Demand and supply of microcredit in presence of selection

Marup Hossain*

January 2, 2018

Please click HERE for the updated version

Abstract

We study whether BRAC, a Microfinance Institution (MFI) based in Bangladesh, uses household performance in a livestock transfer program as a signal of creditworthiness to improve targeting in a subsequent microcredit program. We use data from a Randomized Control Trial (RCT) based livestock transfer program where beneficiary households were also encouraged to take loan from BRAC at the end of the intervention. We find that efficiency in the livestock activity explains subsequent loan approval decision, which implies that BRAC monitors beneficiary households' performance during the intervention period and updates subjective judgment along with other observable characteristics. Results from both sample selection model and Random Forest (RF) based machine learning model show that information from the livestock transfer program benefits loan officers to reject loan application of bad borrowers, hence, reduces adverse selection in the microcredit program. We show the potential of economies of scale and inapplicability of pure machine-based credit scoring model in the Microfinance sector.

*University of Florida. Email: maruphossain@ufl.edu

1 Introduction

Asymmetric information generates moral hazard and adverse selection problems in financial market that can result in credit rationing, mispricing of risk, and market failure (Cull and Morduch, 2017; Crawford et al., 2015; Hulme and Mosley, 1996). Typical financial institutions use collateral requirement and (or) interest rate as screening mechanisms to minimize credit default (Machauer and Weber, 1998; Stiglitz and Weiss, 1981). Microcredit lenders use different strategies as they do not require collateral; they instead use group lending as a tool for peer selection and reinforcement along with regular repayment schedule and non-refinancing threats (Giné et al., 2010; Armendáriz de Aghion and Morduch, 2000; Zeller, 1998; Besley and Coate, 1995). However, microcredit lenders need to rely on credit scoring approach and (or) subjective-judgment of loan officer for individual lending (Agier and Szafarz, 2013). Credit scoring method has advantages of minimizing high data collection cost and reducing subjective bias of loan officers, but a high degree of uniqueness among micro-entrepreneurs and lack of certified financial statements make credit-scoring method mostly inapplicable in microcredit sector especially in developing countries. On the other hand, the success of the subjective judgment approach depends largely on loan officer's capacity in screening out potential good borrower (Armendáriz de Aghion and Morduch, 2000).

The screening problem in microcredit market mainly arises due to a substantial heterogeneity in entrepreneurs' ability and return to capital. For example, De Mel et al. (2008) show that return to capital varies from 0% to 5.3% per month among microenterprises in Sri Lanka. It is also difficult to estimate return to capital accurately for homestead-based or small enterprises due to their low engagement in market transactions. Moreover, a gross return to capital reflect true demand for credit in some instances, for example, Samphantharak and Townsend (2017) argue that enterprise activities are usually associated with aggregate and idiosyncratic risks, therefore, a risk-adjusted return to capital is a better proxy of demand for credit. Several studies use ex-ante peer-referral or subjective indicator to predict credibility of potential borrowers as screening mechanisms. Bryan et al. (2015) show that peer referral can add information on loan approval, but it doesn't necessarily provide information on repayment type in South Africa. Similarly, Maitra et al. (2017) show that local trader agents' referred borrowers gain 27% more in cash crop production and 22% in farm income compared to self-selected group-based borrowers in India. Rigol et al. (2017)

show that entrepreneurs in the top tercile of community predicted growth potential and business profitability have a three-fold higher return than average. An alternative approach to the above ex-ante prediction mechanisms can be the use of ex-post information on performance of potential borrowers from a prior program to improve targeting in a subsequent program. For example, it is common that microfinance institutions (MFI) usually work in different programs, in such case, MFIs can use performance in one intervention as a proxy of credibility for a microcredit program. Interlinking of multiple programs has the potential of gaining economies of scope by MFIs, although it is difficult to quantify how and to what extent one program can benefit another program.

In this study, we estimate whether BRAC, one of the largest MFI in the world, uses performance of beneficiary households in a livestock transfer program as a signal of credibility for its microcredit program. We use information of beneficiary households from a Randomized Control Trial (RCT) based livestock transfer experiment in Bangladesh conducted by Bandiera et al. (2017) where beneficiary households are also encouraged to take a loan from BRAC at the end of the experiment. The program named as "Targeting the Ultra-poor" (TUP) transfers livestock assets, provide entrepreneurship training, and give subsistence stipend to poor women in rural areas of Bangladesh. The TUP program provides training and other services in initial two years, and the beneficiary households are expected to graduate from poverty to a sustainable self-employment based livelihoods by this period. Beneficiary households can take loan from the BRAC-microcredit program if they have potential investment plan during or after the graduation period. During the intervention period, BRAC officers visit beneficiary households periodically to monitor transferred assets, provide training, or discuss potential investment plans. Thus, we assume that BRAC knows the performance of beneficiary households and has the potential of using both observable and unobservable characteristics as a proxy of their borrowing credibility.

We analyze credit market outcomes of 3,677 beneficiary households after four years of the intervention and find that 27% of beneficiary households take at least one loan from BRAC-microcredit program. More importantly, we find that BRAC rejects loan application of another 8% of the beneficiary households, which raises a question of why those households are rejected? If the performance in the livestock activity monitored by BRAC is not a crucial factor in loan approval, we would expect that borrowers' current observable characteristics would be enough to explain loan rejection decision. On the other hand, if performance in the livestock activity is a significant predictor of loan rejection after controlling other observable characteristics, it will imply that individual judgment of BRAC officers as a screening mechanism is an important factor in loan approval by BRAC.

