Sans Joint Liability

A Study of Information Problems in the Absence of Joint Liability and Their Effect on Microfinance Loan Products and Repayment

Jason Wang
Northwestern University
Mathematical Methods in the Social Sciences

Professor Annette Krauss
Kellogg School of Management
Advisor

June 2009
Abstract

Information problems in microfinance have in large part been mitigated by the use of group lending and joint liability mechanisms in microfinance institutions around the world. However, after rethinking the model on a practice standpoint, the Grameen Bank, a prominent microfinance institution which pioneered the group lending process, changed their policies and methods in 2003 as part of the shift to “Grameen Bank II.” In this shift, the long-tested mechanism of joint liability was abandoned while maintaining most aspects of the group lending model. This paper examines the consequences absolving joint liability have on information problems. Our theoretical framework examines the enforcement costs required to prevent moral hazard relating to housing loans. Empirically, we subjected our hypotheses to the tobit model and the logit/probit models: we found that (1) more previous loans taken did not result in higher repayment rates, (2) higher personal savings, a proxy for personal wealth, also did not statistically affect repayment rates, and (3) consumption loans, on average, do not perform as well as income-generating loans. Hence, we find that the transition away from joint liability was successful in maintaining the loan quality of income-generating loans, though consumption loans continue to suffer amid product and banking reforms.
Acknowledgements

This thesis was written under the direction of Professor Annette Krauss at the Kellogg School of Management, who has provided invaluable support and advice throughout the development of this work. Even despite having to communicate long distances, both during my stay in Bangladesh and her return back to Europe, she remained dedicated and willing to help guide me through the ups and downs.

I would also like to give a special thanks to my family. I am grateful to my mother and father, for both the emotional support as well as the financial support that made my trip to Bangladesh possible. To my brother, Stephen, thank you for your work in proofreading my thesis despite your busy schedule.

To the Grameen Bank and all the friends and colleagues I have met there, thank you very much for your hospitality, resources and time. Special thanks to my coordinator, Babor Ali, and the Chief of the International Department, Ratan Nag, for making my research possible. I am also grateful to Samir Chowdhury for his assistance in my research both during and after my stay in Bangladesh.

Special acknowledgement should also be made to AIESEC, both at Northwestern University and in Bangladesh, for their assistance with my travels. At Northwestern University, special thanks to Eric De Feo for coordinating my rather spontaneous plans. In Bangladesh, I would like to thank, in particular, Neda Shakiba and Farihan Rahman, for their incredible hospitality and friendship.

Finally, I would like to thank my friends, most notably, Amol Deshpande, Jonathan Kent and Eric De Feo, for their indispensable laughs. And to Caitlin Daley, thank you for your support throughout the years, and in particular, your help proofreading the drafts.

I dedicate this work to all the friends I have made during this endeavor, without whom I would not have discovered the world that I hope to help change.
**Introduction**

Access to capital has long been essential for countries, corporations, and individuals around the world. These needs have been met by the advent of trusted institutions such as stock markets and banking establishments around the world. However, the poor’s access to capital has only been a recent phenomenon, starting in 1974 with the founding of the Grameen Bank in Bangladesh. Though new institutions have developed other methods of lending to the world’s poorest people, many of their practices are based on the ideas and innovations of the Grameen Bank. We will focus on Grameen’s system, a system that has been replicated at over 143 microfinance institutions in over 38 countries around the world (Grameen Dialogue, p.16). By looking at the Grameen Bank on the micro-level, our conclusions can be applied to microfinance institutions around the world.

The idea of providing capital to the poor challenges many of the fundamental banking mechanisms created to make the economics of lending money make sense. The Grameen Bank faces borrowers who do not have access to financing due to the fact that much of the 10-12 million borrowers have been written off by traditional commercial banks as “unbankable” (Armendáriz and Morduch, p.2). What makes these borrowers “unbankable” is the fact that they do not have collateral to put up against their loan, or that the cost of seizing the collateral, in the case of default, is unprofitable (Yunus BP, p.185). Thus, without collateral, new mechanisms must be put in place to ensure the sustainability of such lending.

Mechanisms such as group lending and joint liability act as the basis for dealing with the issues of adverse selection and moral hazard that result from taking away the collateral requirement; these mechanisms have made traditional microfinance banking a reality for
millions of poor people around the world, the success of which has led to the development of new products designed for the poor, including the housing loan, the microfinance loan, and others.

Though in large demand, the housing loan has not kept up on the supply side. Part of the reason lies with the difference in loan type—the traditional microfinance loan was granted for income-generating purposes, while the housing loan is a consumption loan that does not directly generate income. Thus, much of the criteria used to screen borrowers have been based on their ability to pay back their loans. The Grameen Bank looks closely at their borrowers’ repayment schedules, as well as their current savings as the strongest indicator of the ability to pay back these consumption loans (Barua, p.8).

In 2003, the Grameen Bank stepped back from their innovative group lending and joint liability model and reformed it by removing the joint liability mechanism. By doing so, borrowers were no longer financially responsible for other members in their respective groups—now, borrowers who were having trouble paying back their loans go to the Grameen Bank and renegotiate the terms of their loans (Yunus GBII, p.3). The removal of this important tenant of micro-lending was not only an important step away from microfinance’s dependence on the group structure but also led to the rise of past information challenges.

By using data collected from the Grameen Bank from the Gazipur district in Bangladesh, this paper will look at the revived information problems with regard to the housing loan product, present the issues and concerns microfinance institutions will face in light of these problems, and provide empirical evidence regarding the viability of alternative mechanisms set in place.
Literature Review

Research in the field of housing microfinance is lacking; moreover, most literature tends to focus on the practical implementation of housing programs rather than the economics and viability of housing microfinance. Indeed, housing microfinance has not been in the limelight of international and foundation funding, as Daphnis believes the product is wedged between microfinance loans and mortgage loans (Daphnis). Due to the awkward positioning of the product and the lack of research in the field, the supply of funding for housing microfinance products is small. However, the demand for such products is extremely high, as over half of the world’s population remains unserved by formal housing financing (Ferguson, p.3). According to Ferguson, traditional housing finance fails to reach low-income households because (1) the cost of servicing such housing loans is too high compared to the low-income borrowers’ ability to pay, (2) the long term nature of the housing loans do not match the earning patterns of low-income borrowers, and (3) traditional financiers have no incentive to reach the lower bracket of the population. Thus, a gap is clearly present both within the availability of such financing for low-income households as well as the academic research in the field.

Progressive Lending

Lending without collateral has often been the main goal for many microfinance institutions around the world. With no collateral, two agency problems arise which must be addressed: adverse selection and moral hazard. Adverse selection takes place before a lending contract is agreed upon, underscoring the idea that banks are unable to discriminate between safe and risky borrowers and thus must charge high rates of interest to compensate for the
incomplete information (Stiglitz and Weiss, p.401). This concern is much more important when dealing with housing loans, as the amount of housing loans are larger than traditional microfinance loans on average (Barua, p.4). By using such mechanisms as progressive lending and savings schemes, the Grameen Bank has been able to keep interest rates on housing loans at 8%.

The Grameen Bank uses the progressive lending mechanism to avoid much of the risk regarding incomplete information. Armendáriz and Morduch find that the usage of progressive lending at the Grameen Bank allows the bank to quickly cut the average cost of servicing each loan (Armendáriz and Morduch, p.125). Ghosh and Ray further validate this point, suggesting that the lender “tests” the borrowers with smaller loans and weeds out the risky ones (Ghosh and Ray, p.17). Armendáriz and Morduch suggest a third impact progressive lending has on the adverse selection problem—by offering larger loans to those who are most likely to pay back, progressive lending can “increase the opportunity cost of non-repayment and thereby discourage strategic default even further” (Armendáriz and Morduch, p.125). The Grameen Bank currently requires all housing loan applicants to meet strict criteria—housing loan borrowers must be third time borrowers (that is, they must have taken out and paid back at least 2 previous basic loans) with perfect repayment rates (Barua, p.8-9).

**Moral Hazard**

Moral hazard is the second agency problem that makes the cost of low-income lending very high—it comes about when lenders are unable to observe the borrower’s effort, or actions, in maximizing their objective functions for their respective projects. Traditional banks are able
to give incentive to borrowers to maximize their objective functions, or in most cases, profit, by taking collateral on the loan—if a borrower cannot generate enough cash to pay back the loan and interest payments, then their collateral is seized. Without collateral or other mechanisms in place, borrowers are able to take the money and run. We can examine the moral hazard agency problem in two parts—ex ante moral hazard and ex post moral hazard.

Ex ante moral hazard refers to the efforts and actions a borrower can take after receiving the loan disbursement but before the returns to the loan projects are fully realized. Indeed, these actions can greatly affect the both the performance of the loan as well as the success of the project. Ex post moral hazard, or the enforcement problem, is a condition in which borrowers face difficulties both after the loan is disbursed and after the loan money has been invested. In the case that the bank does not closely observe the profits of the borrower’s business, the borrower can also make false claims regarding the performance of his business—even while a large profit was made, the borrower could claim a loss and thus not repay the loan.

There are unique moral hazard problems that emerge with housing loans due to the nature of housing loans. Housing and shelter have long been main tenants of human rights (United Nations, p.2) and have even been adopted by the Bangladesh constitution as one of the five basic needs of citizens and should be of utmost importance for the government (Khanam, p.1). Thus, interest rates on microfinance housing loans are typically less than interest rates on business-backed microfinance loans—at the Grameen Bank, business-backed loans are disbursed at a 20% interest rate while housing loans are disbursed at an 8% interest rate at the time of writing (Hoek-Smit, p.38). According to Hoek-Smit, had the 10-year housing loans been at a rate of 20%, the target group would not be able to afford this type of product. Thus, the
bank uses cross-subsidization, subsidizing housing loans using funds from other loan products, to allow for feasible interest rates on housing loans.

Due to the difference between the interest rate on microfinance housing loans and business-backed microfinance loans, borrowers are constantly tempted to use funds allocated for housing loans on other aspects of their lives. However, funds from business-backed loans can also be used to pay for housing purposes. In a 1993 study, Rahman and Hasnat found that Grameen Bank housing borrowers were building homes larger than what was specified to the bank (Rahman and Hasnat). While the authors claim that the additional capital came from additional sources of income and savings, they failed to rule out that other loans could have been the sources of the additional capital. Because the Grameen Bank puts a 25,000 taka limit on housing loans, which is only a fraction of what is needed to build a complete house (Barua, p. 4), additional funds are needed to finance these housing projects. Islam, Chowdhury and Ali found in in-depth interviews with 45 borrowers, only 12 borrowers kept the cost within the credit amount, and others spent over 50-100% more than the borrowed amount (Islam, Chowdhury, Ali, p. 17). Thus, we see that the moral hazard problems are even more important when looking at housing loan products, relative to business-backed microloans.

Housing loans are defined as consumption loans—loans in which lenders disperse for a purpose that does not generate income. These loans were developed following the success of the basic loans, which were lent for income-generating purposes—at the Grameen Bank, the housing program began in 1984, almost a decade following the introduction of the first microcredit loan (Yunus BP, p. 128). These consumption loans were different in that they did not have any cash flow to support the loan—all the interest payments had to be from some
other source of income from the household. Because of this problem, these loans are typically riskier to microfinance institutions, which lack information on their borrowers and thus need some reassurance that the borrower will be able to repay the loan plus interest. However, Rahman finds that in studies done in specific microfinance institutions in Bangladesh, around 50% of the borrowers were observed using, on average, 66% of their total income-generating loan for the business purpose, while the rest was spent on consumption purposes (Rahman). Many other studies back findings in which borrowers use part of their loan money for consumption purposes, including a relatively young microfinance institution in Indonesia, where 70% of borrowers interviewed spent their loan money for consumption rather than on their new business. Thus, we see that moral hazard problems can lead to big problems for microfinance institutions; however, it is important to determine the feasibility of enforcing loan money usage and its effect on the quality of the loan portfolio.

