Matching for Credit: Risk and Diversification in Thai Microcredit Groups

Christian Ahlin*

March 2009

Abstract

How has the microcredit movement managed to push financial frontiers? In a context in which borrowers vary in unobservable risk, Ghatak (1999, 2000) shows that group-based, joint liability contracts price for risk more accurately than individual contracts, provided that borrowers match homogeneously by risk-type. This more accurate risk-pricing can attract safe borrowers and invigorate an otherwise dormant credit market. We extend the theory to examine sorting when group size is larger than two and joint liability can take several forms. We also extend borrower heterogeneity to include correlated risk, and show that borrowers will anti-diversify risk within groups, in order to lower chances of facing liability for group members. We directly test risk-matching and intra-group diversification of risk using data on Thai microcredit borrowing groups. We propose a non-parametric univariate methodology for assessing homogeneity of matching; multivariate analysis is carried out using Fox's (2008) matching maximum score estimator. We find evidence of moderately homogeneous sorting by risk, in support of Ghatak's theory. We also see evidence of risk anti-diversification within groups. The anti-diversification results reveal a potentially negative aspect of voluntary group formation and point to limitations of microcredit groups as risk-sharing mechanisms.

*I thank Ashley Ahlin, Pierre-Andre Chiappori, Vivek Dhand, Todd Elder, Yanqin Fan, James Foster, Jeremy Fox, Xavi Gine, Alex Karaivanov, Ethan Ligon, Andrew Newman, Moto Shintani, Rob Townsend, Jonathan Zinman, and seminar participants at Boston University, Chicago, Georgia, Michigan State, Vanderbilt and NEUDC 2005 for valuable input. Jisong Wu provided excellent research assistance. All errors are mine.

1 Introduction

The seemingly unprecedented growth in intermediation and financial services among the world's poor associated with the "microcredit movement" has surprised many.[1] Microcredit has come to be viewed as an effective way to target capital toward productive, entrepreneurial uses by those at the bottom of the world income distribution, and in short, as one of the most promising recent advances in economic development.[2]

Despite its advances, questions about microcredit remain. An obvious and important one – does it work? – seems still to have eluded a definitive answer.[3] However, there seems to be a substantial prima facie case for positive impacts from microcredit: the apparently large number of microcredit institutions lending to poor borrowers but achieving robust repayment rates, financial sustainability, and repeat relationships suggests that positive gains from trade are being realized.

This leads to a second question: *how* does microcredit work? How have lenders managed to solve the repayment problem involved in lending to poor, collateral-less borrowers? The current paper is in the stream of literature focused on this question.

In explaining the apparently new-found success in lending to the world's poor, focus has been on lending *innovations* associated with the microcredit movement, such as group lending. In a group lending contract, borrowers are required to form official groups and to bear some liability for the loans of fellow group members. A body of research has shown theoretically how such joint liability lending can help overcome the informational and enforcement limitations that make uncollateralized lending hard.

[1] Bellman (2006) reports that more than 100 million customers worldwide are borrowing small loans from around 10,000 microfinance institutions. According to the 2006 Nobel Peace Prize Press Release, "Loans to poor people without any financial security had appeared to be an impossible idea".

[2] The UN, for example, declared 2005 the "International Year of Microcredit". The 2006 Nobel Peace Prize was awarded to Bangladesh's Muhammad Yunus and the Grameen Bank for pioneering the microcredit approach. Quoted on Armendariz de Aghion and Morduch (2005), economist Timothy Besley calls microfinance "one of the most significant innovations in development policy of the past twenty-five years". Armendariz de Aghion and Morduch (2005), Ghatak and Guinnane (1999), and Morduch (1999) provide introductions to the topic.

[3] Of course, there is probably no single answer. See Armendariz de Aghion and Morduch (2005) for a discussion of impact studies. Ahlin and Jiang (2008) explore the issue of long-run impact theoretically.

One such theory is by Ghatak (1999, 2000). The context is a standard adverse selection environment (Stiglitz and Weiss, 1981)[4] in which there is no collateral and borrowers' distributions of project returns have identical means but vary in riskiness.[5] In this environment, a lender that cannot observe risk offers all borrowers the same terms, but effectively charges less to risky borrowers, who fail more often, than to safe borrowers. Thus there is cross-subsidization of risky borrowers by safe borrowers, and this may cause safe borrowers to exit the market.

Ghatak (1999, 2000) adds to this context a communal tightness – that is, that borrowers know each other's riskiness – and shows that group lending contracts can harness the borrowers' information to improve the lender's ability to price for risk. The idea is as follows. First, borrowers voluntarily sort into groups that are homogeneous by risk. Second, he shows that given homogeneous matching, the lender can use joint liability contracts to screen (Ghatak, 2000, Gangopadhyay et al., 2005) or pool borrowers (Ghatak, 1999) to increase efficiency. Consider the pooling contract. Even though contract terms are the same for all borrowers, there is effectively a *built-in discount* for safe borrowers: their partners are safer and thus the joint liability clause is less costly for them in expectation. This discount can draw into the market safe borrowers who would have been excluded under standard, individual loans.

The pooling result is appealing in practical terms. It implies that even a very passive or unsophisticated lender that offers a single, standardized group contract is giving implicit discounts to safe borrowers, and hence more accurately pricing for risk than if it used individual contracts. This may help explain the popularity of group lending in microcredit – lenders that use it may be invigorating an otherwise anemic market (even unwittingly) – and could serve as part of the explanation for the growth of credit markets among the poor.

The lynchpin in this analysis is the homogeneous risk-matching of the borrowing groups. Without it, there need be no discount for safe borrowers operating through the group con-

[4]This focus has found empirical backing by the direct and indirect evidence for adverse selection in these lending contexts found in Ahlin and Townsend (2007a,b).

[5]The term "adverse selection" is a slight misnomer here: perfect efficiency is when all agents, *safe and risky*, borrow. Any inefficiency comes from the exclusion of safe borrowers, not the inclusion of risky.

tract. A main contribution of this paper is to test directly for homogeneous risk-matching among borrowing groups in Thailand. We are unaware of other direct tests of this result, though it is the first question on the microfinance empirical research agenda Morduch (1999, p. 1586) lays out: "Is there evidence of assortative matching through group lending as postulated by Ghatak (1999)?" Our broader goal is to extend and test theories of group formation pioneered by Ghatak (1999, 2000), and ultimately to understand better how well Ghatak's theory of how microcredit works fits the data.

The paper first extends theoretical results on group formation in a few ways. First, we demonstrate that depending on how joint liability is operationalized, it can lead to homogeneous or heterogeneous matching. This is not new to the literature; what is new is a characterization of the sorting patterns that can arise when risk-types are substitutes, not complements, and groups have more than two members. An asymmetry between the predictions of risk-type complementarity and substitutability is uncovered: complementarity nearly completely pins down equilibrium groupings for any group size, while substitutability rules out a vanishingly small fraction of groupings as group size gets large. The implication for empirical work is that to reject substitutability, i.e. negative assortative matching, structural estimation that uses an explicit payoff function and borrower types may be necessary.

We also extend the theory to include *correlated* risk. The main theoretical result is that groups sort homogeneously in both dimensions: they match with similar risk-types, and among those, with partners exposed to the same risk. The intuition is straightforward: groups anti-diversify in order to avoid facing liability for their partners. This points to a potentially negative consequence of voluntary group formation. Consider the extreme case where project returns are perfectly correlated within a group. Joint liability is then never executed, and group lending reduces to individual lending. In short, anti-diversification can work against efficiency by hindering the lender's ability to use joint liability effectively.

We test empirically whether groups are homogeneous in both risk-type and risk-exposure. The data come from the Townsend Thai dataset, which includes information on borrowing groups from the Bank for Agriculture and Agricultural Cooperatives (BAAC). The BAAC is the predominant rural lender in Thailand. It offers joint liability contracts to self-formed groups of borrowers with little or no collateral.

To assess homogeneity of matching, we compare across groups within villages. Specifically, for each village and variable, we calculate a variance decomposition, rank correlation, and/or chi-squared test statistic to assess homogeneity of sorting. We put these calculations in perspective by use of a permutation test. That is, we repeat the calculations for all possible groupings of the village borrowers into groups of the observed sizes and map the result from the observed grouping into a *sorting percentile* reflecting how homogeneous or heterogeneous group formation is relative to all possibilities (holding group size and borrowing pool fixed).

For any given variable, villages can be found at both ends of the spectrum – homogeneous sorting (high sorting percentile) and heterogeneous sorting (low sorting percentile). Sorting percentile means and medians (across villages) suggest predominant tendencies. We show that if matching is completely random, then village sorting percentiles are drawn from a uniform distribution. Thus we can statistically compare the overall sorting patterns in the data to random matching by comparing the sample CDF of village sorting percentiles to the uniform distribution CDF. We do so using the Kolmogorov-Smirnov (KS) test.

We find direct evidence for risk-homogeneity within groups. That is, though it is far from perfect homogeneity, the matching pattern can reject random matching in the direction of homogeneity. We also find some evidence for anti-diversification within groups. While groups do not appear occupationally diversified or anti-diversified relative to random matching, in terms of clustering of bad income years and income shocks, they appear anti-diversified. The evidence is consistent with group formation that anti-diversifies group risk in order to circumvent joint liability.

We turn next to a multivariate analysis based on Fox's (2008) matching maximum score estimator and subsampling-based inference. This estimator chooses parameter values that maximize the frequency with which observed groupings yield higher payoffs than feasible, unobserved groupings. In our implementation, the feasible, unobserved groupings are those that result from swapping k borrowers across two groups in the same village. Results offer support for homogeneous matching along both risk-type and correlated risk dimensions.

In sum, Ghatak's (1999) theory receives support from the data: the degree of risk-type homogeneity is far from perfect, but statistically different from random. The evidence suggests that group lending is succeeding in embedding a non-negligible discount for safe borrowers via group matching; this may partly explain how group lending and microcredit has successfully invigorated previously dormant credit markets. However, results on anti-diversification point to a potentially negative aspect of voluntary group formation.

It should be noted that we do not firmly establish causal determinants of group formation. However, to assess whether group lending provides for better pricing for risk by targeting discounts to safe borrowers, we argue that this is not necessary (section 4.2). Whether risk-homogeneity results from purposeful matching or as an unintended consequence, it is by itself sufficient for the improvement in risk-pricing that group lending is theorized to offer.

The paper is organized as follows. The model setup and theoretical sorting results are in section 2. Data are described and key variables defined in section 3. Section 4 presents the methodology behind the nonparametric univariate tests, as well as the results. Section 5 presents the multivariate estimation. Section 6 concludes. Proofs are in the appendix.

