

Application of Logistic Regression Analysis to Household Debt of Bangkok and Metropolitan Area of Thailand

SUKANYA INTARAPAK

Department of Mathematics, Faculty of Science
Srinakharinwirot University
114 Sukhumvit 23, Watthana, Bangkok 10110
THAILAND

THIDAPORN SUPAPAKORN*

Department of Statistics, Faculty of Science
Kasetsart University
50 Ngam Wongwan Rd, Lat Yao, Chatuchak, Bangkok 10900
THAILAND

Abstract: - The objective of this research is to study factors influencing the household debt of people in Bangkok and metropolitan area of Thailand. Using data from the 2017 household socio-economic survey from National Statistics Office (NSO) of Thailand, the results of logistic regression analysis indicate that four variables affecting indebtedness of the household are household sizes, number of people who get wages, remittance receiving, and loan for emergency. Overall, the logistic regression analysis correctly classifies 67.7% of indebtedness of household.

Key-Words: - Household, Indebtedness, Loan, Logistic Regression, Remittance, Wage

Received: January 12, 2020. Revised: June 22, 2020. Accepted: July 2, 2020. Published: July 7, 2020.

1 Introduction

In the past, Thailand had the household debt crisis that affected to the household, particularly, household consumptions. Purchasing powers of household consumers were reduced. Consequently, members of the household might make loans in order to increase their purchasing powers, but it had a drawback, they had burden responsibilities. In the last two years, there were three reasons of increasing in debt as follows: 1) the low saving interest rates, 2) the government policies regarding vehicles or houses that stimulate people in debt and 3) the financial products that are credit cards and personal loans that people can access easily and conveniently.

Household debt picked up again to 78.6% of GDP at the end of 2018 from 78.3% in the previous year. Thai's debt load peaked in 2015 at 81.2%. Thai household debt had an apparent faster growth rate than the country's economic growth in 2018, with the household debt ratio at 78.6 %, which was among the top three countries with the highest household debt in Asia. This directly affected overall economic expansion. The latest data of the Bank of Thailand showed that the household debt increased 80% of the GDP [2]. Therefore, the appropriate levels of the household debt could

support consumptions and economic expansion. If, in fact, the household debt is too high, it definitely affects to household consumptions and expenditures. During deflation, incomes of household may be affected, household risks are increasing and their solvencies are decreasing. These become uncollectible accounts which increase in the risks of economic and financial stability of the country.

Considering the household debt for each region of Thailand in 2017, it was found that Bangkok and Nonthaburi, Pathum Thani and Samut Prakan provinces have the highest average debt per household than other regions [8]. In this paper, the researchers are interested in studying the indebtedness of household in Bangkok and metropolitan area including Nonthaburi, Pathum Thani and Samut Prakan provinces. The objective of this research is mainly to find important factors effecting the household debt of people in Bangkok and metropolitan area of Thailand by using logistic regression analysis.

2 Literature Review

In 2007, Yoon, Jean, and Deanna [10] identified to what extent individuals aged 65 and over hold consumer or mortgage debt and to investigate

factors that influenced the probability of holding consumer or mortgage debt in old age. The results of logistic regression analysis indicated that older individuals who had larger households, had higher levels of education, were aged 65–74, were married, were Black, and were employed, were more likely to hold consumer or mortgage debt. The study concluded that holding one kind of debt was associated with probability of holding another kind of debt. Next, Costa and Farinha (2012) [4] used the Household Finance and Consumption Survey (HFCS) which took place during the second quarter of 2010 to analyze the Portuguese households' indebtedness based on microeconomic information that was particularly useful at the present time, given the high level of debt of this sector and the increase in credit default. The results obtained that low income and young households who had taken mortgages were the most vulnerable groups of the population, for which the probability of materialisation of credit risk was higher. However, the fact that low income households had relatively low participation in the debt market mitigated the impact of their eventual entry into default on the financial situation of banks. For young households, although their market share and the value of their loans were high, their debts were often guaranteed by real estate and the value of the debt service to income ratio for the majority of these households was lower than the usual threshold, used to identify situations of greater vulnerability.

