IJHSR
ijhighschoolresearch.org
■ RESEARCH ARTICLE

Determinants of Loan Repayment: Is Caste a Factor? Evidence from Micro Business Lending in India

Pranav C. Madhukar

Sri Kumaran Children's Home, Mallasandra, Bangalore, Karnataka, 560062, India ; pranav.madhukar@gmail.com

ABSTRACT: There are over 63 million MSMEs (micro, small, and medium enterprises) in India, and most of them depend on loans to sustain and grow. This paper identifies the key determinants of timely repayment of loans by MSMEs by examining data obtained from a Non-Banking Financial Company (NBFC) based in South India focused on lending to micro businesses. I find that backward caste-owned businesses have repaid their loans slightly better than general caste-owned businesses. Furthermore, default rates increased as a consequence of COVID-19. Even in the Post Covid analysis, backward caste-owned businesses have repaid their loans better. Even while estimating the loan outcomes in relative terms, I find that the effect continues to be of economic significance. A positive selection of lower caste borrowers cannot explain these results. This finding is important. While approval of loans might still suffer from biases against backward castes, this paper shows that purely based on on-time repayment track record; lenders would do well to lend to backward caste-owned businesses.

KEYWORDS: Behavioral and Social Sciences; Other; Small business; Small business lending; SME loans and caste; Micro business loans in India; Loan Default in SME businesses; Caste discrimination and loan default.

Introduction

There are over 63 million MSMEs (micro, small, and medium enterprises) in India[1], which cumulatively account for over 30% of the country's GDP[2] and employ more than 110 million people. According to the definition laid out by the Government of India, a micro business is characterized by "investment in plant and machinery or equipment of not more than INR 10 million (~125,000 USD) and annual sales turnover of not more than INR 50 million (~630,000 USD). The vast majority of these businesses (~96%) are proprietary concerns (owned and operated by the same individual), unregistered, do not have high-quality documentation or record keeping, do not file taxes, and for the most part, do not have any traditional credit scores/ formal credit risk metrics associated with them. A 2019 Asian Development Bank Institute study[3] reveals that in India, nearly 80% of people don't have formal credit scores and that such lack of credit scoring for individuals often extends to small businesses that individual proprietors set up. Setting up these small businesses most often happens with family savings or borrowing from friends and family. Even when these businesses need access to capital for operations, working capital, etc., their first resort is to reach out to their community networks.[4] In an important essay, The Economic Lives of the Poor, Nobel laureates Banerjee and Duflo point out that "the data from our 13 countries suggests that the fraction of rural, extremely poor households having an outstanding debt varies between countries, from 11 percent in rural East Timor to 93 percent in Pakistan. But across the surveys, very few poor households get loans from a formal lending source."[5]

For poor households who set up micro-businesses, accessing formal loans is often a last resort. Therefore, this creates a cycle where many of them have no credit history that can be tracked by a credit rating agency and therefore find it difficult to receive credit from formal sources. In 2016, the credit rating firm Moody's published a whitepaper[6] titled, 'Seven Key Challenges in Assessing SME Credit Risk,' reiterating similar issues.

Researchers have studied the importance of 'soft data' in informationally opaque small businesses – such as the information about the character and reliability of the firm's owner – that may be difficult to quantify, verify, and transmit through the layers of management and ownership of a banking organization. "The transactions-based technologies of financial statement lending, asset-based lending, and credit scoring are based primarily on quantitative financial ratios, collateral ratios, and credit scores, respectively. Relationship lending, in contrast, allows informationally opaque small businesses without strong financial ratios, collateral, or credit scores to obtain bank financing by augmenting relatively weak hard information with soft information gained through contact over time."[7]

All of this complicates lending decisions, often requiring loan officers to use a wide variety of proxies to determine creditworthiness. Lending institutions also, therefore, depend on collateral as a way of ensuring that they exert substantial pressure on the borrowers to repay loans on time.

With about 110 million people employed in the MSME sector, making credit available to this sector is an important priority for policymakers. Economists have pointed out that the provision of microloans to existing businesses results in economically meaningful, positive effects on household businesses and consumption.[8]

Other social factors only compound the difficulty lending to micro businesses in India. As Table 1 shows, more than two-

DOI: 10.36838/v5i3.16

thirds of MSME businesses in India are owned by individuals belonging to the so-called 'backward' castes (Scheduled caste – SC, Scheduled tribe – ST, Other backward caste – OBC).