2 Theory

2.1 Baseline Model and Matching Results

Risk-neutral agents are each endowed with no capital and one project. Each project requires one unit of capital and has expected value E. Agents and their projects differ in risk, indexed by $p \in \mathcal{P}$. The project of an agent of type p yields gross returns of Y_p ("succeeds") with probability p and yields zero gross returns ("fails") with probability $1-p$.[6] The higher p, the

[6]This implies that $p \cdot Y_p = E$, for all $p \in \mathcal{P}$.

lower the agent's risk. Agents' types are observable to other agents, but not to the outside lender.

Two different cases may be assumed. First, there may be a unit measure continuum of agents, in which case population risk-types are taken to be distributed with no mass points over $\mathcal{P} = [\underline{p}, \overline{p}]$ according to density function $f(p)$ and CDF $F(p)$, where $0 < \underline{p} < \overline{p} < 1$.[7] Second, the number of agents may be finite, in which case $\mathcal{P}$ is a finite subset of $[\underline{p}, \overline{p}]$.

Following Ghatak (1999, 2000), a lender requires potential borrowers to form groups of size two, each member of which is jointly liable for the other. Specifically, contracts are assumed to take the following form. A borrower who fails pays the lender nothing, since loans are uncollateralized. A borrower who succeeds pays the lender gross interest rate $r > 0$. A borrower who succeeds and whose partner fails makes an additional liability payment $q > 0$. Thus, a borrower of type p_i who matches with a borrower of type p_j has expected payoff

$$\pi_{ij} = E - rp_i - qp_i(1 - p_j), \tag{1}$$

assuming the borrowers' returns are uncorrelated. Note that

$$\frac{\partial^2 \pi_{ij}}{\partial p_i \partial p_j} = q > 0. \tag{2}$$

That is, risk-types are complements in the joint payoff function and homogeneous matching by risk is the stable outcome when there is a continuum of agents, as Ghatak has shown.

Moving slightly closer to the data, in which all groups have more than two borrowers, consider contracts based on groups of fixed size $n \geq 2$.[8] Ghatak (1999) suggests the following generalization of the two-person group contract: assume a borrower owes $q > 0$ for *each* fellow group member that fails. Let $G = (p_1^G, p_2^G, \ldots, p_n^G) \in \mathcal{P}^n$ denote a group of borrowers; let

[7] In this case, as in Legros and Newman (2002), we think of there being a continuum of each type p for which $f(p) > 0$.

[8] Endogenizing group size is beyond the scope of the paper.

$\mathcal{N} = \{1, 2, \ldots, n\}$ and $\mathcal{N}_{-i} = \mathcal{N} \setminus i$. Then the payoff of the ith borrower in group G is

$$\pi_{iG} = E - rp_i^G - qp_i^G \sum_{j \in \mathcal{N}_{-i}} (1 - p_j^G), \tag{3}$$

since (under independence) the expected number of group members besides the ith who fail is just the sum of their failure probabilities.[9]

For any group G, let $\mathbf{\Pi_G}$ be the sum of payoffs of all borrowers in group G:

$$\Pi_G = \sum_{i \in \mathcal{N}} \pi_{iG} = nE - [r + q(n-1)] \sum_{i \in \mathcal{N}} p_i^G + q \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}_{-i}} p_i^G p_j^G. \tag{4}$$

Note that $\partial^2 \Pi_G / \partial p_i^G \partial p_j^G = 2q > 0$; thus risk-type complementarity continues to hold in the case of n-person groups.

By equilibrium matching, we mean[10] that the equilibrium set of borrowing groups cannot be blocked by a subset of n borrowers reorganizing so that each achieves a higher payoff; and that the number or measure of each type of borrower across equilibrium groups is consistent with the number or measure of potential borrowers of each type.

Define $\underline{\mathbf{p}}^{\mathbf{G}}$ ($\overline{\mathbf{p}}^{\mathbf{G}}$) as the minimum (maximum) type in group G: $\underline{p}^G = \min_{i \in \mathcal{N}} p_i^G$ and $\overline{p}^G = \max_{i \in \mathcal{N}} p_i^G$. Given the complementarity of risk-types in group payoffs, we have:[11]

Proposition 1. *Groups are rank-ordered (non-overlapping) in equilibrium. That is, if L and M are two equilibrium groups, then $\overline{p}^L \leq \underline{p}^M$ or $\overline{p}^M \leq \underline{p}^L$.*

This result holds whether there is a continuum or a finite number of borrowers. Obviously, complete intra-group homogeneity may be unattainable when the number of borrowers is

[9]Defining $\tilde{q} \equiv q(n-1)$ and $\overline{p}^G_{\mathcal{N}_{-i}}$ as the average risk-type in $\{p_j^G | j \in \mathcal{N}_{-i}\}$, the payoff can be written

$$\pi_{iG} = E - rp_i^G - \tilde{q} p_i^G (1 - \overline{p}^G_{\mathcal{N}_{-i}}),$$

directly analogous to payoff 1. Hence when matching is homogeneous and for a given n, as Ghatak (1999) has argued, expected payoffs under q and groups of size n are the same as under $\tilde{q}$ and groups of size 2.

[10]For more details, see Legros and Newman (2002). Propositions 1 and 2 use techniques stemming from their approach.

[11]All proofs are in the appendix.

finite; but with a continuum of borrowers, it is the probability-one outcome.

Corollary 1. *If there is a continuum of borrowers, almost every group is homogeneous in equilibrium.*

These results, and the intuition behind them, are quite similar to Ghatak's (1999, 2000).

Hence voluntary matching delivers within-group risk homogeneity, which in turn makes possible the better pricing for risk (via joint liability) that can help rejuvenate the credit market. However, the following two extensions show that voluntary matching can look very different and may work against better pricing for risk.

2.2 Matching Under Direct vs. Dynamic Liability

Sadoulet (1999)[12] makes the point that matching for joint liability loans need not be homogeneous. In particular, there are two forms of group liability in Sadoulet's model, which we term *direct* liability and *dynamic* liability. Direct liability refers to payments due from successful members to cover contemporaneous debts of unsuccessful members (as in Ghatak, 1999, 2000). Dynamic liability refers to the denial of future loans based on current group repayment performance. Both types of liability appear to be used in practice.

Here we parameterize the two kinds of liability to disentangle the matching incentives. We consider a two-period setting with no discounting and a relatively simple contract involving only two group liability parameters. As before, q is the direct liability payment due from a successful borrower in a given period for each group member who fails in that period. Dynamic liability is captured by d, with $d \in [0, 1/n]$: if $k \in \{0, 1, \ldots, n\}$ borrowers in a group fail in period 1, then $d \cdot k$ is each group member's probability of being denied a loan in period 2.[13] Hence, the probability of not getting a period-2 loan is proportional to contract parameter d and to the number of borrowers in one's group who fail in period 1.

[12] See also Guttman (2008) and, in a different environment, Chiappori and Reny (2006).

[13] Sadoulet (1999) analyzes groups of size 2, $q = r$ (i.e. full direct liability), and certain denial of future loans if and only if both borrowers fail (the only case where the bank does not receive $2r$). The n-person group generalization we use is inspired by his, but differs in allowing $q < r$ and probabilistic denial of loans.

Under this group contract with direct and dynamic liability, the expected payoff of borrower i in group G can be written

$$\pi_{iG} = [E - rp_i^G - qp_i^G \sum_{j\in\mathcal{N}_{-i}} (1-p_j^G)]\ [2 - d\sum_{j\in\mathcal{N}}(1-p_j^G)]. \tag{5}$$

The first bracketed term is the expected payoff from receiving a loan in a given period, identical to equation 3. The second bracketed term is the expected number of loans: 2 minus the probability of being denied a loan in period 2, which equals d times the expected number of failures in the group. The sum of payoffs in group G is then

$$\Pi_G = \sum_{i\in\mathcal{N}} \pi_{iG} = \{nE - [r + q(n-1)]\sum_{i\in\mathcal{N}} p_i^G + q\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}_{-i}} p_i^G p_j^G\}\ [2 - d\sum_{j\in\mathcal{N}}(1-p_j^G)], \tag{6}$$

identical to equation 4 except for the bracketed term involving d.

In this case, the cross-partial with respect to p_i^G and p_j^G, $i \neq j$, is

$$\frac{\partial^2 \Pi_G}{\partial p_i^G \partial p_j^G} = 2\left[2q - rd - dq\left(2n - 1 + p_i^G + p_j^G - 3\sum_{i\in\mathcal{N}} p_i^G\right)\right]. \tag{7}$$

Whether risk-types are complements or substitutes depends in part on which kind of group liability is dominant. In particular, by inspection it is clear that

Lemma 1. *Fix $q > 0$; for d small enough, $\partial^2\Pi_G/\partial p_i^G \partial p_j^G \approx 4q > 0$. Fix $d > 0$; for q small enough, $\partial^2\Pi_G/\partial p_i^G \partial p_j^G \approx -2rd < 0$.*

Thus, if dynamic liability (d) is negligible compared to direct liability (q), risk-types are complements and matching is homogeneous; specifically, rank-ordered (proposition 1) and homogeneous in a continuum (corollary 1). On the other hand, if dynamic liability dominates direct liability, risk-types are substitutes and the matching pattern will be different. The intuition for the different interactions is as follows. Under either kind of liability, all borrowers prefer safer partners. Under direct liability, *safe* borrowers' preferences are stronger since they are more likely to be "on the hook" for their partners (since partner bailouts can only be required from borrowers who have succeeded). Under dynamic liability, *risky* borrowers have stronger preferences for safe partners; a safer partner makes a future loan more likely, and risky borrowers benefit more from loans (they are cross-subsidized by safe borrowers).[14]

How will matching occur when risk-types are substitutes? Legros and Newman (2002) show that in the $n = 2$ case, negative assortative matching of the following type occurs: the highest type(s) match with the lowest, the next highest type(s) with the next lowest, and so on (onion-style). Every group has one member above and one member below the median.

However, the literature appears not to contain a derivation of the matching pattern under substitutability in the $n \geq 2$ case. Here we explore which restrictions from the $n = 2$ case generalize to higher dimensions, i.e. larger group sizes.

Say that group L **rank-wise dominates** group M if the jth largest element of L is greater than the jth largest element of M for all $j \in \mathcal{N}$. Define two groups L and M as **nearly rank-wise identical** if there exists an $i \in \mathcal{N}$ such that the jth largest elements of L and M, respectively, are equal for all $j \in \mathcal{N}_{-i}$. That is, L and M are nearly rank-wise identical if at $n-1$ (or n) ranks they have the same value.

Proposition 2. *Assume $\partial^2 \Pi_G / \partial p_i^G \partial p_j^G < 0$. In equilibrium, one group rank-wise dominates another only if the two groups are nearly rank-wise identical.*

This appears to be the main implication of substitutability of risk-types in higher dimensions: groups that are not extremely similar are intertwined, in that each dominates the other at some rank.[15] The following results extend this idea for different settings.

[14]The result is similar but the mechanism is slightly different in Sadoulet (1999). There, only when a borrower fails might he be denied a future loan; if he does fail, he may be spared credit denial if fellow group members' perform well. Since risky borrowers are thus more likely to be at risk for credit denial, they have stronger preferences for safe group members. This force could straightforwardly be added here by assuming the probability of borrower i being denied a loan depends on the total number of failures in the group *and* whether borrower i failed. For simplicity and since the results are not affected, we do not do so here.

