The Impact of Population Mobility on Repayment Rates in Microfinance Institutions

Allison Vernerey

Johan Hörnell

Dr. Genna R. Miller, Faculty Advisor

Dr. Kent Kimbrough, Seminar Instructor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University

Duke University

Durham, North Carolina

2013

Acknowledgments

We would like to thank Dr. Genna R. Miller for her enthusiastic support as well as insightful and invaluable comments throughout the process, without which this project would not have been possible. We would also like to express our sincere gratitude to Dr. Kent Kimbrough, our seminar instructor, for his helpful guidance and very constructive criticism. In addition, we are grateful for the help provided by Dr. Marie Laurie (formerly Godquin) in elaborating on the intricacies of her repayment model. We are also indebted to Mr. Shahid Khandker and Mr. Hussain Samad, both with the World Bank for providing detailed information on the datasets used in our research.

Lastly, we want to thank our classmates in Econ 495-496 for their valuable feedback and many words of encouragement throughout the year.

Abstract

Several studies have attempted to model the determinants of repayment rates[1] for group-based loans administered by microfinance institutions (MFIs). One of the main variables that have been identified as playing a role in determining the repayment rate is social capital. Empirical research however has struggled with quantifying this qualitative variable, resulting in vast inconsistencies across studies, aggravating cross-comparison and objective interpretation. Instead, we argue that the use of a quantitative, cross-country comparable proxy that is intuitively linked to social capital would yield more consistent and reliable results. We hypothesize that population mobility is such a proxy, and that lower population mobility correlates positively with higher social capital and thus higher repayment rates. Using population mobility as a proxy for social capital would allow MFIs to lower their cost of data collection for performance assessments and simplify the process for policy makers trying to evaluate the programs' success. At the village level, we find significant evidence that higher emigration within a community is strongly linked to lower repayment rates in microfinance. These results provide microfinance institutions with a new and more cost effective way to monitor their performance as well as to improve their capacity to make well-informed lending decisions.

JEL classification: G; G21; G02; R23;
Keywords: Microfinance Institutions, Population Mobility, Social Capital, Repayment Rates, Bangladesh

---

[1] Or the closely related concept of delinquency rates
[2] Rotating Savings and Credit Associations (ROSCAs) being one example
[3] Of course, it is important to note that in some cases two populations with the same degree of mobility could have different levels of social capital. This could be the case for example if those groups have different religious

1. Introduction

Background

Throughout the world, groups of people are to a varying extent unable to access and make use of today's advanced financial systems. Lacking traditional access to capital, these groups of people have ingeniously developed alternative ways revolving around their own community and neighborhood to pool money and fund their businesses, whether it is by lending money to one another or putting their money together to create the necessary initial financial capital[2]. Community-based, informal financial agreements are hence not a recent phenomena, but it was still not until the latter half of the 20th century that a wider and more structured approach was adopted to leverage these methods to alleviate poverty on a bigger scale. This formalization of the previously scattered, informal and relatively disorganized small-scale finance programs is what has become known as "Microfinance" (Brau and Woller, 2004). Since the institutionalization of microfinance in the 1980s, the sector has shown a tremendous growth, now serving an estimated 665 million people (Christen et al., 2004).

The intuition behind microfinance is that low-income populations often lack the required credit ratings and financial collaterals necessary to access traditional loans. Without fulfillment of these prerequisites, banks are unable to assess the creditworthiness of these individuals resulting in automatic exclusion from the financial system. The microfinance industry on the other hand has been able to overcome this prevalent paradigm whereby poor people are stereotyped as un-bankable, unreliable and unprofitable individuals, to whom it would be an economically wasteful process to lend money. Instead, the microfinance movement has

[2] Rotating Savings and Credit Associations (ROSCAs) being one example

forcefully proved to the traditional finance industry that it is indeed possible to extend the circle of borrowers, and provide cost-effective financial services to this type of client - with impressive and remarkable social and economic effects as a result.

What is Group-Based Lending

It has to be pointed out, however, that microfinance is not a straightforward adaptation of financial traditions. On the contrary, microfinance has developed parallel to commercial finance, with methods, mechanisms and policies of its own. One of the hallmarks of microfinance is the idea of group based, joint liability lending. Different subtleties exist in the way microfinance institutions (MFIs) implement these types of lending programs but in essence joint liability loans is a practice whereby a group of people collectively take out a loan, and are subsequently jointly responsible for repaying the principal along with the associated interest payments. Group lending was developed as a way for MFIs to pool risks, escape asymmetric information and work around the issue of non-existent collaterals and credit assessments (Woolcock, 2001). With a joint liability program, the lender is effectively able to leverage the existence of social capital within the borrowing entity (Gomez and Santor, 2001). Social capital, defined as "features of social organizations, such as trust, norms and networks that can improve the efficiency of society by facilitating coordinated actions" (Putnam, 1993), influences the way in which the peers interact with each other, and is proved to correlate positively with repayment rates in microfinance (Gomez and Santor, 2001). Generally speaking, social capital has two major impacts on the programs' design: firstly, it serves as a form of social collateral, to some extent reducing the need for traditional financial collateral (Woolcock, 2001). Secondly, there is normally the potential for a collective action or free riding problem with group loans, but social capital internalizes a preventive mechanism that is based off of trust and preexisting relationships, thus mitigating this risk (Ghatak, 1999). In addition, joint liability programs feature peer monitoring within the group, with the possibility of social sanctions and ostracizing of members who do not repay on time. This peer monitoring then allows for higher repayment rates and less delinquency, especially since access to future loans is dependent on the successful repayment of current loans (Gomez and Santor, 2001).

Our Extension of Current Research

Whereas scholars largely agree that social capital plays an instrumental role in the effectiveness of repayment rates and MFIs' success, researchers have experienced significant difficulty quantifying the very qualitative, intangible and abstract concept that social capital constitutes. This has often posed a significant problem for scholars. While we agree that social capital is not easily applicable for quantitative methods and analysis, we feel that completely leaving out the variable neglects an essential part of the analysis. We, instead, argue that the best method to go about this is to find a relevant, quantifiable proxy variable for social capital and integrate it into the model. The variable we have decided to use is population mobility. Our intuition behind this choice is that population mobility impacts joint liability lending because the more time you spend with the people in your community, the more it will affect the strength of the relationships and the social ties formed, as well as the trust and capacity of members to peer-pressure one another within the group (Putnam 2000). As such, we believe that population mobility indirectly impacts MFIs repayment rates. Essentially, population mobility measures how mobile a population is, and how likely members of a specific group are to move. As such, this variable meets diverse criteria that we judge essential to qualify as a good proxy: it is fairly easily accessible in terms of data availability; it is quantifiable; it can be found and used for different populations, cultures and countries thus allowing for cross comparisons; and finally, it is intuitively related to social capital and thus to the repayment rate. This is important because recently MFIs, donors and third-party organization have become increasingly interested in performance assessment, which is greatly simplified if data becomes more easily available and allows for cross-country comparison. Additionally, MFIs – being financial institutions – need to be able to predict repayment rates in order to make well-informed lending decisions. Thus, while repayment rate models have not traditionally included mobility as a determinative variable, we argue that it is a suitable proxy for social capital[3]. Our hypothesis then states that group-lending based MFIs that are located in places with higher population mobility, all other factors influencing repayment being equal, experience lower repayment rates due to a weaker buildup of social capital within the population they serve.

The repayment rate model we have chosen to work off of and extend is one that was originally developed by Godquin (2004). The reason we chose this model is partially because of data availability, but also because of the comprehensiveness of the model, and the number of factors it considers. To maximize comparability, we used the same dataset as Godquin did but added population mobility as a variable to know how it would affect the fit of the model. Godquin already used a variable that was somewhat linked to social capital, 'age of the group' (representing how much time has passed since the group formed), so to make sure we would capture the full effect of population mobility, we performed econometric analysis on the model both with and without her 'age of the group' variable. Mathematically speaking, her model is constructed in four steps. First, she estimates the size of the loan contingent on the incentive structure of group lending. Secondly, she performs an exogeneity test for loan principal in the determination of the repayment performance and concludes that endogeneity of the loan size

[3] Of course, it is important to note that in some cases two populations with the same degree of mobility could have different levels of social capital. This could be the case for example if those groups have different religious affiliations that might affect the way people connect with each other and build social capital

cannot be rejected and hence instruments for the size of the loan. The third step is the actual prediction of the repayment rates whereas the fourth step represents comparison of the observed figures. Our changes are incorporated in step three where we add our own population mobility variables. The actual construction of the population mobility variables, their justification, as well as other econometric details is described in section 5.[4]

Structurally, we will start with reviewing previous theoretical material in our Literature Review (Section 2). Section 3 focuses on describing the theoretical framework of our research. From there, we will go on with elaborating on the datasets that we are using for our analysis. This leads us into the Empirical Specification (Section 5), in which expand on our theoretical framework in a mathematical way, adding in our novel population mobility variables and developing the extended model. Lastly, Section 6 and 7 focus on the results of our regressions and the subsequent discussion and interpretation, coupled with policy implications and a clear-cut answer to our initial hypothesis.