A number of models exist in current literature regarding the information problem of moral hazard, though none explicitly target the housing loan problem. Ghatak and Guinnane look deeply into the information problems, empirically studying how joint liability mechanisms can affect moral hazard problems. The authors find that despite simple joint liability mechanisms, each borrower will still choose the same risk-level project as in an individual liability situation (Ghatak and Guinnane, p. 204). This point adds to the large shift in policy at the Grameen Bank, taking financial joint liability mechanisms away from their lending model and providing flexible loan terms in the case of a defaulted borrower (Yunus GBII, p.2). Because of the shift away from joint liability both at the Grameen Bank and at other institutions, we will
illustrate the points of the moral hazard model that are still relevant to an individual liability situation.

**Ghatak and Guinnane Moral Hazard Model**

A simple setup is used in this particular model where $Y$ takes on either a high value ($Y_H$) or a low value ($Y_L$), with $Y_L$ being normalized to zero. $Y_H$ is the output from which a successful venture produces while $Y_L$ is the output from a failed venture. The model assumes that the bank follows a zero-profit model where the objective of the bank is to provide capital at the cheapest form possible while maintaining sustainability (no losses). Thus, the bank must always satisfy a simple constraint: $pr = \rho$. Here, $r$ is the gross interest rate, that is, principal plus the net interest, which is the amount that the borrower should pay back to the bank; $\rho$ is the amount the lender expects to be paid back; and $p$ is how likely the borrower will pay back the loan. In this model, $p$ can also be thought of as the level of effort the borrower puts into the project and thus translate directly to the probability that the project is successful—thus $p \in [0,1]$. There is a disutility cost of putting in effort, which is expressed as $\frac{1}{2} \gamma p^2$, where $\gamma > 0$.

Since the choice of $p$ by the borrower is unobservable by the bank, $p$ is subject to moral hazard and therefore the choice will depend on the interest rate given by the bank. The borrower faces the following profit function, with the objective to maximize profit:
Examining the interest rate $r$, we see that it is only paid when the venture is successful; if the venture is a failure, the borrower does not pay the interest rate, nor does he repay the principal. Because of this,

$$p^* = \hat{p}(0) \text{ if } \hat{p}(r) < \hat{p}(0)$$

(2.2)

Substituting this result into the bank’s zero-profit condition for $r$, $pr = \rho$, we get that the borrower will choose

$$p = \frac{Y^H + \sqrt{(Y^H)^2 - 4\rho\gamma}}{2\gamma}$$

(2.3)

We see several interesting relationships that develop from this model: the higher the cost of effort, the less likely the borrower will put in enough effort to make the venture successful. Also, we see that the higher the reward for good effort, the more likely the borrower will put enough effort into the venture.

Several changes can be made to Ghatak and Guinnane’s moral hazard model to analyze how the moral hazard problem with regard to housing loans. First of all, instead of having to choose the effort level, the borrower now must choose to take one of two products: a business loan (with a higher interest rate), or a housing loan (with a lower interest rate). The moral hazard problem only arises when a borrower needs a loan for a business, but instead uses the
loan made for housing on his business. The bank is stuck in an interesting situation regarding these two loan products.

In terms of the risk level, housing loans are inherently riskier than business loans, all else equal. From the principles of finance, the risk-return relationship should result in a higher interest rate for the riskier loan, in this case being the housing loan. However, because of the reasons mentioned earlier, housing loans are subsidized and set at a lower interest rate. Therefore, when we look at the probability of success, the probability that one will succeed on a business loan will be higher than that on the housing loan. Moreover, as the housing loan is a consumption loan, we realize that the probability of “success” does not apply here, but rather the probability of the loan being paid back is more important.

Lastly, enforcement remains a problem for the bank, as it can be extremely costly. However, it is necessary to maintain the quality of the portfolio and keep the interest rates low. On a legal standpoint, the banks can only punish borrowers through the threat of not lending to him in the future. Thus, we see that the borrower will only repay if the interest cost is less than the net benefit from all future loans (Ghatak and Guinnane, p.209). By refusing to offer loans in the future, this method can be a viable threat to borrowers if the bank can spare resources to enforce loan usage and delinquent payments.

**Repayment Metrics**

Microfinance institutions look to report repayment statistics as a sign of the health of their loan portfolio as well as a claim on their ability to remain sustainable—however, the methods of reporting such statistics can be deceiving to the observer who is ignorant of the
calculation of such statistics. The most used measure of repayment takes on two distinct types of indicators: collection rates and the portfolio-at-risk rate. Collection rates typically measure amounts that are actually paid against the amount that are supposedly due. Portfolio-at-risk, however, takes the timing of loan repayment, measuring the outstanding balance of loans that have or have had overdue payments against the outstanding loan portfolio.

Collection rates are the most widely reported statistic regarding repayment, though they can have a number of disadvantages. Most microfinance institutions report the collection rate because it uses variables that are absolutely critical to the bookkeeping of most microfinance institutions—the amount collected by the microfinance institution and the amount owed by the borrower (CGAP, p.2). This measurement is static, taken at a specific point in time using relatively simple measures. The Consultative Group to Assist the Poorest (CGAP), a prominent researcher in the field of microfinance, published a paper regarding repayment statistics, presenting several points by which different repayment statistics should be examined. Collection rates fair decently well, passing tests that estimate the actual loan losses likely to result (“bottom line”), do not encourage inappropriate renegotiation or write-off policies (“smoke and mirrors”), and help management estimate cash receipts from the portfolio in future periods (“cash flow”) (CGAP, p.19). However, collection rates do not pass the “red flag” test, which highlights day-to-day operational issues.

The preferred measure of portfolio quality is the portfolio-at-risk measure, though it exhibits a few faults. While the portfolio-at-risk measure did quite well at the red flag, bottom line, and smoke and mirrors tests, the measure fell short when looking at how much it helps management estimate cash receipts from the portfolio in future periods (CGAP, p.19). However,
because this is the international banking standard, it is an important measure that can be used to compare loan performance across countries and regions. The main drawback for microfinance institutions, however, is that the PAR rate requires a well-managed loan and repayment procedure as well as the ability to record transactions accurately and precisely.
Research Framework

A review of the existing literature in the field yields a substantial number of areas that require further study. Using a unique dataset from an established microfinance institution, the Grameen Bank, this paper aims to further the discussion in the following areas with regard to the shift away from joint liability:

1. **Moral Hazard** — the moral hazard problem has always been identified as an important agency problem micro lending must deal with. While there are mechanisms that are put in place to counter this problem, many mechanisms like joint lending are being reviewed and discarded (i.e. Grameen Bank II). When a microfinance institution provides two products, one being an income-generating loan with a relatively high interest rate and the other being a consumption loan (housing loan) with a relatively low interest rate, what concerns should microfinance institutions be aware of when providing these loans?

2. **Adverse Selection** — controlling for relevant variables, progressive lending theory indicates that repayment rates improve with the number of previous loans. The progressive lending mechanism is an extremely important part of the Grameen housing program, as only borrowers with 2 previous loan cycles are eligible to apply for the housing loan. Therefore, while existing literature provides for the theory behind progressive lending, does the data support the theory behind progressive lending?

3. **Repayment** — the largest contributor to a healthy, sustainable microfinance institution is its ability to keep repayment rates high. In fact, besides factors such as the cost of
acquiring and providing capital, the borrowers’ ability to repay their loan can greatly affect the effective interest rates microfinance institutions charge on their products. What characteristics of the borrower and the loan product can help explain a borrower’s ability to repay loans?

We begin answering these questions by introducing a theoretical framework—the current literature on moral hazard and on housing microloans does not address the seemingly important question regarding the moral hazard of housing loan usage. Thus, we will present a model looking at this problem and see how different variables can affect how a microfinance institution will behave when presented with these information problems.

Housing Loan Moral Hazard Model

Let us set up the following game in which a borrower can be either two types: H or L, corresponding to his need of the loan, or what he plans to use his loan on. Type H borrower will be using his loan money on an income-generating business or purpose, while the type L borrower will be using his loan money on a consumption basis like housing improvements.

Suppose the microfinance institution initially offers loans for income-generating business. Thus, as denoted in the Armendáriz and Morduch model, the bank will charge a gross interest rate $R_H$ such that the bank can cover its costs of providing that capital (that is, the full cost of raising money from depositors, donors, and other agencies), $k$ per unit, based on the risk of that loan (Armendáriz and Morduch, p.38): $k = R_H [p_H Y]$, where $Y$ is the output of a
successful venture and \( p_H \in [0,1] \) is the probability of a successful venture. Thus, the bank would charge \( R_H = \frac{k}{[p_H Y]} \), on the assumption that the bank’s objective is to provide capital to borrowers at the cheapest rate.

Now, suppose the microfinance institution would like to add a housing product to their loan products, offering a lower rate \( R_L \) to borrowers based on the ideas presented in the previous chapter. Let \( p_L \in [0,1] \) be the probability of a type L borrower paying back the loan. Since a type L borrower is borrowing for a consumption loan, we would expect that \( p_L < p_H \). Because the risk is higher for a consumption loan, the bank must subsidize the cost of making that loan in order to keep the interest rate low. As mentioned in the previous section, microfinance institutions must use cross-subsidization to keep interest rates low for products such as housing loans. Thus, in our model, the bank must make a profit on the higher rate loan in order to subsidize the lower rate loan—we will allot an amount \( S \) for the cross-subsidization.

Assume that the bank has full information regarding the type of borrower with certainty—the bank will identify the borrower type and offer him the respective rate based on his type. Thus, the type H borrower will receive the following rate:

\[
R_H = \frac{k + S}{[p_H Y]}
\]

Notice that because of the introduction of the housing loan, the income-generating business borrowers now must pay a premium on their loans. The type L borrower will receive the following rate:
\[
k = R_L[p_L Y] + S
\]
\[
R_L = \frac{k - S}{[p_L Y]}
\]

(3.2)

Clearly, the type L borrower pays a lower rate than the type H borrower due to the cross-subsidization.

Now, assume that the bank has no knowledge of the borrowers’ type, and thus must offer two rates based on the probability the borrower will deceive the bank. Let \( q \in [0, 1] \) be the probability a borrower will deceive the bank. If the borrower identifies himself as a type H borrower to the bank, the bank knows that there is no incentive for the borrower to identify as a type H borrower but actually be a type L borrower, and thus lends him money at a rate of

\[
k + S = R_H[p_H Y]
\]
\[
R_H = \frac{k + S}{[p_H Y]}
\]

(3.3)

However, if the borrower identifies as a type L borrower, then the bank must offer a rate that takes into account the incomplete information:

\[
k = R_L[q p_H Y + (1 - q)(p_L Y)] + S
\]
\[
R_L = \frac{k - S}{q p_H Y + (1 - q)(p_L Y)}
\]

(3.4)

The type H borrower is not affected by the bank’s lack of knowledge on the borrower’s type since they are happy to lend at the higher rate. However, comparing \( R_L \) in the case the bank has perfect information and in the case of incomplete information, we see that since \( p_L < p_H \),

\[
R_L^{\text{complete}} = \frac{k - S}{[p_L Y]} > \frac{k - S}{q p_H Y + (1 - q)(p_L Y)} = R_L^{\text{incomplete}}
\]

(3.5)
If \( R_{L, \text{complete}} > R_{L, \text{incomplete}} \), then there is no reason the banks should be worried about this particular moral hazard problem. The case is that, rather, without knowing their borrower type, the banks can achieve a better rate for their consumption loan. We see that this is because \( p_L < p_H \), which defies the standard risk-return principles. In accounting for type H borrowers taking out the lower interest rate loan, the bank is able to decrease the rate since type H borrowers are more likely to pay back their loans.

