# Do Social Ties Trump Collateral In Determining Loan Performance ? Evidence Using Same Day Loan Repayments *

Sumit Agarwal    Prasanna Tantri    Nitin Vishen

February 28, 2020

*Sumit Agarwal is from National University of Singapore. Prasanna Tantri and Nitin Vishen are from Indian School of business. Sumit Agarwal can be reached at ushakri@yahoo.com. Prasanna Tantri can be reached at prasanna_tantri@isb.edu. Nitin Vishen can be reached at nitin_vishen@isb.edu. We thank an anonymous financial institution for providing the necessary data. We are grateful to Indian School of Business for providing the necessary financial assistance for this project. Any remaining errors are ours.

# Do Social Ties Trump Collateral In Determining Loan Performance ? Evidence Using Same Day Loan Repayments

**Abstract**

We compare the performance of collateral based individual loans and joint liability based group loans in situations where the same individual is required to repay both the types of loans on the same day. The group loans out-perform by 10.24 percentage points. The results hold during periods of economic distress indicating co-insurance at work and relatively more for borrowers with scant hard information, indicating better monitoring by the groups. The out-performance exists even when the collateral on individual loans are relatively easily enforceable. Our results show that social ties are more potent than collateral based lending in enforcing loan contracts.

# 1 Introduction

A large financial economics literature on loan contracts has examined both collateral based individual lending and joint liability based group lending extensively.[1] Both types of loan contracts are used in economic environments where perfect monitoring is not possible. Collateral based lending to individuals works well in economic settings where borrowers have assets that can be used as collateral and enforcement of contracts is efficient (Donaldson, Gromb, and Piacentino (2019), Bae and Goyal (2009), Calomiris, Larrain, Liberti, and Sturgess (2017), Agarwal, Ben-David, and Yao (2015), Berger, Frame, and Ioannidou (2011)). Group lending is used in economic settings which do not have the above attributes (Ghatak (1999)). Superior information possessed by group members, their willingness to insure each other, and the possible reluctance on the part of group members to forgo social ties (Wydick (1999), Simmons, Tantisantiwong, et al. (2018), Beaman, Karlan, Thuysbaert, and Udry (2014), Lee and Persson (2016)) are considered substitutes for monitoring and direct enforcement by the bank. Since the two types of loans are directed at different segments of the credit markets and also analyzed by researchers in different silos, we do not yet know which of the two types of loan contracts lead to better loan performance.

Existing studies such as Giné and Karlan (2014) and Carpena, Cole, Shapiro, and Zia (2012) examine the impact of joint liability within group loans . Crucially, in these studies, both the types of loans being compared are not collateralized. Therefore, a comparison between collateralized individual loans and joint liability based group loans, in terms of their loan performance, is an open question.

An apt setting to examine this question is the one where the same individual is required to repay an individual loan and a group loan at the same time. Such a set up will be able to account for all individual level time variant and invariant characteristics, observable and

[1] See for example: Giné and Karlan (2014), Banerjee, Duflo, Glennerster, and Kinnan (2015), Berger and Udell (1990), Bharath, Dahiya, Saunders, and Srinivasan (2009), Jimenez, Salas, and Saurina (2006), Tirole (2010), Berger, Espinosa-Vega, Frame, and Miller (2011), Agarwal, Ben-David, and Yao (2015), Cerqueiro, Ongena, and Roszbach (2016), Agarwal, Murlidharan, Naman, and Tantri (2019), Ghatak and Guinnane (1999), Karlan (2007, 2005).

unobservable to the econometrician, that determine default, and hence, address the concern that group and individual loans are usually given to different types of borrowers. We use such a set up where the same individual is required to repay a group and an individual loan on the same day and compare the default rates between the two types of loans to determine which of the two types of loan contracts lead to better loan repayment behavior.

Default on an individual loan may have consequences such as liquidation of the collateral, attachment of other individual property by the bank,[2] negative impact on credit score, and reduction in access to bank finance in future. Default on group loans is likely to have all the above negative consequences except the loss of collateral and other personal property. In addition, default on group loans is likely to adversely impact social ties as other group members have to bear the burden of default due to joint liability (Karlan (2007)). Further, a defaulting group member may lose access to different forms of support such as additional loans, job referrals, and other forms of insurance from the group. Therefore, it is reasonable to hypothesize that the relative performance of the two types of loans depends on which of the two is valued higher by borrowers–individual property pledged as collateral or social ties?

We examine this question by using a loan-transaction level data that we obtain from a large non banking finance company (NBFC, henceforth) in India. The major difference between a bank and a non-banking finance company is that the later cannot accept deposits from public whereas the former can. In terms of lending technology, NBFCs are similar to a bank. The NBFC that provided the loan level data operates in three large states of India. The loans that we study are loans made to low income borrowers in rural areas for purposes ranging from agriculture to consumption. The loans are required to be repaid in equated installments on or before the due date. The lender uses both weekly and monthly repayment frequencies. Non payment of an installment in full on or before the due date is defined as default. The lender makes both collateral based individual loans and joint liability based group loans. The bank maintains a separate account for each individual in the group.

[2] Note that individual liability is unlimited in India. Even those assets that are not explicitly pledged for a loan can be attached by a lender.

Therefore, we are able to identify individual default even in group loans. In addition, we have information about time varying borrower level characteristics such as age, income, and expenses and also terms of loan such as loan amount, tenure, and interest rates.

We start our analysis by examining loan repayment instances where a single borrower has at least one group loan and one individual loan running simultaneously. In other words, every month, the borrower is required to repay installments on both types of loans. In this sample, we find that the default rate of group loans is lower by 12.57 percentage points. Next, we tighten the identification further by limiting the sample to loan repayment instances where a borrower is required to repay a group and an individual loan on the same day. Here, group loans out perform individual loans by 10.24 percentage points. Finally, to address the concern that the loans that always overlap are special, we restrict the sample to cases which satisfy both the below conditions: (i) a single borrower is required to repay a group loan and an individual loan on the same day, and (ii) the group and individual loans have different repayment frequencies so that they do not always overlap. Within these loans, we consider overlapping (same day) loan repayment instances and find that group loans continue to out-preform individual loans by 8.33 percentage points. The out-performance stated above range between 35% to 61% of the average default rate in the sample, and hence, are economically meaningful. We include borrower level and month X year level fixed effects. Thus, we account for borrower level time invariant factors and also the general time trend.

It is crucial to note that even in cases where we have variation in loan repayment frequencies between loan types within an individual, we do not conduct a difference-in-difference test because such a design will difference out the difference between group and individual loans, which is the question of interest in this study. We elaborate more on this point while discussing our empirical strategy in section 3.

Our thesis is that the borrowers value social ties more than the possible loss of collateral. The literature on collateral has shown that collateral plays a crucial role in mitigating both ex-ante (information asymmetry between borrowers and lenders) and ex-post (moral hazard) credit market inefficiencies (Berger, Frame, and Ioannidou (2011), Berger, Espinosa-Vega, Frame, and Miller (2011), Calomiris, Larrain, Liberti, and Sturgess (2017), Cerqueiro, Ongena, and Roszbach (2016), Catherine, Chaney, Huang, Sraer, and Thesmar (2017)). We now address important concerns relating to our identification strategy and interpretation of results. First, readers may contend that group loans out-perform collateral based individual loans because it is very hard to monitor and enforce collateral in emerging economies (Caballero and Krishnamurthy (2001), Menkhoff, Neuberger, and Suwanaporn (2006), Tantri (2018)) and not because borrowers value loss of social ties more than the loss of collateral. Further, it may be argued that, the main result shown in this paper will flip in cases where collateral can be taken over relatively easily by the lender. Take for instance cases where collateral is standing crop or inventory. It is very hard for the lender to keep track of crop production or inventory and relatively easy for the borrowers to tunnel out such collateral (Gopalan, Martin, and Srinivasan (2017)). In other words, the concern is that the out performance of group loans may disappear if the collateral can be easily monitored and enforced.

To test the above concern, we classify individual loans into those with strong and weak collateral. The classification is done based on the relative ease with which a lender can monitor the collateral and take over and liquidate the same in case of default. In the tightest specification, we consider only gold as strong collateral and all others as weak collateral. This is because the lender has physical possession of the gold pledged as collateral and therefore can easily liquidate it in case of default. Moreover, loan to value ratios are quite low in case of gold loans (Abraham, Chopra, and Tantri (2017)). We find that, within an individual borrower, group loans outperform individual loans even in cases where enforcement of collateral is relatively easy. Therefore, our results are not due to weak enforcement of collateral.

There could be a second concern that group and individual loans, even within a borrower, are borrowed for systematically different purposes (Attanasio, Augsburg, De Haas, Fitzsimons, and Harmgart (2015), Fischer (2011, 2013)). Suppose individual loans are borrowed for risky purposes and group loans are borrowed for relatively safe ones, then our results are likely to follow due to difference in purpose and not due to difference in loan contracts. It is crucial to note that in both group and individual loans, the liability is not source based but is unlimited. An individual is required to repay the loan using all available resources and not just using the funds from the project for which the loan was originally borrowed. Nonetheless, we address the above concern relating to difference in purpose by limiting the sample to within individual pairs of loans borrowed for the same purpose. In other words, we consider only those borrowers who borrow the individual loan and the group loan for the same purpose. Our data base lists 17 purposes ranging from agriculture to consumption. We find that the group loans outperform individual loans by 9 percentage points in the tightest of the three specifications. The results hold even in the other two specifications. For completion, we test and find that our results go through with similar magnitudes even when the purposes are different.

Third, since our main data consists of borrowers having at least one group and one individual loans simultaneously, there could be concerns about selection. To correct for the same, we first obtain data relating to all the loans lent by the lender. We first verify that the distributions of the bigger data set and our data set of simultaneous loans are similar in terms of observable borrower characteristics such as age, income, expenses, land holdings, household size, and income. We then apply the Heckman (1979) two-step correction model. The first step is to estimate a selection model which gives the probability of being selected in the sample from a larger population of borrowers. This probability is used to calculate Inverse Mills Ratio for each borrower. The second step is to include this inverse mills ratio in the main empirical specification, thus correcting for the omitted variable bias. To improve upon the probability prediction in the first stage, we also use a machine learning algorithm (boosted classification trees) to calculate the probability of selection (Mullainathan and Spiess (2017)). We show that this algorithm does a better job at predicting the probability of selection compared to the probit model. We show that the coefficient of the inverse mills ratio turns out to be statistically insignificant, indicating an absence of selection bias, regardless of using probit or machine learning algorithm in the first-stage estimation. Our main result goes through.

We next focus on the mechanism at work. Our results cannot be explained by groups doing a better job of screening (Ghatak and Guinnane (1999), Ghatak (1999)) as we make a within borrower between loan type comparison. The results can be explained by either better monitoring, state verification, and enforcement of repayment by groups or by within group mutual insurance or by both the forces working in tandem. The question is important to understand whether the joint liability feature has any role to play (Giné and Karlan (2014), Carpena, Cole, Shapiro, and Zia (2012)), or, our results can be explained fully by other features of group loans (De Quidt, Fetzer, and Ghatak (2016)). While we cannot disentangle the two types of mechanisms, we can test whether mutual insurance and better monitoring have a role to play. To this end, we examine and find that the out-performance of group loans is higher during times of economic stress. Given that the exposure of different group members to economic shocks is likely to be different, it is reasonable to conclude from the above result that group members bail each other out during distress. The result clearly shows that mutual insurance due to joint liability has a role to play in explaining our results. We also find that the out-performance of group loans is higher in cases where the lender has significantly lower level of hard information about the borrowers, indicating a role for better monitoring and enforcement within the group.

Group loan structure shifts a substantial part of the burden relating to screening, monitoring, and eventual recovery from the lender to the group members. The shift obviously leads to cost saving to the lender. Whether the lender passes on the benefit of cost savings to the borrowers is an empirical question as it depends on the credit market structure (De Quidt, Fetzer, and Ghatak (2016)). Our setting is ideal to test this question as the same borrower is offered a group and an individual loan for approximately the same term. Any difference in pricing is more likely because of the loan structure. We find that the interest rate charged on group loans is 4.76% lower. The difference is economically meaningful as it represents more than a fifth of the average interest rate in the sample. In addition even the loan amount of group loans is higher by Rupees 1,409, which is close to 8% of the average loan amount. It appears that at least a part of the cost savings due to the loan contract contract structure is passed on to the borrowers.

The difference in interest rates and loan amount between group loans and individual loans even when they are lent to the same borrower raises a question about interpretation of our results– whether the difference in loan performance is due to differences in loan terms and not because of difference in loan contract structure. It is likely that the differences in interest rates and loan amount are likely due to expected difference in loan performance and not a cause explaining the difference in loan performance. To address this concern, we first limit the sample to pairs of loans lent to the same borrower on the same day. In this case, the only material difference between the two loans is the loan type. Our results go through. Further, to address the concern that interest rates are also determined by the loan product categorization within a bank and not just by the expected loan performance, we conduct a test in which we consider cases where a borrower is required to repay installments on two individual loans, with varying interest rates and loan amounts, on the same day. We find that the differences in interest rate and loan amount do not impact the performance of such loans materially. Finally, because loan amount could be determined by either supply or demand, we restrict the sample to pairs of loans where the loan amounts of group and individual loans are equal. Our results hold in this sub-sample as well. These results show that the differences in interest rate and loan amount are a consequence of difference in loan structure and they themselves do not cause the difference in loan performance detected in this study.

We contribute to the large literature that deals with loan contract types and their relative efficacy in enforcing loan contracts (Gopalan, Mukherjee, and Singh (2016), Billett, Elkamhi, Popov, and Pungaliya (2016), Giné and Karlan (2014), Puri, Rocholl, and Steffen (2017), Stulz and Williamson (2003), Jimenez, Salas, and Saurina (2006), Ghatak and Guinnane (1999), Besley and Coate (1995)). Using same individual same day repayment group and individual loan pairs, we show that loan performance is better under a group loan structure. Note that Giné and Karlan (2014), Carpena, Cole, Shapiro, and Zia (2012) compare the performance of group loans with or without joint liability. Our focus is on comparison between collateralized individual loans and group loans with joint liability. While our same borrower-same day repayment setting allows us to make this comparison, we are not able to separate clearly the impact of joint liability from other aspects of group loans such as superior information and the power of social sanctions.

We also contribute to the literature that seeks to explore the value of social ties in loan contracts in general and group loans in particular (Lee and Persson (2016), Degryse, Lu, and Ongena (2016), Karaivanov and Kessler (2018)). We show that social ties trump collateral in their ability to instill loan repayment discipline. Finally, we also contribute to the large and growing literature that deals with peculiarities of lending in emerging markets (**?**Tantri (2018), Fisman, Paravisini, and Vig (2017), McKenzie and Woodruff (2008), Giné and Kanz (2017), Cole (2009), Khwaja and Mian (2005), **?**). We show that group loans may be helpful in lowering default rates even among those borrowers who possess collaterizable assets and hence have access to individual loans.

# 2 Institutional Details And Data

As noted in the introduction, we obtain loan transaction level data from a large NBFC in a India. NBFCs are financial institutions that perform all banking transactions except accepting public demand deposits.[3] Our data provider aims to provide diverse range of financial services to the rural poor. It operates in three states of India. The three states are located in South, East and North of the country. Figure 1 depicts the three states on the map of India. The data provider has more than 200 branches. The value of the loan book is in excess of Rupees 30 billion.

The data set contains information about both individual loans and group loans with joint liability. Even for group loans the data are recorded at an individual level and hence it is possible to see whether an individual has defaulted on a group loan or an individual loan.

[3]See this RBI circular for the full definition of NBFC.https://www.rbi.org.in/Scripts/FAQView.aspx?Id=92

The loans are repayable in equated periodical installments. The final data set is therefore organized at a loan-repayment period level. Almost all the loans have a tenure of one year. The NBFC also collects information about income and assets of the individual at regular intervals. The data set does not explicitly provide information about interest rates. We impute interest rates based on other loan terms.

