A Study of Default in Microfinance

Puru Parthasarathy

Abstract

The joint liability model for microfinance has been considered to be a success in terms of achieving higher repayment rates. However, when there are shocks to the microfinance setting, the joint liability model can cause higher rates of default[9]. We examine a case of peer-influenced default in South India, where a committee ruling causes mass defaults among Muslim borrowers in a village. Since the Hindu borrowers are not directly obligated to default, they exhibit strategic default due to peer influences. We develop a model for peer effects using a standard discrete choice modeling framework aided by a support vector scheme for feature selection. We find that although the joint liability model for microfinance reduces idiosyncratic risk, it increases systemic risk significantly.

Mar 2012, MFE Program at UC Berkeley
Applied Finance Project
Advisor: Prof. Dwight M. Jaffee

Contents

1 Introduction

The idea of microfinance was started by a group of young men in Bangladesh in March 1978[2]. These men got together with the idea of fighting rural poverty, and thirty years later microfinance has grown to reach six million villagers in Bangladesh. This organization which now calls itself the Association for Social Advancement (ASA) reaches out to villagers in the poorest parts of the country, and meets their needs by providing small loans of up to $75 without collateral. Other microfinance organizations have emerged within Bangladesh, some names include Grameen Bank and BRAC (Bangladesh Rural Advancement Committee). By the end of 2007, the above three lenders had over twenty million customers. These were customers that other banks deemed 'unbankable'.

If we take a step back and look at the economics behind microcredit (and we distinguish this from microfinance shortly below), we find that economic theory tells us that the marginal returns from a production function that is concave must be higher when the capital invested in the operation is lower[2]. That would mean that the marginal returns would be higher for investments in microcredit. In that case, it should be the invisible auctioneer that allocates capital appropriately - from the affluent centers of London and New York to a low income suburb in Los Angeles, from the sky-scrapers of Mumbai to a remote Indian village. The principle of diminishing marginal returns says that a women selling fruits at an intersection should be able to offer higher returns that an investment in Google.com. However, the paradox resolves itself when we analyze risk. A borrower at the micro-credit level is deemed to risky by a bank, and the cost of administering many small loans is extremely high compared to one large loan to an affluent customer. Some argue that state usury laws prevent banks from offsetting these costs by charging a higher interest rate, but it is yet to be seen if usury laws can be modified (they have too much history)[2]. It is in these circumstances that the idea of microcredit gains attention as being able to overcome the transactional and informational barriers that larger banks have.

Another famous organization that adopted the technique of microcredit was Grameen Bank, which we briefly mentioned earlier. Mohammed Yunus, a Vanderbilt trained economist started this organization and found that not only were the locals able to benefit from these loans, they would also repay loans with almost no default. This led to the idea of group lending, where a smaller group of people were granted loans. In case one member defaulted, the others would help payback her loans and in the event that whole group could not cover the default, the group would not have access to credit in the future. This idea of group lending or joint liability has been a crucial part of microcredit, and although it is usually an efficient mechanism to allocate loans, it may have adverse consequences when a coordinated default event happens.

At this stage we distinguish between microcredit and microfinance. The term microcredit describes making small loans to a borrower without collateral, typically in a group lending model. The term microfinance, however, is a more general term that refers to a broad array of services - micro-insurance, microloans (with interest) and in some cases helping to consolidate savings from low income households. This has led to microfinance as a new revolution in financial services where lenders are commercial for-profit entities which cater to the ‘less poor’. This changing landscape of microfinance presents an array of promises. It also leads to puzzles and debates. It is one such puzzle that we try to explain in this paper.

2 Networks in Microfinance

There are a considerable number of studies that document networks in microfinance and model diffusion effects that lead to the adoption of microfinance[3]. Some studies that evaluate the default characteristics of a microfinance network - Field and Pande[5] conclude that a more flexible schedule in terms of payments can lead to lower defaults in microfinance networks. Further the frequency of repayment in this particular field study has no effect on the default characteristics of the network of households that borrow from the MFI(Microfinance Institution). Another relevant study is by Breza [4] where peer effects are estimated after temporary defaults - where repayments are influenced by whether an individuals peer group has started repaying.

In the paper, we explore peer effects of default in a microfinance setting. Besley and Coates[9] propose that in the joint liability model for lending, any default that would occur would be amplified. However, they point out that since any shock to a community is endogenous to it, (i.e. the behavior of a community is altered by the shock) it is often hard to identify the peer effects of the network. In light of events in the southern Indian state of Karnataka in January 2009, we are lucky to find a data set from a microfinance setting in which there is an exogenous shock to the Muslim community in the form of a *fatwa* announcement against repaying microfinance loans by a local committee. The Hindu communities of these towns are not affected by the committee’s *fatwa* and yet they default - in this we find a unique opportunity to see the effects of peer relationships in a microfinance set up.

The data set that we use comes from a survey that was done across four towns in the affected region. The survey is contained in the Appendix. Figure 1 shows the location of the towns which were surveyed in Karnataka.

3 Data

3.1 Description of the Data Set

An interesting series of events led to defaults of more than 50% of the households in the towns of Kolar and Ramanagaram in July 2009. These towns are located in the state of Karnataka in southern India and the households in these towns have been involved in borrowing from microfinance institutions (henceforth referred to as MFI) for the past several years. Even though Kolar was known for its gold industry, these mines were shut down in 2003 as a result of reduced gold deposits, and now the main industries in the region are agriculture, dairy and sericulture.

An interesting aspect to observe here is that multiple (in fact 11) microfinance organizations were operating in the year 2009 all within 150 miles of each other. The mass defaults had one major cause - a local organization in one of these towns called the Anjuman Committee came to a consensus that the interest rates charged by the microfinance lenders were *haraam*[1] and hence not appropriate. The Anjuman committee then issued a *fatwa* against the microfinance lenders and forced all Muslims in Kolar to stop repaying their loans. This led to a widespread peer induced default among the Muslim households in Kolar and after a point these defaults spread to the nearby town of Ramanagaram.

An organization based out of Washington D.C. (CGAP) collected data from a total of 935 households in these two towns and two more towns, namely Davenagere and Nanjangud, which were relatively default-free in the form of a detailed survey that asked households a wide variety of questions including their income and its seasonality, their religion, their borrowings from various institutions, and even whether the household had any large farm animals. In Table 1, we show a brief snapshot of the data in the four towns surveyed.

Table 1: Default Data From Kolar, Ramanagaram, Davanagere and Nanjangud

	Number of Households Surveyed	Number of Muslim Households
Davanagere	148	86
Nanjangud	129	80
Kolar	356	258
Ramanagaram	306	153

Table 2: Default Data From Kolar, Ramanagaram, Davanagere and Nanjangud

	Households Surveyed	Defaults	Muslim Defaults	Hindu Defaults
Davanagere	148	1	1	0
Nanjangud	129	22	17	5
Kolar	356	201	190	11
Ramanagaram	306	136	106	30

This data set was modified in this study to enumerate some answers and was also scaled in order to run the SVM analysis. For more details on the questions in the survey, please refer to the Appendix.

3.2 Graphical Representations

In order to understand the default behavior across over 900 households in 4 towns, we would like to find ways to visualize these defaults across the towns. Since we are aware

[1]Haraam is an Arabic term meaning forbidden. In Islam , this refers to a sin of the highest order.

