



Munich Personal RePEc Archive

**Predicting risk of credit default using  
discriminant approach: A study of tribal  
dairy farmers from Jharkhand**

MK, Sinha and JP, Dhaka

DWM, Bhubaneswar-751023 India

June 2013

Online at <https://mpra.ub.uni-muenchen.de/55935/>

MPRA Paper No. 55935, posted 15 May 2014 06:35 UTC

# Predicting Risk of Credit Default using Discriminant Approach: A Study of Tribal Dairy Farmers from Jharkhand

MUKESH KUMAR SINHA<sup>a\*</sup> and J. P. DHAKA<sup>b</sup>

<sup>a</sup>Directorate of Water Management, Bhubaneswar-751023

<sup>b</sup>National Dairy Research Institute, Karnal-132001

## ABSTRACT

The study has identified the factors that discriminate defaulters from non-defaulters in the credit market using the survey data from 240 households. A discriminant function was employed to examine the relative importance of different socio-economic factors making borrowers to default. The magnitude of coefficient of the function is an indicator of the relative importance of individual variable. The study has suggested that higher per-capita income from crop production (38.72%), higher per-capita income from dairying (31.62%), percentage of expenditure in total income (16.87%), off-farm income sources (6.43%) and more earning adults in the family (6.36%) are the important factors to make the borrowers non-defaulters and vice versa for defaulters. Further, the confusion matrix of the derived classification analysis has cross-verified the predicted variable and has found the group classified correctly by 68.3 per cent. Hence, the model can be regarded to be valid in predicting a defaulter precisely based on the localized social factors. The study will help in addressing the concern of the credit institutions in assessing the credit risk capital and risk adjusted outcome for serving a larger group of smallholders community.

**Key words:** Discriminant function, Credit, Defaulter and Dairy farmers

JEL Classification: Q14, P25

## Introduction

In India, animal husbandry is an integral part of diversified agricultural systems. It contributes more than one-fourth to the agricultural gross domestic product and supports the livelihood of over 200 million rural poor (World Bank, 1999). The unique feature of animal husbandry is that it generates a continuous stream of income and employment and reduces seasonality in the livelihood patterns (Birthal and Ali, 2005; Sinha and Thombare, 2014). Several empirical studies have indicated that livestock-rearing has a significant positive impact on equity in terms of income and employment generation and poverty reduction in the rural areas (Sinha *et al*, 2012; Enoma, 2010; Abdou, 2009; Singh

---

\* Author for correspondence, Email: drmukesh@scientist.com

and Hazell, 1993; Adams and He 1995; BIRTHAL and Ali 2005) as the distribution of livestock is more egalitarian compared to land (Ahuja, 2000; Taneja and BIRTHAL, 2004; Sinha *et al* 2014). It is more so in the case of tribal agriculture, where the prospects of enhancing crop production are limited for the obvious reasons of topographic and technological impediments. Therefore, animal husbandry assumes special significance to ameliorate poverty among tribal people. Efforts have been made through various schemes and development programmes to provide financial and technical assistance to the tribal people, but there is a recurrent relapse of poverty due to various reasons. As a result, almost half of the tribal population is still in below the poverty line segment (Soudarssanane and Thiruselvakumar, 2009; GOI, 2012).

The provision of credit to the tribal animal husbandry is important in creating livelihood options and better production practices. More the livelihood options the borrowers have, the less are they vulnerable to crises. Further, the purpose of sanctioning credit to the production process is that it will provide sufficient additional resources to meet the repayment obligations and generate a reasonable surplus for the producers. The repayment of loan by the borrowers is the basic requirement for the smooth functioning of the credit institution. It is in this context, that the factors influencing the loan repayment assume significance and identify the potential defaulters based on their social and economic parameters.

