An Application Of Two Binary Choice Models As Credit Scoring Tools for Philippine LGU-Borrowers

by

Winona S. Pasimio*

This paper presents the empirical results of applying two binary choice models as credit scoring tools for evaluating the creditworthiness of local government units (LGUs) in the Philippines. Credit scoring is a quick and economical way of evaluating risk especially when there is a large number of potential borrowers. The study's findings suggest that good LGU payors exhibit quite distinct financial characteristics from poor LGU payors. Both models reveal that the percentage rate of change of total revenue is the single most important factor associated with credit-servicing performance. The findings indicate that quantitative credit scoring models can be employed to assess the creditworthiness of Philippine LGUs.

Introduction

Credit scoring is the quantitative evaluation of the creditworthiness of entities such as private persons, business firms, publicly-owned enterprises, subnational governmental jurisdictions and even sovereign countries. Credit scoring techniques cover a broad range of methods, ranging from simple summing-up procedures to sophisticated statistical models that require the use of automated equipment and specialized software. Rapid advances in computer and communications technology have resulted in a rapid acceleration of the rate at which significant innovations in their form and use have been introduced.

Rapid advances in computer and communications technology have resulted in the greater use of automated credit evaluation programs.

These innovations can be seen in the rapidly growing number and types of credit institutions (e.g., banks, credit card companies, credit unions) that have adopted and employed automated credit evaluation programs. These programs are reported to have provided substantial benefits to their users in the form of sharply reduced loan evaluation costs, faster loan processing, better customer targeting and more effective cross-selling of products and services, to name a few.[1]

---

* *Associate Professor of Finance at the University of the Philippines College of Business Administration.*

[1] See for instance Lawson [November 1995], pp. 61-65 and Stetenfeld, [July 1995], pp.32-35.

The study attempts to develop a credit scoring tool for evaluating the creditworthiness of local government units (LGUs) in the Philippines.

From an economic standpoint, the most socially relevant aspect of recent credit scoring innovations lies in the fact that automation appears to have made it feasible and profitable for various types of lenders to consider underwriting the credit needs of so-called high-risk classes of borrowers, i.e., borrowers that would not otherwise be considered as potential credit recipients. Principally, these consist of various types of small business enterprises[2] which, for various reasons, cannot be subjected to the same credit review process that is usually applied to larger firms.

In view of the benefits that have reportedly been derived from the innovative application of credit scoring models, it would seem that credit institutions in the Philippines should explore their possible/expanded application. It is in this light that the present paper has been undertaken.

This paper presents the empirical results of an attempt to employ two binary choice models for the purpose of developing a credit scoring tool for evaluating the creditworthiness of local government units (LGUs) in the Philippines, based on a review of actual LGU performance as borrowers from the Municipal Development Fund (MDF). The MDF is a special revolving fund which serves as the principal mechanism for making available long-term financing to eligible local governments. It is jointly administered by the Department of Finance (DOF), the Commission on Audit (COA) and the Department of Budget and Management (DBM), but the DOF serves as Administrator through its Bureau of Local Government Finance (BLGF).

LGU borrowers were singled out in this study for two principal reasons:

1) As a general rule, LGUs have made limited use of credit to finance their expenditures, whether it be for operational requirements or for capital investments. The passage of a new Local Government Code (R.A. 7160 or LGC) in 1991, however, empowered LGUs to make more active use of domestic credit finance facilities from both public and private sources.

They would have to participate and compete on the basis of their own creditworthiness which can be assessed, in turn, through a credit scoring system like the binary choice models the present paper seeks to develop.

---

[2] The June 1995 issue of *Catalog Ace*, for example, cites small (mail order) catalog companies as being likely to benefit from credit scoring in that they "may find it easier to get a loan, despite the (state of the) economy, due to the computerization of the small business loan approval process," p.7.

2) From an academic standpoint, the availability of credit performance records as well as other borrower characteristics at the BLGF appeared to present a unique opportunity for this researcher to develop and to test a credit scoring model for a specific class of Philippine credit users.

For both pragmatic and academic purposes, therefore, it was believed that the present study should be undertaken.

Methodology

The present paper is largely empirical. It shall not attempt to present, much less develop, any theoretical model to explain the loan repayment behavior of borrowers, in general, nor of Philippine LGU borrowers, in particular. Instead, it adheres to the credit scoring tradition of employing observed borrower characteristics[3] to calculate the probability of default or to classify borrowers into different default or risk categories.