[15]Consider the simple example with group size $n = 3$ and six borrowers: $p_1 < p_2 < p_3 < p_4 < p_5 < p_6$. There are $\binom{6}{3}/2 = 10$ groupings of these borrowers. In five of these groupings, include the rank-ordered one, one group rank-wise dominates the other: $\frac{(p_4,p_5,p_6)}{(p_1,p_2,p_3)}$, $\frac{(p_3,p_5,p_6)}{(p_1,p_2,p_4)}$, $\frac{(p_3,p_4,p_6)}{(p_1,p_2,p_5)}$, $\frac{(p_2,p_5,p_6)}{(p_1,p_3,p_4)}$, and $\frac{(p_2,p_4,p_6)}{(p_1,p_3,p_5)}$.

Corollary 2. *Assume that* $\partial^2\Pi_G/\partial p_i^G \partial p_j^G < 0$ *and that* $\mathcal{P}$ *is a measure-one continuum with no mass points. If two equilibrium groups are picked at random, one group rank-wise dominates the other with probability zero.*

Corollary 3. *Assume that* $\partial^2\Pi_G/\partial p_i^G \partial p_j^G < 0$*, that* $\mathcal{P}$ *is finite, and that no two agents have the same type.*[16] *In equilibrium, no group rank-wise dominates another.*

Ruling out pair-wise rank-wise dominance sheds light on the overall matching pattern. Analogous to the "matching about the median" of the 2-person case (Legros and Newman, 2002), all groups match around a common type.

Corollary 4. *Assume* $\partial^2\Pi_G/\partial p_i^G \partial p_j^G < 0$*. There exists a type* $\tilde{p} \in \mathcal{P}$ *such that* $\underline{p}^G \leq \tilde{p} \leq \overline{p}^G$ *for any equilibrium group* G*.*

Thus, when group liability is in the form of dynamic rather than direct liability, risk-matching is very different. Rather than rank-ordering by risk-type, groups generally do not rank-wise dominate each other, and they match around a common type. Any pair of equilibrium groups must be intertwined in a substantial way, i.e. relatively heterogeneous.

How restrictive is the no rank-wise dominance result, and is it all that can be said under submodularity without knowing more about the payoff function and distribution of types? These questions in general are beyond the scope of the paper, but we address them in the two-group case.

That is, assume there are $2n$ individuals, each with a unique type p, forming two groups of size n. In the $n = 2$ case, rank-wise dominance occurs in two of the three possible groupings; thus submodularity leaves only the onion-style grouping as stable. In the $n = 3$ case, rank-wise dominance occurs in exactly five of the ten possible groupings (see footnote 15). It

---

In the remaining five, $\frac{(p_2,p_3,p_6)}{(p_1,p_4,p_5)}$, $\frac{(p_2,p_3,p_5)}{(p_1,p_4,p_6)}$, $\frac{(p_2,p_3,p_4)}{(p_1,p_5,p_6)}$, $\frac{(p_2,p_4,p_5)}{(p_1,p_3,p_6)}$, and $\frac{(p_3,p_4,p_5)}{(p_1,p_2,p_6)}$, neither group rank-wise dominates the other. Hence, proposition 2 (and more directly, corollary 3) rules out the first five groupings under substitutability of types, but not the second five.

[16] This can be justified by assuming the finite collection of agents has types drawn from a distribution with no mass points.

is straightforward to show that any of the other five groupings can be the unique surplus maximizer, given the model's production function with $q = 0$ and depending on the values of the p_i's. We have also computationally verified the same result for $n = 4, 5, 6, 7$, that is, that any grouping in which neither group rank-wise dominates the other is the unique surplus maximizer for some set of types. Thus, while we do not have a general sufficiency result, we have established that rank-wise dominance is all that can be ruled out by submodularity in these simple cases.

We next show that ruling out rank-wise dominance is not very restrictive.

Proposition 3. *Consider all groupings of $2n$ uniquely-typed borrowers into two groups of size n (of which there are $\binom{2n}{n}/2$). In exactly a fraction $\frac{2}{n+1}$ of such groupings does one group rank-wise dominate the other.*

This result shows how weak the restriction on rank-wise dominance is in narrowing down the number of potential groupings, as n grows. In the limit, as $n \to \infty$, it rules out virtually no groupings (relatively speaking).[17]

Thus, there is an asymmetry that can be missed by a focus on the $n = 2$ case. When $n = 2$, both supermodularity and submodularity rule out 2 of 3 groupings (all but the rank-ordered one and all but the onion-style one, respectively). For $n > 2$ and with no ties, a fraction $n/(n + 1)$ of groupings (i.e. all but one grouping) is ruled out by the rank-ordered condition, while only a fraction $2/(n + 1)$ of groupings is ruled out by the no rank-wise dominance condition. Only when $n = 2$ are both conditions equally restrictive, while as n moves away from 2 submodularity appears to give far less sharp predictions.

In summary, if dynamic liability is insignificant relative to direct liability, groups will be rank-ordered. However, if direct liability is insignificant relative to dynamic liability, groups will have overlapping ranges and will tend not to have uniformly higher or lower distributions of risk-types. In the latter case, the results suggest that structural estimation that takes into

[17]This of course does not mean there are multiple stable matches, just that the stable match(es) is not pinned down by submodularity alone, but can depend on the distribution of types and other characteristics of the production function.

account the type distribution and details of the payoff function may be desirable, since submodularity alone does not generally pinpoint the stable match.

2.3 Matching over Amount and Type of Risk

The previous section showed that homogeneous matching can break down depending on how joint liability is operationalized. This section suggests another potential pitfall of relying on voluntary matching: not the breakdown of homogeneous matching, but the undesirability of homogeneous matching along a different dimension.

We return to the baseline model (of direct liability only) and in that setting consider matching when borrower risk can be correlated. The goal is to understand the incentives of borrowers to diversify (or anti-diversify) the risk of the group via matching.[18]

Given two borrowers i and j with probabilities of success p_i and p_j, respectively, there is a unique, one-parameter class of joint output distributions (see Ahlin and Townsend, 2007a,b):

	j Succeeds (p_j)	j Fails $(1-p_j)$
i Succeeds (p_i)	$p_i\, p_j + \epsilon_{ij}$	$p_i(1-p_j) - \epsilon_{ij}$
i Fails $(1-p_i)$	$(1-p_i)p_j - \epsilon_{ij}$	$(1-p_i)(1-p_j) + \epsilon_{ij}$

(8)

The case of $\epsilon_{ij} \equiv 0$ is the case of independent returns considered by Ghatak. A positive (negative) ϵ_{ij} gives positive (negative) correlation between borrower returns.

Correlation parameter ϵ_{ij} may differ across pairs of borrowers. We proceed by placing a simple structure on the population's correlatedness. Assume there are two aggregate shocks, A and B, distributed identically and independently; each equals 1 with probability 1/2 and -1 with probability 1/2.

Every agent is assumed to be exposed to risk from either shock A or shock B, or neither;

[18] Though joint liability groups are sometimes thought of as risk-sharing groups, this question need not directly bear on risk-sharing: households may share risk with other households regardless of whether they are in the same joint liability borrowing group. The key issue instead is how groups form in response to joint liability contracts, which may contain implicit rewards for (anti-)diversification.

shock exposure-type is known by all agents but not the lender. The probability of success of an "A-risk" agent of risk-type p_i depends on the realization of A in the following way: $p_i|A = p_i + \gamma A$, for some $\gamma > 0$. That is, if there is a good shock ($A = 1$), an A-risk agent's success probability gets a boost, equal to γ; a bad shock ($A = -1$) lowers the agent's success probability by γ. Exposure to shock A does not change an agent's unconditional probability of success, since this probability has an equal chance of being raised or lowered by γ.

The success of a "B-risk" agent of type p_i depends on the realization of shock B in the exactly analogous way: $p_i|B = p_i + \gamma B$. Fractions α and β of the population are exposed to shocks A and B, respectively, with $\alpha, \beta > 0$ and $\alpha + \beta \leq 1$. The remaining $1 - \alpha - \beta$ agents are exposed to neither aggregate shock ("N-risk"): that is, their probabilities of success are unaffected by the outcomes of A and B. Aggregate shock affiliation is assumed uncorrelated with unconditional probabilities of success, the p_i's; thus, unconditional risk and risk exposure are two orthogonal dimensions along which matching may occur.

With these assumptions, the ϵ_{ij} of equation 8 varies across borrowers i and j in a straightforward way. If borrowers i and j are exposed to the same shock, i.e. are both A-risk or both B-risk, one can show that[19]

$$\epsilon_{ij} = \epsilon \equiv \gamma^2. \tag{9}$$

For similarly exposed borrowers, returns are positively correlated because probabilities of success are pushed in the same direction by the shock. On the other hand, if borrowers i and j are not exposed to the same shock, $\epsilon_{ij} = 0$. This is because the shocks each borrower is exposed to – idiosyncratic and perhaps also aggregate – are independent.[20]

In summary, the correlation structure boils down to $\epsilon_{ij} = \epsilon$ ($\epsilon_{ij} = 0$) for pairs exposed (not exposed) to the same shock. Let $s \in \mathcal{S} \equiv \{A, B, N\}$ denote the potential borrower exposure-types. In this context of two-dimensional types, a group G is a vector

[19] Assume that borrowers i and j are both A-risk. With probability 1/2, the shock is good and the probability of both succeeding is $(p_i + \gamma)(p_j + \gamma)$; similarly, with probability 1/2 the probability of both succeeding is $(p_i - \gamma)(p_j - \gamma)$. The unconditional probability of both succeeding is thus $p_i p_j + \gamma^2$.

[20] Greater scope for diversification would be present if shocks A and B were negatively correlated, which could easily be incorporated without changing results.