However, despite the conclusion that banks would not worry with the moral hazard problem, especially since it seems to imply the problem actually lowers the interest rate the bank can quote, it is extremely important for microfinance institutions to make sure the money it loans is going to where borrowers say the money is going to go. There are three important reasons why banks need to be sure where their loan money is going. First and most importantly, banks face a huge demand for housing loans, especially in countries like Bangladesh in which many houses need improvements or funds for repairs. If the allocated loan portfolio for housing is being used by borrowers who are not using the loans for their specified purpose, then the allocated money is not getting to where it is supposed to go. Second, the bank has an incentive to make sure the information and risk profile on their loans are as accurate as they can be, as banks must monitor and maintain their loan portfolios. Without reliable information on the loans they give, they cannot make decisions and assessments on the loan portfolio level. Third, many microfinance institutions are supported by government grants, government subsidies, and other international organization funding support. Many of these funding sources require that the microfinance institution have the ability to track their money in a reliable way. Because
of these reasons, microfinance institutions must enforce, or monitor, the usage of the loans they disperse.

Suppose now that the bank decides to monitor and enforce loan money usage to crack down on the moral hazard problem. Thus, let \( c \) be the additional monitoring cost per unit that will be the added to the price of maintaining a lower interest rate product. Also, let \( u \) be the probability that a deceiver will be caught by the bank. Since the more the bank invests in monitoring and enforcing, the more likely they will be able to detect deceivers, let \( u = 2c \), where \( u \in [0,1] \) is the probability that a deceiver will be caught. Thus, since \( u = 2c \), we must constrain the domain of \( c \) such that \( c \in [0,0.5] \). Assuming again that the bank cannot see the type of each borrower unless through monitoring and detection, if a borrower identifies himself as a type H borrower, then based on similar reasoning in the previous cases,

\[
R_H = \frac{k + S + c}{[p_H Y]} 
\]

(3.6)

Compared to the case with no monitoring, we see that \( R_H \) is now higher by \( c / [p_H Y] \). If the borrower identifies as a type L borrower, then we see that

\[
k + c = R_L[uqp_H Y + (1-q)(p_L Y)] + S \\
R_L = \frac{k + c - S}{uqp_H Y + (1-q)(p_L Y)}
\]

(3.7)

Notice that the addition of enforcement only adds costs to both the low rate as well as the high rate. Moreover, the increase in \( c \) also decreases the magnitude of the denominator of \( R_L \), thus making \( R_L \) larger. The additional cost not only increased both interest rates, but also increased
even more so. However, this makes sense as no punishment was built into the model; hence, we must look at this game with multiple periods.

Since microfinance institutions traditionally do not take collateral, the only form of punishment they can give is by threatening to cut off the line of credit to the borrower. By looking at the dynamics on a multi-period level, we can find some insight regarding the trade-off between enforcement costs and benefits. Let \( t \) be the period of the game, and for this case, let \( t = 2 \). We will assume again that the borrower is risk neutral and that his objective function is to maximize profit. Thus, defining the profit function of the different types of borrowers,

\[
\pi_{H,H} = p_H (Y - r_H) \\
\pi_{L,L} = p_L (Y - r_L) \\
\pi_{H,L} = p_H (Y - r_L)
\]

where \( \pi_{i,j} \) is the profit of a borrower who identifies himself as a type \( i \) borrower and is actually a type \( j \) borrower. Again, along the same line of reasons in the previous cases, we see that the bank has no reason to expect a borrower who identifies as a type H borrower but is actually a type L borrower. As we are looking at the benefits and costs of enforcement, we would like to find the value of \( c \) that would deter a rational borrower from lying about their type. We will also use \( R_L = (k + c - S) / (p_L Y) \), since the interest rate used will be assuming the bank is certain of the borrower’s type. Thus, for a deceiver, his profit after two periods is

\[
E[\pi] = \pi_{H,L} + (1-u) \pi_{H,L} \left( 1 + \delta \right) 
\]

(3.8)

where \( \delta \in [0,1] \) is the interest rate, or opportunity cost, of the borrower. For the non-deceiver, his profit after two periods is
$$E[\pi] = \frac{\pi_{H,H}}{1+\delta} + \frac{\pi_{H,L}}{1+\delta}$$

(3.9)

Since the borrower is looking to maximize his profits, a type H borrower will choose to identify as type H if

$$\frac{\pi_{H,H}}{1+\delta} + \frac{(1-u)\pi_{H,L}}{1+\delta} > \pi_{H,L}$$

(3.10)

Thus, the bank will choose an enforcement budget such that the borrower will be deterred from deceiving. Substitutions yield the following equation:

$$p_H[Y - \frac{k + S + c}{p_H Y}] + \frac{p_H[Y - k + S + c]}{1+\delta} > p_H[Y - \frac{k - S + c}{p_L Y}] + \frac{p_H[Y - k - S + c]}{1+\delta}$$

(3.11)

Solving for $c$, we get that

$$c > \frac{(2p_Hk + \delta p_H + 2p_HS - p_Ls) \pm \sqrt{(p_Ls - 2p_Hk - \delta p_H - 2p_HS)^2 + 8p_H(\delta(p_H + p_Lk + p_LS - p_Hk))}}{-4p_H}$$

(3.12)

To minimize costs, the bank will choose $c^* = c^*$ such that $c$ is minimized, and thus the value of the expression in equation 3.12.

The bank, however, does not bear the cost of monitoring borrowers, as the cost is passed down to the borrower through the interest rate. In a two period game, given that the bank chooses the level of $c^*$, if

$$R_L(k, c^*, s, p_L, Y) < R_L(k, s, p_H, p_L, q, Y)$$

$$\frac{k + c^* - s}{p_L Y} < \frac{k - s}{q p_H Y + (1-q)(p_L Y)}$$

(3.13)
then borrowers are hurt by the added monitoring since they would have to pay more per loan dollar with monitoring, than they would have without monitoring.

**Comparative Statics**

When this problem is looked at on a comparative static level, we find that the level of $c^*$ can fluctuate with changes in $s$, $k$ and $\delta$. As $c$ has two roots, $c^*$ will be chosen such that it is within the domain and is minimized. As the cross-subsidization amount increases, we see that the cost of enforcement increases as well. This effect can be explained by the extra revenue the bank receives from the higher rate loan in order to make the lower rate loan cheaper. The greater the difference between the two rates, the more of an incentive the borrower has to deceive the bank and take the lower interest rate loan.

An increase in the bank’s cost of attaining and providing capital leads to an increase in the enforcement cost as well. The greater the cost of raising and providing the capital, the more the bank must charge to each borrower, thus reducing the borrower’s profit in the case of a successful loan. The reduction of profit will lead borrowers to pursue a lower rate such that their profits are not threatened. Essentially, borrowers see the trade-off between attaining a higher profit with the lower rate and maintaining a lower profit with a higher rate without the threat of being caught. According to the model, however, borrowers on the margin (that is, those who are making very little profit) do not consider taking the lower rate more than borrowers who are making healthy profits, as the output of a successful venture plays no role in calculating the enforcement cost. This phenomenon can be explained by our assumption of a risk-neutral borrower. A more applied model will assume borrowers on the margin are willing
to take more risk than borrowers who are making healthy profits. While it would be an interesting problem to look at, it is not in the scope of this paper.

An increase in the discount rate, or the opportunity cost, of the borrower has a positive effect on the enforcement cost. The discount rate can be interpreted as the patience of the borrower—the more assets or other forms of income a borrower has, the more patient he is; the less assets or other forms of income a borrower has, the less patient he is. The more impatient the borrower is for capital (higher discount rate), the less he cares about future profits and the more he cares about the current need for cash. We also see that the higher a borrower’s discount rate, the greater of an incentive he has to take the first loan out and run, and example of ex-ante moral hazard. Thus, banks must invest more into the enforcement cost to deter the behavior of impatient borrowers.

As we can see in equation 3.12, the values that $p_H$ and $p_L$ hold can significantly change the value of $c$ and change many of the incentives presented in the model. While our assumption that $p_H > p_L$ has been explained, it is important to see what independent variables these two variables are dependent on. Assume a female borrower who has taken out a loan and must decide whether or not to pay back her loan. Assume further that the female borrower faces a hardship and is on the margin between repaying or not repaying her loan on time. One resource she would look to would be her personal savings—assuming she had enough personal savings, she could take some of those savings to pay off her current loan installment. In fact, this is one of the reasons microfinance institutions provide personal savings (or in Grameen Bank’s case, require personal savings), as it can act as a type of financial collateral (Armendáriz and Morduch, p.136). Another resource the borrower will look to would be other forms of income
she has access to. In fact, borrowers of consumption loans would need to finance all of their installments on either other forms of income or sources of financial assets.

Other issues that would affect the borrower’s decision to repay would be the amount of the installment she was responsible for. The lower the interest rate on the loan, the less she would need to pay this period for the interest installment; the lower the principal amount, the less she would need to pay for the principal installment.

A Numerical Example

In order to flush out and to clarify the model presented, we will use hypothetical data to look at the effect moral hazard has on a bank that offers two products: a basic income-generating loan and a consumption housing loan. Assuming again that we have two types of borrowers, type H and type L, and that the bank has no knowledge of their actual type but only the fraction of borrowers who would deceive the bank, which is this case will be 40%. The bank’s objective is to break even, either because the environment is perfectly competitive or that their charter is to provide capital at the lowest cost possible. The bank’s cost of capital is 40 cents on the dollar, which includes the cost of raising the capital as well as the cost of providing those funds to borrowers. The bank must also take into consideration of the probability of not defaulting, which in this case is 90% for type H borrowers and 70% for type L borrowers. We also assume that the wage per period for a borrower is 10 dollars. This wage is thus the opportunity cost for the borrower, and thus his discount rate for a 100 dollar loan would be 10%. Lastly, in order to keep rates low for the housing loan, the bank puts a 30 cent per dollar
premium on the higher rate loan and uses that 30 cents to subsidize the lower interest rate loan.

These assumptions are summarized in Table 3.1 below.

### Table 3.1: Numerical Example Assumptions

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bank's Cost of Capital</td>
<td>40 cents per dollar</td>
</tr>
<tr>
<td>Worker's Wage = Opportunity Cost</td>
<td>$10 per period</td>
</tr>
<tr>
<td>Cost of Cross-Subsidization</td>
<td>30 cents per dollar</td>
</tr>
<tr>
<td>Percentage of Borrowers Who Will Decieve</td>
<td>40%</td>
</tr>
<tr>
<td>Output of Successful Borrower</td>
<td>2 dollars per $1 loan</td>
</tr>
<tr>
<td>Probability of Success or Repayment</td>
<td></td>
</tr>
<tr>
<td>Type H Borrower</td>
<td>90%</td>
</tr>
<tr>
<td>Type L Borrower</td>
<td>10%</td>
</tr>
</tbody>
</table>

In the first scenario, we will look at the rate the bank would charge in the absence of a second lower-rate loan product. Since the original product is just the income-generating loan,

$$R_H = \frac{k}{p_H Y} = \frac{0.4}{0.9 \times 2} = 22\%$$

Now, consider the addition of a second loan with a lower interest rate. From the equations we derived above, we expect the bank to charge the following interest rates:

$$R_L = \frac{k - S}{qp_H Y + (1-q)(p_L Y)} = \frac{0.4 - 0.3}{0.4(0.9 \times 2) + 0.6(0.7 \times 2)} \approx 6.41\%$$

$$R_H = \frac{k + S}{p_H Y} = \frac{0.4 + 0.3}{0.9 \times 2} = 38.89\%$$

Clearly, through cross-subsidization, the bank is able to reduce the cost of the housing loan at the expense of raising the interest rate on the basic loan.