## 2.1 Sample Construction

Table 1 provides information about sample construction. Our sample consists of borrowers having at least one group loan and one individual loan. There are 14151 unique individuals having two types of loans at the same time. In total, we have 36480 loans in the data-set. Out of this, 20397 are group loans and 16083 are individual loans. There are 1060553 repayment instances, out of which 825193 (235360) relate to group (individual) loans. Note that group loans have a higher proportion of loans repayable on a weekly basis. Finally, in row three, we present the data relating to loan repayments that run simultaneously but not necessarily on the same day. We consider two loans as simultaneous loans if they are active at the same time. They need not be repayable on the same day. However, given that the longest repayment frequency is a month, the maximum gap possible between loan repayment dates of two simultaneous loans is 31 days. An example is in order. Suppose a borrower borrows a group loan on July $11^{th}$, 2010 and an individual loan on October $14^{th}$, 2010, then these loans are not considered as simultaneous loans between July $11^{th}$ and October $13^{th}$. They will be considered as simultaneous loans starting from October $14^{th}$. There are 1030949 loan repayment instances where a borrower simultaneously repays a group loan and an individual loan (which we will call Sample 1), out of which 796916 (234033) relate to group (individual) loans.

In Table 2, we tighten the sample selection further. Here, we consider only those repayment instances where an individual is required to repay a group loan and an individual loan on the same day (this subsample will be referred to as Sample 2). There are 99619 such repayments. In panel A, we consider repayment instances where a borrower is required to repay a group and an individual loan and the two types of loans have different repayment frequencies. As shown in the third row, there are 27300 such loan repayment instances (and we call this Sample 3). Out of these 13704 repayment instances pertain to group loans and 13596 pertain to individual loans. These repayment instances pertain to 9984 group loans and 8919 individual loans borrowed by 8362 unique individuals. Notice that the number of loans and repayment instances for group loans and individual loans are not exactly the same although the sample is restricted to same day repayments. This is because in some cases a single borrower has more than one simultaneously running group loans. The number of borrowers is the same by construction. In Panels B (C), we consider loan repayments to be made on the same day where both group and individual loans are repayable on a weekly (monthly) basis. In Panel B (C), we find 26194 (46186) such instances. These repayment instances pertain to 323 (1988) group loans and 316 (1995) individual loans borrowed by 313 (1835) unique individuals.

## 2.2 Key Variable Definition

### 2.2.1 Equated Periodical Installments

The loans are repayable on an annuity immediate basis where the first installment is paid at the beginning of a period. In other words, the first installment amount which is payable as soon as the loan is disbursed is the equated periodical installment to be repaid though out the loan tenure. There is no separate information about the equated payments to be made in the data set. Therefore, we impute the equated payment based on the first installment. There is a possibility that we underestimate the equated periodical installment in some rare cases where the borrower does not repay the first installment in full. We end up underestimating default in such cases. However, there is no reason to believe that a systematic difference exists between a pair of group and individual loans borrowed by the same borrower and hence, our results are unlikely to be biased because of the above estimation procedure used for calculating periodic payments to be made.

### 2.2.2 Calculation of Interest Rates

The database does not contain direct information on interest rate. So, we impute the same from the repayment records. We use the equated periodical installment amount, initial loan amount and loan tenure to impute interest rate using the annuity immediate formula.

### 2.2.3 Default

The data has loans which are to be repaid in periodical installments. We have the date on which a loan was issued. It is reasonable to assume that the due date for monthly repayments is on the same corresponding date of the following months. For example, consider a loan which was issued on January 19th, 2010; then the first installment is assumed to be due on February 19th, 2010 and so on till the date of maturity of the loan. Similarly, weekly due dates are constructed by adding seven day intervals to the date of issue of a loan. The loan repayment ledger available to us records a repayment whenever it was made. These repayments are not recorded against the particular installment which was due for the given month or a week. The data, in some cases, has records of borrowers making multiple or no repayments in any given month. To calculate timely installment repayment for a loan repayable at monthly (weekly) intervals, we sum up all the repayments in a given month (week) that was made till the due date. A particular installment is categorized as a default when the repayment, calculated as described above, is less than 90% of the amount which is due. We use 90% to allow for any calculation errors. We test the robustness of our results using the 100% threshold as well. As a further robustness, we use a second measure of loan performance– classification of a loan as a non performing asset (NPA). Generally, a loan outstanding for more than 90 days is considered an NPA (Giné and Kanz (2017)). NPA classification triggers higher provisioning.

## 2.3 Summary Statistics

Table 3 provides distribution of key variables in the sample. The average default rate is 23.55%. The average interest rate charged is 22.5%. Such high interest rate is not surprising given that the borrowers are small and risky. The loan amount has an average (median) value of Rupees 18065 (20000) and even its distribution seems to be reasonably well behaved. The loan amount is similar to loan amounts documented in other studies on India such as Tantri (2018), Mukherjee, Subramanian, and Tantri (2018). Almost all loans have a tenure of one year. The actual tenure varies a little bit due to early and delayed repayments. Median age of the borrowers is 40. The average (median) income of Rupees 13156 (10000) is close to the national average during the early part of the decade starting from the year 2010 (Cherodian and Thirlwall (2015)).

# 3 Empirical Strategy

The early banking theory (Diamond (1984)) recognized that banks, unlike individual lenders, can monitor the borrowers and hence enforce repayment discipline. It was soon recognized that such monitoring, at best, can be imperfect and hence, the need for collateral in loan contracts arose (Tirole (2010), Calomiris, Larrain, Liberti, and Sturgess (2017), Agarwal, Ben-David, and Yao (2015)). There are three issues with collateral based lending. First, not all borrowers having positive NPV projects, especially those in emerging markets, have pledgeable collateral (Donaldson, Gromb, and Piacentino (2019)). Information asymmetry between borrowers and the lenders and inability of the borrowers to credibly signal their type leads to a significant difference between the true value of a project and its pledgeable value. Second, weak contract enforcement mechanism and frequent political interventions in debt contracts (Tantri (2018), Mukherjee, Subramanian, and Tantri (2018)) make enforcement of collateral extremely difficult in emerging economies. Finally, recent studies have shown that the presence of collateral leads to some unintended consequences. For instance, Agarwal, Ben-David, and Yao (2015) show that agency consideration within the bank lead to manipulation of the stated value of the collateral. Cerqueiro, Ongena, and Roszbach (2016) show that the presence of collateral leads to reduced monitoring and excessive risk taking.

The presence of the above mentioned failures in collateral based lending lead to the emergence of group lending with joint liability. Here, a group of borrowers approach a lender for a group loan with joint liability. Although each individual group member is given a separate and identifiable loan, the entire group is collectively and severally responsible for all loans lent to the members of the group. The mutual insurance given by group members substitutes for collateral in these loan contracts. In other words, in the event of default by a borrower, the lender has recourse to other members of the group and their personal assets. The group loan structure effectively leverages the social capital present in emerging economies and uses it as a substitute for collateral. As noted by Ghatak (1999), banking functions such as selection, monitoring, state verification, and enforcement of repayment are the responsibilities of the group in group loan contracts. In fact, given the close proximity and the social capital, group members are likely to be in a better position to discharge these functions than the bank. Apart from enforcing joint liability, the bank threatens to cut off future lending to the entire group in case of default by a member (Breza (2012)). Thus, the bank is no longer solely dependent on the formal contract enforcement mechanism.

It is clear from the above discussion that a simple comparison between default rates of group loans and individual loans does not clear the identification bar for determining which of the two loan systems lead to better loan performance. The borrowers of two types of loans are likely to be distinct and hence subject to different kinds of shocks. Therefore, a difference in loan outcomes cannot be attributed to the difference in loan structure alone. The above criticism holds even if both the types of loans are lent by the same lender as the differences are at a borrower level.

Therefore, a good setting to evaluate the relative loan performance of group and individual loans is when (i) the same borrower borrows a group loan and an individual loan from the same lender (ii) the loan installments are repayable at the same time. The design effectively neutralizes both time invariant and time varying differences at the borrowers level as the comparison is within borrower between loan contract types and at the same time.

On the face of it, it may appear that a difference-in-difference strategy using the differences between group and individual loans when they overlap and when they do not overlap is a reasonable way of testing the difference between group and individual loans in terms of their performance. Given that repayment frequencies of loans vary, leading to a pair of within borrower loans overlapping only occasionally, the above strategy is actually implementable. However, when examined carefully, it becomes clear that the difference-in-difference strategy is not helpful in identifying the difference between group and individual loans as this difference itself is differenced out and what is left is the incremental difference due to overlap. In other words, our purpose is not to answer the question- what happens to the difference between loan performance of group and individual loans when both the loans are required to be repaid on the same day when compared to a situation when the two types of loans are required to be repaid with a gap of, say, 3 days. We are interested in the first difference–between group and individual loans–itself and hence cannot use the above difference-in-difference strategy.

The above discussion sets the stage for the description of our empirical strategy. As noted before, each observation represents a loan repayment instance. The frequency could be either monthly or weekly. As described in section 2.3, we start our analysis by considering cases where the same borrower has at least one group loan and one individual loan running simultaneously. We then restrict the the sample to loan repayment instances where a borrower is required to repay a group loan and an individual loan on the same day. Finally, in the tightest specification, we consider only those cases where a borrower has at least one group loan and at least one individual loan and the two types of loans have different repayment frequencies. These are loans which do not always overlap and hence help us rule out any residual selection related concerns. For convenience, we call these samples sample 1, 2, and 3, respectively, henceforth. We depict the three sub-samples in figure 2.

As we run our analysis on borrowers who have a group loan and an individual loan running at the same time, a concern arises that this analysis could be limited only to this specific sub-section of the population. We address concerns relating to sample selection bias by treating it as a specification error (Heckman (1979)). We use the Heckman two-step method, where the first stage is estimating a selection equation using a probit model, and the second stage is running our main empirical regression after including the inverse-mills ratio calculated from stage one as an additional regressor. Additionally, we employ a machine-learning algorithm (boosted decision trees) to replace the probit in first-stage of Heckman correction to get better prediction on the probability of being selected into the sample (Mullainathan and Spiess (2017)).

A discussion on the control variables used is in order here (Spector and Brannick (2011)). As described above, we compare loan repayment instances of at least one group loan and at least one individual loan borrowed by the same individual and repayable at the same time. The research design itself accounts for all time invariant and time variant individual level characteristics as they are the same through out for both types of loans. Therefore, we first present our results without using any time varying borrower level characteristics as control variables. Nonetheless, to address the concern that our results are only due to special situations such as a shock to the assets or income of the borrower, we explicitly account for time varying borrower characteristics such as the income, expense, assets (landholding), number of persons in the borrower household and borrower age. In other words, we test our results with and without these control variables. Finally, to address a concern that our results are due to some special type of borrowers, we include borrower level fixed effects and to account for time related trend effects, we include month X year fixed effects. We cluster the errors at a borrower level as borrowers are heterogeneous in terms of occupation and social background.

It is also important to note that we do not include terms of loan contracts such as interest rates, loan amount, and tenure as control variables. As noted by theory (Ghatak (1999)), these variables are likely to be consequences of difference in loan structure and not a cause. For instance, since in group loans the burden of selection, monitoring, state verification, and collection is borne by the group itself, the bank saves costs on these fronts. Banks are likely to pass on these benefits for group loans. Similarly, banks may be willing to lend higher loan amount for group loans given the joint liability feature. Given these theoretical arguments, we use typical loan contract terms as dependent variables in different tests and not as control variables in our main test.

# 4 Results

## 4.1 Main Result

Given the empirical set-up described in section 3, it is possible to reliably estimate which of the two loan contract types induces higher loan repayment discipline. If individual collateral based lending is more potent, then we expect to see a lower default rate on individual loans. If, on the other hand, social ties matter more than individual collateral, then we expect the opposite result.

We present the univariate results in figure 3. We depict the default rates of group and individual loans within samples 1, 2, and 3. We find that group loans out-perform individual loans by 12.36 percentage points,10.26 percentage points, and 8.48 percentage points in samples 1,2, and 3.

We then move to formal regression based tests. We estimate the following regression equation to resolve the question empirically:

$$Y_{\mathrm{itj}} = \alpha + \nu_i + \gamma_j + \beta_1 * \mathrm{Group}_{itj} + \beta_2 * \mathrm{Borrower\ Characteristics}_{ij} + \varepsilon_{itj} \tag{1}$$

The data are organized at a borrower $i$-loan $t$-repayment frequency $j$ level. As noted before, loans have different repayment frequencies. The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Group*, is a dummy variable that takes the value of one if the loan under consideration is a group loan and zero otherwise. Borrower characteristics include borrower's income, expense, assets (landholding), age, and number of members in the borrower household. $\nu_i$ stands for borrower fixed effects and $\gamma_j$ stands for $monthXyear$ fixed effects.

The results are presented in Table 4. We consider sample 1 in columns 1 and 4, sample 2 in columns 2 and 5, sample 3 in columns 3 and 6. We include time varying borrower characteristics as controls in columns 3 to 6. The results show that group loans default less than individual loans by between 8.33 percentage points to 14.32 percentage points, depending on the sample and the specification used. The co-efficient of interest is significant at conventional statistical levels. Notice that the economic magnitudes of the coefficients is quite close to the univariate difference between group and individual loans in each of the three sub samples. Also, notice that inclusion of borrower level control variables does not move the coefficients beyond couple of percentage points. Moreover, given that the average default rate in the sample is 23.55%, the out-performance of group loans over comparable individual loans works out to anywhere between 35% to 61% of the average default rate. Therefore, the out performance is highly economically meaningful. We also repeat this test after replacing the measure of default with Non-Performing Asset (NPA). A loan account is marked NPA when the repayment is due for at least ninety days. This is shown in the Table A11 of the online appendix.

The bottom line can be summarized in a straight forward way: a group loan structure does lead to higher repayment rates even after accounting for differences in individual characteristics and time trends.

### 4.1.1 Difference in Purpose

A skeptic may argue that the difference in loan performance may be attributed to difference in purpose for which the loans are borrowed. The concern is important due to a debate regarding whether joint liability leads to higher or lower risk taking and also influences the use of funds in any other way (Attanasio, Augsburg, De Haas, Fitzsimons, and Harmgart

(2015), Fischer (2011, 2013)). If group loans are systematically deployed in safer avenues and individual loans are used for relatively riskier projects, then the results may ensue because of the nature of the purpose for which the funds are utilized and not because of difference in contract types. It is important to note that since our analysis is within a borrower and at a point in time, such a systematic difference is unlikely to exist.

Nonetheless, to address any residual concern in this regard, we collect information about the broad purpose for which a loan was borrowed. The data base lists 17 purposes. The purpose for which a loan is borrowed includes purposes such as agriculture, business, consumption among others.[4] To address the concern stated above, we restrict the sample of loans where a borrower has borrowed group and individual loans for the same purpose. We hypothesize that there is unlikely to be systematic difference in the riskiness of the two types of loans in this sub-sample.

We estimate regression equation 1 on the sub-sample where the purpose of all loans borrowed by a borrower is the same. We report the results in Panel A of Table 5. We find that group loans out perform individual loans by between 8.15 percentage points to 14.94 percentage points for different specifications used. Notice that the economic magnitude of the out performance is not very different from our main result. It is clear from the above result that the relative out performance of group loans is not because of difference in purpose. For completion, we also separately consider within borrower loan repayment instances where the purpose of group and individual loans are different. We present these results in Panel B of Table 5. Here, too our results go through.

#### 4.1.2 Difference in Loan Repayment Frequency

As described in Section 2.1, our sample consists of loans having both monthly and weekly repayment frequencies. However, in Table 2, we also note that group loans are more likely to have weekly loan repayment frequency. There could be a concern that the difference in performance is due to a difference in loan repayment frequency. It is possible to argue

[4] A list of purposes is provided in Table A.2 of the online appendix.

that bank obtains information about the borrower relatively quickly in case of loans that have high frequency repayments than those that have low repayment frequency. Such quick information, it may be argued, could lead to better and timely collection efforts, leading to better loan enforcement. In this context, it is crucial to note that since both the loans belong to the same borrower, a bank can use the information that it obtains from the high frequency loan for the other low frequency loan as well. Therefore, difference in repayment frequency should not matter much in our setting.

