Repay as you Earn

Shamika Ravi*
Department of Economics
New York University

March 15, 2005

Abstract

Households that face savings constraints prefer to 'tie' the repayment frequency of a loan to their household income frequency. If households are able to save and transaction cost is large, the optimal number of installments will tend to one per year. However with a low savings ability and a low transaction cost, repayment frequency will tend to income frequency. The study is based on the repayment behavior of 691 rural Indian households and 738 loans. The results uphold the hypothesis. We find that in general income frequency increases the repayment frequency by 32 percent. However, on controlling for savings, we find that this effect reduces to less than 3 percent. That is households which can save, the income frequency does not affect the repayment frequency per year. We also find that transaction costs reduce the repayment frequency per year.

*I am grateful to Jonathan Morduch for guidance and encouragement and GSAS, NYU for funding data collection.
All remaining errors are mine - shamika.ravi@nyu.edu

1 Introduction

In a typical framework, households take a loan, invest into production and then once sufficient revenues are earned, repay the loan with interest. All else the same, if households can save without difficulty, they should be able to follow any repayment frequency. Individuals calculate how much money to spend in each period. In a standard economic model, there is no room for immediate pressures. However, in reality, it is likely that the income gets diverted into miscellaneous consumption expenses. If households realize this, then it is possible that they tie their repayment schedules to their income schedule. That is, repay as soon as the money is earned.

Repayment frequency is an important component of any credit contract. There has been little work done to understand why there is huge variation in repayment frequency across different households. Policy work has focused on the role of subsidies in interest rate to increase outreach and allocation of credit in poor communities. Moneylenders and other informal lenders have an advantage over formal creditors in the form of a wide menu of repayment frequencies that they can offer to borrowers.

Analyzing specific features of microcredit programs, Morduch and Armendariz (2000) have looked at high frequency repayment requirement of these programs. They note that in programs surveyed by Craig Churchill (1999), lenders estimate repayment capacity of a borrower without taking into account expected revenues from the loan in question. Instead, lenders consider income flows provided by all members within the borrowing household. One likely reason being that lenders are aware of the 'decay' of cash in hand of borrowers, and force them to repay as soon as it is earned.

So, what is an optimal repayment frequency? Is this affected by a household's saving opportunities and cost of credit? Do households have preferences over repayment frequencies for a given loan? If so, what are the factors that determine this? These are the questions that we address in this paper. The outline of the study is as follows: section 2 lays out a simple model and section 3 describes the data; in section 4,we have the empirical specification and identification strategy; section 5 has the results and discussion. Section 6 has concluding remarks.

2 Model

We will use a very simple model to elucidate the problem. Consider a borrowing household with income y and consumption expenditure c.After purchasing necessities, the household is left with cash in hand, x,such that

$$y - c = x$$

This cash is received with frequency f and decays at a discount rate s. That is, if this cash in hand is not saved, it gets diverted into miscellaneous expenses with probability $(1 - s)$ every period the income is earned. We will assume that these expenses do not bring any utility to the household, but this can be relaxed and the argument would still hold.

The household borrows a loan l, at an interest rate r. The total amount to be repaid is then L, such that

$$l(1 + r) = L$$

The borrower has to make n installments of repayment, where $n = f/t$ and t is the frequency of installments. If the loan is a year in duration, and the borrower earns monthly income then installments can be either one time $(n = 1, f = 12, t = 12)$ or half yearly $(n = 2, f = 12, t = 6)$ or quarterly $(n = 4, f = 12, t = 3)$ or monthly $(n = 12, f = 12, t = 1)$.

Let the transaction cost associated with each installment be θ.This cost is borne by the borrower. Assuming linear preferences with respect to income and assuming that the loan size is no larger than the fraction of outside revenues that can be secured to repay the bank, the borrower will then choose the frequency of installment, t, to maximize the size of the loan. This is its expected total payment to the bank minus total transaction cost

$$\max_t \left\{ (1 + s + s^2 + ... + s^t)\frac{fx}{t} - \frac{f\theta}{t} \right\} \tag{1}$$

Now borrower has to decide what is the optimal repayment frequency? From the above optimization, we see that for a borrowing household that is able to save easily, that is s is close 1 and faces high transaction costs, that is, θ is large, the optimal installment frequency will tend towards the income frequency of the household and the optimal number of installment, n will tend to one. On the other hand, households that face savings constraints, that is,have low s and low transaction costs, θ,number of installments, n will tend to income frequency.

Proposition: *If households are able to save and s is close to 1 and θ is large, the optimal t will tend towards f. That is the optimal number of installments, n will tend to one. However with a low savings ability and a low transaction cost, n will tend to f.*

[Proof in appendix.]