We first use the bivariate probit model with sample selection to find out factors affect household microcredit demand and loan approval outcomes simultaneously. To test our hypothesis on the role of individual judgment of BRAC officer, we control household efficiency in livestock production in both participation and outcome equations. In addition, we control risk-adjusted income from livestock activity in the participation equation following Samphantharak and Townsend (2017). We find that households with a positive risk-adjusted income are eight percentage points more likely to demand microcredit from BRAC. Presence of an additional male member increases the probability of demand by approximately 11 percentage points. As expected, we find that efficiency in livestock activity is a significant predictor of loan rejection by BRAC. A percentage point increase in efficiency reduces the probability of rejection by 56 percentage points. At the same time, a point increase in aggregate asset index reduces rejection probability by 15 percentage points. We find a consistent result for both time-constant and time-varying efficiency indicators. This result confirms that efficient households are more like to get loan approval from BRAC, which implies that BRAC uses both observable characteristics and performance from livestock experiment in loan approval decisions in the microcredit program.

We next answer a more border question on whether and to what extent the livestock intervention generated additional observable and unobservable information for BRAC? More specifically, could BRAC predict the final credit outcomes using only baseline information or do information from the follow-up rounds have better predictive powers? How predictive are beneficiaries' baseline subjective information on ability to start entrepreneurial activities, their mobility, and community declared asset level to their credit market outcomes? We apply the Random Forest Classifier, a machine learning algorithm, to predict credit outcomes and find out variables those have most predictive power using information from different survey rounds: baseline (objective and subjective indicators), first follow-up, and second follow-up rounds. We find that additional information from follow-up rounds have 100%-200% better prediction ability in truly revealing rejection decision by BRAC than baseline information, which suggests that BRAC gained both observable and unobservable information through the livestock experiment. Our results also show that efficiency in livestock activity, income, asset holding, and age of the household head are most important variables to explain microcredit market outcomes.

Our final hypothesis is to examine if BRAC were able to lend to households with a higher return on capital? Because most of the targeted households had no productive activity before this intervention, we couldnot estimate the return to capital from the baseline data. Therefore, we use information from the first follow-up round and estimate the return to capital in a way that it reflects households pre-program differences in return. We find that return to capital is seven percentage point less for rejected households compared to approved households in livestock activity. Results on returns to capital also reinforce our earlier findings that BRAC was able to identify good borrower to reduce adverse selection problem in microcredit market.

Our study complements earlier studies as well as adds new evidence on using ex-post information from one program to improve targeting in another program. Studies by Bryan et al. (2015), Maitra et al. (2017), and Rigol et al. (2017) show effectiveness of using ex-ante information or referral on return to capital and repayments. We show that monitoring performance of households enriches loan officers' subjective judgment, which in turn improves targeting in the microcredit market. This finding presses the importance of combining both individual judgment and credit scoring methods together in the microfinance sector especially in developing countries. Previous studies by Vogelgesang (2003) and Van Gool et al. (2012) on the feasibility of using only credit scoring method also mention importance of using a combined method for microcredit. Our study also indirectly indicates a potential of gaining economies of scope for the MFIs; large MFIs like BRAC can use one program to improve performance in another program. Hartarska and Parmeter (2009) show that gain in the economics of scope for MFIs mostly comes from savings in fixed cost; MFIs cannot take advantage of using both savings and lending together. Therefore, our study brings a new evidence on potential economies of scope through interlinking multiple programs. Finally, we also add an additional dimension in the literature on the impact of livestock transfer program. Studies by Bandiera et al. (2017), Banerjee et al. (2015a), and Emran et al. (2014) concentrate only on the direct effect of livestock transfer. We show that demand for microcredit is around 35% among the beneficiary households in the livestock program, which is higher than typical microcredit demand Banerjee et al. (2015b). The livestock program also has a potential to work as a pre-screening stage for subsequent large programs such as microcredit.

This study is organized as follows. We discuss the livestock program and transition from the livestock program to the microcredit program in the next section. In the third section, we detail about the data used in this study. In the fourth section, we discuss methodology followed by results in the fifth section. In the sixth section, we introduce Random forest method to show prediction followed by the return to capital in section seven. We conclude in section eight.

2 Context

2.1 *TUP program*

BRAC launched the TUP program in 2002 targeting asset-poor women in rural Bangladesh, a group who are among the hardest to reach through conventional anti-poverty programs. BRAC initiated the TUP program based on its experience from the Income Generation for Vulnerable Group Development (IGVGD) program, a food transfer accompanied with skill training, launched jointly by BRAC and the World Food Program (WFP) in 1985. Although the IGVGD program was successful in increasing income of participant households, it largely failed to generate sustainable impacts due to mistargeting and ineffective service packages (Ahmed et al., 2009; Hashemi, 2001). Based on lessons from the IGVGD program, BRAC introduced the TUP program in 2002 as an asset-based transfer scheme along with an overhauled targeting strategy.

The TUP program transfers productive assets such as livestock or other assets for small-scale retail operation. It also provides skill training for better utilization of transferred assets and bi-weekly subsistence allowance in the first 40 weeks to compensate for any shortfall in income. Additionally, beneficiary households receive complimentary health care and social awareness services from BRAC staffs. Livestock assets are transferred in six different combinations of either one or two types of livestock, for example, cow, goat, or chicken. Bandiera et al. (2017) mention that about 91% of beneficiary households select an asset bundle containing at least one cow. All offered asset bundles had a similar value at around $140. Participants are encouraged to retain transferred assets for at least two years; however, they can exchange current assets for other income generating assets during this period. Skill training includes an initial classroom training, weekly visit of BRAC staffs for the first two years, and monthly or bi-monthly visits of livestock specialists for the first year (Bandiera et al., 2017).