Nonetheless, we test whether the difference in repayment frequency leads to difference in loan performance. We start with sample 1. We estimate regression equation (1) using a sub set from sample 1 where the repayment frequency of the individual loan and the group loan belonging to a borrower is different. We then estimate regression equation (1) using a sub set from sample 1 where the repayment frequency of the individual loan and the group loan belonging to a borrower is the same. We report these results in panel A of Table 6. In columns 1 and 4, the within borrower pair of loans have different frequencies. In panel 2 and 5 (3 and 6), both the type of loans have a weekly (monthly) frequency. As shown in the table, our results go through with similar magnitudes in both the sub samples. For completion, we repeat the above test using sample 2 and report the results in Panel B of Table 6. All other details are same as in Panel A. The results go through here as well. Given the above results, it is reasonable to conclude that the our main result is not caused by systematic difference in repayment frequency of the two type of loans we consider.

## 4.2 Enforceability of Collateral

Our thesis is that the joint liability based group loan structure leads to better loan performance when compared to the collateral based individual loans. However, the effectiveness of collateral based lending heavily depends (i) on enforceability of collateral in case of default and (ii) realizable value of the collateral in case of liquidation.

The enforceability of collateral depends on the overall efficiency of the contract enforcement apparatus in a country. It is well known that contract enforcement is ineffective and slow in many emerging markets, including India (Mukherjee, Subramanian, and Tantri (2018)). In fact, despite significant improvements in overall ease of doing business rankings,[5] India is still ranked a low 164 in terms of speed of enforcement of contracts. Not surprisingly, relationship banking is the preferred mode of banking (Bhue, Prabhala, and Tantri (2015), Vig (2013)) and overall loan delinquency rate in the economy is very high despite there being stringent de-jure creditor rights.

The realizable value of an asset is a function of its nature. Take for instance a case where the collateral is a specialized business asset such as a custom made machine to manufacture specialized products. Only a limited number of buyers are likely to be interested in such machines and hence a bank is likely to receive only a small portion of book value in case it decides to seize and liquidate collateral after default. The realizable value is also a function of a lender's ability to monitor collateral and borrowers' corporate governance practices. Consider a case where the collateral is stock in trade. It is possible that a borrower tunnels (Johnson, La Porta, Lopez-de Silanes, and Shleifer (2000), Atanasov, Black, Ciccotello, and Gyoshev (2010)) out a part of such collateral after pledging them. It is extremely hard for the lender to monitor such collateral.

In this context, it is reasonable to ask whether our results really reflect the triumph of social ties over collateral or are the results due to weak enforceability of collateral. In other words, one may argue that it is not clear whether our results will hold if the collateral was strongly enforceable, and hence, we cannot reasonably claim that social ties are stronger than collateral.

To address this concern, we identify cases where the collateral is likely to be "strong". As a starting point, we consider land and gold as "strong"collateral. We include land as it is very difficult to tunnel it out and also because it is a general asset and not specialized one. Gold loans are a typical Indian phenomenon where jewelery is pledged as a collateral

[5] India is now ranked $77^{th}$ out of 190 countries in the World Bank's ease of doing business rankings. The country has shown an improvement of 65 positions in the last four years. Source:http://pib.nic.in/newsite/PrintRelease.aspx?relid=184513

for loans. Here, the collateral remains in the custody of the lender until the loan is fully repaid. In case of default, the lender can directly liquidate the jewelery pledged without any interference from courts or the government. More importantly, loan to value ratios are low and gold prices are also relatively more stable when compared to other assets used as collateral. Therefore, chances of a lender enforcing collateral and recovering loans is very high in gold loans.

We estimate regression equation (1) by considering only those borrowers whose individual loans are borrowed using either gold or land as collateral. In other words, the sample consists of borrowers who have at least one group loan and at least one individual loans and the collateral for individual loans is either gold or land. We report the results in Panel A of Table 7. The organization of the table mimics the organization of Table 4. The group loans out perform individual loans with strong collateral and lent to the same borrower by between 12.35 to 24.57 percentage points. We tighten the specification further by considering only gold as "strong"collateral. There could be issues with enforcing loan contracts where land is the collateral. There are instances of political interference (Mukherjee, Subramanian, and Tantri (2018), Tantri (2018), **?**), Giné and Kanz (2017)), especially in case of agricultural loans. Also, land ownership is poorly delineated in large number of cases, leading to litigation and consequent delay in enforcement. Gold loans do not face these issues as the physical possession is with the lender. We report the results in Panel B of Table 7. Other details relating to sample selection and organization of the table remain unchanged. We find that group loans out perform individual loans even in this sub-sample. Therefore, it is reasonable to conclude that our main results are not due to week enforcement of collateral. They hold even when collateral is enforceable with ease and without much loss of value in liquidation.

For completion, we estimate regression equation (1) by considering cases where the collateral is not strong. We report the results in Panels A and B of Table A.3 of the online appendix. Our results go through here as well. Finally, we perform a horse race between within borrower group loans, individual loans with strong collateral and individual loans with not so strong collateral. We report the results in Panels A and B of Table A.4. We find that group loans out perform individual loans with strong collateral by a wider margin when compared to individual loans with not so strong collateral. The results hold with both definitions of strong collateral.

## 4.3 Plausible Channel

Here, we attempt to understand why do group loans out perform individual loans even when we compare within the borrower and within loan type. As noted in the Introduction, the difference in the degree of adverse selection or the group's superior ability to overcome adverse selection when compared to the bank (Ghatak (1999)) cannot explain our results as the comparison is within the same borrower and at the same time. Two possible explanations emerge. (i) Groups have a superior ability to monitor and enforce loan contracts (ii) There is mutual insurance at work. Due to joint liability, group members bail out each other. While it is extremely hard to disentangle the impact of the above two factors as they work in tandem, we attempt to examine whether mutual insurance has any role to play.

### 4.3.1 Co-Insurance

Whether joint liability impacts loan performance is a question that is being currently debated in the literature (Giné and Karlan (2014), Carpena, Cole, Shapiro, and Zia (2012))). In fact, may micro-finance lenders, including the famous Grameen bank of Bangaladesh, are introducing group loan products without joint liability. We test whether the mutual insurance induced by the joint liability feature contributes to improvement in loan performance. To this end, we examine whether the out performance of group loans increases during times of economic distress. Mutual insurance is likely to be activated in such times and hence any increase in the relative out performance of group loans during times of economic distress is likely to indicate that mutual insurance does play a role in enhancing loan performance of group loans. No incremental change during times of economic stress will lead to an interpretation similar to that of Giné and Karlan (2014) that joint liability feature does not improve loan performance.

To test the above thesis, we estimate the following regression equation.

$$\begin{aligned} \mathrm{Y_{itj}} &= \alpha + \nu_i + \gamma_j + \beta_1 * \mathrm{Group}_{itj} + \beta_2 * \mathrm{Shock}_{itj} \\ &\quad + \beta_2 * \mathrm{Group}_{itj} * Shock_{itj} + \beta_4 * \text{Borrower Characteristics}_{it} + \varepsilon_{itj} \end{aligned} \tag{2}$$

Here, shock is a dummy variable that takes the value of one if the local area in which the borrower resides undergoes an economic slowdown. All other terms have the same meaning as in equation 1. High frequency data relating to economic performance at a local area level is not available for India (Agarwal, Prasad, Sharma, and Tantri (2018)). Given this constraint, we use the night lights data at the district level to measure economic performance (Henderson, Storeygard, and Weil (2011)).[6]

In our main specification, we consider a 20% month on month decline in nigh lights as economic shock. As a first stage exercise, we test the relationship between our definition of loan performance and economic shock in general. We find that an economic shock, as defined by us, leads to higher default rates in general. We report this result in Table A.5 of the online appendix. For robustness, we vary the threshold for economic shocks to 15% and 10% and find directionally similar results. We report these results in Tables A.6 and A.7 of the online appendix.

With this background, we estimate regression equation (2) and report the results in Table 8. Notice that, in general, group loans out perform individual loans by between 8.72 percentage points to 19.32 percentage points, depending on the specification used. The interaction term between *Group* and *Shock* dummies shows that the relative out performance of group loans over individual loans increases, depending on the specification used, by between 2.76 to 7.4 percentage points. Also, notice that the *shock* dummy has a positive and

[6] We obtain this data from http://india.nightlights.io/nation/2006/12.The data for India are available for years 2009 to 2013.

significant value through out, indicating an increase in default rates, in general, during times of economic stress. This further validates the first stage results presented in Table A.5 of the online appendix. As a robustness exercise, we estimate regression equation (2) by using 15% and 10% month on month decline in night lights as thresholds for categorizing an area as distressed. The results are directionally similar. We present these results in Tables A.8 and A.9 of the online appendix.

### 4.3.2 Monitoring

To test whether superior monitoring and enforcement within groups also plays a role, we use the well known fact that monitoring plays a key role in situations where crucial information about borrowers is not easily verifiable (**??**Fisman, Paravisini, and Vig (2017)). We use the fact the borrowers having a income above the threshold of Rupees. 500,000 are more likely to file a income tax return and are also more likely to be closely monitored by the tax department.[7] Therefore, more credible information is available about such borrowers when compared to those below the threshold. We call borrowers above the threshold as high information borrowers. We hypothesize that, if monitoring plays a role, then the out-performance of group loans should be higher among borrowers who are less likely to file tax returns. Non verifiable information is likely to play a dominant role in such cases.

We estimate a regression equation similar to equation (2) and report the results in Table 9. $HighInformation$ is a dummy variable that takes the value of one for borrowers who are likely to file tax returns and zero for others. We use sample 2 for this test. The results are similar with samples 1 and 3 as well. In columns 1 and 4, we consider all instances when a borrower is required to repay a group and individual loan on the same day. To address the concern that tax filers are likely to be systematically larger and our results are therefore likely to be driven by reasons related to size, in columns 3 and 5 we restrict the sample to borrowers having an income within a bandwidth of 250,000 around the threshold. In

[7] Income upto Rupees 250,000 is exempt from tax. However, tax payers are allowed exemptions and deductions of another Rupees 250,000. Source:https://www.hdfclife.com/insurance-knowledge-centre/tax-saving-insurance/latest-income-tax-slab-and-deductions-fy-2014-15

columns 3 and 5, we further shrink the threshold to Rupees 100,000.

Notice that while the coefficient related to group dummy is negative and significant, the interaction term is positive. In other words, the out-performance of group loans is higher in cases where the borrower is less likely to file a tax return. As discussed above, in the absence of credible verifiable information, better monitoring, state verification, and enforcement by the groups seem to be leading to this out-performance.

Given the results presented in Tables 8 and 9, it is reasonable to conclude that both co-insurance and better monitoring aided by superior information play a role in the out-performance of group loans. However, we are unable to disentangle the impact of the two channels.

## 4.4 Loan Terms

Group loans shift the burden of selection, monitoring, state verification, and enforcement from the bank to the group members (Ghatak (1999)), resulting in significant cost savings for the bank. In case of individual loans, even when they are collateralized, the bank is required to spend resources on the above mentioned activities. Whether the cost savings get passed on to the borrowers is an empirical question as it crucially depends on the competitive structure of the banking industry in the local area.

We test whether the same borrower is charged a lower interest rate for a group loan when compared to an individual loans repayable at the same time. We estimate regression equation (1) using interest rate at a loan-repayment level as the dependent variable. All other details remain the same as before. We report the results in Table 10. We find that group loans are charged, depending on the specification used, by between 3.26% to 6% lower than individual loans in terms of interest rate. The reduction represents between 14% to 26% of the average interest rate and hence, is economically meaningful.

We next examine whether the same borrower obtains larger loans under the group structure when compared to collateral based individual loan structure. If the lender perceives group loan structure as more potent in enforcing a loan when compared to collateral, then it is likely that the loan amount is higher under group loans. We estimate regression equation (1) using loan amount as the dependent variable. We report the results in Table 11. We find that the loan amount is higher, depending on the specification used, by between 1410 to 15379 depending on the specification used.

The results show that at least a part of the savings to the banker due to the group loan structure is passed on to the borrowers.

#### 4.4.1 Impact of Loan Terms On Main Results

There could be a concern that the difference in loan terms pointed out above cause the main result we document in this study. In other words, a skeptic may contend that the group loans out perform individual loans lent to the same borrower and repayable at the same time because they are larger and cheaper. Implicit is an assumption in this criticism that even within a borrower, loan terms cause borrowers prefer to maintain better performance on cheaper and larger loans to avoid being cut off from such loans in future. The criticism is not applicable to our setting as both the borrower and the lender of both types of loans is the same. A default on any one of the two loans, irrespective of the terms, will cause the same damage to the borrower's prospect of obtaining a future loan under both types of loans. The lender considers past performance on group loans while lending individual loans and vice versa. It is more likely that the difference in loan terms reflect endogenously the difference in expected loan performance of the two type of loans. In Table A.12 of the online appendix, we restrict the sample to pairs of loans lent to the same borrower on the same day and find that the results replicate. It is unlikely that factors other than expected default rate are at work here.

Nonetheless, to address the concern that difference in loan terms are due to some exogenous reasons such as product features and not due to expected default rate, we consider cases where a borrower has two simultaneous individual loans. We create three sub samples as before. We estimate the following regression equation.

$$Y_{\mathrm{itj}} \quad = \quad \alpha + \nu_i + \gamma_j + \beta_1 * \text{Loan Terms}_{itj} + \varepsilon_{itj} \tag{3}$$

Here, loan terms refer to interest rate, loan amount and loan tenure. All the other terms have the same meaning as before. We report the results in Table 12. We find that the difference in interest rates have no significant impact on default rates. Even the loan amount has an extremely economically weak relationship with default. In the tightest of the three samples, the coefficient of interest is even statistically insignificant. Even in one case, where we find statistical significance, the economic impact is too low.In samples 1 and 2, the results show that a 1,000,000 Rupees change in loan amount is associated with 1% increase in default rates. The economic magnitude is too small to cause our main result given that the average loan amount is 18,064 Rupees.

Given the slight difference in loan amount, we conduct a further robustness test. We limit the sample to pairs where a within borrower group and an individual loan have the same loan amount and estimate regression equation 1. We find that our main results replicate with similar magnitudes in this sub-sample. We report the results in Table A.10 of the online appendix.

Given the above results, it is reasonable to conclude that the difference in loan terms do not cause difference in loan performance between the two types of loans.

## 4.5 Correction for Selection Bias

The sample on which we run our empirical tests contains only those borrowers who have one individual and one group loan, running at the same time. As these borrowers were non-randomly selected, the results of corresponding tests can thus have selection bias. Even though we cannot find an exogenous event which leads to such a sample selection, we try to correct for the bias using the Heckman two-step method. Heckman (1979) showed that such an issue can be treated as an omitted variable bias where the omitted variable is the inverse-mills ratio calculated using a selection equation. The first stage of this method is a probit equation which gives the probability of selection of a borrower into the sample. We use information of additional 166489 borrowers who have a group or individual loan with monthly or weekly repayment schedule. These additional borrowers do not have one individual and one group loan running at the same time. We use these borrowers combined with the selected sample as the population to estimate the probability of selection into our sample. The regressors in first-stage are the same as the ones used in other tests, along with two other borrower characteristics, *viz.*, borrower's profession and location.

After running the first-stage, we obtain the probability of being selected into the sample for each borrower and use it to calculate the Inverse Mills Ratio (IMR) for each borrower. This ratio goes as an additional control variable into the empirical specification in equation (1).

$$Y_{\mathrm{itj}} = \alpha + \nu_i + \gamma_j + \beta_1 \mathrm{Group}_{ij} + \beta_2 \mathrm{Borrower\ Characteristics}_{itj} + \beta_3 IMR_{ij} + \varepsilon_{itj} \quad (4)$$

We find the coefficient of Inverse Mills Ratio ($\beta_3$) is statistically insignificant and hence, conclude that there is no selection bias in our results, under assumptions of Heckman (1979). The results are shown in Table 13. Here, the first column shows the coefficients of the probit model. Column 2 and 3 show the result of equation (4) with and without time varying borrower characteristics, respectively. Similarly, columns 4 and 5 show the same results with Inverse Mills Ratio calculated when the stage one estimation is done using machine-learning algorithm (boosted decision trees). Our main result goes through.