Suppose the bank now wants to incorporate enforcement of loan usage and decides to punish deceivers by threatening to not lend to them again in the second period. Thus,
borrowers will look at their expected profit from both periods and make a decision on whether it is worth it to deceive the bank for a lower rate on their loan:

\[
p_n[Y - \frac{k + S + c}{p_nY}] + \frac{p_n[Y - \frac{k + S + c}{p_nY}]}{1 + \delta} > p_n[Y - \frac{k - S + c}{p_LY}] + \frac{p_n[Y - \frac{k - S + c}{p_LY}]}{1 + \delta}
\]

\[
0.9[2 - \frac{0.4 + 0.3 + c}{0.9 \times 2}] + \frac{0.9[2 - \frac{0.4 + 0.3 + c}{0.9 \times 2}]}{1 + 0.1} > 0.9[2 - \frac{0.4 - 0.3 + c}{0.7 \times 2}] + \frac{0.9[2 - \frac{0.4 - 0.3 + c}{0.7 \times 2}]}{1 + 0.1}
\]

\[c^* > 0.08\]

The bank will choose \( c^* = 0.08 \) in order to deter deceivers, or that is, the bank will spend 8 cents per dollar loaned on enforcement costs. The cost to the borrowers is positive since the following inequality does not hold:

\[
R_L(k, c^*, S, p_L, Y) < R_L(k, S, p_n, p_L, q, Y)
\]

\[
\frac{k + c^* - S}{p_LY} < \frac{k - S}{qp_nY + (1 - q)(p_LY)}
\]

\[
\frac{0.4 + 0.08 - 0.3}{0.7 \times 2} < \frac{0.4 - 0.3}{0.4 \times 0.9 \times 2 + (1 - 0.4)0.7 \times 2}
\]

\[0.129 < .064\]

In fact, the cost of monitoring for moral hazard is 6.5 cents per dollar loaned by taking the difference of the two rates. Nonetheless, because of the benefits to monitoring mentioned in the previous section, banks will probably monitor loan usage despite the cost.

**Hypotheses**

The removal of some aspects of the joint liability mechanism causes microfinance institutions to rely more on other mechanisms to counter information problems such as adverse selection and moral hazard. Progressive lending is one of the most popular methods that
microfinance institutions have used to deal with the adverse selection problem. Thus, it is important that we test the impact such a method has on repayment rates with our data.

A tenant that most banking institutions follow is that the wealthier the borrower, the more likely he or she will pay back the loan. However, this idea keeps many people who need capital away from the credit markets. Microfinance has taken that tenant and tried to show that the wealth of an individual has little effect on their behavior regarding repayment. Thus, it is important that we test this idea as well.

The model presented in this chapter shows the way banks react to this particular moral hazard problem under certain assumptions without joint liability mechanisms. The values that \( p_H \) and \( p_L \) hold are among the most important in this model, as the ability to repay is one of the most crucial issues microfinance institutions must deal with. Thus, it is important to test these crucial assumptions to see if they are used appropriately.

From the framework presented in this section, we see that the research questions described at the beginning of the section can be rewritten to fit the model. The hypotheses formulated for this study are as follows:

**Hypothesis 1:** The more number of previous loans a borrower has taken out, the more likely the borrower will be in repaying his current loan.

**Hypothesis 2:** The greater the personal wealth of a borrower, the more likely the borrower will be in repaying his current loan.
**Hypothesis 3:** Income-generating loans should have higher rates of repayment than consumption loans due to the nature of the loan.
Research Methodology

This chapter will aim to describe the way in which we have conducted our research and how we will use our research to draw conclusions to our stated hypotheses at the end of Chapter 3. We will begin by describing our test instruments, followed by a detailed description of our data set, including thorough explanations of key variables. Then, we will set the framework for our testing model and discuss any other issues and limitations we will run into in the course of our study. By understanding the methodology behind our model, we will be able to better interpret the analysis and results in the next chapter.

The Data Set

This study’s data is from the internal records of the Grameen Bank—prepared and maintained by Grameen Bank personnel. We use the borrowers’ repayment schedule, savings schedule, and other essential information collected by the Grameen Bank for the bank’s own audit and recordkeeping teams. Personal information is collected by the center managers at each branch in paper form. Loan disbursement, repayment and savings transactions are also collected at the branch level. These documents are then sent to the area offices, where the data is reviewed and inputted into the bank’s computer databases. These databases are kept and maintained at the area offices. As the data used are Grameen’s internal records and audited by the Bank’s internal audit team, the data is extremely high quality, especially for data used in the field of microfinance.

The data was obtained through the International Department at the Grameen Bank and was collected at the area office of the Gazipur district on March 2, 2009. After discussions with
the Grameen Bank staff, the data was made available at the area office in the form of archived database files, each reflecting data and status of Grameen borrowers at the calendar year end. Thus, file formatting and data transformation were necessary to retrieve the data from their archived forms.

As the data are the bank’s internal records, the data is strictly confidential. While statistical data summaries can be made available, the data set must be retrieved from the Grameen Bank.

The dataset is a comprehensive collection of all loans still outstanding or paid off in the calendar years 2007 and 2008. While the bank still classifies types of loans into over 10 items, for our research purposes, we will classify all loans into the following categories: the basic loan, the housing loan, the microenterprise loan, the education loan, the modified loan, and other loans. The first 4 types of loans correspond to the exact definition the bank has defined its loans: the basic loan is the first loan product offered to borrowers for income-generating purposes at a 20% interest rate; the housing loan is offered at an 8% interest rate and may only be used for housing purposes; the microenterprise loans are larger in principal and are charged at a 20% interest rate; and the education loan is a loan used for the education of the children of the borrowers with an undefined interest rate. The modified loan refers to all loans that have had their terms changed since the disperse date of the loan—this includes the flexible loan, the straggling member loan, the contact loan and the contractible loan. According to the Grameen Bank II system, implemented in 2002, all members who are unable to make their payments based on the original loan terms have the option of modifying their loan—payments not received for more than 2 years will be considered defaulted loans and thus dropped from the books (Yunus, p.3).
Thus, our data set includes all modified loans since the modification date of January 2005. The “other” loan category is reserved for all other miscellaneous loans that do not fall within the other categories.

According to the documents published from the Grameen Bank regarding the treatment of an overdue borrower, if a borrower fails to pay her installment for 10 consecutive weeks, or if a borrower is unable to repay the full amount (principal) within a six month period following the missed payment, or does not opt to modify his or her loan, the borrower is officially defaults on the loan (Yunus GBII, p.5). Otherwise, these borrowers remain on the books.

Description of Variables

The data collected from the Gazipur database at the area office contains raw data from the years 2007 and 2008, and thus all non-monthly values are the end-of-the-year tallies by bank officials. While the data obtained contains much more data needed for this study, we have included the relevant variables and descriptions of the variables, including any notable issues the data variable might have.
**Table 4.1: Description of Selected Variables: Raw and Generated**

<table>
<thead>
<tr>
<th>Raw Variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BorrowerCode</strong></td>
<td>the borrower code is generated by putting together the branch code, the center code and the individual borrower code. Thus, every loanee has his or her own unique borrower code.</td>
</tr>
<tr>
<td><strong>LoanTerm</strong></td>
<td>the loan term is the number of previous and current loans the borrower has taken out since joining the Grameen Bank. Note that this refers to the number of times the borrower has taken out the specific type (basic, housing, microenterprise, etc.) of loan.</td>
</tr>
<tr>
<td><strong>DisburseDate</strong></td>
<td>the disperse date is the day in which the loan money is handed to the borrower and thus the official date that the loan begins.</td>
</tr>
<tr>
<td><strong>PrincipalLoan</strong></td>
<td>the principal loan is the total principal amount the borrower lends from the bank.</td>
</tr>
<tr>
<td><strong>PersonalSavings07,08</strong></td>
<td>the personal savings is the total outstanding balance in each individual’s savings account with the Grameen Bank at the year-end of 2007 or 2008.</td>
</tr>
<tr>
<td><strong>LoanInsurance07,08</strong></td>
<td>the loan insurance is the total outstanding balance in each individual’s loan insurance account. The borrower is required to put a minimum amount of money (based on the outstanding loan amount and interest) on the last day of the calendar year. This insures that the loan will be forgiven by the bank if the borrower dies.</td>
</tr>
<tr>
<td><strong>LoanRepaid07,08</strong></td>
<td>the loan repaid is the total amount of the principal that the borrower has repaid as of the end of the calendar year 2007 or 2008.</td>
</tr>
<tr>
<td><strong>LoanInstallment07,08</strong></td>
<td>the loan installment is the amount of the principal the borrower must repay at each installment period. This installment calculation is done by a special method defined by the Grameen Bank such that the installments paid each week are equal throughout the entire repayment process. More often than not, the total amount received at the end of each loan cycle is greater than the total amount that is needed to be repaid due to this calculation, and thus the bank will repay the borrower the amount in which he or she has overpaid.</td>
</tr>
<tr>
<td><strong>InterestInstallment07,08</strong></td>
<td>the interest installment is the amount of the interest the borrower must repay at each installment period. See “LoanInstallment” above for calculation details and notes.</td>
</tr>
<tr>
<td><strong>InterestRate07,08</strong></td>
<td>the interest rate is the effective interest rate the bank charges.</td>
</tr>
</tbody>
</table>
Duration07,08  the duration is the specified length of the loan in weekly terms, specified at the beginning of each loan

WeekPassed07,08  the week passed is the number of weeks that have passed on December 31\textsuperscript{st} of that calendar year since the inception of the loan.

**Generated Variables**

**TotalPersonalSavings**  total personal savings is the most recent record of the personal savings of a borrower

**IntRate**  the interest rate is the effective interest rate on the particular loan; it is an aggregate of the 2007 and 2008 records

**Housing**  a dummy variable that is defined as (Housing Loan = 1; otherwise,0)

**Basic**  a dummy variable that is defined as (Basic Loan = 1; otherwise,0)

**Microenterprise**  a dummy variable that is defined as (Microenterprise Loan = 1; otherwise,0)

**Education**  a dummy variable that is defined as (Education Loan = 1; otherwise,0)

**Modified**  a dummy variable that is defined as (Modified Loan = 1; otherwise,0)

**Other**  a dummy variable that is defined as (Other Loan = 1; otherwise,0)

**HousingPurpose**  a dummy variable that is equal to 1 if the specified purpose of the loan is housing related. This includes the following bank-classified purposes: Housing, House Repairing, House Purchase, House Construction, House repairing, Hotel, and Basement Construction. 0 otherwise.

**Repair**  a dummy variable that is equal to 1 if the purpose of the loan was given for House Repair

**Purchase**  a dummy variable that is equal to 1 if the purpose of the loan was given for House Purchase

**Construction**  a dummy variable that is equal to 1 if the purpose of the loan was given for House Construction

**AmtCollected**  the amount collected is the amount of the principal that has been repaid at the end of 2008

**AmtDue**  the amount due is the amount of the principal that needs to be paid to the bank by the end of 2008
Dependent Variables

The dependent variables, or the variables of interest, are two measures of repayment.