## **Materials and Methods**

The study is based on a sample of 240 farmers, covering 120 beneficiaries and 120 non-beneficiaries of dairy loan, selected through the multi sampling technique from 3 clustered villages spread over three blocks of the Ranchi district in the Jharkhand state. In the present context, the interpretation of a viable small farmer represents the one capable of at least maintaining the prevailing standards of living and is able to pay off the total farm expenses. The farm business in study has included the income from crops, dairying, services, wages, etc. To examine the relative importance of different socio-economic factors in discriminating between non-defaulters and defaulters of the loan repayment, discriminant function analysis was carried out. The coefficient of discriminant function

measures the net effect of an individual variable, when all the other variables are taken as constant. The form of the function used for analysis was:

$$Z = \sum_{n=1}^{11} \beta_n X_n$$

where, Z = Total discriminant score for loan defaulters and non-defaulters,

$\beta$  = Discriminant coefficient or weight for that variable, and

$X_i$  = Set of eleven variables thought to influence the repayment of loan; these included size of operational holding in acres (SIZE), number of milch animals (MILCH), per-capita crop income in rupees (CROPINC), per-capita dairy income in rupees (DAIRINC) and per-capita income from off-farm activities in rupees (NFINC), expenditure to income ratio in rupees (EXPRATIO), investment in dairying in rupees (INVDAIRY), percentage earning adults in family, (EARNADU) per-capita food expenditure in rupees (EXPFOOD), per-capita expenditure on dairy products in rupees (EXPDAIRY), and educational level (EDUC) of family head in schooling years.

For the application of discriminant function, two groups of equal size are required (Balakrishna and Iyer, 1968). In the present study, there were 37 non-defaulters and 83 defaulters in the 120 loan beneficiaries. Therefore, a sub-sample of 37 defaulters from the total 83 defaulters was randomly taken in order to make both the groups alike for analysis. The calculation of discriminant function involved the solution of the following 11 equations shown in the matrix notation (Brandow and Potter, 1953).

$$S I = D$$

where,

$$S = \begin{pmatrix} S_{11} & S_{12} & \dots & S_{1K} \\ S_{21} & S_{22} & \dots & S_{2K} \\ : & : & : & : \\ S_{k1} & S_{k2} & \dots & S_{kK} \end{pmatrix} \quad I = \begin{pmatrix} I_1 \\ I_2 \\ : \\ I_k \end{pmatrix} \quad \text{and } d = \begin{pmatrix} d_1 \\ d_2 \\ : \\ d_k \end{pmatrix}$$

where,

$K = 1, 2, 3, \dots, 11,$

$I_k =$  Vector of coefficient of discriminant functions,

$S_k I_k =$  Pooled dispersion matrix, and

$d_k =$  Elements representing difference between means of two groups.

The discriminant function was tested for significance to examine whether the variables considered together were sufficiently discriminating between the groups of defaulters and non-defaulters or not. The Mahalanobis  $D^2$  test was used to measure the distance between the two groups. After transformation of the  $D^2$  statistics, it becomes an F statistic, which was then used to see the group difference from each other.

$$D^2 = \sum_{i=1}^{11} \sum_{k=1}^{11} C_{ik} d_i d_k$$

$$F = \frac{N_a N_b (N_a + N_b - P - 1)}{P(N_a + N_b) (N_a + N_b - 2)} \times D^2$$

where,  $C_{ik}$  is the inverted matrix for the coefficients,  $D_i d_k$  is the matrix of the product of mean differences,  $P$  is the number of characteristics. The value of  $F$  is to be tested for significance with  $(P)$  and  $(N_a + N_b - P - 1)$  degrees of freedom.

## Results and Discussion

The socio-economic characteristics of the borrowers together with the means and their mean differences for the two groups of non-defaulters and defaulters of the loan-repayment were analyzed. The discriminant function for the data was estimated as:

$$Z = 0.5941 \text{ SIZE} + 0.2390 \text{ MILCH} + 0.0032 \text{ CROPINC} + 0.0021 \text{ DAIRINC} + 0.0287 \text{ NFINC} - 0.2321 \text{ EXPRATIO} + 0.00012 \text{ INVDAIRY} + 0.1162 \text{ EARNADU} - 0.0032 \text{ EXPFOOD} - 0.0188 \text{ EXPDAIRY} + 0.0164 \text{ EDUC}$$

The discriminant function was tested for significance to examine whether or not the characteristics considered together were sufficiently discriminating between the groups of non-defaulters and defaulters of the loan-repayment. The test of significance of discriminant function is a test of hypothesis that there is no difference in the mean values

of the chosen characteristics in the two populations of non-defaulters and defaulters of loan-repayment. The values of  $D^2$  and variance ratio were found to be 5.0654 and 5.0931, respectively. Since the tabulated value of F statistics ( $F_{11, 62}$ ) at 5 per cent level is 2.49, the discriminant function was found to be significant. This means that the eleven characteristics considered together are useful in classifying the borrowers of loan into the groups of non-defaulters and defaulters. To examine the relative importance of characteristics based on their power to discriminate between the two borrowing groups, the percentage contribution of each character to the total distance measured was calculated and the results are exhibited in the Table 1. The results revealed that the characteristics like per-capita income from crop production (17.32%), per-capita income from dairying (19.39%), per-capita off-farm income (15.62%), percentage of expenditure in total income (19.79%) and the percentage of earning adults (11.29%) were the major characteristics, which led to classifying the loan borrowers into two groups of defaulters and non-defaulters.

