Caste *dominance* in rural credit markets: Evidence from India

Roshan Saha, Mykel R. Taylor, Valentina Hartarska

Auburn University

Submission for Selected Paper (Lightning Session)

2024 AAEA Annual Conference

Abstract

This study investigates the role of caste in rural credit markets, examining how broad caste categories (Scheduled Tribes, Scheduled Castes, Other Backward Castes, and Others) influence borrowing capabilities and lenders' willingness to lend. Using nationally representative data from 2018-19, the research focuses on four states: Uttar Pradesh, Bihar, Odisha, and West Bengal. The study identifies two main channels through which caste affects credit access: statistical and/or taste-based discrimination, and the cultural 'proximity' channel influenced by caste diversity and social trust. Lenders use caste as a proxy for borrower reliability due to the difficulty and cost of gathering individual information. Despite policies by the Reserve Bank of India to promote lending to disadvantaged groups, caste-based disparities in credit access persist. Caste diversity can erode social trust, leading to increased group-based differences in credit markets due to perceived out-group threats. Conversely, caste diversity may also foster social learning and reduce prejudices, potentially decreasing these differences. The study extends previous research by examining both formal and informal credit sources and considering the interplay between caste dominance and social/economic conditions. Findings reveal that caste diversity exacerbates credit disparities, particularly in districts dominated by Other Backward Castes or Others, while such disparities are absent in Scheduled Caste or Scheduled Tribe-dominated districts. Evidence suggests both discrimination and cultural proximity influence these outcomes, with dominant castes in districts with higher caste diversity exerting more influence on credit outcomes. The research highlights the complex dynamics of caste and credit in rural India, contributing to a deeper understanding of social influences on economic development.

Keywords: Agricultural Finance; Caste; Discrimination; Caste Diversity, Social and Economic Stratification

JEL Codes: Q14, Z13, J15

Introduction

In agriculture there is a gap between sowing and harvest, and as a result farmers need cash/credit advances during the sowing season (Conning and Udry 2007). The amount of financial capital that is available to an individual farmer, at the time of sowing, is in turn dependent upon their wealth and/or their ability to access loans from formal and informal sources. While farming households represent the demand side of the agricultural credit market, the total available credit is limited by the lender’s willingness to provide credit. Most lenders, especially private and informal, are driven by profit motives and as such they may advance credit to the most productive borrowers. However, gathering information on the *type* of the borrower is extremely difficult and costly in a setting with high information asymmetry- such as the rural Indian economy. Under such circumstances, lenders often use ‘observable’ characteristics to extract information about the type of borrower. One of the most striking features of Indian society is the caste system that has historically segregated society into a complicated web of exclusive and/or interlinked networks/social groups (Munshi 2019). Thus, the caste system can be used by creditors to make lending decisions and we explore the extent to which caste affects the credit choices of several types of lenders.. While existing studies have looked at the role of caste either using dummy variables for different caste groups or caste dominance in a district, in this paper we also examine whether caste diversity has any role to play in rural credit markets.

Caste continues to be an eternal measure of economic status and outcomes. Government reports and surveys broadly classify the Indian populace into four broad social groups- Scheduled Tribes (STs), Scheduled Castes (SCs), Other Backward Castes (OBCs) and ‘Others’. These are broad caste categories and each of these groups encompasses several sub-groups or *jatis*. We do not have access to *jati*-level information and therefore we use the term caste and social groups interchangeably in the paper.

Existing research indicates that there are distinct patterns between caste and several socio-political and economic attributes; such as landownership (Bharti 2019), political representation (Jaffrelot 2010), crime (Sharma 2015), education (B. Banerjee and Knight 1985), migration ((Munshi and Rosenzweig 2016) and housing (Thorat et al. 2015), among several other dimensions.[1] Based on these observed characteristics, lenders can develop prior expectations about the behavior of borrowers from these distinct caste groups. For example, through its impact on the ownership of land, socio-ethnic identities such as caste then goes on to play an important role in determining the supply of credit to an individual or to a collective unit of farmers that belong to the same social group. Caste-based differences in access to credit and credit outcomes have been observed despite the emphasis by the Reserve Bank of India towards lending to disadvantaged groups (RBI, 2011; 2014; 2004; 2008) and extending credit facilities to rural areas.

In this paper, we revisit the issue of group-based differences in lending to agricultural households through the lens of the broad caste categories, caste dominance and caste diversity. Using nationally representative data for 2018-19, we estimate the effect of caste on a household's ability to borrow (or a lenders willingness to lend) across four major states- Uttar Pradesh (UP), Bihar, Odisha and West Bengal. We identify two major channels through which group-based differences in credit access might operate. Firstly, lenders might use observable characteristics, such as gender, race or caste, to extract information regarding the creditworthiness of a borrower. This is directly linked to the theories of statistical (Arrow, 1973) and/or taste-based discrimination (Becker, 1969). We try to see which of these theories are applicable to the existing data. Second,

[1] Refer to Munshi (2019) for a detailed review of the existing literature on caste and the Indian economy.

through the interaction between caste diversity and social trust- the cultural 'proximity' channel- group-based differences might occur. When viewed through this latter perspective, there are two possible effects of caste diversity on credit market outcomes. Firstly, social trust serves to mitigate the information asymmetry in the credit market. Therefore, social groups with higher social trust among themselves are more likely to have better credit access. Evidence of 'in-group' preferences has been shown using dyadic data on borrowers and lenders from an Indian bank (Fisman, Paravisini and Vig, 2017). At the same time, caste diversity has a major role in influencing social trust. Dinesen & Sonderskov (2018) review the existing literature and find that there is a general trend of a negative relation between caste diversity and social trust. In other words, in regions with higher caste diversity social trust gets eroded because of a perception of 'threat' from the 'out-groups'. Such behavior is directly linked to the theories of "group threat" and "conflict theory". Alesina and Ferrara (2005) show that ethnic diversity can also lead to lower economic development because it leads to coordination problems and behavioral biases that inhibit trade across groups (as cited in (Bharathi, Malghan, and Rahman 2023)).

In our context, the in-group equivalent would be the 'dominant' caste in a district. We define this as the group with the highest share in landholding. In districts with higher caste-diversity i.e. households belonging to different caste groups are present in the district, the dominant-castes would be exposed to more "out-groups". According to the group-threat and conflict theories, it would lead to an erosion of social trust in the district. So, caste diversity would result in dominant caste's trying to exert their influence to mitigate the out-group threats. As a result, we would expect group-based differences in credit market outcomes to increase. Kumar (2013) has shown that dominant castes in a village influence the lenders willingness to provide credit, especially for cooperative bank loans. Therefore, in districts where dominant castes have an incentive exercise their control, it is likely that we might observe group-based differences in credit outcomes.

At the same time, another strand of literature has emphasized the positive role of caste and *jati*-diversity on social learning, especially in the context of health and sanitary practices (Ashraf et al, 2023). Therefore, within the second channel through which group-based differences might occur, we find that there could be opposing effects. Caste diversity can increase differences if the group threat and conflict theories hold true, and at the same time, they could promote cultural exchange and knowledge sharing which will act as a catalyst in reducing caste-based prejudices. As a result, we can expect group-based differences in credit market outcomes to decline. Which of these two forces are in play in the rural credit markets in India is an empirical question that we address in this paper. [2]

While earlier studies have examined the effect of caste on formal credit (Kumar, 2013; Karthick & Madheswaran, 2018; Kumar and Venkatachalam, 2019; A. Kumar et al., 2020), these studies have ignored the differences in informal sources of credit. In this study, we not only look at caste-based differences in credit usage across different formal sources, but also examine the effect of caste on informal lending. Additionally, Kumar (2013) examined the differences in the organizational structure of commercial and cooperative banks to explain how caste dominance influences credit market outcomes. While commercial banks operate under a centralized decision-making process, cooperative banks are more prone to local level politics. This is where, as Kumar (2013) argues, district level caste dominance comes into play.

In addition to such factors, we argue that it is not just caste dominance but rather the social

[2] In this paper we confine our analysis only to loans taken for agricultural purposes.

and economic conditions under which such dominance may be exerted that influences the credit market outcomes. For instance, dominant castes in more heterogeneous districts (where more caste groups are present) might assert their dominance more intensely compared to dominant castes in a relatively more homogenous setting. We borrow this argument from the literature on ethnic diversity and social trust (Dinesen and Sønderskov 2018). In addition to that, Bharati, Malghan & Rahman (2023) show that caste diversity is a bane for development under conditions of high spatial segregation.

Our paper contributes to the literature by re-examining the issue of caste dominance in rural credit markets through the lens of caste diversity across districts. We find that caste diversity increases the group-based differences in credit market outcomes, especially in districts dominated by the Others/OBC social groups. At the same time, in SC/ST dominated districts, we find that no such differences exist. We also find evidence of taste-based discrimination in favor of OBC/Others borrowers using Oaxaca-Blinder decomposition techniques. Our findings suggest that there exist group-based differences in access to credit and these can be explained through both discrimination and cultural 'proximity' channels.

In the next section we describe the theoretical basis for group-based differences in credit market outcomes. Section 3 provides some descriptive explanation of the data. We discuss our empirical and identification strategy in sections 4 and 5, respectively. Results are explained in section 6, and section 7 concludes. All tables and figures are in the appendix.

Basis for group-based differences in credit markets

Discrimination

Existing research indicates that family assets[3], land size, education, age, nature of tenancy contracts, social groups and caste "dominance" play an important role in determining a household's ability to access and use formal credit (S. M. Kumar 2013) (A. Kumar et al. 2020) (S. M. Kumar and Venkatachalam 2019) (Haque and Goyal 2021). While farming households represent the demand side of the credit market, the total available credit is limited by the lender's willingness to provide credit. With prevalence of information asymmetry, lenders resort to covert measures of extracting information regarding the creditworthiness of the potential borrowers. Often, information is extracted based on some observable characteristics such as race, gender, color, and we argue in this case - the caste of the individual or household. When lenders differentiate based on such social groupings, it leads to two potential types of discrimination- "*taste-based" or "statistical"*.

Taste-based discrimination, *a la* Becker (1969), occurs when lenders might have some prior "taste" preference for lending to a particular group, and it is captured by the 'residual' difference in outcomes between two groups after controlling all relevant characteristics. Statistical discrimination (Arrow, 1973), on the contrary, occurs when differences are attributable to some unobservable 'objective' characteristic, but agents use observable proxies such as race, gender, color, caste or religion, to extract the expected value of the relevant objective characteristic. Statistical discrimination differs from taste-based discrimination in the sense that statistical

[3] Measured through an index comprising of consumer durables in Kumar (2013),

discrimination may occur due to economic incentives, while taste-based discrimination exists despite economic incentives. For example, a lender may get repaid but because he dislikes a person's race, gender or caste, he may not extend the loan even if it is repaid. If some of the characteristics that are controlled are correlated with the caste groups, then the caste-based differences might in fact be a combination of the two-types of discrimination (S. M. Kumar and Venkatachalam 2019).