2.2 *Targeting strategy of the TUP program*

BRAC follows a three-stage targeting strategy to identify eligible households for the TUP program. In the first stage, BRAC identifies poorest districts and sub-districts in Bangladesh using the WFP poverty mapping for Bangladesh. Later BRAC identifies all poorest communities within each sub-district in consultation with local BRAC staffs. In the second stage, BRAC divides all selected communities into clusters of 80-120 households and conduct a Focus Group Discussion (FGD) in each cluster to rank households based on their wealth. All households in bottom 2/3 categories are termed as "community-defined ultra-poor" and are selected for the final stage of screening. In the final stage, BRAC conducts a household level survey among all selected households to identify who meet at least three out of five inclusion conditions and none of the exclusion conditions. Inclusion criteria are as follows: household is dependent on female domestic work or begging, holds less than 10 decimals of land, has no adult active male member, school-aged children are engaged in paid work, or possesses no productive work. Exclusion criteria are as follows: household has no active adult women, is a microfinance participant, or is a beneficiary of government or non-government development project. Finally, households who meet both inclusion and exclusion criteria are selected to participate in the program.

2.3 *From TUP to Microcredit program*

The objective of the TUP program is to build a sustainable livelihood for beneficiary households, which BRAC terms as "Graduation model". According to BRAC a household graduates from poverty if it satisfies some economic and social indicators: no self-reported food deficit, holds multiple income sources, uses a sanitary latrine, has access to clean drinking water, has cash savings, sends children to school, and adopts family planning etc. The whole graduation process takes a maximum of two years and according to BRAC, 95% of the beneficiary households satisfy those conditions at the end of the two-year-long intervention. Throughout the graduation period, beneficiary households become more capable to utilize financial assets effectively, and BRAC takes initiatives to transfer households from the TUP program to the microcredit program.

After the initial asset transfer, BRAC designates all beneficiary households to the nearest microfinance branch and Village Organization (VO), a platform to collect installment and regular discussion between BRAC and microcredit borrowers. BRAC arranges an orientation meeting about microcredit program for all beneficiary households before their graduation. All TUP and microfinance officers of that area attend the orientation meeting. During the orientation meeting, BRAC provides detailed information about the microcredit program and its terms and conditions to all households. At the end of the TUP intervention, BRAC transfers savings amount of the beneficiary households from the TUP branch to the pre-assigned microcredit branch. At this point, a beneficiary can withdraw all her savings if she does not want to continue savings. A TUP member can apply for microfinance loan from their respective branches after graduation; in some cases, beneficiary households can also apply for a loan before graduation if they have potential investment plans.

3 Data

We use data set from a Randomized Control Trial (RCT) experiment conducted to evaluate the second phase of the TUP program. The second phase of the program named as Challenging the Frontiers of Poverty Reduction (CFPR) was initiated in 40 districts of Bangladesh in 2007. Following the targeting strategy of the TUP program, the research team followed a multistage sampling procedure to identify the sample of this study. First, the research team randomly selected 20 sub-districts from 13 poorest districts of Bangladesh. Within each sub-district, one BRAC branch office is randomly selected as treatment branch and one as control branch. All villages within 8 km radius of each branch are considered as the study area. Following all targeting stages as mentioned above, this study covers 13 districts, 20 sub-districts, 40 branches, and 1,409 communities. Total sample of this experiment is 6,305 who are surveyed in 2007(baseline), 2009(medium term), and 2011(longer term) where 3,691 are in the treatment group and 2,614 are in the control group. We use information of 3,691 households from the treatment group over the period of 2009 and 2011 in this study. Among the treatment households, 14 households had no livestock activity in any round, so we drop those observations. Therefore, this study is based on 3,677 households who were involved in livestock rearing in at least one round.

4 Summary statistics

Table 1 shows summary statistics of the beneficiary households over survey rounds. A substantial proportion of the targeted women, who are respondent in the baseline survey, have no formal education; average years of education is only 0.60 years. The average age of the targeted women is 38 years in the baseline. A male member heads approximately 63% of the beneficiary households, which is lower than the national average (87%) implying that some these households do not have any adult male member at all (Joshi, 2004). Household heads also have less than one years of formation education or no education at all. TUP households are typically less populated; average household size is approximately 3.5, which is 4.50 nationally. Many TUP households have no working aged male member; the average working aged male member is less than one. Lack of working aged member in the targeted households aligns with the targeting strategy of the program; one of the inclusion criteria of the TUP program is that households have no working aged member. TUP households held no productive assets. The average number of cow was only 0.08 and 0.15 for goat. Average land holding is only 1.43 decimals, which reflects most of the TUP households even do not own homestead area.

[Table 1 here]

Our results show only 27% of beneficiary households received at least one loan from BRAC from 2008 through 2011. 12% of the beneficiary households took at least one loan and 15% took a single loan from BRAC. Around 8% households applied for a loan but were rejected by BRAC. The remaining 65% of households did not apply for a loan for distinct reasons. For example, 15% of households stated that they did not find the terms and conditions of BRAC loan program suitable and 13% households said they did not need a loan. We present household loan amounts from various sources by TUP loan status in table 2. It is evident that TUP households are not much involved in financial markets irrespective of their loan status from BRAC. Loans from formal banks are virtually zero for both groups over time. Relative or friends are main sources for loans to both groups, although there is no significant difference in loan amount in the longer run.

[Table 2 here]

5 Methodology

We show in the previous section that only 35% of the beneficiary households applied for at least a loan from BRAC. Conditional on demand for microcredit, 23% of the applicant was rejected from receiving a loan and 77% received at least a loan from BRAC. Both demand and supply of microcredit can be influenced by various socioeconomic, demographic, and other factors, however, these influencing factors may be different on two sides. Therefore, a sample selection model will be more appropriate than a Tobit-type model to model microcredit outcomes in our study. Let's assume that supply of microcredit (S_i) (outcome equation) for household i is a function of vector$X$ and demand of microcredit (D_i) (participation equation) is a function of vector Z. We define both outcome and participation equation as follows

$$S_i = (X_i^{'}\beta + u_{1i} > 0) \tag{1}$$

$$D_i = (Z_i^{'}\beta + u_{2i} > 0) \tag{2}$$

Where $u_1 \sim N(0,1)$, $u_2 \sim N(0,1)$, and $corr(u_1, u_2) = \rho$. We observe S_i if and only if $D_i = 1$; if $D_i = 0$, we have no information about S_i. Therefore, we have three types of households in our study: $D_i = 0,\ D_i = 1,\ S_i = 0$, and $D_i = 1,\ S_i = 1$. The log-likelihood function will be,

$$\begin{gathered}\ln L = \textstyle\sum_{i=1}^{n} \{D_i S_i \ln \Phi_2(X\beta, Z\gamma; \rho) + D_i(1 - S_i) \ln[\Phi(X\beta) - \Phi_2(X\beta, Z\gamma; \rho)] \\ + (1 - D_i) \ln \Phi(-X\beta)\}\end{gathered} \tag{3}$$