2. Literature Review

Objectively Assessing MFI Performance Through Repayment Rate Models

As microfinance has grown over the years, many researchers have tried to figure out what factors are responsible for the success or failure of an MFI. The fact that MFIs inherently feature a double bottom line with both social impact and financial sustainability being equally important complicates the performance evaluation. There are also many potential ways to think about this question; for example by, as previously mentioned, looking at the sustainability of the MFI, the

[4] For a detailed explanation of Godquin's methods, please see her 2004 paper 'Microfinance Repayment Performance in Bangladesh: How to Improve the Allocation of Loans by MFIs', pages 1917 – 1919.

types of businesses created by its micro-borrowers, how profitable they become or the institution's loan repayment rates. To clarify, repayment rate, or the closely related concept of delinquency, in the context of microfinance, is a measure that approximates the extent to which borrowers are able and willing to, in full, repay the principal along with the interest payments. Primarily because of data availability and traceability, the most commonly used approach has been to look at the repayment rates, and by conducting field studies trying to extrapolate the dependent variables. Naturally, there exist many factors greatly impacting repayment rate. One of them can be the interest rate. A higher interest rate essentially means that each business has to be increasingly successful, only to meet their debt inferred obligations. Other factors that are believed to influence repayment rate are varied and include variables such as gender of the borrower, business expertise, amount of loan borrowed, MFI specific characteristics, poverty indicator, education and many others (Oke et al, 2007; Godquin 2004; D'Espallier et al, 2009). Different models have been created and tested to assess the impact of those different variables on repayment rate and their significance (Oke et. al, 2007; Godquin 2004; D'Espallier et al, 2009).

One of the most commonly referenced examples is Oke et al. (2007), who conducted a field experiment in Southwestern Nigeria to test what variables are significant in accounting for repayment rates in microfinance and trying to establish causality between those variables and repayment rates. The study found that ten of their twenty-three initially included variables played a significant role in explaining rates of loan repayments. Those ten variables were: income, distance between dwelling place and bank, amount of business investment, socio-cultural expenses, amount of loan borrowed, access to business information, penalty for lateness to group meetings, membership of cooperative society, number of days between loan application and disbursement and poverty indicators. The study includes an indirect measure of social capital in the socio-cultural expenses and the membership of cooperative society variables. Even though the model was a good fit, the adjusted R^2 of 0.36 implies that there is still a decent amount of variability in the repayment rate that the model was unable to explain.

There have been numerous attempts trying to find other significantly influencing factors, and one of the key aspects studied has been gender differentiation in loan repayment and creditworthiness. It has been assumed for a long time that women were "better" borrowers and more likely to repay their loan. For example, D'Espallier et al. (2009) rigorously analyze the veracity of this idea and use a large global dataset covering 350 MFIs in 70 countries to test whether a gender effect on microfinance repayment rates exists. Using sub indicators such as portfolio-at-risk, loan-loss write-offs and provisions to measure repayment rates, they conclude that their "findings provide compelling evidence that focus on women clients enhances microfinance repayment, and that women in general are a better credit risk" (D'Espallier et al., 2009).

Social Capital

In addition, it has been argued that another crucial component in determining repayment rates is social capital. Whereas plenty of scholars have been able to identify the importance of social collaterals, there still exists an ongoing effort to more closely define this concept and quantify this measure into something economically sensible and meaningful. Many authors talk about social capital and how it is a way to measure the strength of social collaterals (Krishna 2000; Dudwick et Al. 2006). Social capital has many interrelating dimensions but can potentially be summarized to the main idea, expressed by Putnam, that "networks and the associated norms of reciprocity have value. They have value for the people who are in them, and they have, at least in some instances, demonstrable externalities, so that there are both private and public faces of social capital" (Putnam, 2003). Problematically enough, however, most measures of social capital rely on qualitative data, with the measurement only being valid in very close proximity to where it was developed. An example is Dr. Krishna's study of 69 villages in northern India, which found that one of the most significant variables affecting social capital is whether the villagers collectively participate in crop disease prevention or not (Krishna, 2000). This variable, for obvious reasons, is not easily applicable to 'villagers' in for example New York City. Thus, social capital is extremely contextual and may be dependent on region, culture, etc. However, we argue that population mobility does not suffer from the same contextual limitations and is inherently quantitative.

In one of the relatively rare papers actually quantitatively looking at the impact of social capital on group-lending repayment rates in microfinance, Karlan (2007) tries to establish a causal relationship between people being more connected socially (one major aspect of social capital) and better performances through the lens of contract monitoring and/or enforcement using data from FINCA-Peru, a group-lending organization. He uses different measures of geographic distances (between one member to original members of the group or from current member to members of other groups) as a proxy for social connections. Regarding what we are interested in, he "finds evidence to support one hypothesis behind group lending: that monitoring and enforcement activities do improve group lending outcomes, and that social connections, broadly defined, facilitate the monitoring and enforcement of joint liability loan contracts. Social connections might have this effect simply through lowering the cost of gathering information about each other (i.e., a monitoring story), or through a social capital story in which more connected individuals trust each other more and value each other's relationships more" (Karlan, 2007).

As briefly discussed before, looking at creating a very complete model for repayment rate in Bangladesh, Godquin's paper (Godquin, 2004) uses variables such as loan size[5], age of the group at the time the loan was due (as a proxy for intragroup social ties, and thus somewhat accounting for social capital), homogeneity in education and age within groups, how much nonfinancial services the MFI provided the group with and whether borrowers are credit-rationed or not. In addition, the model accounts for exogenous control variables including characteristics of the borrower and his/her household as well as basic information on the loan. Thus this model looks at quantitative variables taking into account many factors, including a proxy for intragroup social capital. Godquin, among other things, finds that the size of the loan shows a negative sign and is significant in all the regressions that she runs. Regarding social capital, she finds that variables used to control for the wealth of the household and wealth of its social network show a positive and significant impact on repayment rate while social ties inside the group (proxied by the age of the group) surprisingly show a negative and significant impact on repayment rate (Godquin, 2004). Given these unexpected findings, social capital may not be adequately measured by these variables, which strongly supports our claim that there may be a better way to measure social capital's effect on the repayment rate.

Social Capital and Population Mobility

As mentioned in the previous section, to better capture the effects of social capital on repayment rate we use population mobility as a proxy for social capital. While population mobility has never directly been used as a proxy for social capital in research, there is plenty of literature confirming the primary intuition that increased mobility leads to lower social capital

[5] Instrumented by the value of previous loan size to work with the endogeneity of the variable

and vice versa. In addition and as aforementioned, population mobility benefits from its ease of use and cost effectiveness as a variable. One of the most famously cited works regarding social capital is that of Robert Putnam who looks at how American society has become recently more and more disconnected from family, neighbors and social structures and has thus lost a lot of its social capital. Putnam tries to identify the causes and effects of this phenomenon. In his chapter devoted to the relationship between social capital and mobility Putnam explicitly states that "mobility undermines civic engagement and community-based social capital." (Putnam, 2000)

Looking at more specific examples, a number of scholars have based full arguments on the assumption that social capital falls with mobility and specifically labor mobility. A good example is that of Schiff (2004), looking at labor and goods market integration in a general-equilibrium model with social capital. Schiff develops his model based on the two assumptions that higher social capital is associated with higher productivity and that social capital decreases as labor mobility increases. To support these assumptions, he relies on the work of Zak and Knack (2001) who show that more mobility has a weakening effect on social ties and that higher transaction costs will result from transactions among less familiar agents. (Schiff, 2004)

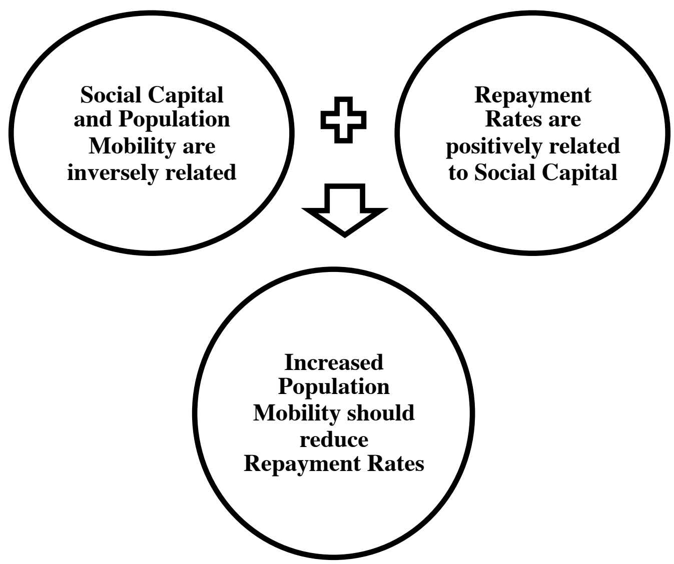
In the same way, in their research on "Social Capital and Growth", Routledge and Von Amsberg (2002) construct a growth model where individuals in a community maximize their lifetime gains to trade, with each trade structured through the framework of the well-known prisoner's dilemma. Using this model as their base, they explore the relations between growth, labor mobility, and social capital. In their analysis, technological change is associated with growth and an increase in labor mobility due to higher labor turnover. While this increased labor mobility leads to a more efficient labor allocation, it also has negative impacts as it decreases the proportion of cooperative trades which define social capital in this model. Thus, while it is not the main goal of the research, this paper shows that increasing mobility is responsible for lower levels of social capital in communities. (Routledge and von Amsberg, 2002) Additionally, looking at rates of mobility from the angle of homeownership versus renters and using data from the US Social Survey, DiPasquale and Glaeser (1999) find that homeownership, which creates high barriers to mobility, is strongly correlated with variables associated with social capital. While their results don't allow them to make any causal statement they do show that homeowners, who have lower mobility, have a tendency to invest more in social capital than renters (DiPasquale and Glaeser, 1999).

