	0.0306	0.0367	0.0592
(Loan type)						
Short-term loan	-0.0201	-0.0515	***-0.0264***	-0.0550		
Medium-term loan	-0.0196	***-0.1175***	-0.0274	***-0.1130***		
Long-term loan	***-0.0582***	***-0.1006***	***-0.0584***	***-0.0931***	***0.1422***	0.0532
(Loan size)						
Medium loan	***-0.0364***	***-0.0327***	***-0.0381***	***-0.0376***	-0.0219	-0.0035
Large loan	***-0.0546***	***-0.1071***	***-0.0722***	***-0.3092***	0.0158	0.0239
(Lending year)						
2001	***0.0279***	0.0013	0.0132	-0.0062		
2002	***-0.0413***	***-0.0443***	***-0.0352***	***-0.0377***	***-0.1054***	***-0.0998***
Constant	-0.0584	0.0216	-0.0347	0.0699	-0.2565	-0.3814
No. of observations	18,310	4,444	16,560	3,965	1,750	479

Note: 1/ Dependent variable is Bank lending decision (good/bad loan).
2/ Marginal effect is at the mean value. For dummy variable, marginal effect is P|1 - P|0.
3/ To avoid singularity problem, a dummy variable is dropped from each group.
Bold and Italic represent 10% significant level or below.

Appendix 2: Relative contribution factor of PNN on bank lending decision models

Independent variables[1/]	Relative Contribution					
	Aggregate		Agriculture		Non-agriculture	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Borrower Characteristics						
Log(asset)	***0.0639***	***0.0567***	***0.0567***	***0.0571***	***0.0703***	0.0293
Age	***0.0526***	***0.0377***	***0.0364***	***0.0544***	0.0126	0.0053
Education	***0.0360***	0.0110	0.0034	***0.0418***	0.0249	0.0139
Credit risk proxies						
Log(collateral)	***0.0506***	***0.0570***	***0.0476***	***0.0376***	***0.0880***	***0.0549***
Return on asset	***0.0548***	0.0108	0.0115	0.0036	0.0058	0.0253
Leverage ratio	0.0158	0.0087	***0.0567***	***0.0409***	0.0072	***0.0562***
Capital turnover ratio	***0.0465***	***0.0583***	0.0115	0.0155	0.0188	0.0301
Relationship indicators						
Borrowing from others	0.0174	0.0268	0.0009	***0.0342***	***0.0725***	0.0029
Duration		***0.0325***		***0.0355***		0.0091
Dummy variables						
Sector	0.0013	0.0018				
(Province)						
Province 1					***0.0905***	0.0205
Province 2	0.0189	***0.0359***	0.0016	***0.0324***	0.0296	0.0029
Province 3	0.0018	***0.0462***	0.0187	***0.0301***	0.0087	0.0314
Province 4	0.0141	0.0066	0.0240	0.0074	0.0267	0.0309
Province 5	***0.0551***	***0.0309***	0.0000	***0.0301***	0.0025	***0.0504***
Province 6	***0.0619***	0.0007	***0.0405***	0.0137	***0.0725***	***0.0642***
Province 7	0.0171	0.0130	***0.0782***	0.0119	0.0148	0.0091
Province 8	0.0224	0.0085	***0.0448***	0.0036	***0.0768***	***0.0610***
Province 9	0.0091	***0.0487***	***0.0601***	***0.0533***	***0.0660***	***0.0669***
Province 10	***0.0362***	0.0007	0.0065	0.0144	0.0126	***0.0384***
Province 11	***0.0619***	0.0185	***0.0604***	***0.0290***		
Province 12	***0.0571***	***0.0565***	0.0224	0.0090	***0.0718***	***0.0634***
Province 13	0.0221	***0.0339***	0.0227	***0.0373***	0.0105	0.0152
Province 14	0.0030	0.0055	***0.0442***	0.0236	0.0256	0.0115
Province 15	0.0224	0.0270	***0.0380***	0.0263	0.0036	***0.0679***
Province 16	0.0141	0.0226	0.0215	***0.0524***		
Province 17	***0.0448***	0.0114	***0.0576***	***0.0565***	0.0198	***0.0472***

Appendix 2: Relative contribution factor of PNN on bank lending decision models (Cont)

Independent variables[1/]	Relative Contribution					
	Aggregate		Agriculture		Non-agriculture	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
(Major production)						
Horticulture	0.0121	0.0124	0.0078	0.0088		
Orchard/Vegetable	***0.0302***	***0.0295***	0.0016	***0.0427***	***0.0451***	***0.0402***
Livestock/Aquaculture	0.0106	***0.0329***	0.0034	0.0155	0.0310	***0.0663***
(Loan type)						
Short-term loan	***0.0322***	0.0078	0.0072	***0.0418***		
Medium-term loan	0.0262	***0.0572***	***0.0405***	0.0263		
Long-term loan	0.0133	***0.0554***	0.0131	***0.0531***	0.0133	0.0035
(Loan size)						
Medium loan	0.0058	0.0149	0.0112	0.0135	0.0014	0.0096
Large loan	0.0028	***0.0288***	***0.0673***	0.0148	0.0198	0.0072
(Lending year)						
2001	***0.0569***	***0.0503***	***0.0744***	0.0252		
2002	0.0093	***0.0428***	0.0075	0.0063	***0.0573***	***0.0653***
No. of inputs	35	36	34	35	28	29
No. of observations	18,310	4,444	16,560	3,965	1,750	479
Mean value of the contribution factor[3/]	0.0286	0.0278	0.0294	0.0286	0.0345	0.0333