Cultural 'proximity'

Cultural proximity, such as shared religion and language, between lenders and borrowers can also influence the likelihood and outcome of a transaction (Fisman, Paravisini, and Vig 2017). Using dyadic data on the religion and caste of officers and borrowers from a national bank in India Fisman, Paravisini and Vig (2017) observe that if the two parties are culturally close to each other then there is a positive effect on the transaction. In-group bias manifests in terms of higher credit access, greater loan size dispersion and lowers collateral requirements. Cultural proximity serves to mitigate asymmetric information between the lender and borrowers.

In the absence of dyadic data on both lenders and borrowers we cannot comment on the in-group preferences for one's own type, but we can nonetheless argue the existence of systematic bias against minority social groups. Borrowing from the literature on social trust and ethnicity we argue that group-based differences in economic outcomes are also explained by the influence of cultural ethnicity and social trust (Dinesen and Sønderskov 2018). Social trust can influence credit market outcomes as they serve to mitigate information asymmetries. The hypothesis follows the "familiarity breeds trust" argument. In districts that have a heterogeneous caste population, i.e. higher caste diversity, it is likely that social trust would be eroded because of higher 'out-group' exposure. As a result, the dominant caste in such a region would show higher proclivity to exert their influence in various spheres of economic and social life. This argument is typically made with reference to ‘conflict theory’ or related “group threat” theory (Dinesen and Sønderskov 2018).

At the same time, higher caste diversity can also positively influence social trust, and thereby mitigate prejudice or information asymmetry, by allowing for greater inter-group interaction. Evidence of caste diversity having a “learning effect” on a socio-economic outcome was observed by (Ashraf et al. 2023) in the context of sanitation practices in India.

There are, thus, two possible and opposing effects of caste diversity on social behavior. With respect to public goods or goods that are non-rival in nature we argue that caste diversity has a positive effect. However, in the case of scarce resources, such as rural credit, caste diversity can have a negative effect on social trust/behavior because the groups are competing for limited credit. Under such circumstances, we would expect group-based differences to increase with caste diversity. Adding to the existing scholarship on caste dominance and credit market outcomes, in this paper we examine whether district-level caste diversity can explain group-based differences in credit market outcomes.

Data and Descriptive Evidence

This article uses household level data from the 77th round of the Land and Livestock Holdings of Households and Situation Assessment of Agricultural Households conducted by the National Sample Survey Office (NSSO) for the year 2019. The survey provides detailed information on the incidence of land ownership, credit usage, source of loans, expenditure on farming, income from farming and other socio-economic variables. An important feature of the 77th round of the survey was that it integrated two schedules (33 and 18.1) of the previous survey round that was conducted in 2013.

Before we proceed with our analysis it is important to clearly define some of the key variables here. Firstly, we have information at the plot level, and therefore to understand household level behavior, we aggregate the plot level data to get household level values. Diversity in agricultural practices, economic attributes, geographic factors and social norms makes it difficult to conduct an in-depth pan-India study.

Selecting the four states

Jaffrelot (2010) provides a typology of the Indian states based on the political representation of the castes in these states. According to Jaffrelot (2010), there has been a steady-rise in the politicization of Other Backward Castes (OBCs) in UP and Bihar; upper castes (Others) continue to be very resilient in West Bengal; and Odisha has a significant Scheduled Tribe (ST) population despite being led by upper caste leaders.[4] Our basis for including these 4 states that

[4] This is an observation made by the author(s) and is not directly attributable the findings of Jaffrelot (2010).

form a contiguous geographic region rests on the heterogeneity in the caste-based political dynamics across these states. We can expect that different social groups enjoy economic and political dominance in these 4 states.

We focus on four states from Eastern India- Bihar, Odisha, West Bengal and Uttar Pradesh. While a large part of Uttar Pradesh falls in western India, it has agricultural practices similar to those in Eastern India and is a major part of the Indo-Gangetic plains. Therefore, it is included it in the study. Furthermore, UP has a lot of similarities with Bihar and can provide a basis for comparison. Inclusion of Odisha into the sample was driven by the fact that the state has experienced significant development in the recent decade, and it also has a significant ST population.

Earlier studies have shown the positive link between caste affiliation of the political leader/party and public good provision (A. Banerjee and Somanathan 2007). Therefore, the trajectory of development, and ownership of assets, is likely to be different for different caste groups across these states. Such variation will help in identification in our empirical framework.

Caste and Class

We make a distinction between class and caste. There is evidence of discrimination across both caste and class lines in the literature (Borooah, Tagat, and Mishra 2019;Nandwani 2016;Srinivas 1959; and Mukherjee, 1999). Although Bardhan, 982) argues that land holdings and labor endowment of a household endogenously leads to the creation of agrarian class structures, we use the classification of farmers according to landholding sizes to define four major classes in this study.

Table I. Distribution of land across different land classes

[insert here]

In Table I we depict the distribution of land across the four major classes of farmers- small, semi-medium, medium and large. This definition has been followed by the Government of India in most of its statistical reports on the agriculture sector. It is a broadly accepted and well-defined measure of class structure. Studies have shown that there exists a positive relationship between these class groups and various measures of economic outcomes such as access to credit, education, to name a few. As table I depicts, the total land asset held by small farmers is 6,799 acres. The *land_asset* variable is defined as the total amount of land owned (including the land that was leased-out). We include leased-out land because it reflects the total land asset of a household. By using land asset as opposed to operational holdings, we deviate slightly from the metric used by the Government of India. However, we believe that this captures the class structure more adequately than operational holding (which also includes leased-in land). According to this definition, we see that small farmers own 25 percent of all land, but account for 69 percent of all households in our sample. At the same time, medium farmers own 35 percent of the total land but account for only 9 percent of households. Similarly, 6 percent of large farmers account for 9 percent of total land. From this cursory glance at the data, it is evident that there exists an underlying inequality in the farmland ownership structure.

Fourth, we define caste based on the social groups that each household belongs to. Most All-India surveys gather information on the four major caste groups- Scheduled Castes (SC), Scheduled Tribes (ST), Other Backward Classes (OBC) and Others. Given our focus on land holdings, we recategorize the social-group variable into a binary variable that equals 1 if SC/ST and 0 otherwise. Instead of creating four separate dummy variables, we create a binary variable as defined above following Rajeev and Nagendran, 2023. Analysis by the authors using data from the 70th round of the Land and Livestock Holdings survey shows dominant landholding castes are primarily the OBC and Others caste groups. Because we are trying to disentangle the role of caste in farmland rental market participation, we prefer using the binary variable for SC/ST. This allows us to estimate the impact of SC/ST households with reference to the OBC and Others caste groups. Rajeev & Nagendran (2023) study the adoption of crop insurance among Indian farmers. They find that households belonging to SC/ST, relative to Others and OBC, are less likely to adopt insurance. Therefore, using SC/ST as binary variable to capture discrimination across social-groups has some precedence. However, we note that this approach abstracts from various layers of complexity in socio-cultural hierarchies that exist in the socio-cultural fabric of rural society in India.

Table II. Caste-Class nexus across States (% households)

[insert here]

Table II shows the caste-class nexus across states. There exist substantial regional variations in the caste-class structure. For example, in UP, Bihar and West Bengal we find that the distribution of farmers belonging to different classes is skewed in favor of the Others/OBC caste. However, in Odisha we find that the distribution of farmers is relatively evenly balanced between Others/OBC and SC/ST. Another striking feature is that across all states, except Odisha, the large farmers predominantly belong to the Others/OBC castes. However, in Odisha, it is the SC/ST caste that forms the large farmer's class.

Dominant Caste

In addition to the variation in political representation across these states, our analysis is grounded in the theory of 'caste dominance' proposed by Srinivas (1959). A 'dominant caste' refers to the caste in a village that is numerically strong and also wields economic and political power. Dumont (1970) later argued that dominance stems from economic power rather than numerical preponderance. And economic power flows essentially from the control of resources like land. Here, we follow literature (Anderson, (2011); Jaffrelot (2011)) and use 'dominant caste' to identify the caste group that owns the largest share of land in each district. We expect that lenders would have different *expected* values of credit worthiness across districts, depending upon which caste is 'dominant'.

Table III. Distribution of dominant caste districts[5]

[insert here]

In fact, we find an overlap between our definition of 'caste-dominance' and the typology of caste-based political representation observed by Jaffrelot (2011). Among the four states studied in the paper, Uttar Pradesh and Bihar would fall under the Hindi-belt pattern that is characterized by the growing dominance of the OBCs vis-à-vis the Others. In West Bengal, the Others (upper caste) exhibit resilience and continue to dominate the political personnel of the state legislature. In figure 1, we see that OBC dominates predominantly in UP (46 out of 71 districts) and Bihar (30 out of 38 districts), while 'Others' dominate in West Bengal (17 out of 35), and districts are dominated by ST households in Odisha (10 out 17) (see table III or figure 1).

[5] Also see figure 1.

In figure 2, we present the distribution of the difference in landholdings from the dominant castes in each district. This allows us to see the variation in the relative 'dominance' of a caste in a district. In the horizontal axis, we have the difference from the dominant caste. Values closer to zero imply a very small difference. We see that there is sufficient variation in the degree of dominance of a caste at the district level.

Figure2. Difference from Dominant Caste

[insert here]

Caste Diversity

We measure caste diversity at the district-level as:

$$caste\ diversity\ index = 1 - \sum_{i} s_{ij}^{2}$$

Where, s_{ij}^{2} is the square of the population share of the i-th social group in j-th district.

Figure3. Caste Diversity Index[6]

[insert here]

The mean value of caste diversity index is 0.75 (see table VIII) implying that districts in our sample of 4 states are relative heterogeneous in terms of caste diversity. Figure 3 depicts the distribution of caste diversity across these states. We see that there is a fair amount of variation in the diversity index across districts in most states.

[6] Also see Map 1.

Finally, the survey collects information on income of the household from wages, rent and pension; and income earned from sale of output. To create a variable that captures both agricultural and non-agricultural income, we add the two components to arrive at the total income of the household (*ln_income*). Additionally, we also create a dummy variable that captures the education level of the head of the household- whether they have at least high school education. We have information on the age, participation in employment guarantee schemes such as the Mahatma Gandhi National Rural Employment Guarantee Act (MGNREGA), and also whether the household has non-farm income.

