Mitigating Agricultural Lending Risk: An Advanced Analytical Approach

Aditi Singh and Nishtha Jain

1 Introduction

As per the Situation Assessment Survey (SAS) for Agricultural Households by NSSO 70th round, in 2012–13, almost 40% of the agricultural households still relied on non-institutional sources for their credit needs, an increase of almost 11% over 1990–91. Moneylenders form a major part, around 26%, of that non-institutional credit. Even with the rising credit disbursements and loan waivers, we have not been able to improve the situation of our farmers. Empirical and situational evidences suggest that generalized loan waivers have made less than marginal contribution toward improving credit situation of farmers (FE Online 2018). It rather creates a situation of moral hazard which affects the loan repayment behavior of all the farmers. Since 2011–12, percentage of bad loans from agriculture sector has climbed every year and the growth rate of loans disbursed to this sector has become close to stagnant. In FY 2018, banks disbursed only an additional 6.37% to this sector which is the lowest in a decade (Iyer 2019).

Agriculture sector poses risks for the banks in multiple forms. Lending to the agriculture sector has been adversely affected in recent times, and it could be indicative of the deteriorating asset quality (Trends and Progress Report, RBI 2017–18).

Reserve Bank of India (RBI), which is India's central bank and regulating body, oversees the functioning of all the banks that operate in the country and has identified certain priority sectors, of which agriculture is also a part, to ensure necessary credit flow to these sectors. However, banks especially the private and foreign banks are not familiar with India's agricultural landscape and feel reluctant to lend in this sector (Jayakumar 2018). Because of inadequate knowledge of the risks pertaining to these sectors, they refrain from direct lending and instead end up investing in Rural Infrastructure Development Fund (RIDF) of NABARD or buying Priority

---

A. Singh (✉) · N. Jain
Data Modeler, Experian Credit Information Company of India, Mumbai, India
e-mail: aditi.s11@gmail.com

Sector Lending Certificates (PSLCs) to meet their priority sector lending targets. This paper aims to shed some light on the financial landscape of agricultural sector in India to help banks understand the market and model the associated credit risk in an improved manner and hence bridge the gap between the borrower and the lender.

Credit risk associated with an individual can be classified into two broad categories:

1. Capacity to pay and
2. Intention to pay.

Capacity to pay is governed by the principle that the individual should have the ability to generate a steady flow of income which depends largely on his demographic features such as age, qualification and profession. These features along with income and existing assets of the individual determine whether the individual has the capacity to pay back the loan. Following Maurer (2014), risks in agriculture finance can be broadly classified into the following categories which influence an agricultural household's capacity to pay:

i.   Production Risks: Agriculture production in India is fraught with the risk of poor monsoon, disease and pests due to which farmer's income suffers. Lack of proper irrigation facilities, immense dependency on monsoon, lack of good quality seeds and chemical fertilizers can lead to suboptimal output and therefore insufficient generation of income to pay back the loans.

ii. Market Risks: There are price uncertainty and volatility associated with farming where farmers do not know at the time of plantation what prices their produce would fetch. The interplay of demand and supply factors in determining market prices causes agriculture income to be volatile. Minimum support price (MSP) plays a crucial role here in defining a floor price at which government would procure crops from farmers. The Cabinet Committee on Economic Affairs (CCEA), Government of India, determines the MSP based on the recommendations of the Commission for Agricultural Cost and Prices (CACP). The objective of MSP is to protect farmers from the price shocks and to ensure food security through buffer stocks and Public Distribution System (PDS). MSP, however, is replete with problems. The 2016 Evaluation Report on Minimum Support Prices released by NITI Aayog underlined that the lack of procurement centers, closed storage facilities and delay in payments were some of the shortcomings of MSP. Lack of knowledge about MSP also contributed to farmers not being able to plan crop growing pattern ahead of sowing season and reap additional benefits from it. According to the report, despite its shortcomings farmers find MSP to be very useful and want it to continue as it provides a floor price for their produce and protects them against price fluctuations.

Despite having the capacity to pay, an individual may not have the intention or discipline to pay back the loans on a timely basis which is costly for banks. This is the behavioral aspect of credit risk and is reflected in his/her credit history. Recent delinquency, on-time payment history, leverage, default and non-default credit accounts are some metrics which give us insights into the behavior of the customer through

which we can evaluate whether he/she has the intention and required discipline to pay back the loans.