Note: 1/ Dependent variable is bank lending decision (good/bad loan).
2/ Equal to 1 over the number of inputs used in the model.
Bold and italic indicate that the relative contribution is above the mean value of the contribution factors.

Appendix 3: Relative contribution factor of GRNN on credit availability models

Independent variables[1/]	Relative Contribution					
	Aggregate		Agriculture		Non-agriculture	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Borrower Characteristics						
Log(asset)	***0.0482***	***0.0421***	***0.0514***	***0.0447***	***0.0469***	***0.0412***
Age	0.0086	0.0117	0.0081	0.0035	***0.0356***	0.0005
Education	0.0076	0.0081	0.0021	0.0232	***0.0477***	0.0000
Credit risk proxies						
Log(collateral)	***0.0639***	***0.0546***	***0.0541***	***0.0555***	***0.0671***	***0.0593***
Return on asset	0.0134	***0.0458***	0.0131	***0.0356***	0.0040	0.0070
Leverage ratio	0.0129	0.0137	***0.0500***	0.0103	***0.0369***	0.0090
Capital turnover ratio	***0.0330***	0.0214	0.0023	0.0152	0.0108	0.0007
Relationship indicators						
Borrowing from others	0.0231	0.0222	0.0032	0.0110	0.0137	0.0169
Duration		0.0123		***0.0333***		0.0245
Dummy variables						
Sector	***0.0348***	0.0161				
(Province)						
Province 1					***0.0650***	***0.0620***
Province 2	***0.0543***	0.0108	***0.0449***	***0.0522***	***0.0555***	***0.0571***
Province 3	0.0008	***0.0390***	0.0092	0.0253	0.0078	***0.0403***
Province 4	***0.0386***	0.0277	0.0247	***0.0396***	0.0278	***0.0583***
Province 5	0.0269	0.0075	***0.0343***	***0.0337***	0.0202	0.0260
Province 6	0.0043	0.0238	***0.0558***	0.0129	***0.0485***	0.0202
Province 7	***0.0502***	***0.0469***	0.0065	0.0145	***0.0566***	***0.0522***
Province 8	***0.0480***	0.0152	0.0274	0.0075	0.0286	***0.0513***
Province 9	0.0167	0.0271	***0.0412***	***0.0349***	***0.0642***	***0.0586***
Province 10	***0.0467***	0.0269	***0.0551***	0.0150	0.0226	0.0109
Province 11	0.0218	0.0053	0.0283	0.0283		
Province 12	***0.0411***	0.0222	0.0281	0.0201	0.0213	0.0270
Province 13	0.0165	0.0273	***0.0454***	0.0237	***0.0550***	0.0287
Province 14	0.0216	***0.0333***	0.0154	***0.0518***	0.0275	0.0226
Province 15	0.0109	***0.0333***	***0.0551***	0.0187	***0.0677***	***0.0457***
Province 16	***0.0388***	0.0258	***0.0378***	***0.0347***		
Province 17	0.0190	***0.0559***	***0.0408***	***0.0344***	***0.0369***	***0.0610***

Appendix 3: Relative contribution factor of GRNN on credit availability models (Cont.)

Independent variables[1/]	Relative Contribution					
	Aggregate		Agriculture		Non-agriculture	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
(Major production)						
Horticulture	0.0025	0.0046	0.0048	0.0091		
Orchard/Vegetable	***0.0490***	***0.0295***	***0.0551***	0.0222	0.0121	***0.0406***
Livestock/Aquaculture	***0.0617***	***0.0562***	***0.0304***	***0.0382***	0.0105	0.0041
(Loan type)						
Short-term loan	0.0094	***0.0407***	0.0046	***0.0309***		
Medium-term loan	0.0129	0.0255	***0.0456***	***0.0393***		
Long-term loan	***0.0325***	***0.0352***	0.0288	***0.0506***	0.0089	***0.0411***
(Loan size)						
Medium loan	***0.0434***	***0.0544***	***0.0468***	***0.0506***	***0.0380***	***0.0457***
Large loan	***0.0546***	***0.0504***	***0.0371***	***0.0445***	***0.0453***	***0.0576***
(Lending year)						
2001	***0.0304***	0.0073	0.0099	0.0157		
2002	0.0018	0.0203	0.0025	0.0192	0.0173	0.0301
No. of inputs	35	36	34	35	29	30
No. of observations	18,310	4,444	16,560	3,965	1,750	479
Mean value of the contribution factor[2/]	0.0286	0.0278	0.0294	0.0286	0.0345	0.0333

Note: 1/ Dependent variable is log(volume of credit).
2/ Equal to 1 over the number of inputs used in the model.
Bold and italic indicate that the relative contribution is above the mean value of the contribution factors.

