

CSAE Working Paper WPS/2015-13

Can community-based microfinance groups match savers with borrowers? Evidence from rural Malawi *

Rachel Cassidy[†] Marcel Fafchamps[‡]

May 21, 2015

*This project was funded by the UK Department for International Development (DFID) as part of the research consortium Improving Institutions for Pro-Poor Growth (iiG), which studies how to improve institutions in Africa and South Asia. The views expressed are not necessarily those of DFID. Rachel Cassidy's work is supported by the Economic and Social Research Council [grant number ES/J500112/1]. We thank Waluza Munthali, James Mkandawire and all at Invest in Knowledge Initiative for excellent surveying assistance (<http://www.investinknowledge.org/>); Hardish Bindra for excellent data entry assistance; and Smit Gade for excellent research assistance. We also thank Helene Bie Lilleør and the Rockwool Foundation for kindly granting access to supplementary data (<http://www.rockwoolfonden.dk/research/the+rockwool+foundation+research+unit>) and Christopher Ksoll, Jonas Helth Lønberg and Ole Dahl Rasmussen for fieldwork advice. We thank Sam Asher and Paul Novosad for sharing their “fuzzy matching” algorithm. Finally, we are grateful to James Fenske, Climent Quintana-Domeque, Simon Quinn, Daniela Scur and participants at the Oxford Firms Group, IIG Conference 2015 and CSAE Conference 2015 for useful comments. All errors are our own.

[†]Corresponding author. University of Oxford, Department of Economics and Centre for the Study of African Economies. Email: rachel.cassidy@economics.ox.ac.uk

[‡]Stanford University, Freeman Spogli Institute for International Studies. Email: fafchamp@stanford.edu

Abstract

This paper examines how members sort across community-based microfinance groups, specifically Village Savings and Loan Associations in rural Malawi. Our central question is to ask whether such groups allow savers (especially commitment savers) to match with potential borrowers, thereby promoting financial intermediation. We analyse novel data in the form of a census of all 3,800 members of 150 VSLA groups. We first develop predictions on sorting in terms of individual members' occupation and present bias, and then test these predictions in a dyadic regression framework. We find evidence that whilst there is positive assortative matching on occupation, suggesting unrealised intermediation possibilities; there is negative assortative matching on present bias, indicating that these groups do at least create a degree of financial intermediation between commitment savers and borrowers. The latter may be welfare-enhancing for both commitment savers and borrowers, given the low access to commitments savings technologies and to credit in these communities.

1 Introduction

Community-based microfinance groups have experienced rapidly-growing popularity among NGOs and policymakers in recent years (Gash and Odell, 2013). The most prominent examples are so-called “Self-Help Groups” — also known as VSLs, VSLAs, SILCs, SfCs and other acronyms depending on the NGO responsible — which are particularly prevalent in India and Sub-Saharan Africa and now have over 100 million members worldwide (Greaney et al., 2013). The key features of Self-Help Groups are as follows. Participants are first invited to self-select into groups, and then receive training from an NGO or NGO-trained field agent.¹ The group then comes together for weekly meetings, at which individuals can make savings deposits. These deposits cannot be withdrawn until the end of the savings cycle (usually a year). Members can also request interest-bearing loans out of the total savings fund, and any profits from lending are shared out at the end of the cycle in proportion to individual members’ savings. Participants thereby obtain a secure way to save, earn dividends on savings and potentially gain access to credit.

Given that some members may use such groups primarily for saving whilst others may use them predominantly to seek credit, these groups may in theory greatly enhance financial intermediation within these villages: reallocating capital from those with a demand for saving towards those with a demand for credit, and potentially benefitting both sides if a fair interest rate on lending can be reached.² Given this, the ways in which individuals sort across such groups may be crucial in determining the level of intermediation and thus the efficiency and equity of the financial outcomes generated. Our central question in this paper is to ask whether such groups allow those with savings needs, in particular those engaged in agriculture and those who are “commitment savers”, to match with potential borrowers, in particular those engaged in non-farm activities. Despite the

¹Such groups typically do not form without NGO involvement for several reasons. Firstly, the NGO can help overcome any inherent coordination problems by bringing people together and highlighting the benefits of financial services and intermediation. Secondly, the system of governance and accounting used to protect members against theft, default or improper conduct by other members is quite sophisticated, and so the NGO training is key. Thirdly, the NGO provides access to account books and safe boxes, although the group is typically asked to purchase these out of its own funds rather than receiving these materials for free.

²See section 6.3 for discussion of interest rates in this context.

potential benefits, it may be that such sorting does not materialise, for example due to difficulties in enforcing the repayment of loans, or if more general social obligations obstruct sorting along financial lines.

We provide an empirical analysis of sorting using data from a specific type of Self-Help Group, known as “Village Savings and Loan Associations” (VSLAs).³ In the Summer of 2013, we revisited all 150 VSLA groups originally trained as part of an intervention in 2009-11 in rural Malawi, as evaluated in Ksoll et al. (2013).⁴ We enumerated a census of all members, past and present, and also collected basic demographic information for each individual as well as their membership history.⁵ Given that we collected this information two to four years after the groups were initially trained, our data is uniquely suited to studying long-run equilibrium sorting patterns of members across groups (and long-term use of the groups use more generally). In particular, individuals have had time to learn about the financial technologies provided by VSLAs and the benefits of grouping with different members. Accordingly they have had chance to join VSLA groups, switch across groups at the end of savings cycles, and indeed drop out of groups altogether.⁶

We develop a stylised conceptual framework to guide predictions on how members might sort across groups, and to frame the logic of our empirical test. We first outline the potential role of sorting on occupation: specifically, the two predominant occupations in our survey villages — agriculture and small business — may imply very different patterns of savings and borrowing needs; therefore it may be efficient for farmers and business-people to group together (*negative* assortative matching on occupation) in order to allow financial intermediation between these two groups. We next highlight the possibility of sorting in terms of savings preferences. We describe how VSLAs act as a multi-faceted commitment savings technology, and thus are particularly likely to attract sophisticated present-biased individuals as members. In terms of sorting, the most interesting scenario

³VSLAs follow the standard “Self-Help Group” model as described earlier. Section 2 outlines the workings of VSLAs in more detail.

⁴See section 2.1.

⁵We also enumerated a group survey with questions on group history and practices, as well as photographing all individual members’ account books. Section 4 describes the data collection in more detail.

⁶We do observe a significant number of late joiners, drop-outs and individuals who switch across groups or become members of more than one group; see section 4.

from a financial intermediation perspective is if these present-biased commitment savers sort into groups with time-consistent individuals who join as a way to borrow. This would lead us to observe *negative* assortative matching on present bias in the data. In contrast, if the credit function of VSLAs becomes negligent in the long run⁷ then two scenarios are possible. If time-consistent savers join as well as commitment savers, and if time-consistent savers make the most desirable savings partners (as is plausibly the case), then we would expect to see *positive* assortative matching on present bias. Alternatively, if interest rates on saving in VSLAs are not high enough to attract time-consistent savers, then only commitment savers participate and so we would expect to see *zero* assortative matching on a binary indicator of present bias.

We then employ a dyadic regression framework to test the sorting predictions empirically. Specifically, we construct all pairs of members i and j from the census, and then estimate multivariate logit equations to examine the determinants of i and j being members of the *same* group (given that, by virtue of being in our dataset, i and j are both members of *some* group). In particular, we are able to test whether *ceteris paribus* there is evidence of positive or negative (or zero) assortative matching on measures of individuals' occupation and present bias, as well as a host of controls. Importantly, our regressors come from matching individuals to survey data from 2009, before VSLA groups were introduced into the area. We therefore eliminate concerns about reverse causation, which might arise if being members of the same group leads individuals to have a greater difference or similarity in their measured characteristics.

Our results provide a mixed picture in terms of how much potential financial intermediation is being realised through VSLAs. In contrast to the prediction of negative assortative matching on occupation, the effect of i and j both having the same primary source of employment *increases* the probability of being in the same group by 16.8 percentage points (i.e. *positive* assortative matching). However, in line with the prediction on efficient matching in terms of savings preferences, there is strong evidence of *negative* assortative matching on present bias. The latter is indicative that (conditional on occupation) these groups do at least create a degree of financial intermediation between commitment

⁷This may happen for several reasons, e.g. if loan repayment is hard to enforce. See section 3.

savers and borrowers. In terms of magnitude, both effects are similar in size: if member i in village v has the same occupation as member j also in village v , then this increases the probability that i and j are members of the *same* group by around 16.8 percentage points; whilst if i is present-biased and j is *not* present-biased (or vice versa), then this increases the probability that i and j are members of the same group by around 16.5 percentage points. Either these effects is equivalent to around 50% of the baseline probability that two VSLA members from the same village are in the same group (29.6%). The coefficients on occupation and present bias persist across a range of specifications and are robust to the inclusion of a host of controls and to various robustness checks.

The lack of negative sorting on occupation suggests that there are unrealised gains from financial intermediation between occupational groups. This may be because individuals find that the informational and enforcement advantages of grouping with people from the same occupation outweigh the potential benefits (in terms of capital reallocation) of matching with individuals who have a demand for borrowing and saving at different times. On the other hand the fact that commitment savers ostensibly match with those seeking to borrow may be welfare-enhancing — both for commitment savers, given the low access to commitment savings technologies in these communities, and for borrowers, given the low access to credit. Nevertheless, NGOs and policymakers might be concerned if present biased individuals’ willingness to pay for commitment leads them to accept a very low interest rate when lending out their savings to borrowers.⁸ We offer preliminary evidence that this is not a serious concern, as interest rates on lending appear to be in line with the “fair” benchmark of individuals’ average monthly discount rates.

1.1 Contribution to existing literature

Our analysis contributes to several related literatures. In terms of the broader micro-finance literature, we add to a more recent body of work which has shifted focus away from micro-lending and towards access to savings — in particular commitment savings

⁸Assessing welfare of course become complicated in the presence of time-inconsistent individuals. See section 6.3 for further discussion.

— as a key part of expanding financial inclusion in developing countries (Ashraf et al., 2006; Dupas and Robinson, 2013; Karlan et al., 2013).⁹ Indeed, there has also been a renewed focus on the role of groups in such a context (Kast et al., 2012). However, to our knowledge we are the first to examine the potential for groups to mediate interaction between savers and borrowers, and in particular commitment savers.

The literature specifically on Self-Help Groups is relatively nascent. Several recent papers have exploited randomized controlled trials in order to assess the impact of introducing Self-Help Groups at the village level on household-level outcomes (Ksoll et al., 2013; Beaman et al., 2014). Effects on consumption smoothing appear to be positive, as both Ksoll et al. (2013) and Beaman et al. (2014) find positive and significant intent-to-treat effects on indicators of food security (Ksoll et al. (2013) suggest that this may be linked to increased agricultural investments). However, similar to randomized evaluations of pure micro-credit, neither study finds significant effects of Self-Help Groups on business profits, health, education or female empowerment. The researchers note however that this may be an artefact of short evaluation time-frames, as most of those who joined Self-Help Groups had completed at most one savings cycle by the time of the endline surveys.

In closer relation to this paper, little work has been done to examine Self-Help Groups at the group level, or indeed to analyse the financial characteristics of those who participate. A key exception is Greaney et al. (2013), who run a field experiment in Kenya, Tanzania and Uganda to compare the performance of groups who pay for their own training to groups whose training is NGO-funded. They find that groups who pay for their own training do better across a number of group-level indicators, such as loan repayment rates and average member business profits. Importantly, they also examine individual selection into the Self-Help Group program, in terms of the characteristics of those who participate compared to those who do not across the two regimes. In particular they model borrowing motives for joining Self-Help Groups, and argue that making groups pay for their own training reduces adverse selection by driving out individuals who pose

⁹There have also been suggestions that micro-credit may be used by many as a substitute for commitment savings products (Bauer et al., 2012).

a bad credit risk.

Our analysis complements Greaney et al. (2013)'s study in two ways. Firstly, we consider the related but separate question of, conditional on participation, how members sort across groups. As outlined above, sorting may be crucial in determining how well Self-Help Groups are able meet the financial needs of their members. We also specifically highlight commitment savings motives for joining such groups, which we believe are a key part of understanding Self-Help Groups but which are overlooked by Greaney et al. (2013) due to their focus on the (also important) issue of adverse selection.