Formal Credit Usage

In 2018 Kisan Credit Card (KCC) facilities were extended to livestock and fish producers. As a result of initiatives such as KCC, along with earlier schemes such as the Jan Dhan Yojana (2014), Doubling Credit Policy (2004), financial sector reforms of the 1990s, and establishment of NABARD in 1982, the flow of institutional credit has increased over time. In 2013, at the all-India level, share of institutional credit increased to 60 percent from 7 percent 5 decades ago. It is almost equal to a 10 percent increase per decade. Nonetheless, widespread disparities across caste, class and geographic regions continue to persist. Eastern states in India have very low share of institutional credit (half of the national average). Among the eastern states, Bihar and Jharkhand have a very high share of informal credit (54 and 63 percent) respectively (2013 estimates). Furthermore, a positive relationship between access to formal credit and farm size indicates that small and marginal farmers may be facing inherent difficulties in accessing formal credit. Ironically, it is the small and marginal farmers that require access to credit more relative to to the larger farmers. At the same time, credit per hectare is low for large farmers and high for small and marginal farmers. This is particularly interesting because although smaller farmers face difficulties in accessing formal sources of credit, they continue to borrow from informal sources (A. Kumar et al. 2020).

Earlier scholarship has highlighted the extortionary nature of informal sources of credit. Interest rates charged by non-institutional sources of credit are higher than the formal sources, and they are disproportionately higher for those belonging to the socially disadvantaged groups, and small and marginal farmers. One of the reasons for lower access to credit by farmers belonging to disadvantaged castes and classes is the low asset value. Despite the introduction of formal credit institutes, the prevalence of informal credit has not declined. On the contrary, the disparity between interest rates charged by formal and informal sources of credit have widened between 2002 and 2013.

Table IV. Interest rates across different lending institutions (2019)

[insert here]

Table IV shows that the trends observed between 2002 and 2013 are still observed in the four eastern states of Uttar Pradesh, Bihar, West Bengal and Odisha in 2019. The average rate of interest charged by informal sources of credit is higher than the formal interest rate across the states. It is the highest in Bihar followed by Odisha, UP and West Bengal. At the same time, formal interest rates are lowest in UP and highest in West Bengal.[7] Households could still be borrowing from informal sources, or may choose not to borrow at all, despite having access (A. Kumar et al.

[7] . It is important to make the distinction between access to credit and actual usage of formal credit. Access to formal credit may not imply that a household is borrowing. Conversely, a household may still be borrowing despite being unable to access formal credit. Even when a household has access to formal credit, the amount borrowed may be zero. Thus, access to formal credit may not imply positive borrowing.

2020). Due to these reasons, we abstain from using a measure of access to credit, such as holding a bank account or a KCC, and instead create a measure of formal credit usage.

We use information on credit and loans outstanding at the time of the survey and select only those loans that were taken for agricultural purposes because we want to estimate the effect of caste on access to agricultural credit. We make a distinction between formal and informal sources of loan/credit. We use the definition provided by NSSO to define institutional and non-institutional sources.[8] We identify a household as a credit user if they have some amount of agricultural loan outstanding at the time of survey. Subsequently, if the loan is taken from a formal source, we identify the household as a formal credit user. This is a binary variable that equals 1 if the household has taken agricultural credit from a formal source, and 0 if they do not have formal agricultural credit or if they have not taken any loan.

We acknowledge that the *formal_credit_user* variable might be concealing information by combining those who take informal credit and no credit into a single category. To facilitate this analysis, the *formal_credit_user* variable adequately captures the required information. We can compare the group of households that use formal credit to those that do not use formal credit.

In addition to examining overall formal sector credit usage, we also carry out a disaggregated analysis for loans taken from two major sources of institutional credit- commercial banks and cooperative banks.

Table V. Distribution of loans from different institutional sources

[8] Institutional sources of credit include scheduled commercial banks, regional rural banks, co-operative society, co-operative bank, insurance companies, provident fund, employer, financial corporation/institution, NBFCs including micro-finance institutions, bank linked Self-Help groups, and other institutional agencies. Non-institutional sources include moneylenders, landlord, input supplier, relatives or friends, market commission agents/traders etc. (*page C-69 of Instructions to Field Staff, Vol-1, NSS 77th Round*).

[insert here]

We see that approximately 70 percent of all institutional loans are from commercial banks while cooperative banks account for 19 percent. The remaining 11 percent of loans is from Self-Help Groups (SHGs) or other sources. We focus only on commercial and cooperative banks for two reasons. First, self-help groups have a very different modus operandi compared to commercial and cooperative banks. Second, the latter two account for approximately 90 percent of all loans from institutional sources.

Commercial Bank

To conduct separate analysis for households that have taken loans from commercial banks we create a variable *commercial_bank* equal to 1 if the household has taken loans only from commercial bank and 0 if the household did not take any loan. We ignore loans from non-institutional sources and non-agricultural loans. Reference category in the binary variable captures those who did not take any loans from commercial banks but might have taken another loan or no loans. This is true because we consider only households that took loans from commercial banks alone. In table VI, we tabulate the average interest rates charged by the commercial banks by caste in each state.

Table VI. Average interest rate for loans from Commercial Banks (%)

[insert here]

Borrowers belonging to the 'Others' social group face the lowest interest rates on average across the four states in our sample. There are inter-state variations in the average interest charged to different social groups. SC households in Odisha face the highest interest rates compared to other groups in the state. At the same time, in Bihar it is the ST and SC households that are charged the highest interest rates, in West Bengal, average interest rates charged to OBC households are higher than the other social groups.

Cooperative Bank

We also include cooperative bank loans in our analysis. Following the same procedure as above, we create a variable *cooperative_bank* equal to 1 if the household has taken loans only from cooperative banks and 0 if they did not take any loans from cooperative banks but may have other loans or don't have other loans. Because we confine our analysis to those households that took loans only from cooperative banks, we can argue that the reference category captures those who did not take any loans from cooperative banks.

Table VII. Average interest rate for loans from Cooperative Banks (%)

[insert here]

We find similar patterns in the cost of lending to different social groups by Cooperative Banks. We can see that it is upper caste borrowers, belonging to the 'Others' social group, that face the lowest interest rates. While the highest interest rates faced by a social group depends upon the state. For example, in West Bengal it is the OBC, in Bihar ST, and in Odisha and UP it's the borrowers belonging to the SC social group.

Empirical Model

We employ a Probit model to estimate the probability of being able to use formal credit by rural households. We are interested in understanding whether castes, age, size of land, education status, income, caste-dominance at the district level, and state and district level controls explain the probability of access to credit. First, we model the probability that a farmer household obtains a production loan from an institutional source. Then, we conduct separate analysis for the two major constituents of institutional credit sources and informal credit sources. We employ univariate probit model specifications to analyse the probability of loans for the four types of outcome variables- namely, formal credit, commercial bank loan, cooperative bank loan, and credit from informal sources.

The univariate model assumes that the binary dependent variable, y_i, is determined according to whether the latent dependent variable y_i^* is positive or negative, where

$$y_i^* = \mathbf{x}_\mathbf{i}'\beta + \epsilon_i$$

And,

$$y_i = \begin{cases} 1, & y_i^* > 0 \\ 0, & y_i^* \leq 0 \end{cases}$$

So that,

$$\mathrm{Prob}(\mathrm{y_i} = 1) = \Phi(\mathbf{x}_\mathbf{i}'\beta)$$

Φ is the cumulative distribution function of a standard normal distribution.

Table VIII. Summary Statistics

[insert here]

In table VIII we report the summary statistics for our sample. 23 percent of households in our sample belong to SC/ST and SC/ST is the dominant caste in only 4 percent of all the districts. Average caste diversity is 0.75 implying sufficient heterogeneity in castes within a district. Average age of the head of the household in our sample is 49 years. We restrict our sample to household heads who are at least 15 years old at the time of the survey. 15 percent of the household heads have attended educational institutions at least till high school. In other words, 85 percent of our sample has education below high-school. Average landholding per household member is 0.34 acres, with the maximum at 13.32 acres. Only 10 percent of our households have a member employed under the MGNREGA scheme. We see that the average loan amount is Rs.87,817 (approx. US$1050 according to exchange rate as of April 8, 2024. 1 INR= 0.012 USD).

Identification strategy

In order to estimate the effect of caste on credit outcomes for rural households we must be able to delineate the effects that emerge due to systematic differences in caste groups. These effects could arise due to supply-side and demand-side factors. Our interest is in disentangling the effect of caste from the perspective of the supply-side i.e., the lenders. Doing that allows us to explain that caste-based differences in credit market outcomes, driven by supply-side bias, still exist and there is need for policy reformulation to address these issues. Furthermore, we control differences in demand for credit that could arise due to differences in the primary occupation of the household. Our paper focuses only on households engaged in agricultural activities and thus we can assume that there are no differences in demand across SC/ST or Others/OBC social groups due to occupational structures.

Kumar (2013) has argued that caste does not play a role in access to credit unless the source of loan is distinguished. Using a bivariate probit model to disaggregate formal loans into commercial and cooperative bank loans, Kumar (2013) finds that marginalized social groups like the SC and OBC are less likely to procure a loan from the cooperative banks. On the other hand, SCs are more likely to have loans from Commercial banks, suggesting that affirmative action policies have been effective. With cooperative banks more prone to local 'elite capture' the effect of affirmative action policies is diminished, and local level politics influences the probability of procuring loans. In the subsequent sections we provide descriptive evidence of the share of loans and interest rates charged to the different social groups across lending institutions. We try to establish that there are no demand side effects and our estimates are indicative of supply side bias in lending.

Identifying the effects of caste requires validation of three major assumptions that ensure that differences arise mainly due to supply-side bias, and not demand-side effects. In other words, we try to rule out the possibility that there are caste-based differences in demand for credit- both formal and informal. The three necessary assumptions, and their justifications, are discussed below.

i. Caste-based migration in search for credit

There is no relocation from one district to another based on availability of credit. Although agriculture is usually location specific, the nature of cross-section data requires that we control migration patterns driven by credit concerns. Munshi (2019) reviews evidence of limited migration due to the presence of caste-based factors. Based on the theory of community networks, Munshi (2019) argues that insurance and credit, in the absence of formal markets, is usually caste-based. Households that decide to migrate to another district risk losing the benefit of insurance or credit, in the event of unforeseen shocks, that the caste-based community network provides to those living within it. Empirical and theoretical evidence is favorable towards the validity of the assumption that caste-based migration in search for credit is limited. There is evidence of caste-based business networks that encourage mobility (Munshi, 2019). However, we focus on households that are engaged in the agricultural sector, and as such are not seeking to migrate outwards to look for new business opportunities. Moreover, Anderson (2011), as argued in Kumar (2013), has shown that the relative proportions of caste at the village, and district levels, have remained largely constant.

Therefore, we argue that households in our sample do not migrate to districts where their probability of getting credit is higher.

ii. Caste-based sorting due to crowding-out and interest rate differentials

Caste-based differences in loans from formal and informal (or commercial and cooperative) banks are not due to crowding out effects. This ensures that there is no caste bias in the demand for loans from specific types of banks. Formal credit institutions account for 78 percent of the loans in the four states covered in the study. Furthermore, we show in table 7 that the differences in loans from different sources are not because of crowding-out. Had it been due to crowding-out, then we would have noticed that due to majority of loans going to OBC and 'Others' in the formal sector, the ST and SC are forced to take loans from the informal sources in UP for example. Instead, what we observe is similar patterns in informal lending as well.