The students 't'- test was conducted for testing the mean difference between the groups for each variable and they exhibited significant 't' values for the above identified variable at 5 per cent level of significance. Since, the variable per-capita income from crop production ( $X_3$ ), per-capita income from dairying ( $X_4$ ), per-capita off-farm income ( $X_5$ ), percentage expenditure to total income ( $X_6$ ), and the percentage of earning adults in total family ( $X_8$ ) were significant. Hence, these variables were judged the major characteristics, which discriminate the borrowers into non-defaulters and defaulters of loan-repayment.

**Table 1 . Factor contribution of individual characteristics to total distance measured**

| Socio-economic variable | Coefficient ( $I_k$ ) | Mean difference ( $d_k$ ) | Contribution of variable ( $I_k \times d_k$ ) | Factor contribution (%) |
|-------------------------|-----------------------|---------------------------|---|-------------------------|
| SIZE ( $X_1$ )          | 0.5941                | 0.312                     | 0.1853  | 5.78                    |
| MILCH ( $X_2$ )         | 0.2390                | 0.270                     | 0.0645  | 2.08                    |
| CROPINC ( $X_3$ )       | 0.0032                | 167.556                   | 0.5362  | 17.32                   |
| DAIRINC ( $X_4$ )       | 0.0021                | 285.970                   | 0.6005  | 19.39                   |

|                            |         |          |         |        |
|----------------------------|---------|----------|---------|--------|
| NFINC (X <sub>5</sub> )    | 0.0287  | 16.850   | 0.4836  | 15.62  |
| EXPRATIO (X <sub>6</sub> ) | -0.2321 | -2.640   | 0.6127  | 19.78  |
| INVDAIRY (X <sub>7</sub> ) | 0.0001  | 1581.973 | 0.1898  | 6.13   |
| EARNADU (X <sub>8</sub> )  | 0.1162  | 2.990    | 0.3474  | 11.22  |
| EXPFOOD(X <sub>9</sub> )   | 0.0032  | -30.108  | -0.0963 | -3.11  |
| EXPDAIRY(X <sub>10</sub> ) | -0.0188 | -8.241   | 0.1549  | 5.00   |
| EDUC (X <sub>11</sub> )    | 0.0164  | -6.092   | 0.0179  | 0.005  |
| Total                      |         |          | 3.0965  | 100.00 |

The discriminant function was re-run by taking only those five significant variables in the equation to see whether these characteristics alone could discriminate the defaulters and non-defaulters significantly. It was concluded from the analysis that these characteristics were very useful for measuring distance in the discriminating power. The new discriminating function, taking only the significant factors, was estimated as follows:

$$Z = 0.0023 \text{ CROPINC} + 0.0011 \text{ DAIRYFINC} + 0.0038 \text{ NFINC} \\ - 0.0636 \text{ EXPRATIO} + 0.0212 \text{ EARNADU}$$

Again, the discriminant function was tested to examine whether these characteristics considered together could significantly discriminate between the groups of defaulters and non-defaulters. The  $D^2$  and variance ratio were worked out to be 4.6241 and 5.3154, respectively. Since the tabular value of  $F_{(5,62)}$  at 5 per cent level is 4.43, the discriminant function was significant. This means that the five characteristics considered together were useful in classifying the borrowers into the groups of non-defaulters and defaulters of loan-repayment. Thus, the difference in the groups was mostly oriented towards per-capita income, percentage expenditure in total income and percentage of earning adult in the family. These discriminating variables obtained were quite contrary to the variables reported by of George *et al.* (1984), Lekshmi *et al.* (1998) and Gandhimathi (2012), while the result was in conformity with the findings of Pouchepparadjou (1998), Bandyopadhyay (2006) and Nawai and Shariff (2010).