3 Data

The data used for this study is from 738 current outstanding loans of 691 rural households from Uttar Pradesh and Kerala. Table1 provides the descriptive statistics. We have divided households into six categories depending on the frequency of income flow. In general, rural households have more than one source of income. Therefore there is more than one frequency with which income flows into these households. For assigning a particular income frequency to a household, we consider the source of largest fraction of the total annual income for each household.

Households that depend on daily earnings are the poorest in our sample. Cost of credit to these households is highest at 37 percent per annum. The loans are repaid at very high frequencies, the average being 8.4 installments per year. This category also has the lowest fraction of households that reported to have some form of savings. Households in this category reported a 'preferred' repayment frequency of 14.4 per year, which implies that on an average these households would prefer to repay their debt through 6 more installments per year. On the other extreme are 36 households that earn their income annually. These generally depend on one harvest, for example households in Uttar Pradesh that have one potato harvest. More than half the number of households within this category have reported to hold some form of savings or another. These households also have the lowest repayment frequency within a year, but would prefer to repay 0.5 fewer installments per year. Households that earn monthly income includes households that run small businesses, rentals and households that are employed in government services. Households that earn quarterly income include traders, mill owners with seasonal operations, rubber cultivators. Households with half yearly income are mostly cultivators with two harvests in a year. Then there are households that do not have a fixed frequency of income. These include households that depend on monetary transfers from members living outside the village.

The observed repayment frequency for each loan is an outcome. It is the result of interplay of repayment frequency offered by a lender and the repayment frequency 'preferred' by a borrower. We have information on the preferred repayment frequency but not on the offered. However, we can

assume that when the actual and preferred repayment frequencies are the same, then the offered frequency must be the preferred one. On the other hand when the actual frequency is different from the preferred frequency, then the offered must be the offered one.

Many microfinance programs target women because they are believed to have greater financial discipline and prove to be 'good' borrowers. In each category of households that we study, we consider percentage of female headed households. The highest percentage of female headed households are included in the 'Not Fixed' category. More than half of these households reported holding some form of savings and the frequency of loan repayment is not very high, with four installments a year.

Figure 1 provides some summary statistics for savings. The first pie chart provides a breakup of the forms of savings that households reported. Several households have savings in more than form. For each household that reports holding some form of savings, we consider the form in which the largest fraction of saving is held. From the breakup, we note that 43 percent of households that report savings, hold them in savings account in banks, cooperatives and post offices. One in every 5 households with savings, holds it in form of cash. One in 6 households reported holding savings in the form of jewelry. 8 percent hold grains as savings while 13 percent save in the form of insurance policies. The second pie chart provides information on the most important motives for saving that households reported. The two most important motives for saving are marriage and loan repayment. 23 percent households with savings reported that they were saving to repay outstanding debts. This is a significant proportion. The other saving motives reported are for medical emergencies, education of children and for investment purposes.

Figures 2 and 3 show the fitted values of income frequency and the observed repayment frequency from the data. Figure2 shows the correlation against saving. We note that households that have reported holding some form of savings have very low correlation between income and repayment frequency, this is almost close to zero. However, the correlation between income frequency and loan repayment frequency is strongly positive, for households that hold savings. More specifically, in Figure 3, we consider the correlation for a particular type of saving, that is holding a savings account, since it is the most common form of savings reported by the households. We find a very strong correlation between the two frequencies for households that do not hold a savings account. However the correlation is close to zero for households that hold a savings account.

One point to highlight is that 2 percent of households in our data have irregular income flow and therefore fall under the 'not fixed' category. The loan data reveals that 21 percent of all

loans have 'not fixed' repayment schedules. These are loans from friends and relatives. However it is interesting to note that when households report their preferred repayment schedules, only 1.2 percent of all households prefer a 'not fixed' schedule. This might imply that households internalize the savings constraint they face and prefer to commit to a particular repayment schedule for a loan.

4 Empirical Specification and Identification Strategy

We are interested in testing the following hypothesis: If households are able to save and transaction cost is large, the optimal repayment frequency will tend to one per year. Conversely, with a low savings ability and a low transaction cost, the optimal repayment frequency will tend to income frequency. In order to test this, the basic model that we estimate is of the following form

$$R_i = \alpha + \beta_1 f_i + \beta_2 S_i + \beta_3 (f_i * S_i) + \beta_4 C_i + \beta_5 L_i + \beta_6 H_i + \epsilon_i \tag{2}$$

where i is a loan. R_i is the repayment frequency for the given loan, f_i is the income frequency of the household with the loan i. S_i is a dummy which takes value 1 if the household with loan i holds some form of savings. We will separately consider different types of savings to see if the results depend on this. For example, we will consider savings account, jewelry, cash and LIC policy. C_i is transaction cost incurred in repayment. This will include distance to nearest bank, cooperative society H_i is a vector of household characteristics like income, asset, years of education and occupation. L_i are the loan characteristics of the loan i such as source/lender, interest rate, amount and purpose of loan.