It is important to note that while some of these papers focus on the more specific case of labor mobility, in our case, the use of this variable would not make as much sense. This is because labor mobility is defined as a population's propensity to change employment, not necessarily geographic location. While there may be an overlap between labor mobility and the general, wider population mobility, the overlap is not complete and the two measures are only comparable to a certain extent. Because social ties need not to be severed primarily from someone changing their job, the measure is inadequate to capture the social capital aspect we are looking for in our proxy variable. In addition, the majority of clients of MFIs are women who in developing countries like Bangladesh are not very likely to experience labor mobility but are still susceptible to relocating for other reasons such as marriage, or following their husbands and families when they relocate. This reinforces how the measurement of social capital is highly contextual, and how the proxy for it needs to be carefully defined. For these reasons, we look at mobility in a broader sense and construct an index of population mobility rather than limiting ourselves to labor mobility.

3. Theoretical Framework

The theoretical framework of our research is based on three different pillars from economics and sociology: we use the fact that repayment rates in joint liability lending programs show a positive correlation with social capital; the idea that population mobility and social capital are tied together through a negative relationship; and finally we connect the first two points to deduce that population mobility will have an impact on repayment rates through social capital. Figure 1 below illustrates this concept.

Figure 1: The Three Pillars



Regarding the first pillar, we are leveraging the fact that the repayment rate for a joint liability lending program is positively correlated with social capital (Karlan, 2007). Rationing credit along with collateral requirements are the traditional methods commercial banks use to overcome the inherent asymmetric information that exists in the banking industry. This methodology however, while effectively regulating and stabilizing the credit market for everyone who can provide the required collateral, excludes the poor and deprives them of their access to capital. It is empirically proven however that microfinance, which focuses entirely on these 'un-bankable' groups of people, is a successful form of providing credit. To describe this, most of the sources cited in the literature review rely on the principal / agent theory to explain how microfinance contracts create a collective liability that mitigates the moral hazard and adverse selection that arises from the information asymmetries. In addition, it is often argued that joint liability lending increases repayment rate by making strategic defaults more costly. Social capital also results in greater trust so that problems of free-riding are largely diminished. Lastly, social ties and the group's homogeneity affect the repayment rates indirectly through the way they facilitate peer pressure and peer monitoring within the group. Thus, when social capital builds up, peer pressure and peer monitoring become increasingly more powerful methods of controlling the group's behavior and hence, all other factors constant, as social capital builds up and increases, so does the repayment rate.

The second pillar revolves around the intuition behind how social capital itself builds up. While social capital is inherently complex, there exist some fundamental and simple prerequisites that have to be in place in order for it to build up. One of these fundamental conditions is that people are physically close enough to each other to actually interact. Logically, one would assume that the more time individuals spend with each other, the higher the chance that they form stronger social ties and relationships. Intuitively, there hence exists a theoretical connection between how likely individuals in a population are to leave their current village and friend group and how large their stock of social capital is. Put differently, the mobility of a population and the magnitude of the social capital within the population and the group are negatively related. This is empirically and scholarly supported by Zak and Knack, 2001; DiPasquale and Glaeser, 1999 and Routledge and von Amsberg, 2002.

From these two pillars, we infer a third one. If repayment rates for joint liability loans are positively correlated with the amount of social capital the group possesses, and the social capital in turn is contingent on the level of mobility, then this implies that repayment rates are indirectly correlated with population mobility. To test this, we run an expanded version of the following regression:

$$Repayment\ Rate = \alpha_1 + \beta_1 x + \beta_2 (population\ mobility) + \varepsilon$$

where x represents all of the other variables that affect the loan repayment rate. Our hypothesis that comes from this regression is mathematically represented as: $H_0 : \beta_2 < 0$

If our results conclude that the population mobility variable is significant, then the next step is to test the fact that adding this variable as a proxy for social capital improves the adjusted R^2 of the repayment rate model we are using. This is formally done by running a Wald test, comparing the different regressions and testing for if the adjusted pseudo R^2 significantly improves in the regressions where we add our mobility variables. This sums up our theoretical framework. Details on the model we use, as well as the construction of the measure of social capital are specified in the Empirical Specification section (section 5).

4. Data

The data in our research comes from two major sources. Firstly, we use the main data set on MFIs' performance from Bangladesh available through the World Bank used in Godquin's paper (Godquin, 2004). Our initial step here was to use this to reconstruct the repayment model with which Godquin was working. The data stems from a quasi-experimental survey on microfinance in Bangladesh in 1991 and 1992 carried out by the World Bank and later deposited into the public domain. Originally, the survey was designed to assess the efficacy of microfinance as a means of alleviating poverty. The survey covered a total of 1798 households from 87 villages in 29 different thanas. A thana is an administrative division in Bangladesh, comparable to a county in the US[6]. Out of these observations, we are specifically interested in the 905 households that fell below the poverty line and were hence eligible for microfinance loans and also made the active decision to participate in one of the microfinance programs offered. These 905 households received a total of 2073 loans, distributed between three major types of credit providers: Bangladesh Rural Development Board (BRDB), BRAC (formerly Bangladesh Rural Advancement Committee) and the Grameen Bank. Few differences exist in terms of characteristics of the loans provided or types of borrowers. All three MFI's initially had a uniform fixed interest rate of 16% per annum with BRAC and the Grameen Bank increasing their interest rate on loans to 20% per annum after 1991. The distribution of loans across these three MFIs is presented in Table 1 below.

In addition to reconstructing Godquin's model, we also needed data to be able to construct our own population mobility variables that serve as the proxy for social capital. The necessary migration data for these variables is based on two sources.

TABLE 1
Loan Origin

Loan Origin	Frequency	Obtained
Bangladesh Rural Development Board	479	23%
BRAC	441	21%
Grameen Bank	1153	56%
Total	**2073**	**100%**

Firstly, the previously discussed MFI study from the World Bank includes a second survey round carried out in 1998 and 1999, where the participating households were asked to fill

[6] The word thana is best translated into 'subdistrict'. There are 1009 thanas in Bangladesh, A zila is the primary administrative division in Bangladesh, and represents a district. There are 64 zilas in Bangladesh.

out a migration roster, specifying which of their former household members had left and the reason for their departure. These migration rosters were used to construct alternative measures of population mobility (described in more details in the Empirical Specification, section 5).

In order to construct yet another population mobility variable, we used census data from the two previous censuses, namely 1991 and 2001. The census data collection and distribution is administered by the Bangladesh Bureau of Statistics. In addition, the Bangladesh Bureau of Statistics also provided us with point statistics, such as mortality and fertility rates.

5. Empirical Specification

As previously mentioned, our model is a modification of the one used in Godquin's paper. The extensions come in the form of our proxy for social capital; the population mobility variables. This section starts out with a broad description of the repayment rate model, after which we carefully explain how it is set up and how it works on a micro level before concluding with explaining what results we expected to get from our regressions before actually running them, based on our initial theory. After explaining Godquin's original model, we go on to firstly describe our population mobility variables, and then discuss how we added them to her regressions.

The Baseline Empirical Model

Before going into the mathematical description of our model, it is important to lay down the process that takes places when a group applies for a loan in the Bengali context.

Step 1: The group submits an application for a loan of a given size, which correlates to the largest size the group can expect, given the projects they are looking at undertaking, and defined by characteristics of the group, its individual members and the surrounding environment.

Step 2: Once the application is received by the MFI, their staff evaluates it in terms of default probability using the information they have available on the group. If the MFI concludes that the risk of default meets the internal standards the MFI has set up, the application is accepted and the loan disbursed.

Step 3: The group receives the money and distributes it to its members, based on internal policies of the group, previously agreed upon by the MFI. The individual members who receive a part of the loan are responsible for repaying their share along with the associated interest. Should one party however fail to repay, the group as a whole becomes liable for the unpaid debt. Until repayment in full has been made, no further loans will be extended to the group or any of its members.