To identify the equation (3), we control an additional binary indicator on whether household's risk-adjusted income is positive or not in the participation equation. Following Samphantharak and Townsend (2017), we argue that demand of microcredit for subsequent investment will depend on beneficiary households risk-adjusted income from the livestock activity. It is also expected that BRAC as large MFI will not consider aggregate risk and, arguably, idiosyncratic risk as an important factor in supplying microcredit to an individual household i. To test our main hypothesis on the role of monitoring by BRAC, we control household efficiency in livestock production in both participation and outcome equations. Next, we explain estimation process of both risk-adjusted income and efficiency, consecutively.

5.1 *Risk-adjusted Income*

We formulate a modified production function to estimate risk-adjusted income, income net of risk premium, for each household i,

$$y_{it} = f(x_{it}; \beta) + \varepsilon_{it} \tag{4}$$

where ε_{it} is the aggregate error term. We estimate conditional heteroscedasticity ($\hat{\varepsilon}^2_{it}$) from the equation (4) as a proxy of variability in income. As the modified production function does not include any market-return variable, $\hat{\varepsilon}^2_{it}$ consists of both aggregate and idiosyncratic risks with an that measurement error is random. The next step is to estimate risk-adjusted income using the following regression specification,

$$y_{it} = \alpha_0 + \alpha_1 \hat{\varepsilon}^2_{it} + \varepsilon_{it} \tag{5}$$

we predict risk premium from the equation (5) as $\hat{\alpha}_1 \hat{\varepsilon}^2_{it}$ and estimate the risk-adjusted income as $\tilde{y}_{it} = y_{it} - \hat{\alpha}_1 \hat{\varepsilon}^2_{it}$.

5.2 *Estimation of efficiency*

We estimate efficiency in livestock production using a stochastic production frontier model. Suppose that income from livestock rearing is a function of both observed inputs and unobserved efficiency[1], which we can present using the following function,

$$y_{it} = f(x_{it}; \beta)\xi_{it} \exp(v_{it}), \qquad i = 1, ..., N;\ t = 1, ..., T \tag{6}$$

where y_{it} is income from livestock production for household $i$at time t, $f(x_{it}; \beta)$ is a linear function of inputs (x_{it}), ξ_{it} presents the level efficiency of household i within an interval of (0,1], and v_{it}captures random shocks in the production technology. Assuming u_{it}=$-\ln(\xi_{it})$ such that $u_{it} \geq 0$, we can

[1]We use a production frontier model instead of a profit frontier model due to unavailability of price of livestock products; the survey has information on input uses and net income from livestock rearing only. Using a production frontier instead of profit frontier will not a big problem in our study as our goal is estimate household inefficiency compared to a benchmark level. Kumbhakar et al. (2015) mention that if profitability indicates the rate of return, or the objective is to find out variability in profits given a reference point, one can simply use the production frontier instead of dealing with specific profit function.

transform the equation (6) into linear logarithm form as follows,

$$\ln(y_{it}) = \beta_0 + \sum_{j=}^{k} \beta_j \ln(X_{jit}) + v_{it} - u_{it} \tag{7}$$

u_{it} is a time-variant truncated-normal random variable defined as: $u_{it} = \exp\{-\eta(t - T_i)\}u_i$, where T_i is the last period in the ith panel and η presents a decay parameter. The assumption of $u_{it} \geq 0$ implies that $0 < \xi_{it} \leq 1$. $u_i \sim iid\ N^+(\mu, \sigma_u^2)$ and $v_{it} \sim iid\ N(0, \sigma_v^2)$; both u_i and v_{it} are independent of each other and other covariates in the model. Technical-efficiency $(\exp -(u_{it})|\varepsilon_{it})$ is estimated from the conditional distribution function of $f(u|\varepsilon)$, where $\varepsilon = y - x\beta$.

6 Findings

We estimate both time-varying and time-constant efficiency indicators using the stochastic production function. Although the livestock intervention lasted only for first two years, we use information of total 4 years to estimate efficiency with an anticipation that when BRAC officers judge a household performance, they consider both current indicators as well as the future prospect. Utilizing both current and future information is especially useful in livestock production because it takes a relatively long time to get a return. While estimating risk-adjusted income, we assume that measurement error is random, if any, and adjusted livestock income for both idiosyncratic and aggregate risks. As the beneficiary households in this study are very poor and most of them just started their first self-employment activity through this intervention, so they likely to account both types of risk in deciding about a potential new investment or taking a loan from a financial institution. Finally, we use the bivariate probit model with sample selection to estimate the model. Table 3 shows results from sample selection model with time-varying and time-constant efficiency indicator.