There are some concerns that we should address before running the above regression. First of all since we observe only the actual repayment schedule and not the offered repayment schedule, it is difficult to identify the effect of savings and income frequency. As an identification strategy, therefore, we will use the 'preferred repayment' schedule for each household, instead of the actual repayment schedule for each loan, on the left hand side.

Secondly, we have loan specific information only for borrowing households and if we include loan specific information in the regression the non-borrowers drop out of out analysis. We will therefore run the above regression first without loan specific data and then including it. This will help us understand whether a household's preference of repayment schedule is affected by loan specific details such as usage of loan or whether it is secured against collateral.

Another concern in the above specification that might arise is the endogeneity of income frequency, f. One can argue that frequency of income is exogenously determined because it depends on the occupation structure within the household. Except for farmers who can choose which crops to cultivate which will determine the frequency of harvest and therefore income, there is no reason to believe that frequency of income is endogenously determined for other occupations like casual labor and regular secured employment. In order to avoid reverse causality arguments, we will run the above regression for farmers separately and test if the results vary significantly.

5 Results and Discussion

Before looking at the results, it is helpful to get some sense of the variables used in the regression. This is available in Table 2. Comparing the preferred and actual repayment frequencies we note that households in our data would prefer to repay more frequently than they actually do. Households would prefer to repay on an average two more installments in a year. Rural households earn income from many sources. Income frequency of a household is therefore calculated as the frequency of the largest proportion in total annual income. On an average, households earn income twenty times in a year. This average is high because nearly one third of the households depend on daily wages as a main source of income. On average a casual laborer works for 45 days in a year. Farmers, on average earn income twice a year with each harvest. Households with regular and secured employment earn monthly salaries. Self employed households either earn daily, monthly or seasonal/quarterly income. 13 percent of our data are female headed households. Looking at savings, we notice that 61 percent of households have reported holding some form of savings or another. But only 34 percent have a savings account in a formal institute like bank or cooperative society. In order to account for transaction costs, we control for distance to nearest metallic road, distance to bank and the presence of cooperative in the village.

First of all lets look at the determinants of savings. Table 3 presents the results of a Probit, where the dependent variable takes value one if the household reports any form of savings, and 0 otherwise. First of all wealthier households have a greater probability of holding savings, wealth measured in terms of income, assets, landholding or value of house. The presence of a cooperative society within the village increases the probability of holding assets. This reaffirms the view that institutions can help people commit to a certain behavior, in this case saving. Occupation of household head and frequency of income flow into the household do not affect the probability of saving. It is interesting

to note that while a household with more dependents has a greater probability of holding savings, households with more members is less likely to do so. Female headed households are more likely to save, though not significantly more.

Nuclear households have a significantly higher probability of holding savings. Nuclear households are not only smaller on average than stem or joint families but there is also one well defined household head who makes financial decisions. This result raises several questions about decision making process within different family structures. Could this result have been due to mis-reporting of savings by joint families? By separately looking at sub samples, Kerala and Uttar Pradesh, as well as subgroups of households by occupation, we find that this result holds across different categories. Nuclear families seem to have more financial discipline and have a greater probability of holding savings.

Moving on to the main results of this study, table 4 shows the results of the Maximum Likelihood Ordered Probit estimation. The dependent variable is preferred repayment frequency of each household, it takes values 1 if annual, 2 if half yearly, 4 if quarterly, 12 if monthly. The first and second columns include all households from data. First column excludes the savings dummy which takes value 1 if the household has reported some form of saving and 0 otherwise. The second includes the savings dummy and it's interaction with income frequency - this is the coefficient of interest to us. The third column includes loan specific information and therefore only households with current outstanding loans. There are 468 borrowing households. All regressions have a vector of household characteristics, variable of credit relations described in Table 2. Column (3) also includes loan specific details and lender dummies as described in Table 2.

Ordered Probit (1) shows that income frequency of a household is a positive and significant determinant of the repayment frequency of a loan. The coefficient is quite high at 0.323, implying that with a one fold increase in income frequency, households would prefer to repay 32 percent faster rate. However, when we include the savings dummy in the second column, we note that this coefficient reduces less than 10 times. Households with savings have significantly lower repayment frequencies. This is in tune with our hypothesis. The coefficient of interest is savings interacted with income frequency - here we note that income frequency does not determine repayment frequency for saving households. The coefficient is significant but close to zero. The third column also reveals a similar result for borrowing households. Income frequency does not affect repayment frequency for households with savings.