$G = (p_1^G, p_2^G, \ldots, p_n^G, s_1^G, s_2^G, \ldots, s_n^G) \in \mathcal{P}^n \times \mathcal{S}^n$. For two borrowers i and j in group G, let $\kappa_{i,j}^G$ equal 1 if $s_i^G = s_j^G = A$ or $s_i^G = s_j^G = B$, and 0 otherwise. Then the payoff of the ith borrower in group G is, using equation 8 and the fact that here $\epsilon_{ij} = \kappa_{i,j}^G \epsilon$,

$$\pi_{iG} = E - rp_i^G - q \sum_{j \in \mathcal{N}_{-i}} [p_i^G(1-p_j^G) - \kappa_{i,j}^G \epsilon] = E - rp_i^G - qp_i^G \sum_{j \in \mathcal{N}_{-i}} (1-p_j^G) + q\epsilon \sum_{j \in \mathcal{N}_{-i}} \kappa_{i,j}^G. \quad (10)$$

The more group members that are exposed to the same risk as borrower i, the lower the expected number of other failures – and thus expected bailout payments – when borrower i succeeds; hence, borrower i's payoff is *increasing* in the amount of risk exposure he shares in common with his group. Group payoffs are then

$$\Pi_G = \sum_{i \in \mathcal{N}} \pi_{iG} = nE - [r + q(n-1)] \sum_{i \in \mathcal{N}} p_i^G + q \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}_{-i}} p_i^G p_j^G + q\epsilon \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}_{-i}} \kappa_{i,j}^G. \quad (11)$$

The following can be shown:

Proposition 4. *Assume a continuum of borrowers. In equilibrium, almost every group is homogeneous in both unconditional risk ($p \in [\underline{p}, \overline{p}]$) and risk exposure ($s \in \{A, B, N\}$).*

Thus, in equilibrium groups are homogeneous in risk-type and risk-exposure; they contain either all A-risk, all B-risk, or all N-risk borrowers. The intuition for the latter result is simple: borrowers choose to anti-diversify their groups so as to lower their chances of facing liability for their partners.

This result holds when there are many borrowers (a continuum). In a finite population, unidimensionally-optimal matching along both dimensions simultaneously may not be feasible. For example, the grouping that is rank-ordered by risk-type may involve sub-maximal anti-diversification, and any grouping that is maximally anti-diversified may violate rank-ordering by risk-type. In this case, tradeoffs between the two dimensions of matching are inevitable. It seems clear, though, that matching will tend toward uniformity along both dimensions.

While a complete analysis of contracting in this environment is beyond the scope of this paper, it appears that borrower sorting along the correlated risk dimension works against efficient lending. First, in a finite population it may divert borrowers from rank-ordering based on risk-type, which is the basis for joint liability's efficiency gains in this context.[21] Second, correlation in a sense lowers the *effective* rate of joint liability. Taking the extreme case of perfect correlation, the effective rate of joint liability is 0 regardless of how the bank sets q, since when one borrower fails, they all do. In general, the greater the correlation, the more irrelevant and blunted is any joint liability stipulation. This takes away from the lender a potentially valuable tool that can be used to increase lending efficiency. Thus, voluntary group sorting may not work in favor of efficiency along every dimension.

3 Data and Variable Descriptions

The empirical goal of the paper is to assess sorting patterns of borrowing groups related to amount and type of risk. To do so, we will compare groups drawn from the same pool of borrowers.

3.1 Data description and environment

The data come from the Townsend Thai survey data. In May 1997, a large cross section of l92 villages was surveyed, covering four provinces from two contrasting regions of Thailand. The Central region is relatively close to Bangkok and partially industrialized. The Northeast region is poorer and semi-arid, containing some of the poorest areas in the country. Both have large agricultural sectors. Within each province, twelve subcounties, or tambons, were chosen. Within each tambon, a cluster of four villages was selected. In each village as many borrowing groups of the Bank for Agriculture and Agricultural Cooperatives (BAAC) as

[21]Of course, other dimensions of borrower heterogeneity may influence matching and work against perfect rank-ordering by risk, such as religion, race, proximity, kinship and friendship. On the other hand, to the extent these other characteristics are correlated with risk, they may work in favor of homogeneous sorting by risk even when borrowers do not know or care about their own or others' risk.

possible were interviewed, up to two. This baseline survey contains data on 262 groups, 200 of which are one of two groups representing their village. Unfortunately for the purposes of this study, the borrower-level data provided in this survey are minimal and are all provided by the group's official leader, not the individual borrowers.

Hence, we turn to a smaller re-survey, conducted in April and May 2000. The resurvey data were collected from a random subset of the same tambons. Included are data on 87 groups, 14 of which are the only groups in their village, 70 of which are one of two groups interviewed from the same village, and 3 of which are one of three groups interviewed from the same village.[22] Though smaller, the resurvey data have two decisive advantages over the baseline data. First, in the resurvey individual group members respond to questions *on their own behalf*, up to five per group and on average 4.5.[23] Second, several resurvey questions were designed explicitly to measure income risk and correlatedness, the key variables in the theory.

The BAAC is a government-operated development bank in Thailand. It was established in l966 and is the primary formal financial institution serving rural households. It has estimated that it serves 4.88 million farm families, in a country with between sixty and seventy million inhabitants, about eighty percent of which live in rural areas. In the Townsend Thai baseline household survey covering the same villages, BAAC loans constitute 34.3% of the total number of loans, as compared with 3.4% for commercial banks, 12.8% for village-level financial institutions, and 39.4% for informal loans and reciprocal gifts (see Kaboski and Townsend, 1998).

The BAAC allows smaller loans[24] to be backed only with social collateral in the form of joint liability. This kind of borrowing is widespread: of the nearly 3000 households in the baseline household survey, just over 20% had a group-guaranteed loan from the BAAC

[22]This was apparently a mistake in implementation of the data collection methodology, which capped responses to two groups per village; we use the three-group village anyway.

[23]Unreported results for this paper strongly suggested that when one person responds for every group member, the responses will tend to exaggerate within-group homogeneity.

[24]The cap on group loans at the time of the baseline survey was 50,000 Thai baht, about $2000. The median group loan was closer to $1000.

outstanding in the previous year. To borrow in this way, a borrower must belong to an official BAAC borrowing group and choose the group-guarantee option on the loan application. The group then faces explicit liability for the loan; that is, in the event of a group member's default on a loan, the BAAC may opt to follow up with the delinquent borrower *or* other group members in search of repayment. Contract terms leave the BAAC leeway, and there are examples not only of this kind of direct liability, but also of dynamic liability: some group members report delays or greater difficulties in getting future loans when a group member is in default. What is not clear is which kind of liability is dominant, if either.

Groups typically have between five and fifteen members; about 15% are larger. Of critical importance to this paper on sorting, group formation is primarily at the discretion of the borrowers themselves. Typically, groups are born when borrowers propose a list of members to the BAAC, and the BAAC then approves some or all members. The BAAC seems to use its veto power sparingly: only about 12% of groups in the baseline survey report that the BAAC struck members from the list.[25] We know of no case where the BAAC adds members to a list or forms a group unilaterally. Thus, while the BAAC has some say about group formation, it is in large part left to the borrowers themselves.

3.2 Variable descriptions

The empirical strategy involves comparing across groups within villages to determine whether homogeneity is greater within groups than across groups. To do so, measures of risk and of correlatedness are necessary. Our main measure of risk takes the theory (section 2.1) quite literally.[26] Group members were asked what their income would be in the coming year if it were a good year (Hi), what their income would be if it were a bad year (Lo), and what they expected their income to be (Ex). Assuming that income can take only one of two values,

[25]This is in response to a free-form question about how original members were determined when the group was founded.

[26]Ahlin and Townsend (2007b) find direct evidence for adverse selection in this credit market using this measure.

Hi and Lo, the probability of success, **prob** or prob-high, works out to be

$$prob = \frac{Ex - Lo}{Hi - Lo},$$

using the fact that $prob * Hi + (1 - prob) * Lo = Ex$. Another measure of risk, less directly related to the model, is the **coefficient of variation** of income.[27] Based on the same projected income distribution, this works out to be

$$\sigma / Ex = \sqrt{Hi/Ex - 1}\sqrt{1 - Lo/Ex}.$$

Correlatedness is proxied in three ways. First, we use information on **occupation**, and more specifically, fraction of revenue coming from various agricultural occupations. Each borrower reports the amount of revenue received in more than thirty categories. Ten of the categories are agriculture-related – "rice farming", "corn farming", "raising shrimp", "raising chicken or ducks", etc. Our measure of occupation is thus a *vector* with ten entries, each giving the fraction of total household revenue accounted for by one agricultural category.[28]

This measure is motivated by the setup of section 2.3, which features household exposure to various shocks. We choose to generalize from that section's scalar measure of occupation/risk-exposure because some degree of within-household occupational heterogeneity is common in our data.[29] We also choose to focus specifically on agricultural revenue components because they arguably entail more exposure to common shocks than the other revenue categories (prevalent among which are wage labor, small business categories, investment income, and remittances). Further, the BAAC explicitly targets farmers and lends to promote agricultural investment.

[27]The coefficient of variation equals the standard deviation normalized by the mean.

[28]The vector sums to one, except when the household has revenue in non-agricultural categories, as is often the case.

[29]Occupational similarity between two borrowers is then the dot product of their respective vectors. The dot product would give the $\kappa_{i,j}^G$ of section 2.3 if we had only two categories and all vector entries had to be 0 or 1, i.e. there were no fractional occupations. With more than two categories and fractional occupations, it is straightforward to show that the dot product is the correct generalization of the theory in section 2.3.

Second, we use timing of bad income years, **worst_year**. Specifically, borrowers are asked which year of the past two was worse for household income: "one year ago", "two years ago", or "neither". If borrowers are exposed to the same aggregate shocks, bad income years are more likely to coincide; thus coincidence of bad years can proxy anti-diversification. One drawback of this measure is its coarseness, as it maps past performance into one of three categories.

Third, we calculate a direct measure of a household income **shock**. Ideally, it captures the percent deviation of this year's income from its expected value. The household's current income comes from a very detailed compilation of realized business and farm revenues and expenses for the just-completed year, Inc. The *expected value* of its current income is proxied by next year's expected income, Ex, mentioned above. The income shock is then $Shock = (Inc - Ex)/Ex$. This measure cleanly captures the income shock if households' income draws are i.i.d. over time. Then Ex is exactly mean income, and $Shock$ is this year's realized random component of income (as a percent of mean income). If incomes are not stagnant over time but growing at the same expected rate across households, the measure would merely have to be adjusted by a constant to continue to capture the income shock. However, if household incomes are growing at different rates, then $Shock$ captures not only the income shock but differential growth rates of income. We view this as the main drawback of this measure: within-group homogeneity may imply sorting aimed at anti-diversifying risk, or it may imply sorting based on income growth rates.

Flaws notwithstanding, the three measures of correlatedness can shed light on whether groups tend to be internally diversified or anti-diversified.

4 Univariate Methodology and Results

4.1 Univariate Methodology

Section 4.1.1 proposes a way to combine permutation testing with standard measures to assess homogeneity of sorting in a given village. Section 4.1.2 discusses how to use these village-level measures to test – nonparametrically, across all villages – whether groupings look more homogeneous or heterogeneous than under random matching. Finally, section 4.1.3 proposes alternative, model-driven metrics of sorting that can be used as the basis for these nonparametric tests against random matching.

4.1.1 Quantifying Sorting

Consider data on variable X from two groups in village v, L and M, of respective sample sizes l and m: $L = (x_1, ..., x_l)$ and $M = (x_{l+1}, ..., x_{l+m})$. In this section, we propose ways of measuring how homogeneously sorted these groups are.