Our motivation is conceptually similar to a body of work which examines sorting in informal financial institutions more broadly: Ghatak (2000) and Ahlin (2009) develop theoretical models of how sorting might enable the efficient pricing of risk in the context of joint-liability microfinance; and Eeckhout and Munshi (2010) examine empirical patterns of sorting into chit funds in India.¹⁰ Our contribution differs insofar as we examine sorting when some members are primarily interested in saving, although others remain primarily interested in borrowing.¹¹

Our empirical strategy builds on work by Arcand and Fafchamps (2012) who also use a dyadic regression framework to study sorting in community-based organisations. In contrast to our work, the groups they consider are predominantly producer cooperatives rather than community-based microfinance groups, and their focus is on program inclusiveness rather than financial intermediation. Other authors have used dyadic frameworks to examine sorting on risk preferences for risk-sharing activities (Attanasio et al., 2012; Barr et al., 2012). Our approach differs in that we study sorting on time preferences for savings and borrowing purposes, and use data from a fully-fledged program setting rather than from a framed field experiment.

The remainder of the paper is organised as follows. Section 2 explains the workings of

¹⁰“Chit fund” is the term used in India to denote a bidding Rotating Savings and Credit Association.

¹¹ROSCAs and chit funds also potentially offer a commitment savings technology, but Eeckhout and Munshi (2010) model all members as joining with the sole motivation of getting an early turn in the rotation and thus obtaining a *de facto* loan.

VSLAs in more detail, and provides information on the specific intervention from which we draw our data. Section 3 outlines a stylised conceptual framework for examining sorting, and draws out the key predictions about how we might expect participants to sort across VSLAs. Section 4 describes key aspects of the data. Section 5 outlines the empirical strategy to test the predictions, and to explore sorting in VSLAs more broadly. Section 6 details the results and robustness checks, and discusses the potential welfare implications. Section 7 concludes.

2 Village Savings and Loan Associations (VSLAs)

Our data comes from a specific type of Self-Help Group known as “Village Savings and Loan Associations” (VSLAs), although VSLAs embody the standard Self-Help Group model and thus to that extent are representative of virtually all types of Self-Help Group.¹² This standard model is as follows. After an initial awareness-raising meeting held by the NGO, individuals are invited to self-select into groups of 15-25 members in order to participate. With the help of the NGO, each group then purchases a cash box with three separate locks and elects three different members to act as key-holders, thereby reducing the probability that any funds placed into the box will be subject to theft. The NGO or an NGO-trained field agent then assists the group in writing a constitution, and trains the group in financial literacy and account-keeping over a period of several weeks. The group then begins to hold weekly meetings, at which each group member can make savings by purchasing between one and five “shares” of a fixed, small value.^{13,14} After a month, members can also start requesting to take loans, to be repaid

¹²VSLAs were arguably the first version of “Self-Help Groups” to be introduced to Sub-Saharan Africa, by CARE International in Niger in 1991. VSLAs alone have seen a dramatic acceleration in outreach over recent years, and now have more than six million active members across 61 countries. See <http://www.vsla.net>.

¹³The median share price observed in 2013 was 100 Malawi Kwacha, equivalent to around \$0.30 (the exchange rate at the start of the July 2013 survey was 330 MWK = 1 USD). Members of the median group can therefore save between \$0.30 and \$1.50 per week.

¹⁴Each member is also required to make a small weekly contribution to the group’s insurance fund, to cover events such as illness or death of a group member’s relatives. However, in practice we found that most groups set the level of such contributions to be very small (around 20 MK, equivalent to \$0.06) and some groups had even dropped this component altogether, citing past disagreements about payouts.

after a month at a fixed monthly rate of interest. At the end of each cycle — usually a year — the group’s total remaining savings fund plus the successfully-recovered loans and the loan interest is “shared out” in proportion to individual members’ savings (hence the term “shares”). Participants thereby obtain a secure way to save, earn dividends on saving and potentially gain access to credit.

Comparing their structure to that of other, more established savings and credit institutions, VSLAs lie somewhere between credit cooperatives and Rotating Savings and Credit Associations (ROSCAs). VSLAs clearly bear a strong resemblance to credit cooperatives and credit unions. However, VSLAs are generally much smaller and much less formalised — for example VSLAs typically have no legal status, unlike some of the larger rural credit unions. As a consequence, VSLAs are likely to be more reliant on interpersonal relationships for monitoring and enforcement of loan repayment (which may have both advantages and disadvantages). Insofar as VSLAs are informal savings and credit groups which rely on “social capital” to sustain them, they are close in spirit to ROSCAs. However, VSLAs are more sophisticated than ROSCAs in that they allow for a much greater degree of flexibility: both on the savings side, as each individual member can choose to buy between one and five shares each week (rather than committing to a fixed payment which is the same for all members and all weeks); and on the lending side, as members can demand the size and timing of the loans they receive (rather than having to wait for their turn in a rotation).¹⁵ Meanwhile, compared to traditional group-based microfinance, VSLAs are clearly distinct insofar as they offer a commitment savings technology as well as credit (section 3 discusses the commitment savings features of VSLAs in more detail). Furthermore, VSLA interventions are also typically targeted at communities which are under-served even by microfinance lenders.

¹⁵Bidding ROSCAs do allow members some choice over when they receive the pot; however, each member can still only receive the pot once and cannot choose the pot size.

2.1 VSLA survey

In the Summer of 2013, we surveyed 150 VSLA groups spread across 44 villages in the Karonga District of northern Malawi (see figures 1 and 2). These VSLAs were originally formed as part of an intervention which ran from 2009 to 2011, which was funded and evaluated by the Rockwool Foundation¹⁶ and implemented by CCAP Synod of Livingstonia Development Department (SOLDEV).^{17,18} The results of the initial evaluation are detailed in Ksoll et al. (2013), as outlined in section 1.1.

We returned to survey all 150 groups that had been trained directly by SOLDEV as part of the initial 2009-11 intervention.¹⁹ Forty-six villages were included in the initial intervention, half of which were invited to form groups and receive VSLA training in 2009 (treated villages), and half of which were not eligible for such treatment until 2011 (control villages). Since we visited the area two years after the control villages were phased into treatment, our sample covers both treatment and control villages.²⁰ However, two of the control villages essentially dropped out of the program in 2011 and never established any groups, most likely due to the extreme remoteness of these two villages. Therefore our 2013 sample contains information on all groups in the remaining 44 villages.

We contacted each group via their field agent, who invited all group members to a meeting.²¹ At the start of each meeting, we first explained the purpose of our survey and

¹⁶See <http://www.rockwoolfonden.dk/programme+areas/food+security+and+poverty+alleviation/saving+to+survive+in+malawi>. We are grateful for the cooperation of the Rockwool Foundation and SOLDEV in our own work in this area.

¹⁷<http://www.ccapsolinia.org/>.

¹⁸The training of these groups was paid for by the Rockwool Foundation, rather than by members themselves. As discussed in section 1.1, Greaney et al. (2013) show that whether NGOs or members pay for training affects who participates in VSLAs. To that extent, our empirical results on sorting may only be representative of groups in which NGOs pay for the training. However, this is still by far the most widely-used model for Self-Help Group interventions.

¹⁹Anecdotally, we learned that a number of “replication” groups did later form as individuals tried to mimic the VSLA system, or were taught by VSLA members who were later trained by SOLDEV to propagate VSLAs. Data on such informal groups was difficult to obtain. However, this is not a major concern since our results are still internally valid for all members of official groups.

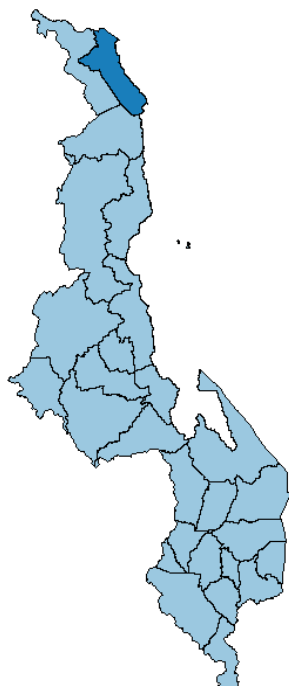
²⁰The analysis in this paper does not exploit the initial randomisation.

²¹To maximize attendance, these meetings were pre-scheduled and were held at the group’s usual meeting place.

Figure 1: Malawi



Figure 2: Location of Karonga District (dark blue) within Malawi



obtained the consent of all group members to be surveyed and to have their individual account books photographed. We then asked that three or more members stay to assist with our survey. In practice, almost all group members chose to remain and to help answer the survey questions. The data collection then proceeded in three steps. First, we used the set of individual account books (as a rule these are stored inside the group’s cashbox) to construct a roster of all group members, past and present. We then elicited basic demographic information for each member as well as their membership history, by reading out each member’s name and then asking a series of questions about that individual. Second, we enumerated a survey about the group’s history and practices, such as the interest rate charged on loans and the typical punishment for late loan repayments. Finally, we photographed each individual’s account book, which details their weekly savings decisions and their borrowing behaviour for the entire current cycle (and often also for previous cycles).

We were later able to match around a fifth of the individual members (722 out of 3,801) by name to the 2009-11 panel dataset which was collected by Ksoll et al. (2013) for the initial impact evaluation.^{22,23} From now on we refer to this subsample as “matched individuals”. Specifically, we matched individual names from our member census to the names of household heads and their spouses from the panel survey households. Importantly, the preference modules of the panel survey were administered separately to both the household head and the spouse (where applicable), thus we are able to obtain preference information at the individual level for the matched individuals. Furthermore, by matching to the 2009 panel wave we obtain this data as measured before the introduction of VSLAs into the area.

²²We thank the authors for sharing this data. We also thank Sam Asher and Paul Novosad for developing a “fuzzy matching” program, based on Levenshtein’s algorithm, which we used to conduct the matching.

²³The panel dataset covers a stratified random sample of households from the treatment and control villages, surveyed at baseline before the introduction of VSLAs (2009), one year into the VSLA program (2010) and at endline just before the control villages were phased into treatment (2011). By construction, the 2009 wave (and indeed the later waves) contains information on some individuals who go on to become members of VSLAs and other individuals who do not. However, since our own census is of VSLA members only, the individuals in our dataset are matched only to individuals from the 2009 panel wave who go on to be members — we do not use information on those who go on to be non-members.

3 Sorting across VSLAs - conceptual framework

We now outline a stylised conceptual framework to make clear the possible role of sorting across VSLAs, and to motivate the logic of our empirical test in section 5. We first note that there are multiple groups per village in the vast majority of villages (see section 4) and that there is considerable scope for members to sort across these groups. As outlined above, when VSLAs are first introduced, individuals must self-select into groups prior to asking for NGO training. Furthermore, individuals may later leave a group at the end of its cycle, and/or join a new group at the start of its next cycle. Indeed, we observe considerable “churn” in terms of individuals joining, leaving and moving across groups, as outlined in section 4.

3.1 Sorting on occupation

Arguably the most salient characteristic on which we might expect to see sorting for financial intermediation purposes is occupation. The population of the survey area is engaged overwhelmingly in one of two occupations — agriculture, or trading/small business (see section 4 for details). These occupations likely have very different cash-flow needs: farming households require a large investment of funds during the planting season in January, and essentially receive one major inflow of funds just after the harvest in April.²⁴ Meanwhile those engaged in business (and other non-farm activities) are likely to have frequent, although fluctuating, income and may require access to capital throughout the year.²⁵

In terms of sorting, from a financial intermediation perspective we might therefore predict to see *negative* assortative matching on occupation — i.e. those in agriculture sorting into groups with those engaged in non-farm activities. If this were to occur, the farmers

²⁴April is the major harvest for the area’s main crop — maize — and indeed for rice. There is also a second, smaller harvest for cassava which takes place in November.

²⁵Typical businesses in the area include selling vegetables and goods from nearby markets, bricklaying and carpentry, driving bicycle taxis, and sewing.

would (in expectation) act as savers for most of the cycle — except for occasional out-of-season agricultural investments and emergencies — whilst those in non-farm activities could frequently borrow out of the farmers’ savings, thereby generating dividends for the farmers through loan interest repayments. Such sorting might be achieved through direct bargaining or through the loan interest rate: competition for borrowing funds amongst non-farmers within a given VSLA would drive up the interest rate on lending, eventually pushing some non-farmers to switch membership to other groups with relatively few non-farmers.