Transaction cost includes distance to nearest metallic road. As expected this negatively affects the repayment frequency. The magnitude is small in all the three regressions but significant.

Let us look at behavior of female headed households. Female headed households repay loans at a significantly higher frequency. In (1), we note that female households repay loans, 24 percent time more within a year. Even after controlling for their savings behavior, we note that the gender effect remains significant at 14 percent. Looking at the interaction term of female with income frequency, we find that income frequency remains a positive and significant determinant of loan repayment frequency for female headed households. This is not surprising as results in table 3 revealed that females are not likely to save more, therefore they have a greater reason to tie repayment frequency with frequency of their income flow.

Do wealthier households behave differently then poorer ones with respect to repayment frequency? The results show that while households with greater income, repay at significantly lower rates, households with greater assets repay more frequently. Controlling for savings, we note that richer households still repay less frequently. Also the income frequencies of these households do not significantly affect the repayment frequencies. This confirms claims made by Rutherford (2000) and Morduch and Armendariz (2000) which explain that in poorer households the opportunity cost of time is relatively low and where mechanisms to enforce financial discipline are relatively low and that these tendencies are reinforced by the fact that small scale businesses like petty trading tend to generate flow of revenue on a daily or frequent basis, making frequent collections by lenders especially desirable in the absence of satisfactory savings facilities. However, in the wealthier households the opportunity costs are likely to be higher and income flows less frequent, implying lower frequency of loan repayment, as the results show.

Loan specific variables also reveal some interesting insights. The results from (3) show that when loans are secured against collateral, households repay loans faster. The usage of the loan, however, doesn't affect the repayment behavior of borrowing households.

6 Robustness

We have information on saving motives of households. 60 households reported that the most important motive for their savings was to repay loans. We will exploit this information to test our hypothesis across different groups of households. The results are in Table5. The dependent variable

again is the preferred repayment frequency for each household. First column, (4) shows the results only for the 60 households that reported loan repayment as the most important saving motive. The second column (5) shows the results for all households that have savings. The third column, (6) shows results for non-savers and 60 households that save for loan repayment. All regressions have a vector of household characteristics, variable of credit relations described in Table 2.

The results for the 60 households that save for loan repayment shows that income frequency does not affect loan repayment frequency. The coefficient is significant but close to zero. This supports our hypothesis, when households are able to save their repayment frequency is not tied to the household income frequency. This result is specially true for households that save for loan repayment. When households are able to save to repay loans, they have no incentive to repay as they earn for the fear of cash decay.

Results for all savers in second column reveals that this relationship is stronger but still very low at 0.15 compared to 0.32 for all households from Table 4. There are many reasons why households save, as Figure 1 depicts. Marriage, education and medical reasons are strong incentives to save. So if we look at the pool of all savers, we note that households' repayment behavior is influenced by their income frequency.

The last column which shows the results for non-savers and the 60 households that save to repay loans. The third column does not include the savings dummy and its interaction terms while column four includes these. The results show that within this group income frequency very strongly affects the loan repayment frequency. Households within this group very strongly tie their repayment frequency with flow of income. However the last column reveals that when we control for savings, this effect falls to nearly one third the original magnitude. The coefficient on savings dummy shows that those that save within this group have very low repayment frequencies. The interaction term reveals that households that hold savings, the income frequency doesn't affect repayment. This coefficient is almost zero and significant. Transaction costs reduce the frequency of repayment. This result holds across all samples.

TO BE ADDED - Results for farmers - to test the endogeneity of income frequency.

7 An Alternate Hypothesis

Households do not save for several reasons. Firstly, poor households do not have sufficient resources to put aside as savings. Secondly, households that can save but do not have opportunities in the form

of access to savings institutions. Thirdly, there are households thave have access to savings facilities but lack the financial discipline required for saving. Of course the most important thing to note is whether people are sophisticated and recognize the immediate pressures and their subsequent behavior or they are naive. There is evidence in the literature to suggest that people choose commitment devices. The difficulty of sticking with a course of action in the presence of immediate pressures also has implications for how people save. In the standard economic model of savings, there is no room for immediate pressures. Behavioral economists have begun to understand the devices people may use to deal with such pressures. This approach of modelling behavior is called hyperbolic discounting.