The Binary Choice Models Employed in the Study

In essence, credit scoring models are employed as screening mechanisms to identify poor credit risks from an undifferentiated group of credit applicants. In their simplest form, they classify individual loan applicants into one of two (or binary) categories - the poor credit risk and the good credit risk. For this reason, binary choice models have been generally considered to be the most appropriate and convenient to employ for this purpose.[4]

Credit scoring models are employed as screening mechanisms to identify credit risks from an undifferentiated group of credit applicants.

In this paper, two such modeling techniques are considered: multiple discriminant analysis (MDA) and multinomial logit analysis. MDA is a multivariate statistical technique which seeks to obtain a specific combination of the observed values of a set of independent variables such that the differences among two or more previously defined groups can be maximized. In other

---

[3] The so-called "five C's of credit" - character, capacity, capital, collateral and conditions - are the traditional factors considered by credit managers in evaluating creditworthiness and are invariably discussed in most finance textbooks. See for instance, *Managerial Finance,* 7th Ed. by J. Fred Weston and Eugene F. Brigham [1981], p. 131.

[4] Other statistical techniques may also be employed for this purpose. A review of the relevant literature and a comparative evaluation of the different statistical techniques available is given, for example, in Llanto and Suleik [1988].

words, it seeks to determine a mathematical rule, the so-called "optimal discriminant function," which can classify an entity (or observation) into one of several (in the present case, two) known groups - the good and the poor payors.

The discriminant model classifies an LGU loan applicant either as a good payor or as a poor payor based on a combination of the observed values of a set of predictor variables. The "optimal" linear discriminant function may thus be defined as one which maximizes the differences between the two groups of payors, or conversely, as one which minimizes the probability of classifying a loan applicant into the "wrong" group or category.

Discriminant analysis, then, consists of finding the values of the coefficients of a linear function of the form

$$Z = a + b_1 X_1 + b_2 X_2 + \dots + b_j X_j + \dots + b_k X_k$$

where X_j is the jth independent variable and Z is the discriminant score that maximizes the differences between the two groups.

Most credit scoring studies have utilized the linear form of multiple discriminant analysis (LMDA) which assumes that explanatory variables are multivariately normally distributed, with the *a priori* groups having equal variance-covariance matrices with different means. When these underlying conditions are not satisfied, the LMDA is not technically appropriate for the analysis and the quadratic form of the model should instead be used. When this situation arises, however, using the more complex and nonlinear form of the discriminant function makes the application of the model more difficult and less intuitively satisfying.

The discriminant model classifies an LGU loan applicant either as a good payor or as a poor payor based on a combination of the observed values of a set of predictor variables.

In the light of these difficulties, more recent studies have employed the logit model as an alternative analytical framework. Based on the cumulative logistic probability function, rather than the cumulative normal probability function, a unique maximum is always obtainable with this model and the problems associated with fulfilling normality conditions are avoided. Multinomial logit analysis seeks to derive or "fit" a linear model by using a transformation of the binary classification variable as dependent variable given a set of predictor variables or characteristics. The fitted model is then used to segregate poor payors from good payors.

The logit model predicts the conditional probability or "odds" that an LGU will be a good payor based on the set of predictor variables for that LGU. That is, $P_i = (\text{Good} | X_1, X_2, \dots, X_k)$ where P_i is the probability that LGU i will be a good payor and X_1 through X_k are predictor variables.

Given that P_i approaches 1 if the LGU borrower is "good" and P_i approaches 0 if the LGU is "poor"[5], the probability that a particular LGU loan applicant will prove to be a good payor may be defined as follows:

$$P_i = P(X)$$

$$= \frac{e^u}{1+e^u}$$

where **X** is a shortcut notation for variables X_1 through X_k, $u = a + \sum_{i=1}^{k} b_i X_i$, a and b_i are parametric coefficients, e^u is the odds ratio that an LGU will be a good payor rather than a poor payor, and u is the log (or transformation) of the odds ratio.

The first phase of the study segregated MDF borrowers into good and poor payors by conducting a general review of the repayment performance of MDF borrowers from March 1984 up to August 1993.

