



Caste and Credit: Discrimination in India's Credit Sector

Navjot Sangwan^{ID} and Bibhas Saha

Abstract: This article examines how caste shapes access to credit in India's formal and informal lending markets. Using nationally representative data from the India Human Development Survey (2011–2012), we analyse loan application rates and loan amounts and compare outcomes between General Castes (GC) and three lower-caste groups: Other Backward Castes (OBC), Scheduled Castes (SC) and Scheduled Tribes (ST). We find that GC households are more likely to apply for and receive larger loans from formal banks, while lower-caste households rely more heavily on informal moneylenders. A substantial share of these credit gaps—particularly in bank lending—remains unexplained by observable characteristics, pointing to potential caste-based discrimination. In contrast, moneylenders do not appear to penalize lower-caste borrowers to the same extent and, in some cases, lend more than expected to OBC households. These findings suggest that entrenched caste hierarchies continue to influence credit access in India, with formal institutions reinforcing rather than correcting social inequalities.

Keywords: Caste discrimination, credit, Blinder–Oaxaca decomposition, quantile decomposition, India

I. Introduction

India's caste system represents one of the world's most enduring forms of social stratification, profoundly shaping economic opportunities and social mobility for over a billion people. Despite constitutional abolition of caste-based discrimination in 1950, this ancient hierarchy continues to influence virtually every aspect of Indian society—from education and employment to housing and financial services (Deshpande, 2000; Kijima, 2006; Thorat and Attewell, 2007). The rigid social hierarchy traditionally assigned occupations, resources and rights along hereditary lines, with upper-caste groups controlling

most resources and well-paying occupations, while lower-caste communities were relegated to menial and low-paying work (Mosse, 2018). Although the most overt practices of untouchability and public discrimination have been outlawed, caste-based divisions persist in more subtle but equally consequential forms across economic, social and political spheres.

Credit access represents a fundamental pathway through which caste continues to influence economic outcomes in contemporary India (Thorat, 2009). Access to financial resources is widely acknowledged as critical



for facilitating upward economic and social mobility, enabling entrepreneurship, agricultural investment and poverty reduction (Duflo and Banerjee, 2011; Kaboski and Townsend, 2012). Conversely, systematic exclusion from credit markets can perpetuate intergenerational cycles of disadvantage, particularly for historically marginalized communities.

The importance of credit access cannot be overstated in the Indian context. Credit serves as one of the most critical constraints to economic advancement for lower castes (Thorat, 2009), with variations in access constituting a major source of income inequality (Demirgüç-Kunt and Levine, 2009). Previous research has demonstrated that lack of credit access constrains entrepreneurship (Banerjee et al., 2017), limits agricultural investment and farm production (Kochar, 1997), restricts spending on education (Doan et al., 2014) and impedes poverty reduction efforts (Chowdhury et al., 2005). These constraints are particularly acute for lower-caste groups who remain socially excluded from mainstream economic opportunities and lack access to assets, public facilities and networks that could improve their economic prospects (Thorat and Neuman, 2012).

Credit access can be shaped by caste through multiple mechanisms. A lower-caste loan applicant may face discrimination from bank officials or moneylenders not just in terms of approval but also regarding loan amounts. This discrimination may be entirely driven by caste prejudice—what Becker (2010) defines as ‘taste-based discrimination’. Alternatively, lenders may believe certain caste groups represent higher default risks, leading to statistical discrimination as conceptualized by Arrow (1973) and Phelps (1972). Complicating matters further, anticipating such discrimination, lower-caste borrowers may apply for smaller loans or may not apply at all, creating a self-fulfilling prophecy where fear of discrimination reduces credit demand.

A growing body of research has explored the relationship between caste and credit in India, though with varying methodological approaches and findings. Kadam et al. (2024) investigated small business credit access and found that entrepreneurs from marginalized castes receive significantly lower loan amounts. Patel et al. (2022) examined caste discrimination in microfinance loan approvals, finding that females from a lower caste, relative to those from a higher caste, have lower odds of receiving loans. Kumar and Venkatachalam (2019) examined caste-based discrimination in agricultural credit, finding marginalized caste borrowers are less likely to apply for bank loans. Khanna and Majumdar (2020) found that some lower castes pay substantially higher interest rates and get smaller loans in the informal lending sector.

In this article, we examine whether a borrower's caste influences their access to credit. We focus on two key dimensions: the probability of applying for a loan and the amount borrowed. Our analysis uses data from the India Human Development Survey (IHDS), conducted in 2011–2012, which covers over 41,000 households. The data give information on households credit outcomes as well as their caste identity, along with an array of other information. We look at four broad caste categories—a grouping of higher caste, which we call General Castes (GC), and three groupings of lower castes, namely Other Backward Castes (OBC), Scheduled Caste (SC) and Scheduled Tribes (ST). The social hierarchy places OBC ahead of SC and ST.

Our empirical strategy compares GC with each of the three lower-caste groups across two types of lenders: formal banks and informal moneylenders. For each, we analyse differences in loan application rates and loan amounts. We use the Blinder–Oaxaca decomposition technique to separate the observed gaps into portions explained by differences in household

characteristics (such as income, land, education and occupation) and those that remain unexplained. The unexplained component, while not conclusive proof of discrimination, could reflect omitted variables correlated with caste (Deshpande and Sharma, 2014). This is a limitation common to all decomposition analyses. To address potential endogeneity in current income, we instrument household income using past income reported in the 2004–2005 wave of IHDS.

We find four main results. First, GC households are more likely to apply for bank loans, while lower-caste households are more inclined to seek credit from moneylenders. Second, GC borrowers receive larger loan amounts from both banks and moneylenders, even though lower-caste households rely more heavily on informal credit. Third, a substantial part of these credit gaps remains unexplained by observable characteristics, suggesting the presence of discrimination or systematic bias in formal lending. Fourth, moneylenders do not disadvantage lower castes to the same degree; in some cases, they even appear to favour OBC borrowers beyond what their observable characteristics would predict.

These findings highlight how caste shapes not only who participates in different segments of the credit market but also how borrowers are treated once they apply. In formal banking, strict rules, paperwork and risk models tend to favour GC households, while lower-caste borrowers may be judged as riskier or discouraged from applying at all. In contrast, moneylenders rely more on local knowledge, relationships and social pressure to ensure repayment. This makes them more willing to lend to lower-caste borrowers and in some cases even prefer them, because repayments can be enforced informally. Together, these findings show that differences in how lenders assess and enforce risk help explain why credit gaps between castes look different in banks compared to moneylenders.

This study makes several important contributions to the existing literature on caste-based credit discrimination in India. First, most studies focus on either formal or informal credit markets but rarely compare discrimination patterns across both sectors systematically. Second, we distinguish between productive and consumption loans, uncovering differential patterns of discrimination based on loan purpose. Third, unlike previous research, this analysis employs Oaxaca–Blinder decomposition to quantify the extent to which observed credit gaps can be attributed to differences in observable characteristics versus potential discrimination, while addressing selection and simultaneity bias.

This article is structured as follows: Section II presents data and uses descriptive evidence to highlight caste differences in India. Section III sets out the methodology for the article. Section IV presents and discusses the results. Section V concludes the article.