Table 1: Percentage of MSME businesses owned by social category.

	Social categories				
Sector	**SC**	**ST**	**OBC**	**Others**	**Not Known**
Micro	12.48	4.11	49.83	32.79	0.79
Small	5.5	1.65	29.64	62.82	0.39
Medium	0	1.09	23.85	70.8	4.27
All	12.45	4.1	49.72	32.95	0.79

It is well-documented that caste-based discrimination is widespread in India.[9] This age-old categorization is still prevalent in modern India.[10] This discrimination has often extended to the provision of formal credit. There is a growing body of literature that discusses such caste-based discrimination with regard to access to credit.[11] Existing literature has examined the effects of 'taste-based' discrimination, in which discrimination occurs purely based on race, caste, or such attribute, and also 'statistical discrimination' accounting for creditworthiness or other objective criteria.[12]

A relatively underexplored aspect of lending to micro businesses is the determinants of on-time repayment. In addition to the financial metrics, the micro business provides, lenders collect several other parameters, such as demographic data, education, marital status, type of business, the purpose of loans, etc., as part of their credit assessment methods. They often combine this with other qualitative factors, which they gather through informal inquiries about the borrower's social standing and ability to pay.

Some useful studies attempt to associate various characteristics of borrowers with the ability to repay loans. Empirical research from several parts of Africa provides interesting insights into this. Education came up as a significant determinant of on-time repayment of loans in a study conducted in Ethiopia.[13] The experience of the entrepreneur running the business also came up as an important determinant for loan repayment.[14] Using behavioral science experiments to nudge loan repayments in the Philippines, researchers found that text message reminders are useful only when they include the account officer's name and only for clients previously serviced by the account officer.[15]

While the context of microfinance in Africa provides valuable insights into determinants of loan repayments, the different social dynamics in India, where caste is an important factor, warrant a special look.

Most of the literature on caste and loans in India focuses on access to credit. An analysis of the 2006-2007 Fourth All India Census for Micro, Small, and Medium Enterprises finds that owners from socially disadvantaged groups have a lesser likelihood of receiving loan approval upon controlling for other factors (such as creditworthiness.)[16] These results are supported by a significant body of analytical all of which indicate lower access to credit for borrowers from lower castes.[17]

One important addition to this body of research in the Indian context is identifying differences in loan repayment patterns by caste. In this paper, I analyze the repayment patterns of loans to micro businesses owned by general and backward caste borrowers to assess whether any difference in default and on-time repayment rates is attributable to borrowers' caste. The results of this paper indicate that the government policies which stimulate lending to lower caste communities could do well both for businesses and society. Furthermore, it challenges the stereotypes in lending institutions that preferentially lend to upper-caste borrowers. The paper shows that lower caste borrowers repay well yet do not get equal credit. This creates a space for policy and perspective shifts.

But before I go any further on this exploration, I would be remiss if I did not mention another very different approach to lending that might somewhat obviate the necessity of collecting the data that traditional lending has warranted, including data on caste. The emergence of fintech lending has been an important feature for lending to individuals and small businesses. As the latest research points out, most SMEs are limited in their availability of collateral, and tighter regulation and higher capital requirements make long-term unsecured loans to SMEs unattractive to banks.[18] Fintech lending provides firms with an alternative to access long-term unsecured financing. A recent academic paper shows that finance companies and fintech lenders increased lending to small businesses after the 2008 financial crisis. And by 2016, this increase had almost entirely offset the decrease in bank lending in the United States.[19]

Hypothesis

Differences in loan repayment and default rates between lower and general caste individuals can be due to many factors.

Positive selection within this sample could be why lower caste individuals might have better loan outcomes. The data could show this outcome because loans have been given very selectively to lower-caste borrowers. In Table 8, I consider this possibility.

A possible explanation for the opposite observation – indicating that lower caste individuals have worse loan outcomes – could be due to exposure to business. People of general castes have, on average more resources, education, and exposure due to centuries of historical discrimination against lower-caste individuals. These metrics are very subjective and varied and cannot be tested easily. However, a variable Education has been controlled for in the analysis.

Another hypothesis could be that the loan's interest rate causing this difference in repayment. Biases may play into the mindsets of loan officers resulting in differential interest rates for similar general and lower caste borrowers. This could affect both repayment and default rates. However, the data for interest rates is not accessible in the dataset and hence has not been analyzed.