Savings behavior of people in poor communities can perhaps be also understood using this perspective. Gugherty (2003) provides some evidence with respect to roscas. Roscas are immensely popular but they often pay no interest. There is also a high risk of default by early winners. Gugherty suggests that one reason for their popularity may be that they serve as a commitment device in several ways. Ashraf, Karlan and Yin (2004) give another illustration. They offer savers at a bank in the Philippines the opportunity to participate in 'SEED' accounts. These are like deposit accounts except that individuals cannot withdraw at will, and it does not pay any extra interest. They find that more than 30 percent offered this account, take them up. Six months later these households showed greater savings rate.

Do households pay higher interest rates in order to choose repayment frequencies that help them commit to savings. That is, are households paying to tie the repayment frequency to their income frequency? [TO BE TESTED]

8 Conclusion

In a typical framework, households take a loan, invest into production and then once sufficient revenues are earned, repay the loan with interest. All else the same, if households can save without difficulty, they should be able to follow any repayment frequency. Individuals calculate how much money to spend in each period. In a standard economic model, there is no room for immediate pressures. However, in reality, it is likely that the income gets diverted into miscellaneous consumption expenses. If households realize this, then it is possible that they tie their repayment schedules to their income schedule. That is, repay as soon as the money is earned.

This paper presents a very simple model the main hypothesis being that if households are able to save and transaction cost is large, the optimal number of installments will tend to one per year. However with a low savings ability and a low transaction cost, repayment frequency will tend to income frequency. The study is based on the repayment behavior of 691 rural Indian households and 738 loans.

The results uphold the hypothesis. We find that in general income frequency increases the repayment frequency by 32 percent. However, on controlling for savings, we find that this effect reduces to less than 3 percent. That is households which can save, the income frequency does not affect the repayment frequency per year. We also find that transaction costs reduce the repayment frequency per year. Nuclear families have great financial discipline in our study as they have a higher probability of saving.

References

[1] Ashraf, Karlan and Yin (2004)

[2] Gughert, M.K. (2003). "You can't save alone"

[3] Maddala, G.S. (1983). *Limited Dependent and Qualitative Variables in Econometrics*. Econometrics Society Monographs, No.3. Cambridge University Press.

[4] Morduch, J. (1994). "Poverty and Vulnerability," *The American Economic Review*, Vol. 84, Issue 2, Papers and Proceedings of the Hundred and Sixth Annual Meeting of the American Economic Association, 221-225.

[5] Morduch and Armendariz de Aghion (2000)"Microfinance Beyond Group Lending"

[6] Ramaswami, B., Shamika Ravi and S.D. Chopra (2003) "Risk Management in Agriculture" Vol. 22 in State of the Indian Farmer, Academic Foundation and Ministry of Agriculture, Government of India.

[7] Ray, Debraj (1998) *Development Economics*. Princeton University Press.

[8] Townsend, R. M. (1995) "Financial Systems in Northern Thai Villages", *The Quarterly Journal of Economics*, Vol. 110, No. 4, 1011-1046.

[9] Rutherford (2000)