II. Data and Sample Characteristics

The data used in the article come from the IHDS—a nationally representative survey of 42,152 households in 2011–2012 collected from 1,503 villages and 971 urban neighbourhoods across India. The survey covers a range of questions relating to economic activity, income and consumption expenditure, assets, social capital, education, health and marriage. Additionally, we use household income data from the earlier round of the IHDS (2004–2005) to construct our instrumental variable.

Table I presents the descriptive statistics of the variables of the four caste groups used in the analysis. The proportion of caste groups used in IHDS 2011–2012 is similar to National Sample Survey (2006), where GC are 30%, OBC are 41.1%, ST are 8.6% and SC are 19.5%. GC households report the highest average loan amounts and the highest share of borrowing from formal banks. In contrast, OBC, SC

Table 1. Summary Statistics.

<i>Variables Description</i>	<i>GC</i>	<i>OBC</i>	<i>ST</i>	<i>SC</i>
Proportion in the Sample	28.57	41.10	8.78	21.74
Loan details:				
Loan amount ¹ (₹)	64,410	52,949	24,611	28,483
Bank (₹)	240,066	154,717	137,227	104,658
% loan from banks	43.78	32.31	26.42	22.90
Moneylenders (₹)	90,334	69,159	34,867	42,400
% loan from moneylenders	19.86	21.45	23.77	21.23
% participation in the market	46	60	44	56
Household characteristics:				
Yearly income (₹)	178,309	114,354	92,998	99,492
Yearly Income 2005	75,420	47,279	39,268	38,676
Female-headed household	0.14	0.14	0.15	0.15
Age of the head	51.88	49.53	47.92	47.83
Education years of the head	7.45	5.46	3.91	4.39
Size of the household	4.82	4.92	4.77	4.83
Amount of land in acres	11.94	11.60	16.52	4.78
Number of years living in the current place	46.22	45.71	45.65	45.20
Experienced any expenditure shock in the last 5 years	0.62	0.63	0.60	0.62
% own land	46	46	59	35
% live in urban area	44	35	14	30

and ST households rely more on informal sources, especially moneylenders. While 46% of GC households participate in the credit market, this share is higher for OBCs (60%) and SCs (56%), and slightly lower for STs (44%).

Caste-based stratification translates into low human capital for lower-caste individuals. The average number of education years completed by ST (head of the households) in the sample was 3.91 years, 4.39 years for SC, 5.46 years for OBC, followed by 7.45 years for the GC. Generally, the differences at lower levels of education (primary) are less pronounced across social groups but start to diverge widely by middle school and higher. For instance, only 3.5% of the SC heads of household have achieved a graduate or postgraduate education compared to 14% of GC individuals.

Substantial caste-based disparities are also evident in income, education and landholding. GC households report the highest average income (₹178,309), followed by OBCs (₹114,354), SCs (₹99,492) and STs (₹92,998). The average number of education years completed by ST (head of the households) in the sample was 3.91 years, 4.39 years for SC, 5.46 years for OBC, followed by 7.45 years for the GC. Generally, the differences at lower levels of education (primary) are less pronounced across social groups but start to diverge widely by middle school and higher. For instance, only 3.5% of the SC heads of household have achieved a graduate or postgraduate education compared to 14% of GC individuals. Although lower-caste households are more likely to live in rural areas, they often own less land: only

35% of SC households and 46% of OBC households report land ownership, compared to 46% among GCs and 59% among STs.

Other household characteristics such as age of the household head, household size and residential stability (years spent in the same location) show little variation across caste groups. Around 14%–15% of households are female-headed in all groups, and over 60% report having experienced an expenditure shock in the past five years due to illness, natural disaster, marriage, unemployment, death or crop failure—suggesting that economic vulnerability is common but not necessarily caste-specific.

III. Methodology

This article presents estimates of the credit differential between castes in the Indian credit sector and the extent to which this differential can be explained by differences in observable characteristics or ‘endowments’ of clients across caste groups. The amount of credit the borrower has arisen from the following equation:

$$\ln Y_{ij} = X_{ij} \beta_j + u_{ij}, \quad (1)$$

where $\ln Y$ is the natural logarithm of the loan amount of i th individual in j th social group ranging from GC, SC, ST and OBC. X_{ij} is a vector of observed characters, and β_j is a coefficient vector to be estimated for each caste type, and u_{ij} is assumed to be a normally distributed error term.

The Blinder–Oaxaca decomposition is employed to decompose the credit amount gap in outcomes between various castes.² Oaxaca (1973) and Blinder (1973) introduced a regression-based decomposition method that separates the observed gap in an outcome between two groups into an ‘explained’ component and an ‘unexplained’ component. The explained component reflects differences in

endowments and observable characteristics (such as education, income or assets) between the groups. The unexplained component, by contrast, captures differences in the returns to these characteristics—often interpreted as unequal treatment or structural bias. While the unexplained component is often used as a measure for discrimination in the literature (Amin et al., 2019), it is very likely that the residual also includes the effects of unobservable or unmeasurable characteristics.³

The difference in the credit amount arises from the following equation:

$$\overline{\ln Y_g} - \overline{\ln Y_l} = \underbrace{\left[(\bar{X}_g - \bar{X}_l) \hat{\beta} \right]}_{\text{EXPLAINED}} + \underbrace{\left[(\hat{\beta}_g - \hat{\beta}^*) \bar{X}_g + (\hat{\beta}^* - \hat{\beta}_l) \bar{X}_l \right]}_{\text{UNEXPLAINED}}, \quad (2)$$

where $\ln Y$ is the natural logarithm of the loan amount, g and l subscripts stand for general caste and lower castes (SC, ST and OBC), respectively. X_g is a vector of observed characters for the general caste, X_l is a vector of observed characters for various lower castes, and β_g is a coefficient vector to be estimated for the general caste, β_l is a coefficient vector to be estimated for the lower caste and β^* is the estimate of the non-discriminatory credit coefficient from a pooled regression over both groups (Neumark, 1988).

Selectivity and Simultaneity Bias

Another methodological problem faced in analysing the caste gap is the existence of endogeneity which can be caused by self-selection and simultaneity bias. Selection bias could occur when individuals with similar characteristics (education, assets or consumption level) have different levels of entrepreneurship, perseverance and ability, which may lead to different probabilities of their participating in the credit market. The Heckman two-step procedure (Heckman, 1979) corrects for sample selection bias by

explicitly modelling two related processes: (a) the selection equation, which estimates the probability that a household chooses to apply for a loan, and (b) the outcome equation, which estimates the amount borrowed or the approval outcome, conditional on having applied. The method assumes the error terms in the selection and outcome equations are jointly normally distributed. The first step is typically estimated using a probit model that includes variables likely to affect the decision to seek credit, such as household composition, landownership, education and other relevant household characteristics. This yields an inverse Mills ratio (IMR), which summarizes the part of the error term correlated with selection. In the second step, this term is included as a regressor in the outcome equation to correct for selection bias.⁴

Using the Blinder–Oaxaca decomposition, the observed earnings differential can be further decomposed into:

$$\ln \bar{Y}_g - \ln \bar{Y}_i + [\gamma_i \bar{\lambda}_i - \gamma_i \bar{\lambda}_g] = \underbrace{[(\bar{X}_g - \bar{X}_i) \hat{\beta}^*]}_{\text{EXPLAINED}} + \underbrace{[(\hat{\beta}_g - \hat{\beta}^*) \bar{X}_g + (\hat{\beta}_i - \hat{\beta}^*) \bar{X}_i]}_{\text{UNEXPLAINED}} \quad (3)$$

where γ is the coefficient of the IMR (λ).