Sample and Data

This study was conducted in two phases. The first phase segregated MDF borrowers into good and poor payors by conducting a general review of the repayment performance of MDF borrowers from the time of the MDF's inception (March 1984) up to August 1993. Good payors were defined as those LGU borrowers which voluntarily remitted amortization payments, promptly and in full, to the BLGF, while poor payors were defined as those which were unable to service part or all of their loan amortizations within the 90-day grace period provided for in the covering loan agreements.

As of August 31, 1993, the MDF had a total of 72 LGU-borrowers consisting of eight highly-urbanized cities (HUCs), one independent component city (ICC), nine component cities (CCs) and 54

[5] The threshold value of P_i is equal to 0.50. This value, however, should not be confused with the p - value which signifies the attained level of significance.

municipalities.[6] Of this total number, 14 borrowers (including five new borrowers) had to be excluded because of data insufficiency. Out of the 58 LGU-borrowers which comprised the final sample, 34 borrowers were identified as good payors while 24 were adjudged poor payors.

The second phase of the study then focused on identifying the financial and economic dimensions of the operations of LGUs which can be associated with their credit-servicing performance. Candidate predictor variables were identified based on the results of previous empirical research. Because of budgetary and time constraints, secondary sources of information were tapped for the principal data requirements. These consisted mainly of the Budget Operations Statements which all LGUs submit annually to the BLGF, from which corresponding Statements of Income and Expenditures were drawn up for each LGU.

The second phase of the study then focused on identifying the financial and economic dimensions of the operations of LGUs which can be associated with their credit-servicing performance.

Based on these statements, more than 100 financial variables were computed for each LGU. To these were added eight extraneous factors which were obtained mainly from published census data. Together, these constituted the potential explanatory variables which were initially grouped into five sub-categories: extraneous factors, revenue characteristics, expenditure variables, financial balance ratios and efficiency ratios[7] Efficiency ratios were subsequently merged with the financial balance ratios, reducing the number of variable sub-groups into just four (4).[8]

---

[6] LGUs are classified according to their annual incomes as provided for in E.O. No. 249, s. 1987. The last reclassification under this order, as of the time of data closure, took effect on July 1, 1991. The categorization given here of the MDF borrowers was based on data supplied by Sosmena [1991] and Jamig [1991].

[7] Each of the individual variables as well as the relevant groupings are discussed and operationally defined in Pasimio [1994]. Because of space constraints, they are no longer discussed here.

[8] In this study, variable sub-groupings are necessitated not only by theoretical considerations but also by the fact that the number of explanatory variables being considered is large relative to the number of LGUs included in the sample. In order to comply with statistical estimation procedures, preliminary iterations had to be conducted on each variable sub-group before the two final sets of explanatory variables were subjected to the final logit and the discriminant analyses. The SPSS and the SAS software packages were employed to operationalize the logit and the MDA procedures employed in this study.

Variables

Because of the financial orientation of this study, fiscal characteristics constituted the largest single group of explanatory variables which were considered. These variables consist of financial ratios which were calculated from the previously mentioned Statements of Income and Expenditures available at the BLGF for each LGU.

In addition, easily obtainable socio-economic characteristics which are unique to each LGU were also considered, such as population size, population growth rate, size of land area, income class, and other similar characteristics.

A limited number of extraneous factors, data for which are readily available, were also included in the analysis such as the inflation rate of the region to which the specific LGU belonged. General economic conditions, which past studies have identified as constituting the most significant determinant of the financial health of local governments [Bahl 1984], were included and operationalized through the "significant year" variable.

Empirical Results

Both the logit and the discriminant models were developed using stepwise selection techniques. Potential explanatory variables were added to, or dropped from the models, one at a time, based on their contribution to the overall fit of the model.

As previously noted, exploratory logit or discriminant procedures were successively performed, initially, on the various sub-groups into which each potential explanatory variable had been originally included. In this manner, a final set of candidate predictor variables, one for each model, was chosen and then subjected to the final logit, or discriminant, procedure.

In the case of the logit model, this iterative exploratory procedure eventually resulted in the identification of ten candidate predictor variables shown in Table 1.

The variables of this study consist of financial ratios, socio-economic characteristics which are unique to each LGU, inflation rate of the region, and general economic conditions.