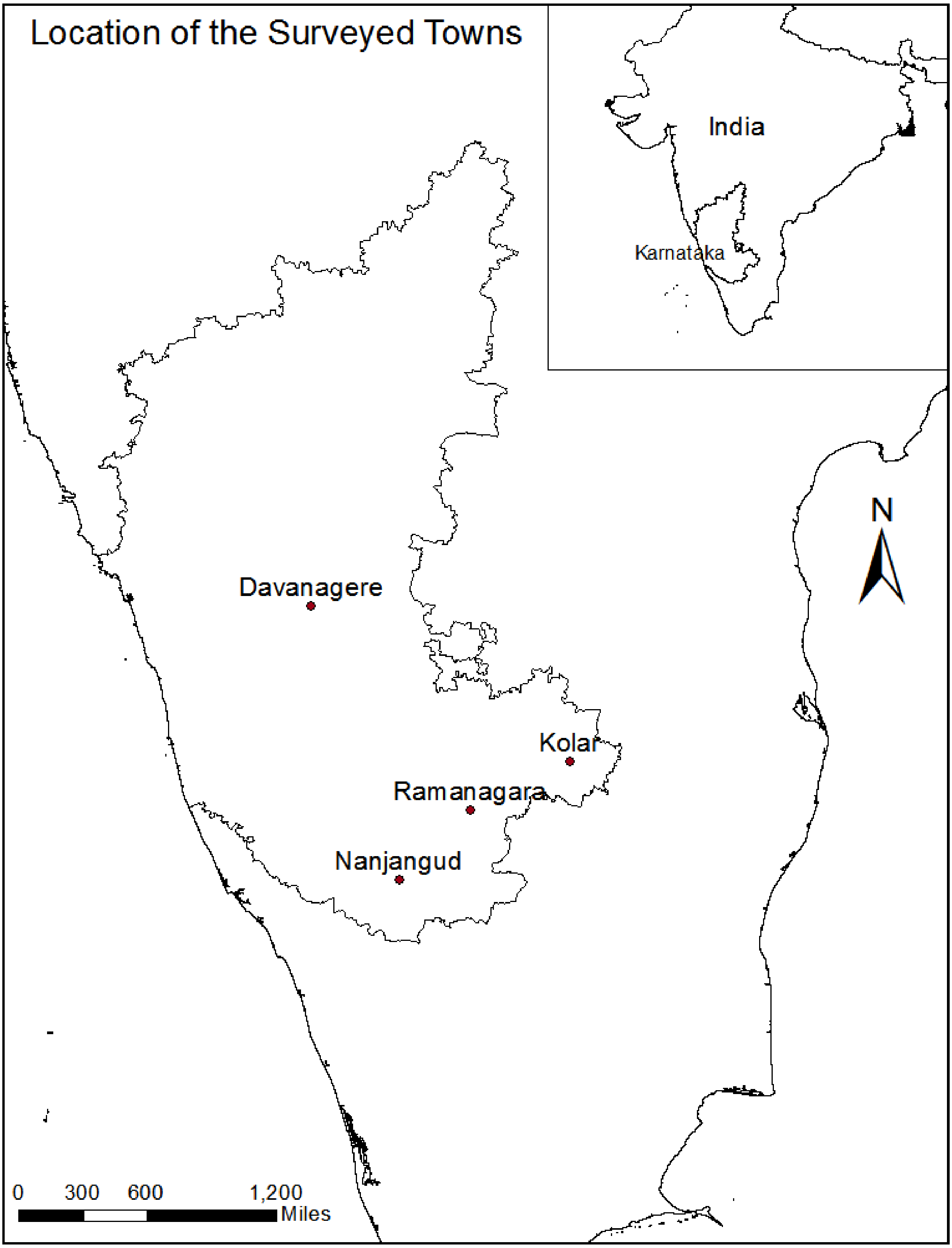

Figure 1: The towns of Kolar, Davenagere, Ramanagaram and Nanjangud

at this point that religion is an important facet of this data, we would like to partition the data according to religion and then explore graphical visualizations of default within each community. One immediate concern is that in case of the Muslim community, the defaults are very large in number. Hence we need to explore some methods of sparse representation of the same. We explore these in this section to understand the dynamics of default better.

Given a dense correlated network, how can we estimate the structure of the network in a sparse fashion? We turn to ideas of conditional independence in networks for representing our data. We make some assumptions on our data to see the structure of the correlated network, and use some established theorems on multivariate distributions which are proven in the Appendix.

Proposition: Consider a multivariate gaussian distribution $X \sim N(0, \Sigma)$ where Σ is $p \times p$ covariance matrix. Then Variables x_i and x_j are conditionally independent iff $\Sigma^{-1}_{ij} = \Sigma^{-1}_{ji} = 0$ where Σ^{-1} is the precision matrix.

Proof: Please see Appendix A.

In order to extract conditional independence information from this network of households, we assume that the response vectors and characteristics across the towns are normally distributed around the mean. We construct a sample covariance matrix:

$$S = \frac{1}{N-1}\left[X - \widehat{X}\right]\left[X - \widehat{X}\right]^T \tag{1}$$

where X is the matrix of household characteristics and $\widehat{X}$ is the mean. Hence S is the sample covariance for this network, in fact it is the *sample peer covariance*, in line with our proposed idea of capturing peer networks. Further, we compute the precision matrix and perform an eigen value decomposition to estimate a low rank version of this matrix:

$$\Sigma^{-1} = VDV^T \tag{2}$$

$$\Sigma^{-1} = \sum_{i=1}^{K} \gamma_i v_i v_i^T \tag{3}$$

$$\Sigma^{-1} = \sum_{i=1}^{k} \gamma_i v_i v_i^T \tag{4}$$

where k is chosen less than K to represent the precision matrix as a low rank matrix. We plot this low rank estimation in Figure 1, by discarding numbers that are below a threshold. This threshold is chosen arbitrarily, since the goal of this section is to visualize data in a sparse manner. The threshold we use in Figure 3 is 0.01. We consistently discard information below a certain threshold and the relationship between covariances and peer influences is justifiably monotonic. In the above plot, a line between two nodes shows a 'strong' correlation between two households. Note that this a plot of only a peer influence network; assuming that household with similar traits belong to a local peer network that

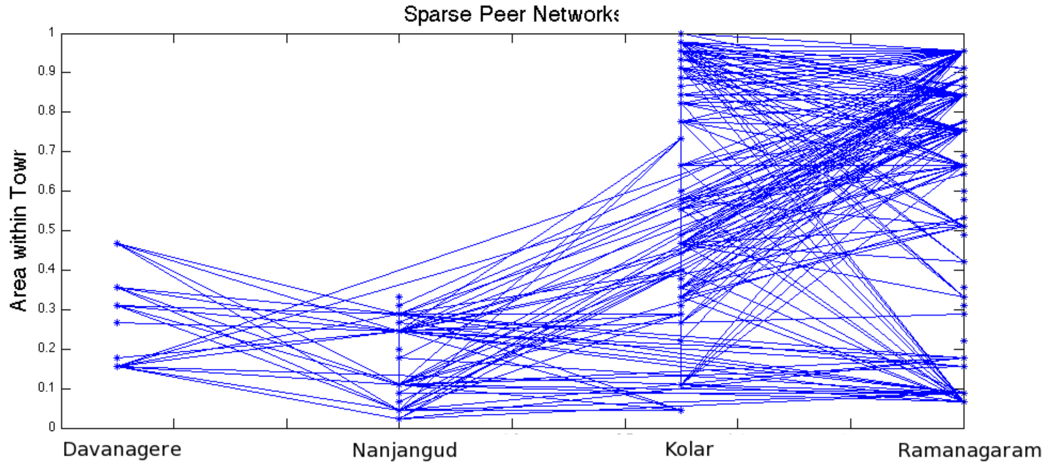

Figure 2: Peer Network in Kolar, Davenagere, Ramanagaram and Nanjangud

would influence their repayment and default behavior.