Simultaneity bias can arise due to the potential endogeneity of household income: Higher income may increase the amount a household borrows, while at the same time, obtaining larger loans could influence income (reverse causality). To address this, we require an instrumental variable (IV)—an exogenous variable that is correlated with current income (the endogenous regressor) but uncorrelated with the error term in the loan amount equation.

In this analysis, we use household income reported in 2004–2005 (from the first round of the India Human Development Survey, IHDS-I) as an instrument for current household income (measured in 2011–2012). Lagged variables are commonly used as instruments in the literature (Ackerberg et al., 2007; Doraszelski and Jaumandreu, 2018; Wang and

Bellemare, 2019), leveraging temporal separation to address endogeneity.

The key identification assumption—the exclusion restriction—requires that past income affects current loan amount only through its effect on current income. Several factors support this assumption. First, the seven-year lag means lenders are highly unlikely to observe or consider historical income when making loan decisions; instead, they focus on current verifiable indicators such as employment, assets and recent earnings. Second, the major life-course traits that could channel past income into borrowing—such as household head's education, land ownership and primary occupation—tend to be fixed over this period and are explicitly included as controls. As a result, any direct effect of past income on current borrowing should be negligible, except insofar as it operates through current income.⁵

Past income is strongly correlated with current income in the first-stage regression (see Table A5) (satisfying the relevance condition) and F-statistics, providing further confidence in the instrument's strength. The Cragg–Donald Wald F-statistic of 949.244 substantially exceeds the Stock–Yogo critical values for weak instrument bias, surpassing even the most stringent threshold of 16.38 for 10% maximal IV size distortion. This indicates that our instrument is sufficiently strong to avoid the bias and inference problems associated with weak instruments (Stock and Yogo, 2005). The Anderson canonical correlation LM statistic (903.661, $p < .001$) confirms that our equation is identified, rejecting the null hypothesis of under-identification.

IV. Results

Decomposing the Differences in Loan Application Rates from Banks and Moneylenders

We begin by estimating the probability of applying to either bank or moneylender.

The dependent variable is 1 if the client has applied or 0 otherwise. The estimates of the probit regressions (see Tables A1 and A2 in the Appendix) are used to construct the IMR for the purpose of correcting the credit amount equation for selection bias as reported in the later section. Since participation in the credit market is a binary outcome, we cannot directly apply the standard Oaxaca–Blinder decomposition. Instead, we use Fairlie’s (2005) method, which adapts the decomposition approach for nonlinear outcomes.

Table 2 reports the decomposition of caste gaps in loan application rates into explained and unexplained parts, separately for formal banks (Columns 1–3) and informal moneylenders (Columns 4–6). The analysis compares GC with three other caste groups. The first panel shows participation gaps in bank credit. GC households are significantly more likely to apply for bank loans than all three disadvantaged groups. The loan application rate difference ranges from 9.2% points between GC and OBC, 15.9% points between GC and ST and 21% points between GC and SC, all significant at the 1% level. The results suggest that most of these gaps are explained by differences in

observable characteristics—such as higher education, greater land ownership and higher income among GC households. For example, in the GC versus SC comparison, 15.8% points of the 21-point gap is explained by these characteristics. The remaining unexplained component—which may reflect discrimination, anticipatory exclusion or unobserved traits—is small but statistically significant in some cases: 5.1% points for GC versus SC and 4.5 points for GC versus ST. For GC versus OBC, the unexplained component is small and not significant, suggesting that most of the gap is driven by observable characteristics.

This pattern is consistent with the institutional features of formal banking: paperwork requirements, collateral demands and documented income history advantage households with higher socio-economic status—traits more common among GC households.

The second panel shows contrasting patterns for the loan application rate in borrowing from moneylenders. Here, lower-caste households are significantly more likely than GC households to apply for loans, as suggested by the negative coefficient. The loan application rate difference ranges from –14.1% points between GC and

Table 2. Decomposition of Loan Application Rates.

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	GC vs. OBC	GC vs. ST	GC vs. SC	GC vs. OBC	GC vs. ST	GC vs. SC
	<i>Panel I: Banks</i>			<i>Panel II: Moneylenders</i>		
Difference	0.092*** (0.009)	0.159*** (0.015)	0.210*** (0.010)	–0.141*** (0.008)	–0.142*** (0.015)	–0.171*** (0.010)
Explained	0.081*** (0.004)	0.114*** (0.009)	0.158*** (0.006)	–0.038*** (0.003)	–0.069*** (0.006)	–0.060*** (0.005)
Unexplained	0.011 (0.009)	0.045** (0.016)	0.051*** (0.011)	–0.103*** (0.008)	–0.073*** (0.016)	–0.111*** (0.011)
Observations	12,950	5,867	8,824	12,950	5,867	8,824

Notes: Robust standard errors are given in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The table shows the decomposition of loan application rates from banks and moneylenders. The dependent variable is 1 if applied for a loan at a bank, 0 otherwise (Columns 1–3). The independent variable is 1 if applied for a loan from a moneylender (Columns 4–6). The dependent variables are the same as those used in the selection model.

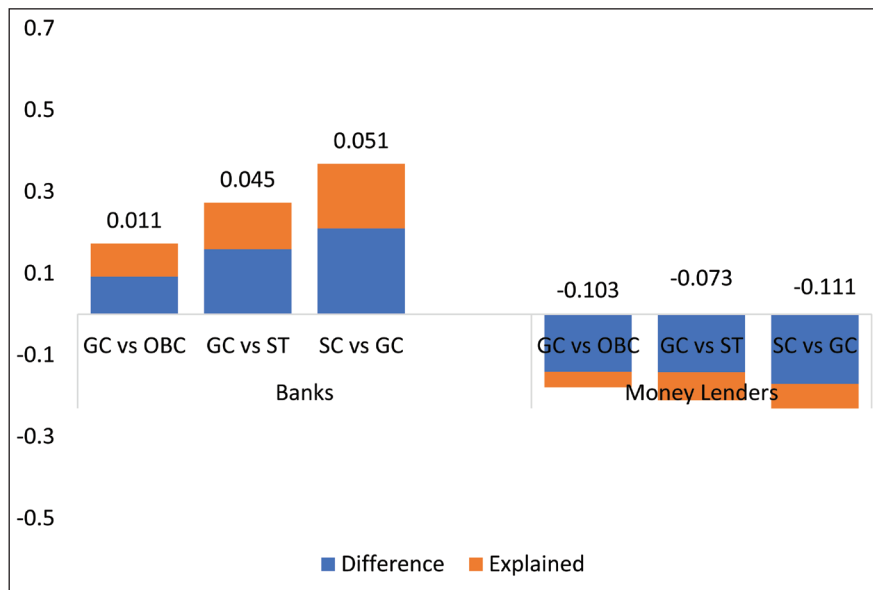


Figure 1. Loan Application Rate Differences Between Castes in Lending from Banks and Moneylenders.

OBC, -14.2% points between GC and ST and -17.1% points between GC and SC. Part of this difference is explained by observed characteristics, but a substantial share remains unexplained and statistically significant, for instance, 11.1% points for GC versus SC, 10.3 for GC versus OBC and 7.3 for GC versus ST, suggesting that moneylenders lend to these groups at higher rates than observable traits alone would predict. Figure 1 visualises these caste differences in loan application rates, breaking down the explained and unexplained components for both banks and informal lenders.