On the other hand, a second possible equilibrium pattern of sorting might arise given that most VSLA lending is done on a limited-liability basis.²⁶ Specifically, if information and ability to monitor borrowers and punish strategic default are much stronger among those with the same occupation than across occupations, then we might expect this to counteract the financial intermediation driver of sorting and instead lead to *positive* assortative matching on occupation. A similar argument would apply if transaction costs are much lower among those with the same occupation compared to across occupations, or indeed if individuals derive much greater social benefit from interacting in a VSLA with others who are engaged in the same occupation. We refer to any of these scenarios as one in which “social frictions” are strong.

Finally, it may be that VSLAs are unable to engage in providing much access to investment credit altogether. This may occur if even the most tightly-knit groups are unable to enforce high rates of loan repayment, or if savers are so risk-averse that they are unwilling to lend out their savings. In this case, members would join VSLAs almost exclusively in order to use them as a savings technology, and to gain the option value of access to emergency loans if these are still available. We would therefore expect to see *negative* assortative matching on occupation if members successfully diversify the risk of such emergency loans within groups, or *zero* assortative matching on occupation if even such idiosyncratic lending is low.

²⁶The member’s own shares can be seized as collateral in the case of non-repayment, but these may not be enough to cover the value of the loan.

3.2 Sorting on time preferences

At the same time, a more subtle type of sorting might also arise based on time preferences, in particular due to the presence of present-biased individuals. We start with the observation that a significant number of individuals who join VSLAs (and Self-Help Groups more generally) likely do so as way to undertake commitment savings (Ksoll et al., 2013). VSLAs can be seen as a multi-faceted commitment savings technology, and may therefore offer an attractive package to individuals who have problems with self-control (or indeed “other-control”, i.e. demands from one’s spouse or relatives) and who are sophisticated enough to recognise these problems and the potential value of commitment.^{27,28} Anecdotally, many members do seem sophisticated in this way: when we asked about reasons for joining, many individuals told us that being in a VSLA “overcomes the temptation of spending savings kept at home”, or “addresses the problems encountered within the household”.

To highlight the multiple commitment features of VSLAs: first, VSLAs offer a withdrawal commitment, since savings made into the box (shares purchased) cannot be liquidated until the end of the savings cycle.²⁹ Second, there is at least *de jure* a deposit commitment, since all members are required to purchase at least one share per week.³⁰ Third, even if this “hard” deposit commitment is sometimes relaxed, the group meetings likely provide other “soft” deposit commitment technologies: active peer pressure to save; self-imposed pressure to save in order to appear reliable to one’s peers (Breza and Chandrasekhar,

²⁷The fact that self-control or other-control problems may generate a demand for commitment savings technologies is well documented, e.g. Ambec and Treich (2007) and Anderson and Baland (2002) in the case of ROSCAs.

²⁸In what follows we focus primarily on self-control problems – i.e. present bias – as a driver of demand for commitment savings. However, in our empirical analysis we also perform analogous tests of sorting along measures of household bargaining, as an attempt to examine possible “other-control” motives.

²⁹There is a provision that individuals can withdraw a small number of their own shares in the case of e.g. a medical emergency. However, in practice we rarely see instances of this when examining the individual account books.

³⁰*De facto* we see from the account books that some groups relax this requirement, allowing individuals to purchase zero shares some weeks. It may be that the groups somehow strike a balance between commitment and liquidity to cover shocks, by relaxing the deposit commitment but keeping the withdrawal commitment strong — allowing access to liquidity out of one’s own shares only very occasionally, and allowing liquidity through borrowing but at a cost. However, we do not attempt to model this here.

2013); and reminders to save (Kast et al., 2012).

Regarding the equilibrium sorting of members across groups in terms of present bias, we focus on three possible scenarios. In the first scenario, present-biased individuals who join as a way to undertake commitment savings sort into groups with time-consistent individuals who join as a way to borrow (*negative* assortative matching on present bias). This is arguably the most interesting scenario from a financial intermediation perspective, since at least in theory it may allow commitment savers' capital to act as a source of credit to borrowers.

The logic of how this pattern of sorting might emerge is as follows. First, aside from commitment savers, it is highly likely that some individuals in these communities are credit constrained and thus join VSLAs primarily as a way to seek credit (Greaney et al., 2013).³¹ Second, even if some of these credit-seeking individuals are present-biased, when we consider all members it is the time-consistent members who are more likely to have joined for borrowing motives (since time-consistent individuals do not have a commitment savings motive for joining).^{32,33} Third, as long as commitment savers are willing to risk lending out their savings to borrowers, then either direct bargaining or sorting mediated through the loan interest rate will logically result in commitment savers and borrowers sorting into groups together (rather than a separating equilibrium). This is because commitment savers have a weakly positive willingness to pay for commitment, and thus are willing to accept a weakly lower interest rate (compared to non-commitment savers or prospective borrowers) when lending out their savings deposits. Thus prospective borrowers who seek out the lowest loan interest rate will end up matched with commitment savers. Furthermore, there may be credit rationing even at the equilibrium loan interest rate, since prospective borrowers within VSLAs make loan requests from a limited pot. Therefore competition for funds even at the equilibrium loan interest rate may further drive prospective borrowers into groups with more commitment savers (rather than more

³¹Indeed, when Ksoll et al. (2013) asked individuals in treated villages in 2011 to report their main reason for joining VSLAs, 50% cited access to credit and 45% cited access to savings (the remaining 5% cited social motives).

³²This is abstracting from "other-control" motives; see footnote 28.

³³One might also expect that a given group is more likely to grant loans to time-consistent rather than present-biased individuals, if groups learn that present-biased individuals are less reliable for repayment.

borrowers), in order to obtain the highest probability of obtaining credit.

Note however that even if all of these forces are at play, we might still observe *zero* assortative matching on present bias if social constraints are simply so strong as to overwhelm sorting along savings and borrowing characteristics *per se*. For example, it may be that the etiquette to match with one’s own social class — or the informational and enforcement advantages of doing so, as described in section 3.1 — outweigh the forces which would otherwise push commitment savers to match with borrowers.

A second possible scenario arises if the credit function of VSLAs becomes negligent in the long run, for reasons as described above in section 3.1. If so, both commitment savers and time-consistent individuals may still join VSLAs, but those time-consistent individuals who do join must also do so primarily for savings motives (although by definition not “commitment” savings). Specifically, there may still be a small probability of earning a positive interest rate on savings, if there is still a small probability that lending will occur to any random member at some point (e.g. for emergency purposes, as also described in section 3.1). Alternatively — commitment motives aside — it may still be worth saving at zero interest if the risk of theft of savings left at home is high, or if there are social benefits to joining a VSLA.

Since in this scenario both types of members join to save, any sorting that occurs is likely to reflect savings motives. Most plausibly, time-consistent savers may be more desirable as fellow group members than present-biased savers, since time-consistent savers are likely to save more reliably and this may produce a positive peer effect on other fellow group members’ savings (e.g. through desire to “keep up with the Joneses”). This would create a “vertical” preference for time-consistent partners (everybody would prefer time-consistent partners to present-biased partners), and as a result we would expect to see *positive* assortative matching on present bias: the most time-consistent (least present-biased) saver would make a group with the second most time-consistent (second least present-biased) saver, and the third, and so on; with a new group then beginning once the first group has reached a maximum feasible membership size.³⁴ Alternatively, if no

³⁴Side payments could theoretically complicate the sorting process, as commitment savers might be

such peer effects exist, then we would expect to see *zero* assortative matching on present bias.

A third and final possible scenario emerges if the credit function of VSLAs becomes negligent *and* the expected interest payments on savings (or the social benefits) are not large enough to attract time-consistent savers. In this case only commitment savers would join VSLAs in the long run.^{35,36} Consequently we would observe *zero* sorting across groups on a binary indicator of individual members' present bias, since all members would be present-biased.³⁷

3.3 Summary of predictions

Table 1 summarises the key predictions on matching in terms of occupation and present bias.

willing to make larger transfers to time-consistent savers in order to obtain their good influence as fellow group members. However, the possibility of sustaining multiple bilateral side payments amongst group members seems unlikely. What appears more likely is that this willingness to pay for time-consistent partners would be reflected in the interest rate on lending, but this takes us back to the first scenario as outlined above.

³⁵One might think of this as analogous to anti-addiction or weight-loss groups, i.e. only those with commitment problems find it beneficial to join.

³⁶The reverse possibility — a scenario in which only prospective borrowers join — seems unlikely, firstly because of the many commitment savings technologies inherent in VSLAs and secondly because there would be intense competition on funds. However, in such a case adverse selection issues as modelled in Greaney et al. (2013) would likely come to the fore, and we would likely observe positive assortative matching on risk aversion or riskiness of investment opportunities (Ghatak and Guinnane, 1999). In our empirical analysis we find no such evidence of sorting on risk preferences.

³⁷We might observe positive assortative matching if we had a more continuous measure of present bias, as it might be that less present-biased individuals are still more attractive partners than more present-biased individuals, similar to the peer effects described in the second scenario.

Table 1: Summary of predictions

Function of VSLAs	Who joins - occupation	Who joins - present bias	Sorting
Commitment saving; possibly regular saving; borrowing	Farmers; non-farmers	Present-biased individuals; non-present-biased individuals	<i>Negative</i> assortative matching on occupation; <i>Negative</i> assortative matching on present bias
Commitment saving; possibly regular saving; borrowing <i>Social frictions strong</i>	Farmers; non-farmers	Present-biased individuals; non-present-biased individuals	<i>Positive</i> assortative matching on occupation; <i>Negative</i> assortative matching on present bias (or <i>zero</i> if social ties overwhelm this)
Commitment saving; regular saving	Farmers; non-farmers	Present-biased individuals; non-present-biased individuals	<i>Negative</i> assortative matching on occupation; <i>Positive</i> assortative matching on present bias (or <i>zero</i> if no benefits of being matched with time-consistent savers)
Commitment saving	Farmers; non-farmers	Present-biased individuals	<i>Zero</i> assortative matching on occupation; <i>Zero</i> assortative matching on present bias (if binary indicator)

4 Data

4.1 Individual member characteristics

Census of members: Table 8 in Appendix A describes some of the key demographic characteristics of the 3,801 individuals in our member census, whilst table 9 (also in Appendix A) provides more detail on the categorical variables of economic activity and schooling. 73% of members report farming as their primary economic activity, whilst 21% work in business (mainly family businesses). 95% have some education (although most have only some primary education) and 83% report being literate (although this is self-reported ability to read and understand a newspaper in Chichewa, and may be over-stated).³⁸ The average age of participants is 36.³⁹

Interestingly, although SOLDEV has no gender-based rules for participation in VSLAs, 75% of the members are female. Furthermore, 21% of members come from female-headed households, which is larger than the average proportion of female-headed households in the local population (16%, as measured from the 2009 panel wave). However, of course this may reflect other correlated factors such as the fact that female-headed households in Malawi are typically poorer, rather than providing *prima facie* evidence that women *per se* have higher demand for VSLAs.

Finally, it should be noted (figures not shown in table) that we observe a large incidence of members joining sometime after the first cycle: specifically 1,262 of the 3,801 members. Some members also drop out or are forced out of groups: 521 of the 3,801 members have left the group at some point by 2013 (102 of whom are in the subsample of matched individuals). 146 individuals are member of more than one group (47 of whom are matched individuals).

³⁸Chichewa is the official national language of Malawi, along with English, and is the language of school instruction nationwide. However, the mother tongue of most individuals in the survey area is Chitumbuka.

³⁹We also asked for ratings of the member's father's wealth and the member's spouse's father's wealth on a scale. These act as proxies of relative exogenous non-income wealth (and inherited social class more generally) to be used as controls in the regression analysis.