In agricultural loan market, loan waivers announced by government severely impact the behavior of the agriculture households and create a moral hazard problem where farmers default on loans in expectation of loan waivers in the future. Such loan waivers undoubtedly relieve distressed farmers of their credit burden, but it negatively impacts the credit culture. Such political risks associated with the agriculture sector make banks reluctant to lend to this sector. Post the 2008 comprehensive loan waiver scheme, a survey showed that one out of every 4 respondents wanted to wait for another loan waiver (Maurer 2014).

Instead of giving out generalized loan waivers to farmers, the need of the hour is to focus on creating robust mechanism to assess credit risk in the agriculture sector which can help banks increase their reach and help bring the farmers into the formal sector. In this paper, we highlight an approach that can make this possible. We show how using farm and household characteristics we can risk rank the agriculture households by assessing their "capacity to pay." Considering the difficulties faced by farmers and banks, our model would help in bridging the gap between them. By reducing the risk associated with farm lending, it would create a potentially profitable market for banks and would make cheaper credit available to the farmers along with reducing their dependency on moneylenders.

2 Literature Review

The economic survey of 2017–18 reveals that India's agricultural sector which employs more than 50% of the population contributes only 17–18% in its total output (Economic Division 2018). Therefore, enhancement of farm mechanization is important to mitigate hidden unemployment in the sector and free up useful labor. Agricultural credit plays a pivotal role in achieving technical innovation, and therefore measures need to be taken to expand the reach of low-cost formal credit to all farmers. Abhiman Das (2009) show that direct agricultural lending has a positive and significant impact on agricultural output whereas indirect credit has an affect after a lag of one year. Therefore, despite its shortcomings like less penetration to small and marginal farmers, and paucity of medium- to long-term lending, agricultural credit plays a critical role in supporting agricultural production and hence farm incomes and livelihood. In order to lend efficiently and minimize defaults on loans, it is imperative to have a sound analytical system in place to assess creditworthiness of borrowers. There have been several studies on credit scoring models for agricultural lending which use bank or credit history data as well as farm's and borrower's characteristics to assess debt repaying capacity of farmers. Identifying low-risk customers using credit risk assessment models is important not only for reducing cost for banks but also to increase the penetration of credit to small and marginal farmers who would have otherwise been left out due to misclassification as bad customers. Bandyopadhyay (2007), using sample data of a public sector bank,

developed a credit risk model for agricultural loan portfolio of the bank. With the help of bank's credit history and borrower's loan characteristics such as loan to value, interest cost on the loan, value of land and crops grown, he arrives at a logistic regression model that predicts the probability of default-defined as per the then NPA norm of the RBI, i.e., if the interest and/or installment of principals remains overdue for two harvest seasons but for a period not exceeding two and half years in the case of an advance granted for agricultural purpose. However, low sample size of the study serves as a major limitation of the model as it renders the model vulnerable to sample biases. Seda Durguner (2006), in their paper, showed that net worth does not play a significant role in predicting probability of default for livestock farms while it does matter significantly for crop farms. They develop separate model for crop and livestock farms in order to prevent misclassification errors that could arise by not differentiating between the farm types. Durguner (2007) showed using a panel data of 264 unique Illinois farmers for a five-year period, 2000–2004, that both debt-to-asset ratio and soil productivity are highly correlated with coverage ratio (cash inflow/cash outflow). Using a binomial logit regression model on 756 agricultural loan applications of French banks, Amelie Jouault (2006) show that leverage, profitability and liquidity at loan origination are good indicators of probability of default.

The studies mentioned above suffer from some severe limitations which need to be addressed for obtaining a robust credit risk model:

(1) No differentiation on geographical location and farm type: The ability of a farmer to repay depends on the income that he generates which is highly dependent on where he lives, rainfall pattern in that location, the soil type, crop grown, etc. Therefore, considering such agro-climatic factors is necessary in the model building process.

(2) Limited data sources: Bank's data would not be helpful for assessing risk of the farmers who are new to formal credit or if banks expand their direct lending to agriculture to new locations. Alternative methods to score farmers for their riskiness need to be identified as opposed to relying just on their past performance.