[Table 3 here]

Form the demand equation, we find that positive risk-adjusted income, asset index, and the number of male working aged members in the household have a significant and positive effect on microcredit demand from BRAC. Surprisingly, we find that pre-treatment subjective indicators on women entrepreneurial prospect, for example, communication with outside male person, mobility outside of the homesteaded area, and bargaining capacity to initiate a new business reduce the probability of demand for microcredit. Positive effects of risk-adjusted income and profit are expected as both indicators likely to incentivize households for further investment and, hence, getting a loan from BRAC. Importance of male member for economic activities is expected as well given the social and cultural context in Bangladesh. Earlier literature shows that female participation in economic activities is significantly lower than that of men, according to BBS (2010), the rate of labor force participation among the population aged 15 years or older is only 36% for females compared to 83% for males. Female workers in the agricultural sector work in activities such as vegetable gardening, livestock production, and aquaculture within or near their homesteads, and many of them tend to be unpaid laborers (Khan et al., 2009; Kabeer, 2012). Although women started livestock activity due to the transfer from BRAC, they might depend on the presence of male member when it's about taking a loan or investing in economic activities. In fact, Roy et al. (2015) find that livestock transfer program increased women's ownership of productive assets, but new investment is largely owned by men. They also find that the livestock transfer program reduces women's movement outside of home and control over income. This might explain part of our finding on negative association among women entrepreneurial indicators and demand for microcredit.

Turning to the loan supply decision by BRAC, we find that efficiency in livestock production significantly reduces the probability of loan rejection. These results are similar for both time-varying and fixed efficiency indicators. This confirms our hypothesis that BRAC uses performance in livestock activity as a signal of creditworthiness. Therefore, monitoring the beneficiary households throughout the intervention period helped loan officer to reject loan application of inefficient households. Among other observable characteristics, we find that the aggregate asset index, household income, and communication easiness of targeted women reduce the probability of loan rejection. Surprisingly, we find that presence of additional male members in the households and women bargaining capacity to initiate a new business increases the probability of loan rejection.

Overall, we notice that both monitoring or performance in livestock production and other observable characteristics are important to explain loan supply decision by BRAC. It presses the importance of combining both individual judgment and credit scoring methods together in the microfinance sector. Previous studies on the feasibility of using only credit scoring methods also conclude in a similar fashion. For example, Vogelgesang (2003) finds that credit scoring model predicts poorly in microfinance loan decision in Bolivia; he finds that model correct prediction rate is only 59% even with a threshold value of 15%. Van Gool et al. (2012) find that performance of automated credit scoring method as an alternative to human-intensive process is very weak due to significant risk associated with microcredit client in Bosnia–Herzegovinian.

7 How much information did BRAC gain through the livestock experiment?

In the previous section, we find that BRAC uses performance in the livestock experiment along with other observable characteristics in loan approval decision. This raises a broader question that could BRAC use only baseline information to predict loan demand and supply outcomes? If baseline information is enough to predict the final credit market outcomes, there is no need of doing the livestock experiment if BRAC's ultimate goal is to lend microcredit to these targeted households. It also intriguing to know whether baseline subjective or objective information has better predictability of credit market outcomes. In this section, we use information from different survey rounds to predict supply decision of microcredit. We estimate prediction as well as variables of importance in prediction using the Random Forest (RF) classifier, a machine learning method, which has an advantage over other tree-based classifiers in reducing over-fitness in prediction. We briefly explain RF method below closely following Zhang and Ma (2012).

Random Forest (RF) is a tree-based classifier that generates a set of decision trees based on a collection of random variables to predict an outcome variable. Let $f(X)$ is a prediction function using a real-valued vector of input $X = (X_1, ...X_p)^T$ on a response variable Y. $f(X)$ minimizes the expected value of the loss function, $E_{XY}(L(Y, f(X)))$, where $L(Y, f(X))$ for the binary indicator of loan approval (S) is as follows,

$$L(Y, f(X)) = I(Y \neq f(X)) = \begin{cases} 0 \text{ if } \; Y = f(X) \\ 1 \;\; \text{otherwise} \end{cases} \tag{8}$$

The minimized expected loss function sets $f(X)$ equal to the maximum value of prediction (arg $\max\limits_{y \in S} P(Y = y|X = x)$) based on a set of "base learners", $h_1(x), ..., h_j(x)$. The jth base learner, alternatively jth tree, is defined as $h_j(X, \Theta_j)$, where $\Theta_j$is a collection of random variables. Therefore, in a classification algorithm $f(X)$ is the most frequently predicted class,$f(X) = \arg \max_{y \in S} \sum_{j=1}^{J} I\ (Y = h_j(x))$. RF Method randomly splits data into training and validation subsets and perform the following steps,

1. Draw a bootstrapped sample of size N with replacement from the train data and fit a tree using binary recursive partitioning as follows,

2. Select m predictors at random out of p available predictors and find the best binary split among all splits by m predictors.

3. For prediction at x, split the node into descendants using criteria until a terminal node is defined using GINI index.

4. Make predictions as $\hat{f}(X) = \arg \max_{y} \sum_{j=1}^{J} I\ (\hat{h}_j(x) = y)$ for jth tree.

RF use observations out of bootstrapped sample in Step 1, called "Out-of-bag data", in two steps to estimate variable of importance: first, for each tree the out-of-bag observations are passed down the tree and predicted values are stored, and second, values of variable k are randomly permuted and predicted values are stored again holding all other variables fixed. The difference between the error rate of the predictions provides the importance of the variable k. Overall variable importance is estimated using an average of all observations.

We randomly split the sample by half into training and validation subsets. We use observable characteristics that BRAC uses in loan approval such as aggregate asset index, land holdings, income, the age of household head, and the number of working-aged members from different survey rounds as the basic information set. We add time (fixed) constant efficiency with follow-up round's basic set to estimate its role in overall prediction. In addition, we use a set of subjective indicators such as community predicted asset level of the household, targeted women ability to initiate self-employment activity, communication, and outside mobility from the baseline survey as an additional set indicator. Results are presented in table 4. Column 1 shows the overall accuracy of prediction using the validation data set. Column 2 shows sensitivity of prediction (loan approval) calculated as true positive/ (true positive + false negative). Column 3 shows specificity of prediction (loan approval) calculated as true negative/ (true negative + false positive). Column 4-6 show 3 most important variable in predicting credit market outcome in descending order.