Table IX. Distribution of Formal and Informal loans by social group (%)

[insert here]

In table IX, we see that the share of loans from formal and informal sectors in each state does not present any underlying pattern that may suggest crowding-out effects. Caste-based clustering in loans from formal and informal sources can, thus, be ruled out. What we instead find is that in OBC-dominated states such as UP and Bihar, they account for majority of loans, while in West Bengal, dominant 'Others' account for majority of the loans. In Odisha, where most districts are dominated by ST households, OBC still accounts for a majority of loans, but the ST has a higher share than SC populations. ST population is relatively small (9 percent) compared to the other groups. Nonetheless, their share in total loans (16.6 percent) is more than their share in the population in ST-dominated Odisha. We conclude that while there is no caste-based sorting due to crowding out effect in loans from different sources, there is some evidence of sorting based on the *dominant* caste in the district.

Furthermore, interest rate differentials should not lead to caste-based sorting across formal and informal credit institutions or cooperative and commercial banks. If interest differentials lead to caste-based sorting, then a higher interest rate for a caste group by commercial banks would lead to lower demand for loans by them, and they may then approach cooperative banks for loans. However, we do not observe the demand for loans, and thus our result would be interpreted as lower willingness to sanction loans on the part of the lender. Our estimate of the effect of caste would not be identified. To ensure that such caste-based sorting does not exist, we look at the average interest rates charged to different social groups, across lending institutions. In table VI and VII, we show that for commercial and cooperative bank loans, 'Others' have lower interest across all the four states compared to the rest of the social groups. This implies that the 'Others' should have a higher demand for loans, and the share in observed lending could be due to both lender-bias and higher demand by the borrowers.

Table X. Share of Commercial Bank Loans by Social Group (%)

[insert here]

When we look at the share of loans from commercial and cooperative banks, we find that ST and SC do have lower share in commercial bank loans compared to OBC and 'Others', on average (see table X). But, when we look at the state-wide variations we find that the overwhelming majority of loans are taken by the 'OBC' borrowers. This occurs even though OBCs face a higher interest rate compared to 'Others'. Furthermore, if there was caste-based sorting, then higher loans to OBC and 'Others' borrowers by commercial banks should have crowded out lower caste borrowers who are then forced to approach the Cooperative banks.[9]

Table XI. Share of Cooperative Bank Loans by Social Group (%)

[insert here]

Table XI reports the share of loans from cooperative banks by the social groups. Once again, we find that a majority of these loans are taken by the OBCs despite higher interest rates compared to 'Others'. We do not find any underlying pattern in the share of cooperative bank loans by social groups (table XI) that suggests any crowding-out effects.[10]

We do not find evidence of caste-based sorting due to differences in interest rates. From the above discussion, it is fair to assume that there are no caste-based demand effects on the credit market outcomes in these four states. Rather, we can find underlying evidence of bias in the lending markets that emerge from the relative dominance of a particular caste in the district or state. Therefore, caste-based differences in credit market outcomes in rural India can be attributed to

[9] There are institutional differences and as Kumar (2013) has argued, cooperative banks are more prone to 'elite capture' the outcomes may be different. But the reason why I am arguing that lower caste borrowers, upon being crowded out of commercial banks, might approach cooperative banks for loans. But, the evidence I table XI suggests that even for cooperative bank loan, the share of SC/ST borrowers is lower than Others/OBC groups. So it helps to justify that demand side effects are absent.

[10] Interestingly, in Bihar there might have been some crowding out between cooperative and commercial banks. SC households account for 6 percent of commercial bank loans, while OBC and Others account for approximately 92 percent of the loans. Because SCs are crowded out in the commercial bank, they approach the cooperative banks and account for 45 percent of the loans from the cooperative banks while and the 'Others' account for only 10 percent. So there might be some crowding effect in Bihar, in particular between SC and 'Others' households. But once again, the OBCs continue to dominate both commercial and cooperative bank loans.

bias on the supply-side. This bias could be due to the lender's discriminatory behavior or the influence of the 'caste dominance' on bank operation.

Results

Extensive Margin (Access to Loans)

In this section we explain the results showing the factors affecting a household's probability of getting credit from institutional and non-institutional lending institutions. This is the extensive margin or the availability of loans. We present the probit model coefficients for the different sources of loans, separately. In columns I-IV of table XII we present the estimates for the probability of obtaining loans from all institutional sources, commercial banks, cooperative bank loans, and informal sources respectively.[11]

Lending support to our hypothesis regarding systematic differences in credit market outcomes for minority social groups, we find that SC/ST households have a lower probability of formal credit usage compared to "Others"/OBC social group. This result holds true when we conduct a disaggregated analysis of the source of institutional loans. Interestingly, for loans from non-institutional sources, we find that the effect of SC/ST is positive but not statistically significant. It would be erroneous to draw conclusions regarding social-group based discrimination in the credit markets from the association between SC/ST variable and our outcome variable(s). Although for alternative model specifications we find that this effect is consistent (refer to Appendix). Nonetheless, we control for potential moderators that could confound the effect of social group on probability of credit from various sources. Following the literature, we use caste

[11] We present detailed estimates for estimates for each loan source in Appendix. Refer to table XIII-XVI.

dominance and caste diversity index at the district level to estimate the effect of social group on a household's probability of obtaining credit from both institutional and non-institutional sources. Our hypothesis is that belonging to the dominant social group at the district level will have a positive effect on credit market outcomes.

We find this to be true for institutional loans in general, and also for loans taken from commercial and cooperative banks. This finding is consistent with the literature (Kumar, 2013; cite other papers). In the informal lending sector, on the other hand, we do not find any evidence of caste-based differences in access to credit or loans.

We find that SC/ST households, in districts where they are also the dominant land holding caste, have a higher probability of securing loans from Institutional sources compared to the SC/ST households residing in districts dominated by the Others/OBC social groups. Local or regional dominance of a social group is known to influence credit market outcomes (Kumar, 2013; Anderson and some other papers). However, limited explanation is provided for the possible mechanism through which this effect plays out. A potential moderator of the influence of social group dominance on credit market and other economic outcomes is the diversity of social groups in the regional context. Controlling for caste diversity at the district level, we find that the correspondence between the household social group and dominant social group in the district has a positive impact on the probability of getting credit from an institutional source. We include several household level controls such as age of the head of the household, education status of the head, presence of non-farm income, land per family member of the household, if a member of the household is employed under the national rural employment guarantee scheme, and income (log).

We conduct a disaggregated analysis for two major sources of institutional loans- commercial and cooperative banks. We find social group dominance has a positive effect on the probability of loans from commercial banks for SC/ST social group. However, contrary to the findings in Kumar (2013), we do not find any significant effect of dominant social group on the likelihood of procuring cooperative bank loans if the household belongs to the SC/ST social group. However, when it comes to the informal sources of credit, we see that social group, social group dominance or caste diversity does not play any role in influencing a household's probability of getting credit.

We find age to have a non-linear effect on the decision to participate in the formal credit market. These findings are similar to what Li et al (2020) find for farmers in rural China. Age has a positive impact on formal credit usage and could reflect that experience and credit history both plays a positive role in the credit market participation. However, a negative and significant coefficient for the squared terms suggests that this effect is declining with age. Once again, it follows the expected sign because the positive effect of age at higher values might decline owing to the decline in farming capability beyond a certain age.

Table XII. Probit Model Coefficients

[insert here]

Another important variable in this model is the education level of the head of the household. We find that if the household head has at least attended high-school, there is a positive and significant effect on the probability of credit from institutional sources and commercial banks in particular. However, we find this effect to be negative for cooperative and informal sources of credit, although the estimates are not statistically significant.

Land is an important measure of household wealth and also one of the most commonly used asset that banks or lending institutions prefer as collaterals when extending loans. We find a positive effect of land per family member on the probability of obtaining loans across the different types of institutional credit sources. Examining the variables that capture the income of the households we find that both log income and whether a member of the household is employed in the MGBNREGA[12] program has a positive effect the obtaining loans from formal and informal sources.

Table XIII. Average Marginal Effect of caste diversity index

[insert here]

So far, we have explained the coefficients of the probit model estimates. However, these estimates do not convey the marginal effects. Our focus is on the effect of caste-diversity on caste dominance and probability of credit. We believe that this effect plays out through the regional diversity in caste. In table XIII we report the average marginal effect of caste diversity on agricultural credit. Contrasting Others/OBC and SC/ST dominated districts, we find that for a unit increase in caste diversity, the increase in probability of institutional credit in general is larger among the latter (SC/ST). In particular, the effect of diversity index is higher among SC/ST households in SC/ST dominated regions (10 percent) compared to SC/ST households in OBC/Other dominated districts. However, for both SC/ST and Others/OBC households in SC/ST dominated districts, the effect of caste diversity is the same. But, when we compare households in OBC/Other dominated districts we find that OBC/Other households have a higher probability of getting credit compared to their SC/ST counterparts. We can argue that the effect of caste diversity is more salient in OBC/Other dominated districts compared to SC/ST districts when it comes to institutional credit in general.

[12] World's largest work-fare scheme that self-targets poorer households offering what is dignified as "government work" at the national minimum wage (Mosse, 2018).

We can conclude that diversity creates more friction in credit market transactions in districts dominated by OBC/Others. The average marginal effect of caste diversity on the probability of loans from institutional sources for OBC/Others households in OBC/Others dominated districts is 7 percent while it is 6 percent for SC/ST households. When we look at the same figures for SC/ST dominated districts, we see that the probability of formal credit increases by 10 percent for both groups. Others/OBC are relatively placed higher up in the social hierarchy compared to SC/ST. We can see that the perception of out-group exposure is very different in districts dominated by the two groups. The relatively higher ranked groups find out-group exposure more conflicting and therefore it results in higher group-based differences in credit market outcomes.

This is also reflected in the lower predicted probability of formal loans for households belonging to both groups in OBC/Others dominated districts when compared to the households in SC/ST dominated districts (see figure 4a).

Figure4a. Predicted probability of institutional vs non-institutional credit

[insert here]

Looking at the effect of social group diversity on commercial banks credit, we find that OBC/Other households have a higher probability of credit compared to SC/ST in districts dominated by the former. The effect of caste diversity in SC/ST dominated districts, however, is not statistically significant.

Figure4b. Predicted probability of commercial vs cooperative bank credit

[insert here]

Thus, we see that it is in the OBC/Others dominated districts that the effect of caste or social group is more salient in influencing probability of loans from institutional sources, and commercial banks in particular. These effects are negative and not statistically significant for cooperative banks loans and credit from non-institutional sources.