**Table 2. Relative importance of significant characteristics for defaulters and non-defaulters of loan-repayment**

| Socio-economic variable | Coefficient ( $I_k$ ) | Mean difference ( $d_k$ ) | Contribution of variable ( $I_k \times d_k$ ) | Factor contribution (%) |
|-------------------------|-----------------------|---------------------------|---|-------------------------|
| CROPINC ( $X_3$ )       | 0.0023                | 167.556                   | 0.3853  | 38.72                   |
| DAIRYINC ( $X_4$ )      | 0.0011                | 285.970                   | 0.3146  | 31.62                   |
| NFINC ( $X_5$ )         | 0.0038                | 16.850                    | 0.0640  | 6.43                    |
| EXPRATIO ( $X_6$ )      | -0.0636               | -2.640                    | 0.1679  | 16.87                   |
| EARNADU ( $X_8$ )       | 0.0212                | 2.990                     | 0.0633  | 6.36                    |
| Total                   |                       |                           | 0.9951  | 100.00                  |

Further, the relative importance of the characteristics to discriminate between the two groups of borrowers, the percentage contribution of each variable to the total distance measured was examined and the results are presented in Table 2. The magnitude of coefficient of a function is an indicator of the relative importance of an individual variable. The coefficients in the Z equation suggest that higher per-capita income from crop production, higher income from dairying, percentage of expenditure in total income, off-farm income sources and more earning adults in the family contributed high value of Z, explained a major share in discriminating the non-defaulters from defaulters followed by the percentage earning adults and off-farm income. The weights associated with these characteristics to the total distance measured were obtained as 38.72, 31.62, 16.87, 6.43 and 6.36, respectively.

The discriminant function was later used to predict whether a borrower is likely to be a non-defaulter or defaulter of loan-repayment. The mean discriminant score,  $Z_1$  for the non-defaulters and  $Z_2$  for defaulters were found to be 0.316 and -1.322, respectively.



The critical mean discriminant score ( $Z$ ) for the two groups was found to be  $-0.503$ . This implies that, if the discriminant score for a respondent on the basis of significant variable for his data is found to be more than  $-0.503$ , he can be predicted to be a non-defaulter, otherwise he is likely to be a defaulter. The high value of  $Z$  corresponds to a non-defaulter and low value to a defaulter. This can be shown as:

| Discriminating score<br>for non-defaulter ( $Z_1$ ) | Separating mean score<br>( $Z$ ) | Discriminating score<br>for defaulter ( $Z_2$ ) |
|---|----------------------------------|---|
| 0.316   | -0.503                           | -1.322  |
|   |                                  |   |

It was interesting to see as to what proportion of respondents considered in the study was classified rightly by the function. With this view, the whole sample of 120 respondents was classified into defaulters and non-defaulters. Then, it was compared with the actual classification. It is called as derived classification analysis.

**Table 3. Classification results (Confusion matrix) of loan borrower groups**

| Loan repayment group | Number of cases | Predicted group membership |      |                |      |
|----------------------|-----------------|----------------------------|------|----------------|------|
|                      |                 | Defaulters                 | %    | Non-defaulters | %    |
| Defaulters           | 83              | 50                         | 60.2 | 33             | 39.8 |
| Non-defaulters       | 37              | 5                          | 13.5 | 32             | 86.5 |

The percentage of cases classified correctly is the productive power of fitted discriminant function. It is also important to consider the observed misclassification rate to that by chance, while evaluating the measure. It was seen from the Table 3 that 50 out of 83 defaulters (60.2%) and 32 out of 37 non-defaulters (86.5%) were classified rightly in the  $Z$  function. The number of wrongly-classified respondents was 38 in 120 respondents. Therefore, grouped cases classified correctly were 68.3 per cent. Thus, the model has been found to be valid to predict whether a borrower is likely to be a defaulter or non-defaulter, more precisely.

## Conclusions

The credit support to the smallholder tribal families for the dairy development is an important source of viable mix of livelihood options and improving quality of life. The study has investigated the localized social factors that can be used by the credit institutions in risk rating a farmer-customer. The study has suggested that lower income from crop production, dairy and off-farm activities coupled with high expenditure proportionate to income and smaller number of earning adults in the family leaves only a smaller surplus with farmers making them to be a defaulter. The derived classification analysis has cross-verified the predicted variable and has found that the group classified correctly by 68.3 per cent as factors of default. Thus, model has been found valid in predicting a defaulter of loan-repayment based on the localized factors precisely. The study will address the concern of the credit institution in advance to assess the credit risk capital and risk adjusted outcome for serving a larger group of smallholder's community.