Matched individuals: As described above, we successfully matched 722 individuals by name to the 2009-11 panel dataset. Table 10 in Appendix A shows the distribution of matched individuals across groups. We were able to match individuals in all but seven groups, and often to match a sizeable number: on average 4.76 individuals across all groups, or 5.0 conditional on matching at least one member.⁴⁰

Matching to the 2009-11 panel dataset yields a much richer set of data on individual characteristics for the matched individuals, as summarized in table 2. We report values from the 2009 panel wave, since these were measured at baseline before the VSLA project was introduced into the sample area.⁴¹

Crucially, detailed preference modules were administered to a random subset of the panel data households.⁴² In particular, discount rates were elicited using standard list choices: participants were asked whether they would prefer to receive 2000 Kwacha (around \$14)⁴³ now or varying amounts in one month; then whether they would prefer to receive 2000 Kwacha in one year or the same varying amounts in one year and one month.⁴⁴ A higher discount rate in the immediate time frame than in the one-year time frame is taken to indicate present bias, and vice versa to indicate future bias. Overall 11% of individuals are recorded as present biased. This might seem somewhat low; although it is consistent with the proportion found by other studies measuring present bias in rural Malawi using a similar elicitation method (e.g. Bruné et al. (2011) find 10% of individuals to be present biased). Furthermore, we might expect the level of present bias in our survey population to be low given that the survey was enumerated shortly after the harvest and thus in a time of relatively little hunger and stress, which have

⁴⁰The groups with no matched individuals are mechanically dropped from the main specifications; see section 6.2 for further discussion.

⁴¹In a few instances, values were missing from the 2009 wave and thus we report values from the 2010 wave. However, by the time of the 2010 wave only a very small number of the eventual VSLA groups had formed; thus these characteristics are still plausibly exogenous to the characteristics of the other members that an individual eventually ends up in a group with.

⁴²For budgetary reasons, preference modules were not enumerated to all households.

⁴³The exchange rate at the start of the 2009 survey was 145 MWK = 1 USD.

⁴⁴Following Andreoni and Sprenger (2010) we impute discount rates by taking the lowest value of deferred payment in the interval, i.e. the value that makes the individuals seem most patient. This partially mitigates concerns that imputations assuming linear utility tend to upward-bias the measured discount rate.

been shown otherwise to exacerbate impatience and in particular present bias (Ashton, 2014). However we are somewhat concerned that the discount rates measured from the one-year time frame appear particularly high, which may mean that the true number of present-biased individuals is greater than 11% if some individuals' true discount rates in the one-year frame are actually lower than their discount rates as measured in the near frame. We therefore note that our analysis is based on a conservative measure of who is present biased, and may not be capturing some more mildly present-biased individuals. Furthermore, in light of the concern about the one-year-frame discount rates — which would normally be used to proxy the “true” discount rate — we instead use the near-frame discount rates and discount factors in all of the analysis that follows.

We also have measures of the matched individuals' risk aversion, which were elicited using standard Binswanger lotteries. 58% of matched individuals are classified in the two most risk-averse bands, which is perhaps unsurprising given that these are very poor individuals: average monthly consumption per capita is 2177 Malawi Kwacha, which in June 2009 was equivalent to \$15.⁴⁵ There is prima facie evidence of non-cooperative household behaviour, with 44% of matched individuals reporting that they ever hide money from their spouse. Female decision-making power in these households also appears to be low. We construct an index from a series of four questions in the panel survey, which asked who usually makes the decision on topics such as family planning and large expenditures. We score one if the female has some say and two if she has complete control, thus the maximum possible score is eight (although a score of between four and eight may be the most desirable, insofar as it implies female agency on all issues but also shared decision-making). On average, households score just 2.87.

⁴⁵Malawi's GNI per capita in 2009 was \$26.6/month (<http://data.worldbank.org>) but our figure is plausible given that these are particularly poor households in a very remote region.

Table 2: Detailed individual member characteristics, matched subsample

	Matched individuals				
	Mean	Std dev	Min	Max	N
Occupation					
Farmer	0.82	(0.38)	0.0	1.0	722
Businessperson	0.14	(0.35)	0.0	1.0	722
Other	0.03	(0.18)	0.0	1.0	722
Time Preferences					
Present biased	0.11	(0.31)	0.0	1.0	350
Monthly discount factor (near frame)	0.87	(0.12)	0.7	1.0	352
Risk Preferences					
Risk aversion extreme or severe	0.58	(0.49)	0.0	1.0	330
Risk aversion intermediate or moderate	0.31	(0.46)	0.0	1.0	330
Intra-Household Bargaining					
Ever hides money from spouse	0.44	(0.50)	0.0	1.0	307
Female HH decisionmaking power (Index 0-8)	2.87	(1.81)	0.0	8.0	377
Social Variables					
HH important in village decisions (Index 1-6)	3.27	(1.14)	1.0	6.0	721
HH ever speaks at village meetings	0.57	(0.50)	0.0	1.0	718
Demographic Variables					
Male	0.25	(0.44)	0.0	1.0	722
Female-headed household	0.23	(0.42)	0.0	1.0	722
Age	38.13	(11.44)	18.0	83.0	722
Education					
Some primary educ.	0.80	(0.40)	0.0	1.0	722
Some post-primary educ.	0.15	(0.36)	0.0	1.0	722
Literate (read & understand newspaper)	0.81	(0.40)	0.0	1.0	722
Wealth					
Father well-off in village (Index 1-5)	3.29	(1.31)	1.0	5.0	679
Spouse's father well-off in village (Index 1-5)	3.38	(1.31)	1.0	5.0	707
Income					
Monthly consumption per capita, MK	2176.95	(973.09)	648.7	7811.1	721
Food security poor (dummy)	0.28	(0.45)	0.0	1.0	722

4.2 Group characteristics

Table 3 describes the distribution of groups across villages. Thirty-five villages have at least two groups.⁴⁶ It is interesting to note that the presence of more than one group per village is in a sense *prima facie* evidence of inefficiency in the level of financial intermediation created by VSLAs: given that there is no secondary market for capital in these villages, VSLAs with excess capital cannot lend to other VSLAs; thus the most efficient arrangement in terms maximizing the possibility of capital reallocation across individuals (and probably also in terms of mitigating risk) would be to have just one large VSLA per village. However, issues such as trust and ability to monitor loans, and indeed transaction costs and quorum requirements for VSLA meetings, likely become binding constraints before a VSLA grows to the size of a whole-village institution.⁴⁷ This therefore motivates our analysis of sorting, as a key driver of efficiency in this “second-best” scenario with more than one VSLA per village. Indeed, the scope for sorting across groups appears to be considerable, with some villages having up to 14 groups.

Table 14 in Appendix A describes how groups are composed in terms of member characteristics. The average group size is 25 members, although groups range in size from 10 to 45 members. Groups also range in gender composition from all-male to all-female, although most groups (i.e. groups within one standard deviation of the mean) are mixed but with a majority of female members. In some groups, as many as 62% of members come from female-headed households. There is also clear heterogeneity across groups in terms of occupational composition: some groups consist purely of farmers, whereas others contain almost no farmers. However, only dyadic regression analysis can determine whether such heterogeneity is evidence of individuals sorting across groups within villages, or whether it represents differences in population characteristics across villages (see section 5).

⁴⁶Nine villages have just one group, and thus sorting is not identified in these villages when we include village fixed effects (see section 5).

⁴⁷Indeed, anecdotally we were told that one VSLA of around 45 members had to split into two groups, because the members found it too difficult to reach consensus decisions in such a large group.

Table 3: Number of VSLA groups per village

# groups in village	# villages	% of villages
1	9	20.5%
2	17	38.6%
3	4	9.1%
4	4	9.1%
5	3	6.8%
6	2	4.6%
7	2	4.6%
11	1	2.3%
13	1	2.3%
14	1	2.3%
Total	44	100%

From our VSLA group survey conducted in July-August 2013. VSLA training was offered in 46 villages: 23 villages in 2009 and 23 villages in 2011. However, no groups were formed in two of the 2011 wave of villages; most likely due to logistical problems as these two villages are extremely remote.

4.3 Use of savings and loan funds

To give a broad picture of the kinds of financial needs that individuals are meeting by participating in VSLAs, figure 3 describes the most important uses of saving and loan funds as reported by the group during the group survey.⁴⁸ The patterns of savings funds usage compared to loan fund usage appear quite distinct. The most-reported use of savings is for agricultural inputs (and some agricultural investments), which is logical since the groups all choose to share out in January which is during the planting season.⁴⁹ January is also the lean season, which may explain why savings are also reportedly used by over 40% of groups to buy food. The other most prominent uses of savings are livestock, household items (typically durables like kitchenware) and building materials. There is also some use of savings for trading/business, although much smaller than loan use.

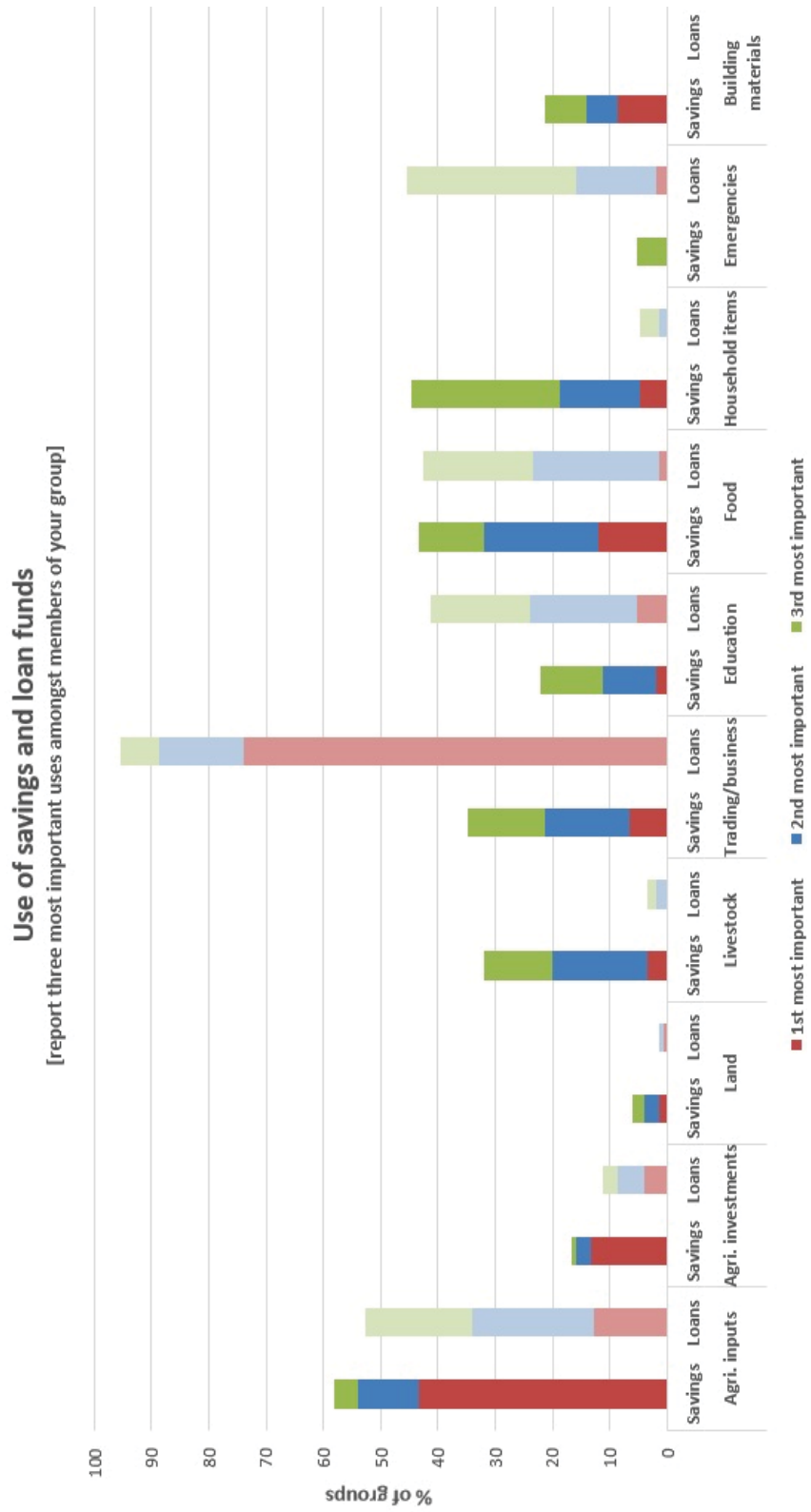
⁴⁸“Most important” was explained to the groups as meaning broadly the most common uses, but also with some weighting by the amount of money that each use typically involved.