Table 1
Candidate Predictor Variables : Logit Model

Ratio	Formula	Description
1. RPTTRGRO	(RPTTRC - RPTTRP) / RPTTRP	Change in the ratio of real property taxes to total revenue from the prior year to current year
2. TTRTRC	TTRC / TRC	Total tax revenue of the current year to total revenue of the current year
3. BANATRG	BANATRC / BANATRP	Rate of change in ratio of BIR allotments and national aid to total revenue, from prior year to current year
4. LSRTEGRO	LSRTEC / LSRTEP	Rate of change in the ratio of locally-sourced revenues to total expenditures from prior year to current year
5. LSRAAGRO	(LSRAAC - LSRAAP) / LSRAAP	Rate of change in locally-sourced revenues plus aid and allotments from prior year to current year
6. RAR3C	TEC / TRC	Ratio of total expenditures to total revenue for the current year
7. INFLATE	n. a.	Inflation rate during relevant year of region to which LGU belongs
8. TTRGRO	TTRC / TTRP	Rate of change in total tax revenues from prior year's level
9. TNTRGRO	TNTRC / TNTRP	Rate of change in total non-tax revenues from the prior year's level
10. LSRGRO	(LSRC - LSRP) / LSRP	Rate of change in locally-sourced revenues from prior year to current year

The Final Logistic Regression Equation

The final logit procedure that was undertaken consisted of only two steps. After the first iteration, the following solution to the logistic regression equation was obtained:

$$U = -0.5792 + 0.1146 \text{ TTRGRO}$$

TTRGRO = the percentage change in total tax revenues from the prior year to the current year.

The regression coefficients, the p- value and the classification summary of the first logit model are shown in Tables 2 and 3. The statistics show significant overall discriminatory power for the model and the overall classification accuracy amounted to 89.66 per cent. It is worthwhile noting that the model is perfectly accurate in identifying poor payors but is less efficient in identifying good payors.

Table 2
Selected Statistics : First Logit Model

Variable	Coefficient	P - value
TTRGRO	0.1146	0.0013^{s}
constant	- 0.5792	—

s = significant at 5 % level

Table 3
Classification Results : First Logit Model

	Predicted Poor (P)	Predicted Good (G)	Percent Correct
Observed			
Poor (P)	24	0	100.00%
Good (G)	6	28	82.35%
		Overall	89.66%

At the next iteration, another variable, LSRTEGRO, entered the equation for U as follows:

$$U \;=\; -\,42.5990 - 40.5133 \text{ LSRTEGRO} + 36.4945 \text{ TTRGRO}$$

where LSRTEGRO = the percentage change in the ratio of locally-sourced revenues to total expenditures from the prior year to the current year,

and TTRGRO = the percentage change in total tax revenues from the prior year to the current year.

The regression coefficients, the p - value and the model's overall classification accuracy are shown in Tables 4 and 5.

Table 4
Selected Statistics : Second Logit Model

Variable	Coefficient	P-value
LSRTEGRO	- 40.5133	0.6952
TTRGRO	36.4945	0.6948
Constant	- 42.5990	—

Table 5
Classification Results : Second Logit Model

	Predicted		
	Poor (P)	Good (G)	Percent Correct
Observed			
Poor (P)	23	1	98.83%
Good (G)	0	34	100.00%
		Overall	98.28%

The second logit equation clearly yields a higher level of overall classification accuracy. In terms of the individual coefficients of the predictor variables, however, the Wald test indicates that the coefficients are not significant at the 5 percent level. For this reason, it would seem that the first logit equation is to be preferred over the second logit equation.

The logit modeling exercise was able to yield relatively accurate results although the first logit equation does not do very well in identifying "good" payors.

Based on these results, the logit modeling exercise may be said to be quite satisfactory in that it was able to yield very parsimonious and relatively accurate results. Both logit models are parsimonious in that they include, respectively, only one and two out of the more than one hundred variables which were initially considered. Both models also registered high overall levels of accuracy, although the first logit equation does not do very well in identifying "good" payors.

It should be noted, however, that these classification results are highly biased because the same set of sample data which was used to "fit" the logit function was also used to evaluate its performance.

The Optimal Discriminant Function

Satisfactory results have also been obtained with respect to the discriminant model. An optimal linear discriminant function was obtained that included nine predictor variables. These variables, presented in the order that they entered the discriminant model, have the following unstandardized[9] and standardized coefficients (Table 6):

[9] These coefficients will serve as the multipliers of the variables when they are expressed in the original units.