All the above hypotheses are in the context that borrowers of lower caste are less able to access credit. There are two possible explanations for this. Taste-based discrimination (i.e., lenders have personal biases against lower castes) has been proven to exist in existing literature (as discussed in the section above). The second explanation is belief-based discrimination (i.e., lenders believe that lower caste borrowers do not repay well).

DOI: 10.36838/v5i3.16

This paper deals with the second explanation and demonstrates that it is unjustifiably held in the context analyzed.

This analyzed dataset is unique because I can test the entire picture of the discrimination hypotheses: both the quality of the applicant (basic economic fields, loan outcomes) and the characteristics (caste) of the applicant.

Methods

The dataset used in this paper was acquired from a leading non-banking financial company (NBFC) based in South India that lends to thousands of micro-business owners in several states. While micro-businesses predominantly borrow from informal sources, there is a growing push by the government to encourage traditional banks and NBFCs to lend to micro-businesses. In India, an NBFC is a company that provides specific banking services to people without a banking license and has separate regulations governing such entities. NBFCs typically rely on various credit assessment mechanisms, including sector analysis of loan applicants (i.e., dairy, service, automobiles, etc.), income checks through non-traditional channels, background checks through references, and collateral.

I obtained data for 1050 loan records from 2016 – 2021 from one NBFC based in South India. All the loan records are unique – the dataset does not include any repeat borrowers. The dataset comprises 350 records of defaulted loans, 350 records of delayed repayments, and 350 records of timely repayments. The dataset also includes over two dozen additional parameters ranging from education level and caste to type of business, the purpose of the loan, EMI, total business liabilities, etc. These additional parameters fall into both categories of categorical variables as well as continuous variables.

Table 2: Year-wise distribution of raw data.

	Counts		
Disbursement Year	**On-time**	**Delay**	**Default**
2016	155	185	38
2017	110	97	67
2018	54	41	82
2019	31	26	137
2020	0	1	22
2021	0	0	4
	350	350	350

The data I have obtained is a combination of (a) the information that loan applicants fill out as part of their application form and (b) the data that the NBFC collects and compiles as part of its loan approval process and tracking the performance of the loan.

This data has been randomly selected by the NBFC and shared with us. I do not have information about how the NBFC received many applications in each category and what percentage of applications were approved. This data is a sample from one NBFC, and therefore the analysis that this paper contains is relevant to this sample. It is unclear whether and how the results in the paper can be extrapolated to other loans or regions in India.

The following is a detailed discussion of some important parameters which I use in conducting our analysis:

Caste:

The NBFC collates caste data into six categories. The dataset provided contains six categories of caste – general, OBC, minority, SC, ST, and others. Considering that the bulk of the observations was distributed between General and OBC, I have reclassified the Caste variable into a simpler categorical variable with two modalities: general and non-general (backward) castes. Due to the small sample of records in SC, ST, and Minority categories, they have been grouped with OBC under the 'nongeneral' category for the data analysis.

Education:

The dataset originally includes five categories of education (signifying the highest degree received) – school, high school, graduation, professional degree, post-graduation, and doctorate. The education variable was notably concentrated at the lower end of the distribution (with 810 recipients of 1050 in total having just attained 'school' level). Therefore, I recategorize the variable into two modalities: school and more than school.

Type of Business:

The type of business is a critical factor to control for while studying the relevance of various predictors on loan repayment. The dataset classifies businesses broadly into eight categories: service, power/ auto/ handloom, trading (others), hotel, dairy, trading (essentials), manufacturing, and home-based enterprise. I leave this classification untouched since the distribution is, for the most part, uniform across the dataset.

Loan Purpose:

The purpose of borrowing by the micro business can inherently have differing success rates, directly impacting the ability to repay the loan. The dataset from the NBFC defined twelve categories of loan purposes; however, over 90% of the loans were categorized as one of two loan purpose categories. The remaining minority of loans were spread between ten loan purpose categories. Therefore, I opt to simplify the variable by giving it only three modalities: business asset creation, working capital, and others.