To check if the above network is an accurate picture of default behavior, we populate a sparse link matrix of default. Let A_{ij}^{m} be 1 if both households are Muslim and they both defaulted. We also build A_{ij}^{nm} in the same fashion for non-Muslims in the four survey. For the non-Muslim default network, graphical representation is easy since the defaults are few. However, to visualize the Muslim-defaults across the four towns, we need a sparse representation. We make the A matrix sparse by an equi-probable assignment to join two links when all the criteria is satisfied. Again, this is only a technique to visualize the sparse version of the Muslim default network. We plot these in Figure 2 (non Muslim) and Figure 3 (Muslim).

We observe that our sparse peer network from the precision matrix captures the realized default network fairly well. This suggests that a network of peer default would be fairly successful in capturing the dynamics of this complex social system. We will approach this problem of modeling default in a network later in this paper.

Now that we described the data using graphical models, we proceed to more formal models to characterize default. In the first step, to guide our analysis and to help fine-tune our hypotheses, we investigate the SVM kernel technique to unearth systematic issues in the data. As a first test, we group the two towns with large defaults, Kolar and Ramanagaram and test if the SVM can:

- Pick up that the default was largely a community-centric event.
- Classify the test set better than a naive linear regression based classifier.

Next we propose a technique to measure peer-influence by constructing peer-scores. In the last section of the paper, we use SVM technique to aid in feature selection, and we model the peer effects of the Muslim default on the Hindu community. We use the method

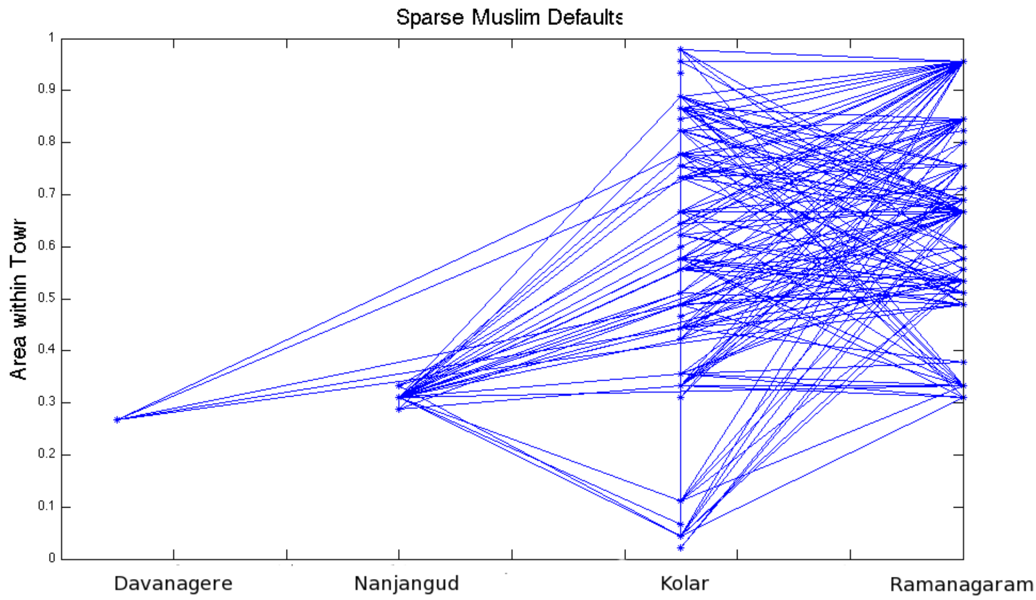

Figure 3: Muslim Default Network in Kolar, Davenagere, Ramanagaram and Nanjangud

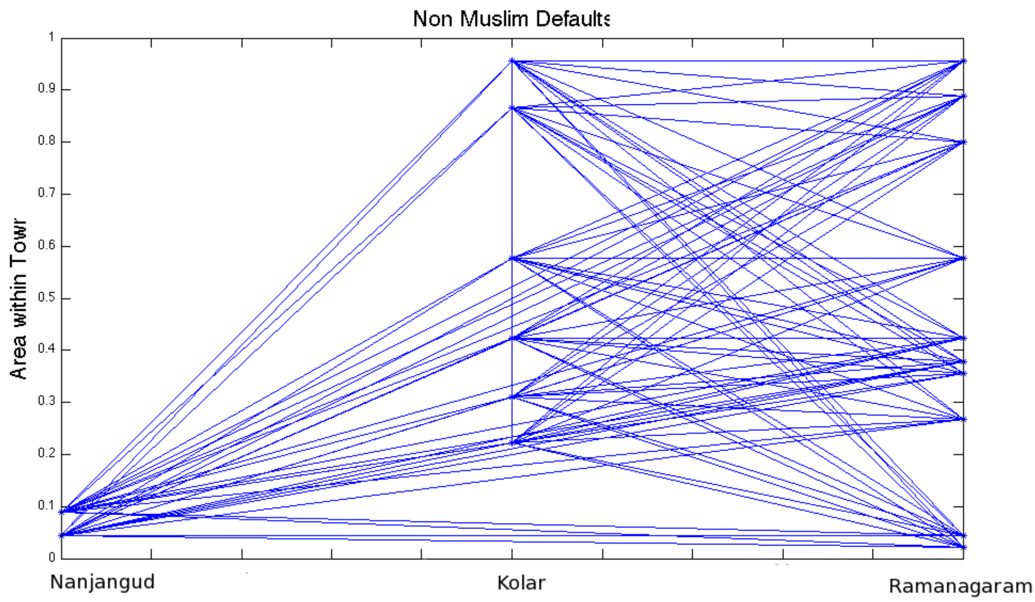

Figure 4: Non-Muslim Default Network in Kolar, Davenagere, Ramanagaram and Nanjangud

of simulated moments(MSM) to estimate the effects of peer networks and conclude by comparing our results to existing work by Gine, Krishnaswamy and Ponce[10].

4 Support Vector Technique To Predict Defaults

One of the appealing characteristics of the support vector technique is its ability to perform feature selection in an elegant manner[11]. Selecting what features are most important for default seems to be a question of natural interest; and hence the support vector technique is particularly useful given the data set we have. The survey questions as shown in Appendix A are wide in their coverage of the details of the household, including questions about the frequency of income, the number of farm animals in the house and whether the household owns a refrigerator. Furthermore, one of the classic problems in machine learning is the binary classifier problem. Since we would like to study default, the linear SVM as a binary classifier is our first choice.

We test the data from households in Kolar and Ramanagaram which exhibit a large number of defaults to see if the linear SVM is able to learn the structure of the data. The microfinance default problem can be cast as a standard classification problem in machine learning. The problem specification is: given a vector x_n that characterizes a household, given we have a training set such that $\{x_i, y_i\}$, where y_i is the default or not-default variable on $\{-1, 1\}$, can we come up with a function that classifies a new household into the two categories?

In particular , we explore the linear SVM classifier problem. The problem is that of training a linear model of the form:

$$\widetilde{Y} = sgn(w^T K(\widetilde{X}) + b) \tag{5}$$

where sgn is the sign function, $K(\widetilde{X})$ is a kernel matrix representation of characteristics of a household and $\widetilde{Y}$ is the label vector that determines whether the household defaulted or not. The kernel matrix representation of the data is simply a higher dimensional projection of the data so as to be able to find a linear hyperplane in the higher dimensional space. In our analysis, we will denote default by a -1 and +1 otherwise. Typically, the above problem is readily solved by writing the problem as that of finding a hyperplane separator with the maximum possible margin between the points that are relatively hard to classify.