First, assume X is an ordered variable. A relatively atheoretic way to assess within-group homogeneity is to calculate a *variance decomposition* of $X = (x_1, ..., x_{l+m})$ into between-group and within-group components. To illustrate the approach, compare village 1 with $L = (2, 5, 6, 8)$ and $M = (1, 4, 7, 9)$ to village 2 with $L' = (1, 2, 5, 6)$ and $M' = (4, 7, 8, 9)$. Of the overall variance, 0% in village 1 and 44% in village 2 would be attributed to between-group differences. The higher value reflects more homogeneous sorting, and the lower value more mixing.

A disadvantage of the variance decomposition is that it uses the cardinality of the data, while in the theory only ordinality matters (under supermodularity at least; see proposition 1). On these grounds, rank correlations are preferred since they use only the ordinality of the data. In particular, one can calculate the Kendall's tau$_b$[30] between the data X and a group index variable, Y, where for example $y_1 = ... = y_l = 1$ and $y_{l+1} = ... = y_{l+m} = 2$.

[30] Results using Spearman's rho end up nearly identical, so we do not report them. Formulas for this and Spearman's rho can be found in Gibbons and Chakraborti (2003, pp. 419-20, 422-3).

Using the villages of the previous example, Kendall's tau$_b$ between $(2, 5, 6, 8, 1, 4, 7, 9)$ and group index $(1, 1, 1, 1, 2, 2, 2, 2)$ is 0%, and between $(1, 2, 5, 6, 4, 7, 8, 9)$ and the same group index is 57%.[31] Again, the higher value reflects more homogeneous sorting, and the lower value more mixing.

For categorical variables, e.g. occupation or worst_year, neither of the preceding sorting metrics can be used. An atheoretic approach for this case is to use the chi-squared independence (or homogeneity) test statistic.[32] This statistic quantifies deviations from the grouping in which each group has the same proportion of responses in each category as the village population. For example, letting A and B be two occupations, compare village 1 with $L = (A, A, B, B)$ and $M = (A, A, B, B)$ to village 2 with $L' = (A, A, A, B)$ and $M' = (A, B, B, B)$. The chi-squared test statistic for village 1 is 0 and for village 2 is 2.[33] Again, the higher value reflects more homogeneous sorting.

We thus have three metrics for sorting, two for ordered variables (variance decomposition, rank correlation) and one for categorical variables (chi-squared statistic). To move toward a statistical test for homogeneous sorting, we use the same permutation test to scale each metric.

Specifically, consider again observed data $X = (x_1, ..., x_{l+m})$ from two groups in village v, L and M, of respective sizes l and m. We assume that borrowers can match with any others in their village – a reasonable assumption since villages are relatively small and geographically concentrated. Hence, we form all possible combinations of the $l+m$ borrowers into two groups of respective sizes l and m and perform the same calculation – variance decomposition, rank correlation, and/or chi-squared statistic – on each one. The *observed* village grouping can then be assigned a "sorting percentile" (or sorting percentile range, given ties and a finite population) based on where its calculated value falls relative to this universe of possibilities.

[31] The correlations would be the same but negative if the group indices were reversed, i.e. if we used group index vector $(2, 2, 2, 2, 1, 1, 1, 1)$. Since group index is arbitrary, we take the absolute value of the rank correlation (more generally, the maximum across all potential group indexings).

[32] The formula can be found in DeGroot (1986, pp. 536-7, 542-3).

[33] The formula generalizes in an obvious way to fractional occupations that may not sum to one.

In this way, the sorting score for each village is assigned a value (or range) in $[0, 1]$, with higher numbers representing greater homogeneity in sorting and lower numbers representing more heterogeneous sorting. This permutation scaling is applied for each variable and each metric.

To illustrate, consider again a village 1 with groups $L = (2, 5, 6, 8)$ and $M = (1, 4, 7, 9)$. There are $\binom{8}{4}/2 = 35$ ways to sort eight borrowers into two groups of size four. Of these 35 combinations, 32 register higher between-group variance while 3 (including the observed combination) register exactly the same, i.e. zero between-group variance. Thus village 1 is somewhere between the 0th and 8.6th percentiles in terms of group homogeneity; its sorting percentile range is $[0, 8.6]$. The somewhat wide range reflects the fact that there are ties and that the number of groupings is relatively small. Consider village 2 with groups $L' = (1, 2, 5, 6)$ and $M' = (4, 7, 8, 9)$. There are two combinations of borrowers higher, thirty one combinations lower, and two combinations tied, in terms of between-group variance. Village 2's sorting percentile range is thus $[88.6, 94.3]$. Similarly, applying this permutation test to Kendall's tau$-b$ rank correlation measure gives a slightly wider sorting percentile range to village 1, $[0, 11.4]$, and the same sorting percentile range to village 2, $[88.6, 94.3]$.

The same approach can be used with the chi-squared test statistic.[34] In village 1 with $L = (A, A, B, B)$ and $M = (A, A, B, B)$, there are seventeen combinations with a larger chi-squared test statistic and eighteen combinations tied. This village's percentile range is then $[0, 51.4]$. In village 2 with $L' = (A, A, A, B)$ and $M' = (A, B, B, B)$, eighteen combinations have less, one combination has greater, and sixteen combinations have the same chi-squared test statistic. Thus this village's sorting percentile range is $[51.4, 97.1]$.

Thus for a given variable and sorting metric, each village is assigned a sorting percentile range. A higher sorting percentile range reflects more homogeneous sorting, according to the metric employed, while a lower sorting percentile reflects more heterogeneous sorting. One can then interpret villages with percentiles above the 95th as exhibiting homogeneous

[34]Using p-values based on the chi-squared distribution seems undesirable due to the small group sample sizes.

sorting at the 5% confidence level, for example.

4.1.2 A Nonparametric Test

Rather than test sorting village by village, however, we combine villages in a single test (per variable and sorting metric) of the overall tendency to sort homogeneously. Each village's sorting percentile is treated as a draw from the same distribution, and this distribution is compared using the Kolmogorov-Smirnov test to a benchmark distribution. An advantage of this approach is that it is non-parametric and requires no distributional assumptions.

The benchmark comparison distribution we use is the one that would obtain if sorting with respect to the given variable were completely random in all villages: the uniform distribution on $[0, 1]$. The basic idea can be illustrated by considering the case of a large number of borrowers in a village, no two groupings of which result in a tie using the given sorting metric. If each of the N, say, possible groupings is equally likely, as is the case under random matching, then each $1/N$th sorting percentile is equally likely to be realized by a given village. That is, a village's sorting percentile is drawn from the uniform distribution – approximately, with the difference getting arbitrarily small as N increases.

With smaller numbers of borrowers and, especially, with ties, villages are assigned non-negligibly wide sorting percentile *ranges*, not sorting percentiles (see previous section). Consider drawing a sorting percentile randomly from the village's sorting percentile range via the uniform distribution. For example, if a village's sorting percentile range is calculated to be $[88.6, 94.3]$, the sorting percentile would then be drawn randomly from the uniform distribution on this interval.

To summarize, let a village's sorting percentile range be calculated by the permutation methods described in the previous section; and let its sorting percentile (point estimate) be drawn at random from the uniform distribution on its sorting percentile range. Then the exact distribution of a village's sorting percentile under random matching, regardless of the sorting metric, is the uniform on $[0, 1]$.

Proposition 5. *Under random matching, a village's sorting percentile z is drawn from the uniform distribution on* $[0, 1]$.

The procedure is then to construct a sample CDF from the village sorting percentiles, and compare it using the Kolmogorov-Smirnov (KS) test to the uniform distribution, i.e. random matching. If the sample CDF stochastically dominates the uniform, this means villages' sorting percentiles tend to be higher than random matching would give rise to and provides statistical evidence for homogeneous sorting. On the other hand, if the sample CDF is stochastically dominated by the uniform, this means villages' sorting percentiles tend to be lower than what random matching would produce, suggesting heterogeneous matching.

We thus can report p-values for these KS one-sided tests of stochastic dominance. Note, however, that one such p-value involves a number of random choices: the random draws that pick villages' sorting percentiles out of their sorting percentile ranges. Thus, even given the data, the p-value is a random variable. So, we repeat the test 1 million times under 1 million different sets of random draws, and report the average p-value across all draws.[35]

4.1.3 The test with structural sorting measures

The sorting metrics of section 4.1.1 – variance decomposition, rank correlation, and chi-squared test statistic – have the advantage of being well-known and intuitive, but their connection to the theory is not always clear.[36] So, we turn next to alternative metrics derived directly from the theory. The theory predicts that any grouping we observe must

[35] Since each p-value can be interpreted as a probability, and since each p-value is an equally valid assessment of this probability, taking the mean p-value across a large number of random draws appears to be a reasonable approach.

[36] It is clear that both the between-component of the variance and Kendall's τ_b is higher for any rank-ordered grouping than it is for any non rank-ordered grouping (into two groups of size n). So, these two measures of homogeneity are in principle equally effective in testing rank-ordered grouping, i.e. complementarity of types in the payoff function. Less clear is how well they do at detecting substitutability in the payoff function: given how few groupings substitutability by itself rules out (see section 2.2), some of the groupings that are permissible under substitutability are likely to rank relatively high in terms of rank correlations or between-group variances. Using the atheoretic measures thus gives rise to the danger of *falsely* interpreting high sorting percentiles as evidence against substitutability of types in the payoff function, i.e. as the predominance of group liability. However, while substitutability by itself may rule out few groupings, knowledge of the production function and the types generally leaves only one optimal grouping.

maximize the sum of group payoffs; otherwise, the groups could re-organize with everyone being made better off.[37] This suggests that the group payoff function itself can serve as the measure of sorting. Comparing the payoffs achieved in the observed grouping with the payoffs from alternative, unobserved groupings will then give a sense of how well sorted the grouping we observe is with respect to maximizing the payoff function of the theory (which, we have shown, leads to homogeneous matching under direct liability).

Consider first the baseline model with uncorrelated risk. Given existing groups L and M in a village, we would like to use $\Pi_L + \Pi_M$ as the measure of sorting. Using group payoff function 4, this equals:

$$\Pi_L + \Pi_M = 2nE - [r + q(n-1)] \left(\sum_{i \in \mathcal{N}} p_i^L + \sum_{i \in \mathcal{N}} p_i^M \right) + q \left(\sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}_{-i}} p_i^L p_j^L + \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}_{-i}} p_i^M p_j^M \right).$$

Note that the only part of this total payoff that may differ across groupings of this set of $2n$ borrowers is the interaction terms (the double sums). Hence, given our ultimate purpose of comparing $\Pi_L + \Pi_M$ against alternative groupings of the same set of borrowers, we can ignore all but the interaction terms:[38] $\sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}_{-i}} p_i^L p_j^L + \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}_{-i}} p_i^M p_j^M$. Note also that we do not typically have the entire group's data, in part due to missing data but moreso because a maximum of five group members are sampled. To account for this, we define $\overline{p}^G_{\mathcal{N}_{-i}}$ as the average risk-type in $\{p_j^G | j \in \mathcal{N}_{-i}\}$ and scale payoffs by $1/(n-1)$ to get:

$$\sum_{i \in \mathcal{N}} p_i^L \overline{p}^L_{\mathcal{N}_{-i}} + \sum_{i \in \mathcal{N}} p_i^M \overline{p}^M_{\mathcal{N}_{-i}}. \tag{12}$$

Letting $\mathcal{S}^G \subseteq \mathcal{N}$ be the sampled subset of group G and $\overline{p}^G_{\mathcal{S}^G_{-i}}$ be the average risk-type in

[37] This is the idea underlying the multivariate matching estimator of Fox (2008), which we use in section 5.