(3) Narrow focus of study: Credit risk from farm sector, as mentioned in the above section, can result from inability to pay that can be influenced by price risk and market risk or it could be due to indiscipline and fraudulent behavior which could result from political risk. Focusing only on the behavior of farmers on their credit account will not take into account a complete picture of the situation of the farmers, and most importantly, it would leave out those who are new to credit.

(4) Small sample size: Given the nature of diversity in India's agricultural landscape, a single bank's data cannot capture all the dimensionalities of the sector and a small sample size can lead to sample biases and cannot be applied universally.

3 Agriculture Credit in India: Trends and Current Scenario

Current mandate for Priority Sector Lending (PSL) by RBI requires all scheduled commercial banks and foreign banks to lend 18% of their Adjusted Net Bank Credit (ANBC) or Off Balance Sheet Exposure, whichever is higher, to the agriculture sector. Out of this, a sub-target requires them to lend 8% to the small and marginal farmers. As per the RBI guidelines, a small farmer is one who holds less than or equal to 1 ha of land whereas any farmer with more than 1 ha but up to 2.5 ha of land is considered to be a marginal farmer. These guidelines hold for all Scheduled Commercial Banks (SCBs) including foreign banks (RBI 2016).

Additional measures taken by the government to improve the farm credit situation include Kisan Credit Card (KCC) and Agricultural Debt Waiver and Debt Relief Scheme though their effectiveness can be debated and most of the experts consider them to be an unnecessary fiscal burden.

Despite taking the policy measures mentioned above, year-on-year growth of farm loans has gone down in past few years. After seeing a close to 40% growth rate in 2014, increase in farm credit went down to below 10% which is lowest since 2012 (Trends and Progress Report, RBI 2017–18).

As per the All India Rural Financial Inclusion Survey (NAFIS), 2016–17, by NABARD, 52.5% agricultural households had an outstanding debt at the time of the survey and out of these almost 40% households still went to non-institutional sources for their credit needs. Similar results are shown by the Situation Assessment Survey (SAS), 2013, by NSSO which shows a dependence of 44% households on non-institutional sources (please refer to Table 1). Even though two surveys have different samples, this indicates that the share of non-institutional sources has remained almost same from 2013 to 2016–17 and additionally corroborates the fact that growth in institutional credit has remained stagnant.

With flexible lending terms and often no collateral required, agricultural households continue to borrow from informal sources (moneylenders, friends and family).

Despite the exorbitant interest rates, which can go as high as 4 times the interest rates charged by the formal sources (refer to Table 2), moneylenders continue to cater to the credit needs of close to 11% of the farm borrowers (NAFIS 2016–17). This, including the reasons mentioned above, could be due to various factors including the availability of credit for personal reasons such as marriage. Another reason for this could be the unavailability of formal sources of credit. As per SAS 2013, for the agricultural households which owned less than 0.01 ha of land, only about 15% of

Table 1 Distribution of rural credit across institutional and non-institutional sources (in %)

Type of credit	1951	1961	1971	1981	1991	2002	2012
Institutional credit	7.2	14.8	29.2	61.2	64	57.1	56
Non-institutional credit	92.8	85.2	70.8	38.8	36	42.9	44

Table 2 Distribution of outstanding cash debt as per the rate of interest charged

Rate of interest (%)	Distribution of outstanding cash debt			
	Rural		Urban	
	Institutional	Non-institutional	Institutional	Non-institutional
Nil	0.8	18.3	0.4	27
<6	7.1	2.3	1.5	1.1
10-June	26	0.4	14.5	0.9
12-October	12.9	0.7	41.6	1.2
15-December	42.6	4.1	34.1	7.7
15–20	7.3	5.6	6.2	4.3
20–25	**2.1**	**33.9**	**1.2**	**27.3**
25–30	**0.1**	**0.6**	**0.2**	**0.3**
>30	**1**	**34.1**	**0.4**	**30.2**

loans were sourced from institutional lenders. On the other hand, this number was as high as 79% for farmers with more than 10 ha of land.

Most of this farm lending continues to be done by public sector banks. As of December 2016, private sector lent out 9.5% of the total loans whereas public sector lent out 85% of the total loans (Credit Bureau Database 2018). Private players, including foreign banks, have been reluctant to lend to farmers. For the year 2017–18, private and foreign banks met their PSL targets but did not meet their sub-targets of 8% lending to the small and marginal farmers (Trends and Progress Report, RBI 2017–18).