Equation 5 refers to a kernel that projects a training set into a higher dimension. One of the interesting issues in using the SVM is the choice of the kernel. The software we use provides us the ability to use the kernel of our choice and also offers certain presets. Since we hope that our data will be linearly separable to a large extent, we try using the given presets in the LIBSVM. We use three types of preset kernels in our preliminary analysis - linear, gaussian and polynomial. In choosing the kernels, we inherently impose

Table 3: Results of the Kernel Learning Techniques on **Kolar and Ramanagaram**

	Gaussian	Linear	Polynomial
Overall Accuracy	56.27%	74.82%	44.27%
Default Prediction Accuracy	98.27%	83.62%	100.00%
NonDefault Prediction Accuracy	22.60%	65.75%	0%

Table 4: Results of the Kernel Learning Techniques on **Davanagere and Nanjangud**

	Gaussian	Linear	Polynomial
Overall Accuracy	91.53%	90.40%	91.53%
Default Prediction Accuracy	0%	6.67%	0.00%
NonDefault Prediction Accuracy	91.53%	98.15%	91.53%

the following structure on the data projection:

$$K(\widetilde{X}) = \widetilde{X}\widetilde{X}^T \tag{6}$$

$$K(x_i, x_j) = exp(-\frac{\|x_i - x_j\|}{2\sigma^2}) \tag{7}$$

$$K(x_i, x_j) = (ax_i^T x_j + b)^p \tag{8}$$

In equations 7 and 8, parameter selection is often the most crucial part of model fitting. In the gaussian kernel, the parameter σ has to be calibrated, and in the polynomial kernels the parameters a, b and p are selected. We use a standard off-the-shelf product called LIBSVM[6] to train the model with 400 of the 600 classified households in Kolar and Ramanagaram. In training the model, we cross validate our model using 200 points - this is usually necessary for parameter selection of the gaussian and polynomial kernels. Finally, we test it on the remaining 200 households to construct estimates of accuracy. Our results are tabulated in Table 3.

The last two measures above are non-standard and next we motivate these. The Default Prediction Rate measures accuracy only for the defaulted households. We measure this to indicate how good the predictor is for default events - however, a good predictor has to do well in predicting the households that do not default correctly too, and that is given by the last row in the tables. Table 3 is more important for our study since it contains a large number of defaults. Table 4 covers the two sample towns that do not participate in the mass defaults, hence the default prediction of the SVM performs poorly as well. The significance of the second table is clarified when we present the linear regression classifier in the following paragraph.

From the above results we see that the linear kernel shows best overall accuracy and hence we pick the linear kernel. It is to be noted that the polynomial kernel assigns a 'default' to everything, and hence has 0% predictive power for the households that do not default. The gaussian kernel does a little better; the linear kernel does a fairly good job of picking out the one important difference between the defaulters and the households that do not default. We expect to see this since the households seem to be divided on how they react to the *fatwa* announcement.

At this point we ask the question: If the SVM technique can find the best separating hyperplane, can we get any economic insight from this learning model? Consider a thought experiment where we have a data set on which we have no prior except that we know what each column stands for. We would like to form a hypothesis in the genre of causal discovery. For this we order the weights that contribute the most to the -1 default label of the defaulting household, and to the ones that contribute the most to a household paying back its loans. Note that this is not necessarily causality, it might be only association that we discover in this process. Nevertheless, when we compute the weights implied by the SVM on the different factors in the default data, we find that the top three questions in the survey *in decreasing order of importance* that are associated with default are:

- What is your religious affiliation?
- Has your economic condition changed after joining the MFI?
- Are all your credit needs met by the MFI?

These are the questions with a negative weight, and hence these are the ones that push a household towards defaulting as seen by the SVM classifier. The questions that help the classifier predict an absence in default are(in decreasing order of importance):

- How often do you ask your husband or family for money to repay the installment?
- Does the household have savings?
- Do you feel the MFIs make money like any other moneylender?

The above results give plenty of scope for discussion. The SVM captures the central theme of default - religion. It appears that when a household reported that its credit needs were not met by the MFI, it had a higher chance of default. Another interesting observation on the other side is that caste seems to play a role in a household not defaulting, but we remove this from the list since it is simply a proxy for religion (Muslims do not have a caste to report). The factors that have a positive weight are equally interesting - the SVM learns that savings are important, and so are the subjective responses to questions about the profit model of the microfinance institutions.

A benchmark for this is the standard linear regression to predict the value of the label Y. If Y is negative, assign it as a default event and vice-versa. We first apply this linear regression classifier to the Kolar and Ramanagaram data, and find that the accuracy is 74%, the SVM classifier only slightly outperforming it. However, when we use the

Table 5: SVM & linear regression classifier: **Kolar and Ramanagaram**

	SVM Classfier	Linear Regression Classifier
Overall Accuracy	74.82%	74.01%
Default Prediction Accuracy	83.62%	75.00%
Non-Default Prediction Accuracy	65.75%	73.97%

Table 6: SVM & linear regression classifier: **Davenagere and Nanjangud**

	SVM Classfier	Linear Regression Classifier
Overall Accuracy	92.09%	70.62%
Default Prediction Accuracy	6.67%	40.67%
Non-Default Prediction Accuracy	98.15%	73.46%

classifier for the other two towns which display much smaller defaults, we see that the linear regression classifier does poorly. In summary, the comparison between the linear SVM and the linear regression based classifier is shown in Table 5 and Table 6.

Further, the questions that are most important to the regression classifier for default are (i.e. these are the weights with a negative sign):

- What is your religious affiliation?
- What is the total amount you are repaying to the MFIs every month?
- What is your income during low months?

The top weights with a positive sign are The top questions that influence households not defaulting are:

- The entry that records if the household borrowed from MFI 7.
- What is your husband's average income?
- What is your income during the low months (seasonality)?

As seen above, the SVM classifier does as well as the naive regression classifier in the case of large defaults (Kolar and Ramanagaram) and beats the regression classifier by a considerable margin when it comes to evaluating the default behavior of borrowers in Davenagere and Nanjangud. More importantly, both the SVM technique and linear regression capture the central case in point: that the defaults had religious underpinnings. The outperformance of the linear SVM is much clearer when we analyze the sparse defaults in the other two towns which were not subject to shocks, as shown in Table 6.

We think that the SVM classifier outperforms the linear regression since the classifier solves an optimization problem in which there is an adjustable weight for penalizing wrongly classified points. It can be shown that solving the maximum margin problem is equivalent to[7]:

$$\min_{w,b} C \sum_{i=1}^{l} max\{0, 1 - y_i(w^T x_i + b)\} + 0.5w^T w \tag{9}$$

where w is vector of weights, and C is the parameter that chooses penalty on misclassification. In both our SVM runs, we find that the optimal C value chosen is 10. Since a linear regression classifier does not have this feature of adjusting the penalty, we think that the SVM classifier is more general in terms of adapting to the dataset. We can also ask if this difference is statistically significant - unfortunately we do not have enough data to compare how the two classifiers perform in different settings as of this point.

5 Estimating Peer Effects of Default

5.1 Peer-Scores for Hindu Borrowers

In this section, building on work by Gine, Krishnaswamy and Ponce[10], we attempt to capture the effects of peer networks in default. In line with Coates and Brandes[9], it is well known that peer social effects are significant in determining how an ethnic group behaves; however, since a shock to a community in most cases is inseparable from the characteristics of the social group, it is almost always impossible to perform a linear regression without the curse of endogeneity. However, as discussed above in the data section, we have a situation when the Hindu population is not bound to the *fatwa* issued by the Anjuman Committe, and hence exhibit strategic default as a result of Muslim influence. One open question at this point is whether Muslim influences completely capture the default behavior of Hindus, or whether there are other factors that drive it. We will answer this question in due course of time as we develop our analysis.