Aniola-Mikołajczak (2016) [1] applied the method of logistic regression to study the factors influencing the use of credit by Polish households. The source material was individual data of the Household Budget Survey in 2011 conducted by the Central Statistical Office of Poland. The research results showed that the use of credits by Polish households was most influenced by the householder's sex, age, socio-occupational status, family type, and income in the household. Similarly, Chounlakorn and Kittichotipanit (2016) [3] studied factors influencing the family debts of civil servants of three types, namely, general officials, academic officials, and administrative and management officials, by using secondary data from National Statistics Office, under the title of the 'Survey on Living Conditions of Civil Servants in 2012'. The data were analyzed with multiple linear regression analysis technique. The results revealed that there were three common factors that affected the general officials to incur family debts: the capability of making installed payments, the number of family members, and the cost of living. The three main independent variables were wages and supportive

money for personnel serving the families, educational expenses, and educational backgrounds using as the bases for job positioning and current positions. The variation of the dependent variables could be explained to be 57.7. The common factor that influenced the academic officials to incur family debts was the capability of making installed payments. The two main independent variables were wages and supportive money for personnel serving the families, and educational expenses. The variation of the dependent variables was able to explain to be 50.1. The common factor that involved the administrative and management officials to incur family debts was the capability of making installed payments. The two main independent variables were expenses related to savings, and genders. The variation of the dependent variables could be described to be 63.5.

Recently, Rayasawath (2018) [9] determined the factors affecting the succession of youths from farming households in agricultural occupations from 400 farm households within Nakhon Ratchasima Province, Thailand. By binary logistic regression analysis, the results showed that five variables significantly influenced the decision of the new generation on succession in agricultural occupation: their experience in agricultural work, attitude towards agriculture as an occupation, the number of agricultural labor within a household, marital status, and if they faced problems with the agricultural resources in the past. Moreover, Haq, Ismail, and Satar (2018) [5] investigated the relationships between different socioeconomic and demographic characteristics and households' debt decision and demand. The data of Pakistan household integrated expenditure survey (HIES) 2001 to 2014 and multilevel models were used. The decision of taking household debt varied 22% at primary sampling unit level and 18% at provincial level due to unobserved variables. The results found that households having higher financial assets, higher income and larger household sizes tended to have a higher percentage of debt. The amount of debt also increased with education and age. In the case of demand for debt, the variation was 12% at the provincial level.

3 Methodology

The aim of this study is to identify factors influencing the indebtedness of household. The study is based on data of the 2017 household socio-economic survey from National Statistics Office (NSO) of Thailand [8]. From the sample of 2368 households in Bangkok and metropolitan area, there is 35.09% of household with debt (see Table 1).

Binary logistic regression is a method used with the dichotomous dependent variable. Independent variables can be both qualitative and quantitative [6]. Based on logistic regression analysis, the methods are as follows:

(1) Check the assumptions of logistic regression

- The independent variables are categorical variables, or a mix of continuous and categorical variables, logistic regression is preferred.

- Logistic regression requires each observation to be independent. So, the model should have no multicollinearity.

- Logistic regression requires large sample sizes to provide sufficient numbers in both categories of the outcome variable that is sample size $30p$ where p is the number of independent variables.

(2) Select the variable to be included in the analysis by using the backward stepwise method. All of the variables that are specified as independent variables are initially entered into the model together and then tested for removal one by one. The removal of the independent variables from the model is based on the significance of the Wald statistic.

(3) Create the logistic regression equation which can be written in term of the log of the odds:

$$\log\left(\frac{P_y}{1-P_y}\right) = \beta_0 + \beta_1x_1 + \dots + \beta_px_p \quad (1)$$

where y is the dependent variable, P_y is the probability of interested outcome, x_i is the i^{th} independent variable, β_0 is the intercept parameter, β_i is the i^{th} regression coefficient, $i = 1, 2, \dots, p$.

(4) Assess whether the model fits the data by considering the values of Cox and Snell's R^2 and Nagelkerke's R^2 .

4 Results

In this study, the sample size of household in Bangkok and vicinity is 2368. The dependent variable is the household indebtedness (1 if having debt; 0 otherwise). There are 8 independent variables which are both continuous and categorical variables. There are three categorical independent variables; remittance receiving (X_3), loan for business (X_7) where X_{71} is in case of loan is rejected, X_{72} is getting partial loan, X_{73} is getting complete loan and X_{74} is in case of no need of business loan and loan for emergency (X_8) where X_{81} is in case of loan is rejected, X_{82} is getting

partial loan, X_{83} is getting complete loan and X_{84} is in case of no need of emergency loan. There are five continuous independent variables; household size (X_1), number of people who get wages (X_2), income per month per household (Baht) (X_4), consumption expense (Baht) (X_5) and cost of goods and services (Baht) (X_6).