The breakdown of the process into three steps is meant to give an intuitive overview of the possibility of endogeneity between the independent variables and the repayment rate. More specifically, the determination of the loan size from the MFI's perspective is most likely dependent on the same omitted variables as the group's ability to repay the loan. Based on this, Godquin, in her construction of the model, decided to use an individual dummy variable for on-time repayment and a probit model to estimate the probability for a borrower to repay the loan on time. Testing for endogeneity, Godquin could not reject the possibility of endogeneity and thus instrumented the loan size. All in all, this leads her to the following estimation strategy:

Step 1: Estimating the Loan Size

$$(1)\ P_i = \hat{P} + \varepsilon_i^p = \alpha^p + \sum \beta_j^p X_{ij} + \sum \rho_j^p Y_{ij} + \lambda^p Z_i + \sum \sigma_j^p W_{ij} + \gamma^p IVp + \varepsilon_i^p$$

P_i *and* $\hat{P}$ respectively represent the actual and the predicted size of the loans disbursed to the individuals. The variable X_j aggregates the different measures of the incentive structure of group lending, social ties and group homogeneity. The *i* index denotes the different observations, that is the different groups to which a loan has been extended. The *j* index on the other hand corresponds to the different variables that are clustered together under the incentive structure of group lending. Godquin used the age of the group at the time the loan matured for intragroup social ties and assumed that as the age of the borrowing group increases (AGEGP), social capital in the form of more deeply rooted knowledge of the other members in the group and greater levels of intragroup dependency is likely to build up[7]. Based on this, the intuition is that the variable AGEGP should be positively correlated with repayment rates (Although this turned out not to be true in the regressions she ran). Group homogeneity is approximated using shared characteristics (i.e. age and education level) within the group. As discussed in the theoretical framework, group homogeneity (SAMEEDU, SAMEAGE) should have a positive impact on the level of social capital with positive repayment rate correlation as a result. Y_j depicts the nonfinancial services offered. For the non-financial services offered, Godquin uses access to basic literacy (NFSL) and to primary health (NFSH) as the independent covariates. Just as with X_j, these services are expected to positively impact the repayment performance. Z_i measures the dynamic incentives the MFI makes use of. Godquin proxies this with credit rationing, more specifically how disbursement of new and larger loans is conditional on previous on-time repayment. Credit-rationing, (CRD) is expected to have a positive impact on repayment since

[7] See Table 9 for a full list of the variables in the model

borrowers who are being credit rationed will be incentivized to a higher degree than those who are not. W_j takes into account the various exogenous control variables. The control variables encompass characteristics of the borrowers and their households as well as fundamental details of the loan. The final variable is the instrument variable, IVp, used for the loan size. Godquin uses the size of the previous loan as an instrument for the current loan size. The argument behind this is that size of the previous loan should not affect the repayment of the present loan while it was, at the same time, determined on the basis of the same unobservable and omitted variables as the present loan is.

Step 2: Test for Exogeneity

The second stage in Godquin's approach is to test for whether the loan size is exogenous to the repayment performance. She does this using Smith and Blundell's (1986) exogeneity test. The test methodology states that exogeneity is rejected if the instrumented regression of the loan size has a statistically significant coefficient for the error term (η). Mathematically, this implies that the actual structure of the error term would be $\varepsilon_i = \chi\varepsilon_i^p + \mu$

$$(2) \qquad R_i = \alpha + \omega\widehat{P}_i + \sum\beta_j \; X_{ij} + \sum\rho_j \; Y_{ij} + \lambda \; Z_i + \sum\sigma_j \; W_{ij} + \eta\varepsilon_i^p + \varepsilon_i$$

The equation above is the full, instrumented regression that Godquin uses to assess the appropriateness of the instrument variable and the possible endogeneity. R_i is the model's latent variable and measures the group's capacity to generate cash in excess of the principal plus interests that they have to repay on or before the due date. The observed variable is R, which takes on the value 1 if $R_i > 0$ and 0 if $R_i < 0$. Running the instrumented regression, Godquin found that the error term's coefficient η was significant and endogeneity could hence not be rejected while the choice of the IV_p for the size of the loan was appropriate.

Step 3: Estimating the Repayment Rate

In the third step, Godquin puts it all together in her final repayment rate equation:

$$(3) \qquad R_i = \alpha + \omega\widehat{P_i} + \sum \beta_j \; X_{ij} + \sum \rho_j \; Y_{ij} + \lambda \; Z_i + \sum \sigma_j \; W_{ij} + \varepsilon_i$$

The Population Mobility Variables

As stated in our theoretical framework, we extend Godquin's research with variables for population mobility meant to proxy the social capital existing within a spatial region. Unfortunately, there is no set standard for how to measure mobility and scholars use different methods depending on the context and the data available. In our research, we are looking to find a measure that is as easily constructed as possible, and where data is as easily collected as possible – ideally without compromising the proxy's validity. For this reason, we develop two different variables that we separately introduce to the regression. This method gave us more leeway and flexibility in terms of finding the optimal proxy.

Net Migration Rate (NMR)

The first variable we designed is constructed using one of the few standard methodologies currently available. This measurement was developed using guidelines from UN's Department of Economic and Social Affairs' Manual on Methods of Measuring Internal Migration. The manual suggests using census population data from two different censuses combined with birth and mortality rates to extrapolate a number for the internal migration. Table 2 below summarizes the fertility and mortality rates used to calculate the NMR.

TABLE 2
MORTALITY & FERTILITY RATES

Variable	Value
Fertility Rate*	22.52
Mortality Rate**	5.71

* Per 1000 population per year
** Per 1000 population per year

Source: Bangladesh Bureau of Statistics

The first full equation looks as follows:

$$(4)\ Net\ M' = (POP_t - POP_{t-1}) - \left(Birth\ Rate * \frac{(POP_t + POP_{t-1})}{2} - Mortality\ Rate * \frac{(POP_t + POP_{t-1})}{2}\right),$$

Where POP_t is the population at time t and *Net M'* is the Net Internal Migration between time t and t-1. From this equation, we derive the Net Migration Rate variable by dividing it with the average population in the thana over the census period.

$$(5) \qquad NMR = \frac{Net\ M'}{\left(\frac{POP_t + POP_{t-1}}{2}\right)}$$

This variable is denoted as NMR in the regressions[8]. Table 3 below shows the summary statistics for the Net Migration Rate (NMR) for the thanas used in Godquin's dataset:

TABLE 3
NMR SUMMARY STATISTICS

Summary Statistic	Value
Mean	11%
Median	11%
Standard deviation	5%

Though cross-country comparison is hard due to the number of different measures out there, it does indeed come across as high to have a median internal migration of 11% of the population in a thana. As a quick comparison, the average inter-state migration in the United States between

[8] One of the main limitations of the variable however is that, unfortunately, Bangladesh lacks an accurate citizen registration system, meaning there exist some uncertainty associated with the population measurements in between the different census.

2001 and 2010 is less than 2% (Molloy et al, 2011). The thanas also feature a sizeable variability of 5%. However, based on the closeness of the mean and the median, the dataset does not seem to contain any real outliers that would cause the mean and the median to diverge. A closer look at the full table in Table 4 confirms this.

The Expected Rate of Emigration (ERE)

As our first variable shows, all of the thanas measured have seen a steady inflow of people over the course of the two censuses. Because we are looking at how population mobility will affect repayment rates in joint liability lending, we are focusing our investigative efforts solely on looking at outflows of people and how they impact repayment rates. The logic behind this principle is that people who have emigrated were part of the stock of social capital that the community, the microfinance groups within the community, as well as the individuals possessed. Inferred from this is the notion that as somebody leaves a community, they will have an immediate impact on the level of social capital within that community, and subsequently the repayment rates. In contrast, one would not expect immigrants to disrupt existing social ties.

However, emigration would break down and disrupt group ties that, presumably, had been built up and fostered over years. Therefore, we expect emigration to be negatively related to the amount of social capital. Since our theory is more focused around how social capital is depleted when somebody emigrates – rather than the possible build up that may or may not take place when somebody moves into a thana, we develop a variable that simply measures the probability of somebody emigrating. This is exactly what our second mobility variable is meant to capture. In other words, the Expected Rate of Emigration (ERE) measures the expected value that someone from a specific geographical location will emigrate.