One major reason for this reluctance is the rise in bad loans coming from this sector. Between 2012 and 2017, bad loans in agriculture sector have jumped by 142.74% (Financial Express Online 2018). One reason behind this jump is the farm loan waiver announced by the central government. Subvention schemes, a subsidy provided by the government on interest rate, are another reason why private banks find PSL challenging. Banks are mandated to charge 7% interest on loans up to 3 lakhs. A further 3% subvention is provided in case of timely payments. So effectively these loans become available to farmers at 4% interest rate (PIB, DSM/SBS/KA, release ID 169414). This scheme has recently been made available to the private sector banks since 2013–14, prior to which it was only available to public sector banks.

Another reason for the meager farm lending by the private and foreign banks is the lack of understanding of the agriculture sector as a whole. This also leads to the inability to effectively assess risk in this sector. Without a proper understanding of the sector and the understanding of risk, operating in rural and semi-urban areas can be very expensive for banks. Entering a new market requires opening of new branches, launching market-specific products and huge operating costs. Due to all these reasons, these banks stay away from the agriculture sector or do marginal amount of lending in urban areas.

4   Research Methodology

Given the challenges faced by banks in lending to the agriculture sector, we propose in this paper a holistic approach to assess credit risk of farmers using alternative data and advanced analytical techniques. The focus of this study is the farmers who are not a part of the formal credit and still rely on non-institutional sources. These farmers would not have a credit footprint available to assess their riskiness, and therefore we focus on "capacity to pay" of the farmers rather than their default behavior on their credit accounts. In this study, we have used NSSO 70th round data—Key Indicators of Situation of Agricultural Households in India to identify complete characteristics of farmers in India. This is a comprehensive dataset of agricultural households in India which are defined as households receiving at least Rs. 3000 of value from agricultural activities (e.g., cultivation of field crops, horticultural crops, fodder crops, plantation, animal husbandry, poultry, fishery, piggery, beekeeping, vermiculture, sericulture, etc.) during last 365 days, and it encompasses all the factors that reflect the then situation of farmers. The survey was conducted in two visits. Visit 1 comprises data collected in the period January 2013 to July 2013 with information collected with reference to period July 2012 to December 2012, and Visit-2 comprises data collected between August 2013 and December 2013 with information with reference to period January 2013 to July 2013. This way it covers both kharif and rabi cropping seasons. However, for our modeling purpose we use only Visit-1 data as the information on outstanding loans is captured only in the Visit-1 Survey. The NSSO data captures variables such as the kind of dwelling unit of farmers, status of ownership of land, primary and subsidiary activity of farmers, whether the household has MGNREG job cards, no. of dependents and their employment status, the kind of crop grown on the farm, size of land under irrigation, the value of sale of crop, the agency the crops are sold through (dealers, mandi, cooperative agency and government), details of expenses in inputs and whether the farmer avails MSP or not. Such a detailed dataset of farmer characteristics is very helpful in assessing whether the farmer will be able to "afford" the loan or not. Following Seda Durguner (2006), we used debt to income as a proxy to judge creditworthiness. The mean debt to income in the population is 14, while the median is 1.5. Below table shows the distribution of debt to income in the data (Refer to Table 3).

We use debt-to-income ratio of 4 as the threshold; i.e., farmers whose ratio of outstanding debt is more than 4 times the income of one cropping season are classified as bad (farmers who would default), and logistic regression technique is used to predict the probability of default for these farmers. The overall bad rate of the population with the given threshold is 26%.

We build two models in our analysis. First, we use the variables that are captured by banks (Model 1) in their agricultural loan application form. For this purpose, we use standardized loan application form for agricultural credit devised by Indian Bank's Association (IBA). This form contains the required details that need to be collected from agri-credit loan applicants. This helps banks and customers maintain uniformity in the loan applications for agricultural needs. The second model (Model

Table 3 Distribution of debt-to-income ratio in our data

Quantile	Estimate
100% max	21,300
0.99	186.96
0.95	51.07
0.9	18.18
75% Q3	4.8
50% median	1.55
25% Q1	0.53
0.1	0.19
0.05	0.1
0.01	0.03
0% min	0

2) that we built considered both, the details already captured by the bank along with the additional features created from the NSS 70th round data. We use information value (IV), which tells how well my variable is able to distinguish between good and bad customers, to select important or predictive variables in the model. The variables whose IV was between 0.02 and 0.5 were then binned using weight of evidence (WOE). Variables with similar WOE were combined in a bin because they have similar distribution of events and nonevents. In this way, we transformed continuous independent variable to a set of groups/bins. We then built a logistic regression model to obtain probability of default using WOE of independent variables.