Several mechanisms could explain this. First, moneylenders operate in local markets where they have better information about borrowers' true financial status, reputation and repayment history, reducing informational asymmetry. Second, moneylenders often rely on social and community-based enforcement, making it easier to extract repayment from households who are locally embedded and less legally protected. This

aligns with the idea that marginalized borrowers may be seen as easier to discipline informally or, as Kuran and Rubin (2018) argue, more 'financially powerless'—paradoxically increasing their attractiveness to lenders.

Overall, the results highlight a dual pattern: GC households are better positioned to access formal credit markets, largely because of favourable socio-economic characteristics, while lower-caste households disproportionately rely on informal lenders, partly due to structural disadvantage and partly because moneylenders may be more willing to engage with them beyond due to social networks and better information about their financial status.

Decomposing Credit Differential with the Selection Effect

Loan Amount

In the next stage, we examine caste differences in the loan amount taken from banks separately for production and consumption loans.

To account for self-selection into borrowing and the potential endogeneity of income, we estimate Heckman two-step models with instrumental variables, as described in the methodology section (see Tables A3 and A4 in the Appendix). We then apply Blinder–Oaxaca decomposition to the loan amounts to disentangle how much of the caste gap is explained by observed characteristics (such as income, land and education) and how much remains unexplained.

Table 3 presents these results, comparing GC households with OBC, ST and SC households, separately for production loans (Columns 1–3) and consumption loans (Columns 4–6).

The adjusted credit differential shows that GC has an advantage of 0.39 log points (~48%) over OBC, 1.12 log points (~207%) over ST and 0.87 log points (~139%) over SC for production loans. A substantial portion of these credit differentials remains unexplained: 0.21 log points (~24%) between GC and OBC, 0.61 log points (~85%) between GC and ST, and 0.55 log points (~73%) between GC and SC. This unexplained portion suggests that even after controlling for measurable economic advantages,

GC households still receive significantly higher loan amounts, possibly reflecting differential treatment by banks, unobserved creditworthiness or potential discrimination in loan allocation.

A similar pattern emerges for consumption loans. GC households again have a higher loan amount, with differences of 0.36 log points (~43%) for GC versus OBC, 0.85 log points (~134%) for GC versus ST, and 0.68 log points (~97%) for GC versus SC. Here too, part of the gap is explained by observable characteristics, but a sizeable unexplained component remains: 0.43 log points (~54%) for GC versus ST and 0.26 log points (~30%) for GC versus SC. Notably, the unexplained portion for GC versus OBC is smaller (0.097 log points or ~10%) but still statistically significant. Figure 2 provides a graphical summary of caste-based credit gaps in bank lending for both production and consumption loans, decomposed into explained and unexplained components.

These findings imply that banks lend larger amounts to GC households not solely because they are objectively ‘better risks’ (e.g., higher income or collateral), but also due to factors that are not fully captured by observable characteristics. Possible mechanisms include implicit biases

Table 3. Credit Differences Between Castes in Bank Lending.

Variables	Production Loan			Consumption Loan		
	GC vs. OBC	GC vs. ST	GC vs. SC	GC vs. OBC	GC vs. ST	GC vs. SC
Difference	0.392*** (0.049)	1.119*** (0.093)	0.867*** (0.067)	0.361*** (0.046)	0.850*** (0.103)	0.683*** (0.054)
Explained	0.173*** (0.026)	0.503*** (0.059)	0.312*** (0.056)	0.264*** (0.025)	0.422*** (0.058)	0.421*** (0.040)
Unexplained	0.219*** (0.045)	0.616*** (0.093)	0.554*** (0.075)	0.097** (0.043)	0.429*** (0.096)	0.261*** (0.053)
Observations	2,885	1,383	1,555	3,242	1,536	2,255
	0.392***	1.119***	0.867***	0.361***	0.850***	0.683***

Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The table shows the decomposition results corrected for selection and endogeneity for a sample of borrowers from banks. The covariates include log of income, age, education, land ownership, sex, urban or rural area, numbers of years living in the current place, size of the household, IMR and occupation.

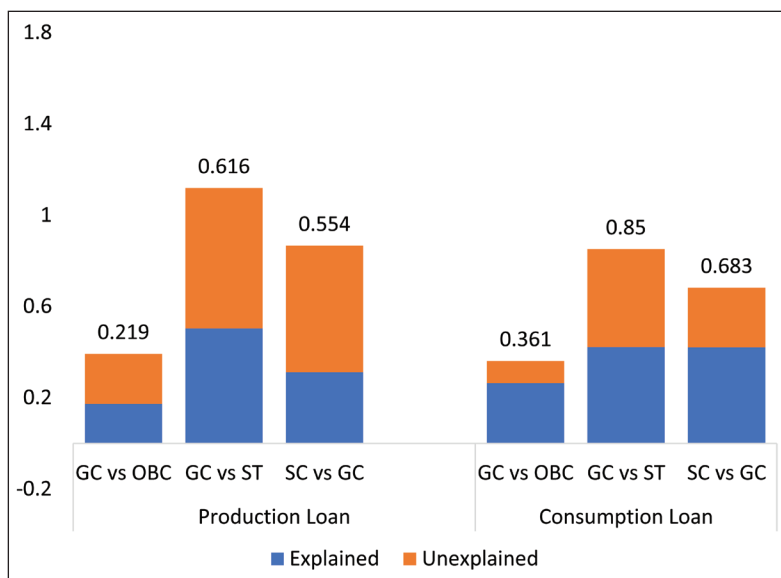


Figure 2. Credit Differences Between Castes in Bank Lending for Production and Consumption Loans.

by loan officers, institutional norms that favour traditionally privileged groups or greater confidence in GC borrowers' repayment capacity based on historical patterns.

Moneylenders

Although the share of moneylenders has reduced significantly over the past few years, they still play a major role in financing lower-caste borrowers. However, Table 4 shows that there are large credit differentials between the general caste and other lower castes.

The adjusted credit differential, focusing on production loans obtained from moneylenders, indicates that GC households borrow significantly larger amounts than lower-caste households: 0.34 log points (~40%) more than OBC, 1.21 log points (~235%) more than ST, and 0.85 log points (~134%) more than SC.

For GC versus OBC, the explained component (0.84 log points or ~132%) is larger

than the observed gap (0.34 log points or ~40%), while the unexplained component is negative and significant (−0.50 log points or approximately −39%). This means that based solely on observable traits like income, assets and education, GC households should borrow substantially more. Instead, moneylenders lend relatively larger amounts to OBC borrowers than their characteristics alone would predict, suggesting informal lenders may actually favour OBC borrowers or at least do not penalize them as formal lenders might.

For GC versus ST, both the explained (0.39 log points or ~48%) and unexplained (0.82 log points or ~127%) components are positive and significant, suggesting moneylenders lend more to GC borrowers partly due to better economic characteristics and partly due to unobserved factors or discriminatory preferences.

For GC versus SC, the unexplained component is small (0.07 log points or ~7%) and

Table 4. Credit Differences Between Castes in Lending from Moneylenders.