Impact of COVID:

COVID was clearly an unprecedented global calamity that upended businesses across the world. Therefore, this paper organizes the dataset into pre-Covid and post-Covid data. For ease of making this distinction between these two categories, I assumed that all loans disbursed in any year were made in the month of April of that year. Then I added the tenure in months to each loan. For any loan that closed before April 2020, I have categorized it as a pre-Covid loan. Loans that were closed after April 2020 were classified as Post-Covid loans.

This categorization is important throughout this paper since I will test for determinants of on-time repayment on both data sets. To the extent that there are differences in the determinants of on-time repayment in these two scenarios, the paper will further investigate the factors that explain these differences.

Loan Repayment:

The data has three categories of loan repayment: on-time repayment, delayed repayment, and default. For clarity of analysis, I group these variables into two further categories

–- Default_v1 and Default_v2. Default_v1 is a test for timely repayment of loans –- in this variable; I group delayed repayment with defaults as a single category for not on-time repayment. Default_v2 is a test for default of loans –- I group on-time and delayed repayment together to signify loans that do not default. I perform all our regressions on both of these categories and list both of the results below.

Financial Metrics:

The dataset includes various financial parameters such as monthly sales, monthly net income, total liabilities, loan amount, and EMI (Equated Monthly Installments). These parameters are in the form of continuous numerical variables and are key considerations in determining a company's ability to repay loans. It is also possible to create secondary variables as a combination of these primary variables. For instance, in our analysis, I use EMI/monthly sales as an indicator in determining the capacity to repay a loan.

Method

One of the main objectives of this paper is to identify the characteristics of borrowers that are most indicative of timely repayment. As a starting point, I consider the differences in mean outcomes of on-time repayment, delayed repayment, and default samples across various predictors. The predictor variables that I analyze include Caste, Purpose of Loan, Type of Business, Level of Education, PostCovid, Net Income, and Disbursement amount. However, some of the differences from simple comparisons could be due to factors associated with the predictors –- omitted variables. To address the possibility that caste is correlated with factors that affect loan repayment, I use ordinary least squares:

$$Yi = \beta 0 + \beta 1 castei + \beta 3 PostCovidi + \beta 2 xi + \varepsilon i$$

This regression controls for various factors that could be considered omitted variables. I control for Covid and caste in all regressions I perform. The control variables are loan purpose, type of business, level of education, monthly sales / EMI, and an interaction term. Covid has been accounted for in all regressions because the random effects of Covid could significantly alter the results observed due to fixed characteristics such as Caste. Hence, a Covid*Caste interaction term has been included in 4 of the regressions performed. Additionally, caste was found to be a significant predictor of loan repayment, and hence it has been accounted for in all the regressions.

With this framework, I run two sets of regressions for measuring loan outcomes (Yi): First, with timely repayment as the outcome and then with default as the outcome variable. In each case, I run a set of 5 multivariate OLS regressions, including 2 with difference-in-differences.

Results and Discussion

Descriptive Analysis:

The descriptive analysis of the full dataset pointed to some useful ways of thinking about the analysis in this paper. I list some key tables from the descriptive analysis and explain some relevant context.

Table 3: Default rates of general and backward caste borrowers.

Caste	Counts by Category	Fully repaid on time	Fully repaid after delays	Default in repayment
OBC	345	38%	37%	25%
GENERAL	597	31%	31%	38%
MN	72	29%	42%	29%
SC	21	24%	33%	43%
OTH	10	60%	20%	20%
ST	5	40%	20%	40%

From Table 3, just by comparing the default rates of OBC and general categories (categories where a majority of the counts are concentrated), I find some interesting results –- OBC borrowers seem to be defaulting at lower rates when compared to borrowers from the general category.

Table 4: Caste and Purpose of Loan.

Caste	Purpose of Loan	Counts	Fully repaid on time	Fully repaid after delays	Default in repayment
OBC	Business Asset Creation	196	38%	41%	21%
OBC	Working Capital	126	37%	30%	33%
GENERAL	Business Asset Creation	189	35%	38%	27%
GENERAL	Working Capital	334	28%	25%	47%

To further analyze the caste factor, I grouped it with the purpose of the loan. It is clear from the counts column of Table 4 that General category borrowers have received more working capital loans than OBC borrowers. Based on the default column, I can also infer that, on average, working capital loans seem to default at higher rates than business asset creation loans.

Table 5: Business Type and Purpose of Loan.