We find that Model 2 performs better than Model 1 in terms of Gini, KS and rank ordering. The results of the model are discussed in the next section.

5 Results

Our model gives a comprehensive set of variables which includes farmer's demographic features, agro-climatic factors and cropping patterns that describe his/her ability to pay. Variables like highest value crop grown and whether the farmer faced crop loss during the last one year capture the farming pattern for the farmer and explain how the recent trend of farming has been for the farmer. Whether the farmer has taken technical advice or not shows if farmer has access and willingness to incorporate new techniques in his farming. Our model covers both the endowment and behavior-related variables of the farmers.

The following tables give the resultant significant variables in both the models:

1. Model 1:

Analysis of maximum likelihood estimate		
Parameter	Sign of coefficient	Pr > ChiSq
Intercept	Positive	<0.0001
Primary income source	Negative	<0.0001
Percentage of land cultivated	Negative	<0.0001
Percentage of expense on machine hiring	Negative	<0.0001
Percentage of expense on fertilizers and chemicals	Negative	<0.0001
Percentage of expense on seeds	Negative	<0.0001
Number of male members in the family	Negative	<0.0001
Count of members between the age of 18 and 60 years	Negative	<0.0001
Age of the household head	Negative	<0.0001

2. Model 2:

Analysis of maximum likelihood estimate		
Parameter	Sign of coefficient	Pr > ChiSq
Intercept	Negative	<0.0001
Whether technical advice taken or not	Negative	<0.0001
Whether farmer suffered crop loss in the last season or not	Negative	<0.0001
Primary income source	Negative	<0.0001
Segment of the highest value crop grown by the farmer	Negative	<0.0001
Rainfall as a percentage of average rainfall in the district	Negative	<0.0001
Percentage of cultivated land	Negative	<0.0001
Percentage of expense on machine hiring	Negative	<0.0001
Percentage of expense on seeds	Negative	<0.0001
Number of male members in the family	Negative	<0.0001
Age of the household head	Negative	<0.0001

Capturing these additional variables in our model for assessing the risk of farmers gives us a lift of almost 7% in the Gini coefficient (from 41.7 to 48%); i.e., it improves the model accuracy from 70 to 75%. Bad rate distribution for our model goes from 58.87% in the lowest decile (highest risk decile) to 5.48% in the highest decile (lowest risk decile). On the other hand, using the variables already captured by banks, bad rate ranged from 55.27 to 8.02%. The below graph shows the risk ranking across deciles for both the models. We observe a break in the risk ranking of Model 1 at decile 7, whereas Model 2 holds perfectly across all the deciles. Refer to Appendix for detailed tables and to Appendix 2 for a note on Gini and KS Summary Statistic.
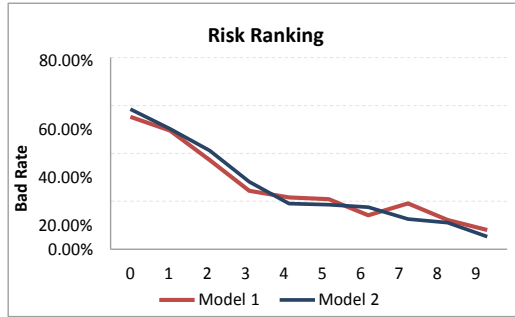
Model 2 provides a significant decrease in risk levels in comparison with Model 1 assuming that banks keep their approval rates constant across models. For example, if a bank decides to approve 19% of the credit applications received, it would face a 20% lower risk of default using Model 2 as compared to Model 1. This would allow banks to curb their bad rates and would be welfare generating for both the banks and the farmers.