Intensive Margin (Loan amounts)

In addition to the differences in credit outcomes at the extensive margin, we also estimate the effect on the intensive margin. We further examine if there are differences based on social groups in the amount of loans that households get from the different sources. We test the following empirical specification:

$$\log loan_{hd} = \beta_0 + \beta_1 \mathbf{1}(SC\ or\ ST)_{hd} + \beta_2 \mathbf{1}(Dominant\ Caste = SC\ or\ ST)_d + \beta_3\ \mathbf{1}(SC\ or\ ST * Dominant\ Caste)_{hd} + \beta_4\ Diversity\ Index_d + \boldsymbol{\tau X_{hd}} + \boldsymbol{\gamma_D} + \epsilon_{hd}$$

Table XIV. Estimates of intensive margins

[insert here]

Firstly, in table XIV we report the estimates of the intensive margin models for institutional, commercial, cooperative and informal loans. We find that the belonging to an SC/ST household has a negative impact on the log loan amounts across all sources of credit. To be precise, compared to OBC/Other households, average formal loan amounts is 30 percent lower for SC/ST households. The amount of loans received by SC/ST households from commercial bank loans, cooperative bank loans and informal sources are 32 percent, 31 percent and 23 percent lesser than Others/OBC households, respectively. We find that social group dominance has positive impact on the credit market outcomes at the intensive margin as well. Our results are significant for institutional loans, and commercial bank loans particularly. We do not find any statistically significant difference between the SC/ST or Others/OBC social groups for loans from cooperative banks and informal sources.

Figure 5 depicts the predicted loan amounts (log) from institutional sources. We see that in districts dominated by Others/OBC social groups, households belonging to those social groups are likely to receive larger loan amounts compared to SC/ST households. Although the predicted loan amounts for SC/ST households in SC/ST dominated districts are higher than the predicted loan amounts for Others/OBC households, the difference is not statistically significant. This once again reaffirms the conclusion that it is in the OBC/Others dominated districts where the effect of caste on credit market outcomes is more salient.

Figure5. Predicted probability of (log) loan amounts (institutional loans)

[insert here]

Robustness checks

Our primary focus was on whether minority social groups face systematic differences in the rural credit markets. We find evidence of group-based differences against the minority social groups across both extensive and intensive margins of rural credit. The effect is more pronounced in regions that are dominated by the Others/OBC social groups. Relative to the SC/ST group, the Others/OBC are economically, socially and politically well-positioned in society, and therefore they can exert their ‘caste dominance’ upon the SC/STs more than what the SC/STs can exert in districts dominated by the latter.

Table XV. Interest rate

[insert here]

We extend our analysis to examine if such differences also extend to the interest rates faced by these two groups across the four lending institutions that we study in this paper. We report the results in table XV. We find that compared to OBC/Others, SC and ST households face higher interest rates on average across formal lending agencies. We find that the interaction term between household social group and district caste dominance is negative for formal and commercial loans, although it not statistically significant. Nonetheless, we can see that across various measures of credit accessibility and usage, SC/ST communities face credit constraints, either in the form of lower lending i.e. credit rationing, or higher cost of credit.

We also examine the other channel through which group-based differences in the credit market outcomes may occur- namely statistical discrimination. To further understand whether the differences in credit market outcomes are due to systematic differences in borrower's characteristics, or due to differences in the coefficients of the two groups, we use Oaxaca-Blinder decomposition methods. We estimate the differences in getting loans from formal lending institutions and commercial banks, between OBC/ "Others" (group A) and SC/ST borrowers (group B). The differences in coefficients are the "unexplained" component of the decomposition and it captures the differences in levels of unobservable variables and their differentiating effects (Yun 2004). It represents the change in outcome if SC/ST had the coefficients of the OBC/ "Others" and it is also referred to as the discrimination coefficient.

Table XVI. Oaxaca-Blinder Decomposition

In table XVI we present the Oaxaca-Blinder decomposition estimates. In column I, we decompose the effect for all institutional sources of credit, and in column to we specifically look at the decomposition of the differences in commercial bank loans. For all loans from institutional sources, we see that OBC/Others households are more 11 percent more likely to get credit compared to SC/ST households. 76 percent of this difference can be explained by the difference in coefficients. If SC/ST households had the coefficient of the Others/OBC households, their average probability of getting loans from institutional sources would increase by 9 percent. We can argue that there is some evidence of taste-based discrimination against SC/ST borrowers.

Similarly, the difference in average probability of getting a loan from commercial banks for OBC/ "Others" is 7.6 percent higher than SC/ST borrowers. 60 percent of this difference is attributable to the difference in coefficients. In other words, if SC/ST had the coefficients of the OBC/ "Others" borrowers, their average probability of getting a commercial bank loan would increase by 4.5 percent.

Therefore, we find evidence for group-based differences in credit market outcomes through both channels of discrimination and cultural 'proximity'. Of the two opposing forces in the channel of cultural proximity, we find that the negative effect of caste diversity operating through mechanisms of group threat and conflict theory is stronger than the positive effect of cultural exchange and social learning. This is true in particular for the districts dominated by the OBC/Others social groups.

Conclusion

We study the effect of caste dominance, across heterogeneous caste diversity at the district-level, and its implications on credit market outcomes in rural areas of Uttar Pradesh, Bihar, West Bengal and Odisha. Caste diversity can increase differences if the group threat and conflict theories hold true, and at the same time, they could promote cultural exchange and knowledge sharing which will act as a catalyst in reducing caste-based prejudices. As a result, we can expect group-based differences in credit market outcomes to decline. Which of these two forces are in play in the rural credit markets in India is an empirical question that we address in this paper. We find that it is the latter effect that dominates in our study.

Our findings shed light on the differential impact of caste diversity on credit accessibility and loan amounts across various social groups, particularly highlighting the contrasting scenarios in districts dominated by SC/ST and OBC/Other communities. Notably, we observe that while caste diversity tends to have a more pronounced effect on credit market transactions in OBC/Other dominated districts, the outcomes for SC/ST households are less affected by caste diversity in districts where SC/ST hold dominance. This effect of caste diversity and dominance on credit market outcomes operates through the channel of cultural 'proximity'. We find that group-based differences increase as a result of an increase in caste diversity. Therefore, we can argue that for a service that is competitive in nature, such as credit, it is the negative effect of caste diversity that operates, unlike the case of public goods such as sanitation and health practices, where caste diversity has a positive effect (Ashraf et al, 2023).

Our analysis extends beyond accessibility to credit, delving into the intensive margin by examining loan amounts. Here, we uncover disparities in the amounts of loans received by different social groups, with SC/ST households consistently receiving lower loan amounts across various sources of credit compared to their OBC/Other counterparts. These findings underscore the group-based differences in credit outcomes prevalent in rural credit markets, reflective of broader social and economic disparities. We explain two underlying mechanisms contributing to these disparities, exploring both channels of discrimination and cultural "proximity". Through Oaxaca-Blinder decomposition analyses, we demonstrate that discrimination also contributes to group-based differences in credit market outcomes.

We do not find evidence of caste-based differences in informal credit markets. However, our data is limited in terms of the number of observations that capture the informal credit markets. The result should therefore be interpreted with caution. We also do find evidence of any group-based differences in loans from cooperative banks. Our findings are robust to other measures of credit access or usage, such as interest rates. These preliminary findings reaffirm caste dominance in rural credit markets, especially in commercial banks, thereby suggesting the need to re-examine affirmative action policies at the regional level.

References

Alesina, Alberto, and Eliana La Ferrara. 2005. “Ethnic Diversity and Economic Performance.” *Journal of Economic Literature* 43 (3): 762–800. https://doi.org/10.1257/002205105774431243.

Anderson, Siwan. 2011. “Caste as an Impediment to Trade.” *American Economic Journal: Applied Economics* 3 (1): 239–63. https://doi.org/10.1257/app.3.1.239.

Ashraf, Sania, Cristina Bicchieri, Upasak Das, Tanu Gupta, and Alex Shpenev. 2023. “Learning from Diversity: Jati Fractionalization, Social Expectations and Improved Sanitation Practices in India.” arXiv. https://doi.org/10.48550/arXiv.2312.15221.

Banerjee, Abhijit, and Rohini Somanathan. 2007. “The Political Economy of Public Goods: Some Evidence from India.” *Journal of Development Economics* 82 (2): 287–314. https://doi.org/10.1016/j.jdeveco.2006.04.005.

Banerjee, Biswajit, and J.B. Knight. 1985. “Caste Discrimination in the Indian Urban Labour Market.” *Journal of Development Economics* 17 (3): 277–307. https://doi.org/10.1016/0304-3878(85)90094-X.

Bardhan, Pranab. 1982. “Agrarian Class Formation in India.” *The Journal of Peasant Studies* 10 (1): 73–94. https://doi.org/10.1080/03066158208438190.

Bharathi, Naveen, Deepak Malghan, and Andaleeb Rahman. 2023. “Ethnic Diversity and Economic Development with Spatial Segregation.” *Economics Letters* 222 (January):110951. https://doi.org/10.1016/j.econlet.2022.110951.

Bharti, Nitin Kumar. 2019. “Local Level Land Inequality in India.”

Borooah, Vani Kant, Anirudh Tagat, and Vinod Mishra. 2019. “Conflict, Caste and Resolution: A Quantitative Analysis for Indian Villages.” *Indian Growth and Development Review* 13 (2): 319–38. https://doi.org/10.1108/IGDR-08-2019-0087.

Conning, Jonathan, and Christopher Udry. 2007. “Chapter 56 Rural Financial Markets in Developing Countries.” In *Handbook of Agricultural Economics*, edited by R. Evenson and P. Pingali, 3:2857–2908. Agricultural Development: Farmers, Farm Production and Farm Markets. Elsevier. https://doi.org/10.1016/S1574-0072(06)03056-8.

Dinesen, Peter Thisted, and Kim Mannemar Sønderskov. 2018. “Ethnic Diversity and Social Trust: A Critical Review of the Literature and Suggestions for a Research Agenda.” In *The Oxford Handbook of Social and Political Trust*, edited by Eric M. Uslaner, 0. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190274801.013.13.

Fisman, Raymond, Daniel Paravisini, and Vikrant Vig. 2017. “Cultural Proximity and Loan Outcomes.” *The American Economic Review* 107 (2): 457–92.

Haque, T., and Ankita Goyal. 2021. “Access to Institutional Credit by Farmers in Eastern India.” *Journal of Asian Development Research*, August, 2633190X211040622. https://doi.org/10.1177/2633190X211040622.

Jaffrelot, Christophe. 2010. *Religion, Caste, and Politics in India*. New Delhi: Primus Books.

Kumar, Anjani, Raya Das, S Aditya K, Seema Bathla, and Girish K. Jha. 2020. “Examining Institutional Credit Access among Agricultural Households in Eastern India: Trends, Patterns and Determinants.” *Agricultural Finance Review* 81 (2): 250–64. https://doi.org/10.1108/AFR-04-2020-0054.

Kumar, Sunil Mitra. 2013. “Does Access to Formal Agricultural Credit Depend on Caste?” *World Development* 43 (March):315–28. https://doi.org/10.1016/j.worlddev.2012.11.001.