The first measure of repayment is $BankLoss$. This variable is calculated on each individual account in the following manner:

$$
BankLoss = \begin{cases} 
0 & \text{if } AmtCollected - AmtDue < 0 \\
AmtCollected - AmtDue & \text{otherwise}
\end{cases}
$$

(4.1)

The $AmtCollected$ is the amount of the principal that has been repaid at the end of 2008 and is calculated by taking the maximum of the $LoanRepaid$ variable for either 2007 or 2008. The $AmtDue$ is calculated by observing the number of weeks that have passed since the disbursement date of the loan and December 31, 2008 and multiplying that by the principal installment, or that is, $AmtDue = WeeksPassed \times LoanInstallment$. However, if $WeeksPassed$ is greater than $Duration$, which means that the loan is overdue, $AmtDue = PrincipalLoan$. If the difference of $AmtCollected$ and $AmtDue$ is negative, then it indicates that after the final payment of the loan, the bank will pay back to the borrower the amount in which the borrower has overpaid. This happens because in the calculation of principal and interest installments, the
bank bases these installments off of their effective interest rates but often round to the nearest 5
taka in order to make repayment of the loans more easily to handle as well as to simply the
repayment schedules for borrowers.

While this metric is not a standard way of calculating repayment as discussed in the
literature review, there are several advantages to this method of calculation as well. Other
metrics such as portfolio-at-risk require frequent repayment schedule data to calculate and can
be quite effective in describing the quality of the microfinance institution portfolio. However,
such a frequent repayment schedule can run into several issues regarding how repayments are
recorded on paper. First off, while microfinance institutions are very strict in maintaining a
consistent form of recordkeeping (as such recordkeeping is essential if the microfinance
institution would like to be sustainable), we cannot overlook the possibility of variance in the
way each center leader marks down the repayments they receive. Second, because center
leaders and branch managers have an incentive to keep books with perfect repayment
schedules, it can be possible that late payments may be recorded as being paid on time. Since
the timing of such payments can highly alter the portfolio-at-risk metric, it is important to keep
such recordkeeping issues in mind when dealing with such metrics.

However, BankLoss, a collection rate metric, has little or no dependence on late payments.
Because this metric is based off of the year-end tally, branch managers and center leaders have
little room to “fudge” the repayment schedule. If a borrower misses a payment on the second
week of December but pledges to repay all the installments up till the middle of January the
following year, branch level personnel would find it quite difficult to record in the books that
she has paid in full up till the end of December. Because of the use of year-end data, we can side step this problem described for the portfolio-at-risk metric.

While it would be worthwhile to use the portfolio-at-risk metric to analyze our data, our dataset lacks the variables (week-to-week repayment schedules) to calculate this metric.

The second dependent variable we will look at is Repay. This variable is simply a binary variable that is based off of our BankLoss metric—0 if BankLoss is positive and 1 if BankLoss has a value of 0. Repay is a measure to see if this particular loan has been paid off or not, with no measure of magnitude. BankLoss, however, does include magnitude, and thus we will use different econometric models to see the effects other variables have on these two measures.

Sample Population

It is important to describe the population of the dataset and to analyze both the merits of the sample population as well as the limitations. As explained above, the dataset is the collection of data gathered from the Grameen Bank at the Gazipur district. As such, all borrowers in the dataset are from the Gazipur district. Gazipur is located just north of Dhaka, Bangladesh’s capital city and the center of the Bangladesh economy. According to interviews with Mr. Babor Ali and Mr. Ratan Nag of the International Department at the Grameen Bank, the Gazipur district is has the highest concentration of textile manufacturing factories in Bangladesh, of which manufacturing contributes to over 17% of the nation’s GDP, the largest of the industry sectors defined by the Bangladesh Bureau of Statistics. Because of the lack of publicly available data by the district (zila) level, we cannot provide statistical data to support these claims. However, because the literacy rate is a strong indicator of quality of life as well as
economic activity (Narayana, p.259), and since the Bangladesh Bureau of Statistics provides data on literacy rates at the district level, we can see below that the Gazipur district is one of the most literate districts in Bangladesh in Table 4.2.
As such, because of the indicators of strong economic activities, we expect loan

<table>
<thead>
<tr>
<th>Zila</th>
<th>All Religions</th>
<th>Muslim</th>
<th>Hindu</th>
<th>Buddhist</th>
<th>Christian</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Barisal Division</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jhalokathi</td>
<td>65.35</td>
<td>64.43</td>
<td>72.63</td>
<td>56.79</td>
<td>70.99</td>
<td>59.21</td>
</tr>
<tr>
<td>Pirojpur</td>
<td>64.31</td>
<td>62.62</td>
<td>71.24</td>
<td>45.53</td>
<td>67.44</td>
<td>55.70</td>
</tr>
<tr>
<td>Patuakhali</td>
<td>51.65</td>
<td>50.77</td>
<td>61.87</td>
<td>61.88</td>
<td>75.96</td>
<td>57.65</td>
</tr>
<tr>
<td>Bhola</td>
<td>36.89</td>
<td>35.68</td>
<td>60.96</td>
<td>53.97</td>
<td>55.56</td>
<td>33.15</td>
</tr>
<tr>
<td><strong>Chittagong Division</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chittagong</td>
<td>55.55</td>
<td>53.79</td>
<td>64.88</td>
<td>70.75</td>
<td>79.45</td>
<td>34.81</td>
</tr>
<tr>
<td>Feni</td>
<td>54.27</td>
<td>53.51</td>
<td>64.57</td>
<td>69.44</td>
<td>78.85</td>
<td>41.33</td>
</tr>
<tr>
<td>Noakhali</td>
<td>51.67</td>
<td>50.99</td>
<td>62.76</td>
<td>74.43</td>
<td>86.02</td>
<td>36.11</td>
</tr>
<tr>
<td>Khagrachhari</td>
<td>41.81</td>
<td>48.93</td>
<td>33.11</td>
<td>37.63</td>
<td>42.71</td>
<td>37.68</td>
</tr>
<tr>
<td>Brammonbaria</td>
<td>39.46</td>
<td>38.88</td>
<td>45.30</td>
<td>80.49</td>
<td>64.94</td>
<td>32.26</td>
</tr>
<tr>
<td>Cox's Bazar</td>
<td>30.18</td>
<td>29.21</td>
<td>42.79</td>
<td>38.20</td>
<td>74.82</td>
<td>19.02</td>
</tr>
<tr>
<td><strong>Dhaka Division</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dhaka</td>
<td>64.79</td>
<td>64.62</td>
<td>65.18</td>
<td>91.44</td>
<td>87.67</td>
<td>62.77</td>
</tr>
<tr>
<td><strong>Gazipur</strong></td>
<td>56.40</td>
<td>56.16</td>
<td>56.45</td>
<td>88.50</td>
<td>79.28</td>
<td>34.15</td>
</tr>
<tr>
<td>Narayanganj</td>
<td>51.75</td>
<td>51.20</td>
<td>60.92</td>
<td>92.03</td>
<td>82.01</td>
<td>40.28</td>
</tr>
<tr>
<td>Munshiganj</td>
<td>51.62</td>
<td>51.21</td>
<td>55.35</td>
<td>45.56</td>
<td>82.60</td>
<td>45.90</td>
</tr>
<tr>
<td>Narsingdi</td>
<td>42.91</td>
<td>41.88</td>
<td>58.10</td>
<td>80.52</td>
<td>61.19</td>
<td>18.41</td>
</tr>
<tr>
<td>Faridpur</td>
<td>40.85</td>
<td>39.30</td>
<td>53.50</td>
<td>66.04</td>
<td>71.89</td>
<td>30.60</td>
</tr>
<tr>
<td>Mymensingh</td>
<td>39.11</td>
<td>38.21</td>
<td>57.89</td>
<td>39.72</td>
<td>56.09</td>
<td>20.78</td>
</tr>
<tr>
<td>Shariatpur</td>
<td>38.95</td>
<td>38.21</td>
<td>56.43</td>
<td>50.00</td>
<td>56.94</td>
<td>23.15</td>
</tr>
<tr>
<td>Jamalpur</td>
<td>31.80</td>
<td>31.50</td>
<td>50.30</td>
<td>13.50</td>
<td>47.10</td>
<td>11.70</td>
</tr>
<tr>
<td><strong>Khulna Division</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bagerhat</td>
<td>58.73</td>
<td>56.40</td>
<td>67.49</td>
<td>79.17</td>
<td>63.74</td>
<td>34.47</td>
</tr>
<tr>
<td>Chuadanga</td>
<td>57.81</td>
<td>56.99</td>
<td>60.20</td>
<td>83.07</td>
<td>67.83</td>
<td>29.70</td>
</tr>
<tr>
<td>Jhenaidah</td>
<td>48.56</td>
<td>47.15</td>
<td>53.46</td>
<td>73.68</td>
<td>62.09</td>
<td>44.38</td>
</tr>
<tr>
<td>Magura</td>
<td>44.66</td>
<td>44.08</td>
<td>49.51</td>
<td>83.66</td>
<td>55.34</td>
<td>31.30</td>
</tr>
<tr>
<td>Meherpur</td>
<td>40.88</td>
<td>40.70</td>
<td>46.69</td>
<td>64.58</td>
<td>58.68</td>
<td>21.08</td>
</tr>
<tr>
<td>Narail</td>
<td>40.37</td>
<td>39.81</td>
<td>56.26</td>
<td>66.13</td>
<td>75.36</td>
<td>32.14</td>
</tr>
<tr>
<td>Satkhira</td>
<td>37.80</td>
<td>37.54</td>
<td>41.79</td>
<td>80.95</td>
<td>54.64</td>
<td>15.57</td>
</tr>
<tr>
<td><strong>Rajshai Division</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jaipurhat</td>
<td>49.62</td>
<td>49.90</td>
<td>47.68</td>
<td>40.00</td>
<td>58.41</td>
<td>35.14</td>
</tr>
<tr>
<td>Rajshahi</td>
<td>47.54</td>
<td>47.67</td>
<td>50.39</td>
<td>39.09</td>
<td>36.71</td>
<td>19.33</td>
</tr>
<tr>
<td>Naogaon</td>
<td>44.39</td>
<td>45.19</td>
<td>44.70</td>
<td>20.34</td>
<td>33.55</td>
<td>14.98</td>
</tr>
<tr>
<td>Bogra</td>
<td>42.89</td>
<td>42.69</td>
<td>45.66</td>
<td>71.54</td>
<td>82.14</td>
<td>21.56</td>
</tr>
<tr>
<td>Kurigram</td>
<td>33.45</td>
<td>32.66</td>
<td>43.52</td>
<td>45.04</td>
<td>64.23</td>
<td>19.96</td>
</tr>
<tr>
<td><strong>Sylhet Division</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moulvibazar</td>
<td>42.06</td>
<td>42.06</td>
<td>42.06</td>
<td>62.50</td>
<td>42.39</td>
<td>41.12</td>
</tr>
<tr>
<td>Habiganj</td>
<td>37.72</td>
<td>35.50</td>
<td>47.16</td>
<td>32.39</td>
<td>44.94</td>
<td>19.73</td>
</tr>
<tr>
<td>Sunamganj</td>
<td>34.37</td>
<td>31.85</td>
<td>47.72</td>
<td>41.51</td>
<td>53.49</td>
<td>16.88</td>
</tr>
</tbody>
</table>

*Source: SVRS 2004, Bangladesh Bureau of Statistics*
performance in the Gazipur district to have relatively better repayment records than other districts. This expectation is confirmed by Nag, who mentioned that the district had the highest number of five-star branches, or branches that achieved 100 percent repayment records (Yunus GBII, p.3).

Therefore, our data will carry a degree of selection bias since it is coming from a relatively successful district for the Grameen Bank. The conclusions reached in this study should be considered with the fact that this data may not be representative of the entire Grameen Bank operation in mind.