3. Out-of-sample validation results: To assess the stability of our models across samples, we validated them on a randomly selected 30% sample of the development data. The result is given below:

Samples	Gini (%)	KS (%)
Model 1	41.51	32.76
Model 2	47.68	37.48

4. To check the applicability of Model 2 on different farmer segments based on their land holding, we validated the model on marginal, small and other farmers as defined by RBI. The model holds well in these segments in terms of rank ordering, Gini and KS, but some variables do not rank order in "Other Farmers" segment. The result is given below:

Samples	Gini (%)	KS (%)
Small farmers	47.15	36.27
Marginal farmers	47.24	37.30
Others	51.62	40.13

6 Policy Implications

A policy aspect that comes out from this analysis is that this model would allow the government to figure out the population they need to focus on for their policy measures. Farmers who get identified with a lower capability to pay using Model 2 become the target population for the government policies. Also, the model variables on which they did not do well define the areas where government needs to focus to bring those farmers to the formal financial sector. For example, if a large number of farmers in a district are identified to have a lower capability to repay due to having not taken technical advice or because of having suffered crop loss in the last farming season, this defines the focus area for the government to work upon. Here, they need to improve the availability of technical advice to the farmers and work on reasons of crop loss. Hence, above model would serve the dual purpose of helping both the banks and the government. Even though banks need to keep lending at the reduced interest rates as per the government policies, using this model they can identify the population with a lower risk of default. At the same time, government can form specific policies based on the needs of the farmers and help bring them to the formal credit market.

7 Limitations and Conclusion

Even though our model brings out results which can help both the banking sector and government, we do not claim that our model is free of any limitations or has no scope for improvement. Considering the type of data that has been used for this model, there is an inherent risk of endogeneity to occur in the analysis and it needs to be accounted for in the model building process. Also, the variables in Model 2 are not easily verifiable and it would require banks to invest in proper due diligence of their agricultural loan applicants.

On the basis of above information, it is understandable that farming sector needs special attention when it comes to credit facilities. Existing schemes and facilities have been unable to fulfill the credit needs of this sector. Generalized loan waivers announced time, again have put a financial burden on the economy and are not a solution in the long run. Our model shows that if banks capture specific information about farmer characteristics and consider agro-climatic conditions like rainfall in their lending decisions, they can reduce the delinquencies from this sector. In this way, agricultural lending can be made much more efficient and the level of financial inclusion of farmers can be improved.

Appendix 1

Results and approval strategy for Model 1:

Decile	Good	Bad	Total	Bad rate (%)	% population
0	16,971	20,974	37,945	55.27	9.48
1	20,799	20,403	41,202	49.52	10.30
2	26,315	15,600	41,915	37.22	10.47
3	25,543	8251	33,794	24.42	8.45
4	42,643	11,776	54,419	21.64	13.60
5	23,761	6272	30,033	20.88	7.51
6	32,621	5360	37,981	14.11	9.49
7	36,168	8574	44,742	19.16	11.18
8	32,855	4569	37,424	12.21	9.35
9	37,436	3262	40,698	8.02	10.17
			Overall bad rate		**26.25**
				Gini	**41.4**

Approval Strategy:

Decile	Good	Bad	% population	Default rate (%)	Approved population (%)
0	16,971	20,974	9.48	26	100.00
1	20,799	20,403	10.30	23	90.52
2	26,315	15,600	10.47	20	80.22
3	25,543	8251	8.45	17	69.75
4	42,643	11,776	13.60	16	61.30
5	23,761	6272	7.51	15	47.70
6	32,621	5360	9.49	14	40.20
7	36,168	8574	11.18	13	30.70
8	32,855	4569	9.35	10	19.52
9	37,436	3262	10.17	8	10.17
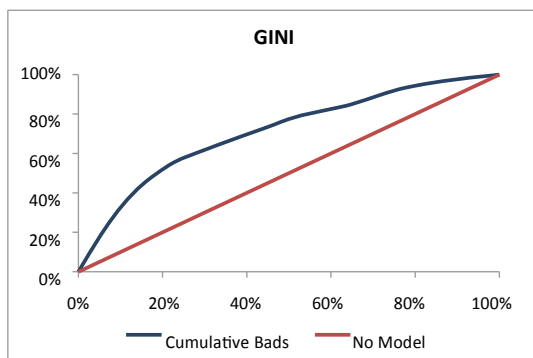
Results and approval strategy for Model 2:

Decile	Good	Bad	Total	Bad rate (%)	% population
0	16,596	23,349	39,945	58.45	9.98
1	20,542	20,808	41,350	50.32	10.33
2	21,034	14,688	35,722	41.12	8.93
3	34,241	13,398	47,639	28.12	11.91
4	28,493	6698	35,191	19.03	8.79
5	30,892	7029	37,921	18.54	9.48
6	35,153	7484	42,637	17.55	10.66
7	36,267	5212	41,479	12.57	10.37
8	34,434	4302	38,736	11.11	9.68
9	37,460	2073	39,533	5.24	9.88
			Overall bad rate		**26.25**
				Gini	**47.9**

Approval Strategy:

Decile	Good	Bad	% population	Default rate (%)	Approved population (%)
0	16,596	23,349	9.98	26	100.00
1	20,542	20,808	10.33	23	90.02
2	21,034	14,688	8.93	19	79.68
3	34,241	13,398	11.91	16	70.76
4	28,493	6698	8.79	14	58.85
5	30,892	7029	9.48	13	50.06
6	35,153	7484	10.66	12	40.58
7	36,267	5212	10.37	10	29.93
8	34,434	4302	9.68	8	19.56
9	37,460	2073	9.88	5	9.68
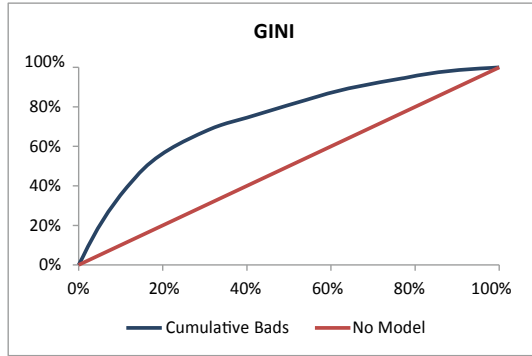
Appendix 2: Key Summary Statistic

Kolmogorov–Smirnov Test

The test was based on the following hypothesis:

$H0$: The Validation and Development Samples Have the Same Distribution.

$H1$: The Validation and Development Samples Do not Have the Same Distribution.

D_{max} is the maximum absolute difference between the score distributions for the two unweighted samples.

Applying the Kolmogorov–Smirnov test, we accept H0 if

$$D_{max} < D_{crit} \quad \text{where } D_{crit} = 100 * K \sqrt{\frac{1}{M} + \frac{1}{N}}$$

where

M is the total number in the development sample.

N is the total number in the validation sample.

K is the Kolmogorov–Smirnov statistic.

Gini Coefficient

The Gini coefficient is a measure of the power of a scorecard. The higher the Gini, the stronger the scorecard. A scorecard with no discrimination would have a Gini of zero; a perfect scorecard would have a Gini of 100%. A Gini is calculated by comparing the cumulative number of goods and bads by score. Graphically, it is the area between the two lines on the curve below (*XYZ*) expressed as a percentage of the

maximum possible (*XYW*). The two axes on the graph are cumulative percentage of goods (*y*-axis) and cumulative percentage of bads (*x*-axis). Graphical representation of the Gini coefficient:
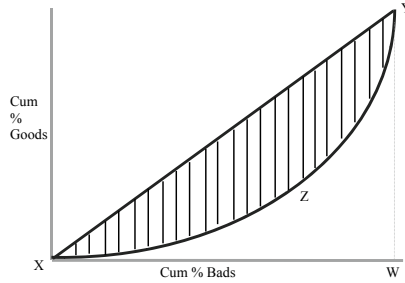
The Gini coefficient is calculated as follows:

Given that:

b_i Cumulative percentage of bads at a given score

g_i Cumulative percentage of goods at a given score

S_n Nth score in the score distribution.

The area under the curve (the unshaded area, not enclosed within Z) for a given score is defined as:

$$A_i = \frac{1}{2}(b_i - b_{i-1}) * (g_i + g_{i-1})$$

The total area not defined by the curve is equal to:

$$A_g = \sum_{i=S_1}^{S_n} A_i$$

And the area defined by the triangle *XYW* is equal to:

$$A_{X\hat{Y}W} = \frac{1}{2}(100 * 100) = 5000$$

The Gini coefficient is then calculated as:

$$g = \frac{\left(A_{X\hat{Y}W} - A_g\right)}{A_{X\hat{Y}W}}$$

References

Abhiman Das, M. S. (2009). *Impact of agricultural credit on agricultural production: An empirical analysis in India.* Reserve Bank of India Occasional Papers.