Kumar, Sunil Mitra, and Ragupathy Venkatachalam. 2019. "Caste and Credit: A Woeful Tale?" *The Journal of Development Studies* 55 (8): 1816–33. https://doi.org/10.1080/00220388.2018.1425397.

Mukherjee, Ramkrishna. 1999. "Caste in Itself, Caste and Class, or Caste in Class." *Economic and Political Weekly* 34 (27): 1759–61.

Munshi, Kaivan. 2019. "Caste and the Indian Economy." *Journal of Economic Literature* 57 (4): 781–834. https://doi.org/10.1257/jel.20171307.

Munshi, Kaivan, and Mark Rosenzweig. 2016. "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap." *American Economic Review* 106 (1): 46–98. https://doi.org/10.1257/aer.20131365.

Nandwani, Bharti. 2016. "Caste and Class." *Review of Market Integration* 8 (3): 135–51. https://doi.org/10.1177/0974929217706807.

Rajeev, Meenakshi, and Pranav Nagendran. 2023. "Protecting Land and Livelihood under Climate Risks: What Hinders Crop Insurance Adoption?" *Land Use Policy* 131 (August):106711. https://doi.org/10.1016/j.landusepol.2023.106711.

Sharma, Smriti. 2015. "Caste-Based Crimes and Economic Status: Evidence from India." *Journal of Comparative Economics* 43 (1): 204–26. https://doi.org/10.1016/j.jce.2014.10.005.

Srinivas, M. N. 1959. "The Dominant Caste in Rampura." *American Anthropologist* 61 (1): 1–16.

Thorat, Sukhadeo, Anuradha Banerjee, Vinod K Mishra, and Firdaus Rizvi. 2015. "Urban Rental Housing Market: Caste and Religion Matters in Access." *Economic and Political Weekly* 50 (26/27): 47–53.

Appendix

Table I. Distribution of land across different land classes

Land Class	Total Land Asset (acres)	Percent (of total land assets)	Number of Households	Percent (of households)
small (<=2 acres)	6799	25	12823	69
semi-medium (2-4 acres)	8422	31	3020	16
medium (4-10 acres)	9461	35	1698	9
large (>10 acres)	2492	9	1025	6
Total	27174	100	18566	100

Table II. Caste-Class nexus across States (% households)

	UP		Bihar		West Bengal		Odisha	
	Others	SC/ST	Others	SC/ST	Others	SC/ST	Others	SC/ST
small	69.5	30.5	73.36	26.64	64.25	35.75	53.57	46.43
semi-medium	87.15	12.85	92.23	7.77	75.68	24.32	54.49	45.51
medium	92.41	7.59	94.82	5.18	80.29	19.71	63.55	36.45
large	80.3	19.7	53.42	46.58	61.61	38.39	47.88	52.12
Total	76.31	23.69	76.78	23.22	66.52	33.48	54.03	45.97

Table III. Distribution of dominant caste districts

	Dominant Castes across districts (by states)				
State	ST	SC	OBC	Others	Total
Uttar Pradesh	-	2	46	23	71
Bihar	-	-	30	8	38
West Bengal	2	3	13	17	35
Odisha	10	-	-	7	17
Total	12	5	89	55	161

Figure1. Distribution of dominant castes

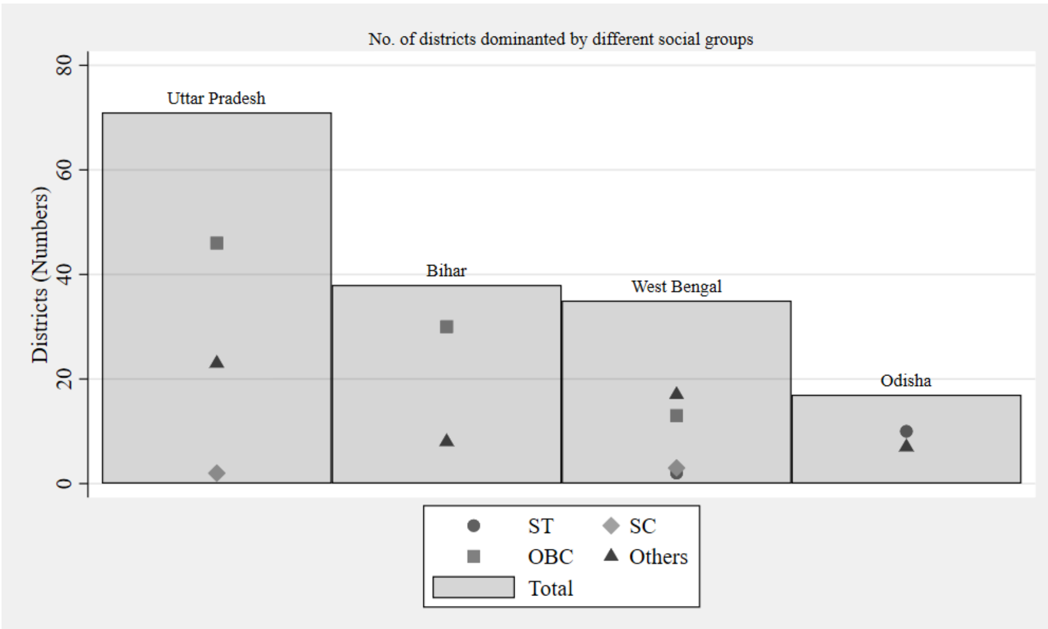


Figure2. Difference (or social distance) from Dominant caste

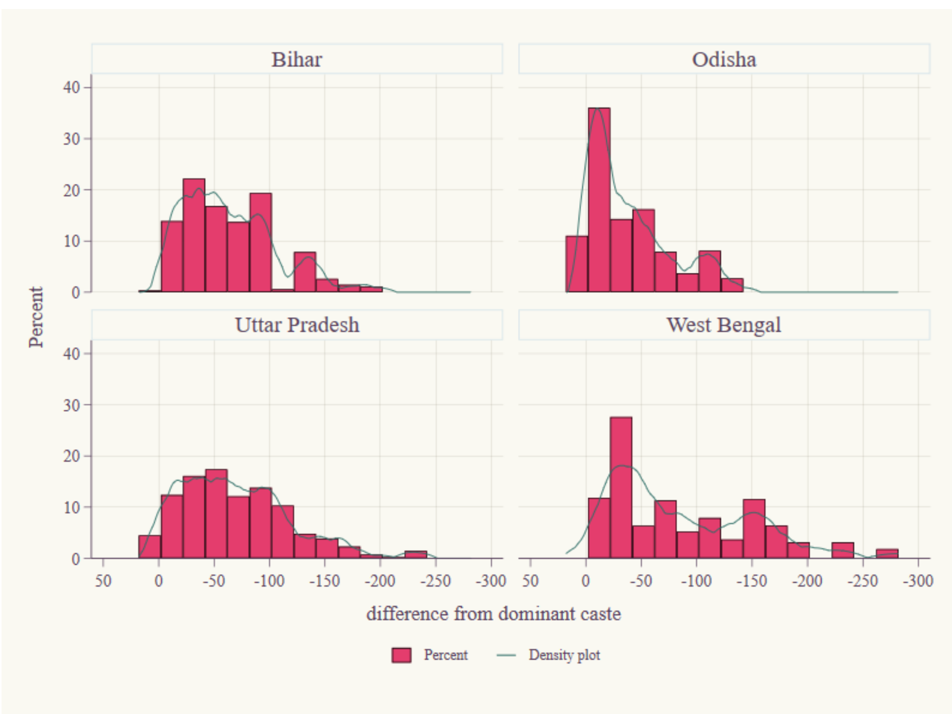
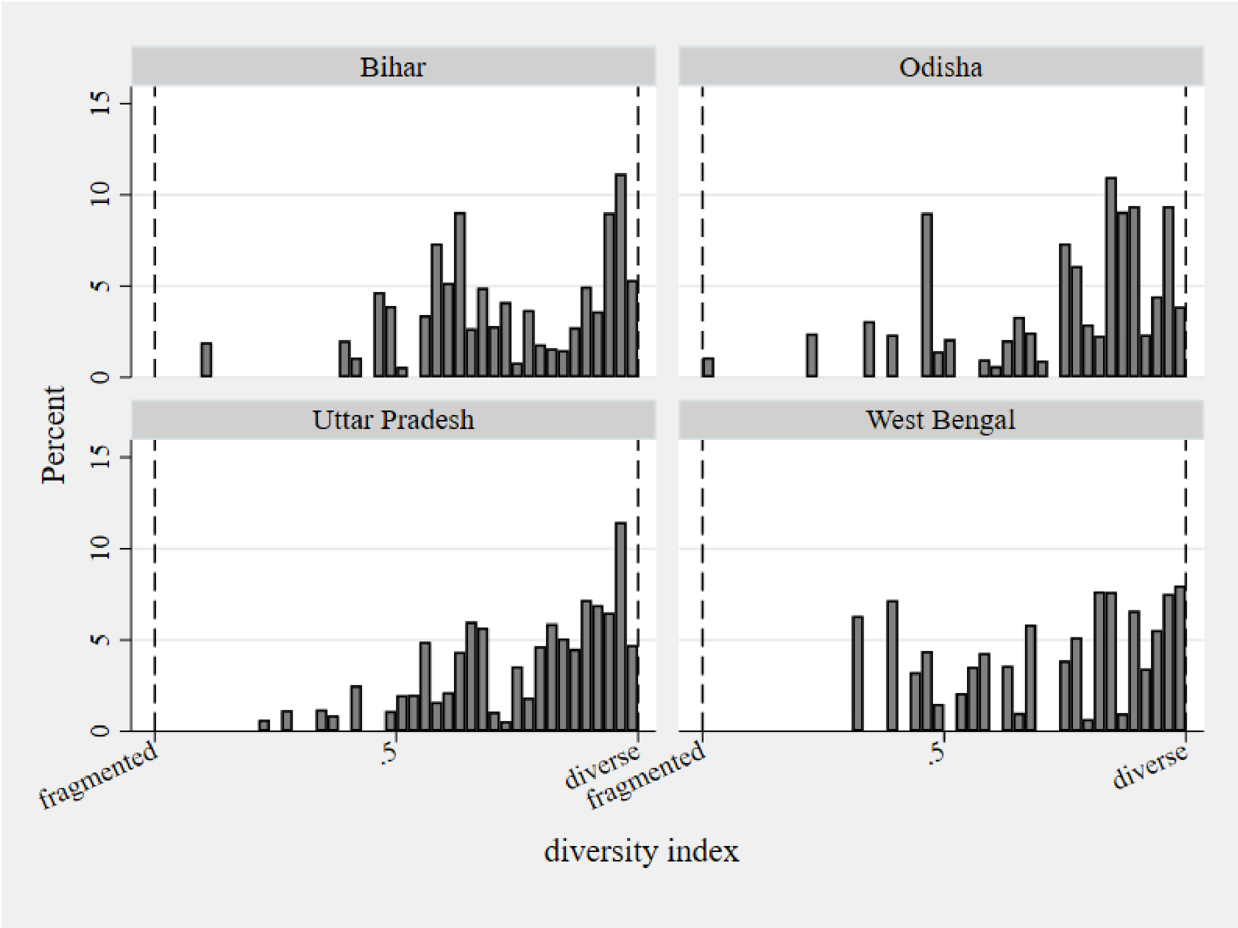