⁴⁹This choice of timing itself may reflect sophistication about the effect of having cash-in-hand just before input purchase is required (Duflo et al., 2011; Bruné et al., 2011).

Loans on the other hand are highly concentrated on trading/business purposes: 74% of groups say this is the most important use of loans, and altogether 95% say this is amongst the three most important uses of loan funds. Loans are also used for agricultural inputs, perhaps as a complement to savings or perhaps for different inputs which require purchase at times other than January. Loans are used more than savings to cover education expenses, which may reflect the fact that school fees are typically paid at the start of the school year in September rather than closer to the share-out in January. Perhaps unsurprisingly given that VSLA savings are illiquid until the end of the cycle, we see that households ostensibly use loans to deal with consumption shocks: over 40% of groups also report using loans to buy food, and for emergencies.

The overall picture is that there does appear to be some financial intermediation taking place *within* groups: some members save for the planting (and lean) season, and meanwhile their savings are lent out to members who engage in a degree of trading/business activities. The question we now examine with the dyadic regression analysis is whether there is intermediation on a larger scale at the village level: specifically, whether commitment savers/those employed in farming sort into groups with borrowers/those employed in business; or whether any financial intermediation is at best only conditional on being in a group of all-commitment-savers or all-farmers.

Figure 3: Reported use of savings and loan funds from VSLAs



5 Empirical strategy

5.1 Dyadic regression framework

To test the key pair-level predictions from section 3.3, we employ a dyadic regression framework. The intuition behind this approach is as follows: if there are multiple groups in a village, and if there is positive (negative) sorting on a given characteristic, then in equilibrium two members who are more similar on that characteristic are, *ceteris paribus*, more (less) likely to be observed as members of the *same* group. In order to distinguish between the scenarios outlined above, we therefore examine whether individuals who are more similar in terms of occupation and present bias are, *ceteris paribus*, more (or less) likely to be observed in the same group in 2013. We take this as evidence of positive (or negative) sorting on occupation and present bias.

To implement this, we first construct all possible pairs – dyads – of individuals in the 2013 member census, conditional on both individuals being members of groups in the same village.⁵⁰ Table 11 in Appendix A describes features of the universe of such dyads (including dyads which contain unmatched as well as matched members). When we restrict attention to the dyads in which both individuals are matched to the 2009-11 panel dataset, this gives us a sample size of 7,326 dyads. Of these, 1,641 are dyads in which both individuals are matched to the full version of the 2009-11 survey, which included the detailed preference modules.⁵¹

Table 4 highlights some of the key characteristics of the matched dyads, in terms of the absolute differences and the sums of i and j 's characteristics (tables 12 and 13 in Appendix A describe the characteristics of the matched dyads in full). Note that on average 30%

⁵⁰In practice we found it to be extremely rare that an individual would join a group outside of his/her village of residence; thus *de facto* only the other individual members from an individual's village of residence are candidates to be members of the same group as that individual.

⁵¹The dyads which can be matched to the full version of the panel survey are a slightly unrepresentative sample of the whole universe of dyads from the 2013 member census, as table 15 in Appendix A make clear. Therefore, we later re-run all of our main specifications weighting each dyad by the inverse probability of that dyad being matched to the full 2009-11 survey. This does not change our results (see section 6.2).

of dyads comprise individuals who are members of the same group, implying that 70% comprise individuals who are members of different VSLA groups within the same village.

Table 4: Key dyadic regressors, matched subsample

	Key dyad characteristics - matched dyads				
	Mean	Std dev	Min	Max	N
Membership					
Same VSLA group (dummy)	0.30	(0.46)	0.0	1.0	7326
Absolute differences					
Same economic activity (dummy)	0.70	(0.46)	0.0	1.0	7326
Farmer	0.27	(0.44)	0.0	1.0	7326
Businessperson	0.23	(0.42)	0.0	1.0	7326
Present biased	0.16	(0.37)	0.0	1.0	1641
Monthly discount factor (near frame)	0.12	(0.11)	0.0	0.3	1655
Risk aversion extreme or severe	0.43	(0.49)	0.0	1.0	1513
Risk aversion intermediate or moderate	0.39	(0.49)	0.0	1.0	1513
Ever hides money from spouse	0.45	(0.50)	0.0	1.0	1269
Female HH decisionmaking power (Index 0-8)	1.95	(1.57)	0.0	8.0	1914
HH important in village decisions (Index 1-6)	1.08	(1.11)	0.0	5.0	7314
HH speaks at village meetings	0.47	(0.50)	0.0	1.0	7266
Sums					
Farmer	1.61	(0.60)	0.0	2.0	7326
Businessperson	0.32	(0.55)	0.0	2.0	7326
Present biased	0.19	(0.42)	0.0	2.0	1641
Monthly discount factor (near frame)	1.70	(0.18)	1.4	2.0	1655
Risk aversion extreme or severe	1.18	(0.74)	0.0	2.0	1513
Risk aversion intermediate or moderate	0.60	(0.67)	0.0	2.0	1513
Ever hides money from spouse	0.79	(0.71)	0.0	2.0	1269
Female HH decisionmaking power (Index 0-8)	5.90	(2.71)	0.0	16.0	1914
HH important in village decisions (Index 1-6)	6.54	(1.69)	2.0	12.0	7314
HH ever speaks at village meetings	1.16	(0.71)	0.0	2.0	7266

Our main estimating equations are undirected dyadic logit models with observations at the dyad level. These take the following form:

$$\begin{aligned} \Pr(D_{ij} = 1 | D_i = 1 \ \& \ D_j = 1; \mathbf{Z}_i, \mathbf{Z}_j, \mathbf{W}_{ij}, v) \\ = \Pr(\alpha + \beta|\mathbf{Z}_i - \mathbf{Z}_j| + \gamma(\mathbf{Z}_i + \mathbf{Z}_j) + \delta\mathbf{W}_{ij} + \mu_v + \varepsilon_{ij} > 0) \quad (1) \end{aligned}$$

where D_i and D_j denote dummies equal to one if i and j respectively are members of some VSLA group,⁵² and D_{ij} is a dummy equal to one if i and j are members of the *same* group. \mathbf{Z}_i and \mathbf{Z}_j are vectors of i 's and j 's individual characteristics respectively, which in our central specifications include measures of present-bias and other aspects of time and risk preferences. \mathbf{W}_{ij} is a vector of characteristics of the dyad (such as whether i and j share the same category of occupation, when occupations are listed in full rather than in the form of a dummy variable for farm vs. non-farm employment). μ_v is a village fixed effect. ε_{ij} is a dyad-specific error term, which we assume takes a logistic distribution. When estimating equation 1, we cluster standard errors at the village level.^{53,54}

It follows from the logic outlined above that an estimate of $\hat{\beta} < 0$ indicates *positive* assortative matching on the characteristic in question; an estimate of $\hat{\beta} > 0$ indicates *negative* assortative matching on that characteristic; and an estimate of $\hat{\beta} = 0$ indicates random matching *ceteris paribus* (i.e. an absence of evidence of sorting on that characteristic).

Since we estimate equation 1 on a sample which only includes individuals who are members of *at least one* group, an estimate of $\hat{\gamma} > 0$ (< 0) indicates that conditional on being member of at least one group, individuals with a high value of that particular variable are more (less) likely to be members of *more than one* group. This increases (decreases) the probability that such individuals are in the same group as a randomly-chosen other member, simply because such individuals are members of more (fewer) groups.

5.2 Identification

A direct test between the scenarios outlined in section 3.3 is possible if we include a dummy variable for i and j having the same occupation into \mathbf{W}_{ij} , and the absolute difference in i and j 's present bias into $|\mathbf{Z}_i - \mathbf{Z}_j|$, and examine the resulting values of

⁵²This is for notational completeness, as by construction both dummies will always be equal to one in our analysis since our data contains only members.

⁵³This is more flexible than the more common method of clustering by dyad (Fafchamps and Gubert, 2007).

⁵⁴For notational convenience, we omit v subscripts from all regressors except the fixed effects.

$\hat{\beta}$.^{55,56} Data on occupation is available for all matched individuals, and a proxy of present bias as described in section 4 is available for the 1,641 dyads matched to the full 2009-11 survey. Importantly this proxy and indeed the other regressors are taken from the 2009 wave and thus were measured before VSLA projects were introduced to the sample area. We therefore eliminate concerns about reverse causation, which might arise if being members of the same group leads individuals to have a greater difference or similarity in their measured characteristics (e.g. if group activities lead to correlation in individuals' incomes, Fafchamps and Ferrara (2011)).

We also seek to minimize omitted variable bias by controlling for a rich set of characteristics which might also drive sorting and might be correlated with present bias: for example, risk preferences, gender and income. Section 6 details the full set of controls used in each specification. The village fixed effects control for the average probability of matching in the village, which depends on the number of groups and also on the relative size of each group: for example, if there is one large group and one small group then the probability of matching is high since most members are both in the large group. The inclusion of village fixed effects also absorbs a range of factors which might affect the probability of being in the same group but remain constant at the village level, for example whether the village is served by other NGO programs.

In sum therefore, when testing the coefficients on the dummy for i and j sharing the occupation and on the absolute difference in i 's and j 's present bias, we make the identifying assumption that after the inclusion of a rich set of controls and village fixed effects, there are no further unobserved sources of heterogeneity at the dyad level which are correlated both with the difference between i and j 's present bias in 2009 and with their propensity to be a members of the same group by 2013.

⁵⁵Instead of the dummy for sharing the same occupation, we also run alternative specifications with the absolute difference in i and j 's dummy variable for being a farmer. This does not change our results.

⁵⁶Strictly speaking, in order to proxy a demand for commitment savings we need to proxy both present bias *and sophistication*. However, given that the present-biased individuals we observe have already joined VSLAs, we assume that they are already sophisticated enough to have understood the value of VSLAs as a commitment savings technology; see section 3.

6 Results

6.1 Dyadic regressions

Table 5 outlines the results of our preferred specification in full.⁵⁷ Firstly we see a large, positive coefficient on i and j both having the same primary source of employment, which increases the probability of being in the same group by 16.8 percentage points (*positive* assortative matching). This is equivalent to 57% of the baseline probability of two members being in the same group, conditional on both being members of some VSLA group in the same village (29.6%). Secondly there is a similarly large, positive coefficient on the absolute difference between i 's and j 's present bias in determining the probability that i and j are members of the same group (*negative* assortative matching). Taken together, these coefficients suggest that we are observing sorting in line with the second scenario in table 1: commitment savers are sorting into groups with borrowers, but social frictions dominate sorting on occupation.

The magnitude of the coefficient on present bias is very similar to the estimated positive assortative matching on occupation: if member i in village v is present-biased and member j also in village v is not present-biased (or vice versa), this increases the probability that i and j are members of the *same* group by around 16.5 percentage points. However, as explained in section 4 we do not place too much weight on the exact coefficient value here, given that we have reason to believe that our estimate of the number of individuals who are present biased may be conservative. Relatedly, the coefficient on the monthly discount factor does not come out as significant, but we caution against over-interpreting this since we are using the near-frame monthly discount factor (given the perceived unreliability of the far-frame monthly discount factor) and this may be correlated with present bias.

The coefficients on the absolute differences in various controls also indicate further patterns of sorting, which may be taken as indicative of the kinds of social frictions described

⁵⁷In our main analysis as we consider all individuals who have ever been a member of the group. Section 6.2 also details specifications considering just current members.

in section 3 (and may therefore partially explain why individuals appear to assort positively rather than negatively on occupation). Perhaps unsurprisingly, we observe positive assortative matching on proxies of exogenous non-income wealth (or social class more generally): specifically, we observe positive assortative matching on whether a member's spouse's father is well off (marginal effect size of 5.1 pp, significant at the 1% level), and a smaller effect of a member's own father being well-off (marginal effect size of 2.6 pp, significant at the 5% level). The fact that a member's spousal family wealth appears to play a greater role than the member's own family wealth may reflect the fact that most VSLA members are female, and given the patrilineal nature of the local society it is generally the male's family wealth which is a stronger determinant of a household's own position. The significant coefficients on wealth contrast with an insignificant coefficient on consumption, although this may reflect the fact that consumption is likely measured with much more noise.