#### 4.5.1 The Use of Machine Learning For First stage

In the first stage above, we also used a gradient boosted classification trees to calculate the probability of being selected into a sample (Mullainathan and Spiess (2017)). The outcome variable of this supervised learning algorithm is "Selected", which takes a value of one for the observations which are a part of our sample and zero otherwise. Then we use seven borrower level characteristics like household size, income, expense, land, age, profession and location to train the algorithm. We use Gradient Boosted Classification Trees, and run a five-fold cross-validation on twenty percent of randomly chosen observations (Kleinberg, Lakkaraju, Leskovec, Ludwig, and Mullainathan (2017)). Rest of the observations, we use for out-of-sample verification. We plot the actual frequency of selection corresponding to the probability calculated by the algorithm. Graphically, it seems that the machine learning algorithm does a better job compared to the probit model in predicting probability of being selected into the sample, which is displayed in Figure 4. This is despite the fact that the area under the curve of the Receiver Operating Characteristic (ROC) curve is quite high for both the probit model (0.96) and the machine-learning algorithm (0.98).[8]

#### 4.5.2 Comparison of Distributions

In figures 5 to 9, we compare the entire distribution of the overall loan book and the distribution of our sample consisting of borrowers having simultaneous loans. We compare the two distributions based on variables such as age, income, expenses, house size, and land holing. Prima facie, the distributions seem similar. Therefore, it is unlikely that our main sample is "special".

[8]Another thing to note here is that as profession and location are text, we use vector representations of these words as control variables (Mikolov, Sutskever, Chen, Corrado, and Dean (2013)). Using this algorithm, we find two vectors Profession1 and Profession2 which represent profession and District which represents location of the borrowers.

# 5 Conclusion

A lender's inability to monitor the borrowers led to collateral based lending and the difficulties in enforcing collateral and realizing value, that arose mostly in emerging economies, led to group lending with joint liability. Although, these two loan contract types are prevalent in many economies, they have not been compared in terms of their ability to enforce loan repayment discipline. Empirically, such a comparison is difficult as group loans with joint liability and collateral based individual loans are made to different type of individuals in different locations.

We overcome the above identification problem by comparing the loan performance of group and individual loans lent to the same individual and repayable at the same time. We obtain loan transaction level data from a NBFC in India. The data contains instances where an individual is required to repay a group loan and an individual collateral based loan on the same day.

Using the above set up, we find that among such pair of loans, group loans out-perform in terms of default rates. We hypothesize that the strength of social ties trumps enforceability of collateral in its impact on loan performance. Further, the results hold even when collateral on individual loans are relatively easily enforceable. We then examine whether the relative out-performance of group loans changes during times when borrower faces economic distress. If group loans are seen as insurance during times of distress, the out performance should increase. We find that group loans out-perform even more during times of economic distress. The results hold irrespective of the purpose for which the loans are borrowed.

Our findings show that social ties have a stronger impact than collateral in enforcing loan repayment discipline even among borrowers who have access to bank finance. Given the above findings, it is reasonable to infer that group loans play a crucial role in expanding access to finance in emerging economies.

# References


Abraham, J. K., Y. Chopra, and P. L. Tantri (2017): "In-kind transfers as safety nets for the poor: Do they expand or contract during economic shocks? evidence from india," *Evidence from India (October 11, 2017)*.

Agarwal, S., I. Ben-David, and V. Yao (2015): "Collateral valuation and borrower financial constraints: Evidence from the residential real estate market," *Management Science*, 61(9), 2220–2240.

Agarwal, S., A. Murlidharan, N. Naman, and P. Tantri (2019): "Information V/s Diversification: What leads to better loan Outcome ?," *Working Paper*.

Agarwal, S., S. P. Prasad, N. Sharma, and P. L. Tantri (2018): "A Friend Indeed: Does The Use of Biometric Digital Identity Make Welfare Programs Counter Cyclical?," .

Atanasov, V., B. Black, C. Ciccotello, and S. Gyoshev (2010): "How does law affect finance? An examination of equity tunneling in Bulgaria," *Journal of Financial Economics*, 96(1), 155–173.

Attanasio, O., B. Augsburg, R. De Haas, E. Fitzsimons, and H. Harmgart (2015): "The impacts of microfinance: Evidence from joint-liability lending in Mongolia," *American Economic Journal: Applied Economics*, 7(1), 90–122.

Bae, K.-H., and V. K. Goyal (2009): "Creditor rights, enforcement, and bank loans," *The Journal of Finance*, 64(2), 823–860.

Banerjee, A., E. Duflo, R. Glennerster, and C. Kinnan (2015): "The miracle of microfinance? Evidence from a randomized evaluation," *American Economic Journal: Applied Economics*, 7(1), 22–53.

Beaman, L., D. Karlan, B. Thuysbaert, and C. Udry (2014): "Self-selection into credit markets: Evidence from agriculture in mali," Discussion paper, National Bureau of Economic Research.

Berger, A. N., M. A. Espinosa-Vega, W. S. Frame, and N. H. Miller (2011): "Why do borrowers pledge collateral? New empirical evidence on the role of asymmetric information," *Journal of Financial Intermediation*, 20(1), 55–70.

Berger, A. N., W. S. Frame, and V. Ioannidou (2011): "Tests of ex ante versus ex post theories of collateral using private and public information," *Journal of Financial Economics*, 100(1), 85–97.

Berger, A. N., and G. F. Udell (1990): "Collateral, loan quality and bank risk," *Journal of Monetary Economics*, 25(1), 21–42.

Besley, T., and S. Coate (1995): "Group lending, repayment incentives and social collateral," *Journal of development economics*, 46(1), 1–18.

Bharath, S. T., S. Dahiya, A. Saunders, and A. Srinivasan (2009): "Lending relationships and loan contract terms," *The Review of Financial Studies*, 24(4), 1141–1203.

Bhue, G. S., N. Prabhala, and P. Tantri (2015): "Creditor rights and relationship banking: Evidence from a policy experiment," .

Billett, M. T., R. Elkamhi, L. Popov, and R. S. Pungaliya (2016): "Bank skin in the game and loan contract design: evidence from covenant-lite loans," *Journal of Financial and Quantitative Analysis*, 51(3), 839–873.

Breza, E. (2012): "Peer effects and loan repayment: Evidence from the krishna default crisis," *Job Market Paper MIT*, 1.

Caballero, R. J., and A. Krishnamurthy (2001): "International and domestic collateral constraints in a model of emerging market crises," *Journal of monetary Economics*, 48(3), 513–548.

Calomiris, C. W., M. Larrain, J. Liberti, and J. Sturgess (2017): "How collateral laws shape lending and sectoral activity," *Journal of Financial Economics*, 123(1), 163–188.

Carpena, F., S. Cole, J. Shapiro, and B. Zia (2012): "Liability structure in small-scale finance: evidence from a natural experiment," *The World Bank Economic Review*, 27(3), 437–469.

Catherine, S., T. Chaney, Z. Huang, D. Sraer, and D. Thesmar (2017): "Aggregate effects of collateral constraints," *NBER Working Paper*.

Cerqueiro, G., S. Ongena, and K. Roszbach (2016): "Collateralization, bank loan rates, and monitoring," *The Journal of Finance*, 71(3), 1295–1322.

Cherodian, R., and A. Thirlwall (2015): "Regional disparities in per capita income in India: convergence or divergence?," *Journal of Post Keynesian Economics*, 37(3), 384–407.

Cole, S. (2009): "Fixing market failures or fixing elections? Agricultural credit in India," *American Economic Journal: Applied Economics*, 1(1), 219–50.

De Quidt, J., T. Fetzer, and M. Ghatak (2016): "Group lending without joint liability," *Journal of Development Economics*, 121, 217–236.

Degryse, H., L. Lu, and S. Ongena (2016): "Informal or formal financing? Evidence on the co-funding of Chinese firms," *Journal of Financial Intermediation*, 27, 31–50.

Diamond, D. W. (1984): "Financial intermediation and delegated monitoring," *The review of economic studies*, 51(3), 393–414.

Donaldson, J. R., D. Gromb, and G. Piacentino (2019): "The paradox of pledgeability," *Journal of Financial Economics*.

Fischer, G. (2011): ""Contract structure, risk sharing and investment choice", Econ. Org," *Public Policy Discuss. Pap*, 23.

Fischer, G. (2013): "Contract Structure, Risk-Sharing, and Investment Choice," *Econometrica*, 81(3), 883–939.

Fisman, R., D. Paravisini, and V. Vig (2017): "Cultural proximity and loan outcomes," *American Economic Review*, 107(2), 457–92.

Ghatak, M. (1999): "Group lending, local information and peer selection," *Journal of development Economics*, 60(1), 27–50.

Ghatak, M., and T. W. Guinnane (1999): "The economics of lending with joint liability: theory and practice," *Journal of development economics*, 60(1), 195–228.

Giné, X., and M. Kanz (2017): "The economic effects of a borrower bailout: evidence from an emerging market," *The Review of Financial Studies*, 31(5), 1752–1783.

Giné, X., and D. S. Karlan (2014): "Group versus individual liability: Short and long term evidence from Philippine microcredit lending groups," *Journal of development Economics*, 107, 65–83.

Gopalan, R., X. Martin, and K. Srinivasan (2017): "Weak Creditor Rights and Insider Opportunism: Evidence from an Emerging Market," .

Gopalan, R., A. Mukherjee, and M. Singh (2016): "Do debt contract enforcement costs affect financing and asset structure?," *The Review of Financial Studies*, 29(10), 2774–2813.

Heckman, J. J. (1979): "Sample selection bias as a specification error," *Econometrica: Journal of the econometric society*, pp. 153–161.

Henderson, V., A. Storeygard, and D. N. Weil (2011): "A bright idea for measuring economic growth," *American Economic Review*, 101(3), 194–99.

Jimenez, G., V. Salas, and J. Saurina (2006): "Determinants of collateral," *Journal of financial economics*, 81(2), 255–281.

Johnson, S., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2000): "Tunneling," *American economic review*, 90(2), 22–27.

Karaivanov, A., and A. Kessler (2018): "(Dis) advantages of informal loans–Theory and evidence," *European Economic Review*, 102, 100–128.

Karlan, D. S. (2005): "Using experimental economics to measure social capital and predict financial decisions," *American Economic Review*, 95(5), 1688–1699.

——— (2007): "Social connections and group banking," *The Economic Journal*, 117(517), F52–F84.

Khwaja, A. I., and A. Mian (2005): "Do lenders favor politically connected firms? Rent provision in an emerging financial market," *The Quarterly Journal of Economics*, 120(4), 1371–1411.

Kleinberg, J., H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan (2017): "Human decisions and machine predictions," *The quarterly journal of economics*, 133(1), 237–293.

Lee, S., and P. Persson (2016): "Financing from family and friends," *The Review of Financial Studies*, 29(9), 2341–2386.

McKenzie, D., and C. Woodruff (2008): "Experimental evidence on returns to capital and access to finance in Mexico," *The World Bank Economic Review*, 22(3), 457–482.

Menkhoff, L., D. Neuberger, and C. Suwanaporn (2006): "Collateral-based lending in emerging markets: Evidence from Thailand," *Journal of Banking & Finance*, 30(1), 1–21.

Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado, and J. Dean (2013): "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, pp. 3111–3119.

Mukherjee, S., K. Subramanian, and P. Tantri (2018): "Borrowers Distress and Debt Relief: Evidence from a Natural Experiment," *The Journal of Law and Economics*, 61(4), 607–635.

Mullainathan, S., and J. Spiess (2017): "Machine learning: an applied econometric approach," *Journal of Economic Perspectives*, 31(2), 87–106.

Puri, M., J. Rocholl, and S. Steffen (2017): "What do a million observations have to say about loan defaults? Opening the black box of relationships," *Journal of Financial Intermediation*, 31, 1–15.

Simmons, P., N. Tantisantiwong, et al. (2018): "Evaluation of Individual and Group Lending under Asymmetric information," Discussion paper.

Spector, P. E., and M. T. Brannick (2011): "Methodological urban legends: The misuse of statistical control variables," *Organizational Research Methods*, 14(2), 287–305.

Stulz, R. M., and R. Williamson (2003): "Culture, openness, and finance," *Journal of financial Economics*, 70(3), 313–349.

Tantri, P. L. (2018): "Contagious Effects of a Political Intervention in Debt Contracts: Evidence Using Loan-Level Data," *The Review of Financial Studies*, 31(11), 4556–4592.

Tirole, J. (2010): *The theory of corporate finance*. Princeton University Press.

Vig, V. (2013): "Access to collateral and corporate debt structure: Evidence from a natural experiment," *Journal of Finance*, 68(3), 881–928.

Wydick, B. (1999): "Can social cohesion be harnessed to repair market failures? Evidence from group lending in Guatemala," *The Economic Journal*, 109(457), 463–475.

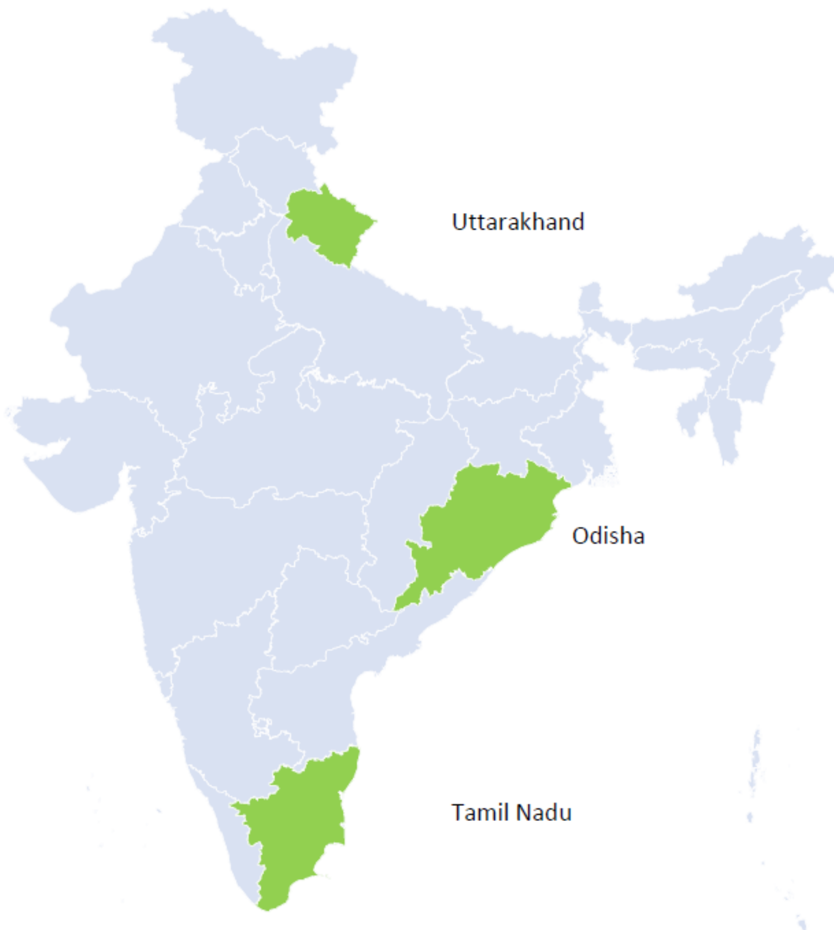


Figure 1: This figure depicts the location of the area of operation of the data provider within India.

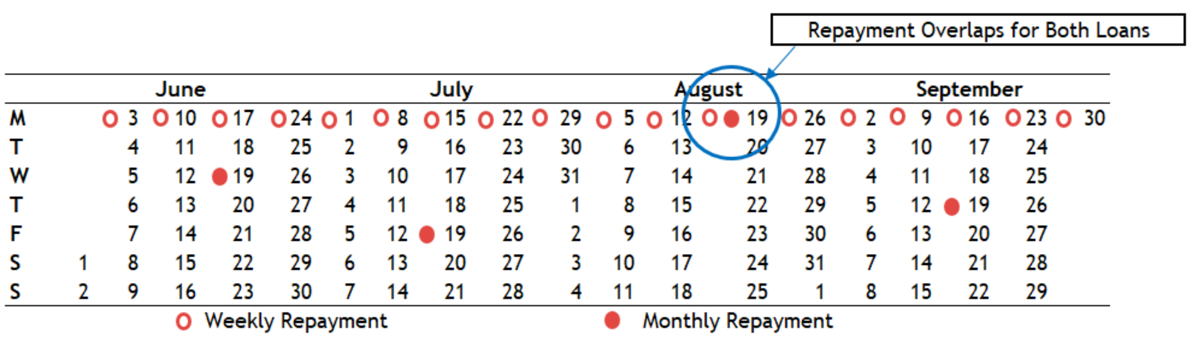


Figure 2: This image shows the repayment dates for a borrower who has two loans running simultaneously. Hollow circles represent weekly repayment dates (Monday of every week) and solid circles represents monthly repayment dates ($19^{th}$ of every month). In this example, in the month of August, monthly repayment due-date coincides with weekly repayment due-date (as $19^{th}$ is a Monday), hence the borrower has both the loans due on the same day. This does not happen in the other months shown in this example.