TABLE 4

NET MIGRATION RATE (NMR)

Thana Id	Thana	NMR
1	DHAMRAI	9%
2	GAIBANDHA SADAR	7%
3	HOBIGONJ SADAR	18%
4	KOLAROA	13%
5	MANIKGONJ SADAR	8%
6	NARSINGDI SADAR	23%
7	RANGPUR SADAR	18%
8	SRIBORDI	4%
9	BIRGANJ	13%
10	DUMURIA	7%
11	FAKIRHAT	6%
12	KARIMGANJ	7%
13	MUKTAGACHA	11%
14	MATHBARIA	2%
15	SATKHIRA SADAR	16%
16	SHIBGANG	10%
17	ZAKIGONJ	12%
18	JALDHAKA	15%
19	PATUAKHALI SADAR	3%
20	PIRGONJ	12%
21	ROYGONJ	16%
22	SAKHIPUR	8%
23	SONARGAON	14%
24	SREEPUR	3%
25	BRAHMANPARA	10%
26	DAULATPUR	10%
27	FARIDPUR SADAR	19%
28	JHENAIDAH SADAR	15%
29	MANDA	5%
30	BEGUMGONJ	11%
31	MIRSARAI	11%
32	SATKANIA	10%

We arrive at the equation by adding up, year by year, everyone who has moved from one geographical unit to another. We then simply divide this number first by the number of people surveyed and then by the number of years between the two survey rounds. Mathematically:

$$(6) \qquad ERE = E(M) = \frac{\Sigma_{t=0}^{7}\left(\frac{\#\ of\ emigrants\ from\ households\ surveyed}{\#\ of\ people\ surveyed}\right)_t}{total\ number\ of\ years\ (7)}$$

The data for this variable comes from the follow-up survey round that was part of the original MFI study. In this round, taking place in 1998, the same households were asked to specify which household members had left them between the two rounds, what year they left and what the reasons were. For increased granularity, we constructed two versions of this measure. Firstly, we looked at the expected emigration on a thana level and secondly on the village level. The thana version obviously contains more households, people and emigrants per thana but as the tables below show, the summary statistics for the two measurements are relatively similar:

TABLE 5
ERE THANA SUMMARY STATISTICS

Summary Statistic	Value
Mean	3.00%
Median	3.10%
Standard deviation	0.47%

TABLE 6
ERE VILLAGE SUMMARY STATISTICS

Summary Statistic	Value
Mean	2.98%
Median	2.97%
Standard deviation	0.58%

As we can see, the means, medians, and standard deviations are substantially lower for the ERE than they are for the NRM (11%, 11% and 5% respectively). This could be the result of first and foremost the fact that we are now only looking at emigration, giving rise to lower degrees of variability than the net migration would. On top of that, this is a direct measure where we do not have to rely on statistics such as birth and mortality rates to estimate the migration levels, but instead we have a direct account of the degree of outbound mobility. Table 7 below shows ERE values for the 32 thanas included in the study. Under that, Table 8 shows the same thing but at the village level.

Our Extended Model

The final step is to put together Godquin's model and our population mobility variables to see if we are able to better account for the impact of social capital on repayment rates. As Godquin used the "age of the group" (AGEGP) variable to proxy social ties, we ran regressions with each of our variables (NRM, ERE VILLAGE and ERE THANA) separately in two different ways. For simplicity, the variables for mobility are all denoted as [MOB] in the next two equations. First, as shown in equation 7, we simply added [MOB] as a variable in the original probit model while keeping the AGEGP variable as part of the regression and thus leaving the incentive structure denoted by X_{ij} in the equation intact.

$$(7) \qquad R_i = \alpha + \omega\widehat{P_i} + \omega^m[MOB] + \sum\beta_j \ X_{ij} + \sum\rho_j \ Y_{ij} + \lambda \ Z_i + \sum\sigma_j \ W_{ij} + \varepsilon_i$$

Secondly, we added our [MOB] variable and changed the incentive structure by taking out the AGEGP variable to see if the variables were actually picking up similar effects. In all essence, X_{ij} is now comprised of the incentive structure, minus the age group and is noted as $X_{ij}^{new} = X_{ij} - AGEGP$ in equation 8 below.

$$(8) \qquad R_i = \alpha + \omega\widehat{P_i} + \omega^m[MOB] + \sum\beta_j \ X_{ij}^{new} + \sum\rho_j \ Y_{ij} + \lambda \ Z_i + \sum\sigma_j \ W_{ij} + \varepsilon_i$$

Before running the six regressions, our expectation in terms of the *[MOB]* coefficient was that a higher degree of mobility would have a negative impact on the repayment rate.

TABLE 7
EXPECTED RATE OF EMIGRATION (ERE THANA)

Thana Id	Thana	ERE THANA
1	DHAMRAI	3%
2	GAIBANDHA SADAR	3%
3	HOBIGONJ SADAR	2%
4	KOLAROA	3%
5	MANIKGONJ SADAR	3%
6	NARSINGDI SADAR	3%
7	RANGPUR SADAR	3%
8	SRIBORDI	3%
9	BIRGANJ	3%
10	DUMURIA	3%
11	FAKIRHAT	3%
12	KARIMGANJ	4%
13	MUKTAGACHA	3%
14	MATHBARIA	3%
15	SATKHIRA SADAR	3%
16	SHIBGANG	4%
17	ZAKIGONJ	2%
18	JALDHAKA	3%
19	PATUAKHALI SADAR	3%
20	PIRGONJ	4%
21	ROYGONJ	4%
22	SAKHIPUR	4%
23	SONARGAON	3%
24	SREEPUR	2%
25	BRAHMANPARA	2%
26	DAULATPUR	3%
27	FARIDPUR SADAR	3%
28	JHENAIDAH SADAR	3%
29	MANDA	3%
30	BEGUMGONJ	2%
31	MIRSARAI	2%
32	SATKANIA	2%

TABLE 8
EXPECTED RATE OF EMIGRATION (ERE VILLAGE)

Village Id	Village Name	ERE VILLAGE	Village Id	Village Name	ERE VILLAGE
1	Agzathail	3%	45	Srirampur	3%
2	Boro Bakulia	3%	46	Sultanpur	2%
3	Rambhadrapara	4%	47	Chandpara	4%
4	Tengarjani	3%	48	Bat Tala	3%
5	Barmatat	2%	49	Sonashar	4%
6	Kumarervita	4%	50	Pathanchak	4%
7	Shayestanagar	3%	51	Gordishpur	3%
8	Mohonpur	2%	52	Aaraji Shimulbari	3%
9	Purba Bahula	3%	53	Khanapara	2%
10	Kauria	2%	54	Cheoradangi	3%
11	Dakshin Bahura	4%	55	Haqtullah	2%
12	Jhikra (Uttar)	3%	56	Lebukhali	2%
13	Boro Burundi	4%	57	Kuripaika	3%
14	Sadinagar	3%	58	Prathamdanga	3%
15	Bhikora	3%	59	Hasanpur	3%
16	Kandapara	2%	60	Kataduar	3%
17	Khalpar	3%	61	Dumrai	2%
18	Algi Tanpara	3%	62	Sengati	3%
19	Amashu	3%	63	Sonakhara	2%
20	Sardarpara	3%	64	Borochowna	4%
21	Mudikhana	3%	65	Chambaltala	4%
22	Jhalupara	3%	66	Kachua	3%
23	Khamariapara	3%	67	Bari Majlish	3%
24	Tarakandi	3%	68	Kafurdi	4%
25	Chak Mohonpur	3%	69	Noyapur	4%
26	Bhognagar	4%	70	Bekashara	3%
27	Krishnanagar	3%	71	Kathali	4%
28	Khorsonda	3%	72	Vitipara	3%
29	Rajapur	3%	73	Ashabari	3%
30	Araji Dumuria	3%	74	Chhoto Dhusia	3%
31	Town Nowapara	4%	75	Chandipur	3%
32	Khajura	3%	76	Begunbaria	3%
33	Kachua	3%	77	Philipnager	2%
34	Atkapara	3%	78	Purba Bajumara	3%
35	Palaikanda	3%	79	Brahmankanda	2%
36	Tamnee	3%	80	Raghunandanpur	3%

	Islampur				
37	Krishnabari	4%	81	Boro Madhabpur	2%
38	Dighaigaon	3%	82	Dakatia	3%
39	Kandigaon	3%	83	Narikelbaria	3%
40	Senertikikata	3%	84	Purba Narayanpur	3%
41	Sabujnagar	3%	85	Bathail	3%
42	Fuljhuri	4%	86	Kotoktoil	3%
43	Gobindakathi	3%	87	Raipur	3%
44	Bakal	3%			

6. Results

In this section, we focus on presenting the results of our regressions as well as briefly discussing summary statistics. The next section deals with the economic interpretation of the results and compares and contrasts the findings of the different models.

The first step of our research was to recreate Godquin's original results to then be able to assess the effect of population mobility variables on her model. Firstly, Table 9 below gives a description of each variable and their meaning. Table 10 shows how our results match up to Godquin's. One of the important findings is that the social proxy Godquin uses, the AGEGP variable shows up with a negative sign and is indeed significant just as in Godquin's findings. As table 10 illustrates, we, apart from just the AGEGP variable, were able to recreate Godquin's result with a relatively high degree of accuracy, both in terms of the variables' values, their signs, and their significance. Along with this, all variables excluding the AGEGP, just like in Godquin's results, show up with the sign that one would intuitively expect.