Approaching the problem from the perspective that Gine, Krishnaswamy and Ponce[10] take, we construct artificial indices that reflect the Muslim influence around a Hindu borrower. We construct four different kinds of indices. We should mention here that for each borrower, we have data on her town, her area and the microfinance center she is part of. The indices are:

- Peer Score (Aggregate Level): A ratio that measures the average density of Muslims in the borrower's area, town and center.
- Peer Score (Area Level): A ratio that measures the density of Muslims in the borrower's area.
- Peer Score (Center Level): A ratio that measures the density of Muslims in the borrower's microfinance center.

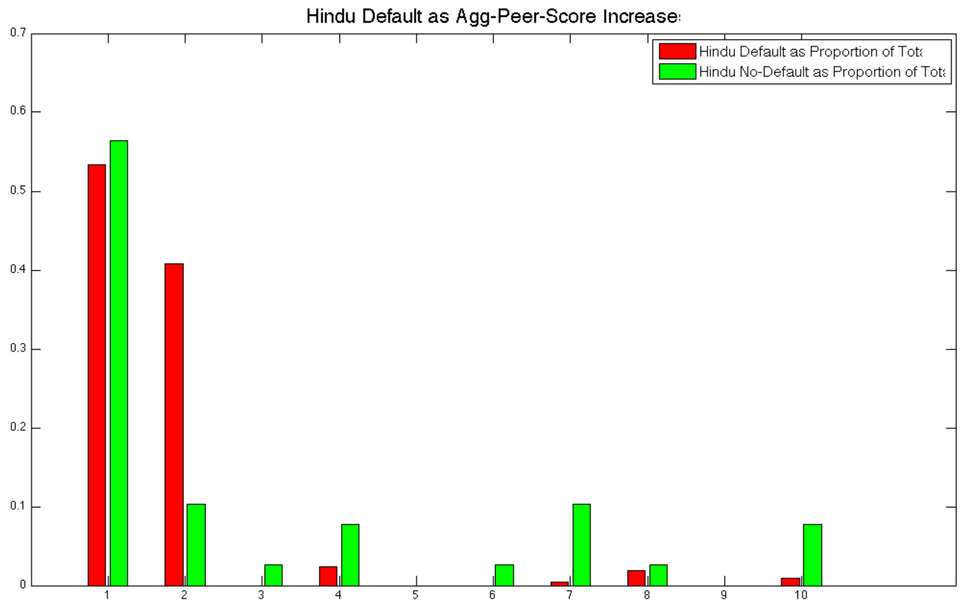

Figure 5: No of Hindu defaulters bucketed by agg-peer score

- Peer Score (Town Level): A ratio that measures the density of Muslims in the borrower's town.

In the above approach, we should clarify that we are in-effect constructing proxies for peer defaulters in the vicinity of the Hindu defaulter. Before we present any of the data, we develop some priors on the relationship between the peer score and the distribution of defaulters and non-defaulters. If we have a prior that lower peer-score leads to lower default, clearly, when the peer score is low, the number of the Hindu re-payers should outweigh the number of Hindu defaulters. When the peer score is high, these should flip over in that high peer score should see a high number of defaulters and low number of re-payers. We plot the bar graphs for the four different scores in Figure 5-8.

From the above bar graphs we notice that the distributions in the number of defaults for Hindus bucketed by peer-scores seem to have downward sloping trends. This seems opposite to what our priors are. We expect the defaults to increase as the peer-score influence increases. But we soon realize that the correct way of thinking about default is to look at the ratio of defaults to no-defaults. When we plot these in Figures 9 and 10, we notice a surprising fact - the trends on center-score and town-score are opposite! The default ratios have an increasing trend with center-score and the opposite trend with town-score. This leads us to an idea; we examine the data and compare a Hindu borrower in Kolar and a Hindu borrower in Ramanagaram. When we look for differences outside of just location, we find an interesting fact. The demographics of Ramanagaram are such