Another possible form of selection bias in our dataset is the lack of data on rejected applications for loans, or data on loans that were never disbursed. Because of the procedure in which borrowers apply for loans, there is no data on these loan requests. Grameen borrowers meet weekly at their center meeting with all borrowers from the center. If a member would like to apply for a loan, she must first gain approval from her group, then the center leader, before anything is put in writing. Even when the loan application is filled out, rejected applications are discarded and never recorded onto the computers. This procedure was observed by the author on visits to Centers 47 and 70 of the Shekherchar Narsingdi branch as part of the internship program at the Grameen Bank.

Omissions from the Dataset

The dataset was truncated because of certain concerns regarding selection biases. The first omission was records of all loans with disburse dates prior to January 1st, 1997. Because most observations with disburse dates before 2005 are housing loans (due to their long
duration), it is important to keep observations that are able to have perfect repayment records.

Since our dataset includes all loans fully paid or overdue still on the books as of January 1st, 2007, and housing loans have a typical duration of 520 weeks, or 10 years, it was appropriate to truncate all observations before January 1st, 1997.

All education loans were also eliminated from the dataset used in our analysis. Education loans were left out of our analysis because a high percentage of education loans had 0% interest rates. Moreover, education loans were given as scholarships to borrowers who qualified and thus many of these loans remain unpaid (Barua and Dowla, p.234). Also, because these loans are not part of the scope of this paper, they were left out of our analysis.

**Summary of the Data**

Below in Table 4.3, a summary of our dataset is given. For more detailed information regarding the dataset, please see Appendix [A].

<table>
<thead>
<tr>
<th>Table 4.3: Summary of the Dataset by Loan Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Housing</td>
</tr>
<tr>
<td># of Obs</td>
</tr>
<tr>
<td>869</td>
</tr>
<tr>
<td>Avg Loan Size</td>
</tr>
<tr>
<td>24083</td>
</tr>
<tr>
<td>Avg Duration</td>
</tr>
<tr>
<td>360</td>
</tr>
<tr>
<td>Int Rate</td>
</tr>
<tr>
<td>8.00</td>
</tr>
<tr>
<td>Avg Loan Term</td>
</tr>
<tr>
<td>1.01</td>
</tr>
<tr>
<td>Avg Savings 07</td>
</tr>
<tr>
<td>41441</td>
</tr>
<tr>
<td>Avg Savings 08</td>
</tr>
<tr>
<td>59601</td>
</tr>
<tr>
<td># of Loans On Time</td>
</tr>
<tr>
<td>644</td>
</tr>
</tbody>
</table>
Data Analysis Plan

The characteristics of our data set as well as the distribution of our dependent variables are our initial indicators that a simple linear OLS regression would not be the best option at analyzing our dataset and testing our hypotheses. The distribution of the dependent variable BankLoss has a minimum of 0 taka and a maximum of 396,691. We also see that almost 95% of the values have a value of 0. Because of this strong frequency of observations at 0, OLS will treat the 0’s as actual values of the ability to repay. However, just because a borrower is able to repay his or her loan does not mean that they have the same ability to repay. Because of this issue, we see can see why a linear OLS regression would not be the best choice of econometric models to use for this analysis. The second dependent variable, Repay, is distributed between 0 and 1 and only takes on the values of 0 and 1. We would also possibly violate the assumptions of OLS as we would be using a non-variable outcome variable. Thus, usage of the linear OLS regression does not apply in this case.

When working with the BankLoss dependent variable, one option would be to use either a truncated or censored normal regression model. The underlying assumption of the truncated normal regression model is that error term must be independent of all independent variables as well as normally distributed, a rather strong assumption; more importantly, however, use of truncation would be justified if the values we were to truncate (the 0’s) are inappropriate. This is not the case since the 0 values indicate perfect repayment and therefore the truncated normal regression model would not be the ideal model to use. The censored regression model works under the assumption that we do not know the true values of some of the observations, and
because this assumption does not hold in this case, the censored regression model is not the ideal model to use either.

Therefore, the tobit model is the most appropriate model when examining the BankLoss variable. Used in many applied cases, it deals with the fact that many applications have corner solution responses. In our case, there is a lower bound at 0, a corner solution, while there are also values going up to 396,691. While a linear OLS model might work, it would be possible to obtain negative fitted values which would imply negative predictive values for y. However, most importantly, the large percentage of values at 0 would probably show that the dependent variable does not have a conditional normal distribution, which would be needed to avoid asymptotic justification.

The Tobit model uses the form

\[ y^* = \beta_0 + \mathbf{x} \beta + u, \text{ where } u \mid x \sim \text{Normal}(0, \sigma^2) \]  

(4.2)

\[ y = \max(0, y^*) \]  

(4.3)

We see that \( y^* \), the latent variable, satisfies the assumptions for simple OLS regression (equation 4.2) since it has a homoskedastic and normal distribution.

The Repay dependent variable can be analyzed using a number of models as well. A basic model, the linear probability model has several drawbacks which discourage its use. The first drawback is that the probabilities fitted to the estimators can be greater than 1 or less than 0, which can lead to awkward interpretations of the coefficient estimators. The second drawback is that the partial effect of any independent variable would be constant.
Our binary \textit{Repay} dependent variable can be best analyzed with logit and probit models, where the logistic function and standard normal cumulative distribution function is used, respectively. Both models have the form

\[
P(\text{repay} = 1 | x) = G(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k)
\]

(4.4)

where \( G \) is a function that takes on values from 0 to 1. By making sure that \( G \) can strictly only take values between 0 and 1, we are certain that our estimated probabilities will remain between 0 and 1 as well. In the logit model,

\[
G(z) = e^z / (1 + e^z)
\]

(4.5)

while in the probit model,

\[
G(z) = \int_{-\infty}^{z} 2\pi^{-1/2} e^{-z'^2/2} dz'
\]

(4.6)

The goal of these two models is to explain the effects our independent variables have on the response probability \( P(\text{repay} = 1 | x) \) (Wooldridge, p. 585). Because \( G \) is non-linear, there are partial effects on \( P(\text{repay} = 1 | x) \) and these effects are

\[
\frac{\partial P(\text{repay} = 1 | x)}{\partial x_j} = g(\beta_0 + \mathbf{x}\beta), \text{where } g(z) \equiv \frac{dG}{dz}(z)
\]

(4.7)

Thus, the relative effect that any two independent variables will not be dependent on \( \mathbf{x} \).
Empirical Analysis and Results

This chapter will aim to test our hypotheses presented in the “Model Framework” chapter, assess the significance of our results, and provide context within the broader topics discussed in this paper. Using the models developed in the “Research Methodology” chapter, we will test our two dependent variables, $BankLoss$ and $Repay$, and analyze what effects the different characteristics of loans and borrowers have on these variables.

The Tobit Regression

We begin by regressing the $BankLoss$ dependent variable on our explanatory variables. As discussed in the previous section, OLS, truncated and censored regressions are not the ideal methods to use with our dataset. Therefore, the tobit method is used and the results are shown in Table 5.1 below.
Table 5.1: The Tobit Model with BankLoss as the dependent variable

| BankLoss               | Coefficient (Std Err) | t-stat | P>|t|     | 95% Confidence Interval       |
|------------------------|-----------------------|--------|--------|-----------------------------|
| Housing**              | 2914.858 (1219.639)   | 2.39   | 0.017  | 524.378 - 5305.338          |
| Basic***               | -10873.710 (288.763)  | -37.66 | 0.000  | -21747.420 - -21747.420     |
| Microenterprise***     | -6621.788 (335.228)   | -19.75 | 0.000  | -13243.570 - 13243.577      |
| Modified***            | 14275.380 (588.066)   | 24.28  | 0.000  | 13122.770 - 15427.980       |
| TotalPersonalSavings***| 0.005 (0.001)         | 3.00   | 0.003  | 0.002 - 0.008               |
| TotalLoanInsurance***  | -0.515 (0.029)        | -17.26 | 0.000  | -1.031 - 1.031              |
| PrincipalLoan***       | 0.079 (0.003)         | 22.00  | 0.000  | 0.072 - 0.086               |
| Duration               | -0.001 (0.006)        | -0.15  | 0.878  | -0.014 - 0.012              |
| LoanTerm***            | 432.215 (46.768)      | 9.24   | 0.000  | 340.550 - 523.880           |
| HousingPurrpose        | -1234.247 (1225.871)  | -1.01  | 0.314  | -3636.942 - 1168.447        |
| Repair                 | 647.234 (1256.429)    | 0.52   | 0.606  | -1815.355 - 3109.822        |
| Construction           | 256.823 (1063.534)    | 0.24   | 0.809  | -1827.692 - 2341.338        |
| Constant               | -12838.510 (284.777)  | -45.08 | 0.000  | -25677.020 - -25677.020     |
| /sigma                 | 12320.250 (126.726)   |        |        | 12071.860 - 12568.630       |

Number of observations 92954

LR chi2(13) 4161.350 *** Statistically Significant at the 1% Level
Prob > chi2 0.000 ** Statistically Significant at the 5% Level
Pseudo R2 0.027 * Statistically Significant at the 10% Level
Log likelihood -75025.196

44
With 92,954 observations, the tobit regression predicting the \textit{BankLoss} variable from housing loan, basic loan, microenterprise loan, modified loan, the borrower’s personal savings, loan insurance, principal loan amount, loan’s duration, loan’s term, and housing purposes, as a whole, was statistically significant—the p-value for chi-squared with 13 degrees of freedom is virtually 0. A lower bound was set at 0 for \textit{BankLoss}, as there are 87,146 left-censored observations at $\text{BankLoss} \leq 0$, with 5,808 observations uncensored. The reported “goodness-of-fit” ratio, the Pseudo $R^2$, is 0.027. The pseudo $R^2$ in the output is McFadden’s pseudo $R^2$ and is calculated by $R^2 = 1 - LL(\text{full model}) / LL(\text{constant only model})$. The pseudo $R^2$ implies that these predictors accounted for almost 3% of the overall variance in the dependent variable \textit{BankLoss}.

Of the independent variables, only 7 of them were statistically significant at the 1% level and 1 of them, \textit{Housing}, was significant at the 5% level. Thus, a unit increase in a borrower’s total personal savings, total loan insurance, the number of previous loans he or she has taken out, and the loan’s principal amount, lead to a .004 taka increase, .515 taka decrease, 432.215 taka increase, and a .079 taka increase in \textit{BankLoss}, respectively. Stated simply, although the borrower’s personal savings, total loan insurance, and the loan’s principal amount are statistically significant, the effects these variables have on the amount the borrower would owe to the bank are quite small. However, for each additional loan a borrower has previously taken out, the model predicts that the borrower would owe an additional 432 taka to the bank.

The model also predicts that if the loan type is housing, basic, microenterprise, or modified, the amount of money owed to the bank increases by 2,915 taka, decreases by 10,874 taka, decreases by 6,622 taka, and increases by 14,275 taka, respectively.
Using this model, we will consider the hypotheses we proposed in Chapter 3 of this paper and see if our model can or cannot reject them. Consider Hypothesis 1, in which the more number of previous loans a borrower has taken out, the more likely the borrower will be in repaying his current loan. For this model, we can rewrite Hypothesis 1 as:

\[
H_0 : \beta_{\text{LoanTerm}} = 0 \\
H_1 : \beta_{\text{LoanTerm}} < 0
\]  

(5.1)

For our null hypothesis, we set the coefficient equal to 0 to see if there is any statistically significant effect \( \text{LoanTerm} \) would have on our dependent variable. Because \( \text{LoanTerm} \) is statistically significant, we can reject our null hypothesis. Our two-sided alternative hypothesis is rejected as well, since the amount a borrower would likely owe to the microfinance institution actually goes up with the number of previous loans taken.