Table 1. Characteristics of Categorical Variables in Logistic Regression Model

Categorical variable	Response category	Count	%
Household Indebtedness (Y)	without debt	1537	64.91
	with debt	831	35.09
Remittance Receiving (X_3)	Not receive	2008	84.80
	Receive	360	15.20
Loan for Business (X_7)	Loan is rejected	283	11.95
	Partial loan	373	15.75
	Complete loan	409	17.27
	No need	1303	55.03
Loan for Emergency (X_8)	Loan is rejected	248	10.47
	Partial loan	424	17.91
	Complete loan	498	21.03
	No need	1198	50.59

Table 2. Characteristics of Continuous Variables in Logistic Regression Model

Continuous variable	Mean	Standard deviation
Household size (X_1)	2.80	1.50
Number of People who get wages (X_2)	1.11	0.93
Income per month per household (Baht) (X_4)	47,363.28	68,462.71
Consumption Expense (Baht) (X_5)	11,396.56	7,817.54
Cost of Goods and Services (Baht) (X_6)	32,121.91	28,640.21

The descriptive statistics of all variables are separately shown in Table 1 and Table 2. In Bangkok and metropolitan area, there are about 3 people live in the house with only one person get wage. The average monthly income per household is 47,363.28 Baht but the monthly expense of consumption, goods and services per household are

totally 43,518.47 Baht. In each month, there is insufficient money remaining which it may occur loan and indebtedness. There is only 15.20% of household getting the remittance receiving. Moreover, there are 44.97% and 49.41% of household having loan for business and emergency, respectively.

For checking assumption of multicollinearity, the maximum of Variance Inflation Factor (VIF) of continuous independent variables is 2.773 less than 10 [7], that is, there is no multicollinearity problem, as shown in Table 3.

Table 3. Multicollinearity Diagnostics

Independent Variables	VIF
Household Size (X ₁)	1.226
Number of People who get wages (X ₂)	1.2.1
Income per month per household (X ₄)	1.6.1
Consumption Expense (X ₅)	2...2
Cost of Goods and Services (X ₆)	2.6..

After selecting the independent variables by using the backward stepwise method and the appropriate variables are obtained for the model, it can be concluded that the factors affecting household debt are the household size, number of people who get wages, remittance receiving and loan for emergency. The results are shown in Table 4.

One of the factors determining indebtedness of the household is the household size, according to Haq, Ismail, and Satar [5], Chounlakorn and Kittichotipanit [3] and Yoon, Jean, and Deanna [10]. The results of logistic model indicate that with the increase of the household size and the number of people who get wages, the chances of indebtedness increase. The coefficients of the model show that the size of household increases one person, the chances of indebtedness increases 30%. As expected, the probability of holding debt in later life increases as the number of family increases. Similarly, the number of people who get wages increases one person, the chances of indebtedness increases 19%. Hence, employed persons with higher expected income would have a better access to loans. Employed persons are more inclined towards debt as they easily apply for loans in the

formal sector. Another factor significantly differentiating indebtedness is the remittance receiving. The coefficient of the logistic model indicates that the chances of having debt of the households which receive remittance are 38% lower than that which do not receive remittance. The last important factor included on the model differentiating the indebtedness is the loan for emergency which can be categorized into partial and complete loans where the group of no need of emergency loan and the group of households which got rejected for emergency loan are set to be the reference. The logistic model shows that households with the loan for emergency increase the probability of having debt. When the other independent variables are constant, the results reveal that the opportunity of making debts for household with completed emergency loan is 2.437 times the reference group. In the same manner, holding all other independent variables constant, the opportunity of making debts for household with partial emergency loan is 2.862 times the reference group. Thereby, the emergency expenses may be occurred for example, medical expenses in case of illness and emergency money for an uncertain event. As expected, the probability of holding debt in later life increases as the number of emergency loan increases.

Two additional descriptive measures of goodness-of-fit in Table 4 are Cox and Snell's R² and Nagelkerke's R². The results indicate that the model which includes the four independents variables explains 10.3% and 14.2% of the variation in indebtedness of household. When considering the Hosmer and Lemeshow test statistics, the chi-square of 15.555, with 8 df is not statistically significant (p-value =0.049) at the 0.01 level.