TABLE 9
VARIABLE DESCRIPTIONS

VARIABLE	DESCRIPTION
PPRIN	Size of the loan (in taka)
DURATION	Duration of the loan (in days)
BRAC	Dummy = 1 if the loan was extended by BRAC
BRDB	Dummy = 1 if the loan was extended by BRDB
SEX	Dummy = 1 if the borrower is male; 2 if the borrower is female
PASSET	Value of the productive assets (in taka) of the borrower's household
SELEAGR	Dummy = 1 if the borrower received income from agricultural self-employment
NBLR	Number of relatives lended to
NFSL	Dummy = 1 if the loan program provided the borrower with access to basic literacy
NFSH	Dummy = 1 if the loan program provided the borrower with access to primary health facility
AGEGP	Age of the borrowing group at the due date (in months)
SAMEEDU	Dummy = 1 if the borrower and the group leader had the same education level (+/- 2 yrs)
SAMEAGE	Dummy = 1 if the borrower and the group leader are of the same age
CRD	Dummy = 1 if the borrower would have liked to borrow more at the same interest rate

Adding the Mobility Variables

Following the recreation of Godquin's result, the next step is to add our mobility variables and compare these new regressions to the baseline results. From that point on, we use the results we found when reproducing Godquin's regressions. We started out adding the NMR, followed by the ERE VILLAGE, and concluding with the ERE THANA. The results are laid out below.

Adding the NMR

Table 11 shows the results of adding the NMR variable to the regression both when keeping the incentive structure as it was created originally (i.e. without taking out the AGEGP variable) and with the modified incentive structure (i.e. taking out the AGEGP variable).

TABLE 10

REGRESSION COMPARISON

Variable	Godquin's results	Baseline results
PPRIN	-0.0004	-0.0002
	(-7.72)***	(-7.72)***
DURATION	0.0018	0.0014
	(5.14)***	(4.03)***
BRAC	-0.7247	-0.6666
	(-7.36)***	(-7.02)***
BRDB	-0.2866	-0.2461
	(-2.58)***	(-2.25)***
SEX	0.1025	0.1601
	(1.32)	(2.06)**
PASSET	0.0000	0.000
	(3.08)**	(2.47)**
SELEAGR	0.2805	0.3055
	(-3.62)***	(3.94)***
NBLR	0.0551	0.0529
	(5.09)***	(4.90)***
NFSL	0.1957	0.1425
	(2.56)**	(1.89)*
NFSH	0.1215	0.0708
	(1.46)	-0.86
AGEGP	-0.0057	-0.0066
	(-2.65)**	(-3.18)**
SAMEEDU	0.0048	0.0496
	(0.07)	(0.71)
SAMEAGE	0.0898	0.0625
	(1.31)	(0.92)
CRD	0.0419	0.0439
	(0.57)	(0.6)
CONS	-0.0672	-0.7203
	(-0.26)	(-3.03)***

Estimates using a probit, t statistics are given in parenthesis

*10% significance level

**5% significance level

***1% significance level

The addition of the Net Migration Rate variable results in a slight increase of the pseudo R^2 but that change, as shown later in the paper, is not significant. In addition, while the negative sign of the NMR's coefficient goes along with our theory, the variable is far from significant. It is important to note however that adding the mobility variable has only a very minor impact on the other variables from the baseline regression. Additionally, the results end up being very similar when the AGEGP variable is taken out, letting us think that the variables don't pick up very similar effects anyways.

The ERE Village Variable (ERE VILLAGE)

Secondly, we added the ERE VILLAGE variable. The results are summarized in Table 12. While the NMR estimations showed results that weren't significant, the findings for the ERE VILLAGE variable support our theory. The coefficient is negative, and the variable is significant to the 1% level. Additionally the regression shows an increase in the pseudo R^2 from 0.1217 to 0.1297. Once again, there was no notable difference in the results when taking out the AGEGP variable.

The ERE THANA Variable

Table 13 shows the regression results from adding the ERE THANA. While the ERE VILLAGE results matched up with our hypothesis, the addition of the ERE THANA is more puzzling. The variable is significant to the 1% level, but the sign goes against what our theory suggests. In short, the regression hints that a greater outflow of people from a thana would increase an individual's repayment rate. We discuss these surprising findings more fully in the discussion section (Section 7). In addition, we find a significant increase in the pseudo R^2. As before, there was no significant difference in the results when taking out the AGEGP variable.

TABLE 11
REGRESSION WITH NMR

Variable	W/O mobility variable	W/ NMR	W/ NMR and W/O AGEGP
FNM		**-0.1137**	**-0.0309**
		(-0.16)	***(-0.04)***
PPRIN	-0.0002	-0.0002	-0.0003
	(-7.72)***	(-7.88)***	(-11.93)***
DURATION	0.0014	0.0014	0.0012
	(4.03)***	(3.87)***	(3.44)***
BRAC	-0.6666	-0.6690	-0.7416
	(-7.02)***	(-6.98)***	(-8.14)***
BRDB	-0.2461	-0.2190	-0.2542
	(-2.25)***	(-2.0)**	(-2.39)***
SEX	0.1601	0.1563	0.2095
	(2.06)**	(1.97)**	(2.73)***
PASSET	0.0000	0.0000	0.0000
	(2.47)**	(2.53)**	(2.61)**
SELEAGR	0.3055	0.3151	0.2955
	(3.94***)	(4.04)***	(3.83)***
NBLR	0.0529	0.0534	0.0523
	(4.90)***	(4.78)***	(4.75)***
NFSL	0.1425	0.1632	0.1908
	(1.89)*	(2.12)**	(2.56)**
NFSH	0.0708	0.0564	0.0971
	(0.86)	(0.68)	(1.22)
AGEGP	-0.0066	-0.0062	
	(-3.18)**	(-2.97)***	
SAMEEDU	0.0496	0.0402	0.0172
	(0.71)	(0.57)	(0.25)
SAMEAGE	0.0625	0.0535	0.0526
	(0.92)	(0.78)	(0.78)
CRD	0.0439	0.0413	0.0274
	(0.6)	(0.56)	(0.37)
CONS	-0.7203	-0.6850	-0.8217
	(-3.03)***	(-2.84)***	(-3.53)***
Pseudo R2	0.1217	0.1233	0.1183

Estimates using a probit, t statistics are given in parenthesis
*10% significance level
**5% significance level
***1% significance level

TABLE 12
REGRESSION WITH ERE VILLAGE

Variable	W/O Mobility Variable	W/ ERE VILLAGE	W/ ERE VILLAGE and W/O AGEGP
ERE VILLAGE		**-22.1081**	**-19.7803**
		(-3.82)* **	**(-3.5)*** **
PPRIN	-0.0002	-0.0002	-0.0003
	(-7.72)***	(-7.23)***	(-11.58)***
DURATION	0.0014	0.0013	0.0011
	(4.03)***	(3.69)***	(3.17)***
BRAC	-0.6666	-0.6916	-0.7692
	(-7.02)***	(-7.25)***	(-8.45)***
BRDB	-0.2461	-0.1817	-0.2230
	(-2.25)***	(-1.65)*	(-2.08)**
SEX	0.1601	0.1261	0.1895
	(2.06)**	(1.6)*	(2.50)**
PASSET	0.0000	0.0000	0.0000
	(2.47)**	(2.55)**	(2.63)**
SELEAGR	0.3055	0.2972	0.2776
	(3.94***)	(3.8)***	(3.58)***
NBLR	0.0529	0.0631	0.0602
	(4.90)***	(5.66)***	(5.52)***
NFSL	0.1425	0.1705	0.2048
	(1.89)*	(2.24)**	(2.77)**
NFSH	0.0708	0.0428	0.0927
	(0.86)	(0.52)	(1.16)
AGEGP	-0.0066	-0.0073	
	(-3.18)**	(-3.45)***	
SAMEEDU	0.0496	0.0217	-0.0008
	(0.71)	(0.31)	(-0.01)
SAMEAGE	0.0625	0.0618	0.0597
	(0.92)	(0.9)	(0.88)
CRD	0.0439	0.0498	0.0353
	(0.6)	(0.67)	(-0.48)
CONS	-0.7203	0.0703	-0.1636
	(-3.03)***	(0.23)	(-0.55)
Pseudo R2	0.1217	0.1297	0.1236

Estimates using a probit, t statistics are given in parenthesis
*10% significance level
**5% significance level
***1% significance level