[10] Department of Statistics, Government of India, (2001), *Handbook 2001.*

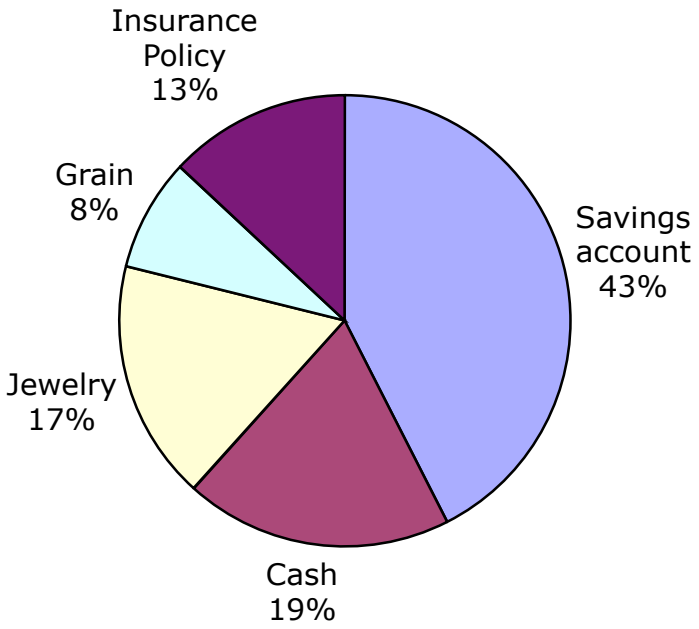
Table1: Some Descriptive Statistics

Income Frequency (No. hh)	Wealth in Rs.	Loan Amount in Rs.	Annual Interest Rate	Actual Repayment frequency per year	Preferred Repayment Frequency per year	Female %	Percentage Reported Savings
Daily	1,73,415	12,311	37.2	8.4	14.21	14.4	43
[323]	[2,20,406]	[23,334]	[48.8]	[4.1]	[6.46]		
Monthly	4,08,761	26,181	20.8	3.2	8.2	13	62
[164]	[4,07,958]	[47,558]	[28.8]	[4.1]	[3.22]		
Quarterly	4,72,974	15,120	21	3.4	3.8	6	35
[20]	[4,64,028	[13,883]	[22.8]	[3.9]	[2.3]		
Half yearly	5,28,058	12,138	32.7	2.8	2.25	11	63
[160]	[6,36,423]	[24,134]	[35.7]	[4.09]	[1.38]		
Annual	4,66,998	8,630	25	2.1	1.46	9	59
[36]	[4,19,031]	[9,485]	[42.1]	[4.2]	[1.12]		
Not Fixed	5,50,343	12,527	32	4	2.68	34	60
[14]	[5,64,754]	[18,218]	[35.2]	[4.9]	[1.28]		

Income frequency for household is assigned depending on the frequency of largest fraction in the annual income. E.g. household whose largest fraction of income is from daily wages is assigned 'daily' as income frquency. 76% of households have multiple sources of income.