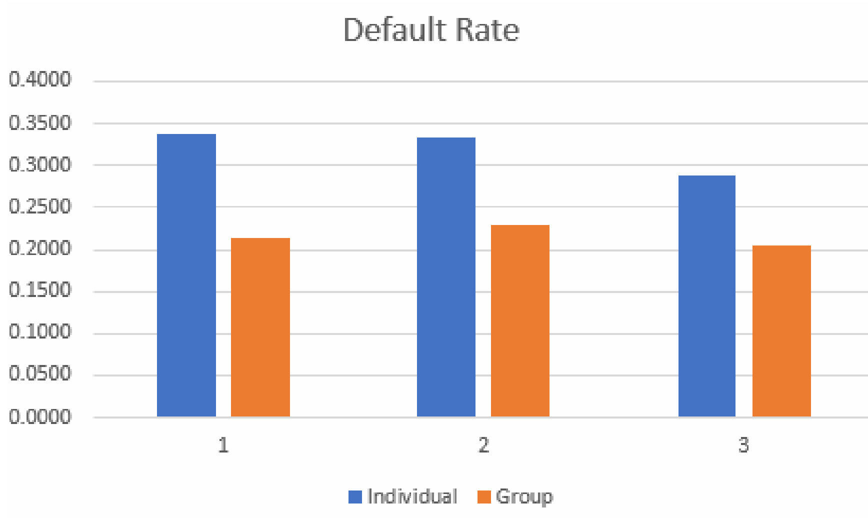


Figure 3: This graph shows the difference between default rates between Group and Individual Loans in the three samples. Sample 1 consists of all the repayments in which group and individual loans are running simultaneously. Sample 2 has the repayment instances where an individual is required to repay a group loan and an individual loan on the same day. Sample 3 has the repayment instances where a borrower is required to repay a group and an individual loan on the same day, but the two types of loans have different repayment frequencies.

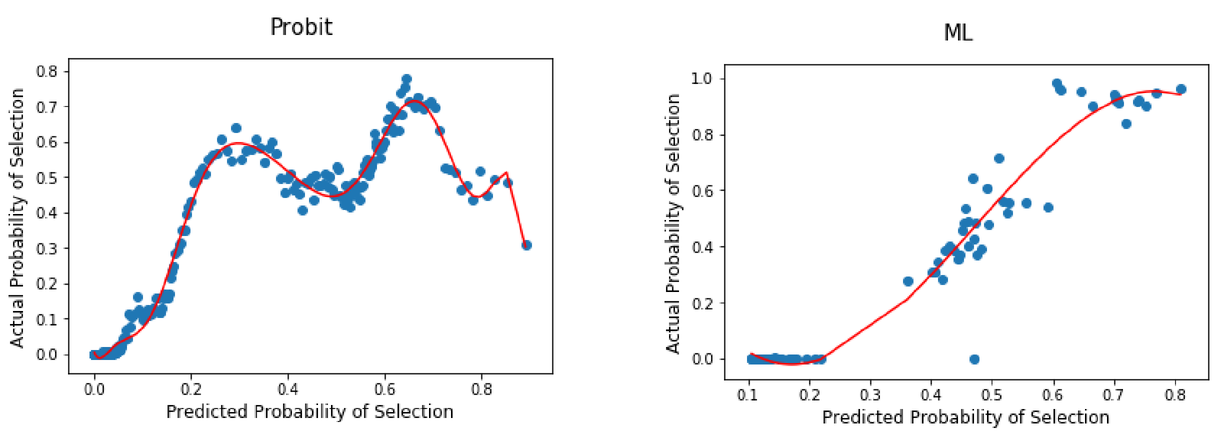


Figure 4: This graph shows actual frequency of selection corresponding to the probability calculated by the selection equation for being selected into the sample. The left shows the plot for the selection equation being estimated using a Probit Model, and the right panel shows the plot for the selection equation being estimated using a Machine Learning algorithm.

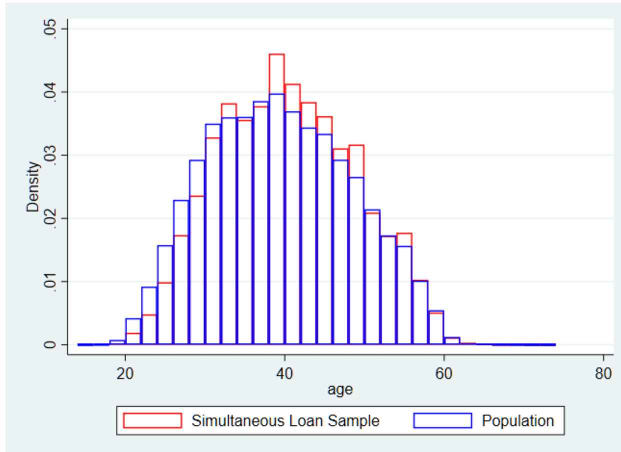


Figure 5: This graph plots the age distribution of the borrowers for our main sample of simultaneous loans (red color) and the entire loan portfolio of the lender (blue color).

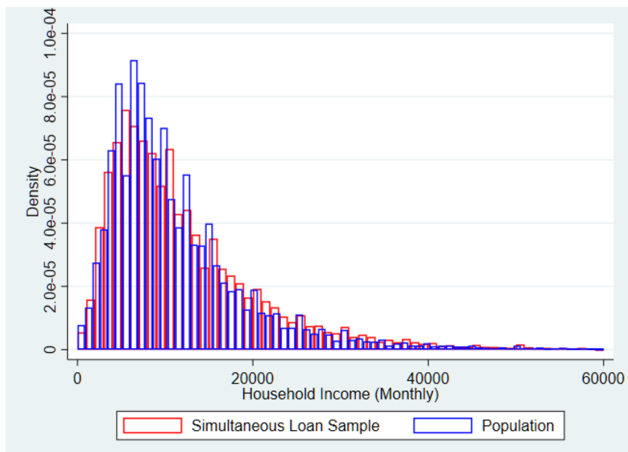


Figure 6: This graph plots the income distribution of the borrowers for our main sample of simultaneous loans (red color) and the entire loan portfolio of the lender (blue color).

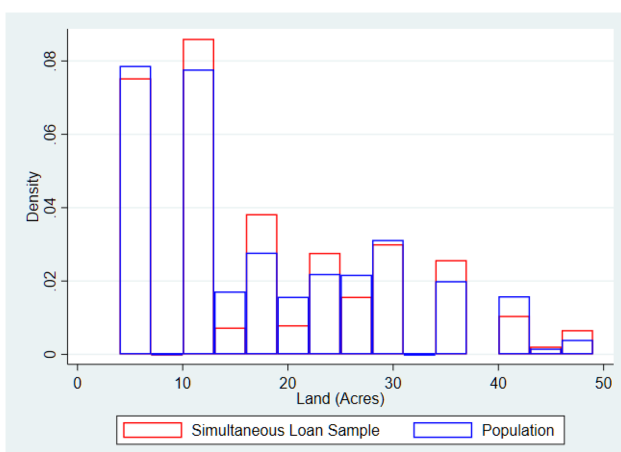


Figure 7: This graph plots the landholding distribution of the borrowers for our main sample of simultaneous loans (red color) and the entire loan portfolio of the lender (blue color).

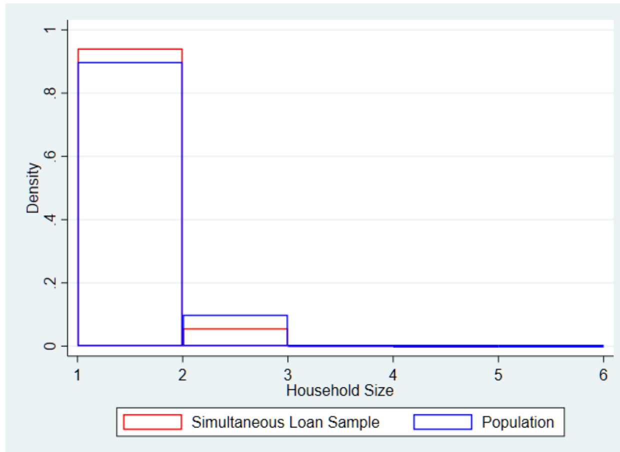


Figure 8: This graph plots the house size distribution of the borrowers for our main sample of simultaneous loans (red color) and the entire loan portfolio of the lender (blue color).

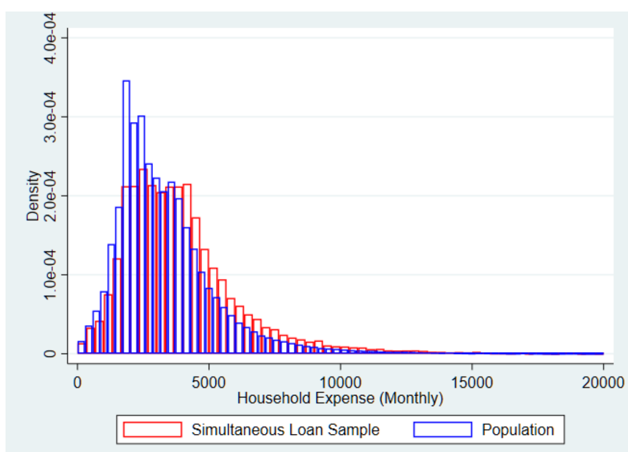


Figure 9: This graph plots the expenses distribution of the borrowers for our main sample of simultaneous loans (red color) and the entire loan portfolio of the lender (blue color).

**T**ABLE 1: Sample Construction: This table contains details regarding sample construction. The sample represents borrowers who have two loans running at the same time; a Group and an Individual Loan.

| Sample Period | Jun-09 | to | Jul-15 |
|---|---|---|---|
| | Group | Individual | Total |
| Number of Borrowers | 14151 | 14151 | 14151 |
| Number of Loans | 20397 | 16083 | 36480 |
| Number of Installment Repayments | 825193 | 235360 | 1060553 |
| Number of Simultaneous Installment Repayments | 796916 | 234033 | 1030949 |

**T**ABLE 2: Sample Construction: In this data-set, all borrowers have at least two simultaneously running loans. Each loan is repayable either at monthly frequency or at weekly frequency. This table shows the number of borrowers who have to repay both the loans at different frequencies in Panel A and those who have to repay them at same frequency in Panel B and Panel C. In this table, we only consider the repayments which have same due date for both the group and individual loans of a borrower.

| Panel A: Group and Individual Loans have different repayment frequencies; one weekly and the other monthly | | | |
|---|---|---|---|
| Loan Type | Number of Borrowers | Number of Loans | Number of Installment Repayments |
| Group | 8362 | 9984 | 13704 |
| Individual | 8362 | 8919 | 13596 |
| Total | 8362 | 18903 | 27300 |

| Panel B: Group and Individual Loans both have weekly repayment frequencies | | | |
|---|---|---|---|
| Group | 313 | 323 | 13104 |
| Individual | 313 | 316 | 13090 |
| Total | 313 | 639 | 26194 |

| Panel C: Group and Individual Loans both have monthly repayment frequencies | | | |
|---|---|---|---|
| Group | 1835 | 1988 | 23065 |
| Individual | 1835 | 1995 | 23121 |
| Total | 1835 | 3983 | 46186 |

There are sixty-one days when three simultaneously running loans have to be repaid on the same day by a borrower. Thus, adding total repayments of panels A, B and C gives 99680 instead of 99619.

**T**ABLE 3: Summary Statistics: This table shows the key summary statistics.

| | Mean | Std Dev | Percentiles 1% | 25% | 50% | 75% | 99% |
|---|---|---|---|---|---|---|---|
| Borrower Characteristics | | | | | | | |
| Age (Years) | 39.96 | 8.55 | 23 | 33 | 40 | 46 | 58 |
| Monthly Household Income (INR) | 13516.53 | 26807.27 | 1458 | 6000 | 10000 | 16145 | 62500 |
| Monthly Household Expense (INR) | 4025.78 | 2621.56 | 524 | 2391 | 3549 | 4866 | 8506 |
| Land Area (sq m) | 4.00 | 55.05 | 0 | 0 | 0 | 0 | 122.08 |
| Loan Statistics | | | | | | | |
| Default | 0.2355 | 0.4243 | 0 | 0 | 0 | 1 | 1 |
| Loan Aount (INR) | 18064.61 | 7386.40 | 1000 | 15000 | 20000 | 24000 | 35000 |
| Interest Rate | 0.2254 | 0.1160 | 0.0000 | 0.2194 | 0.2397 | 0.2519 | 0.3507 |
| Tenure (Years) | 1.02 | 0.20 | 0.87 | 0.96 | 0.96 | 1 | 2 |
| Group Loans | | | | | | | |
| Default | .2045 | .4033 | 0 | 0 | 0 | 0 | 1 |
| Loan Amount (INR) | 18663.17 | 5471.41 | 5000 | 15000 | 20000 | 20000 | 35000 |
| Interest Rate | .2170 | .0946 | 0 | .2194 | .2397 | .2397 | .2593 |
| Tenure (years) | .9867 | .0736 | .9615 | .9615 | .9615 | .9615 | 1.1923 |
| Individual Loans | | | | | | | |
| Default | .3442 | .4751 | 0 | 0 | 0 | 1 | 1 |
| Loan Amount (INR) | 15965.99 | 11628.73 | 1000 | 2000 | 20000 | 25000 | 45000 |
| Interest Rate | .2605 | .1752 | 0 | .2490 | .2605 | .2696 | 1.0359 |
| Tenure (years) | 1.14 | .38 | .5 | 1 | 1 | 1 | 2 |

**T**ABLE 4: Loan Repayment: Group and Individual Loans. The table shows results of comparison between loan repayment rates of group loans and individual loans lent to the same individual. The data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously. Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency. The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Group* is a dummy variable that takes the value of one if the loan under consideration is a Joint Liability Group loan and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | | | | Default | | |
| Group | -0.1257*** | -0.1024*** | -0.0832*** | -0.1432*** | -0.1139*** | -0.1097*** |
| | (-35.24) | (-13.39) | (-12.08) | (-31.16) | (-12.16) | (-9.38) |
| Household Size | | | | -0.0033 | -0.0493*** | -0.0028 |
| | | | | (-0.51) | (-3.42) | (-0.11) |
| log (Land Area) | | | | 0.0001 | -0.0002 | 0.0002 |
| | | | | (0.08) | (-0.08) | (0.05) |
| log (Household Income) | | | | 0.0372*** | 0.0440*** | 0.0392*** |
| | | | | (9.85) | (5.84) | (3.81) |
| log (Household Expense) | | | | -0.0049 | -0.0426*** | -0.0167 |
| | | | | (-0.85) | (-3.30) | (-1.09) |
| Age | | | | 0.0048*** | 0.0062** | 0.0067** |
| | | | | (3.30) | (2.55) | (2.48) |
| Constant | 0.5626*** | -0.0247 | 0.0716 | 0.1339 | -0.1908 | -0.3323 |
| | (4.38) | (-0.16) | (0.37) | (0.89) | (-0.95) | (-1.35) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1030949 | 99619 | 27300 | 1012199 | 94984 | 27277 |
| $R^2$ | 0.394 | 0.540 | 0.442 | 0.403 | 0.554 | 0.444 |