TABLE 13

REGRESSION WITH ERE THANA

Variable	W/O Mobility Variable	W/ ERE THANA	W/ ERE THANA and W/O AGEGP
ERE THANA		**35.3145**	**31.2085**
		(4.43)***	**(4.01)*****
PPRIN	-0.0002	-0.0002	-0.0003
	(-7.72)***	(-7.34)***	(-11.73)***
DURATION	0.0014	0.0013	0.0011
	(4.03)***	(3.71)***	(3.18)
BRAC	-0.6666	-0.6151	-0.7108
	(-7.02)***	(-6.42)***	(-7.83)
BRDB	-0.2461	-0.2464	-0.2905
	(-2.25)***	(-2.24)**	(-2.72)**
SEX	0.1601	0.1272	0.1949
	(2.06)**	(1.62)	(2.57)**
PASSET	0.0000	0.0000	0.0000
	(2.47)**	(2.76)***	(2.84)***
SELEAGR	0.3055	0.2819	0.2652
	(3.94***)	(3.6)***	(3.42)***
NBLR	0.0529	0.0533	0.0521
	(4.90)***	(4.86)***	(4.86)***
NFSL	0.1425	0.2280	0.2466
	(1.89)*	(2.94)***	(3.28)***
NFSH	0.0708	0.0780	0.1128
	(0.86)	(0.94)	(1.42)
AGEGP	-0.0066	-0.0075	
	(-3.18)**	(-3.52)***	
SAMEEDU	0.0496	0.0180	-0.0062
	(0.71)	(0.25)	(-0.09)
SAMEAGE	0.0625	0.0569	0.0537
	(0.92)	(0.83)	(0.79)
CRD	0.0439	0.0021	-0.0094
	(0.6)	(0.03)	(-0.13)
CONS	-0.7203	-1.7055	-1.7269
	(-3.03)***	(-5.15)***	(-5.35)***
Pseudo R2	0.1217	0.1319	0.1252

Estimates using a probit, t statistics are given in parenthesis
*10% significance level
**5% significance level
***1% significance level

Regression Fit Comparisons

The last part of the regression analysis is to test whether or not the addition of our mobility variables has improved the fit of the original model. Table 14 below summarizes the fit of the various regressions.

TABLE 14
PSEUDO R^2 COMPARISON

	Baseline	NMR	ERE THANA	ERE VILLAGE
Pseudo R^2	0.1217	0.1233	0.1319	0.1297
Difference from original R^2	0	0.0016	0.0102	0.008

As shown above, the pseudo R^2 increases with any of the mobility variables. It reaches its maximum fit with the ERE THANA variable at 13.19%, followed by the ERE VILLAGE and the NRM. To conclude if these results are significant, a Wald Test was conducted. Table 15 summarizes its results.

TABLE 15
Wald Tests

Regression	Wald Chi2	df	Pr > F
Baseline	248.48	14	0.0000
NMR	0.03	1	0.8707
ERE THANA	21.54	1	0.0000
ERE VILLAGE	13.8	1	0.0002

As the table shows, the increase in pseudo R^2 is significant for all the regressions except for the NRM. Most importantly, the better fit of the ERE VILLAGE regression compared to the original one is significant to the 98% level.

7. Discussion

The results we acquired tell a very interesting story about how population mobility is inherently intertwined with the notion of social capital. The insignificant results of the NMR can most likely be attributed to two things. Firstly, the NMR measures both inflows and outflows (total net migration) of people and while we have a theoretical justification for claiming that somebody leaving a borrowing group might negatively affect social capital, we have no evidence justifying that people moving into a thana, and eventually joining a borrowing group once new ones are formed, has a positive or negative effect on the social capital within the original group. More specifically, we do not have proof showing that the causality runs both ways. Secondly, it is likely that measuring population mobility on a thana level is just not granular enough, especially when we are looking at everyone in the thana, not just the subset of the population that has received MFI loans. This gives a consistent and contextually appropriate explanation for why the NMR provides us with insignificant results.

Based on the intuition that in reality only outflows of people are relevant for our measurement, we move on to our two ERE variables. As aforementioned, the ERE VILLAGE differs from the NMR in three very important ways. First of all, it is only measuring outflows, meaning it is more directly tied to social capital because of emigration's immediate impact on social capital (as explained in Section 5). Secondly, because of limitations in the data source, it only takes into account people who live or have lived in households in which someone had a joint liability loan extended to them through an MFI. Lastly, the variable is aggregated on the village level which means that the resolution, compared to measuring social capital on the thana level, has been greatly improved. These three differences are integral to the results we are seeing in the ERE VILLAGE variable. As Table 12 shows, results of the ERE VILLAGE coefficient are consistent with the theory that higher mobility leads to lower levels of social capital and thus lower repayment rates. This goes hand in hand with the theory we put forward in the framework, and confirms our initial hypothesis that social capital can be proxied by population mobility. Further evidence supporting this claim can be found in the predicted values that the repayment rate model provides. Table 16 below shows summary statistics for the predicted median repayment rates for the different villages.

TABLE 16
ERE Village Repayment Predictions Summary Statistics

Variable	Median	Std. Dev.	Min	Max
ERE VILLAGE Median Repayment Predictions	0.5480	0.1399	0.1423	0.7347

As evident by table 16, the median repayment rates in the villages is slightly shy of 55%, with minimum and maximum values at 14% and 73%. In addition, we found that the correlation between the ERE VILLAGE median repayment predictions and the ERE VILLAGE is -32.29%. Considering the number of other variables going into the repayment rate predictions and the variability of those, we consider this number a strong indicator of the fact that we are indeed picking up the depletion of social capital with our ERE VILLAGE variable. Another important quality of the ERE VILLAGE is that adding the variable to the baseline regression has very limited impact on the signs, coefficients and significances of the other factors measured in the model.

Moving on to the ERE THANA, we expected to see either an insignificant result, which would go along with the granularity argument from the NMR, or a negative significant result which would have supported our original theory. As table 13 suggests however, the variable came out to be positive and significant. While these results remain somewhat of a puzzle, perhaps requiring further research, we do have some thoughts as to how to reconcile them with the ERE VILLAGE results. The explanation that we put forward is the following: The ERE mobility variables pick up two types of effect. Firstly, the social capital aspect, which, as proved by the ERE VILLAGE results, diminishes as mobility goes up. Secondly however, there may be some other, unobserved variables that also get picked up by the EREs. On the village level, the social capital aspect dominates these unobserved variables, leading to a result consistent with our theory. On the thana level however, the results are reversed, because the social capital aspect is dominated by these latent variables, giving rise to results that are inconsistent with our theory. In general, it seems that as the size of the group whose migration is being measured increases, the link to the social capital should become weaker and weaker. This means that one would expect those other, unobserved, effects to begin to dominate.

To further strengthen this argument, it is crucial to understand the intrinsic differences between a thana and a village. As aforementioned, a thana is a geographical and administrative region comparable to a county in the US. There are 1009 thanas in total in Bangladesh. A thana is comprised of multiple villages, and it is from here that we infer a crucial assumption: people from different villages but in the same thana most likely do not know each other as well as people within a given village in the thana, and their social capital is most likely not impacted if someone from another village moves.

The villages, on the other hand, are much smaller, ranging between 150-450 people in size. This leads us to assume, opposite of the ties on the thana level, that everyone in a village at least know of each other, and that when someone emigrates out of a village, that has a negative, depleting impact on the social capital of that community. This assumption is strongly supported by research in sociology. Hill and Dunbar (2003) examined the size of social networks in humans by studying the exchange of Christmas cards and found the average maximum network size per person to be 153.5 individuals. It then seems reasonable to think that a lending group of about 2 to 4 people would have a social network large enough to know of all people in the village.

So, the difference in the resolution between the two measurements, the village level and the thana level, leads to different effects dominating. On the thana level, because people from different villages don't necessarily know each other, the level of social capital depletion is minimized in comparison to the effect of the other, unobserved, variables. Going back to the ERE VILLAGE results however, we see that when we have stronger social ties and connections, the mobility proxy truly captures the severing effect outbound mobility has on those ties, subsequently negatively affecting the repayment rates.

Additionally, both ERE variables significantly improve the fit of the model, which confirms that in both cases we are picking up some effect on the repayment rate that was not accounted for in the original model. As mentioned earlier, we do believe that the improved fit stemming from the addition of the ERE THANA variable is the case of one or many unobserved, latent variables being inferred through the ERE THANA. On the other hand, the improved fit that we are able to observe when adding the ERE VILLAGE variable seems much more in line with our main theory.

To summarize, our findings confirm our initial hypothesis that social capital can be proxied using a population mobility variable. If attention is only paid to outflows of people, and the population of the group inspected is sufficiently small, with contemporary sociological research providing good guidelines as to where the social ties start getting weaker, then that measure improves the fit of Godquin's repayment model. In addition, our research shows that at the appropriate level (here the village level) higher population mobility is linked to lower repayment rates.

8. Conclusion

Based on the lack of an easy, simple way of adding a social capital component to the repayment rate models used in microfinance, we constructed a hypothesis that stated that population mobility could serve as such a proxy. After theoretically justifying our claim, we went on to extend Godquin's repayment rate model with three different mobility variables, each capturing different nuances of the social capital phenomenon. After running the regressions, it appears that the best way to proxy for social capital is to look at population outflows rather than at net migration levels. In addition, it is crucial to get information on a small enough scale, here at the village level. This is because, on a larger scale, other effects that stem from the population size differences between a thana and a village interfere too much, and dominate the social capital effect.