Considering Hypothesis 2, that is, the greater the personal wealth of a borrower, the more likely the borrower will be in repaying his current loan. While we have no exact measure for a borrower’s personal wealth, as we do not have records from other possible savings borrowers may have, the personal savings account that borrowers keep with the Grameen Bank should be a good proxy for a borrower’s wealth. Thus, Hypothesis 2 can be rewritten as

\[
H_0 : \beta_{\text{TotalPersonalSavings}} = 0 \\
H_1 : \beta_{\text{TotalPersonalSavings}} < 0
\]  

(5.2)

Because our estimator for \( \text{TotalPersonalSavings} \) is statistically significant, we can safely reject our null hypothesis using this test. Our alternative hypothesis, that is that the estimator should be negative, can be rejected as well, though the estimator probably indicates, with significance, that the level of personal savings really has no real effect on how much borrowers
owe the bank—a taka increase in a personal savings account only predicts a .004 taka increase in how much the borrower owes.

Hypothesis 3, which claims that income-generating loans should have higher rates of repayment than that of consumption loans due to the nature of the loan, can be written as the following test:

\[
H_0 : \beta_{\text{Housing}} = 0 \\
H_1 : \beta_{\text{Housing}} > \beta_{\text{Basic}} \\
H_2 : \beta_{\text{Housing}} > \beta_{\text{Microenterprise}}
\]  
(5.3)

Since the \textit{Housing} coefficient is statistically significant, we can reject the null hypothesis. However, since the coefficients for basic and microenterprise loans are negative while the housing coefficient is positive, we cannot reject both alternative hypotheses. Thus, the model implies that consumption loans are more likely to create losses for microfinance institutions than income-generating loans.

When we apply robust standard errors to our tobit model, the results to Hypothesis 3 change drastically. Since our sample size is rather large (n=92,954), it makes sense to use robust standard errors to test our hypotheses since our regression is more than likely to exhibit heteroskedasticity (in order to exhibit perfect homoskedasticity, the error term, conditional on the independent variables, must be constant). The results of the tobit regression with robust standard errors are presented below in Table 5.2.
Table 5.2: The Tobit Model with Robust Standard Errors

| BankLoss          | Coefficient (Robust Std Err) | Wald Stat | P>|z| | 95% Confidence Interval |
|-------------------|------------------------------|-----------|-----|--------------------------|
| Housing***        | -4445.170 (861.585)          | -5.16     | 0.00 | -6133.846 -2756.494     |
| Basic***          | 3502.973 (409.223)           | 8.56      | 0.00 | 2700.910 4305.037      |
| Microenterprise*  | -1797.553 (1091.296)         | -1.65     | 0.10 | -3936.454 341.349      |
| Modified***       | -2905.566 (838.124)          | -3.47     | 0.00 | -4548.260 -1262.873    |
| TotalPersonalSavings | 0.002 (0.005)           | 0.40      | 0.69 | -0.009 0.014           |
| TotalLoanInsurance| -0.034 (0.089)              | -0.38     | 0.71 | -0.210 0.142           |
| PrincipalLoan***  | 0.141 (0.040)               | 3.49      | 0.00 | 0.062 0.220           |
| Duration***       | 0.000 (0.000)               | 4.26      | 0.00 | 0.000 0.001           |
| LoanTerm***       | -337.104 (73.679)           | -4.58     | 0.00 | -481.513 -192.695     |
| HousingPurpuse    | -229.153 (715.204)          | -0.32     | 0.75 | -1630.929 1172.623    |
| Repair            | -163.947 (724.943)          | -0.23     | 0.82 | -1584.811 1256.916    |
| Purchase***       | 22072.190 (2802.934)        | 7.87      | 0.00 | 16578.540 27565.840   |
| Construction      | -790.091 (497.613)          | -1.59     | 0.11 | -1765.395 185.214    |
| Constant          | 8528.764 (753.558)          | 11.32     | 0.00 | 7051.817 10005.710   |
| /Insigma          | 8.822 (0.107)               | 82.38     | 0.00 | 8.612 9.032           |
| sigma             | 6781.274 (726.187)          |           |     | 5497.415 8364.963    |

Number of observations: 92954  
Wald chi2(13): 437.880  
Prob > chi2: 0.000  
Log pseudo-likelihood: -68614.074

*** Statistically Significant at the 1% Level  
** Statistically Significant at the 5% Level  
* Statistically Significant at the 10% Level
Using robust standard errors, a couple results change from our original tobit regression. Hypothesis 1 retains the same conclusion as the original regression, but Hypothesis 2 and Hypothesis 3 lead to different results. With regard Hypothesis 2, we notice that we can no longer reject our null hypothesis—TotalPersonalSavings is no longer statistically significant. This result echoes our analysis of the second hypothesis in the original tobit regression, since the magnitude of the coefficient had a small impact despite its significance. With regard to the Hypothesis 3, we must reject $H_1: \beta_{Housing} > \beta_{Basic}$ and $H_2: \beta_{Housing} > \beta_{Microenterprise}$, as the coefficients have changed signs and magnitudes.

While the logit model gave us some initial conclusions regarding our hypotheses, our “goodness-of-fit” calculation of .027 is quite weak. Another reason the logit model might be flawed is the extreme amount of skew to 0 in our dependent variable—over 95% of our dependent values are 0, while only the remaining 5% actually have variance. Therefore, by looking at our other dependent variable and its binary features, we may be able to draw some stronger conclusions regarding our hypotheses.

The Logit and Probit Models

Since the logit and probit models are quite similar to one another, and because our results do not vary between the two models, we will present results from both the logit (Table 5.3) and probit (Table 5.4) models but refer to the logit model when analyzing the results, as the coefficients are much more intuitive with this model.
### Table 5.3: The Logit Model with *Repay* as the dependent variable

|                                | Coefficient | Std Err | Wald Stat | P>|z| | 95% Confidence Interval |
|--------------------------------|-------------|---------|-----------|-----|-------------------------|
| Housing***                     | -0.767      | 0.196   | -3.89     | 0.000 | -1.153 -0.381           |
| Basic***                       | 1.822       | 0.042   | 42.58     | 0.000 | 1.739 1.906             |
| Microenterprise***             | 0.809       | 0.053   | 15.11     | 0.000 | 0.704 0.914             |
| Modified***                    | -2.651      | 0.103   | -25.57    | 0.000 | -2.854 -2.447           |
| TotalPersonalSavings**         | 0.000       | 0.000   | -2.37     | 0.018 | 0.000 0.000             |
| TotalLoanInsurance***          | 0.000       | 0.000   | 13.56     | 0.000 | 0.000 0.000             |
| PrincipalLoan***              | 0.000       | 0.000   | -3.04     | 0.002 | 0.000 0.000             |
| Duration                       | 0.000       | 0.000   | 0.13      | 0.893 | 0.000 0.000             |
| LoanTerm***                   | -0.086      | 0.007   | -10.87    | 0.000 | -0.101 -0.070           |
| HousingPurpose                 | 0.180       | 0.200   | 0.90      | 0.369 | -0.213 0.574            |
| Repair                        | -0.268      | 0.205   | -1.30     | 0.193 | -0.671 0.135            |
| Construction                  | -0.106      | 0.155   | -0.68     | 0.497 | -0.411 0.200            |
| Constant                      | 1.521       | 0.037   | 40.56     | 0.000 | 1.447 1.594             |

Number of observations 92946

LR chi2(12) 4209.730 *** Statistically Significant at the 1% Level
Prob > chi2 0.000 ** Statistically Significant at the 5% Level
Pseudo R2 0.097 * Statistically Significant at the 10% Level

Log likelihood -19622.089
Table 5.4: The Probit Model with *Repay* as the dependent variable

| Repay                  | Coefficient (Std Err) | Wald Stat | P>|z| | 95% Confidence Interval |
|------------------------|-----------------------|-----------|------|----------------------------|
| Housing***             | -0.431 (0.098)        | -4.36     | 0.000 | -0.862 - 0.862             |
| Basic***               | 0.895 (0.022)         | 40.12     | 0.000 | 0.852 - 0.939              |
| Microenterprise***     | 0.435 (0.027)         | 15.57     | 0.000 | 0.381 - 0.490              |
| Modified***            | -1.632 (0.060)        | -27.18    | 0.000 | -3.265 - -3.265            |
| TotalPersonalSavings*  | 0.000 (0.000)         | -1.79     | 0.073 | 0.000 - 0.000              |
| TotalLoanInsurance***  | 0.000 (0.000)         | 13.31     | 0.000 | 0.000 - 0.000              |
| PrincipalLoan***       | 0.000 (0.000)         | -3.67     | 0.000 | 0.000 - 2557.000           |
| Duration               | 0.000 (0.000)         | 0.15      | 0.878 | 0.000 - 0.000              |
| LoanTerm***            | -0.040 (0.003)        | -10.55    | 0.000 | -0.079 - 0.079             |
| HousingPurrpose        | 0.091 (0.099)         | 0.92      | 0.359 | -0.104 - 0.286             |
| Repair                 | -0.093 (0.101)        | -0.92     | 0.359 | -0.293 - 0.106             |
| Construction           | -0.066 (0.087)        | -0.76     | 0.448 | -0.237 - 0.105             |
| Constant               | 0.931 (0.020)         | 45.72     | 0.000 | 0.891 - 0.971              |

Number of observations 92946
LR chi2(12) 4187.090 *** Statistically Significant at the 1% Level
Prob > chi2 0.000 ** Statistically Significant at the 5% Level
Pseudo R2 0.096 * Statistically Significant at the 10% Level
Log likelihood -19633.407
Looking at the regression results in Table 5.3 of the logit model, 92,946 observations were analyzed using the logit model, and as a whole, the model was statistically significant with a p-value for chi-squared with 13 degrees of freedom of 0. The independent variables in the model were all significant on the 5% level other than the duration of the loan, and the housing purposes of the loan. The “goodness-of-fit” measure in this model is significantly higher than our tobit model, with a pseudo $R^2$ of .097, implying that almost 10% of the variance in the dependent variable is explained by the logit model.

As the interpretation of the coefficients can be awkward using the logit regression, a more intuitive way of interpreting the coefficients is by exponentiating the coefficients, thus interpreting the results as odds-ratios. The output of this conversion is provided below in Table 5.5.
Considering Hypothesis 1, we rewrite it in terms of the logit model as we did in equation 5.1, but with several changes to match the dependent variable in this model:

Table 5.5: The Logit Model with the Odds-Ratio

| Repay                     | Odds Ratio (Std Err) | Wald Stat | P>|z|   | 95% Confidence Interval |
|---------------------------|----------------------|-----------|------|-------------------------|
| Housing**                 | 0.465 (0.091)        | -3.89     | 0.000| 0.316 - 0.683           |
| Basic***                  | 6.187 (0.264)        | 42.58     | 0.000| 5.689 - 6.728           |
| Microenterprise***        | 2.245 (0.120)        | 15.11     | 0.000| 2.022 - 2.493           |
| Modified***               | 0.071 (0.007)        | -25.57    | 0.000| 0.058 - 0.087           |
| TotalPersonalSavings**    | 1.000 (0.000)        | -2.37     | 0.018| 1.000 - 1.000           |
| TotalLoanInsurance***     | 1.000 (0.000)        | 13.56     | 0.000| 1.000 - 1.000           |
| PrincipalLoan***          | 1.000 (0.000)        | -3.04     | 0.021| 1.000 - 1.000           |
| Duration                  | 1.000 (0.000)        | 0.13      | 0.893| 1.000 - 1.000           |
| LoanTerm***               | 0.918 (0.007)        | -10.87    | 0.000| 0.904 - 0.932           |
| HousingPurrpose           | 1.198 (0.240)        | 0.90      | 0.369| 0.808 - 1.775           |
| Repair                    | 0.765 (0.157)        | -1.30     | 0.193| 0.511 - 1.145           |
| Construction              | 0.900 (0.140)        | -0.68     | 0.497| 0.663 - 1.221           |

Number of observations: 92946
LR chi2(12): 4209.730 *** Statistically Significant at the 1% Level
Prob > chi2: 0.000 ** Statistically Significant at the 5% Level
Pseudo R2: 0.097 * Statistically Significant at the 10% Level
Log likelihood: -19622.089

Considering Hypothesis 1, we rewrite it in terms of the logit model as we did in equation 5.1, but with several changes to match the dependent variable in this model:
We will test these hypotheses with the non-odds-ratio regression, while using the odds-ratio when interpreting the coefficients. The Wald-statistic test, similar to a t-test, shows that the \( \text{LoanTerm} \) coefficient is statistically significant and we can safely reject our null hypothesis. However, we also must reject our alternative hypothesis. The odds-ratio can be interpreted in the following fashion—for every additional previous loan a borrower has taken, the odds of the borrower repaying his current loan (vs. not repaying) increases by only .918. If it were to not reject our alternative hypothesis, an odds increase of at least 1 would be necessary.