Figure1: Descriptive Statistics on Savings

Forms of Saving



Most Important Reason for Saving

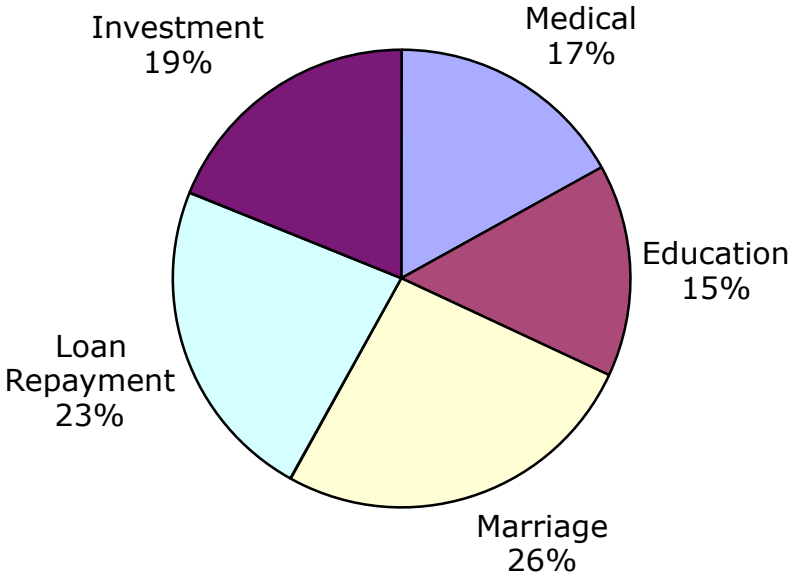


Figure 2: Fitted Values - Income and Loan Repayment Frequencies

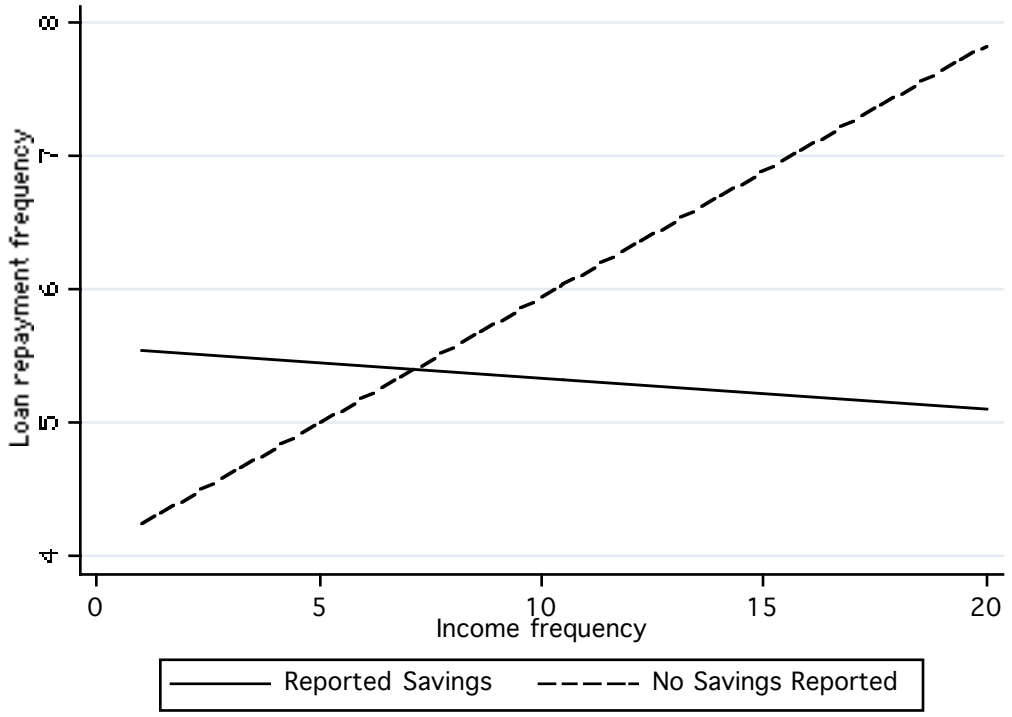


Figure 3: Fitted Values- Income and Loan Repayment Frequencies

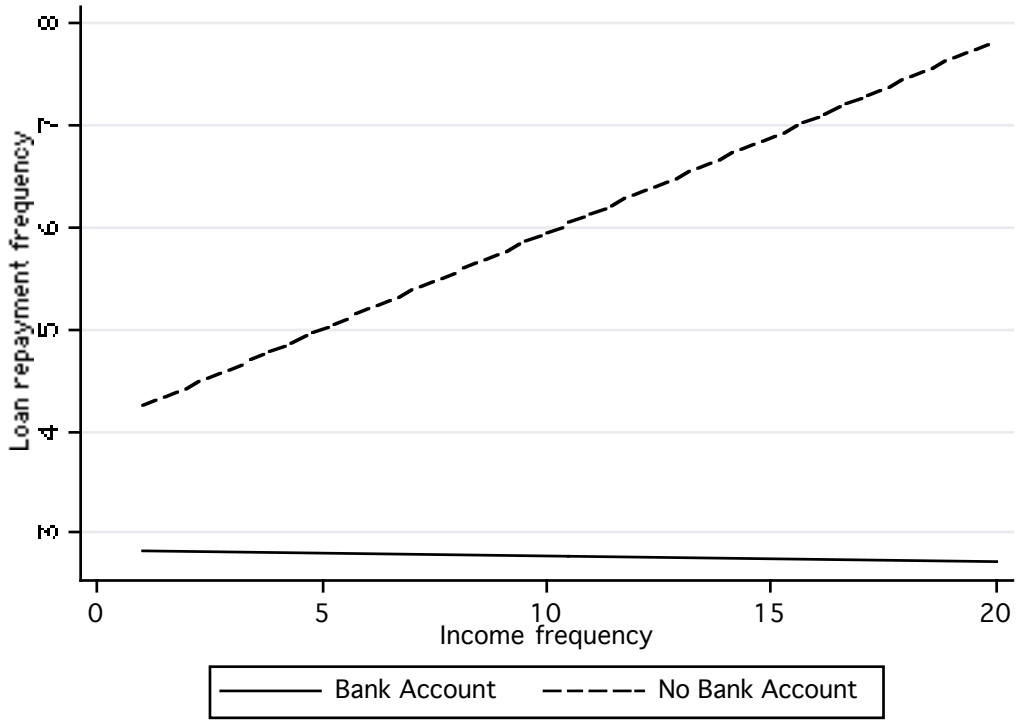


Table 2: Statistics of Variables in Regressions

Variables	Mean	Standard	Sample size
Preferred repayment frequency per year	7.4	4.2	691
Actual repayment frequency per year	5.7	4.3	738
Income frequency per year	20.3	18.2	691
		Household Characteristics	
Logarithm of annual income	10.04	1.38	691
Logarithm of total assets	11.78	1.48	691
Number of years of education	6.07	0.83	691
Age of household head	52.08	13.55	691
Total number of members	6.18	3.53	691
Number of dependents	1.69	1.98	691
Female headed household	0.13	0.34	691
Household head> 60 years	0.36	0.48	691
Schedule caste/tribe category	0.21	0.4	691
Muslim	0.09	0.28	691
Farmer	0.39	0.49	691
Casual labor	0.37	0.48	691
Regular salaried job	0.13	0.39	691
Self employed	0.08	0.27	691
Unemployed	0.03	0.14	691
		Loan Characteristics	
Logarithm of size of loan	8.806	1.3	738
Length of loan (years)	1.99	2.7	738
Use of collateral to secure loan	0.617	0.48	738
Loan use =production	0.53	0.46	738
Loan use=consumption/medical	0.451	0.47	738
		Credit Relations	
Reported some form of saving	0.61	0.49	691
Savings account in any formal institution	0.345	0.47	691
Number of moneylenders in village	4.38	6.33	691
Distance to nearest bank (kms)	3.4	2.08	691
Presence of co-operative society in village	0.72	0.27	691
Borrow from one lender	0.81	0.5	691
Borrow from more than one lender	0.19	0.38	691
		Type of Lender	
Friends and relatives	0.21	0.41	738
Bank	0.21	0.41	738
Co-operative society	0.28	0.45	738
Moneylender	0.24	0.44	738
Trader-employer-landlord	0.04	0.21	738