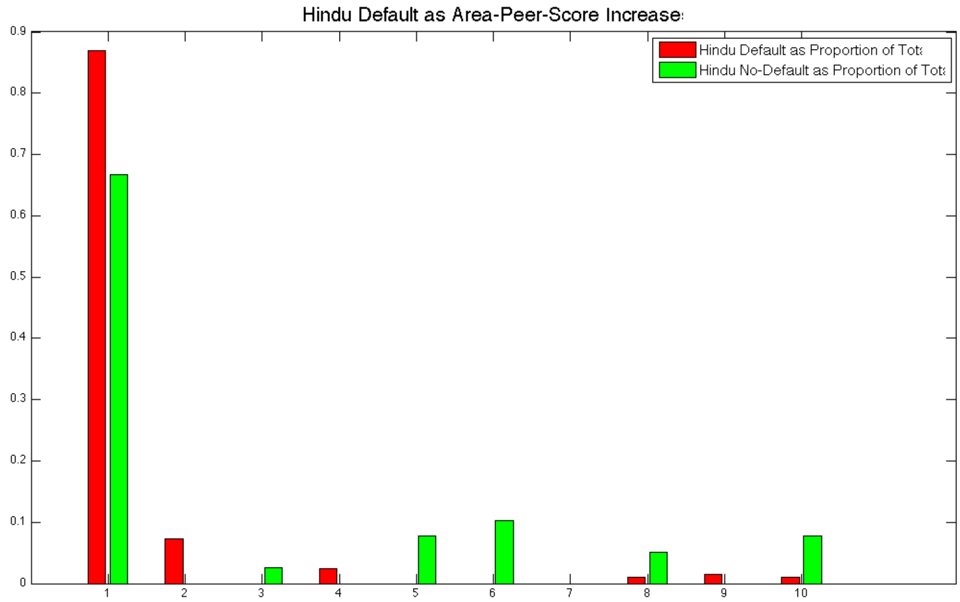

Figure 6: No of Hindu defaulters bucketed by area-peer score

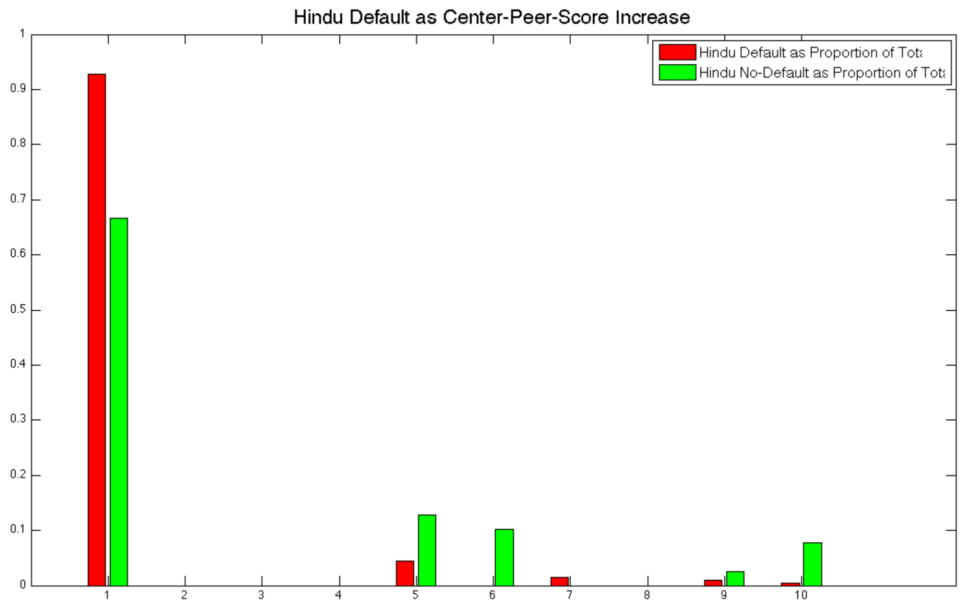

Figure 7: No of Hindu defaulters bucketed by center-peer score

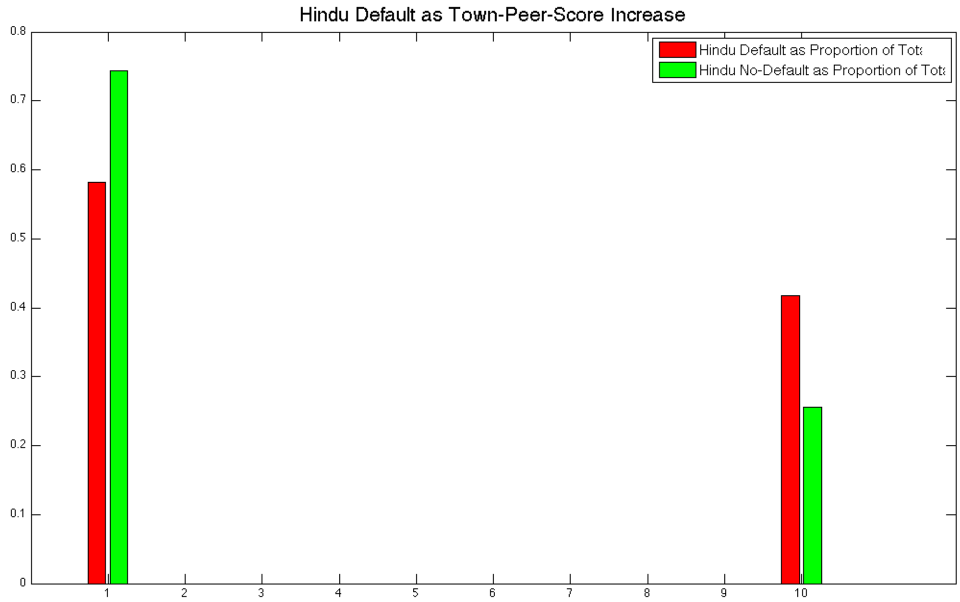

Figure 8: No of Hindu defaulters bucketed by town-peer score

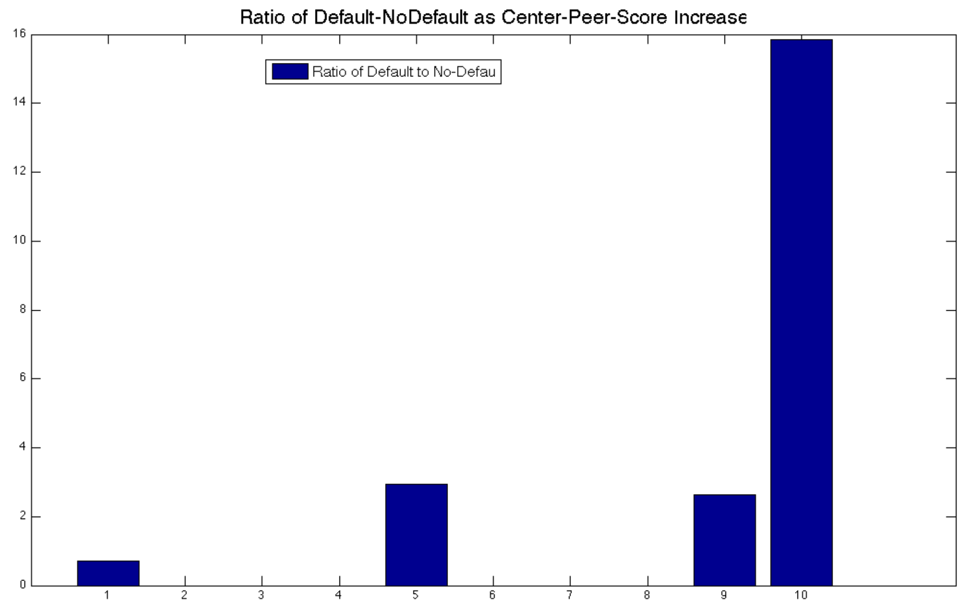

Figure 9: Ratio of Default to No-Default with increasing center-peer-score

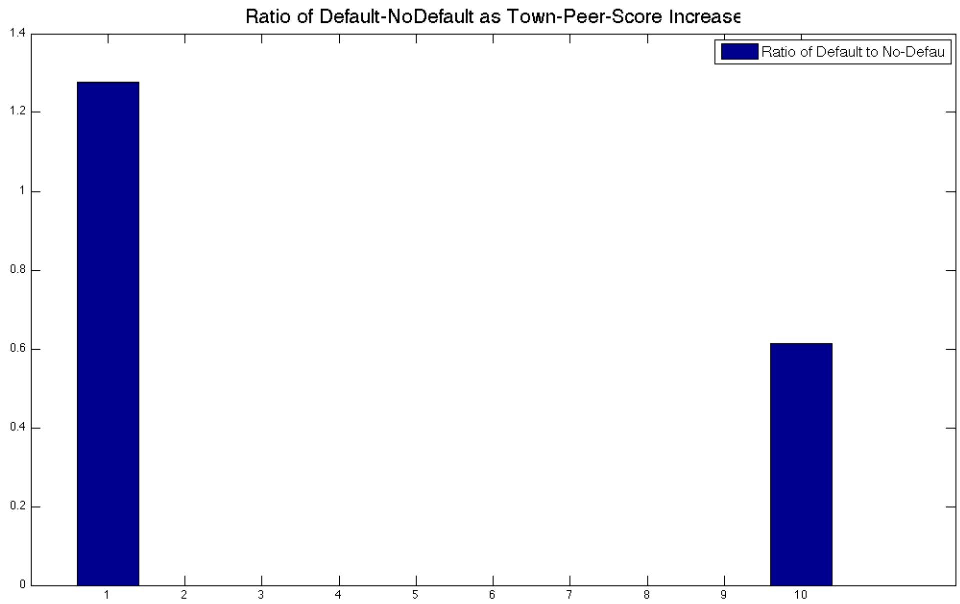

Figure 10: Ratio of Default to No-Default with increasing town-peer-score

that at the town level, the Hindu population in Ramanagaram is in larger proportion to that in Kolar (which is not surprising, the word Rama is related to Lord Rama, a Hindu deity). This creates a thresholding effect at the town level in which, since the Hindu population is above a certain critical mass, defaults in the peer group are more likely. On the other hand, Kolar, where the *fatwa* was issued, does not have a sufficient density of Hindus at the town level to breach this threshold and hence they do not default in a strategic fashion to the extent they do in Ramanagaram. Note that at the center level, there is no anomaly as Figure 9 suggests that as the center-peer influence increases, there is higher default among Hindu borrowers. This is well in accordance to what Gine, Krishnaswamy and Ponce find[10] in their paper.