Business Type	Purpose of Loan	Total Count	% Fully repaid on time	% Fully repaid after delays	% Default in repayment
Service	Business Asset Creation	17	41%	29%	29%
Service	Working Capital	73	27%	21%	52%
Power/Auto/Handloom	Business Asset Creation	69	35%	29%	36%
Power/Auto/Handloom	Working Capital	39	10%	15%	74%
Trading-Others	Business Asset Creation	75	33%	37%	29%
Trading-Others	Working Capital	156	27%	29%	44%
Hotel	Business Asset Creation	12	17%	58%	25%
Hotel	Working Capital	40	45%	23%	33%
Dairy	Business Asset Creation	132	40%	48%	12%
Dairy	Working Capital	58	40%	36%	24%
Trading-Essentials	Business Asset Creation	63	41%	32%	27%
Trading-Essentials	Working Capital	94	31%	36%	33%
Manufacturing	Business Asset Creation	30	33%	37%	30%
Manufacturing	Working Capital	62	29%	31%	40%
Home Based Enterprise	Business Asset Creation	11	36%	36%	27%
Home Based Enterprise	Working Capital	15	47%	20%	33%

I then explored whether the difference in default rates between business asset creation and working capital loans depends on the type of business. In the total counts column of Table 5, I find that certain business types were more likely to take loans for a particular reason. For instance, in Table 5, Dairy businesses have a higher proportion of business asset creation loans, whereas Service businesses have a greater proportion of working capital loans. Another important observation is that for each business type, the default rates for business asset creation loans are lower than for working capital loans, once again reaffirming the results from Table 4.

The results of this descriptive analysis indicate that Caste, Purpose of Loan, and Type of Business may be significant predictors of loan defaults. In the following subsection, I analyze these variables along with additional parameters to identify the relevant predictors.

Timely Repayment vs. Untimely Repayment:

The regression results yield several interesting insights into the relevance of predictors in determining the timeliness of loan repayment. The difference-in-differences estimate between general and non-

DOI: 10.36838/v5i3.16

Table 6: Regression results of timely repayment.

Dependent Variable = Default V1	On-Time Repayment	On-Time Repayment	On-Time Repayment	On-Time Repayment	On-Time Repayment
Constraints	DID (1)	DID (2)	Pre Covid (3)	Pre Covid (4)	Cumulative (5)
Constant	0.510 ***	0.523 ***	0.375 ***	0.747 ***	0.474 ***
	(0.037)	(0.047)	(0.078)	(0.041)	(0.042)
Caste - General	-0.080	-0.082	-0.083	0.068**	0.025
	(0.054)	0.054	0.061	(0.033)	(0.029)
Education - School+		0.001	0.015	-0.053	0.000
		(0.034)	(0.081)	(0.037)	(0.034)
Loan Purpose		0.005	0.041	-0.009	0.001
		(0.023)	(0.045)	(0.027)	(0.023)
Type of Business		-0.004	0.026	-0.013	-0.004
		(0.007)	(0.017)	(0.008)	(0.007)
Ratio		0.003	0.073	0.002	0.003
(=EMI / Monthly Sales)		(0.005)	(0.045)	(0.005)	(0.005)
Covid - Post Covid	0.192 ***	0.187 ***			0.263 ***
	(0.045)	(0.046)			(0.032)
Interaction	0.145 **	0.149 *			
(=Caste* PostCovid)	(0.063)	(0.064)			
R-squared	0.070	0.071	0.026	0.01	0.066
Adjusted R-squared	0.067	0.065	0.009	0.003	0.061
No. of Observations	1050	1040	290	758	1048
Note: Standard Errors are reported in parentheses. *, **, and *** indicate significance at the 90%, 95% and 99% levels respectively.					

general castes pre- and post-Covid is 0.145 (the coefficient on the interaction term), which is significant at the 5% level, as noted in Column (1). Even after making the regression conditional on other factors, the interaction coefficient changes very little, as indicated in Column (2). This suggests that untimely repayment of general caste borrowers rose by 14.5% more than non-general caste borrowers when comparing pre-Covid and post-Covid data. This result indicates a differential impact of Covid on timely repayment rates amongst general caste borrowers.