[Table 4 here]

We find that overall accuracy in prediction stays around 75% with no major differences among different information sets. Sensitivity indicator lies within an interval of 91% to 95%. However, specificity indicator has a relatively large interval from 8% to 30%. From specificity indicator, we find that information in follow-up rounds perform better in predicting credit supply or rejection decisions compared to baseline information. As expected, we also find that adding efficiency indicator with follow-up round information improve prediction of rejection decision. It is also apparent that baseline subjective information does worse in prediction credit rejection decision. Both subjective and objective information from the baseline survey have also some predictability power. We can conclude that follow-up information has better predictability power over baseline information, therefore, the livestock intervention adds additional observable and otherwise unobservable (efficiency) information for BRAC, which help to reduce adverse selection in the credit market.

Column 4-6 list most variables in a descending order in predicting credit outcomes for all information sets. We find that Income, aggregate asset index, and age of the household head are most important variables from the baseline subjective and objective variable list. Same variables also remain as top three predictors when we use information from the follow-up rounds. As expected, we find that efficiency indicator becomes one of the most important variables when we add it to the model. Among the subjective indicator from the baseline survey, we find that women bargaining capacity to initiate new business activity, communication freedom, and mobility index are top predicting variables. From all the results, it is apparent that income, asset index, the age of household head, and efficiency and most important variable in predicting microcredit market outcomes of BRAC. It is important to note that RF method list variable of importance in terms of their prediction power, it does not tell us whether a variable is positively or negatively related to the outcome variables.

8 Does return to capital differ by the selection?

In last two sections, we show that BRAC selectively supplied microcredit to creditworthy borrowers depending on their observable and efficiency indicators. As a lender BRAC's objective is to lend to households who have the prospect of a high return to capital. If results from the last two sections are valid, it will also indicate that households with a high return are more likely to get a loan from BRAC. As the targeted households in this experiment have no productive asset or entrepreneurial activity before the intervention, we cannot examine this hypothesis using pre-intervention information. We instead use information on return to capital only from the first follow-up round to check to minimize the possibility that return to capital is driven by access to credit. We use the following regression model,

$$profit_{it} = \alpha_0 + \alpha_1 \times capital_{it} + \alpha_2 \times rejected_i + \alpha_3 \times capital_{it} \times rejected_i + \lambda X' + \varepsilon_{it} \qquad (9)$$

We consider estimating the return to capital for livestock and other self-employment activities separately. Profit is calculated as household income net of input cost in each sector. Capital reflects value of all physical capital at replacement cost excluding building or infrastructural items. One potential problem with the above specification is that rejection by BRAC can be endogenous. From the demand side, there is no selection bias as both approved and rejected households had the intention to take microcredit from BRAC. On the supply side, if none of these households received a loan before 2009, we could have argued that there is no endogeneity bias; BRAC sanctioned loan based on the return from the pre-microcredit period. However, as some households received a loan before 2009, so access to credit can be multiplied by unobserved ability, which can affect profit. If this is true, there will be a biased estimation. In absence of a credible instrument for rejection, we control inefficiency in livestock production in the model. Controlling inefficiency along with other observable characteristics in the model will serve the goal of getting a robust estimate of the difference in return to capital by loan rejection category (α_3), which will proxy the pre-program differences in return to capital. Access to credit can increase the productivity of household labor in self-employment activities, which can bias estimation results. To tackle this problem, we follow De Mel et al. (2008) and adjust profit by deducting household labor payment in each sector. We use branch level average wage rate from the baseline survey to estimate payment for family labor and deduct it from gross profit.

Results are presented in table 4. Column 1-3 show results for livestock activity and columns 4-6 shows result for other self-employment activities such as non-farm business, nursery, fishery, and small retail shop. We find that rejected households have a significantly lower return to capital in livestock activities compared to loan approved households. The results remain same irrespective of different specifications with controlling efficiency indicator and adjusting profit from family labor hours. We do not find any significant differences in return to capital by loan rejection indicator in non-livestock activities.

[Table 5 here]

9 Conclusion

We study whether BRAC uses performance of beneficiary households in a livestock transfer program as a signal of credibility for its microcredit program. We use information from a livestock transfer experiment in Bangladesh where beneficiary households were also encouraged to take a loan from BRAC at the end of the experiment. We use efficiency in the livestock sector as a proxy for their performance and show that it significantly explains loan approval decision by BRAC. It indicates that BRAC used information from the livestock program to improve targeting in the microcredit program. We also show that BRAC was able to identify households with a high return to capital to reduce adverse selection problem in the microcredit program.

Our study complements earlier studies as well as add new evidence on targeting in the microcredit program. We show that monitoring the performance of households enriches loan officers' subjective judgment, which in turn improves targeting to reduce adverse selection. We show that combining both individual judgment and credit scoring methods together in the microfinance sector is important for better targeting especially in developing countries. Our study also indicates a potential of economies of scope for the MFIs.

References

Agier, I. and Szafarz, A. (2013). Subjectivity in credit allocation to micro-entrepreneurs: Evidence from brazil. *Small Business Economics*, 41(1):263–275.

Ahmed, A. U., Rabbani, M., Sulaiman, M., and Das, N. C. (2009). The impact of asset transfer on livelihoods of the ultra poor in bangladesh. *BRAC, monograph Series*, 39.

Armendáriz de Aghion, B. and Morduch, J. (2000). Microfinance beyond group lending. *Economics of transition*, 8(2):401–420.

Bandiera, O., Burgess, R., Das, N., Gulesci, S., Rasul, I., and Sulaiman, M. (2017). Labor markets and poverty in village economies. *The Quarterly Journal of Economics*, 132(2):811–870.

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienté, W., Shapiro, J., Thuysbaert, B., and Udry, C. (2015a). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science*, 348(6236):1260799.