5.2 Other Factors Using the SVM Technique

With the above ideas to set the stage to model Hindu default influenced by peer effects, we perform one last test to determine what the correct set of features in the discrete choice model is. For this we refer to our earlier section of the SVM technique, where we have the ability to perform feature selection. We run the SVM technique on the data set now comprising only of the Hindu borrowers. This procedure is exactly as described in our previous section. We find the following factors to be most important from the linear kernel based SVM:

- Town-Peer-Score
- Frequency of Income
- Center-ID

To briefly describe the 'Frequency of Income' section in the data, it is a scale from 1 to 7 that has the following options:

1. Daily
2. Weekly
3. Fortnightly
4. Monthly
5. Employer pays every so many days(regular)
6. Employer pays every so many days(irregular)
7. Only after completion of contracted task

We observe that from choice 1 to 4, the higher the number, the smoother the income stream of the household. This is is because a monthly wage household is more likely to be employed at a steady job compared to day worker who may not get paid on certain days. This means that a borrower who has a higher response on this question is bound to have a lower probability of default, and since we mostly see answers in the top four categories the break in monotonicity (from 5 to 7) should not be a problem. We will return to this important observation when we estimate the discrete choice model. Now that we have a set of features to set up our discrete choice model - peer scores and frequency of income, we are in a position to describe it formally.

6 A Discrete Choice Model for Default

6.1 The Model

Consider a borrower who derives a utility from her behavior, in this case either defaulting or choosing not to default. This utility may be a negative number, but since we develop the model as a difference in utilities, this should not matter. Let U_{i0} be the utility of default. Since we want to model this as a function of peer effects and the frequency of income for each borrower, we utilize the scores we developed in the last section for each Hindu participant at the center level s_c and the town level s_t and combine these with the frequency of income factor in writing the utility function:

$$U_{i0} = \alpha_{i0} + \Theta S_i + \Phi F_i + \epsilon_{i0} \tag{10}$$

where the subscript i stands for a household, and the subscript 0 refers to a default event. The above equation says that a household that defaults obtains a utility of α_0 and an extra utility of default which is positively correlated with the factors of interest: the Muslim defaulter influence around it in the form of two peer scores and the frequency of income of the household. Hence Θ is a vector of parameters, θ_1 and θ_2 stand for coefficients on the Muslim score at the center level and Muslim score at the town level. To refresh the memory of the reader, the scores are simply ratios of the number of Muslim borrowers at the center level and the town level. S_i is a vector that contains these scores.

From this point onwards, for the sake of brevity, we assume that Φ, which is the coefficient for the frequency of income, is absorbed in Θ and F_i, the frequency of income, in S_i.

We write a similar equation for no-default; clearly there is no extra utility from Muslim borrowers around the household when the household repays. Hence the utility is written as:

$$U_{i1} = \alpha_{i1} + \epsilon_{i1} \tag{11}$$

We assume ϵ_0 and ϵ_1 to be normally distributed. Further, the utility differential between no-default and default is:

$$U_{i1} - U_{i0} = \alpha^* - \Theta S_i + \epsilon_i^* \tag{12}$$

And the household defaults if the utility differential stated above is larger than 0. Hence, we have:

$$y_i = \mathbb{1}(U_{i1} - U_{i0} > 0) = \mathbb{1}(\alpha^* - \Theta S_i + \epsilon_i^* > 0) \tag{13}$$

where y_i at 0 indicates default and y_i at 1 indicates that the household did not default. The indicator function is denoted by $\mathbb{1}$ in the above expression. We show in Appendix A that the model parameters are identified(Proposition 2).

6.2 Estimation Using Non Invertible Generalized Method Of Simulated Moments

In a generic parametric econometric model[1]:

$$y_i = f(x_i, \theta_0, \epsilon_i) \tag{14}$$

where x_i is an observed variable, ϵ_i is an unobserved variable, θ_0 is the true parameter of interest that we would like to estimate and y_i is the observed dependent variable. We would also suppose that the conditional distribution of ϵ_i is specified either in a simple parametric fashion (normal or logistic) or as a mixture of parametric distributions. Let us call this conditional distribution $p(\epsilon_i|x_i)$.[1]

Given data observed as $\{x_i, y_i\}_{i=1}^N$, we would like to set up a sample moment in the following fashion:

$$\mathbb{E}\left[y_i - \mathbb{E}\left[f(x_i,\theta,\epsilon_i)|x_i\right]|x_i\right] \tag{15}$$

The expectation of any function $h(x_i)$ at the true parameter θ_0 multiplied by the above expression should be identically zero:

$$\mathbb{E}\left[y_i - \mathbb{E}\left[f(x_i,\theta,\epsilon_i)|x_i\right]|x_i \otimes h(x_i)\right] = 0 \text{ at } \theta = \theta_0 \tag{16}$$

As a result of the above expression, a value $\widehat{\theta}$ that sets the sample analog of the above moment as close as possible to zero is a consistent estimator of θ_0[6]. The expectation inside the outer expectation is obtained by simulating ϵ_i and computing the average of $f(x_i,\theta,\epsilon_i^s)$ over a fixed number of ϵ draws from the distribution $p(\epsilon|x_i)$:

$$\mathbb{E}\left[f(x_i,\theta,\epsilon_i)\right] = \frac{1}{S}\sum_{s=1}^{S} f(x_i,\theta,\epsilon_s) \tag{17}$$

And finally the estimator for θ_0 takes the form

$$\widehat{G_N(\theta)} = \frac{1}{N}\sum_{i=1}^{N}\left[(y_i - \mathbb{E}\left[f_i(\theta)\right]\otimes h(x_i))\right] \tag{18}$$

An MSM estimator is one that uses the above moment and sets it as close as possible to zero. These estimators are known to be consistent for finite S[8]. In fact, we minimize a form of the above moment to find the estimator for θ_0:

$$\widehat{\theta} = \underset{\theta}{\operatorname{argmin}}\ G_N(\theta)^T A G_N(\theta) \tag{19}$$

where A is the inverse of the variance of $G_N(\theta)$ matrix evalued at θ_{init}:

$$A = Var(G_N(\theta_{init}))^{-1} \tag{20}$$

where the variance is given by:

$$\frac{1}{N}\frac{1}{N}\sum_i\left(g_i(\theta_{init}) - \frac{1}{N}\sum_i g_i(\theta_{init})\right)\left(g_i(\theta_{init}) - \frac{1}{N}\sum_i g_i(\theta_{init})\right)^T \tag{21}$$

Finally, once we obtain the estimate for θ_0, we compute the standard deviation of our estimates by using standard results[6]:

$$\text{Asymptotic Variance} = (\Gamma^T A\Gamma)^{-1}(\Gamma^T A V A\Gamma)(\Gamma^T A\Gamma)^{-1} \tag{22}$$

where Γ is computed as:

$$\Gamma = \mathbb{E}\left[\frac{\partial G(\theta_0)}{\partial\theta^T}\right] \approx \frac{\partial G_N(\hat{\theta})}{\partial\theta^T} \tag{23}$$

and V is equation (20) computed at $\hat{\theta}$.

7 Results and Conclusions

The discussion above is general for any function f. In fact in our case, this function is the indicator function. We use Matlab to implement the simulated GMM technique described in the previous section. We minimize:

$$G_N(\theta)^T A G_N(\theta)$$

where

$$G_N(\theta) = \frac{1}{N}\sum_i \left((y_i - \mathbb{1}_s(\alpha^* - \theta^c S_i^c - \theta^t S_i^t - \theta^f S_i^f + \epsilon_i > 0)) \otimes \begin{pmatrix} 1 \\ S_i^c \\ S_i^t \\ S_i^f \end{pmatrix} \right) \tag{24}$$

where the indicator function is simulated over ϵ, a white noise process. Note our simple choice for the function h. A is the inverse variance matrix as described in the previous section.

Our estimated parameters with their standard deviations are tabulated in Table 7.