Therefore, the binary logistic regression equation to predict the household debt is P(household with debt)

$$= \frac{e^{-1.942+0.261X_1+0.174X_2-0.478X_3+1.051X_{82}+0.891X_{83}}}{1 + e^{-1.942+0.261X_1+0.174X_2-0.478X_3+1.051X_{82}+0.891X_{83}}}$$

The two rows in Table 5 provide the actual values of the dependent variable, and the two columns the predicted values of the dependent variable, based on the model. Here, there is 1537 household without debt. Using the binary logistic model, the predicted indebtedness is 1346 household without debt and 191 household with debt, consequently, the success rate is 87.6%. Of the 831 household with debt, the predicted indebtedness is 573 household without debt and 258 household

with debt, hence, the success rate is 31.0%. Overall, according to the Table 5, the success rate is 67.7%.

Table 4. Coefficients of Binary Logistic Regression Model for Households with Debt

Independent Variable	$\hat{\beta}$	<i>p</i> -value	Exp($\hat{\beta}$)
Constant	-1.942	0.000*	0.143
Household Size (X_1)	0.261	0.000*	1.298
Number of People who get wages (X_2)	0.174	0.001*	1.190
Remittance Receiving (X_3)	-0.478	0.001*	0.620
Partial loan for Emergency (X_{82})	1.051	0.000*	2.862
Complete loan for Emergency (X_{83})	0.891	0.000*	2.437
Sample size	2368		
Cox and Snell's R^2	0.103		
Nagelkerke's R^2	0.142		

* *p*-value \leq 0.01

Table 5. Coefficients of Binary Logistic Regression Model for Households with Debt

Observed		Predicted		
		Household Indebtedness		% Correct
		without debt	with debt	
Household Indebtedness	without debt	1346	191	87.6
	with debt	573	258	31.0
Overall Percentage				67.7

5 Conclusion

Due to the government continuously resolving household debt problems, it can be seen that it is still not effective and unable to achieve the goal sustainably. This research focuses the factors effecting the household debt in Bangkok and maropitant area (Nonthaburi, Pathum Thani and Samut Prakan provinces) of Thailand using the binary logistic regression model. The results show that there are household sizes, number of people who get wages, remittance receiving, and loan for emergency influencing indebtedness of the household of Bangkok and vicinity in Thailand. The overall correct prediction is 67.7% which is better than chance (50%), but not much. Consequently, the

Thai government should consider these factors in creating the policies and measures to solve household debt problem. However, the household sector must be aware of adjusting the spending behavior appropriately, establishing financial discipline and having the suitable financial planning in the long term, in addition, people should regularly maintain health care to decrease the risk of illness in order to reduce financial problems for household.

References:

- [1] P. Anioła-Mikołajczak, An Application of Logistic Regression in Identification of the Factors Influencing the Use of Credit by Polish Households, *Handel Wewnêtrzny*, Vol.6, No.365, 2016, pp. 5-15.
- [2] S. Chanchoksan, S. Suwanik, and T. Suksumek, *BOT-Nielsen Household Financial Survey*, Bank of Thailand, 2019.
- [3] S. Chounlakorn, and N. Kittichotipanit, Factors Influencing on Family's Debt of Civil Servants in Bangkok and Metropolitan Area, *Journal of Applied Science*, Vol.15, No.2, 2016, pp. 52-67.
- [4] S. Costa, and F. Luísa, Households' Indebtedness: a Microeconomic Analysis Based on the Results of the Households' Financial and Consumption Survey, *Financial Stability Report*, 2012, pp. 133-157.
- [5] W. Haq, N. A. Ismail, and N. M. Satar, Investigation of Household Debt through Multilevel Multivariate Analysis: Case of a Developing Country, *Journal of Reviews on Global Economics*, Vol.7, 2018, pp. 297-316.
- [6] M. H. Kutner, C. J. Nachtsheim, J. Neter, and W. Li, *Applied Linear Statistical Models*, 5th edition, Mc Graw Hill, 2005.
- [7] S. W. Menard, *Applied Logistic Regression Analysis*, 2nd edition, Thousand Oaks, CA: Sage Publications, 2002.
- [8] National Statistical Office, The 2017 Household Socio-Economic Survey whole Kingdom. Bangkok, <http://www.nso.go.th/sites/2014en/Pages/survey/Social/Household/The-2017-Household-Socio-Economic-Survey.aspx>, 2017.
- [9] C. Rayasawath, Factors Affecting the Household Succession in Agricultural Occupation in Nakhon Ratchasima Province, Thailand, *Agriculture*, Vol.8, No.109, 2018, pp. 1-14.
- [10] G. L. Yoon, M. L. Jean, and L. S. Deanna, Predictors of Holding Consumer and Mortgage Debt among Older Americans, *Journal of Family and Economic*, Vol.28, 2007, pp. 305-320.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en_US