**T**ABLE 5: Loan Repayment: Comparison Between Group and Individual Loans Based On Loan Purpose. The table shows results of comparison between loan repayment rates of group loans and individual loans lent to an individual based on the purpose of the loan. There are 17 different purposes for which loans are borrowed like agriculture, repayment of other debt, buying jewelery, business capital, etc. The data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Group* is a dummy variable that takes the value of one if the loan under consideration is a Joint Liability Group loan and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| Panel A: Comparison Between Group And Individual Loans Lent For Same Purpose | | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dependent Variable | Default | | | | | |
| Group | -0.1271*** | -0.0815*** | -0.0904*** | -0.1494*** | -0.0866*** | -0.1437*** |
| | (-18.26) | (-6.78) | (-6.04) | (-17.30) | (-6.46) | (-5.88) |
| Household Size | | | | -0.0052 | -0.0449* | 0.0436 |
| | | | | (-0.43) | (-1.91) | (0.88) |
| log (Land Area) | | | | -0.0048 | -0.0042 | -0.0042 |
| | | | | (-1.45) | (-0.70) | (-0.60) |
| log (Household Income) | | | | 0.0475*** | 0.0399*** | 0.0696*** |
| | | | | (6.04) | (3.15) | (3.17) |
| log (Household Expense) | | | | -0.0061 | -0.0571*** | 0.0119 |
| | | | | (-0.51) | (-2.65) | (0.37) |
| Age | | | | 0.0031 | 0.0058 | 0.0055 |
| | | | | (1.13) | (1.16) | (0.93) |
| Constant | 0.3013*** | -0.0596 | 0.0471 | -0.1184 | -0.0840 | -0.7953* |
| | (2.87) | (-0.39) | (0.23) | (-0.58) | (-0.28) | (-1.95) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 244384 | 45364 | 6511 | 238659 | 43010 | 6504 |
| $R^2$ | 0.443 | 0.584 | 0.496 | 0.453 | 0.598 | 0.501 |

| Panel B: Comparison Between Group And Individual Loans Lent For Different Purposes | | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dependent Variable | | | Default | | | |
| Group | -0.1239*** | -0.1184*** | -0.0817*** | -0.1368*** | -0.1402*** | -0.0954*** |
| | (-31.52) | (-12.32) | (-10.63) | (-26.38) | (-10.89) | (-7.43) |
| Household Size | | | | 0.0013 | -0.0481*** | -0.0098 |
| | | | | (0.18) | (-2.63) | (-0.34) |
| log (Land Area) | | | | 0.0009 | -0.0007 | 0.0009 |
| | | | | (0.56) | (-0.22) | (0.27) |
| log (Household Income) | | | | 0.0300*** | 0.0571*** | 0.0244** |
| | | | | (7.18) | (5.75) | (2.20) |
| log (Household Expense) | | | | -0.0013 | -0.0388** | -0.0207 |
| | | | | (-0.19) | (-2.44) | (-1.20) |
| Age | | | | 0.0055*** | 0.0053* | 0.0067** |
| | | | | (3.53) | (1.96) | (2.27) |
| Constant | 0.7329*** | 0.2332 | 0.3330 | 0.2869** | -0.0521 | 0.0595 |
| | (6.12) | (0.61) | (0.69) | (1.97) | (-0.13) | (0.12) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 818111 | 58704 | 22853 | 804530 | 56164 | 22833 |
| $R^2$ | 0.392 | 0.517 | 0.452 | 0.399 | 0.529 | 0.453 |

**T**ABLE 6: Loan Repayment: Comparison Between Group and Individual Loans Based On Repayment Frequencies. The table shows results of comparison between loan repayment rates of group loans and individual loans lent to the same individual. The data is organized at a borrower-loan-repayment frequency level. The observations included in this table are the repayments for loans which are running simultaneously. In columns 1 and 4, the sample consists of all the repayments in which group and individual loans are to be repaid at different frequencies, one is to be repaid at monthly and the other at weekly frequency. Columns 2 and 5 have the cases where both the group loan and the individual loan borrowed by an individual have to be repaid at weekly frequency. Columns 3 and 6 have the cases where both the group loan and the individual loan borrowed by an individual have to be repaid at monthly frequency. The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Group* is a dummy variable that takes the value of one if the loan under consideration is a Joint Liability Group loan and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| Panel A: Comparison Between Group And Individual Loans For Simultaneously Running Loans | | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dependent Variable | | | Default | | | |
| Group | -0.1118*** | -0.2334*** | -0.0954*** | -0.1222*** | -0.2327*** | -0.0940*** |
| | (-27.08) | (-12.77) | (-23.50) | (-20.37) | (-12.71) | (-21.47) |
| Household Size | | | | 0.0052 | -0.0697*** | -0.0454*** |
| | | | | (0.73) | (-3.17) | (-2.92) |
| log (Land Area) | | | | 0.0012 | 0.0078 | -0.0006 |
| | | | | (0.85) | (0.72) | (-0.08) |
| log (Household Income) | | | | 0.0204*** | 0.0754*** | 0.0269*** |
| | | | | (4.64) | (3.46) | (3.52) |
| log (Household Expense) | | | | 0.0008 | -0.0863** | 0.0194** |
| | | | | (0.13) | (-2.31) | (2.02) |
| Age | | | | 0.0055*** | 0.0887*** | 0.2080*** |
| | | | | (3.67) | (3.49) | (28.35) |
| Constant | 0.7143*** | 0.3490*** | 0.3149 | 0.3546* | -2.7530** | -8.4150*** |
| | (3.65) | (2.64) | (0.64) | (1.80) | (-2.58) | (-28.42) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 858557 | 61400 | 133245 | 858340 | 61400 | 114712 |
| $R^2$ | 0.395 | 0.586 | 0.334 | 0.395 | 0.588 | 0.384 |

| Panel B: Comparison Between Group And Individual Loans When Repayment Dates Coincide | | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dependent Variable | Default | | | | | |
| Group | -0.0832*** | -0.2000*** | -0.0583*** | -0.1097*** | -0.2000*** | -0.0586*** |
| | (-12.08) | (-7.62) | (-20.79) | (-9.38) | (-7.62) | (-19.94) |
| Household Size | | | | -0.0028 | -0.0502** | -0.0693*** |
| | | | | (-0.11) | (-2.10) | (-3.21) |
| log (Land Area) | | | | 0.0002 | 0.0005 | 0.0121 |
| | | | | (0.05) | (0.03) | (0.94) |
| log (Household Income) | | | | 0.0392*** | 0.0219 | 0.0522*** |
| | | | | (3.81) | (0.69) | (3.49) |
| log (Household Expense) | | | | -0.0167 | -0.1973*** | -0.0035 |
| | | | | (-1.09) | (-3.29) | (-0.18) |
| Age | | | | 0.0067** | 0.0850*** | 0.4431*** |
| | | | | (2.48) | (3.69) | (15.14) |
| Constant | 0.0716 | 0.3039** | 0.1904*** | -0.3323 | -1.3349* | -18.1168*** |
| | (0.37) | (2.52) | (8.10) | (-1.35) | (-1.76) | (-15.13) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 27300 | 26194 | 46186 | 27277 | 26194 | 41574 |
| $R^2$ | 0.442 | 0.605 | 0.524 | 0.444 | 0.607 | 0.560 |

**T**ABLE 7: Loan Repayment: Group Loans vs Individual Loans With Strong Collateral. The table shows results of comparison between loan repayment rates of group loans and individual loans lent to the same individual. Here the individual loans have a strong collateral which is easy to confiscate like land or gold (Panel A) only gold (Panel B). The data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Group* is a dummy variable that takes the value of one for Group loans and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| Panel A: Group Loans vs Individual Loans With Strong Collateral | | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dependent Variable | Default | | | | | |
| Group | -0.1921*** | -0.1265*** | -0.1893*** | -0.2015*** | -0.1235*** | -0.2457*** |
| | (-27.12) | (-11.63) | (-14.06) | (-22.76) | (-8.50) | (-10.69) |
| Household Size | | | | -0.0000 | 0.0072 | -0.0138 |
| | | | | (-0.00) | (0.31) | (-0.37) |
| log (Land Area) | | | | -0.0071*** | -0.0089*** | -0.0139*** |
| | | | | (-4.52) | (-3.08) | (-4.21) |
| log (Household Income) | | | | 0.0114* | -0.0155 | 0.0470*** |
| | | | | (1.90) | (-1.18) | (2.62) |
| log (Household Expense) | | | | 0.0073 | 0.0111 | 0.0149 |
| | | | | (0.84) | (0.63) | (0.59) |
| Age | | | | 0.0091*** | 0.0153*** | 0.0096** |
| | | | | (3.57) | (3.72) | (2.12) |
| Constant | 0.4175*** | 0.0667 | 0.2643 | -0.0303 | -0.4481 | -0.5207 |
| | (4.72) | (0.30) | (1.23) | (-0.20) | (-1.41) | (-1.47) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 309606 | 21789 | 8548 | 304484 | 20721 | 8542 |
| $R^2$ | 0.355 | 0.506 | 0.452 | 0.362 | 0.517 | 0.459 |

| Panel B: Group Loans vs Individual Loans With Jewelery As Collateral | | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dependent Variable | | | Default | | | |
| Group | -0.9293*** | -0.9353*** | -0.9361*** | -0.9300*** | -0.9583*** | -0.9612*** |
| | (-112.36) | (-63.63) | (-63.06) | (-81.75) | (-38.93) | (-38.40) |
| Household Size | | | | 0.0079 | 0.0054 | 0.0055 |
| | | | | (0.96) | (0.19) | (0.20) |
| log (Land Area) | | | | 0.0010 | -0.0008 | -0.0010 |
| | | | | (0.35) | (-0.17) | (-0.20) |
| log (Household Income) | | | | -0.0034 | 0.0191 | 0.0215 |
| | | | | (-0.46) | (0.82) | (0.92) |
| log (Household Expense) | | | | -0.0157 | -0.0377 | -0.0361 |
| | | | | (-1.07) | (-0.57) | (-0.54) |
| Age | | | | -0.0014 | -0.0031 | -0.0032 |
| | | | | (-0.38) | (-0.51) | (-0.53) |
| Constant | 0.9363*** | 0.9794*** | 0.9773*** | 1.1310*** | 1.2336** | 1.2026** |
| | (98.72) | (60.17) | (58.92) | (6.06) | (2.15) | (2.11) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 50258 | 1568 | 1549 | 50182 | 1565 | 1546 |
| $R^2$ | 0.842 | 0.928 | 0.929 | 0.841 | 0.928 | 0.930 |

**T**ABLE 8: Shock And Default: Group vs Individual Loans. This table shows the differential effect of an economic shock on repayment rates of group and individual loans taken by a borrower. We measure economic shock using decline in night lights data in a district in a given month compared to previous month. Night Lights data for India is available publicly only till the year 2013, and is organized at a district-month level. Rest of the data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Shock* is a dummy variable which takes a value of one when there is a decline of twenty percentage points in median value of night-lights in a district for a month compared to previous month's median.The other important explanatory variable, *Group* is a dummy variable that takes the value of one for Group loans and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | Default | | | | | |
| Group | -0.1932*** | -0.1089*** | -0.1434*** | -0.1870*** | -0.0872*** | -0.1650*** |
| | (-36.74) | (-8.65) | (-9.70) | (-29.52) | (-5.15) | (-8.80) |
| Shock | 0.0792*** | 0.0720*** | 0.1109*** | 0.0796*** | 0.0717*** | 0.1125*** |
| | (22.19) | (4.11) | (5.11) | (22.34) | (4.05) | (5.19) |
| Shock x Group | -0.0065** | -0.0292** | -0.0322** | -0.0074** | -0.0276** | -0.0339** |
| | (-2.01) | (-2.52) | (-2.02) | (-2.29) | (-2.39) | (-2.12) |
| Household Size | | | | 0.0038 | 0.0070 | 0.0397 |
| | | | | (0.63) | (0.41) | (1.40) |
| log (Land Area) | | | | 0.0039** | 0.0038 | 0.0011 |
| | | | | (2.24) | (1.12) | (0.29) |
| log (Household Income) | | | | 0.0072* | -0.0225 | 0.0338*** |
| | | | | (1.90) | (-1.61) | (2.70) |
| log (Household Expense) | | | | 0.0327*** | 0.0387* | 0.0398** |
| | | | | (6.37) | (1.86) | (2.33) |
| Age | | | | 0.0075*** | 0.0077** | 0.0041 |
| | | | | (3.92) | (2.57) | (1.29) |
| Constant | 0.2856*** | 0.1129 | 0.0930 | -0.3218*** | -0.3059 | -0.6758** |
| | (7.56) | (0.65) | (0.46) | (-3.42) | (-1.20) | (-2.57) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 548015 | 20531 | 14773 | 547780 | 20504 | 14758 |
| $R^2$ | 0.377 | 0.509 | 0.524 | 0.378 | 0.512 | 0.526 |

**T**ABLE 9: The Impact of Monitoring: High Information is a dummy variable which equals one when the household's annual income is above 500000. Columns 1 and 4 have all the observations for group and individual loan repayments which were due on the same day, columns 2 and 5 have observations which are a subset of columns 1 and 2 where the annual household income is in between 250000 and 750000. Columns 3 and 6 are also a subset of columns 1 and 2 with income range on 400000 to 600000.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | | | Default | | | |
| Group | -0.1038*** | -0.2314*** | -0.2984** | -0.1155*** | -0.2360*** | -0.3193** |
| | (-13.36) | (-4.73) | (-2.27) | (-12.16) | (-4.71) | (-2.22) |
| High Information | -0.0291 | -0.0451 | -0.0325 | -0.0966** | -0.1025 | -0.2793 |
| | (-0.76) | (-0.35) | (-0.23) | (-2.32) | (-0.75) | (-1.51) |
| Group X High Information | 0.0705*** | 0.1876*** | 0.2373* | 0.0836*** | 0.2020*** | 0.2629* |
| | (4.05) | (3.44) | (1.88) | (4.49) | (3.74) | (1.89) |
| Household Size | | | | -0.0488*** | -0.1152** | 0.2563 |
| | | | | (-3.40) | (-2.10) | (1.51) |
| log (Land Area) | | | | -0.0001 | -0.0061 | 0.0000 |
| | | | | (-0.05) | (-0.30) | (.) |
| log (Household Income) | | | | 0.0461*** | 0.0123 | 1.2323 |
| | | | | (5.86) | (0.08) | (0.83) |
| log (Household Expense) | | | | -0.0437*** | -0.2423*** | -0.2551 |
| | | | | (-3.38) | (-3.58) | (-1.16) |
| Age | | | | 0.0063*** | -0.0180 | 0.0120 |
| | | | | (2.63) | (-0.84) | (0.30) |
| Constant | 0.0041 | 0.3381** | 1.3486*** | -0.1773 | 3.0278 | -10.5255 |
| | (0.02) | (2.57) | (13.41) | (-0.83) | (1.51) | (-0.61) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 99729 | 10767 | 1645 | 95094 | 10590 | 1575 |
| $R^2$ | 0.540 | 0.659 | 0.731 | 0.554 | 0.663 | 0.742 |

**T**ABLE 10: Interest Rate: Group Loans vs Individual Loans. The table shows results of comparison between the interest rate charged to a borrower for a group loan vs an individual loan. Since, the interest rates were not available to us in the data-set, we had to impute it using annuity formula. The calculations were not possible for all the loans as we assume the first repayment as the cash flow per period. If a borrower underpays or overpays on first installment, the imputed interest rate is unreliable. Hence, the number of observations are fewer compared to other tables. The data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Group* is a dummy variable that takes the value of one for Group loans and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | | | Interest Rate | | | |
| Group | -0.0426*** | -0.0590*** | -0.0476*** | -0.0326*** | -0.0603*** | -0.0339*** |
| | (-20.72) | (-10.25) | (-14.59) | (-12.94) | (-8.68) | (-7.29) |
| Household Size | | | | 0.0011 | 0.0020 | 0.0073** |
| | | | | (0.50) | (1.40) | (2.05) |
| log (Land Area) | | | | 0.0025*** | 0.0027** | 0.0036** |
| | | | | (3.58) | (2.00) | (2.25) |
| log (Household Income) | | | | -0.0134*** | 0.0034 | -0.0124*** |
| | | | | (-8.30) | (0.83) | (-2.90) |
| log (Household Expense) | | | | 0.0044** | -0.0036 | -0.0004 |
| | | | | (1.96) | (-1.45) | (-0.07) |
| Age | | | | 0.0018** | -0.0005 | 0.0005 |
| | | | | (2.43) | (-0.47) | (0.44) |
| Constant | 0.1247*** | 0.2005*** | 0.1668*** | 0.1279*** | 0.2166*** | 0.2289*** |
| | (7.51) | (19.74) | (15.06) | (3.47) | (4.87) | (3.58) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 979791 | 90000 | 24297 | 964557 | 86032 | 24278 |
| $R^2$ | 0.494 | 0.548 | 0.499 | 0.502 | 0.560 | 0.501 |