[38] An alternative way of stating this is that our ultimate results, the KS p-values (reported in section 4.2), will not change when the entire group payoff function is used for any choice of n, E, r, and $q > 0$.

$\{p_j^G | j \in \mathcal{S}^G \setminus i\}$, we use the analog estimator from the data:

$$\sum_{i \in \mathcal{S}^L} p_i^L \overline{p}^L_{\mathcal{S}^L_{-i}} + \sum_{i \in \mathcal{S}^M} p_i^M \overline{p}^M_{\mathcal{S}^M_{-i}}. \tag{13}$$

In the end, this estimate of the sum of group payoffs is simply the sum, over all village borrowers, of the borrower's risk-type multiplied by the average risk-type of borrowers he is matched with. To illustrate, village 1 with $L = (2, 5, 6, 8)$ and $M = (1, 4, 7, 9)$ has sum of group payoffs of 202,[39] compared to 234.67 for more homogeneously matched village 2 with $L' = (1, 2, 5, 6)$ and $M' = (4, 7, 8, 9)$. Using the data we have, we can directly calculate this expression substituting our empirical measure *prob* (see section 3) for the p's.

This expression was derived from the baseline payoff function 4. Consider instead payoff function 6, which incorporates both direct and dynamic liability. If dynamic liability is set to zero ($d = 0$), the same basic analysis as above applies: the relevant part of the sum of payoffs is the double sums above. If direct liability is instead set to zero ($q = 0$), again the same basic analysis as above applies except that the above double sums are multiplied by -1. That is, under direct liability only, the grouping should maximize the above payoff, while under dynamic liability only, the grouping should minimize it.[40] Thus, low (high) sorting percentiles would provide evidence for the existence of dynamic liability (direct liability).

Consider next payoff function 11, which incorporates correlated risk. There are two types of interaction terms in that payoff function, involving unconditional risk-type (p) and risk exposure-type. To test for (anti-)diversification using the univariate techniques of this section, we ignore the unconditional risk-type interaction terms and focus on interactions in risk exposure-type.[41] Using techniques as above and defining $\overline{\kappa}^G_{i,\mathcal{S}^G_{-i}}$ as the average correlatedness

[39] This comes from $2 * 19/3 + 5 * 16/3 + ... + 7 * 14/3 + 9 * 4$.

[40] This is true if we have the population, rather than just a sample, of members of each group sampled. That is, if types are substitutes and the theory holds, so that the true grouping of $2n$ borrowers minimizes expression 12, *a sample* from the true groups may not minimize expression 13 across all groupings of the sampled borrowers. Though we expect the sample analog to at least come close to minimizing its expression if the full-population grouping does, this is a caveat to our results.

[41] Both are incorporated simultaneously in section 5.

dummy in $\{\kappa^G_{i,j} | j \in \mathcal{S}^G \setminus i\}$, the key expression for sum of group payoffs due to correlated risk is

$$\sum_{i \in \mathcal{S}^L} \overline{\kappa}^L_{i,\mathcal{S}^L_{-i}} + \sum_{i \in \mathcal{S}^M} \overline{\kappa}^M_{i,\mathcal{S}^M_{-i}}.$$

The interpretation is simple. For a given borrower, $\overline{\kappa}^G_{i,\mathcal{S}^G_{-i}}$ is the fraction of other same-group, sampled borrowers that have the same risk exposure-type; this fraction is then summed over all village borrowers to get the relevant part of the sum of group payoffs. Again, compare village 1 with $L = (A, A, B, B)$ and $M = (A, A, B, B)$ to village 2 with $L' = (A, A, A, B)$ and $M' = (A, B, B, B)$. The correlation-related payoffs sum to 2.67 in the first village and to 4 in the second, more anti-diversified (homogeneous) village.[42]

Using these model-based sorting metrics, the procedure is as before: use permutation tests to calculate sorting percentile ranges for each village's observed grouping, then use the KS test to compare the sample CDFs of village sorting percentiles to the uniform distribution.

The remaining question in this approach is how to use our data to proxy for $\kappa^G_{i,j}$, the correlatedness dummy. In the case of worst_year (see section 3 for descriptions of these measures), we proxy $\kappa^G_{i,j}$ simply by $1\{worst\_year^G_i = worst\_year^G_j\}$; that is, if the two borrowers (do not) name the same year as worst, we say they are (not) exposed to the same risk. In the case of shock, a continuous variable, we proxy $\kappa^G_{i,j}$ by $e^{-|shock^G_i - shock^G_j|}$; that is, our assessment of the probability two borrowers are exposed to the same risk is one if their shocks exactly coincided and decreasing toward zero in the distance between their shocks. In the case of occupation, a ten-entry vector with the fraction of total revenue coming from each of ten agricultural areas, $\kappa^G_{i,j}$ is measured as the dot product of the borrower's vectors; see section 3.2 for explanation.

4.2 Univariate Results

Sorting by risk-type. The probability of achieving the high income realization, **prob**, is a

[42]In village 1, for example, 1/3 of each borrower's fellow group members is exposed to the same shock; summing 1/3 across 8 borrowers gives 2.67.

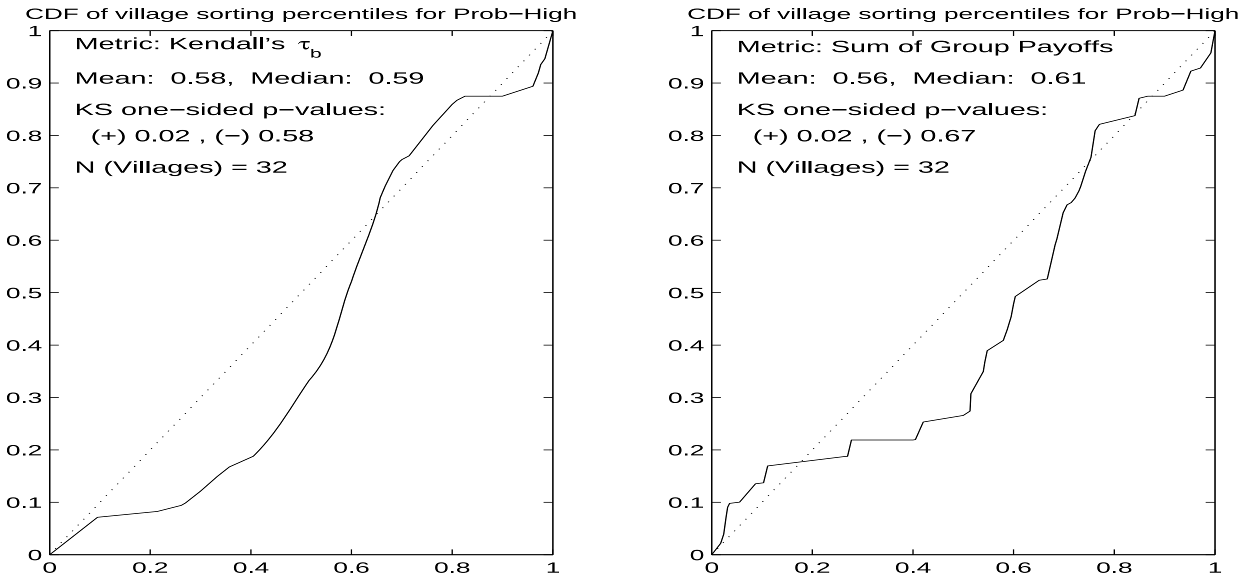
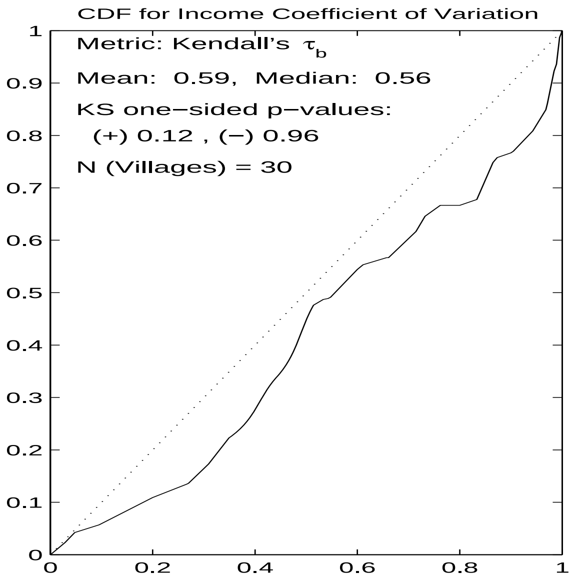


Figure 1: Sample CDF of villages' sorting percentiles based on rank correlation Kendall's tau$_b$ (left panel) and the sum of group payoff functions (right panel) for *prob*, the probability of realizing high income.

close analog to the risk-type variable in the theory and is thus the focus of our empirical tests for homogeneous risk-matching. The sample CDFs of village sorting percentile ranges for *prob* based on the Kendall's tau$_b$ and the structural sorting metric, respectively, are graphed in Figure 1.[43] Based on percentile ranges from the rank correlation, the mean (median) village is more homogeneously sorted than 58% (59%) of all possible combinations of borrowers into groups of the observed sizes. The random-matching benchmark, the uniform, is graphed as a dotted line. Using a one-sided KS test, we reject at the 5% level the hypothesis of heterogeneous sorting, that is, that the true distribution of village sorting percentiles is first-order stochastically dominated by the uniform.[44] These results point to risk-matching that, while not rank-ordered, is statistically distinguishable from random matching in the direction of homogeneity.

[43] The reported p-values are averages over 1 million KS p-values based on random draws from each village's sorting percentile range. The sample CDFs graphed are essentially averages over an infinite number of sample CDFs constructed based on these random draws; equivalently, they incorporate the sorting percentile *range* of each village directly. Means and medians are computed using these sample CDFs.

[44] Results using the variance decomposition sorting metric are similar: mean (median) of 57% (62%), and KS one-sided (+) p-value of 0.01.