This led us to try and understand which parameters are most relevant in determining the timeliness of repayment in a pre-Covid setting and then in a post-Covid setting. The results from the multivariate linear regression in Column (3) surprisingly indicate that none of these five variables are statistically significant at even the 10% level. Moreover, I see that these five variables cumulatively only explain a mere 2.6% (referring to the R-squared value) of the variation in timely loan repayment. However, what is interesting to note, despite its insignificance, is that the coefficient on caste has a negative sign, suggesting that borrowers belonging to general castes were more likely to repay on time.

Now I repeat the same multivariate linear regression for a post-Covid world. As tabulated in column (4), the regression results indicate that caste is the only statistically significant (at a 5% level) predictor of the five variables. The coefficient on the caste variable implies that those belonging to general castes were 6.8% more likely to not repay on time (or 6.8% less likely to repay on time). The results from the first column (which suggests Covid was a significant event) motivated us to run a multivariate regression of timely repayment on six variables (five from earlier plus Covid) cumulatively on data pre- and post-Covid (as defined in the Data section). The results of the regression land not too far off from the previous ones in terms of coefficients. The only variable which ends up being significant here is Post Covid (at the 1% level). This can be interpreted as borrowers were 26.3% more likely to not repay on time in a post-Covid world. One last observation: although insignificant, the coefficient on the caste variable is positive.

5.2.3 Repayment vs. Default:

Table 7: Regression results of default.

Dependent Variable = Default V2	Loan Default	Loan Default	Loan Default	Loan Default	Loan Default
Constraints	DID (1)	DID (2)	Pre Covid (3)	Post Covid (4)	Cumulative (5)
Constant	0.019	0.084	0.017	0.45 ***	0.044
	(0.034)	(0.043)	(0.016)	(0.046)	(0.039)
Caste - General	-0.019	-0.032	-0.016	0.090 **	0.058 **
	(0.050)	(0.050)	(0.012)	(0.037)	(0.027)
Education - School+		0.039	0.020	0.045	0.039
		(0.032)	(0.016)	(0.041)	(0.032)
Loan Purpose		0.020	-0.012	0.035	0.016
		(0.021)	(0.009)	(0.030)	(0.021)
Type of Business		-0.022 **	0.001	-0.028 ***	-0.022 **
		(0.007)	(0.004)	(0.009)	(0.007)
Ratio		0.004	-0.003	0.004	0.004
(= EMI / Monthly Sales)		(0.005)	(0.009)	(0.005)	0.005
Covid - Post Covid	0.377 ***	0.357 ***			0.421 ***
	(0.042)	(0.043)			(0.039)
Interaction	0.119 **	0.126 *			
(=Caste*PostCovid)	(0.059)	(0.059)			
R - Squared	0.187	0.199	0.019	0.027	0.195
Adjusted R-squared	0.185	0.193	0.002	0.021	0.191
No. of Observations	1050	1040	290	758	1048
Note: Standard Errors are reported in parentheses. *, **, and *** indicate the significance at the 90%, 95%, and 99% levels respectively.					

These regression results yield new insights into the relevance of various parameters in determining loan default. Column (1) captures a difference-in-difference regression that allows us to estimate the mean effect of caste as a predictor of default in a pre-Covid and post-Covid world. The difference-in-difference estimate is 0.119 (the coefficient on the interaction term), which is significant at the 5% level. This suggests that Covid altered the likelihood of default by The results from column (3) of Table 7 closely match those from column (3) in the previous table (Table 6) in that here, too, I conclude that none of the predictors are significant even at the 10% level. Like in the previous table, the coefficient on caste is also statistically insignificant, but the coefficient value suggests that borrowers belonging to general castes were 1.6% less likely to default in the pre-Covid sample.

Column (4) is the same multivariate linear regression, but for a post-Covid world. The regression results indicate that caste (at the 5% level) and type of business (at the 1% level) are the only statistically significant variables. Notably, the five variables account for a tiny percentage (2.7%, see R-squared) of variation in loan defaults. The coefficient on the caste variable implies that those belonging to general castes were 9% more likely to default. The coefficient on the type of business is less insightful as it implies that moving one unit up the encoded list of business categories decreases the likelihood of default by an average of 2.8%.

In column (5), I run a cumulative regression of loan defaults on six variables (five of the same variables as in the previous two columns plus a Covid variable). The caste (5% level), type of business (5% level), and PostCovid (1% level) variables end up being significant here. This can be interpreted as follows: borrowers belonging to general castes were 5.8% more likely to default, and borrowers were 42.1% more likely to default in a post-Covid world. Now the R-squared value stands out because it jumps up to a staggering 0.195, meaning these six variables now cumulatively explain 19.5% of the variation in loan defaults (much of which is being presented by the inclusion of Post-Covid as a variable).