Variables	Production Loan			Consumption Loan		
	GC vs. OBC	GC vs. ST	GC vs. SC	GC vs. OBC	GC vs. ST	GC vs. SC
Difference	0.337*** (0.090)	1.207*** (0.146)	0.852*** (0.107)	0.091 (0.061)	0.875*** (0.100)	0.498*** (0.064)
Explained	0.838*** (0.165)	0.389** (0.157)	0.783*** (0.221)	0.644*** (0.103)	0.322*** (0.092)	0.631*** (0.115)
Unexplained	-0.501*** (0.172)	0.818*** (0.189)	0.070 (0.231)	-0.553*** (0.114)	0.552*** (0.123)	-0.133 (0.122)
Observations	1,124	404	567	2,732	990	1,983
	0.337***	1.207***	0.852***	0.091	0.875***	0.498***

Notes: Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The table shows the decomposition results corrected for selection and endogeneity for a sample of borrowers from a moneylender. The covariates include log of income, age, education, land ownership, sex, urban or rural area, no of years living in the current place, size of the household, IMR and occupation.

not statistically significant, indicating the gap is primarily driven by observable differences in characteristics.

In the case of consumption loans, the overall gaps are smaller: GC households borrow only 0.09 log points (~9%) more than OBC households (not significant), but substantially larger amounts compared to ST (0.88 log points or ~141%) and SC (0.50 log points or ~65%).

Again, the unexplained components vary. For GC versus OBC, the unexplained part is large, negative and significant (-0.55 log points or about -42%), implying that moneylenders extend larger loans to OBC households than their observable characteristics predict. For GC versus ST, the unexplained component is positive and significant (0.55 log points or ~73%), indicating possible unobserved favouring of GC borrowers. For GC versus SC, the unexplained component is negative (-0.13 log points or approximately -12%) but not significant, suggesting this gap can largely be explained by observable household characteristics. Figure 3 provides a graphical summary of caste-based credit gaps in informal lending for both production and consumption loans, decomposed into explained and unexplained components.

These patterns suggest that, unlike banks, moneylenders do not systematically favour GC borrowers. In fact, for OBC households, moneylenders lend larger amounts than observables would predict—reflected in negative unexplained components. The absence of a significant unexplained component for SC also suggests that the moneylender does not discriminate against SC. Several mechanisms might explain this. First, moneylenders often know their borrowers personally, which reduces perceived risk beyond what formal indicators capture. Second, marginalized borrowers may be easier to pressure or discipline if they default, making them attractive clients despite weaker formal profiles. As argued by Kuran and Rubin (2018), the legally privileged (like GC borrowers) may be harder to discipline, making lower-caste borrowers paradoxically more attractive from a risk-management perspective.

Discussion of the Results

The results show that GC households obtain larger loans than lower-caste households across banks and moneylenders, for both production and consumption loans (with the exception of OBCs in one model, where the gap is positive

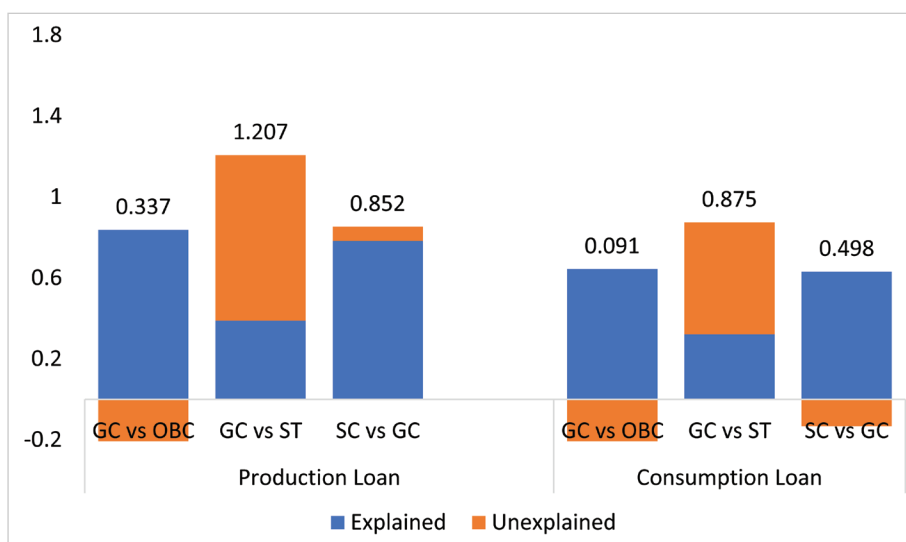


Figure 3. Credit Differences Between Castes in Lending from Moneylenders for Production and Consumption Loans.

but statistically insignificant). The critical difference lies not in the credit differentials but in the unexplained component of these gaps, which suggests how lenders treat caste groups beyond what observable characteristics justify.

For banks, the unexplained component is consistently positive and significant across caste groups and loan types. This implies that GC borrowers receive larger loans than their observed profiles would predict, pointing to potential discrimination against lower castes. In contrast, for moneylenders, the unexplained component is negative and significant for OBCs, indicating that moneylenders lend more to OBC borrowers than their observed characteristics alone would justify. For SC borrowers, the unexplained component is generally small and insignificant, suggesting an absence of discrimination in lending.

One plausible mechanism for this pattern is the enforceability of financial contracts, following the argument of Kuran and Rubin (2018). In their study of Ottoman Istanbul,

socially privileged groups paid higher interest rates because biased courts made it harder for lenders to enforce repayment, increasing credit risk. Conversely, disadvantaged groups, lacking legal protections, were easier to pressure or punish, lowering lenders' perceived risk.

In the Indian informal credit market, moneylenders rely on local enforcement: panchayats (village councils typically dominated by upper-caste men, whose decisions often favour the lenders), reputational sanctions or direct seizure of collateral. These mechanisms make lending to lower-caste borrowers less risky despite weaker formal credit profiles. This enforcement advantage reduces lenders' need to discriminate as a risk-management tool, explaining why moneylenders may favour OBC borrowers or at least not penalize SC borrowers.

By contrast, formal banks operate within a regulated credit system heavily influenced by state-mandated policies aimed at uplifting disadvantaged castes. Successive governments

have introduced schemes prioritizing loans for SC and ST at concessional or interest-free rates. While these measures are well-intentioned, they often produce unintended consequences. Banks perceive lending to SC/ST borrowers as riskier due to diminished incentives for repayment—amplified by practices such as loan waivers—and the absence of robust punitive measures for default. Unlike moneylenders, banks cannot easily use informal pressure; instead, they rely on collateral requirements or credit histories—all areas where lower-caste households typically face structural disadvantage. To manage perceived risk, banks may restrict loan sizes, leading to persistent unexplained gaps.

Another important mechanism is relational lending. Moneylenders often draw on personal knowledge, long-term relationships and local reputation when assessing borrowers. These informal sources of information can compensate for weaker formal indicators of creditworthiness among lower-caste borrowers, reducing uncertainty and lowering risk premia. Banks, in contrast, depend on documentation, standardized scoring models and formal records. Lower-caste households—often with shorter credit histories, fewer assets and weaker documentation—may therefore be disadvantaged even before any conscious bias comes into play.

The unexplained component in our decomposition could reflect either taste-based discrimination (prejudice against lower castes) or statistical discrimination (lenders using caste as a proxy for unobserved risk). Distinguishing between these two mechanisms is empirically challenging and beyond the scope of our study due to data limitations. In credit markets where borrower records are incomplete and credit histories remain patchy—as is often the case in India—statistical discrimination is plausible: Lenders may infer risk from caste even after controlling for observable factors such as income, education and assets.