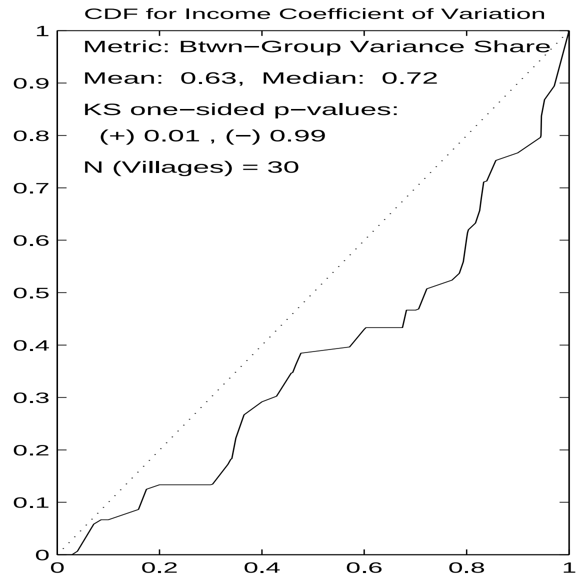


Figure 2: Sample CDF of villages' sorting percentiles based on rank correlation Kendall's tau_b (left panel) and variance decomposition (right panel) of the coefficient of variation for income (standard deviation / mean).

The results using the structural sorting metric, which is based on the model's specific payoff function, are quite similar in this case to the atheoretic results. The mean (median) sorting percentile is 56% (61%) and substitutability of types (heterogeneous matching) is rejected at the 5% level. These results help alleviate concerns that, by relying on an atheoretic sorting metric that captures heterogeneity in general but not the specific kind defined by the actual payoff function and borrower types, we are falsely rejecting negative assortative matching (which, as section 2.2 makes clear, does not by itself restrict the outcome much).

A second measure of risk, though in a way not as closely related to the theory, is the **coefficient of variation** of projected income, described in section 3. The sample CDFs of village sorting percentile ranges for the coefficient of variation based on Kendall's tau_b and variance decomposition are graphed in Figure 2. Here, the variance decomposition gives strong evidence of homogeneous sorting (compared to random matching): the mean (median) village is more homogeneously sorted than 63% (72%) of all possible borrower groupings, and heterogeneous sorting is rejected at the 5% level. However, when judged by the rank

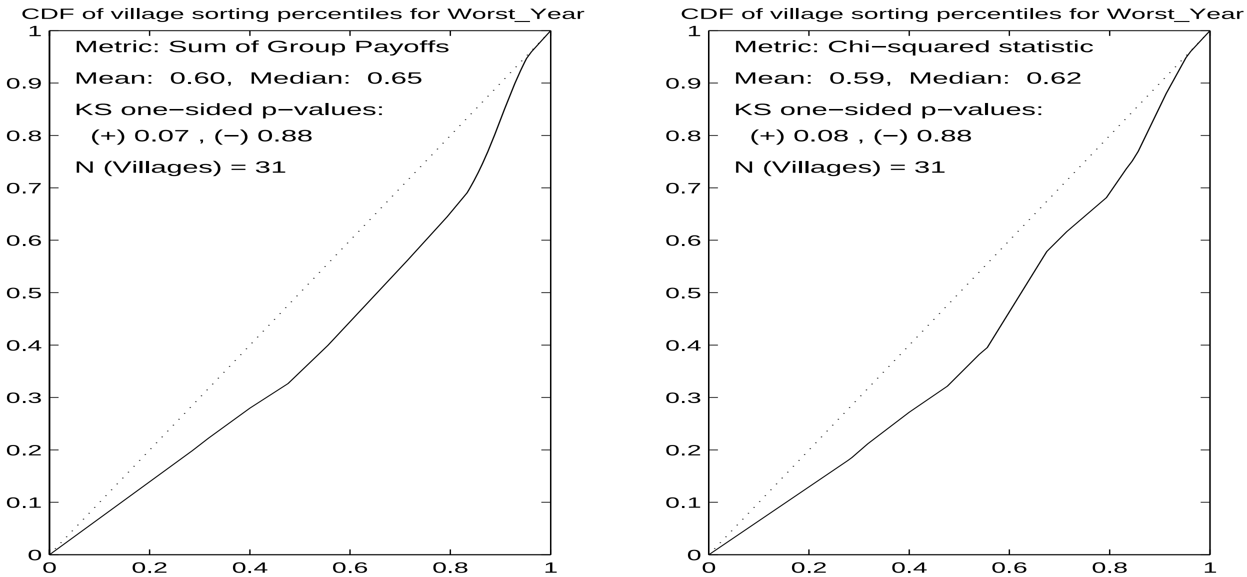


Figure 3: Sample CDF of villages' sorting percentiles based on the sum of group payoff functions (left panel) and the chi-squared statistic (right panel) of the worst_year for income.

correlation there is less evidence for homogeneous sorting by coefficient of variation. The means and medians drop to 59% and 56%, respectively, and the KS tests come somewhat close but fail to reject heterogeneous sorting at the 10% level. While the coefficient of variation measure gives weaker results, we view it as auxiliary to the *prob* measure, and somewhat supportive.

Overall, the data give solid evidence for a non-negligible degree of homogeneous risk-matching, and are typically able to reject heterogeneous matching. Based on the results of section 2.2, this offers support to Ghatak's (1999) formulation of group liability and suggests that direct liability is not negligible relative to dynamic liability in this context. One explanation for this, even if dynamic liability is a preferable approach, is that the BAAC may lack credibility when threatening to deny future loans, given its political mandate to maximize outreach to Thai farmers.

Sorting by correlated risk. We next examine diversification within groups. Consider the **worst_year** measure. This is a categorical variable, so the atheoretic sorting metric is the chi-squared test statistic. The structural sorting metric boils down to the sum of each

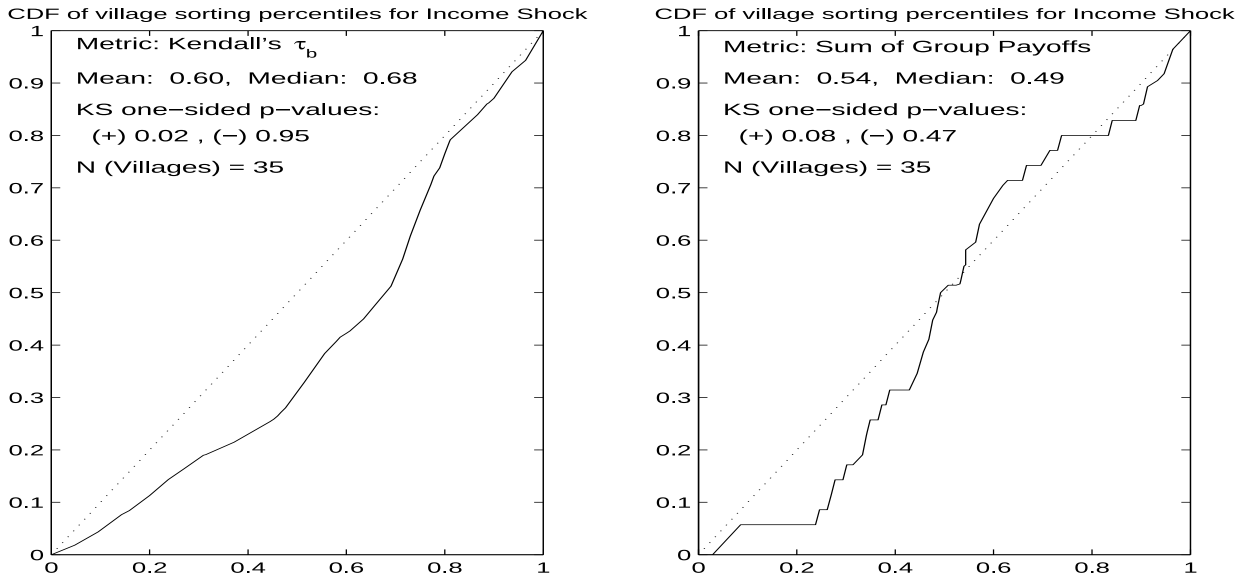


Figure 4: Sample CDF of villages' sorting percentiles based on rank correlation Kendall's tau$_b$ (left panel) and the sum of group payoff functions (right panel) for income shocks.

borrower's fraction of fellow group members that report the same worst year for income. Both results are reported in figure 3. Using the structural metric, the average (median) village is more homogeneously sorted than 60% (65%) of villages; using the chi-squared metric, the average (median) village is more homogeneously sorted than 59% (62%) of villages. In both cases, heteregenous matching (diversification) is rejected by the KS test at the 10% level.

Next, consider coincidence of income **shock**s, where the shock is measured by the (signed) percent deviation of this year's realized income from next year's expected income. Results using the rank correlation and the structural sorting metric are presented in Figure 4. The rank correlation metric yields a mean (median) of 60% (68%) and rejects diversification at the 5% level.[45] The structural metric produces a mean (median) of 54% (49%), and rejects heterogeneous matching (diversification) at the 10% level.

We turn finally to **occupation**al diversification, using the chi-squared and the structural sorting metrics. Results are graphed in Figure 5. Interestingly, they suggest that matching is close to random based on occupation, where occupation is measured by shares of total

[45] Results using the variance decomposition sorting metric are similar, if slightly weaker statistically: mean 57%, median 66%, KS one-sided (+) p-value 0.06.

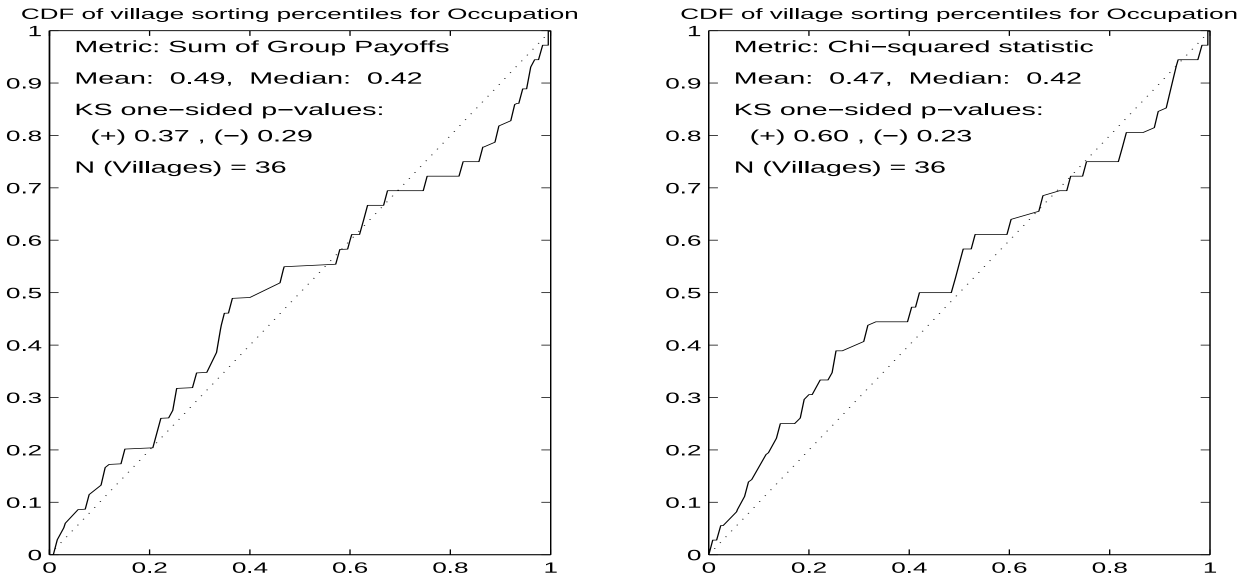


Figure 5: Sample CDF of villages' sorting percentiles based on the sum of group payoff functions (left panel) and the chi-squared statistic (right panel) for occupation.

revenue coming from ten different agricultural categories. The means and medians are in the 40%'s, and the KS p-values are lower in the test against homogeneous matching (anti-diversification); however, neither diversification nor anti-diversification can be rejected at better than a 20% significance level.