Hypothesis 2, which states that the greater the personal wealth of a borrower, the more likely the borrower will be in repaying his current loan, can also be written to fit our model:

\[
\begin{align*}
H_0 & : \beta_{\text{TotalPersonalSavings}} = 0 \\
H_1 & : \beta_{\text{TotalPersonalSavings}} > 0
\end{align*}
\]

Here, again, we can safely reject our null hypothesis as well as our alternative hypothesis. Looking at the odds-ratio, a 1 taka increase in personal savings only translate to the predicted odds of being able to repay by .999. Thus, we must reject Hypothesis 2 using this test.

When considering Hypothesis 3, that is, income-generating loans should have higher rates of repayment than those of consumption loans due to the nature of the loan, the hypothesis can be again rewritten as follows:

\[
\begin{align*}
H_0 & : \beta_{\text{Housing}} = 0 \\
H_1 & : \beta_{\text{Housing}} < \beta_{\text{Basic}} \\
H_2 & : \beta_{\text{Housing}} < \beta_{\text{Microenterprise}}
\end{align*}
\]

Clearly, from the original logit model, the Wald-statistic test rejects the null hypothesis, and thus our coefficient is statistically significant. The two other alternative hypotheses are not
rejected from our test, as the housing coefficient is less than both the basic and microenterprise coefficients. The odds-ratio results imply that while housing loans (vs. “other” loans) increase the odds of repayment by only .465, basic loans (vs. “other” loans) increase the odds of repayment by 6.187 and microenterprise loans increase the odds of repayment by 2.245.

According to our regression, those who take out basic and microenterprise loans are more likely to repay than those who take out housing loans, which is consistent with Hypothesis 3.

### Table 5.6: Summary of Hypothesis Testing

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Tobit Model</th>
<th>Tobit Model with Robust SE</th>
<th>Logit/Probit Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothesis 1: Progressive Lending*</td>
<td>Reject</td>
<td>Reject</td>
<td>Reject</td>
</tr>
<tr>
<td>Hypothesis 2: Personal Wealth**</td>
<td>Cannot Reject</td>
<td>Reject</td>
<td>Reject</td>
</tr>
<tr>
<td>Hypothesis 3: Income-generating vs. Consumption***</td>
<td>Cannot Reject</td>
<td>Reject</td>
<td>Cannot Reject</td>
</tr>
</tbody>
</table>

* The more number of previous loans a borrower has taken out, the more likely the borrower will be in repaying his current loan

** The greater the personal wealth of a borrower, the more likely the borrower will be in repaying his current loan

*** Income-generating loans should have higher rates of repayment than those of consumption loans

### The Big Picture

The results to our regressions, as summarized above in Table 5.6, answer some important questions in the microfinance debate. Microfinance institutions consistently struggle between two objectives—providing cheap capital to needy borrowers while maintaining a model that will be sustainable for years to come. In pursuit of providing capital to a socio-
economic group that is in dire need of such funds, microfinance institutions must maintain a trustworthy administration as well as provide innovative loan products so that the right borrowers have access to the capital at a suitable price. Mechanisms such as joint lending and cross-subsidization are used in addition to other methods in order to provide loan products that will be used.

However, in order to be a sustainable institution, microfinance institutions must also make sure that their portfolios are healthy and that its procedures are well-regarded and carry minimal risk. The most important factor of maintaining a healthy loan portfolio is making sure that disbursed loans will be paid back on time with interest, along with other mechanisms such as product diversification and strong deposit systems. In order to manage an institution with minimal risk, checks and balances must be put in place along with proper enforcement of loan fund usage.

The three hypotheses we tested in this paper contribute to these issues by looking at how certain mechanisms or characteristics of loans and borrowers would affect repayment. Hypothesis 1 looks at the theory behind progressive lending, examining the effect that previous good repayment history would lead to better repayment rates. All of our models presented in this paper rejected this hypothesis, which was an interesting result considering progressive lending is one of the most accepted and well-used mechanism to overcome the adverse selection problem. The models presented in this paper have several flaws when looking at this problem, most notable being the lack of variance in the dependent variable. While a high percentage of loans with a loan term greater than 1 had good repayment, other loans that were
first-time loans also had good repayment. The lack of variance in the repayment variable could affect the results of this model.

Hypothesis 2 examines the effect a borrower’s personal wealth would have on his or her repayment ability. While this hypothesis was not rejected by the tobit model, the other models used (the tobit model with robust standard errors, the logit model and the probit model) rejected the hypothesis. The borrower’s personal wealth is very important when looking at the accessibility of loans to borrowers—most microfinance institutions regard personal wealth as a sign that borrowers can repay their loans, since they can draw from assets to help support difficult payments. While the personal savings data we use in our study does not match up to the total personal wealth of a borrower (the borrower could have cash in hand, accounts with other institutions, or material assets), the personal savings account with Grameen Bank is still a good indicator of a borrower’s wealth. While there is no underlying data to support the claim, it is not too strong of an assumption that the savings accounts that borrowers keep with the Grameen Bank are proportionate with their total wealth. For example, if a borrower has a total wealth of $100, she may keep $50, or 50% of that in the savings account. Similarly, a borrower with a total wealth of $200 would probably keep a similar proportion ($100) in the savings account. Because the interest rates that the Grameen Bank provides for savings accounts are quite competitive (at the time of writing, 8%), this claim would be a safe assumption. The rejection of Hypothesis 2 has strong implications regarding the accessibility of loans by asset-less borrowers—if personal wealth was a minimal condition on the borrower’s ability to repay, then such wealth characteristics of the borrower can be looked over and other considerations can determine the granting of loans. The Grameen Bank has been operating under the
assumption that personal wealth does not have a huge effect on repayments, and our model shows that this may be the case.

Hypothesis 3 tests the differences behind income-generating loans and consumption loans, looking at how repayment differs between the two. The intuition behind the idea that income-generating loans should perform better than consumption loans is that there is an underlying cash flow with income-generating loans that can help support repayments, while consumption loans are usually paid off with other sources of income. Our logit and probit models clearly support this hypothesis, as does our tobit model with regular standard errors. However, the tobit model with robust standard errors, perhaps a more suitable test for such datasets with a large sample size, rejected the hypothesis. Despite the results to the tobit model with robust standard errors, in practice, microfinance institutions have had issues adding consumption loans to their loan product line. Building in consumption loans into the loan portfolio has been a common headache for microfinance institutions because of this theoretical problem.
Further Discussion

This chapter aims at discussing issues that arise from the process and results of our study, as well as other topics that are related but not within the scope of the paper. Among the most striking issues surrounding the Grameen Bank housing program is the recent lack of disbursed housing loans. In the dataset collected from this study, of all the loans disbursed after January 1st, 2006, only 1 housing loan was approved in the Gazipur district after 2006. This issue is not just unique to the Gazipur district. A similar situation was observed in the adjacent Narsingdi district. In interviews with Grameen borrowers at Center 40, 53, and 47 of the Shekherchar Narsingdi branch, when asked what more the Grameen Bank could do for them, at least 5 borrowers from each center suggested that the Grameen Bank should approve more housing loans. This experience shows that there is strong demand for housing loans from borrowers, yet the Grameen Bank has recently been hesitant to approve these loans.

Our data shows that although only 1 housing loan was approved after 2006, basic and microenterprise loans were being approved for housing purposes during that period. In fact, of 67,124 loans approved after January 2nd, 2006, just over 9,000 of them were for housing purposes—8,638 for house repair, 8 for house purchases and 135 for house construction. While loans were being approved for housing purposes, and thus consumption purposes, these loans were being disbursed at a higher interest rate of 20% and a shorter duration of 54 weeks, on average, as opposed to the 8% interest rate and 520 week duration borrowers would have received on a housing loan.

While not a specific focus of this paper, the reluctance to issue housing loans during the relevant period is a critical issue of interest as several variables are at play. Despite being basic
loans, these loans were disbursed for housing purposes and thus a consumption loan. However, instead of being charged a lower interest rate, the Grameen Bank receives the higher rate for that loan. This change in loan disbursements elicits concerns that because the cross-subsidized housing loan is no longer feasible, banks must now charge higher interest rates to make up for the implied increase of default.

Another interesting issue we sidestepped in our paper is the idea of the saturation of both the borrowers and the microcredit market, which would have affected our analysis of progressive lending. The saturation of borrowers refers to the concept that not all borrowers need constant access to credit and thus may not take out 2nd or 3rd time loans. The saturation of the market is much more of a concern, now that many economists believe that many of the countries in which microfinance has been successful in have passed their expansionary phase and have now reached a point of saturation (Charitonenko and Rahman, p.12). Many issues arise when the markets become oversaturated with microfinance funds, as borrowers are then able to select from multiple microfinance institutions, and competitors. This ability to choose and have multiple sources of credit disables the microfinance institution’s threat to stop lending if previous loans are left unpaid. Moreover, borrowers may use credit from competitor institutions to pay off old or overdue loans.
Conclusion

Poverty has certainly not been eliminated by the movement in microfinance, nor has microfinance been identified as the cure for poverty. However, it has brought capitalistic ideas to the poor, allowing them access to financial inventions that have brought prosperity to the millions of people on the other end of the spectrum. Microfinance institutions around the world have seen that despite the wealth of their borrowers, the poor will pay back their debts (provided the correct mechanisms are in place) to the best of their abilities. Group lending, joint liability, progressive lending, and other mechanisms have made the access to credit feasible and possible. However, as markets become saturated with such funding, by reexamining the mechanisms that had sparked the growth of microfinance, some find that these mechanisms put unneeded hardships on certain borrowers. Thus, abandoning joint liability on as large a scale as the Grameen Bank has certainly brought new problems and hardships on the bank. The move away from joint liability put new strains on the theories of progressive lending as well as other key mechanisms that kept the economic model of microfinance alive. Though our analysis found that progressive lending was not particularly effective at promoting repayment, it is certainly a strong theoretical element that would require further analysis. The move away from joint lending has also brought on hardships on the distribution of consumption loans: both in theory and in the data, consumption loans fare worse than the standard income-generation loans. A deeper look at how the consumption loan products are designed is required if microfinance institutions would like to maintain these product lines and remain sustainable.

Because of these issues, the Grameen Bank has chosen to stop lending in certain product lines in the Gazipur district, and this effect can be observed in other regions of Bangladesh as well.
Professor Yunus found in his early trial studies that the poor do pay back, and thirty-or-so years later, his statement remains true and is reality. However, if this reality is to stay on course, the same type of innovation that built Wall Street needs to be applied to lending on the dusty paths as well.
Bibliography