Sample size - 691 households and 738 loans

Table 3: Determinants of Savings

	Household holds savings	
Explanatory Variables	Coefficient	z-statistic
Nuclear Family	0.134	(4.10)**
Female	0.027	-0.37
Total number of members	-0.017	(2.12)*
Number of dependent mems	0.069	(4.86)**
Age of household head	0.03	(2.57)*
Farmer	-0.047	-0.45
Casual labor	-0.071	-0.73
Unemployed	0.045	-0.24
Regular job	0.157	-1.43
Education	0.015	(2.73)**
Log_income	0.047	(2.93)**
Log_(house value)	0.042	(2.73)**
Log _land	0.017	(3.15)**
Log_asset	0.081	(10.96)**
Income_frequency	0.001	-1.21
Number of moneylenders in village	-0.001	-0.31
Distance to nearest bank	-0.008	-0.66
Cooperative society	0.411	(4.33)**
Kerala dummy	-0.422	(5.34)**
Constant	-2.735	(6.88)**
Observations	691	
Pseudo R-squared	0.33	
Chi-squared	144.5	

Probit estimation with robust standard errors, dependent variable takes value 1 if household holds savings, 0 otherwise. Absolute value of z statistics in parentheses; * significant at 5%, ** significant at 1%

Table 4: Preferred Repayment Frequency

	Preferred repayment frequency		
Explanatory Variables	**(I)**	**(II)**	**(III)**
Income Frequency	0.322	0.026	0.02
	(2.20)*	1.03	1.38
Saving Dummy		-0.24	-0.262
		(2.57)*	-1.41
Saving*Income_frequency		-0.009	0.002
		(2.16)*	-0.16
Transaction_cost	-0.02	-0.026	-0.006
	(8.85)**	(10.57)**	(4.21)**
Female	0.24	0.14	0.024
	(2.02)*	(2.17)*	-1.71
Income	-0.22	-0.041	0.07
	(2.3)*	(2.08)*	-1.38
Total Asset	0.18	0.031	0.074
	(2.34)*	(2.56)*	(2.29)*
Nuclear family	-0.44	-0.24	-0.156
	(3.89)**	(5.85)**	-1.39
Female*Income_freq	0.2	0.18	0.18
	(2.65)**	(2.94)**	(2.36)*
Nuclear*Income_frequency	0.41	0.017	0.028
	(1.99)*	(2.43)*	-1.14
Total Asset*Income_Frequency	-0.026	-0.002	-0.013
	-1.18	-0.98	(2.79)**
Collateral*Income_Frequency			0.027
			(3.08)**
(LoanUse=Production)*Income_Frequency			0.007
			-1.26
Constant	2.582	1.495	0.067
	(2.19)*	(3.81)**	-0.07
Observations	691	691	468
Pseudo R-squared	0.28	0.32	0.3
Chi-squared	88.34	73.7	43.2

Maximum Likelihood Ordered Probit estimation with robust standard errors clustered at the household level, Dependent variable is Preferred Repayment Frequency for each hosuehold, (I) is without savings dummy and interactions; (II) includes savings dummy and interactions; (III) is for households with current outstanding loans; Absolute value of z-statistics in parentheses;*significant at 5%, ** significant at 1%

Table 5: Preferred Repayment Frequency

Explanatory Variables	**(IV)**	**(V)**	**(VI)**	
Income Frequency	0.006 (2.3)*	0.15 (2.72)**	0.42 (1.99)*	0.18 (2.62)**
Saving Dummy				-0.83 (2.93)**
Saving Dummy*Income frequency				-0.002 (3.12)**
Transaction cost	-0.09 -1.4	-0.13 -1.2	-0.05 (2.2)*	-0.02 (2.79)**
Constant	1.38 -0.88	0.776 -0.51	0.45 -1.22	0.372 -1.28
Observations	60	432	348	348
Pseudo R-squared	0.33	0.47	0.38	0.42
Chi-squared	72.8	167.4	129.4	133.9

Maximum Likelihood Ordered Probit estimation, Dependent variable is Preferred Repayment Frequency for each hosuehold, (IV) is for 60 households which reported 'Loan Repayment'as the most important motive for saving; (V) includes all households with savings; (VI) includes non-saving households and households that save for loan repayment; Absolute value of z statistics in parentheses;*significant at 5%, ** significant at 1%