**T**ABLE 11: Loan Amount: Group Loans vs Individual Loans. The table shows results of comparison between the loan amount lent to a borrower for a group loan vs an individual loan. The data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Group* is a dummy variable that takes the value of one for Group loans and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | Loan Amount | | | | | |
| Group | 3479.92*** | 13436.58*** | 1409.56*** | 4654.48*** | 15379.08*** | 2439.51*** |
| | (31.31) | (70.73) | (8.24) | (33.59) | (72.53) | (9.10) |
| Household Size | | | | 130.26 | 627.03** | -266.80 |
| | | | | (1.12) | (2.34) | (-0.56) |
| log (Land Area) | | | | 15.15 | 369.24*** | -148.01** |
| | | | | (0.45) | (6.08) | (-2.31) |
| log (Household Income) | | | | -2049.82*** | -8233.02*** | -1783.64*** |
| | | | | (-21.35) | (-30.38) | (-7.97) |
| log (Household Expense) | | | | 589.16*** | 2663.46*** | 761.73** |
| | | | | (4.92) | (7.05) | (2.41) |
| Age | | | | -173.47*** | 293.84*** | -388.30*** |
| | | | | (-4.59) | (4.52) | (-6.13) |
| Constant | 8307.84*** | 657.22 | 10636.77*** | 25806.90*** | 33545.10*** | 32679.21*** |
| | (6.80) | (0.80) | (7.83) | (11.89) | (7.03) | (8.34) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1030949 | 99619 | 27300 | 1012199 | 94984 | 27277 |
| $R^2$ | 0.426 | 0.635 | 0.447 | 0.450 | 0.699 | 0.455 |

**T**ABLE 12: Loan Repayment: Simultaneous Individual Loans. The table shows results of comparison between loan repayment rates of two individual loans lent to the same individual. The two loans are running together for the borrower, at least for one repayment. The data is organized at a borrower-loan-repayment frequency level. In column 1, the sample consists of all the repayments that a borrower has to make on individual loans which are running simultaneously. Column 2 has the cases where both the loans borrowed by an individual have to be repaid on the same day. This sample does not segregate observations based on repayment frequency. Column 3 represents loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency. This sample consists of only those repayments which are repayable on the same day. The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variables are loan terms, like loan amount (in Indian Rupees) borrowed, interest rate charged and tenure (in years) of the loan. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| Dependent Variable | | Default | |
|---|---|---|---|
| Interest Rate | 0.031580 | 0.106594 | 0.156632 |
| | (0.82) | (0.63) | (0.80) |
| Loan Amount | -0.000001* | -0.000004* | -0.000006 |
| | (-1.78) | (-1.75) | (-1.45) |
| Tenure (years) | -0.050401* | -0.076109 | -0.349148*** |
| | (-1.85) | (-0.81) | (-3.51) |
| Constant | 0.530759*** | -0.053412 | -0.100196 |
| | (5.78) | (-0.23) | (-0.51) |
| Borrower Fixed Effect | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes |
| Observations | 53234 | 6365 | 1939 |
| $R^2$ | 0.364 | 0.364 | 0.349 |

**T**ABLE 13: Correcting for Selection Bias - Heckman Two-step Method. The column 1 represents selection into the sample of borrowers having a group and an individual loan at the same time from a larger population of borrowers. The population is comprised of borrowers who do not form the sample. Column one shows coefficients of a probit model used for selection. Columns 2 and 3 show the effect on the coefficient estimates of equation (1) on incorporating Inverse Mills Ration calculated using the probit model. Columns 4 and 5 show the same results using Inverse Mills Ratio calculated using Machine Learning Algorithm. The number of observations here are slightly lower compared to column 1 in Table 4 because of unavailability of information on Profession and Location of some borrowers. The dependent variable is a dummy variable that takes the value of one if a loan under consideration is selected in sample (defaults) and zero otherwise in column 1 (column 2,3,4,5). The main explanatory variable, *Group* is a dummy variable that takes the value of one if the loan under consideration is a Joint Liability Group loan and zero otherwise. IMR is Inverse Mills Ratio calculated using the probit model and IMR_ML is Inverse Mills Ratio calculated using the machine learning algorithm. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) Selected | (2) Default | (3) Default | (4) Default | (5) Default |
|---|---|---|---|---|---|
| Household Size | -0.2117*** | | -0.0008 | | -0.0020 |
| | (-63.11) | | (-0.11) | | (-0.24) |
| log (Household Income) | -0.0263*** | | 0.0373*** | | 0.0373*** |
| | (-23.83) | | (9.69) | | (9.62) |
| log (Household Expense) | 0.2000*** | | -0.0056 | | -0.0067 |
| | (133.02) | | (-0.89) | | (-0.97) |
| Land (sq km) | 0.0015*** | | -0.0000 | | -0.0000* |
| | (41.01) | | (-1.63) | | (-1.87) |
| Age | 0.0069*** | | 0.0050*** | | 0.0051*** |
| | (74.23) | | (3.36) | | (3.34) |
| Profession1 | -0.0042*** | | | | |
| | (-1351.16) | | | | |
| Profession2 | -0.0021*** | | | | |
| | (-626.25) | | | | |
| District | -0.1406*** | | | | |
| | (-721.27) | | | | |
| Group | | -0.1236*** | -0.1429*** | -0.1233*** | -0.1422*** |
| | | (-33.04) | (-30.64) | (-32.73) | (-30.24) |
| IMR | | -0.0134 | -0.0170 | | |
| | | (-0.98) | (-1.15) | | |
| IMR_ML | | | | -0.0229 | -0.0277 |
| | | | | (-0.45) | (-0.42) |
| Constant | 0.2157*** | 0.2291*** | -0.2190*** | 0.2302*** | -0.2116* |
| | (14.83) | (12.33) | (-2.64) | (5.28) | (-1.96) |
| Borrower Fixed Effect | No | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | No | Yes | Yes | Yes | Yes |
| Observations | 11520510 | 981584 | 981584 | 955368 | 955368 |
| $R^2$ | | 0.403 | 0.404 | 0.405 | 0.406 |
| Pseudo $R^2$ | 0.551 | | | | |

# 6 Appendix

**T**ABLE A1: Loan Repayment: Group and Individual Loans. Default = repayment is less than of 100% of first repayment. The table shows results of comparison between loan repayment rates of group loans and individual loans lent to the same individual. The data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency. The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Group* is a dummy variable that takes the value of one if the loan under consideration is a Joint Liability Group loan and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | | | Default | | | |
| Group | -0.1571*** | -0.0916*** | -0.1275*** | -0.1666*** | -0.0960*** | -0.1526*** |
| | (-44.38) | (-11.78) | (-18.13) | (-36.93) | (-10.10) | (-13.28) |
| Household Size | | | | -0.0064 | -0.0556*** | -0.0299 |
| | | | | (-0.99) | (-3.45) | (-1.18) |
| log (Land Area) | | | | 0.0013 | 0.0029 | 0.0002 |
| | | | | (0.92) | (1.07) | (0.07) |
| log (Household Income) | | | | 0.0253*** | 0.0157** | 0.0361*** |
| | | | | (6.95) | (2.00) | (3.59) |
| log (Household Expense) | | | | 0.0024 | -0.0250* | -0.0252 |
| | | | | (0.41) | (-1.76) | (-1.60) |
| Age | | | | 0.0056*** | 0.0089*** | 0.0056** |
| | | | | (3.87) | (3.55) | (1.99) |
| Constant | 0.5669*** | -0.1205 | 0.0774 | 0.1442 | -0.2927* | -0.1633 |
| | (4.39) | (-1.43) | (0.67) | (0.96) | (-1.70) | (-0.81) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1030948 | 99619 | 27300 | 1012198 | 94984 | 27277 |
| $R^2$ | 0.395 | 0.534 | 0.460 | 0.403 | 0.547 | 0.462 |

TABLE A2: List of Purposes For Which A Loan Is Borrowed In The Sample

| ID | Purpose |
|---|---|
| 0 | miscellaneous |
| 1 | animal |
| 2 | business |
| 3 | repayment |
| 4 | vehicle |
| 5 | agri |
| 6 | house |
| 7 | insurance |
| 8 | social |
| 9 | household |
| 10 | jewel |
| 11 | education |
| 12 | travel |
| 13 | liquidity |
| 14 | medical |
| 15 | Fishery |
| 16 | land |

**T**ABLE A3: Loan Repayment: Group Loans vs Individual Loans With Weak Collateral. The table shows results of comparison between loan repayment rates of group loans and individual loans lent to the same individual. Here the individual loans have a weak collateral which is not so easy to confiscate like anything other than land or gold (Panel A), or anything other than gold (Panel B). The data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency. The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Group* is a dummy variable that takes the value of one for Group loans and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| Panel A: Comparison Between Group And Individual Loans When Individual Loan Has Weak Collateral | | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dependent Variable | Default | | | | | |
| Group | -0.0989*** | -0.0959*** | -0.0366*** | -0.1259*** | -0.1120*** | -0.0644*** |
| | (-24.30) | (-10.38) | (-4.63) | (-23.54) | (-10.15) | (-4.66) |
| Household Size | | | | -0.0044 | -0.0711*** | -0.0004 |
| | | | | (-0.52) | (-4.26) | (-0.01) |
| log (Land Area) | | | | 0.0053 | 0.0056 | 0.0046 |
| | | | | (1.06) | (0.34) | (0.26) |
| log (Household Income) | | | | 0.0490*** | 0.0622*** | 0.0328*** |
| | | | | (10.59) | (6.97) | (2.59) |
| log (Household Expense) | | | | -0.0141* | -0.0721*** | -0.0295 |
| | | | | (-1.85) | (-4.24) | (-1.49) |
| Age | | | | 0.0006 | -0.0010 | 0.0012 |
| | | | | (0.38) | (-0.33) | (0.37) |
| Constant | 0.5905*** | -0.1750*** | -0.0935** | 0.2741 | 0.0409 | -0.1372 |
| | (3.56) | (-3.48) | (-2.51) | (1.44) | (0.21) | (-0.63) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 736953 | 78736 | 19080 | 723260 | 75166 | 19063 |
| $R^2$ | 0.413 | 0.554 | 0.456 | 0.422 | 0.569 | 0.457 |

| Panel B: Comparison Between Group And Individual Loans When Individual Loan Has Non-Gold Collateral | | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dependent Variable | Default | | | | | |
| Group | -0.0926*** | -0.0896*** | -0.0337*** | -0.1161*** | -0.1058*** | -0.0544*** |
| | (-26.95) | (-11.57) | (-5.13) | (-25.53) | (-11.23) | (-4.67) |
| Household Size | | | | -0.0063 | -0.0548*** | -0.0134 |
| | | | | (-0.92) | (-3.81) | (-0.54) |
| log (Land Area) | | | | -0.0008 | -0.0024 | -0.0005 |
| | | | | (-0.64) | (-0.94) | (-0.17) |
| log (Household Income) | | | | 0.0410*** | 0.0569*** | 0.0251** |
| | | | | (10.83) | (7.54) | (2.42) |
| log (Household Expense) | | | | -0.0086 | -0.0566*** | -0.0294* |
| | | | | (-1.45) | (-4.33) | (-1.93) |
| Age | | | | -0.0005 | -0.0008 | 0.0017 |
| | | | | (-0.42) | (-0.36) | (0.67) |
| Constant | 0.5311*** | -0.0526 | 0.0626 | 0.2961** | 0.0432 | 0.0602 |
| | (4.20) | (-0.30) | (0.25) | (2.05) | (0.20) | (0.21) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 983603 | 98161 | 25856 | 964929 | 93529 | 25836 |
| $R^2$ | 0.399 | 0.545 | 0.455 | 0.409 | 0.561 | 0.456 |

TABLE A4: Difference Based On The Type Of Collateral: The table shows results of comparison between loan repayment rates of group loans and individual loans lent to the same individual. The data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously. Columns 2 and 5 have the cases where both the group loan and the individual loan borrowed by an individual have to be repaid on the same day. This sample does not segregate observations based on repayment frequency. Columns 3 and 6 is a subset of their respective previous samples where group and individual loans are repayable at different frequencies, one at weekly and the other at monthly. In Panel A (Panel B), the dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The regression equation is changed from equation (1) to:

$Y_{itj} = \alpha + \nu_i + \gamma_j + \beta_1 StrongCollateral_{ij} + \beta_2 WeakCollateral_{ij} + \beta_3 BorrowerCharacteristics_{itj} + \varepsilon_{itj}$

The main explanatory variable, *Strong Collateral*(*Jewel*) is a dummy variable that takes the value of one for individual loans where the collateral used is either land or gold (gold) and zero otherwise. The other important explanatory variable, *Weak Collateral*(*Non-Jewel*) is a dummy variable that takes the value of one for individual loans where the collateral used is anything other than land or gold (anything other than gold) and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, and age. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

Panel A: Comparison Between Group And Individual Loans; Strong vs Weak Collateral

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | | | | Default | | |
| Strong Collateral | 0.2046*** | 0.1892*** | 0.3141*** | 0.4133*** | 0.6059*** | 0.5851*** |
| | (19.60) | (7.67) | (14.19) | (23.21) | (23.11) | (19.99) |
| Weak Collateral | 0.1018*** | 0.0990*** | 0.0376*** | 0.1044*** | 0.1074*** | 0.0401*** |
| | (25.13) | (10.71) | (4.78) | (21.17) | (10.03) | (3.57) |
| Household Size | | | | -0.0029 | -0.0541*** | -0.0033 |
| | | | | (-0.43) | (-3.62) | (-0.14) |
| log (Land Area) | | | | -0.0390*** | -0.0552*** | -0.0551*** |
| | | | | (-15.05) | (-14.95) | (-13.21) |
| log (Household Income) | | | | 0.0176*** | 0.0421*** | 0.0061 |
| | | | | (4.80) | (5.86) | (0.67) |
| log (Household Expense) | | | | 0.0011 | -0.0456*** | -0.0117 |
| | | | | (0.19) | (-3.52) | (-0.75) |
| Age | | | | 0.0058*** | 0.0024 | 0.0037 |
| | | | | (4.20) | (1.02) | (1.44) |
| Constant | 0.3489** | -0.1492 | -0.0588 | -0.0274 | -0.2011 | -0.1797 |
| | (2.42) | (-0.98) | (-0.31) | (-0.16) | (-0.97) | (-0.69) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1030949 | 99619 | 27300 | 1012199 | 94984 | 27277 |
| $R^2$ | 0.392 | 0.539 | 0.453 | 0.404 | 0.558 | 0.465 |

| Panel B: Comparison Between Group And Individual Loans; Collateral Being Jewelery vs Anything Else | | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dependent Variable | | | Default | | | |
| Jewel Loan | 0.9282*** | 0.9342*** | 0.9322*** | 0.9607*** | 0.9816*** | 0.9508*** |
| | (111.01) | (68.90) | (60.74) | (109.04) | (65.82) | (55.64) |
| Non-Jewel Loan | 0.0928*** | 0.0896*** | 0.0339*** | 0.1154*** | 0.1058*** | 0.0548*** |
| | (26.98) | (11.56) | (5.17) | (25.61) | (11.25) | (4.84) |
| Household Size | | | | -0.0056 | -0.0528*** | -0.0113 |
| | | | | (-0.88) | (-3.76) | (-0.51) |
| log (Land Area) | | | | -0.0007 | -0.0023 | -0.0005 |
| | | | | (-0.57) | (-0.96) | (-0.19) |
| log (Household Income) | | | | 0.0399*** | 0.0563*** | 0.0249** |
| | | | | (10.86) | (7.62) | (2.51) |
| log (Household Expense) | | | | -0.0090 | -0.0561*** | -0.0298** |
| | | | | (-1.57) | (-4.38) | (-2.03) |
| Age | | | | -0.0004 | -0.0010 | 0.0014 |
| | | | | (-0.38) | (-0.46) | (0.58) |
| Constant | 0.4379*** | -0.1316 | 0.0360 | 0.1891 | -0.0495 | 0.0267 |
| | (3.48) | (-0.73) | (0.15) | (1.33) | (-0.23) | (0.09) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1030949 | 99619 | 27300 | 1012199 | 94984 | 27277 |
| $R^2$ | 0.416 | 0.554 | 0.502 | 0.426 | 0.570 | 0.503 |