Amelie Jouault, A. M. (2006). *Determining the probability of default of agricultural loans in a French bank.*

Bandyopadhyay, D. A. (2007). *Credit risk models for managing bank's agricultural loan portfolio.* MPRA.

Dinesh, U. (2013, March 17). *Private banks reluctant about rural lending.* Retrieved from Livemint: https://www.livemint.com/Industry/TWPW6KvEmOksgvzfLElCdO/Private-banks-reluctant-about-rural-lending.html.

Durguner, S. (2007). A panel data analysis of the repayment capacity of farmers. In *2007 Annual Meeting*, July 29-August 1, 2007, Portland, Oregon, No 9361, American Agricultural Economics Association (New Name 2008: Agricultural and Applied Economics Association). https://EconPapers.repec.org/RePEc:ags:aaea07:9361.

Economic Division, D. o. (2018). Economic Survey 2017–18 Volume 1. Ministry of Finance, Government of India.

Experian India Pvt Ltd. (2018). Credit Bureau Database. Experian India Pvt Ltd.

Financial Express Online. (2018, January 10). Retrieved from Financial Express: https://www.financialexpress.com/industry/banking-finance/rbi-data-agriculture-bad-loans-jump-by-23-per cent-thanks-to-farmers-loan-waiver-demonetisation/1008842/.

Indiastat. (2008, 2017). Banks/sector wise NPAs of private sector banks in India. Indiastat.

Iyer, A. (2019, January 02). *Farm loan defaults rise as banks brace for big write-offs.* Retrieved from Livemint: www.livemint.com.

Jayakumar, T. (2018, May 03). *Time to do away with priority lending norms.* Retrieved from liveMint: https://www.livemint.com/Opinion/QwGGfDgozR6COUYFyW3PTP/Time-to-do-away-with-priority-lending-norms.html.

Lahiri, A., & Mookherjee, D. (2015, December 14). *Transforming Indian agriculture: The role of credit policy.* Retrieved from Ideas for India: https://www.ideasforindia.in/topics/agriculture/tra nsforming-indian-agriculture-the-role-of-credit-policy.html.

Maurer, K. (2014). Where is the risk? Is agricultural banking really more difficult than other sectors? In *Finance for food: Towards new agricultural and rural finance.*

NAFIS. (2016–17). All India Rural Financial Inclusion Survey (NAFIS) 2016–17, NABARD.

Niti Aayog. (2016). *Evaluation report on efficacy of minimum support prices on farmers.* New Delhi: Development Monitoring and Evaluation Office.

Peter Binswanger, H., & Khandkar, R. S. (1992, October). The impact of formal finance on the rural economy of India. *Journal of Development Studies.*

Pradhan, N. C. (2013, May). Persistence of informal credit in Rurak India: Evidence from 'All India Debt and investment survey'. *RBI WPS (DEPR).*

Press Information Bureau (PIB), DSM/SBS/KA, release ID 169414.

PTI. (2014, August 26). *Nabard's rural infrastructure development fund needs a relook, says RBI official.* Retrieved from Economictimes: https://economictimes.indiatimes.com/news/eco nomy/policy/nabards-rural-infrastructure-development-fund-needs-a-relook-says-rbi-official/ articleshow/40928706.cms.

RBI, F. S. (2016). *Financial stability report.*

Sayantan, B. (2015, March 19). *Surge in credit not benefiting small farmers.* Retrieved from Livemint: https://www.livemint.com/Politics/xbk7sf9N4gy0jjStXyLoiJ/Surge-in-credit-not-ben efiting-small-farmers.html.

Seda Durguner, P. J. (2006, July). *Credit scoring models: A comparison between crop and livestock farms.* Long Beach, California, United States of America.

Standardized common loan application form for agricultural credit, Indian bank's Association (IBA), No. SB/Cir/AGRI/480

Trends and Progress Report, RBI. (2017–18). *Trends and progress report.* RBI.