In informal lending, where moneylenders rely heavily on local knowledge, social ties and relational enforcement, information about borrowers is richer. If discrimination persists despite this informational advantage, it could be indicative of taste-based preferences. However, our data do not allow for a direct test of this hypothesis. This difficulty in separating statistical from taste-based discrimination is well established in the broader literature (Bertrand and Mullainathan, 2004), and both forms may coexist within credit markets

V. Conclusion

This article has examined how caste shapes credit outcomes in India across formal banks and informal moneylenders. The analysis shows that, despite decades of policy aimed at financial inclusion, deeply entrenched caste disparities persist—not only in participation but also in the amounts borrowed.

Prior studies have shown that lower-caste households face disadvantages in formal lending (Kadam et al., 2024; Kumar and Venkatachalam, 2019; Patel et al., 2022; Thorat et al., 2010) and are often pushed into informal credit markets (Guérin et al., 2013). Some evidence also suggests that even in informal lending, lower-caste borrowers receive smaller loans or pay higher costs (Khanna and Majumdar, 2020). Our results broadly confirm this pattern: GC households remain more likely to apply for and secure larger loans from banks, while lower-caste households turn more often to moneylenders. Yet by comparing both formal and informal lenders and decomposing credit gaps into explained and unexplained components, this article goes further than previous work in three ways.

First, we show that credit gaps vary markedly between formal banks and informal lenders—a perspective often missing from studies that analyse these sectors separately. Second, we move beyond describing average

gaps to quantify how much of the difference is due to structural disadvantages (such as lower income, education or land) versus unexplained factors that may signal discrimination. Third, our results reveal that although GC borrowers generally receive larger loans even from moneylenders, the unexplained component is small, negative or insignificant—suggesting moneylenders do not systematically penalize lower-caste borrowers and, in some cases, may even prefer OBC households.

Formal banks operate within bureaucratic and regulated frameworks, relying heavily on documented income, collateral and credit histories—dimensions in which historically advantaged groups are better positioned. Conversely, informal moneylenders leverage local information, reputation and social enforcement mechanisms, reducing their reliance on formal risk indicators. Following Kuran and Rubin (2018) argument, socially marginalized borrowers may, paradoxically, be seen as more ‘enforceable’ because of weaker legal protection, making them relatively attractive to informal lenders.

These results have policy implications. Financial inclusion initiatives that merely expand formal banking infrastructure without addressing underlying social stratification may reproduce or even deepen historical inequalities. Efforts must go beyond physical access to tackle informational asymmetries, build credit histories for disadvantaged groups, and reform practices that embed bias into lending decisions. At the same time, recognizing the role of informal lenders—not as mere market failures but as part of local social networks that help people access credit.

Overall, our findings demonstrate that caste remains a central axis of financial inequality in India. Formal and informal credit markets do not operate in isolation but reflect and reinforce the broader social structure. Addressing credit gaps, therefore, requires confronting the

underlying hierarchies that shape who feels able to apply, who is trusted and ultimately who is served.

Authors' Affiliations

Navjot Sangwan is the corresponding author (n.sangwan@gre.ac.uk) and is affiliated with the School of Accounting, Finance and Economics, University of Greenwich, London, UK. Bibhas Saha is affiliated with Durham University Business School, Durham University, Durham, UK.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The authors received no financial support for the research, authorship and/or publication of this article.

ORCID iD

Navjot Sangwan  <https://orcid.org/0000-0001-7263-4502>

Notes

1. Largest amount of loan taken by the borrowers in the last five years.
2. Blinder–Oaxaca decomposition has been used to measure differences between castes in health outcomes (Maity, 2018), labour market (Hnatkovska et al., 2012), poverty (Borooah, 2005), school enrolment (Borooah and Iyer, 2005) and access to energy (Saxena and Bhattacharya, 2018) in the Indian context.
3. All decomposition analyses are subject to this caveat (Deshpande and Sharma, 2014). However, we have taken several steps to minimize such bias. Our models include a rich set of covariates—including income (instrumented to address endogeneity), education, land, occupation, urban/rural status and household demographics—to account for key determinants of credit outcomes. We also address selection bias through a Heckman correction. By controlling for these factors, the unexplained component represents disparities that persist beyond observable characteristics and selection effects, making it a plausible proxy for unequal treatment or lender bias.
4. Identification of the Heckman model requires at least one variable in the selection equation that influences the decision to apply for credit but does *not* directly

affect the loan amount. In this study, we use recent household expenditure shocks (e.g., drought, illness or marriage) as such instruments. These shocks plausibly increase the need to borrow, raising the probability of applying for a loan, but should not systematically affect lenders' assessment of creditworthiness once the household applies, especially after controlling for observable financial and demographic characteristics.

5. While we recognize the concern that past income could influence current financial decisions via channels such as social capital or entrepreneurial drive, we believe this risk is limited. The instrument (income in 2004–2005) is measured only six to seven years before the outcome, a period over which most key household traits—such as the household head's education, land ownership and occupation—are highly stable. Moreover, these characteristics are included as controls in our model. The exclusion restriction therefore relies on the reasonable assumption that, conditional on these controls, income from 2004–2005 affects current borrowing only by shaping current income, rather than through direct effects on credit demand or lender behaviour.

References

- Ackerberg, D., Benkard, C.L., Berry, S. et al.** 2007: Econometric tools for analysing market outcomes. In Heckman, J. and Leamer E., editors, *Handbook of econometrics*. North-Holland, 4171–276.
- Amin, M., Islam, A. and Khalid, U.** 2019: Decomposing the labor productivity gap between upper-middle-income and high-income countries. *World Bank Policy Research Working Paper* 9073.
- Arrow, K.** 1973: The theory of discrimination. *Discrimination in Labor Markets* 3(10), 3–33.
- Banerjee, A.V., Breza, E., Duflo, E. et al.** 2017: *Do credit constraints limit entrepreneurship? Heterogeneity in the returns to microfinance*. Buffett Institute Global Poverty Research Lab.
- Becker, G.S.** 2010: *The economics of discrimination*. University of Chicago Press.
- Bertrand, M., and Mullainathan, S.** 2004: Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review* 94(4), 991–1013.
- Blinder, A.S.** 1973: Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources* 8(4), 436–55.
- Borooh, V.K.** 2005: Caste, inequality, and poverty in India. *Review of Development Economics* 9(3), 399–414.
- Borooh, V.K. and Iyer, S.** 2005: Vidya, Veda, and Varna: The influence of religion and caste on education in rural India. *The Journal of Development Studies* 41(8), 1369–1404.
- Chowdhury, M.J.A., Ghosh, D. and Wright, R.E.** 2005: The impact of micro-credit on poverty: Evidence from Bangladesh. *Progress in Development studies* 5(4), 298–309.
- Demirgüç-Kunt, A. and Levine, R.** 2009: Finance and inequality: Theory and evidence. *Annual Review of Financial Economics* 1, 287–318.
- Deshpande, A.** 2000: Does caste still define disparity? A look at inequality in Kerala, India. *American Economic Review* 90(2), 322–25.
- Deshpande, A. and Sharma, S.** 2014: Is self-employment the answer to caste discrimination? Decomposing the earnings gap in Indian household nonfarm businesses. <http://dx.doi.org/10.2139/ssrn.2395700>
- Doan, T., Gibson, J. and Holmes, M.** 2014: Impact of household credit on education and healthcare spending by the poor in peri-urban areas, Vietnam. *Journal of Southeast Asian Economics (JSEAE)* 31(1), 87–103.
- Doraszelski, U. and Jaumandreu, J.** 2018: Measuring the bias of technological change. *Journal of Political Economy*, 126(3), 1027–84.
- Duflo, E. and Banerjee, A.** 2011: *Poor economics* (Vol. 619). Public Affairs.
- Fairlie, R.W.** 2005: An extension of the Blinder–Oaxaca decomposition technique to logit and probit models. *Journal of Economic and Social Measurement* 30(4), 305–16.
- Guérin, I., d'Espallier, B. and Venkatasubramanian, G.** 2013: Debt in rural South India: Fragmentation, social regulation and discrimination. *The Journal of Development Studies* 49(9), 1155–71.
- Heckman, J.J.** 1979: Sample selection bias as a specification error. *Econometrica* 47(1), 153–61.
- Hnatkovska, V., Lahiri, A. and Paul, S.** 2012: Castes and labor mobility. *American Economic Journal: Applied Economics* 4(2), 274–307.
- Kaboski, J.P. and Townsend, R.M.** 2012: The impact of credit on village economies. *American Economic Journal: Applied Economics* 4(2), 98–133.
- Kadam, A., Singh, P. and Chatterjee, J.** 2024: Does the owner's caste affect access to credit for enterprises in India's unorganized sector? *Applied Economics Letters* 31(14), 1310–15.
- Khanna, M. and Majumdar, S.** 2020: Caste-ing wider nets of credit: A mixed methods analysis of informal lending and caste relations in Bihar. *World Development Perspectives* 20, 100265.
- Kijima, Y.** 2006: Caste and tribe inequality: Evidence from India, 1983–1999. *Economic Development and Cultural Change* 54(2), 369–404.
- Kocher, A.** 1997: An empirical investigation of rationing