**T**ABLE A5: Shock And Default - First Stage (20%). This table shows the effect of an economic shock on repayment rates of the borrowers. We measure economic shock using decline in night lights data in a district in a given month compared to previous month. Night Lights data for India is available publicly only till the year 2013, and is organized at a district-month level. Rest of the data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Shock* is a dummy variable which takes a value of one when there is a decline of twenty percentage points in median value of night-lights in a district for a month compared to previous month's median. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | | | | Default | | |
| Shock | 0.0720*** | 0.0575*** | 0.0951*** | 0.0715*** | 0.0577*** | 0.0964*** |
| | (31.49) | (3.51) | (4.81) | (31.05) | (3.45) | (4.81) |
| Household Size | | | | 0.0162*** | 0.0111 | 0.0585** |
| | | | | (2.62) | (0.63) | (1.99) |
| log (Land Area) | | | | 0.0115*** | 0.0076** | 0.0075** |
| | | | | (6.40) | (2.21) | (2.02) |
| log (Household Income) | | | | -0.0812*** | -0.0971*** | -0.1044*** |
| | | | | (-21.47) | (-9.90) | (-11.95) |
| log (Household Expense) | | | | 0.0525*** | 0.0535** | 0.0720*** |
| | | | | (9.63) | (2.27) | (4.02) |
| Age | | | | 0.0209*** | 0.0127*** | 0.0121*** |
| | | | | (9.88) | (4.06) | (3.68) |
| Constant | 0.0603 | 0.0588 | 0.0218 | -0.4201*** | -0.0272 | -0.1963 |
| | (1.52) | (0.34) | (0.11) | (-3.97) | (-0.10) | (-0.65) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 548015 | 20531 | 14773 | 547780 | 20504 | 14758 |
| $R^2$ | 0.344 | 0.476 | 0.478 | 0.361 | 0.500 | 0.508 |

**T**ABLE A6: Shock And Default - First Stage (15%). This table shows the effect of an economic shock on repayment rates of the borrowers. We measure economic shock using decline in night lights data in a district in a given month compared to previous month. Night Lights data for India is available publicly only till the year 2013, and is organized at a district-month level. Rest of the data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Shock* is a dummy variable which takes a value of one when there is a decline of fifteen percentage points in median value of night-lights in a district for a month compared to previous month's median. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | | | | Default | | |
| Shock | 0.0632*** | 0.0287 | 0.0701*** | 0.0631*** | 0.0291 | 0.0720*** |
| | (28.77) | (1.56) | (2.98) | (28.63) | (1.55) | (3.03) |
| Household Size | | | | 0.0167*** | 0.0117 | 0.0594** |
| | | | | (2.70) | (0.66) | (2.03) |
| log (Land Area) | | | | 0.0114*** | 0.0076** | 0.0075** |
| | | | | (6.37) | (2.22) | (2.03) |
| log (Household Income) | | | | -0.0816*** | -0.0972*** | -0.1044*** |
| | | | | (-21.56) | (-9.91) | (-11.93) |
| log (Household Expense) | | | | 0.0521*** | 0.0529** | 0.0706*** |
| | | | | (9.57) | (2.25) | (3.95) |
| Age | | | | 0.0208*** | 0.0127*** | 0.0121*** |
| | | | | (9.85) | (4.06) | (3.68) |
| Constant | 0.0692* | 0.0881 | 0.0447 | -0.4030*** | 0.0065 | -0.1644 |
| | (1.75) | (0.49) | (0.21) | (-3.82) | (0.02) | (-0.53) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 548015 | 20531 | 14773 | 547780 | 20504 | 14758 |
| $R^2$ | 0.343 | 0.475 | 0.476 | 0.361 | 0.499 | 0.506 |

**T**ABLE A7: Shock And Default - First Stage (10%). This table shows the effect of an economic shock on repayment rates of the borrowers. We measure economic shock using decline in night lights data in a district in a given month compared to previous month. Night Lights data for India is available publicly only till the year 2013, and is organized at a district-month level. Rest of the data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Shock* is a dummy variable which takes a value of one when there is a decline of ten percentage points in median value of night-lights in a district for a month compared to previous month's median. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | | | | Default | | |
| Shock | 0.0635*** | 0.0289 | 0.0675*** | 0.0636*** | 0.0295 | 0.0692*** |
| | (31.80) | (1.63) | (2.93) | (31.81) | (1.63) | (2.97) |
| Household Size | | | | 0.0170*** | 0.0118 | 0.0597** |
| | | | | (2.76) | (0.67) | (2.04) |
| log (Land Area) | | | | 0.0114*** | 0.0076** | 0.0076** |
| | | | | (6.35) | (2.22) | (2.03) |
| log (Household Income) | | | | -0.0819*** | -0.0972*** | -0.1044*** |
| | | | | (-21.65) | (-9.92) | (-11.93) |
| log (Household Expense) | | | | 0.0517*** | 0.0529** | 0.0705*** |
| | | | | (9.51) | (2.25) | (3.95) |
| Age | | | | 0.0208*** | 0.0127*** | 0.0121*** |
| | | | | (9.82) | (4.06) | (3.68) |
| Constant | 0.0697* | 0.0882 | 0.0484 | -0.3946*** | 0.0063 | -0.1604 |
| | (1.76) | (0.49) | (0.23) | (-3.74) | (0.02) | (-0.52) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 548015 | 20531 | 14773 | 547780 | 20504 | 14758 |
| $R^2$ | 0.344 | 0.475 | 0.476 | 0.361 | 0.499 | 0.506 |

**T**ABLE A8: Shock And Default : Group vs Individual Loans (15%). This table shows the differential effect of an economic shock on repayment rates of group and individual loans taken by a borrower. We measure economic shock using decline in night lights data in a district in a given month compared to previous month. Night Lights data for India is available publicly only till the year 2013, and is organized at a district-month level. Rest of the data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Shock* is a dummy variable which takes a value of one when there is a decline of fifteen percentage points in median value of night-lights in a district for a month compared to previous month's median.The other important explanatory variable, *Group* is a dummy variable that takes the value of one for Group loans and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | | | Default | | | |
| Group | -0.1981*** | -0.1084*** | -0.1441*** | -0.1919*** | -0.0870*** | -0.1661*** |
| | (-39.15) | (-8.93) | (-10.04) | (-31.09) | (-5.29) | (-9.03) |
| Shock | 0.0645*** | 0.0441** | 0.0859*** | 0.0642*** | 0.0434** | 0.0870*** |
| | (18.67) | (2.29) | (3.43) | (18.59) | (2.23) | (3.47) |
| Shock x Group | 0.0004 | -0.0311*** | -0.0323** | 0.0001 | -0.0290*** | -0.0335** |
| | (0.13) | (-2.77) | (-2.05) | (0.03) | (-2.59) | (-2.13) |
| Household Size | | | | 0.0044 | 0.0076 | 0.0406 |
| | | | | (0.73) | (0.45) | (1.43) |
| log (Land Area) | | | | 0.0039** | 0.0038 | 0.0010 |
| | | | | (2.22) | (1.11) | (0.28) |
| log (Household Income) | | | | 0.0066* | -0.0226 | 0.0337*** |
| | | | | (1.73) | (-1.61) | (2.68) |
| log (Household Expense) | | | | 0.0322*** | 0.0383* | 0.0388** |
| | | | | (6.30) | (1.85) | (2.27) |
| Age | | | | 0.0075*** | 0.0077** | 0.0040 |
| | | | | (3.90) | (2.55) | (1.27) |
| Constant | 0.2984*** | 0.1419 | 0.1162 | -0.2988*** | -0.2715 | -0.6411** |
| | (7.91) | (0.79) | (0.55) | (-3.18) | (-1.05) | (-2.36) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 548015 | 20531 | 14773 | 547780 | 20504 | 14758 |
| $R^2$ | 0.376 | 0.509 | 0.523 | 0.377 | 0.511 | 0.525 |

**T**ABLE A9: Shock And Default : Group vs Individual Loans (10%). This table shows the differential effect of an economic shock on repayment rates of group and individual loans taken by a borrower. We measure economic shock using decline in night lights data in a district in a given month compared to previous month. Night Lights data for India is available publicly only till the year 2013, and is organized at a district-month level. Rest of the data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Shock* is a dummy variable which takes a value of one when there is a decline of ten percentage points in median value of night-lights in a district for a month compared to previous month's median.The other important explanatory variable, *Group* is a dummy variable that takes the value of one for Group loans and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | | | | Default | | |
| Group | -0.1953*** | -0.1051*** | -0.1423*** | -0.1894*** | -0.0847*** | -0.1644*** |
| | (-39.53) | (-9.02) | (-10.16) | (-31.20) | (-5.33) | (-9.06) |
| Shock | 0.0687*** | 0.0475** | 0.0851*** | 0.0678*** | 0.0465** | 0.0865*** |
| | (20.14) | (2.55) | (3.48) | (19.90) | (2.47) | (3.53) |
| Shock x Group | -0.0039 | -0.0373*** | -0.0360** | -0.0037 | -0.0345*** | -0.0375** |
| | (-1.23) | (-3.42) | (-2.33) | (-1.17) | (-3.15) | (-2.42) |
| Household Size | | | | 0.0047 | 0.0076 | 0.0407 |
| | | | | (0.79) | (0.44) | (1.43) |
| log (Land Area) | | | | 0.0038** | 0.0038 | 0.0011 |
| | | | | (2.19) | (1.13) | (0.30) |
| log (Household Income) | | | | 0.0063* | -0.0221 | 0.0339*** |
| | | | | (1.66) | (-1.58) | (2.69) |
| log (Household Expense) | | | | 0.0319*** | 0.0389* | 0.0394** |
| | | | | (6.24) | (1.89) | (2.31) |
| Age | | | | 0.0074*** | 0.0075** | 0.0039 |
| | | | | (3.85) | (2.50) | (1.23) |
| Constant | 0.2965*** | 0.1404 | 0.1191 | -0.2920*** | -0.2759 | -0.6401** |
| | (7.88) | (0.78) | (0.56) | (-3.11) | (-1.06) | (-2.36) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 548015 | 20531 | 14773 | 547780 | 20504 | 14758 |
| $R^2$ | 0.376 | 0.509 | 0.523 | 0.377 | 0.511 | 0.525 |

**T**ABLE A10: Loan Repayment: Group and Individual Loans With Equal Loan Amounts. The table shows results of comparison between loan repayment rates of group loans and individual loans lent to the same individual. Both the loans are of exactly the same amount (Indian Rupees). The data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Group* is a dummy variable that takes the value of one if the loan under consideration is a Joint Liability Group loan and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | | | | Default | | |
| Group | -0.1152*** | -0.0932*** | -0.0859*** | -0.1577*** | -0.1450*** | -0.1783*** |
| | (-10.65) | (-5.03) | (-4.21) | (-10.54) | (-4.80) | (-4.98) |
| Household Size | | | | -0.0199 | 0.0246 | 0.0531 |
| | | | | (-1.45) | (0.49) | (0.57) |
| log (Land Area) | | | | -0.0060* | -0.0005 | -0.0008 |
| | | | | (-1.76) | (-0.05) | (-0.08) |
| log (Household Income) | | | | 0.0574*** | 0.0855*** | 0.1051*** |
| | | | | (5.94) | (3.49) | (3.69) |
| log (Household Expense) | | | | -0.0037 | 0.0141 | -0.0075 |
| | | | | (-0.27) | (0.44) | (-0.22) |
| Age | | | | -0.0023 | 0.0041 | 0.0021 |
| | | | | (-1.06) | (0.55) | (0.26) |
| Constant | -0.2074 | -0.3654*** | -0.6088*** | -0.5655* | -1.3135*** | -1.4179*** |
| | (-0.86) | (-3.91) | (-4.20) | (-1.81) | (-2.97) | (-2.84) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 91883 | 3953 | 2864 | 90657 | 3921 | 2862 |
| $R^2$ | 0.419 | 0.550 | 0.493 | 0.427 | 0.559 | 0.504 |

**T**ABLE A11: Loan Repayment (measured as Non-performing Assets): Group Loans vs Individual Loans. The table shows results of comparison between loan repayment rates of group loans and individual loans lent to the same individual. The data is organized at a borrower-loan-repayment frequency level. In columns 1 and 4, the sample consists of all the repayments that a borrower has to make on loans which are running simultaneously.Sub-samples in rest of the columns have only those repayments where the group and individual loans borrowed by an individual have to be repaid on the same day. Further, columns 2 and 5 have sample which does not segregate observations based on repayment frequency. Columns 3 and 6 represent group and individual loan repayments of a borrower which are repayable at different frequencies, one at weekly and the other at monthly frequency The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Group* is a dummy variable that takes the value of one for Group loans and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dependent Variable | | | NPA | | | |
| Group | -0.0475*** | -0.0811*** | -0.0355*** | -0.0651*** | -0.0963*** | -0.0533*** |
| | (-24.04) | (-14.69) | (-8.77) | (-27.56) | (-14.32) | (-8.58) |
| Household Size | | | | -0.0080** | -0.0428*** | 0.0308* |
| | | | | (-2.29) | (-3.29) | (1.91) |
| log (Land Area) | | | | -0.0006 | -0.0015 | -0.0002 |
| | | | | (-0.80) | (-1.05) | (-0.13) |
| log (Household Income) | | | | 0.0321*** | 0.0540*** | 0.0252*** |
| | | | | (16.33) | (10.77) | (4.86) |
| log (Household Expense) | | | | 0.0024 | -0.0396*** | 0.0017 |
| | | | | (0.76) | (-3.86) | (0.21) |
| Age | | | | 0.0016* | 0.0015 | 0.0036* |
| | | | | (1.85) | (0.93) | (1.96) |
| Constant | 0.4434*** | -0.1023 | -0.0868 | 0.1209 | -0.2160* | -0.4558*** |
| | (3.38) | (-1.60) | (-1.00) | (0.86) | (-1.76) | (-3.17) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1030949 | 99619 | 27300 | 1012199 | 94984 | 27277 |
| $R^2$ | 0.162 | 0.316 | 0.364 | 0.167 | 0.325 | 0.366 |

**T**ABLE A12: Loan Repayment: Group and Individual Loans Lent On The Same Day. The table shows results of comparison between loan repayment rates of group loans and individual loans lent to the same individual on the same date. The data is organized at a borrower-loan-repayment frequency level. Repayment frequency is monthly in columns 1 and 2, and it is weekly in columns 3 and 4. In all columns the sample consists of all the repayments where both the group loan and the individual loan borrowed by an individual have to be repaid on the same day. Both individual and group loans were lent to the borrower on the same day. The dependent variable is a dummy variable that takes the value of one if a loan under consideration defaults and zero otherwise. The main explanatory variable, *Group* is a dummy variable that takes the value of one if the loan under consideration is a Joint Liability Group loan and zero otherwise. We control for time varying borrower characteristics like income, assets (land), expenses, age, and number of members in a household. We also control for borrower fixed effects and *month* x *year* (time) fixed effects. Errors are clustered at borrower level and robust t-statistics are reported in parenthesis. ***, ** and * represents significance at 1%, 5% and 10% respectively.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependant Variable | Default | | | |
| Group | -0.0536*** | -0.0542*** | -0.1003*** | -0.1003*** |
| | (-20.25) | (-19.56) | (-3.04) | (-3.04) |
| Household Size | | -0.0700*** | | -0.0705** |
| | | (-3.25) | | (-2.41) |
| log (Land Area) | | 0.0120 | | 0.0002 |
| | | (0.94) | | (0.01) |
| log (Household Income) | | 0.0490*** | | 0.0430 |
| | | (3.02) | | (0.39) |
| log (Household Expense) | | -0.0195 | | -0.2736*** |
| | | (-1.01) | | (-3.71) |
| Age | | 0.4433*** | | -0.0164 |
| | | (15.14) | | (-1.07) |
| Constant | 0.6468*** | -17.9468*** | 0.5092*** | 3.0407** |
| | (23.65) | (-15.00) | (25.50) | (2.49) |
| Borrower Fixed Effect | Yes | Yes | Yes | Yes |
| Month x Year Fixed Effect | Yes | Yes | Yes | Yes |
| Observations | 44670 | 40430 | 15944 | 15944 |
| $R^2$ | 0.532 | 0.567 | 0.581 | 0.583 |