There is also evidence of positive assortative matching on age (marginal effect size of 0.7 pp per year's difference, significant at the 1% level) which may reflect social proximity. Finally, there is positive assortative matching on being a female-headed household (marginal effect size of 17.6 pp, significant at the 5% level), which may reflect attempts at solidarity or female-headed households having different financial needs to male-headed households; or may indicate that female-headed households find it difficult to be accepted into other groups. Perhaps surprisingly though, there is no evidence of sorting on gender once these other factors have been controlled for. Overall, our controls are able to explain around 21% of the observed variation in the probability of two members being in the same group (pseudo R-squared of 0.213).

The sum terms are all insignificant, likely reflecting our lack of power to identify these coefficients since only 42 matched individuals are member of more than one group.⁵⁸ However, we still include the sum terms as controls in our preferred specification, since excluding them could lead to omitted variable bias if they do have a small effect and if the size of certain individual characteristics is correlated with their variance (meaning the sum $Z_i + Z_j$ is correlated with the absolute difference $|Z_i - Z_j|$).

⁵⁸146 of the full 3,801 member census individuals are member of more than one group.

Table 5: Preferred specification – matched subsample

	(1)	
	β / (s.e.)	Mfx
Dependent variable: $samegroup_{ij} = 1$		
Absolute differences		
Same economic activity (dummy)	0.924** (0.419)	0.168**
Present biased	0.907** (0.417)	0.165**
Monthly discount factor (near frame)	-0.408 (0.798)	-0.074
Male	-0.009 (0.193)	-0.002
Female-headed household	-0.969** (0.406)	-0.176**
Age	-0.039*** (0.014)	-0.007***
Schooling level	-0.022 (0.129)	-0.004
Literate (read & understand newspaper)	-0.224 (0.259)	-0.041
Father well-off in village (Index 1-5)	-0.140** (0.071)	-0.026**
Spouse's father well-off in village (Index 1-5)	-0.282*** (0.086)	-0.051***
Monthly consumption per capita, 1000 MK	-0.232 (0.151)	-0.042
Food security poor (dummy)	-0.048 (0.130)	-0.009
Sums of all the above controls	✓	
Observations	1292	
Pseudo R^2	0.213	
Baseline predicted probability	0.296	

6.2 Robustness

Additional controls: The coefficients on economic activity and present bias are robust to the inclusion of a host of controls as shown in table 6. Note that column (2) of table 6 corresponds to our preferred specification with full controls for time preferences, as shown in full in table 5.

Comparing columns (1)-(2), as mentioned above, the absolute difference in i and j 's monthly discount factor does not show up as significant, and does not significantly change the coefficients on whether i and j share the same economic activity and the absolute difference their in present bias. Turning to column (3), the absolute difference in risk aversion bears a negative but small and insignificant coefficient. The coefficient on same economic activity if anything increases when risk aversion is controlled for. The coefficient on present bias does drop slightly (although this is not significantly different from (2), given the large standard errors) and loses one significance level; however, rather than reflecting an omitted variable bias in (2), this may instead reflect the smaller sample size in (3), since measures of risk-aversion are missing for some of the matched individuals.

Looking at column (4), the coefficients on the intra-household bargaining measures (whether the participant ever hides money from their spouse, and the measure of female household decision-making power) are all very small and insignificant. This suggests that self-control (present bias) rather than “other-control” is indeed the driver of commitment savings that is important for sorting. When the household bargaining controls are included, the coefficient on same economic activity drops very slightly compared to (2) and once again the coefficient on present bias also drops compared to (2), however, neither change is significant given the large standard errors. Both coefficients become significant at only the 10% level; but this is likely because the sample size becomes much smaller in column (4), mainly because the intra-household questions were only asked of currently married individuals.

In column (5), one further control which does show up as significant is whether the individual comes from a household that speaks at village meetings — an indicator of how

“active” the household is in local civil society. The absolute difference in this variable has a positive and highly significant coefficient, indicative of negative assortative matching. This makes intuitive sense, as it is likely that certain “leader” households or individuals encourage other “followers” to join within each group. However, such an effect does not appear to have been confounding the coefficients on same economic activity and present bias in (2), as these remain large, positive and significant (and not significantly different from (2)). Again the lower level of significance on both coefficients likely reflects sample size issues. Overall we therefore prefer specification (2), since it is estimated on a much larger sample size.

Weighting: The 2009-11 survey was enumerated on a random sample of the population in each of the villages in the survey area.⁵⁹ However, there is still a potential sampling bias from estimating our main specifications on only those 2013 dyads that can be matched back to the full 2009-11 survey, if these are not a random subsample of all 2013 dyads. This may occur because the 2013 members census population is inherently different from the 2009-11 sample, insofar as the 2013 population consists of members only. For example, whilst 50% of the respondents in the representative 2009-11 survey were male, only 25% of members in our 2013 member census are male; thus from the perspective of our 2013 census population, we are disproportionately likely to match a male member back to the 2009-11 survey compared to a female member. Table 15 shows that the matched dyads are indeed slightly unrepresentative of the full population of 2013 dyads.

To correct for this, we re-run all of our central specifications weighting observations by the inverse probability of that particular dyad being matched to the full 2009-11 survey. To construct the weights, we first estimate a probit equation for the probability of each dyad in the 2013 data also appearing in 2009 data, as function of the full set of dyad characteristics listed in table 11 and village fixed effects (results available on request).⁶⁰

⁵⁹The sampling was stratified insofar as households who expressed an interest in joining VSLAs at the initial NGO awareness-raising meeting were over-sampled. For the purposes of our matching, this only increases the overall probability that we are able to match any randomly-chosen member of our 2013 member census back to the 2009 data.

⁶⁰This is not exactly equivalent to the the product of the separate probabilities that individual i and individual j are matched to the 2009 data, because of differences in the number of dyads across villages

We then use the estimated coefficients to generate the predicted probability that each 2013 dyad is matched to the 2009 data, and take the inverse of this predicted probability as the dyad-specific weight. We then include these weights into the various estimations of 1 — estimated as before on the sample of dyads matched to the 2009 data — in order to recover the coefficients that hypothetically would have been obtained if the matched sample had been fully representative of the dyads in the 2013 population (results available on request).⁶¹ Re-weighting does not change the signs or the significance of the coefficients in the preferred specification; nor does it change the signs, significance and patterns of coefficient movements across the various sets of controls in table 6, except that the coefficients on occupation and present bias become marginally insignificant in column (4). Across most of the specifications, the coefficients on having the same economic activity and the absolute difference in present bias actually increase after the inclusion of sampling weights.

Other robustness checks: We also re-run the main specifications restricting the sample to all-female dyads, since there may be different patterns of sorting across genders. Specifically, the fact that only 25% of members are male may indicate differential patterns of participation across genders, which may in turn influence sorting. Moreover, the effect in particular of the household bargaining measures, especially the female empowerment measure, may logically be different across genders. Reassuringly, the results from our preferred specification (2) remain unchanged in terms of the sign, size and significance of all the coefficients (results available on request). Unfortunately when we run specification (4) with only all-female dyads we run into issues of sample size, since the number of observations falls to just 380. As a result, the coefficients on age and spouse’s father’s wealth retain their significance but the other coefficients — whilst retaining the same sign — drop somewhat in size and lose their significance.

We also re-run all of our specifications restricting the sample to dyads in which neither

and groups.

⁶¹To our knowledge this is the first use of inverse probability weighting in a dyadic specification. However, our proposed technique would appear valid since inverse probability weighting is valid for all M-estimators, of which logit models are a type.

member has ever left the group. The sign and pattern of all coefficients remains the same (results available on request). However, we again face issues of smaller sample size as we lose around 200 dyadic observations. As a result, the coefficient on present bias is only significant (at the 10% level) for the specifications corresponding to (1) and (2) from table 6, and the coefficient on economic activity becomes insignificant across all specifications (1)-(5) although it is still positive and large in magnitude.

Full 2013 member census: Finally, to supplement our main analysis we briefly examine patterns of sorting amongst the 3,801 individuals who comprise the full population of the 2013 member census. Given that we enumerated the census to the whole group simultaneously, and in some cases relied on the group members who were present to answer questions about others who were absent, we were restricted to asking about a reasonably small set of individual characteristics which could be accurately reported by an individual's fellow group members. We therefore did not attempt to measure preferences, and so cannot offer a test of the predictions on present bias from section 3.3. However, it is nonetheless interesting to see whether sorting on the characteristics that we do measure for all individuals is broadly in line with the patterns observed for matched individuals.

Table 7 describes the results. Reassuringly, the sign and significance of the coefficient on having the same economic activity matches that estimated on the subsample of matched dyads, as do the coefficients on the absolute difference in own father's wealth, spouse's father's wealth, being from a female-headed household and age all. Gender and literacy also appear to display positive assortative matching patterns, which were not observed in the specifications estimated on the subsample of matched dyads. However, this is likely because we have far more dyadic observations here, and a far less demanding specification with regard to the number of controls. Indeed, all of these estimates should be treated with caution, since the included variables may of course in part be proxying the effect of preference variables which we are unable to control for in this full sample.⁶²

⁶²The number of goats and bicycles a member's household possesses proxy asset wealth. We also asked groups to rank their members in terms of how well-off they are. Given that this is a within-group ranking, we would expect its coefficient to be biased towards a positive value. The negative coefficient therefore suggests that individuals understood this question to be more about absolute consumption.

Table 6: Dyadic regressions – matched subsample

	(1)	(2)	(3)	(4)	(5)
Dependent variable: $samegroup_{ij} = 1$	β / (s.e.)	β / (s.e.)	β / (s.e.)	β / (s.e.)	β / (s.e.)
Absolute differences					
Same economic activity (dummy)	0.941** (0.407)	0.924** (0.419)	1.093** (0.528)	0.902* (0.541)	0.957* (0.519)
Present biased	0.901** (0.418)	0.907** (0.417)	0.608* (0.346)	0.776* (0.463)	0.789* (0.415)
Monthly discount factor (near frame)		-0.408 (0.798)	-0.935 (0.895)	-1.504 (1.101)	-1.350 (1.081)
Risk aversion			-0.064 (0.082)	-0.144 (0.097)	-0.114 (0.095)
Ever hides money from spouse				-0.038 (0.227)	-0.080 (0.271)
Female HH decisionmaking power (Index 0-8)				0.060 (0.079)	0.081 (0.075)
HH important in village decisions (Index 1-6)					-0.243 (0.161)
HH speaks at village meetings					0.423*** (0.146)
Wealth & income controls (abs. diffs)	✓	✓	✓	✓	✓
Demographic controls (abs. diffs)	✓	✓	✓	✓	✓
Sums of all the above controls					
Observations	1292	1292	1031	744	731
Pseudo R^2	0.213	0.213	0.210	0.215	0.225
Baseline predicted probability	0.296	0.296	0.285	0.292	0.287

Table 7: Dyadic regressions – full 2013 member census

	(1)	
	β / (s.e.)	Mfx
Dependent variable: $samegroup_{ij} = 1$		
Dyads - Absolute differences		
Same economic activity (dummy)	0.767*** (0.070)	0.0857***
Male	-0.474*** (0.134)	-0.0530***
Female-headed household	-0.104*** (0.039)	-0.0116***
Age	-0.007*** (0.001)	-0.0008***
Schooling level	-0.003 (0.017)	-0.0003
Literate (read & understand newspaper)	-0.120** (0.061)	-0.0134**
Father well-off in village (Index 1-5)	-0.114*** (0.014)	-0.0127***
Spouse's father well-off in village (Index 1-5)	-0.114*** (0.027)	-0.0127***
Household well-off in group (Index 1-9)	-0.086*** (0.022)	-0.0096***
HH owns a bicycle	-0.027 (0.030)	-0.0030
Goats	-0.049*** (0.013)	-0.0055***
Dyads - Sums		
Farmer	-0.370*** (0.034)	-0.0414***
Male	0.223*** (0.059)	0.0249***
Female-headed household	0.048* (0.029)	0.0054*
Age	0.003*** (0.001)	0.0004***
Schooling level	0.020* (0.011)	0.0022*
Literate (read & understand newspaper)	-0.049 (0.037)	-0.0055
Father well-off in village (Index 1-5)	-0.006 (0.005)	-0.0007
Spouse's father well-off in village (Index 1-5)	-0.001 (0.007)	-0.0002
Household well-off in group (Index 1-9)	-0.021*** (0.007)	-0.0024***
HH owns a bicycle	-0.016* (0.009)	-0.0018*
Goats	0.027*** (0.009)	0.0030***
Observations	219747	
Pseudo R^2	0.129	
LR χ^2	.	
Prob > χ^2	.	
Baseline predicted probability	0.161	