- constraints in rural credit markets in India. *Journal of Development Economics* 53(2), 339–71.
- Kumar, S.M.** and **Venkatachalam, R.** 2019: Caste and credit: A woeful tale? *The Journal of Development Studies* 55(8), 1816–33.
- Kuran, T.** and **Rubin, J.** 2018: The financial power of the powerless: Socio-economic status and interest rates under partial rule of law. *The Economic Journal* 128(609), 758–96.
- Mosse, D.** 2018: Caste and development: Contemporary perspectives on a structure of discrimination and advantage. *World Development* 110, 422–36.
- Neumark, D.** 1988: Employers' discriminatory behavior and the estimation of wage discrimination. *Journal of Human Resources* 23(3), 279–95.
- Oaxaca, R.** 1973: Male–female wage differentials in urban labor markets. *International Economic Review* 14(3), 693–709.
- Patel, P.C., Lenka, S.** and **Parida, V.** 2022: Caste-based discrimination, microfinance credit scores, and microfinance loan approvals among females in India. *Business & Society* 61(2), 372–88.
- Phelps, E.S.** 1972: The statistical theory of racism and sexism. *American Economic Review* 62(4), 659–61.
- Saxena, V.** and **Bhattacharya, P.C.** 2018: Inequalities in LPG and electricity consumption in India: The role of caste, tribe, and religion. *Energy for Sustainable Development* 42, 44–53.
- Stock, J.** and **Yogo, M.** 2005: Asymptotic distributions of instrumental variables statistics with many instruments. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* 6, 109–20.
- Thorat, S.** 2009: *Dalits in India: Search for a common destiny*. Sage Publications Ltd.
- Thorat, S.** and **Attewell, P.** 2007: The legacy of social exclusion: A correspondence study of job discrimination in India. *Economic & Political Weekly* 42(41), 4141–45
- Thorat, S.** and **Neuman, K.S.** 2012: *Blocked by caste: Economic discrimination in modern India*. Oxford University Press.
- Wang, Y.** and **Bellemare, M.F.** 2019: Lagged variables as instruments. *Working paper*.

Appendix: Tables

Table A1. Probit Model (Banks).

Variables	(1)	(2)	(3)	(4)
	GC	OBC	ST	SC
Log income	0.57*** (0.06)	0.57*** (0.06)	0.34 (0.21)	0.63*** (0.10)
Sex of the head	–0.15** (0.06)	–0.08* (0.05)	–0.08 (0.12)	0.04 (0.07)
Age Sq	–0.00** (0.00)	–0.00 (0.00)	–0.00 (0.00)	0.00 (0.00)
Age	0.04*** (0.01)	0.02* (0.01)	0.03 (0.02)	0.00 (0.01)
Education year	0.03*** (0.01)	0.04*** (0.00)	0.04*** (0.01)	0.04*** (0.01)
Land ownership (1/0)	0.27*** (0.06)	0.23*** (0.04)	0.48*** (0.11)	0.38*** (0.05)
Years in place	–0.00 (0.00)	0.00 (0.00)	–0.00 (0.00)	–0.00** (0.00)
Urban area	–0.19*** (0.06)	–0.14*** (0.04)	–0.09 (0.17)	–0.16** (0.07)

(Table A1 continued)

(Table A1 continued)

	(1)	(2)	(3)	(4)
Variables	GC	OBC	ST	SC
Size of HH	-0.04*** (0.01)	-0.06*** (0.01)	0.01 (0.03)	-0.05*** (0.01)
Shock	-0.06 (0.04)	0.01 (0.03)	0.02 (0.08)	0.01 (0.05)
Occupational dummies	Yes	Yes	Yes	Yes
Constant	-7.86*** (0.57)	-7.19*** (0.50)	-5.03*** (1.83)	-7.52*** (0.85)
Observations	4,586	8,364	1,281	4,238

Notes: Estimates from an Instrumental Variable Probit model. The model instruments for current income using household income from IHDS-I (2004–05). Robust standard errors are given in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable is 1 if the client has taken a loan or 0 otherwise.

Table A2. Probit Model (Moneylenders).

	(1)	(2)	(3)	(4)
Variables	GC	OBC	ST	SC
Log income	-0.15** (0.07)	-0.15** (0.06)	-0.17 (0.19)	-0.19* (0.11)
Sex of the head	0.00 (0.07)	-0.17*** (0.05)	-0.02 (0.12)	-0.16** (0.06)
Age Sq	-0.00 (0.00)	-0.00** (0.00)	0.00 (0.00)	-0.00 (0.00)
Age	-0.00 (0.01)	0.02* (0.01)	-0.01 (0.02)	0.01 (0.01)
Education year	-0.02*** (0.01)	-0.03*** (0.00)	-0.02* (0.01)	-0.03*** (0.01)
Land ownership (1/0)	-0.13* (0.07)	-0.01 (0.04)	0.01 (0.10)	-0.07 (0.05)
Years in place	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
Urban area	-0.08 (0.07)	0.09** (0.04)	-0.06 (0.16)	0.13** (0.06)
Size of HH	0.01 (0.01)	0.01 (0.01)	0.03 (0.03)	-0.00 (0.02)
Shock	0.08* (0.05)	0.08** (0.03)	0.32*** (0.08)	0.10** (0.05)
Occupational dummies				
Constant	1.30* (0.72)	1.20** (0.58)	1.25 (1.68)	1.39 (0.98)
Observations	4,586	8,364	1,272	4,238

Notes: Estimates from an Instrumental Variable Probit model. The model instruments for current income using household income from IHDS-I (2004–05). Robust standard errors are given in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable is 1 if the client has taken a loan or 0 otherwise.