Banerjee, A., Karlan, D., and Zinman, J. (2015b). Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics*, 7(1):1–21.

BBS (2010). Report of the income & expenditure survey 2010. *Bangladesh Bureau of statistics, Bangladesh*.

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

Bryan, G., Karlan, D., and Zinman, J. (2015). Referrals: Peer screening and enforcement in a consumer credit field experiment. *American Economic Journal: Microeconomics*, 7(3):174–204.

Crawford, G. S., Pavanini, N., and Schivardi, F. (2015). Asymmetric information and imperfect competition in lending markets.

Cull, R. and Morduch, J. (2017). Microfinance and economic development.

De Mel, S., McKenzie, D., and Woodruff, C. (2008). Returns to capital in microenterprises: evidence from a field experiment. *The quarterly journal of Economics*, 123(4):1329–1372.

Emran, M. S., Robano, V., and Smith, S. C. (2014). Assessing the frontiers of ultrapoverty reduction: evidence from challenging the frontiers of poverty reduction/targeting the ultra-poor, an innovative program in bangladesh. *Economic Development and Cultural Change*, 62(2):339–380.

Giné, X., Jakiela, P., Karlan, D., and Morduch, J. (2010). Microfinance games. *American Economic Journal: Applied Economics*, 2(3):60–95.

Hartarska, V. and Parmeter, C. (2009). Scope economies of lending and collecting deposits in microfinance institutions. In *2009 Conference, August 16-22, 2009, Beijing, China*, number 51799. International Association of Agricultural Economists.

Hashemi, S. (2001). Including the poorest: linking microfinance and safety net programs. *CGAP Occasional Paper*, 20.

Hulme, D. and Mosley, P. (1996). *Finance against poverty*, volume 2. Psychology Press.

Joshi, S. (2004). Female household-headship in rural bangladesh: incidence, determinants and impact on children's schooling.

Kabeer, N. (2012). Women's economic empowerment and inclusive growth: labour markets and enterprise development. *International Development Research Centre*, 44(10):1–70.

Khan, M., Ali, M. Y., Quayyum, M., Nazrul, M., and Hossain, M. (2009). Year round homestead vegetable production: a means of reducing poverty and nutritional deficiency for small farm. *Bangladesh Journal of Agricultural Research*, 34(1):169–174.

Machauer, A. and Weber, M. (1998). Bank behavior based on internal credit ratings of borrowers. *Journal of Banking & Finance*, 22(10):1355–1383.

Maitra, P., Mitra, S., Mookherjee, D., Motta, A., and Visaria, S. (2017). Financing smallholder agriculture: An experiment with agent-intermediated microloans in india. *Journal of Development Economics*, 127:306–337.

Rigol, N., Hussam, R., and Roth, B. (2017). Targeting high ability entrepreneurs using community information: Mechanism design in the field.

Roy, S., Ara, J., Das, N., and Quisumbing, A. R. (2015). “flypaper effects” in transfers targeted to women: Evidence from brac’s “targeting the ultra poor” program in bangladesh. *Journal of Development Economics*, 117:1–19.

Samphantharak, K. and Townsend, R. M. (2017). Risk and return in village economies.

Stiglitz, J. E. and Weiss, A. (1981). Credit rationing in markets with imperfect information. *The American economic review*, 71(3):393–410.

Van Gool, J., Verbeke, W., Sercu, P., and Baesens, B. (2012). Credit scoring for microfinance: is it worth it? *International Journal of Finance & Economics*, 17(2):103–123.

Vogelgesang, U. (2003). Microfinance in times of crisis: The effects of competition, rising indebtedness, and economic crisis on repayment behavior. *World Development*, 31(12):2085–2114.

Zeller, M. (1998). Determinants of repayment performance in credit groups: The role of program design, intragroup risk pooling, and social cohesion. *Economic development and cultural change*, 46(3):599–620.

Zhang, C. and Ma, Y. (2012). *Ensemble machine learning: methods and applications*. Springer.

Table 1. Summary statistics by year

	Mean	SD	Mean	SD	Mean	SD
	2007		2009		2011	
Education of respondent (years)	0.60	1.68	0.77	1.85	0.66	1.72
Age of respondent (years)	38.79	13.51	40.83	13.57	42.85	13.57
Gender of head (Male=1)	0.63	0.48	0.64	0.48	0.63	0.48
Education of head (years)	0.63	1.81	0.72	1.91	0.69	1.88
Age of head (years)	44.33	13.91	46.28	13.94	48.34	13.93
Household size	3.41	1.67	3.54	1.71	3.99	1.92
Male working aged member	0.77	0.66	0.85	0.70	0.91	0.75
Female working aged member	1.15	0.49	1.19	0.54	1.25	0.59
Land holdings (decimal)	1.43	8.35	1.69	8.93	1.81	8.52
Value of livestock asset (USD/PPP)	914.91	3249.57	14461.84	12035.54	18124.19	18637.43
Value of non-livestock asset (USD/PPP)	328.86	1198.69	855.34	2398.93	1939.21	8274.58
Number of Cow	0.08	0.38	1.34	0.74	1.34	1.06
Number of Goat	0.15	0.61	0.90	1.49	0.61	1.27
Number of Chicken	1.78	3.30	4.43	5.59	3.30	5.23
Time in livestock activity (Hours/year)	711.40	578.51	1377.73	633.69	1389.48	761.50
Time in non-livestock activity (Hours/year)	1398.72	997.26	1115.89	1020.37	1398.64	1137.01
Enterprenurial ability	0.84	0.27	0.84	0.27	0.84	0.27
Communication	1.84	0.36	1.84	0.36	1.84	0.36
Mobility	0.42	0.59	0.42	0.59	0.42	0.59
Income from livestock (USD/PPP)	26.42	137.29	289.16	561.47	372.10	911.80
Income from other sectors (USD/PPP)	978.09	685.75	1307.33	1050.48	1722.78	1332.66
Total income (USD/PPP)	1004.51	711.12	1596.50	1238.74	2094.87	1706.32

Note: Total sample 11,073; equal in each round.