Table 7: Parameter Estimates from the Simulated GMM

	α	θ^c	θ^t	θ^f
Value	-0.369	2.13	-1.79	-0.435
Standard Deviation	0.021	0.019	0.012	0.015

As we see in Table 7, the coefficient θ^c, which is the microfinance center level effect, is positive. This points to the existence of an increased risk of default as the number of muslim borrowers in the center increases. This shows that although the joint liability model reduces idiosyncratic risk by design (since the risk of each borrower individually defaulting is low), it leads to higher systemic risk in the form of a peer effect that we measure through θ^c.

The coefficient θ^t that represents town level peer effects is of the opposite sign than that of θ^c. This is an interesting finding since we expect default influences to always sway borrowers in the same direction; however, we find from the above analysis that default influence at the *center* level works in an expected manner (increases systemic risk) (θ_c has a positive sign), where as default influence at the *town* level works in an opposite manner (θ_t has a negative sign). That is, Hindu borrowers who comprise of a smaller fraction of the total population in their town appear to demonstrate a reverse effect of not defaulting, since a higher Muslim density in their towns naturally means a lower density of Hindus and hence does not give strategic default enough momentum. On the other hand, at the center level, peer-influences sway the Hindu borrowers towards default. This ties well to Gine, Krishnaswamy and Ponce's work[10]. This is also in accordance to what we see in the data in Figure 9 and 10, that even though the *fatwa* was issued in Kolar, Ramanagaram sees a higher proportion of Hindu defaulters.

The coefficient for the frequency of income has a negative sign. This shows that the daily wage households get a larger benefit from default. This could be because daily wage workers usually end up seeing larger volatility in their income streams. Further, it is the daily workers that see lower wages overall as compared to people with a monthly income, possibly at a nearby factory.

In summary, we show that a linear classifier based on a support vector technique learns the cause of the default event in a microfinance setting. We construct graphical representations for peer networks through the towns surveyed. We use the SVM technique on Hindu borrower data to identify important factors in the default behavior - peer scores (which are constructed) and frequency of income (which is from the survey). We set up a utility-based discrete choice model for default, model it using a simulated GMM approach and find evidence for increased systemic risk at the center level. Our findings on peer-influences at the center level are well aligned with past work[10]. We find counter-intuitive results at the town level - when population ratios for a community are small, lack of sufficient momentum among a community of defaulters leads to lower rates of default. Finally, we point out that frequency of income for the borrowers is also a significant factor in determining default in the towns of Kolar and Ramanagaram.

Appendix A

Proposition 1: Consider a multivariate gaussian distribution $X \sim N(0, \Sigma)$ where Σ is $p \times p$ covariance matrix. Then Variables X_i and X_j are independent iff $\Sigma_{ij}^{-1} = \Sigma_{ji}^{-1} = 0$ where Σ^{-1} is the precision matrix.

Proof: Consider the joint distribution of x_i and x_j in a k-variable multivariate normal distribution:

$$f(X_i X_j) = \frac{1}{2\pi|\Sigma|} e^{(X-\bar{X})^T \Sigma^{-1} (X-\bar{X})} \tag{25}$$

Further we see that:

$$f(X_i) = \frac{1}{\sqrt{2\pi}|\sigma_i|} e^{-\frac{1}{2}(X_i - \bar{X}_i)^2 \sigma_i^{-2}}$$

$$f(X_j) = \frac{1}{\sqrt{2\pi}|\sigma_j|} e^{-\frac{1}{2}(X_j - \bar{X}_j)^2 \sigma_j^{-2}}$$

And trivially:

$$f(X_i)f(X_j) = \frac{1}{2\pi|\sigma_1\sigma_2|}e^{(X_i-\bar{X}_i)^2\sigma_i^{-2}+(X_j-\bar{X}_j)^2\sigma_j^{-2}} = \frac{1}{2\pi|\Sigma|}e^{(X-\bar{X})^2\Sigma^{-1}(X-\bar{X})} = f(X_iX_j) \tag{26}$$

Hence X_i and X_j are independent. If Σ_{ij}^{-1} is not zero,

$$f(X_i)f(X_j) = \frac{1}{2\pi|\sigma_1\sigma_2|}e^{(X_i-\bar{X}_i)^2\sigma_i^{-2}+(X_j-\bar{X}_j)^2\sigma_j^{-2}} \neq f(X_iX_j) \tag{27}$$

■

Proposition 2: In the model stated in (13), the model parameters θ_c and θ_t are identified.
Proof:

$$\begin{aligned} Pr(y_i = 1|S_i) &= Pr(\mathbb{1}(\alpha^* - \theta_c S_i^c - \theta_t S_i^t + \theta_f S_i^f + \epsilon_i^*) = 1) \\ &= Pr(\alpha^* - \theta_c S_i^c - \theta_t S_i^t - \theta_f S_i^f + \epsilon_i^* > 0) \\ &= Pr(\epsilon_i^* > -\alpha^* + \theta_c S_i^c + \theta_t S_i^t + \theta_f S_i^f) \\ &= 1 - F_\epsilon(-\alpha^* + \theta_c S_i^c + \theta_t S_i^t + \theta_f S_i^f) \end{aligned}$$

Further,

$$-\alpha^* + \theta_c S_i^c + \theta_t S_i^t + \theta_f S_i^f = 1 - F_\epsilon^{-1}\left[1 - Pr(y_i = 1|S_i)\right]$$

And finally we differentiate by S_i^t and S_i^c to identify all parameters:

$$\theta_c = \frac{\partial}{\partial S_i^c} F_\epsilon^{-1}\left[1 - Pr(y_i = 1|S_i)\right]$$

$$\theta_t = \frac{\partial}{\partial S_i^t} F_\epsilon^{-1}\left[1 - Pr(y_i = 1|S_i)\right]$$

$$\theta_f = \frac{\partial}{\partial S_i^f} F_\epsilon^{-1}\left[1 - Pr(y_i = 1|S_i)\right]$$

And α^* is now identified. Hence we have shown that given the conditional cumulative distribution, all the parameters of the model are identified. ■

References

[1] Daniel A. Ackerberg. A new use of importance sampling to reduce computational burden in simulation estimation. *Quant Mark Econ*, 7:343–376, 2009.

[2] Beatriz Armandariz and Jonathan Morduch. *The Economics of Microfinance*. The MIT Press, 2010.

[3] Abhijit Bannerjee, Arun G. Chandrasekhar, Esther Duflo, and Mathew O. Jackson. Diffusion of microfinance. *Working Paper*, December 2011.

[4] Emily Breza. Peer effects: Evidence from the krishna default crisis. *Preprint*, 2011.

[5] Erica Field and Rohini Pande. Repayment frequency and default in microfinance: Evidence from india. *Journal of the European Economic Association*, 6(2-3):501–509, April-May 2008.

[6] Lars Peter Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4), July 1982.

[7] Tibshirani R. Hastie, T. and J.H. Friedman. *Elements of Statistical Learning*. Springer, 2001.

[8] Daniel McFadden. A method of simulated moments for estimation of discrete reponse models without numerical integration. *Econometrica*, 57(5):995–1026, 1989.

[9] S. Coates T. Besley. Group lending, repayment incentives and social collateral. *Journal of Development Economics*, 46:1–18, 1995.

[10] Karuna Krishnaswamy Xavier Gine and Alejandro Ponce. Strategic default in joint liability groups: Evidence from a natural experiment in india. *Draft*, August 2011.

[11] Jan O. Pedersen Yiming Yan. A comparative study on feature selection in text categorization. *Morgan Kaufmann Publishers*, 1997.