Appendix 4: Relative contribution factor of GRNN on loan pricing models

Independent variables[1/, 2/]	Relative Contribution					
	Aggregated		Agriculture		Non-agriculture	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Borrower Characteristics						
Log(asset)	***0.0466***	0.0100	0.0136	0.0101	***0.0495***	***0.0459***
Age	0.0278	0.0135	0.0187	0.0060	0.0108	***0.0364***
Education	0.0027	0.0046	0.0088	***0.0345***	0.0210	0.0045
Credit risk proxies						
Log(collateral)	***0.0369***	0.0254	0.0224	***0.0467***	***0.0361***	***0.0379***
Return on asset	0.0082	0.0086	***0.0445***	0.0009	0.0077	0.0058
Leverage ratio	0.0103	0.0073	0.0171	0.0170	0.0011	0.0184
Capital turnover ratio	0.0203	0.0168	0.0296	***0.0400***	0.0011	0.0324
Log(volume of credit)	***0.0483***	***0.0549***	***0.0667***	***0.0582***	***0.0720***	***0.0571***
Relationship indicators						
Borrowing from others	0.0228	0.0071	0.0139	0.0203	0.0003	0.0232
Duration		***0.0409***		***0.0555***		0.0297
Dummy variables						
Sector	***0.0411***	***0.0303***				
(Province)						
Province 1					***0.0501***	0.0236
Province 2	0.0116	0.0239	0.0248	***0.0384***	***0.0384***	***0.0423***
Province 3	***0.0333***	0.0232	0.0256	0.0032	0.0145	***0.0409***
Province 4	0.0272	0.0206	0.0275	***0.0559***	***0.0714***	0.0079
Province 5	***0.0418***	***0.0493***	***0.0355***	***0.0493***	0.0185	0.0177
Province 6	0.0262	0.0206	***0.0507***	***0.0522***	***0.0725***	***0.0553***
Province 7	0.0048	***0.0487***	***0.0459***	0.0226	0.0319	***0.0529***
Province 8	***0.0430***	***0.0323***	0.0003	***0.0430***	***0.0575***	0.0207
Province 9	0.0207	0.0186	0.0056	***0.0516***	***0.0654***	***0.0511***
Province 10	0.0127	0.0181	0.0275	***0.0557***	***0.0594***	***0.0556***
Province 11	***0.0367***	***0.0321***	0.0301	0.0005		
Province 12	***0.0331***	***0.0558***	0.0211	***0.0513***	***0.0609***	0.0162
Province 13	0.0036	0.0270	0.0171	0.0159	0.0336	0.0238
Province 14	0.0228	0.0237	0.0272	0.0193	***0.0600***	***0.0522***
Province 15	***0.0405***	0.0069	***0.0397***	***0.0532***	***0.0720***	***0.0481***
Province 16	***0.0405***	0.0199	***0.0589***	***0.0304***		
Province 17	***0.0318***	***0.0564***	***0.0365***	***0.0334***	***0.0629***	***0.0522***

Appendix 4: Relative contribution factor of GRNN on loan pricing models (Cont)

Independent variables[1/, 2/]	Relative Contribution					
	Aggregated		Agriculture		Non-agriculture	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
(Major production)						
Horticulture	***0.0392***	***0.0520***	***0.0339***	0.0099		
Orchard/Vegetable	***0.0346***	***0.0538***	***0.0589***	0.0283	0.0145	***0.0562***
Livestock/Aquaculture	***0.0304***	0.0285	***0.0621***	***0.0090***	0.0026	***0.0540***
(Loan type)						
Short-term loan	***0.0458***	0.0000	0.0099	***0.0391***		
Medium-term loan	***0.0308***	***0.0527***	0.0016	0.0028		
Long-term loan	***0.0466***	***0.0555***	***0.0515***	0.0265	0.0026	***0.0353***
(Loan size)						
Medium loan	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Large loan	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
(Lending year)						
2001	***0.0378***	0.0173	0.0219	0.0012		
2002	***0.0394***	***0.0436***	***0.0512***	0.0182	0.0119	0.0025
No. of inputs	34	35	33	34	28	29
No. of observations	18,310	4,444	16,560	3,965	1,750	479
Mean value of the contribution factor[3/]	0.0294	0.0286	0.0303	0.0294	0.0357	0.0345

Note: 1/ Dependent variable is loan price (interest rate).
2/ Assume that volume of credit is determined before the loan price. Therefore, dummy variables on the loan size are excluded and replaced by log(volume of credit).
3/ Equal to 1 over the number of inputs used in the model.
Bold and italic indicate that the relative contribution is above the mean value of the contribution factors.