Figure3. Distribution of caste diversity across districts in the 4 states



Map1. Caste diversity and Dominant Caste

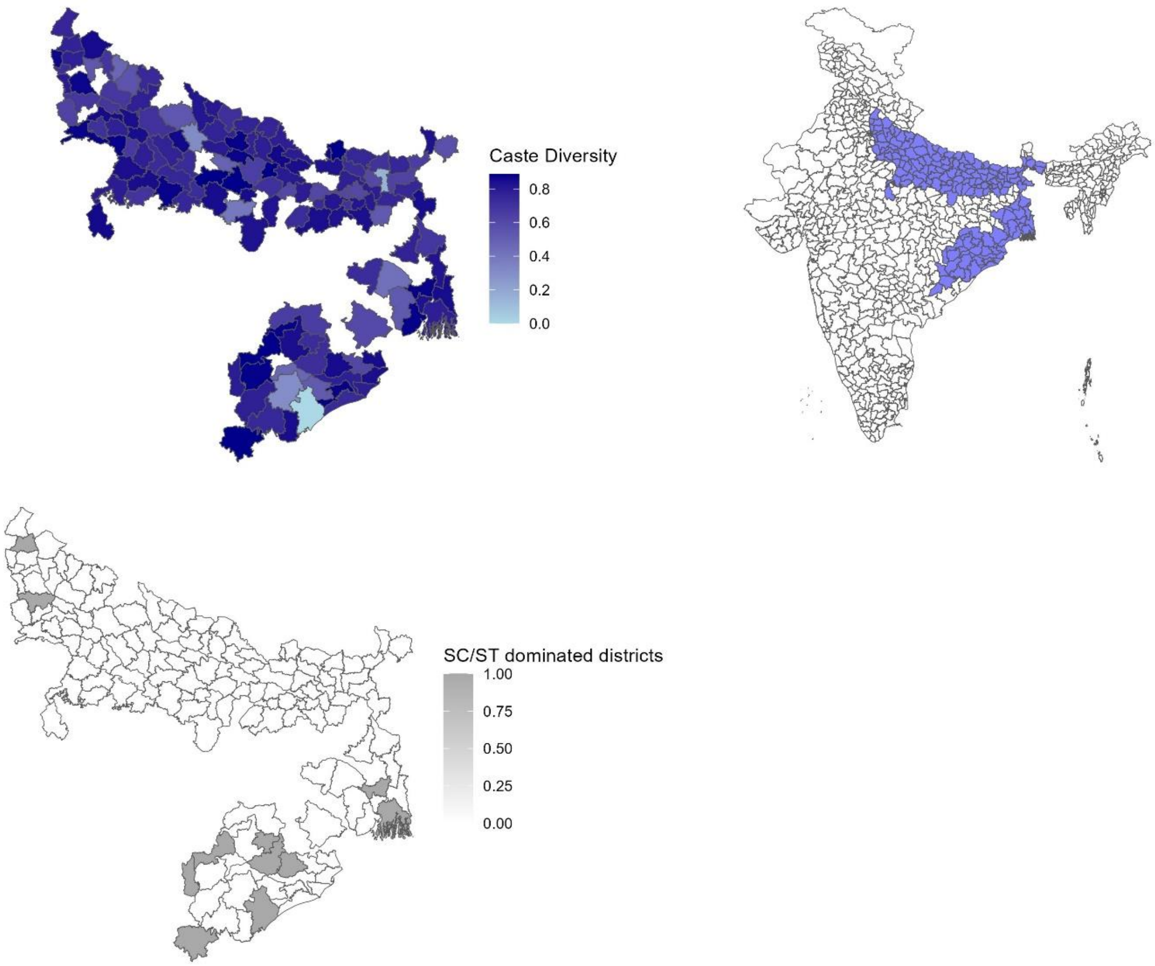


Table IV. Interest rates across different credit agencies (2019)

	Average interest rates (%)			
Source loan	UP	Bihar	West Bengal	Odisha
institutional	7.37	9.18	9.97	9.13
non-institutional	31.95	42.71	24.53	39.78
Total	9.96	17.67	11.21	13.17

Table V. Distribution of loans from different institutional sources

Source	Households	Percent
commercial banks only	2,518	69.75
cooperative banks only	669	18.53
SHG only	415	11.5
other institutional sources	8	0.22

Table VI. Average interest rate for loans from Commercial Banks (%)

Social Group	UP	Bihar	West Bengal	Odisha
ST	10.8	26.87	12.40	12.56
SC	10.98	23.14	11.53	15.02
OBC	8.70	17.64	13.24	12.75
Others	7.99	11.88	9.99	9.98
Total	8.76	17.12	10.99	12.35

Table VII. Average interest rate for loans from Cooperative Banks (%)

Social Group	UP	Bihar	West Bengal	Odisha
ST	10.8	26.87	12.5	12.6
SC	11.83	23.70	11.38	15.11
OBC	9.32	18.34	13.53	12.45
Others	8.15	12.32	9.87	9.75
Total	9.25	17.72	10.95	12.18

Table VIII. Summary Statistics

Explanatory variables	Observations	Mean	SD	Min	Max
Social group (SC/ST=1)	17227	0.23	0.42	0	1
Dominant caste (SC/ST=1)	17227	0.04	0.21	0	1
Caste Diversity index	17227	0.75	0.20	0	1.00
Age	17224	49.34	13.86	15	105
Age^2	17224	2626.63	1426.07	225	11025
High school (yes=1)	17227	0.15	0.36	0	1
Non-farm income (yes=1)	17227	0.06	0.24	0	1
Land per family member (in acres)	16520	0.34	0.58	0	13.32
MGNREGA (yes=1)	17227	0.10	0.29	0	1
Income (log)	16095	10.37	1.08	2.77	14.86
Loan amount (Rs.)	4074	87817.95	126788.40	1	1888467
Interest rate (log)	3772	2.16	0.74	0	4.79

Table IX. Distribution of Formal and Informal loans by social group (%)

	Formal	Informal	Total
		UP	
ST	0.39	0	0.32
SC	11.65	18.34	12.86
OBC	52.75	58.54	53.79
Others	35.22	23.12	33.03
		Bihar	
ST	1.31	4.73	2.37
SC	11.8	14.18	12.54
OBC	56.89	63.64	58.98
Others	30	17.45	26.1
		West Bengal	
ST	5.11	7.58	5.66
SC	21.47	26.71	22.63
OBC	16.05	16.61	16.18
Others	57.36	49.1	55.54
		Odisha	
ST	16.46	17.29	16.64
SC	14.14	13.16	13.92
OBC	42.72	49.62	44.23
Others	26.69	19.92	25.21
		Total	
ST	5.09	6.58	5.42
SC	14.43	18.17	15.25
OBC	42.87	48.19	44.03
Others	37.61	27.06	35.3

Table X. Share of Commercial Bank Loans by Social Group (%)

	UP	Bihar	West Bengal	Odisha	Total
ST	0.29	1.23	5.25	16.57	2.49
SC	12.21	6.16	24.93	13.81	13.34
OBC	53	57.64	16.8	41.44	47.09
Others	34.5	34.98	53.02	28.18	37.08

Table XI. Share of Cooperative Bank Loans by Social Group (%)

	UP	Bihar	West Bengal	Odisha	Total
ST	1.96	0	2.12	16.21	9.31
SC	11.76	45	14.41	8.26	11.99
OBC	49.02	45	11.44	44.95	32.81
Others	37.25	10	72.03	30.58	45.9

Table XII. Probit Model Coefficients

Variable	Institutional credit (I)	commercial banks (II)	cooperative banks (III)	informal credit (IV)
SC/ST	-0.319***	-0.364***	-0.254**	0.022
dominant caste (SC/ST=1)	0.662	-0.395	0.214	-0.255
SC/ST*dominant caste	0.329*	0.426**	-0.301	-0.027
diversity index	0.310**	0.338**	-0.059	-0.241
age	0.049***	0.050***	0.038**	0.033***
age_2	-0.000***	-0.000***	-0.000*	-0.000***
high school or above (yes=1)	0.089*	0.154***	-0.016	-0.113
Non-farm income (yes=1)	0.230***	0.187***	0.261***	0.026
Land per family member (acres)	0.551***	0.498***	0.326***	0.145***
MGNREGA (yes=1)	0.297***	0.128*	0.272***	0.170**
Log income	0.097***	0.066***	0.160***	0
_cons	-3.330***	-3.273***	-4.805***	-1.949***
District Fixed Effects	Yes	Yes	Yes	Yes
N	15300	15127	8687	14288

* p<0.05 **p<0.01 ***p<0.001

errors are clustered at the First Stage Sampling Units

Table XIII. Average Marginal Effect of Caste Diversity Index

Dominant Social Group in the District	Household Social Group	Average Marginal Effect of Caste Diversity Index: Institutional Loan	Commercial Loan	Cooperative Loan	Non-institutional Loan
OBC/Others	OBC/Others	0.0774**	0.0706**	-0.00688	-0.0265
OBC/Others	SC/ST	0.0603**	0.0489**	-0.00509	-0.0273
SC/ST	OBC/Others	0.103**	0.0472	-0.00859	-0.0181
SC/ST	SC/ST	0.103**	0.0506	-0.00454	-0.0179

* p<0.05, ** p<0.01, *** p<0.001

Table XIV. Intensive Margin (loan amount)

Dependent variable (log)	formal loan	commercial loan	cooperative loan	informal loan
SC/ST	-0.371***	-0.390***	-0.372***	-0.259*
dominant caste (SC/ST=1)	-0.786**	-0.754**	-0.701*	0.725
SC/ST*dominant caste	0.438*	0.446*	0.289	-0.41
diversity index	0.306*	0.316*	0.258	-0.248
age	0.013	0.014	0.015	0.014
age_2	0	0	0	0
high school or above (yes=1)	0.044	0.042	0.062	0.082
Non-farm income (yes=1)	0.063	0.063	0.092	-0.025
Land per family member (acres)	0.452***	0.444***	0.460***	0.196*
MGNREGA (yes=1)	-0.092	-0.08	-0.103	-0.138
Log income	0.109***	0.107***	0.110***	0.145***
_cons	8.991***	9.008***	8.898***	7.953***
District F.E	Yes	Yes	Yes	Yes
N	3668	3659	3876	834

* p<0.05 **p<0.01 ***p<0.001
errors are clustered at the First Stage Sampling Units