Once we ran geographically localized data with only outflows considered through our model, we obtained useful results. Hence it seems that, conditional on the two aforementioned factors, social capital can be proxied by population mobility, with improved regression fit and repayment predictability as a result. We were able to conclude that an increase in mobility on the village level is linked to a decrease in the repayment rates of the individuals in the village.

Obviously, it is appealing to be able to say what kinds of unobserved variables dominate social capital on a larger geographical scale, and while we don't have the data to thoroughly investigate the matter, we do have a few candidates that are likely to play a role. One possible explanation is the fact that there exist great heterogeneity within a thana in terms of both population mobility and social capital. A thana, apart from villages, may contain one or more larger cities in which mobility is heightened compared to the surrounding more rural areas, resulting in lower levels of social capital. Conversely, the surrounding villages within that same thana might see considerably lower mobility, with more and closer social ties, resulting in higher levels of social capital and thus increased repayment rates. In aggregate, this uneven distribution of urbanization within certain thanas might help explain why the results on the village and thana level differ so greatly. For example, within the thana of Rangpur Sadar, the city of Rangpur, one of the largest cities in Bangladesh, lies, whereas the thana also contains the village of Amashu (with a population of less than 450 people).

Another possible explanation to our unexpected results is based on the concept of scarce resources. As people emigrate out of a thana, the resources that those people allocated become freed up, which decreases competitiveness and theoretically should increase the economic profits the remaining businesses make. Thus, as profitability goes up, so do repayment rates. To investigate this further, one would probably benefit from using a dynamic rather than static model, in which the degree of competitiveness, or level of scarcity of the necessary inputs is fluctuating along with people's decision to emigrate.

Lastly, another potential argument is that of non-rivalry of ideas. When an individual comes up with an idea that earns superior returns in a given place, the idea will stay there and keep on making the same kinds of returns, regardless of whether the innovator of the idea emigrates. This may be an important factor for MFIs to consider in order to increase the repayment rates of lending groups even though population outflows in this location might be quite high.

It is important to note that these are all notions that play a role not just on the thana level but on the village level as well. However, as discussed earlier, the social capital effect dominates on the village level, overpowering these unobserved variables. Getting to the bottom of why we are seeing a positive correlation between population mobility and repayment rates on the thana level obviously requires both more research and data collection to be thoroughly investigated, and as such it constitutes an interesting opportunity for future research. In the end, what our research really points out is that higher outflows of population are linked to lower repayment rates in microfinance when the measured aggregation unit, in our case the village, is sufficiently small.

As far as policy implications go, and in order for MFIs to obtain higher repayment rates, it seems crucial to reduce outbound migration at the smallest geographical level possible. The ways in which this could be achieved are numerous, but intuitively, adding services that make the community more self-sustainable should constitute a good effort. Such things could be health centers, schools, after-school daycare or work training.

In terms of scholarly implications, our results serve as the foundation for further studies on microfinance performance where less time would have to be spent on developing localized, expensive, subjective and qualitative measures of social capital, and where the researchers could instead focus on simple migration rosters, keeping track of population changes within the spatial region the research focuses on. This will most likely significantly improve the reliability of the data set and reduce the overall cost of the data collection.

Our research raises additional questions that would serve as interesting topics for further research. Most notable is how population mobility becomes increasingly important for social capital as the geographical resolution increases. As such, an interesting study would be to get migration rosters for every specific borrowing group rather than on a village level and see how this impacts the coefficient and its significance. Furthermore, one could expand on the case of the ERE THANA and conduct a study to see if MFI supported businesses, on the margin actually become more profitable as people leave the region. If this is proved to hold, it could help develop better methods for poverty alleviation.

Sources

Literature Sources

Brau, James C.; Woller, Gary M. (2004). "Microfinance: A comprehensive review of the existing literature, Journal of Entrepreneurial Finance", JEF, ISSN 1551-9570, Vol.9, Iss. 1, pp. 1-27, http://hdl.handle.net/10419/55971

Bryceson, Deborah Fahy; Bradbury, Annabel and Trevor. (2006), "Roads to Poverty Reduction? Dissecting Rural Roads' Impact on Mobility in Africa and Asia", Development Policy Review; 26(4):459-482.

Christen, Robert Peck, Rosenberg, Richard, and Jayadeva Richard. (2004), "Financial institutions with a double-bottom line: implications for the future of microfinance", CGAP Occasional Paper, July 2004.

Christen, Robert Peck, Elisabeth Rhyne, Robert Vogel, and Cressida McKean. (1995). "Maximizing the Outreach of Microenterprise Finance: An Analysis of Successful Microfinance Programs", USAID Program and Operations Assessment Report No. 10. Washington, D.C.: U.S. Agency for International Development

D'Espallier et Al. 2009. "Women and Repayment in Microfinance". Microfinance Gateway Working Paper. Oct 8th 2012 http://www.microfinancegateway.org/p/site/m//template.rc/1.9.40253

DiPasquale Denise; Glaeser Edward L. (1999), "Incentives and Social Capital: Are Homeowners Better Citizens?", Journal of Urban Economics 45, p354-384

Dudwick et Al. 2006. "Analyzing Social Capital in Context: A Guide to Using Qualitative Methods and Data". The World Bank. Stock No. 37260. Oct 8th 2012 http://siteresources.worldbank.org/WBI/Resources/Analyzing_Social_Capital_in_Context-FINAL.pdf

Field, Erica & Pande, Rohini. 2008. "Repayment Frequency and Default in Micro-Finance: Evidence from India". Journal of the European Economic Association, Vol 6 2008: 501-509.

Ghatak, Maitreesh. 1999. "Group lending, local information and peer selection". Journal of Development Economics, Vol. 60 1999: 27–50

Godquin Marie, 2004, "Microfinance repayment performance in Bangladesh: How to improve the allocation of loans by MFIs", World Development, 32 (11) (2004), pp. 1909–1926

Gomez, R. and Santor, E. (2001), "Membership has its privileges: the effect of social capital and neighborhood characteristics on the earnings of microfinance borrowers", London: LSE Research Articles Online.

Hill, R. A. & Dunbar, R. I. M. 2003. "Social Network Size In Humans", Human Nature, Vol. 14, No. 1, pp. 53-72.

Karlan, Dean. 2007. "Social Connections and Group Banking". The Economic Journal, 117: F52-F84

Krishna, Anirudh & Shrader, Elizabeth. 2000. "Cross-cultural Measures of Social Capital: A Tool and Results from India and Panama". The World Bank Social Capital Initiative Working Paper Series. Oct 8th 2012 http://siteresources.worldbank.org/INTSOCIALCAPITAL/Resources/Social-Capital-Initiative-Working-Paper-Series/SCI-WPS-21-paper.pdf

Lee, S. Everett (1966), "A Theory of Migration", Demography, Vol. 3, No. 1, pp. 47-57.

Lubotsky, Darren; Wittenberg, Martin (2006), "Interpretation of Regressions with Multiple Proxies," The Review of Economics and Statistics, MIT Press, vol. 88(3), pages 549-562, August.

Molloy, Robert, Smith, Christopher L., Wozniak, Abigail, (2011), "Internal Migration in the United States", Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Vol. 2011-30, pp. 34.

Oke et Al. (2007), "An Empirical Analysis of Microcredit Repayment in Southwestern Nigeria". Humanity & Social Sciences Journal 2 (1): 63-74.

Putnam, Robert (2000), "Bowling Alone: the collapse and revival of American Community", New York: Simon & Schuster

Putnam, Robert. 2003. "Social Capital: Measurements and Consequences". OECD Directorate for Education. Oct 8th 2012 http://www.oecd.org/edu/educationeconomyandsociety/1825848.pdf

Routledge Bryan; von Amsberg Joachim (2002), "Social Capital and Growth", *Tepper School of Business*, Paper 419, http://repository.cmu.edu/tepper/419

Schiff Maurice, (September 2004) "Labor Mobility, Trade and Social Capital", Review of International Economics, Vol. 12, No. 4, pp. 630-642

United Nations (1970), "Methods of Measuring Internal Migration", Department of Economic and Social Affairs

Woolcock Michael, “Microenterprise and social capital:: A framework for theory, research, and policy”, The Journal of Socio-Economics, Volume 30, Issue 2, 4 March 2001, Pages 193-198, ISSN 1053-5357, 10.1016/S1053-5357(00)00106-2. http://www.sciencedirect.com/science/article/pii/S1053535700001062

Zak Paul; Knack Stephen (2001), “Trust and Growth”, Economic Journal 111, p295-321

Data Sources

Micro-credit Impact Studies in Bangladesh Survey, Shahidur R. Khandker, conducted by Bangladesh Institute of Development (BIDS) and the World Bank, "http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/0,,contentMDK:21470820~pagePK:64214825~piPK:64214943~theSitePK:469382,00.html"

Bangladesh Population and Housing Census (1991, 2001 and 2011), conducted by Bangladesh Bureau of Statistics, http://www.bbs.gov.bd/PageSecureReport.aspx