6.3 Discussion

Our regression estimates therefore provide evidence in line with the second scenario described in table 1. In particular, there is strong evidence of positive assortative matching on occupation, which suggests that potential gains from financial intermediation may be lost because of social frictions such as the inability to monitor or enforce loans lent to individuals in other occupations, or high transaction costs between individuals from different occupations. Furthermore, in terms of risk, positive assortative matching on occupation implies that these groups may be heavily exposed to common shocks affecting the amount that individuals members are able to contribute, demands for emergency loans, and also the broader loan default risk. There is also little evidence of sorting on individual risk preferences, which may be a key component of efficiency in this context since most investments made through borrowing from VSLAs are likely to be uninsured. Furthermore we see positive assortative matching on wealth, which may reduce the scope for VSLAs to allow capital reallocation across village members since it implies a tendency for those with the most capital to group together, rather than grouping with those who may be more capital constrained. This may raise concerns about equity; although we cannot say whether it is *prima facie* evidence of inefficiency, since it is also plausible that those with the most capital still have the highest returns to capital if they also have access to better investment opportunities.⁶³

On the other hand, we do observe negative assortative matching on present bias, which we take as evidence that individuals appear to be sorting across VSLAs such that present-biased commitment savers match with time-consistent individuals including prospective borrowers. As explained, this may be promising in terms of capital allocation, since it effectively brings together those who want to save with those who seek to borrow. However, as noted, we caution against placing too much weight on the exact coefficients estimated on the present bias term, given the aforementioned concerns that our results are based on a conservative estimate of the number of individuals who are truly present-biased.

⁶³A more thorough analysis of efficiency would also of course need to examine the returns of the projects typically funded by borrowing from VSLAs.

If commitment savers are indeed matching with borrowers, NGOs and policymakers might still have a separate concern about equity if present-biased individuals' willingness to pay for commitment leads them to accept very low interest rates when lending out their savings to borrowers. Of course, welfare statements are difficult to make when present-biased individuals are involved, and there is nothing exploitative *per se* in allowing sophisticated present-biased individuals to pay for commitment (in this case by accepting a lower interest rate).⁶⁴ It is also not clear what commitment savers' outside option would be in the absence of VSLAs, and thus what the correct counterfactual is in assessing whether commitment savers are receiving "too low" an interest rate. However, if savers were receiving a return of close to zero or even negative values, or were experiencing high rates of default from borrowers without apparently being compensated for that risk through a high loan interest rate, then intuitively this would be grounds for concern.

Reassuringly, we see from the group survey that in the 2013 cycle, the average monthly interest rate on lending set by groups was 17%. This is in line with the most natural benchmark for a "fair" interest rate that we can identify: namely, individuals' average monthly discount rate as measured in the 2009 preference modules, which comes out at 13.2%.^{65,66} It therefore seems that commitment savers are not being "exploited" within

⁶⁴Heidhues and Kőszegi (2010) show that present-biased individuals who are only partially sophisticated may be exploited through penalties for failure to meet mortgage payments, since lenders may take advantage of the fact that these individuals underestimate their own degree of present bias. However, in the context of VSLAs such exploitation of partial naiveté is less of a concern, since there are no penalties for failing to meet the minimum share-purchase requirement each week.

⁶⁵We take the average discount rate as measured from the near frame (comparing now to one month from now), since as discussed we are concerned about the reliability of the measurement of choices in the one-year time-frame. Given that we use the near time frame, which may be subject to additional discounting by individuals with present bias, it may therefore be that this overstates individuals' "true" discount rate. However, this would only point to the interest rate on loans being even "fairer" in terms of being above the average discount rate.

⁶⁶Inflation in 2009 vs. 2013 should of course be taken into account in order to compare these rates in real terms. Inflation in Malawi was fairly stable at around 8% y-o-y (corresponding 0.64% per month) from the beginning of 2009 until the beginning of 2012. However, after a devaluation of the Kwacha by 33% in May 2012, inflation spiked and ran at an average of 28% in 2013 overall (corresponding to 2.1% per month). Groups do not appear to have taken this into account in the nominal loan interest rate set at the beginning of the 2013 cycle, which in most groups remained unchanged from previous cycles. However, even the high 2013 inflation rate is still negligible compared to such a high monthly loan interest rate, and therefore accounting for inflation does not alter our conclusion that the interest rate on loans appears to be "fair": the real monthly interest rate on loans in 2013 is around 17% - 2.1% = 14.9%, whilst the real discount rate as measured in 2009 is around 13.2% - 0.64% = 12.56%.

VSLAs in this sense.⁶⁷

7 Conclusion

Overall, our main contribution in this paper has been to highlight the potential role of VSLA-type groups for financial intermediation, especially given their features as a commitment savings technology. In light of this, we have examined the important question of how individuals might sort across such groups. This issue has hitherto been overlooked in the literature on VSLA-type groups, but is likely to be crucial in determining the efficiency and equity of the financial outcomes that they generate for members.

The unique data we have collected — a census of all members of a VSLA project in rural Malawi — has allowed us to perform novel and detailed tests of sorting. We find strong results of positive assortative matching on occupation, suggesting that the potential gains from financial intermediation may be unrealised due to social frictions. However, there is tentative evidence of negative assortative matching on present bias, which we interpret as implying that *ceteris paribus* commitment savers are sorting into groups with time-consistent borrowers. In terms of equity, the interest rate that commitment savers are able to charge when lending out their savings to borrowers appears to be “fair” insofar as it is commensurate with the average discount rates of individuals in the survey area.

Going beyond these results, more detailed welfare analysis is needed in order to ascertain whether VSLA-type groups do increase the efficiency of capital allocation within villages, and exactly how the benefits of VSLA-type groups are spread across their members. We hope to address some of these questions in future work, drawing on other aspects of the data we have collected.

⁶⁷It is also theoretically possible that non-commitment savers might engage in speculative behaviour, if they seek to “pile in” towards the end of the savings cycle in an attempt to suck out any profits from lending up until that point. However, the scope for such behaviour is limited by the rule of a maximum share purchase of five shares per week.

References

- C. Ahlin. Matching for Credit: Risk and Diversification in Thai Microcredit Groups. Technical report, BREAD Working Paper, 2009.
- S. Ambec and N. Treich. ROSCAs as Financial Agreements to Cope with Self-control Problems. *Journal of development economics*, 82(1):120–137, 2007.
- S. Anderson and J.-M. Baland. The Economics of ROSCAs and Intrahousehold Resource Allocation. *The Quarterly Journal of Economics*, 117(3):963–995, 2002.
- J. Andreoni and C. Sprenger. Estimating Time Preferences from Convex Budgets. Technical report, National Bureau of Economic Research, 2010.
- J.-L. Arcand and M. Fafchamps. Matching in Community-Based Organizations. *Journal of Development Economics*, 98(2):203–219, 2012.
- N. Ashraf, D. Karlan, and W. Yin. Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines. *The Quarterly Journal of Economics*, 121(2):635–672, 2006.
- L. Ashton. Hunger games: Does hunger and cognitive fatigue affect time preferences? Available at SSRN 2538740, 2014.
- O. Attanasio, A. Barr, J. C. Cardenas, G. Genicot, and C. Meghir. Risk Pooling, Risk Preferences, and Social Networks. *American Economic Journal: Applied Economics*, 4(2):134–67, April 2012.
- A. Barr, M. Dekker, and M. Fafchamps. Who Shares Risk with Whom under Different Enforcement Mechanisms? *Economic Development and Cultural Change*, 60(4):677–706, 2012.
- M. Bauer, J. Chytilova, and J. Morduch. Behavioral Foundations of Microcredit: Experimental and Survey Evidence from Rural India. *American Economic Review*, 102(2):1118–39, April 2012.

- L. A. Beaman, D. Karlan, and B. Thuysbaert. Saving for a (not so) rainy day: A randomized evaluation of savings groups in mali. *Yale University Economic Growth Center Discussion Paper*, (1043), 2014.
- E. Breza and A. G. Chandrasekhar. Savings monitors. *Columbia Business School*, 2013.
- L. Bruné, X. Gine, J. Goldberg, and D. Yang. Commitments to Save : a Field Experiment in Rural Malawi. Policy Research Working Paper Series 5748, The World Bank, Aug. 2011.
- E. Duflo, M. Kremer, and J. Robinson. Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. *American Economic Review*, 101(6):2350–90, October 2011. URL <http://ideas.repec.org/a/aea/aecrev/v101y2011i6p2350-90.html>.
- P. Dupas and J. Robinson. Why Don't the Poor Save More? Evidence from Health Savings Experiments. *American Economic Review*, 103(4):1138–71, June 2013.
- J. Eeckhout and K. Munshi. Matching in Informal Financial Institutions. *Journal of the European Economic Association*, 8(5):947–988, 2010.
- M. Fafchamps and E. L. Ferrara. Self-help Groups and Mutual Assistance: Evidence from Urban Kenya. Working Papers 401, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University, 2011.
- M. Fafchamps and F. Gubert. The Formation of Risk Sharing Networks. *Journal of Development Economics*, 83(2):326–350, July 2007.
- M. Gash and K. Odell. The Evidence-Based Story of Savings Groups: A Synthesis of Seven Randomized Control Trials. Technical report, SEEP Network, 2013.
- M. Ghatak. Screening by the Company You Keep: Joint Liability Lending and the Peer Selection Effect. *Economic Journal*, 110(465):601–31, July 2000.
- M. Ghatak and T. W. Guinnane. The Economics of Lending with Joint Liability: Theory and Practice. *Journal of development economics*, 60(1):195–228, 1999.

- B. Greaney, J. P. Kaboski, and E. V. Leemput. Can Self-help Groups Really be Self-help? Working Papers 2013-014, Federal Reserve Bank of St. Louis, 2013.
- P. Heidhues and B. Kőszegi. Exploiting naivete about self-control in the credit market. *The American Economic Review*, pages 2279–2303, 2010.
- D. Karlan, A. L. Ratan, and J. Zinman. Savings by and for the Poor: A Research Review and Agenda. Working Papers 1027, Economic Growth Center, Yale University, July 2013.
- F. Kast, S. Meier, and D. Pomeranz. Under-Savers Anonymous: Evidence on Self-Help Groups and Peer Pressure as a Savings Commitment Device. NBER Working Papers 18417, National Bureau of Economic Research, Inc, Sept. 2012.
- C. Ksoll, H. Bie Lilleør, J. H. Lønberg, and O. D. Rasmussen. Impact of Village Savings and Loans Associations: Evidence from a Cluster Randomized Trial. Technical report, Rockwool Foundation, 2013.

A Descriptive tables

Table 8: Individual member characteristics, full 2013 member census

	Full CSAE Sample				
	Mean	Std dev	Min	Max	N
Occupation					
Farmer	0.73	(0.44)	0.0	1.0	3801
Businessperson	0.21	(0.41)	0.0	1.0	3801
Other	0.06	(0.23)	0.0	1.0	3801
Demographic Variables					
Male	0.25	(0.43)	0.0	1.0	3799
Female-headed household	0.21	(0.41)	0.0	1.0	3796
Age	36.15	(12.05)	12.0	83.0	3785
Education					
Some primary educ.	0.76	(0.43)	0.0	1.0	3801
Some post-primary educ.	0.18	(0.39)	0.0	1.0	3801
Literate (read & understand newspaper)	0.83	(0.38)	0.0	1.0	3795
Wealth					
Father well-off in village (Index 1-5)	3.34	(1.31)	1.0	5.0	3573
Spouse's father well-off in village (Index 1-5)	3.42	(1.32)	1.0	5.0	3566
Income Poverty Indicators					
Household well-off in group (Index 1-9)	7.52	(1.28)	1.0	9.0	3770
HH owns a bicycle	0.51	(0.50)	0.0	1.0	3796
Goats	1.29	(2.59)	0.0	40.0	3792

Table 9: Individual economic activity and schooling, full 2013 member census

Category	N	% of re- sponses
Economic activity	3780	100%
Farmer	2,786	73.34%
Business	793	20.87%
<i>Self-employed</i>	108	2.84%
<i>Family business worker</i>	685	18.03%
Fishing	61	1.61%
<i>Fishing, employed</i>	17	0.45%
<i>Fishing, self-employed</i>	44	1.16%
Employee	89	2.34%
Casual labour (ganyu)	24	0.63%
Student	4	0.11%
Unemployed, not seeking work	9	0.24%
Other	14	0.37%
Schooling	3795	100%
None/pre-school	195	5.1%
Primary	2897	76.3%
Secondary	675	17.8%
Tertiary and higher	28	0.7%

Total membership is 3,801; number of observations for each variable reflects missing values or non-applicability.