The standout observations from all these results for us are threefold:

(1) People of backward castes have better loan outcomes in the dataset. To estimate the economic significance of the results, I calculate the relative effect of caste on default rates. From Table 7, column 4, we see that general caste defaulted on 54% of loans, while the default rate of the non-general caste is 45% in absolute terms. When I estimate the effect of caste in relative terms, I find that the default rate drops by 16.7% [= (45 - 54) * 100 / 54] when moving from the general caste to the non-general caste. This is a huge number, emphasizing the strength of the findings. The incredibly prevalent preference for general caste borrowers in lending is unjustified.

(2) Covid was a major event that impacted the timeliness of loan repayment. The adverse effects of COVID on loan repayment also seemed to be greater on general caste borrowers.

(3) The regressions indicate that the variables, especially Post Covid, could better explain loan defaults than timely repayment.

(4) In the Pre Covid sample, on both metrics of loan repayment, we find that the coefficient for the caste variable (column 3) is negative, indicating that general caste borrowers repay better. However, there are two important things to note here. First, both the corresponding results are statistically insignificant, perhaps due to a lack of power. Second, the descriptive analysis (Table 4) shows that the general caste individuals seem to be taking a significantly larger number of working capital loans than those of lower caste individuals. Furthermore, certain industries (Table 5), such as services that take more working capital loans, seem to show a greater propensity to default. Perhaps the type of business could be the reason for the results obtained, as the Covid shock has hit different business types with differing intensities. Although the type of business has been controlled for in the multivariate regressions, that may be insufficient to draw a satisfactory conclusion because of the multiple categories, which prevents obtaining a satisfactory grouping of the type of business.

Table 8: Regression for positive selection.

Dependent Variable = Caste	**Caste**	**Caste**
Regression	Univariate OLS	Univariate OLS
Constant	0.564 ***	0.571 ***
	(0.016)	(0.015)
Ratio		-0.0067
(= EMI / Monthly Sales		(0.0053)
Total Liabilities	0.000001	
	(0.00000007)	
R-squared	0.0022	0.0015
Adjusted R-squared	0.0012	0.0006
No. of Observations	1039	1046
Note: Standard Errors are reported in parentheses. *, **, and *** indicate significance at the 90%, 95%, and 99% levels respectively		

Based on the data given to us by the NBFC, it is not possible to conclusively test whether the better repayment of non-general caste individuals is due to positive selection by the NBFC. However, I regress Caste against basic economic factors to identify if there is a significant difference in economic situations of non-general and general categories. Columns (1) and (2) contain OLS regressions of caste against the total liabilities and ratio EMI / monthly sales. The coefficients shown in both columns in Table 8 are very small, almost negligible, indicating that there does not seem to be any notable difference in the economic situation of the castes.

There is little evidence to show that this NBFC preferentially lends to certain caste groups. However, it is important to note that I do not have data about loan officers' on-ground interaction and decision-making calculus. As a result, I cannot conclusively rule out the possibility of some positive selection.

Conclusion

Discrimination against backward castes in lending is widely prevalent. Numerous studies show that backward caste borrowers face significant disadvantages in accessing credit. Even when they do get loans, it is often from exploitative informal sources. This has enormous implications for the millions of people who require micro-business loans to maintain their livelihoods. In my paper, I show that borrowers from backward castes repay loans at similar rates as general caste people, if not marginally better. This result is important in combating discriminatory access to formal lending for backward caste-owned micro businesses.

This finding that backward caste-owned micro businesses repay better than general caste borrowers could have policy implications. The government could formulate policies that incentivize lenders to give more loans to micro businesses owned by backward caste groups. One example of such a policy could be in the form of a First Loss Guarantee program for borrowers from backward castes. Such a policy measure could provide a strong signal to lenders to treat backward caste borrowers as creditworthy.

The continuation of this research could be to separately consider the effects of the caste of the loan officer from a formal lender and find out if there is any preferential treatment in making the loan to borrowers belonging to the same caste as the loan officer and then compare repayment outcomes. Additionally, using a different data source with more fields for analyzing positive selection to reconfirm the results of this paper might be useful. Furthermore, conducting this analysis with the additional controls of (i) Loan interest rate and (ii) Tenure of loan could improve the strength of the results. Lastly, as digital lending through smartphone-based apps gains momentum globally, it would be useful to explore how purely data-based transaction lending, agnostic of caste or other social factors, impacts repayment rates.