Table A3. Selection Corrected Loan Amount Equation Estimates (Banks).

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	GC	OBC	ST	SC	GC	OBC	ST	SC
Log income	0.09 (0.10)	0.12 (0.07)	0.07 (0.29)	0.15 (0.21)	0.16* (0.08)	-0.04 (0.09)	-0.11 (0.32)	-0.07 (0.13)
Sex of the head	0.35** (0.15)	-0.02 (0.12)	0.43 (0.28)	-0.05 (0.18)	0.24** (0.10)	0.18** (0.09)	0.29 (0.33)	-0.00 (0.11)
Age Sq	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	-0.00** (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00*** (0.00)
Age	-0.01 (0.02)	-0.00 (0.02)	-0.01 (0.07)	0.07** (0.03)	0.01 (0.02)	0.02 (0.02)	0.02 (0.07)	0.05** (0.02)
Education year	0.01 (0.01)	0.02** (0.01)	-0.00 (0.05)	0.00 (0.02)	0.03*** (0.01)	0.03** (0.01)	-0.07 (0.05)	0.03** (0.01)
Land ownership (1/0)	-0.44*** (0.14)	-0.36*** (0.10)	-0.89* (0.49)	-0.29 (0.19)	-0.03 (0.10)	-0.18** (0.08)	-0.97* (0.54)	-0.13 (0.10)
Years in place	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.01)	-0.00 (0.01)	-0.00 (0.00)	-0.00* (0.00)	-0.00 (0.01)	-0.00 (0.00)
Urban area	0.32** (0.12)	0.51*** (0.09)	0.89** (0.36)	0.34** (0.16)	0.34*** (0.09)	0.31*** (0.07)	0.57* (0.30)	0.38*** (0.09)
Size of HH	0.03** (0.01)	0.05*** (0.01)	0.05 (0.05)	0.05** (0.02)	0.03* (0.01)	0.04*** (0.01)	-0.04 (0.06)	0.06*** (0.02)
Mills ratio (bank)	-0.94*** (0.29)	-0.86*** (0.22)	-1.64 (1.26)	-0.36 (0.49)	-0.91*** (0.24)	-0.91*** (0.24)	-3.35** (1.46)	-0.95*** (0.31)
Constant	11.91*** (0.97)	11.97*** (0.69)	12.79*** (3.55)	9.92*** (1.36)	11.17*** (0.76)	11.29*** (0.78)	13.97*** (4.04)	10.42*** (0.96)
Observations	1,122	1,763	261	433	1,320	1,922	216	935
R-squared	0.22	0.22	0.35	0.21	0.26	0.20	0.30	0.35

Notes: Robust standard errors are given in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable is the log of the amount of the loan.

Table A4. Selection Corrected Loan Amount Equation Estimates (Moneylenders).

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	GC	OBC	ST	SC	GC	OBC	ST	SC
Log income	-0.04 (0.19)	-0.40*** (0.13)	0.77*** (0.19)	0.10 (0.23)	-0.02 (0.13)	-0.06 (0.09)	0.87*** (0.14)	0.26** (0.11)
Sex of the head	-0.16 (0.29)	-0.94*** (0.21)	-0.13 (0.55)	0.08 (0.31)	-0.04 (0.14)	-0.35*** (0.12)	0.44* (0.24)	-0.13 (0.12)
Age Sq	-0.00 (0.00)	-0.00*** (0.00)	-0.00 (0.00)	-0.00** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)
Age	0.00 (0.04)	0.11*** (0.03)	0.09 (0.07)	0.11** (0.04)	0.07*** (0.02)	0.11*** (0.02)	0.14*** (0.04)	0.12*** (0.02)

(Table A4 continued)

(Table A4 continued)

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	GC	OBC	ST	SC	GC	OBC	ST	SC
Education year	-0.04 (0.04)	-0.12*** (0.03)	0.03 (0.04)	0.01 (0.04)	-0.03 (0.03)	-0.03 (0.02)	0.12*** (0.03)	0.02 (0.02)
Land ownership (1/0)	-0.58* (0.34)	-0.12 (0.13)	-0.16 (0.43)	-0.23 (0.20)	-0.21 (0.18)	-0.06 (0.07)	-0.06 (0.21)	-0.04 (0.09)
Years in place	-0.02** (0.01)	-0.01** (0.00)	0.01 (0.03)	-0.01 (0.01)	-0.01** (0.00)	-0.01*** (0.00)	-0.00 (0.01)	-0.01* (0.00)
Urban area	-0.21 (0.28)	0.62*** (0.13)	1.98*** (0.54)	0.40* (0.22)	-0.33* (0.17)	0.33*** (0.08)	1.29*** (0.32)	0.26*** (0.09)
Size of HH	0.07** (0.03)	-0.02 (0.02)	0.07 (0.05)	0.03 (0.04)	0.03 (0.02)	0.02* (0.01)	0.03 (0.04)	0.04** (0.02)
Mills ratio (ML)	4.04*** (1.52)	7.05*** (1.12)	-1.05 (1.14)	1.97 (1.56)	4.36*** (1.07)	3.66*** (0.77)	-2.37*** (0.80)	1.50* (0.79)
Occupation dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	6.82*** (1.68)	2.06 (1.38)	9.08*** (2.55)	6.36*** (2.16)	1.88 (1.54)	4.02*** (0.99)	7.00*** (1.40)	5.91*** (1.24)
Observations	287	837	117	280	663	2,069	327	1,320
R-squared	0.25	0.20	0.35	0.22	0.26	0.14	0.24	0.19

Notes: Robust standard errors are given in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable is the log of the amount of the loan.

Table A5. First-stage Regression.

Variables	(1)	(2)	(3)	(4)
	GC	OBC	ST	SC
Log income (2005)	0.29*** (0.01)	0.27*** (0.01)	0.25*** (0.02)	0.26*** (0.01)
Sex of the head	-0.03 (0.03)	-0.06** (0.02)	-0.03 (0.05)	-0.09*** (0.03)
Age Sq	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)	-0.00*** (0.00)
Age	0.04*** (0.00)	0.05*** (0.00)	0.05*** (0.01)	0.07*** (0.00)
Education year	0.04*** (0.00)	0.03*** (0.00)	0.03*** (0.00)	0.02*** (0.00)
Land ownership (1/0)	0.18*** (0.03)	0.06*** (0.02)	0.02 (0.04)	-0.02 (0.02)
Years in place	-0.00*** (0.00)	-0.00*** (0.00)	-0.00 (0.00)	-0.00** (0.00)
Urban area	0.15*** (0.03)	0.21*** (0.02)	0.29*** (0.06)	0.20*** (0.02)

(Table A5 continued)

(Table A5 continued)

	(1)	(2)	(3)	(4)
<i>Variables</i>	<i>GC</i>	<i>OBC</i>	<i>ST</i>	<i>SC</i>
Size of HH	0.09*** (0.00)	0.11*** (0.00)	0.10*** (0.01)	0.11*** (0.00)
Constant	6.31*** (0.18)	5.96*** (0.14)	7.17*** (0.40)	6.41*** (0.21)
Observations	9,994	13,895	2,562	7,440
<i>R</i> -squared	0.36	0.31	0.35	0.36

Notes: Robust standard errors are given in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable is the log of income in 2011.