Table 10: Distribution of matched individuals across groups

# matched individuals in group	# groups	# groups as % of total groups
0	7	4.7%
1	18	12%
2	19	12.7%
3	21	14%
4	14	9.3%
5	17	11.3%
6	11	7.3%
7	15	10%
8	6	4%
9	7	4.7%
10	5	3.3%
11	3	2%
12	3	2%
13	3	2%
14	1	0.7%
Total	150	100%

Average number of members per group = 25.34

Average number of matched individuals per group = 4.76

Total number of matched individuals = 722 (out of 3,801 individuals in the 2013 member census).

Table 11: Dyad characteristics, full 2013 member census

	Full CSAE Sample				N
	Mean	Std dev	Min	Max	
Membership					
Same VSLA group (dummy)	0.17	(0.37)	0.0	1.0	289914
Absolute differences - Occupation					
Same economic activity (dummy)	0.56	(0.50)	0.0	1.0	289914
Farmer	0.37	(0.48)	0.0	1.0	289914
Businessperson	0.33	(0.47)	0.0	1.0	289914
Absolute differences - Demographic Variables					
Male	0.38	(0.48)	0.0	1.0	289467
Female-headed household	0.32	(0.47)	0.0	1.0	288740
Age	12.64	(10.25)	0.0	65.0	286763
Absolute differences - Education					
Schooling level	0.52	(0.72)	0.0	4.0	288342
Some primary educ.	0.38	(0.49)	0.0	1.0	289914
Some post-primary educ.	0.33	(0.47)	0.0	1.0	289914
Literate (read & understand newspaper)	0.26	(0.44)	0.0	1.0	288342
Absolute differences - Wealth					
Father well-off in village (Index 1-5)	1.41	(1.10)	0.0	4.0	253041
Spouse's father well-off in village (Index 1-5)	1.43	(1.13)	0.0	4.0	252485
Absolute differences - Income and Poverty					
Household well-off in group (Index 1-9)	1.40	(1.15)	0.0	8.0	284485
HH owns a bicycle	0.46	(0.50)	0.0	1.0	288740
Goats	1.84	(2.73)	0.0	40.0	288240
Sum - Occupation					
Farmer	1.37	(0.70)	0.0	2.0	289914
Businessperson	0.48	(0.63)	0.0	2.0	289914
Sum - Demographic Variables					
Male	0.53	(0.63)	0.0	2.0	289467
Female-headed household	0.40	(0.57)	0.0	2.0	288740
Age	72.12	(16.69)	24.0	163.0	286763
Sum - Education					
Schooling level	6.28	(0.91)	2.0	10.0	288342
Some primary educ.	1.48	(0.62)	0.0	2.0	289914
Some post-primary educ.	0.43	(0.58)	0.0	2.0	289914
Literate (read & understand newspaper)	1.68	(0.52)	0.0	2.0	288342
Sum - Wealth					
Father well-off in village (Index 1-5)	6.76	(1.85)	2.0	10.0	253041
Spouse's father well-off in village (Index 1-5)	6.85	(1.88)	2.0	10.0	252485
Sum - Income and Poverty					
Household well-off in group (Index 1-9)	15.04	(1.83)	3.0	18.0	284485
HH owns a bicycle	0.99	(0.73)	0.0	2.0	288740
Goats	2.44	(3.38)	0.0	70.0	288240

Table 12: Detailed dyad characteristics (absolute differences), matched subsample

	Absolute Differences - Matched dyads				
	Mean	Std dev	Min	Max	N
Membership					
Same VSLA group (dummy)	0.30	(0.46)	0.0	1.0	7326
Occupation					
Same economic activity (dummy)	0.70	(0.46)	0.0	1.0	7326
Farmer	0.27	(0.44)	0.0	1.0	7326
Businessperson	0.23	(0.42)	0.0	1.0	7326
Time Preferences					
Present biased	0.16	(0.37)	0.0	1.0	1641
Monthly discount factor (near frame)	0.12	(0.11)	0.0	0.3	1655
Risk Preferences					
Risk aversion extreme or severe	0.43	(0.49)	0.0	1.0	1513
Risk aversion intermediate or moderate	0.39	(0.49)	0.0	1.0	1513
Intra-Household Bargaining					
Ever hides money from spouse	0.45	(0.50)	0.0	1.0	1269
Female HH decisionmaking power (Index 0-8)	1.95	(1.57)	0.0	8.0	1914
Social Variables					
HH important in village decisions (Index 1-6)	1.08	(1.11)	0.0	5.0	7314
HH speaks at village meetings	0.47	(0.50)	0.0	1.0	7266
Demographic Variables					
Male	0.39	(0.49)	0.0	1.0	7326
Female-headed household	0.34	(0.48)	0.0	1.0	7326
Age	12.17	(9.87)	0.0	60.0	7326
Education					
Schooling level	0.45	(0.71)	0.0	4.0	7326
Some primary educ.	0.32	(0.47)	0.0	1.0	7326
Some post-primary educ.	0.26	(0.44)	0.0	1.0	7326
Literate (read & understand newspaper)	0.29	(0.45)	0.0	1.0	7326
Wealth					
Father well-off in village (Index 1-5)	1.41	(1.12)	0.0	4.0	6429
Spouse's father well-off in village (Index 1-5)	1.39	(1.12)	0.0	4.0	6970
Income					
Monthly consumption per capita, 1000 MK	1.01	(0.95)	0.0	7.0	7314
Food security poor (dummy)	0.39	(0.49)	0.0	1.0	7326

Table 13: Detailed dyad characteristics (sums), matched subsample

	Sums - Matched dyads				
	Mean	Std dev	Min	Max	N
Occupation					
Farmer	1.61	(0.60)	0.0	2.0	7326
Businessperson	0.32	(0.55)	0.0	2.0	7326
Time Preferences					
Present biased	0.19	(0.42)	0.0	2.0	1641
Monthly discount factor (near frame)	1.70	(0.18)	1.4	2.0	1655
Risk Preferences					
Risk aversion extreme or severe	1.18	(0.74)	0.0	2.0	1513
Risk aversion intermediate or moderate	0.60	(0.67)	0.0	2.0	1513
Intra-Household Bargaining					
Ever hides money from spouse	0.79	(0.71)	0.0	2.0	1269
Female HH decisionmaking power (Index 0-8)	5.90	(2.71)	0.0	16.0	1914
Social Variables					
HH important in village decisions (Index 1-6)	6.54	(1.69)	2.0	12.0	7314
HH ever speaks at village meetings	1.16	(0.71)	0.0	2.0	7266
Demographic Variables					
Male	0.55	(0.63)	0.0	2.0	7326
Female-headed household	0.45	(0.59)	0.0	2.0	7326
Age	77.04	(16.24)	39.0	157.0	7326
Education					
Schooling level	6.13	(0.88)	2.0	9.0	7326
Some primary educ.	1.59	(0.57)	0.0	2.0	7326
Some post-primary educ.	0.31	(0.52)	0.0	2.0	7326
Literate (read & understand newspaper)	1.64	(0.55)	0.0	2.0	7326
Wealth					
Father well-off in village (Index 1-5)	6.64	(1.89)	2.0	10.0	6429
Spouse's father well-off in village (Index 1-5)	6.88	(1.88)	2.0	10.0	6970
Income					
Monthly consumption per capita, 1000 MK	4.36	(1.43)	1.5	11.2	7314
Food security poor (dummy)	0.55	(0.64)	0.0	2.0	7326

Table 14: Group composition in terms of member characteristics

Variable	Across groups				# groups
	Average	Std Dev	Min	Max	
# members	25.34	5.69	10	45	150
% members farmers	74.2	26.3	3.3	100.0	150
% members businessperson	20.2	22.2	0.0	93.3	150
% members fisherman/woman	1.6	5.5	0.0	30.4	150
% members female	74.9	19.0	0.0	100.0	150
% members from female-headed hh's	20.9	13.0	0.0	61.9	150
Mean age of members	36	4.6	23	49	150
% members literate	82.5	12.2	40.9	100.0	150
% members primary educated	81.6	11.4	38.7	100.0	150
% members own bicycle	52.1	18.6	0.0	95.0	150
Mean # goats owned by members	1.32	0.98	0.04	7.48	150

From our VSLA member census conducted in July-August 2013. Total membership is 3,801 individuals.

Table 15: Test of representativeness – matched dyads vs. all 2013 dyads

	Mean All Dyads	Mean Matched Dyads	Diff.	Std. Error	T stat	Full N	Matched N
Group membership							
Same VSLA group (dummy)	0.17	0.30	-0.13***	(0.01)	-24.27	289914	7309
Absolute differences							
Same economic activity (dummy)	0.56	0.71	-0.15***	(0.01)	-27.11	289914	7309
Farmer	0.37	0.27	0.11***	(0.01)	20.59	289914	7309
Businessperson	0.33	0.23	0.10***	(0.00)	21.00	289914	7309
Male	0.38	0.39	-0.02**	(0.01)	-2.89	289467	7309
Female-headed household	0.32	0.34	-0.03***	(0.01)	-4.64	288740	7309
Age	12.64	12.17	0.46***	(0.12)	3.94	286763	7309
Schooling level	0.52	0.45	0.07***	(0.01)	8.23	288342	7309
Some primary educ.	0.38	0.32	0.06***	(0.01)	10.51	289914	7309
Some post-primary educ.	0.33	0.26	0.07***	(0.01)	13.29	289914	7309
Literate (read & understand newspaper)	0.26	0.29	-0.03***	(0.01)	-5.12	288342	7309
Father well-off in village (Index 1-5)	1.41	1.42	-0.00	(0.01)	-0.12	253041	6415
Spouse's father well-off in village (Index 1-5)	1.43	1.40	0.03*	(0.01)	2.44	252485	6954
Household well-off in group (Index 1-9)	1.40	1.29	0.11***	(0.01)	8.31	284485	7309
HH owns a bicycle	0.46	0.46	0.01	(0.01)	1.21	288740	7309
Goats	1.84	2.30	-0.45***	(0.04)	-11.14	288240	7309
Sums							
Farmer	1.37	1.61	-0.24***	(0.01)	-33.75	289914	7309
Businessperson	0.48	0.32	0.16***	(0.01)	24.67	289914	7309
Male	0.53	0.55	-0.02**	(0.01)	-2.68	289467	7309
Female-headed household	0.40	0.45	-0.05***	(0.01)	-6.69	288740	7309
Age	72.12	77.04	-4.92***	(0.19)	-25.52	286763	7309
Schooling level	6.28	6.13	0.15***	(0.01)	14.68	288342	7309
Some primary educ.	1.48	1.59	-0.11***	(0.01)	-16.17	289914	7309
Some post-primary educ.	0.43	0.31	0.11***	(0.01)	18.76	289914	7309
Literate (read & understand newspaper)	1.68	1.63	0.04***	(0.01)	6.87	288342	7309
Father well-off in village (Index 1-5)	6.76	6.64	0.12***	(0.02)	5.11	253041	6415
Spouse's father well-off in village (Index 1-5)	6.85	6.87	-0.02	(0.02)	-0.84	252485	6954
Household well-off in group (Index 1-9)	15.04	15.41	-0.37***	(0.02)	-18.06	284485	7309
HH owns a bicycle	0.99	1.09	-0.10***	(0.01)	-12.09	288740	7309
Goats	2.44	3.32	-0.89***	(0.05)	-16.90	288240	7309