Table XV. Interest rate

Dependent variable	interest rate (formal)	interest rate (commercial)	interest rate (cooperative)	interest rate (informal)
SC/ST	0.188***	0.192***	0.115**	0.075
dominant caste (SC/ST=1)	0.493**	0.459***	0.291	0.718
SC/ST*dominant caste	-0.042	-0.038	0.14	-1.103
diversity index	-0.02	-0.016	0.077	-0.093
age	-0.013*	-0.011	-0.011*	0.017
age_2	0	0	0	0
high school or above (yes=1)	-0.065*	-0.062*	-0.046	-0.142
Non-farm income (yes=1)	-0.058	-0.056	-0.04	0.05
Land per family member (acres)	-0.129***	-0.133***	-0.090***	-0.082
MGNREGA (yes=1)	0.016	0.012	0.038	-0.043
Log income	-0.019	-0.024	-0.02	0.018
_cons	2.776***	2.714***	2.584***	3.424***
District F.E	Yes	Yes	Yes	Yes
N	3570	3570	3625	438

Figure4a. Predicted probability of extensive margin estimates (formal vs informal)

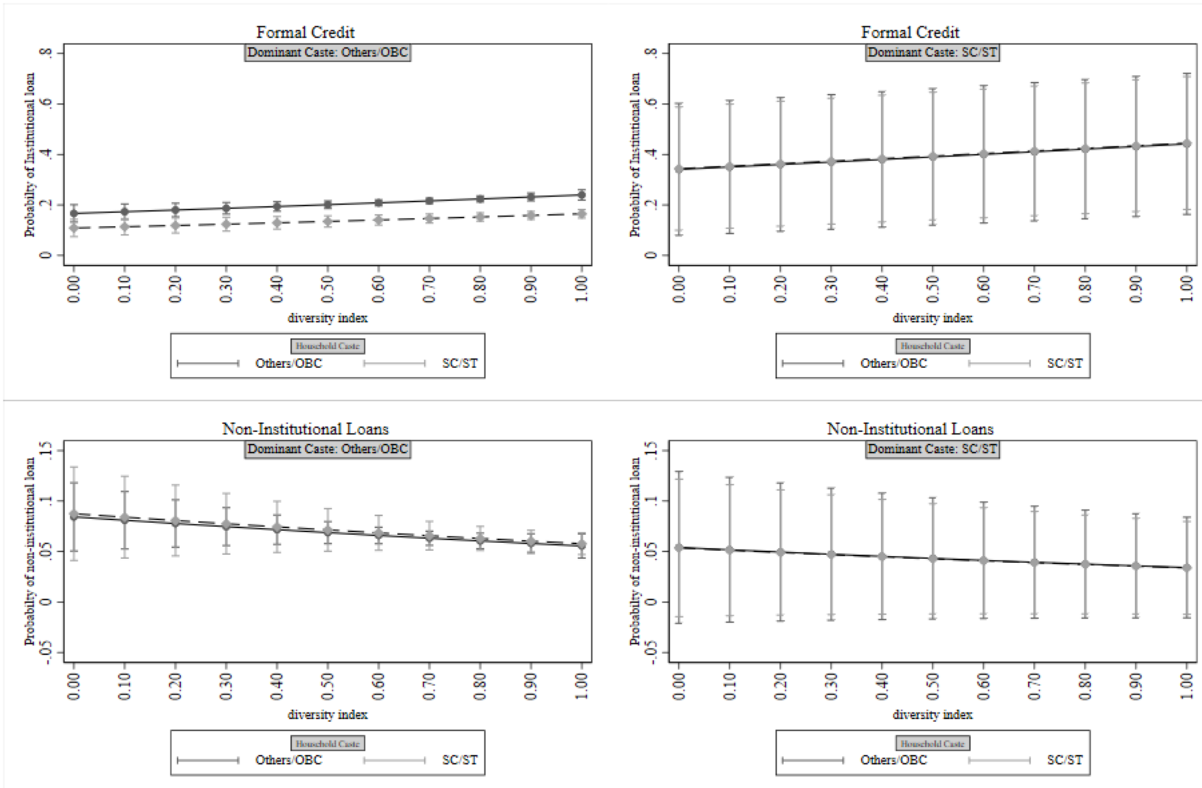


Figure4b. Predicted probability of extensive margin estimates (commercial vs cooperative)

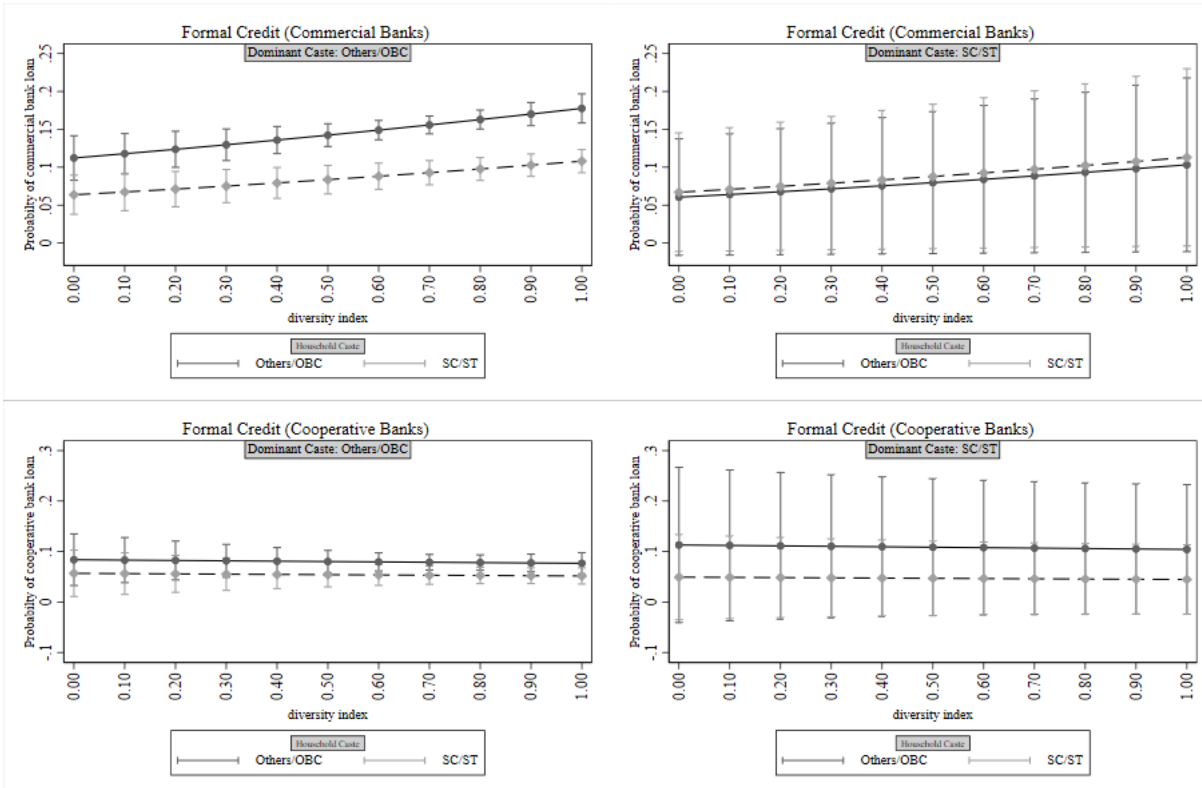


Figure 5. Predicted probability log loan amounts (formal credit in general)

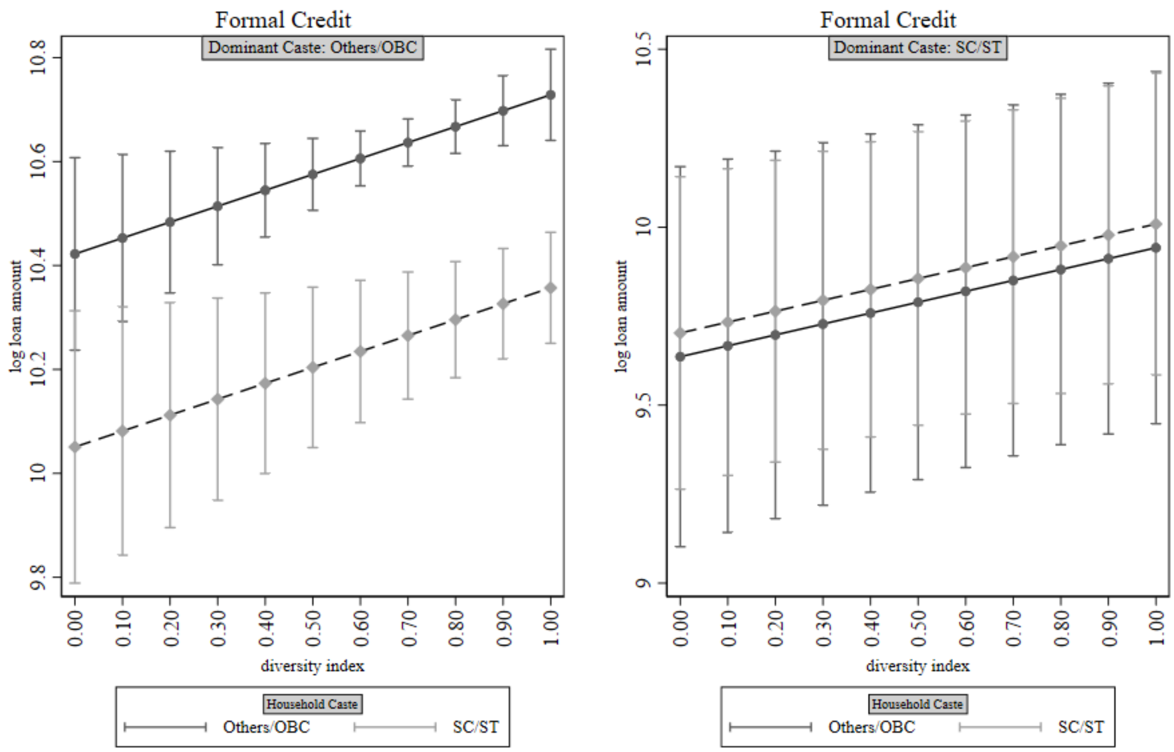


Table XVI. Oaxaca-Blinder Decomposition

	Formal Credit User (I)	Commercial Bank Loans (II)	interpretation
Differential			
Prediction_1	0.352***	0.195***	Social group= Others/OBC
	67.72	44.16	
Prediction_2	0.234***	0.119***	Social group= SC/ST
	29.14	17.78	
Difference	0.118***	0.0762***	Difference between Others and SC/ST
	12.4	9.51	
Decomposition			
Endowments	0.0298***	0.0273***	Increase in outcome of SC/ST if they had the endowments of Others
	8	5.82	
Coefficients	0.0904***	0.0445***	Increase in outcome of SC/ST if they had the coefficients of Others
	9.53	5.64	
Interaction	-0.00175	0.00439	
	0.49	0.89	
N	11241	10232	

t statistics in parentheses

* p<0.05 ** p<0.01 *** p<0.001"