Acknowledgements

The author is grateful to Mr. Brahmanand Hegde, co-founder of Vistaar Finance, who provided access to the dataset for this study.

DOI: 10.36838/v5i3.16

References

1. Ministry of MSME, Annual Report 2020-21, p. 24: and p.26 https://msme.gov.in/sites/default/files/MSME-ANNUAL-REPORT-ENGLISH%202020-21.pdf
2. https://www.pib.gov.in/PressReleasePage.aspx?PRID=1744032
3. Nemoto, Naoko and Yoshino, Naoyuki, Fintech for Asian SMEs (2022, Edited volume), Asian Development Bank Institute
4. Takashino, Nina & Maharjan, Keshav & Fukui, Seiichi. (2014). Impacts of Caste, Risk, and Time Preference on Borrowing Behaviour: A Case Study in West Bengal, India. 10.1007/978-81-322-12 84-3_3.
5. Banerjee, Abhijit V, and Esther Duflo. "The Economic Lives of the Poor." The Journal of Economic Perspectives: *A Journal of the American Economic Association* vol. 21,1 (2007): 141-167.
6. Moody's Analytics, Helene Page, 2016
7. Berger, Allen N., and Gregory F. Udell. "Small Business Credit Availability and Relationship Lending: The Importance of Bank Organisational Structure." The Economic Journal 112, no. 477 (2002): F32–53.
8. Banerjee, Abhijit & Breza, Emily & Duflo, Esther & Kinnan, Cynthia 2019. "Can Microfinance Unlock a Poverty Trap for Some Entrepreneurs?" Discussion Papers Series, Department of Economics, Tufts University
9. Beteille, Andre. "Varna and Jati." Sociological Bulletin 45, no. 1 (March 1996): 15–27.
10. Desai S, Dubey A. Caste in 21st Century India: Competing Narratives. Econ Polit Wkly. 2012 Mar 12;46(11):40-49. PMID: 2273 6803; PMCID: PMC3379882.
11. Sangwan, Navjot 3000 Years of Discrimination and Counting: How Caste Still Matters in the Indian Credit Sector. Greenwich Papers in Political Economy. GPERC77, Institute of Political Economy, Governance, Finance and Accountability, University of Greenwich.
12. Kumar, S. Mitra (2016) Why does caste still influence access to agricultural credit? WIDER Working Paper 2016/86. Helsinki: UNU-WIDER.
13. Kebede, Tadele Tafese, M. 2016 Aug 17. Factors Affecting Loan Repayment Performance of Small Scale Enterprises Financed by Micro Finance Institutions: Study on Private Borrowers around Wolaita and Dawuro Zone. Global Journal of Management and Business Research.
14. Acquah, H. D., and J. Addo. "Determinants of loan repayment performance of fishermen: Empirical evidence from Ghana" (2011).
15. Karlan, D., Morten, M., & Zinman, J. (2016). A personal touch in text messaging can improve microloan repayment. Behavioral Science & Policy, 1(2), pp. 31–39.
16. Raj, S. R., Sasidharan, S. (2018). Does the caste of the firm owner play a role in access to finance for small enterprises? Evidence from India. The Developing Economies, 56(4), 267–296.
17. Patel PC, Lenka S, Parida V. Caste-Based Discrimination, Microfinance Credit Scores, and Microfinance Loan Approvals Among Females in India. Business & Society. 2022;61(2):372-388.
18. Ferreira, Miguel A. & Eça, Afonso & Prado, Melissa Porras & Rizzo, A. Emanuele, 2022."The real effects of FinTech lending on SMEs: evidence from loan applications" Working Paper Series 26 39, European Central Bank
19. Gopal, Manasa and Schnabl, Philipp, The Rise of Finance Companies and FinTech Lenders in Small Business Lending (May 13, 2020). NYU Stern School of Business

Authors

Pranav C. Madhukar is a senior in high school in Bangalore, India, at Sri Kumaran Children's Home - CBSE. He represents India in international parliamentary debates and has a keen desire to focus on issues of equity across caste, race, gender, and economic divides.