## References

- Abdou, H. A. (2009) An evaluation of alternative scoring models in private banking. *The Journal of Risk Finance*, **10**: 38-53.
- Adams R. H. (Jr) and He, J. J. (1995) *Sources of income inequality and poverty in rural Pakistan*, Research Report 102. International Food Policy Research Institute, Washington, D.C. <http://www.ifpri.org/pubd/abstract/abstr102.htm>
- Ahuja, V., George, P.S., Ray, S., Kurup, M.P.G. and Gandhi, V. (2000) *Agricultural services and the poor: case of livestock health and breeding services in India*, IIM, Ahmedabad, The World Bank, and SDC, Bern, Switzerland.
- Soudarssanane, M. Bala, Thiruselvakumar, D. (2009) Overcoming problems in the practice of public health among tribals of India. *Indian Journal of Community Medicine*, **34** (4):283-287.
- Balakrishna, S. and Radha Iyer, B. (1968) Characteristics of adopters of family planning methods in Panjab: A discriminant function approach. *Behavioural Science and Community Development*, **2** (1) : p16.
- Bandyopadhyay, A. (2006) Predicting probability of default of Indian corporate bonds: Logistic and Z-score model approaches. *The Journal of Risk Finance*, **7**(3) : 255-272.

- Birthal, P. S. and Ali, J. (2005) Potential of livestock sector in rural transformation. *In: Rural Transformation in India: The Role of Non-farm Sector* Eds: Rohini Nayyar and A.N.Sharma. Institute for Human Development and Manohar Publishers and Distributors, New Delhi.
- Brandow, G.E. and Potter, A.K. (1953) An application of linear discriminant function. *Rural Sociology*, **18**(4): 321-326.
- Enoma, A. (2010) Agricultural credit and economic growth in Nigeria: An empirical analysis. *Business Economics Journal*, **14** (7) : 1-7.
- Gandhimathi, S. (2012) Determinants of repayment and overdues in agricultural sector, *International Journal of Economics and Business Research*, **4** (5) :590-602.
- George, P.T., Namasirayam, D. and Ramachandraiah, G. (1984) Application of discriminant function in the farmer's repayment performance. *Journal of Rural Development*, **3** (3) : 149-167.
- Lekshmi, S., Rugmini, P. and Thomas, J. (1998) Characteristics of defaulters in agricultural credit use – A macro level analysis with reference to Kerala. *Indian Journal Agricultural Economics*, **53**(4): 640-67.
- Nawai, N. and Shariff, M.N.M. (2010) Determinants of repayment performance in microcredit programs: A review of literature. *International Journal of business and Social Science*, **1**(2) : 152-161.
- GOI (Government of India) (2012) *Press Note on Poverty Estimates, 2009-10*, Press Information Bureau, Planning Commission, Govt. of India, New Delhi.
- Pouchepparadjou, A.(1998) Socio-economic characteristics of non-defaulting farmer borrower of commercial banks. *State Bank of India Monthly Review*, **37**(6): 321-327.
- Singh, R. P. and Hazell, P. B. R. (1993) Rural poverty in the semi-arid tropics of India: Identification, determinants and policy interventions, *Economic and Political Weekly*, 28: (12 and 13) : pp A-9:A-15.

- Sinha, Mukesh K., Dhaka, J.P. and Meena, M.S. (2012) Milk production economics and micro financing impacts in Chhotanagpur plateau of Jharkhand, *Indian Journal of Animal Science*, **82** (7) : 757–761.
- Sinha, Mukesh K., Dhaka, J.P. and Mondal, B. (2014) Analyzing social attributes of loan default among small Indian dairy farms: A discriminant approach, *Scientific Research and Essays*, 9 (2), 2354-2358.
- Sinha, M.K. and Thombare, N.N. (2014) Incidence and impacts of clinical mastitis in dairy cattle farms: Case of Maharashtra Farmer, *Indian Journal of Dairy Sciences*, 67 (1), Jan-Feb. 2014 Issue.
- Taneja, V. K. and Birtal, P. S, (2004) Role of buffalo in food security in Asia, *Asian Buffalo Magazine*, **1**(1): 4-13.
- World Bank (1999) *India Livestock Sector Review: Enhancing Growth and Development*, Washington DC.