Table 6
Unstandardized and Standardized Discriminant Function Coefficients of the Optimal Linear Discriminant Function

Variable	Unstandardized Coefficient	Standardized Coefficient
TTRGRO	-0.9130306E-01	-1.55516
TTRTRGRO	0.6909325E-01	1.27140
BANATRG	-0.2549367E-01	-0.53070
LSRGRO	0.1136331E-01	0.30292
SIGYEAR	0.1517585	0.21208
LANDAREA	0.3038720E-03	0.23683
TTRTRC	-1.608299	-0.25605
TNTRCAP	0.3859065E-02	0.25220
BLTGRO	0.7372560E-02	0.23314
(constant)	0.5475327	

where TTRTRGRO = the rate of change in the ratio of total tax revenues to total revenues from the prior year to the current year;

BANATRG = the rate of change in the ratio of BIR allotments and national aid to total revenues from the prior year's level;

LANDAREA = land area in square kilometers;

TNTRCAP = total non-tax revenues per capita;

BLTGRO = the rate of change in business and local taxes from the prior year to the current year,

and all the other variables are as previously defined.

The classification summary of the discriminant model is shown in Table 7.

Table 7
Classification Results of The Discriminant Model

Actual Group Membership	Number of Cases	Predicted Group Membership	
		Poor	Good
Poor Payors	24	24	0
		100 %	0 %
Good Payors	34	0	34
		0 %	100 %

Percent of "group" cases correctly classified : 100.00 %

As Table 7 shows, the optimal discriminant function has correctly classified all of the sample LGU borrowers into their respective actual groups, resulting in a 100 percent overall classification rate. As previously pointed out, however, these results are quite biased, since the estimation sample was also used to evaluate the model's performance.

The classification results of the discriminant model which were validated using the Lachenbruch technique yielded almost unbiased estimates of the misclassification probabilities.

Unlike the logit function which could not be validated because of the small size of the sample population, the classification results of the discriminant model were validated with the use of the Lachenbruch procedure. This technique, also called the "jackknife," is performed by holding out one observation at a time, estimating the discriminant function based on the remaining ($n_1 + n_2 - 1$) observations, and then classifying the held out observation. The same procedure is repeated until all observations are classified. This method is believed to yield almost unbiased estimates of the misclassification probabilities, and is especially suited for cases like the present investigation where the smallness of the sample being analyzed prevents the use of a hold-out group.

The Lachenbruch procedure was performed on the optimal linear discriminant function in two ways, based on two alternative estimates of the prior probability that an LGU would belong to either the good or poor group when no information about it is available. Proportionate prior probability was estimated based on the proportionate number of good and poor payors among the original sample of 58 LGUs. Equal prior probabilities, on the other hand, as the term implies, was estimated based on the assumption that a new borrower is equally likely to be a good or a poor payor in every case that a loan is approved, else such a loan would not have been approved in the first place.

In the case where equal prior probabilities were assigned to the good and poor payors, the linear discriminant function performed with 88.48 per cent accuracy. In the case where proportionate probabilities were assigned, the linear function performed with 93.10 per cent accuracy.[10]

Conclusion

The statistical results that have been obtained in this study suggest that good LGU payors exhibit quite distinct financial characteristics from poor LGU payors. Both the logit and the discriminant analyses reveal, for example, that the percentage rate of change of total tax revenues of individual LGU- borrowers is the single most important factor that may be associated with credit-servicing performance.

The statistical results also indicate that quantitative credit scoring models, like the logit and the discriminant models that have been developed here, can be meaningfully employed for the purpose of assessing the creditworthiness of Philippine LGUs, about which little credit information and experience have been accumulated.

The results of the study also indicate that indices of growth (or decline), rather than indices of absolute size, however defined or quantified, are more meaningful indicators of credit quality and debt-servicing performance. This is a finding of the present study that has been consistently validated across all the significant size categories that had been considered, which include, among others, LGU income classification, population size, population growth rate, and other similar factors.

This conclusion is further supported by the separate finding that income classification has not performed well as an indicator of creditworthiness or of bankability, based on the MDF experience

The logit and the discriminant analyses reveal that the percentage rate of change of total tax revenues of individual LGU- borrowers is the single most important factor that may be associated with credit-servicing performance.