Table 2. Credit market outcomes

	2009			2011		
	Loan	No loan	Diff.	Loan	No loan	Diff.
BRAC	0.00	8411.94	-8411.94***	0.00	8278.12	-8278.12***
Bank	0.00	3.24	-3.24	0.00	6.31	-6.31
Other NGOs	42.10	35.96	6.14	185.69	206.00	-20.31
Informal lender	29.34	94.52	-65.18	3.28	25.98	-22.70
Shop	16.89	67.76	-50.87***	1.74	38.86	-37.12
Relative	742.13	483.59	258.55*	438.03	711.76	-273.73
Others	29.08	161.86	-132.78	0.00	5.67	-5.67

Note: N=933 in 2009 and N=1,098 in 2011.

Table 3. Microcredit market participation and loan approval

	Rejected	Applied	Rejected	Applied
	(1)	(2)	(3)	(4)
Efficiency score	-0.5821**	-0.1346	-0.5683***	-0.1089
	(0.2416)	(0.2053)	(0.1817)	(0.1578)
Asset Index	-0.1500***	0.0863***	-0.1476***	0.0874***
	(0.0379)	(0.0188)	(0.0379)	(0.0187)
No. of male working aged member	0.1431***	0.1094***	0.1411***	0.1093***
	(0.0429)	(0.0291)	(0.0430)	(0.0291)
No. of female working aged member	-0.0058	-0.0310	-0.0076	-0.0311
	(0.0437)	(0.0324)	(0.0437)	(0.0324)
Age of household head (years)	-0.0012	-0.0017	-0.0012	-0.0017
	(0.0017)	(0.0013)	(0.0017)	(0.0013)
Land holding(Decimal)	0.0012	-0.0021	0.0012	-0.0021
	(0.0024)	(0.0020)	(0.0024)	(0.0020)
Household income(Yearly/USD)	-0.0001***	0.0000	-0.0001***	0.0000
	0.0000	0.0000	0.0000	0.0000
Enterprenurieal bargaining power	0.2465**	-0.1246*	0.2449**	-0.1242*
	(0.1041)	(0.0671)	(0.1037)	(0.0671)
Communication	-0.4312***	-0.1741***	-0.4302***	-0.1735***
	(0.0621)	(0.0462)	(0.0618)	(0.0462)
Mobility	-0.0053	-0.1117***	-0.0027	-0.1115***
	(0.0492)	(0.0337)	(0.0491)	(0.0337)
Year (2011=1)	0.0115	0.1540***	-0.0730	0.1373***
	(0.0527)	(0.0366)	(0.0629)	(0.0419)
Risk-adjusted Income (Positive=1)		0.0855**		0.0796*
		(0.0432)		(0.0426)
Constant	-0.4290**	-0.1945	-0.4271**	-0.2012
	(0.2001)	(0.1495)	(0.1870)	(0.1400)

Note: Robust standard errors in parentheses. ***, **, and * indicate significance at 1 % level, 5 % and 10 % levels, respectively.

Table 4. Microcredit outcome prediction using the Random Forest Classifier

	Accuracy	Sensitivity	Specificity	Variables of Importance		
Basic	0.76	0.94	0.10	Income	Age of Head	Asset Index
Basic + Subjective	0.75	0.93	0.14	Income	Age of Head	Asset Index
First Follow-up	0.74	0.92	0.08	Income	Asset Index	Age of Head
First Follow-up+ Time varying efficiency	0.76	0.95	0.18	Efficiency	Income	Asset Index
Second Follow-up	0.78	0.91	0.30	Income	Asset Index	Age of Head
Second Follow-up+ Time varying efficiency	0.79	0.93	0.27	Income	Efficiency	Asset Index

Note: Baseline objective, first follow-up, and second follow-up include age and education of household head, land holding, livestock asset index, household income, and number of male and female working aged members in households. Subjective indicators consist of women ability to initiate or influence new economic activity (ability), mobility index(mobility), and communication with outside business people (communication).

Table 5. Return to capital and loan approval decision

	Livestock activity			Non-livestock activity		
	(1)	(2)	(3)	(4)	(5)	(6)
Capital (USD/PPP)	0.017	0.007	0.010	0.988***	0.985***	0.206
	(0.024)	(0.023)	(0.020)	(0.220)	(0.221)	(0.167)
Rejected (Yes=1)	45.825	78.235	50.890	-238.461***	-235.805***	-118.122*
	(57.399)	(55.138)	(56.730)	(88.063)	(88.287)	(66.818)
Rejected* Capital	-0.075*	-0.075*	-0.07*	0.274	0.269	0.340
	(0.042)	(0.040)	(0.040)	(0.545)	(0.546)	(0.414)
Male working aged member	100.988***	89.356***	80.97***	202.305***	201.398***	79.107***
	(21.188)	(20.352)	(20.940)	(36.248)	(36.315)	(27.503)
Female working aged member	75.581***	63.952**	55.89**	103.073**	102.174**	41.833
	(27.530)	(26.427)	(27.210)	(47.262)	(47.318)	(35.860)
Age of head (years)	0.721	0.702	0.570	-4.879***	-4.881***	-0.853
	(1.056)	(1.013)	(1.040)	(1.812)	(1.813)	(1.375)
Efficiency score		1,146.878***			88.053	
		(108.744)			(194.770)	
Constant	266.438**	-247.289**	117.380	143.791	103.957	-37.867
	(104.121)	(111.111)	(102.910)	(177.159)	(197.911)	(134.420)

Note: Standard errors in parentheses.***, **, and * indicate significance at 1 % level, 5 % and 10 % levels, respectively. All specification control branch level fixed effect.