[10] When the sample data was subjected to Box's M test, results indicated that group covariance matrices were unequal, indicating that the quadratic, rather than the linear form of the discriminant model should be employed. A quadratic function was therefore estimated, which performed with 100.00 per cent classification accuracy, regardless of how the prior probability of being a good payor was estimated. When subjected to the Lachenbruch procedure, the quadratic function performed with 97.92 per cent and 98.28 per cent accuracy, respectively, when equal and proportionate prior probabilities were assumed.

documented in this study. In the Philippines, this observation validates a similar finding that was reported by Yoingco [1986].

From a practical standpoint, this particular observation is considered one of the more significant findings of this study, specially in the light of the results of the survey conducted by Sanchez [1992], who reported that some banks have "resorted to LGU targeting solely on the basis of income class criterion" because of the scarcity of information concerning LGUs.[11]

It should be emphasized that quantitative credit scoring models should not be employed as the sole or even primary criterion for deciding whether or not a particular LGU borrower should be granted credit because there are other factors which must considered.

From a statistical standpoint, it should be noted that the satisfactory statistical results that have been reported here are limited by the lack of representativeness of the present sample and the relatively small size of the sample population on which they have been based.

Finally, it should be emphasized that quantitative credit scoring models, like the ones presented here, should not be employed as the sole or even primary criterion for deciding whether or not a particular LGU borrower should be granted credit, for the simple reason that there are other more or equally weighty considerations that the model does not take into account. It is strongly recommended, however, that such analytical techniques, as well as others that may still be developed, should be employed in evaluating Philippine LGU loan applicants in order to facilitate the loan evaluation process.

---

[11] A similar observation is reported by Sullivan [1994] who cites borrower income as an "example of a traditional underwriting standard that does not hold up well under empirical scrutiny." She cites the fact that in most of the six or seven very large mortgage portfolios analyzed by Fair Isaac, income has not appeared as a good predictor of credit risk, and that in a number of cases, a negative correlation has been observed between the two variables. She reports that very high-income borrowers, as a group, showed the greatest likelihood to default on a loan.

References

1. Asher, J., 1994. "Look What Credit Scoring Can Do Now," *ABA Banking Journal* (May): 51-52.
2. Bahl, R., 1984. *Financing State and Local Government in the 1980s.* New York : Oxford University Press.
3. Borowsky, M., 1995. "Scoring Puts Up Higher Numbers," *US Banker* (January) : 44.
4. Carrol, J. R., 1995. "A Smarter Approach to Collections," *Mortgage Banking* (February): 59-64.
5. *Catalog Ace,* June 1995.
6. Espahbodi, P., 1991. "Identification of Problem Banks and Binary Choice Models," *Journal of Banking & Finance,* vol. 15: 53-71.
7. Jamig, C. L., 1991. *Local Governments in the Philippines: Basic Data and Information.* Quezon City : Local Government Center, College of Public Administration, University of the Philippines.
8. Lawson, J. C., 1995. "Knowing the Score," *US Banker* (November) : 61-65.
9. Llanto, G. and M. Suleik, 1988. "Early Warning of Bank Failure: The Role of Predictive Models." Unpublished (November).
10. Pasimio, W. S., 1994. *A Discriminant Model for Assessing the Creditworthiness of Local Government Units,* Unpublished doctoral dissertation, University of the Philippines, Quezon City.
11. Rowland, J., 1995. "Credit Scoring," *Credit World* (September/October): 21-24.
12. Sanchez, A., 1992. "Regulatory-Institutional Framework of Financial Markets in the Philippines," *Local Government Credit Finance Overall Study - Book I,* Associates in Rural Development Inc. (May 27, 1992).
13. Sosmena, G. C. Jr., 1991. *Decentralization and Empowerment.* Local Government Development Foundation, Inc., Makati, M.M.
14. Stetenfeld, B., 1995. "Push-button Loans," *Credit Union Management* (July): 32-35.
15. Sullivan, D., 1994. "Scoring Borrower Risk," *Mortgage Banking* (November): 94-98.
16. Yoingco, A., 1986. "Credit Financing for Local Governments in the Philippines," *Regional Development Dialogue,* (Autumn) : 7: 2.