With regard to correlated risk, the results for worst_year and shock suggest that borrowers have incomes that are somewhat anti-diversified along group lines; but the results for occupation suggest that this anti-diversification does not take the form of (agricultural-)occupational homogeneity. A potential interpretation is that the lender encourages (observable) diversification within groups, including by agricultural occupation, but that the borrowers are able to achieve some anti-diversification by exploiting other, unobservable traits.

Discussion. The univariate tests suggest that group composition is homogeneous – not perfectly, but moreso than under random matching – along dimensions of unconditional risk and of correlated risk exposure, as predicted by the baseline Ghatak (1999) theory and the extended version with correlated risk. They typically reject heterogeneous risk-matching, which, in light of the dynamic extension (section 2.2) based on Sadoulet (1999), suggests that direct liability is not insignificant relative to dynamic liability in this context.

Of course, the evidence is not proof of causality running from risk amount or exposure type to sorting. For example, it may simply be that friends or relatives group together, friends or relatives that are alike in certain regards, including along risk dimensions. Or, perhaps monitoring is easier within a group of similarly-occupied individuals, who by nature of their occupation face similar amounts and types of risk (though the lack of observed occupational homogeneity casts doubt on this particular story).

However, for testing certain key aspects of the baseline Ghatak (1999) model – in particular, for assessing whether this story of group lending is an empirically plausible (partial) explanation of its popularity and of any ability it has to revive credit markets – these simple univariate results are in some ways preferable to alternatives. The reason is that the safe borrower discount embedded in lending to risk-homogeneous groups exists *regardless of how* groups end up homogeneously sorted by risk. Borrowers may have consciously considered the risk of their partners in forming groups, or they may have simply formed groups with friends or relatives who happened to have similar risk characteristics; either way, safe borrowers end up with safer partners. Given this homogeneous risk-matching, then, the joint liability stipulation is less onerous for safe borrowers, meaning they get an implicit discount in their borrowing rate. It is precisely this discount, which prices risk more accurately, that is the essential advantage of group lending and allows it to draw more borrowers into the market. The point is that, in this framework, matching that is homogeneous by risk – by whatever mechanism – is all that is needed for group lending to offer an improvement in contracting.

Thus, testing directly the degree of risk homogeneity is arguably the most appropriate approach to testing the main idea of the Ghatak (1999) model. As a corollary, rejecting Ghatak's main idea based on causally identifying, e.g., kinship and not risk-type as the key sorting determinant would appear to be misguided, if the evidence pointed to risk-homogeneous groups (as it does here). Similarly, rejecting homogeneous risk-sorting based on a zero coefficient in a multivariate regression does not necessarily reject the main idea of the Ghatak model if the coefficient is positive in a univariate regression.

A similar argument can be made about the extended model that incorporates correlated risk. If there is unconditional evidence for anti-diversification of risk, then that is enough to raise the concern that some of the contractually stipulated joint liability is being undone – whether or not the anti-diversification is a conscious choice on the part of borrowers.

All this is not to say that for testing other aspects of the Ghatak (1999) model and extensions, namely the predictions about what drives sorting behavior, establishing causal links would be ideal. Understanding sorting behavior seems an important question in its own right, as well as in the context of the Ghatak model, since his overall argument seems more robust if the risk-homogeneity is being achieved purposefully rather than as an unintended consequence. Unfortunately, establishing causal sorting links in an incontrovertible way is beyond the scope of this paper. However, in the next section we take a step in that direction by employing a multivariate approach.

5 Multivariate Methodology and Results

The univariate results are consistent with both dimensions of risk – amount of risk and type of exposure – being important for matching. One might wonder, though, if one dimension of homogeneity is driving the other. For example, if some types of shocks are more volatile than others, then matching based on similar shock exposure would give rise to homogeneous matching based on unconditional risk levels also. Is there evidence that both dimensions of risk enter significantly into payoffs and affect borrowers' sorting behavior?

We turn next to a multivariate approach that allows both dimensions of risk simultaneously to affect payoffs and sorting behavior. In particular, we use the matching maximum score estimator of Fox (2008).[46] This estimator can be used with an atheoretic or a structural

[46] Fafchamps and Gubert (2007) pioneer a different multivariate empirical approach to group formation based on the dyadic regression.

approach; here we take the latter route and derive the specification from the exact payoff function of the model.[47]

The estimator works by choosing parameters that most frequently give observed agent groupings higher joint surplus (sum of payoffs) than feasible, unobserved agent groupings. Thus the estimator exploits the idea that in an environment with no search frictions and transferable utility, like ours, observed groupings maximize total surplus, relative to feasible alternatives.

Consider observed groups L and M in village v. Let L' and M' denote an alternative arrangement of the borrowers from L and M into two groups of the original sizes. As in section 4.1.1, we assume that borrowers can match with any others in their village; thus L' and M' represent a feasible, unobserved grouping. If $\Pi_G(\phi)$ gives the sum of payoffs of any group G as a function of parameters ϕ, theory predicts

$$\Pi_L(\phi) + \Pi_M(\phi) \geq \Pi_{L'}(\phi) + \Pi_{M'}(\phi). \tag{14}$$

The matching maximum score estimator chooses parameters ϕ that maximize the score, i.e. the number of inequalities of the form 14 that are true, where each inequality corresponds to a different unobserved grouping L', M'.

Our set of unobserved groupings, and thus inequalities, comes from all k-for-k borrower swaps across two groups in the same village.[48] For example, if we have data on five borrowers in each of two groups in the same village, there are $5\times5 = 25$ one-for-one swaps, $10\times10 = 100$ two-for-two swaps, and so on. One could contemplate other kinds of unobserved groupings, for example those arising from a k-borrower *transfer*. We choose not to use transfers because they change group size, which was held fixed in the theory.

Consider the model's group payoff function from section 2.3, reproduced from equation 11

[47] A reduced-form estimation that included more controls than we use could also be interesting. However, this is not as attractive in part due to the limited scope of our (re-survey) data; for example, it lacks any data on social networks.

[48] If the larger group in a village has sample size m and the smaller group has sample size n, k is capped at $\min\{n, m-1\}$.

here:

$$\Pi_G = nE - [r + q(n-1)] \sum_{i \in \mathcal{N}} p_i^G + q \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}_{-i}} p_i^G p_j^G + q\epsilon \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}_{-i}} \kappa_{i,j}^G \ .$$

Note that all terms in the group payoff function that do not involve interactions between borrower characteristics drop out of inequality 14, since they appear identically on both sides;[49] hence, we can ignore the non-interaction terms.

We proceed as in section 4.1.3. There, $\overline{p}_{\mathcal{N}_{-i}}^G$ is defined as the average risk-type in group G excluding borrower i, and $\overline{\kappa}_{i,G_{-i}}^G$ is defined as the average across fellow group members of borrower i's correlatedness dummy. Scaling by $1/(n-1)$, payoffs can be written

$$\Pi_G = q \sum_{i \in \mathcal{N}} p_i^G \overline{p}_{\mathcal{N}_{-i}}^G + q\epsilon \sum_{i \in \mathcal{N}} \overline{\kappa}_{i,G_{-i}}^G \ .$$

Since our data for each group represent only a subset of the group (up to 5 members), we use a sample analog expression for the payoff function. Again using notation from section 4.1.3, let $\mathcal{S}^G \subseteq \mathcal{N}$ be the sampled subset of group G, and $\overline{p}_{\mathcal{S}_{-i}^G}^G$ and $\overline{\kappa}_{i,\mathcal{S}_{-i}^G}^G$ be defined analogously to $\overline{p}_{\mathcal{N}_{-i}}^G$ and $\overline{\kappa}_{i,G_{-i}}^G$. Then, the sample analog to the payoff function is

$$\Pi_{\mathcal{S}^G} = q \sum_{i \in \mathcal{S}^G} p_i^G \overline{p}_{\mathcal{S}_{-i}^G}^G + q\epsilon \sum_{i \in \mathcal{S}^G} \overline{\kappa}_{i,\mathcal{S}_{-i}^G}^G \ . \tag{15}$$

It is this expression that we use for group payoffs in the inequality 14.[50]

Given data on borrower probabilities of success (p_i^G's) and correlatedness ($\kappa_{i,j}^G$'s), the parameters q and $\tilde{\beta} \equiv q\epsilon$ can be estimated, but only up to scale, since multiplication by any positive scalar would preserve the inequality. Note that ϵ would be identified as $\tilde{\beta}/q$.

[49] Thus coefficients on non-interaction payoff function terms (e.g. E, r) cannot be estimated.

[50] That is, interaction terms involving only sampled borrowers are used to estimate the group payoff function. Similarly, the counterfactual groups are formed via k-for-k borrower swaps across these sampled sub-groups. As discussed in footnote 40, sorting optimality conditions (here, inequality 14) need not hold for subsets of groups. Specifically, even if inequality 14 holds for groups L, M and every counterfactual grouping L', M', it need not hold for randomly selected subsets of L, M and every counterfactual grouping of these subsets. This is a caveat to our results. However, though Fox (2008) does not address this issue, it seems a reasonable conjecture that the parameters that maximize the probability that inequality 14 holds using all group members, i.e. the true parameters, also maximize the probability that inequality 14 holds using randomly selected subsets of groups.

This approach, however, requires data that can capture the existence of correlation ($\kappa_{i,j}^G$) as distinct from the intensiveness (ϵ) of correlation. That is, to identify ϵ, $\kappa_{i,j}^G$ should reflect the similarity of shocks to which borrowers are exposed, but not the degree of exposure to those shocks. Our measures of correlatedness (coincidence of income shocks and occupation) probably cannot be assumed to distinguish existence and intensiveness of correlatedness.

Rather than attempt to identify ϵ separately from $\kappa_{i,j}^G$, we focus on the overall correlation between borrowers i and j in group G, call it $C_{i,j}^G \equiv \epsilon\kappa_{i,j}^G$. $C_{i,j}^G$ is proxied in different ways, depending on the variable used (see section 4.1.3). When worst_year is used, $C_{i,j}^G = \phi_{wst}\,1\{worst\_year_i^G = worst\_year_j^G\}$. When income shock is used, $C_{i,j}^G = \phi_{shk}\,e^{-|shock_i^G - shock_j^G|}$. When occupation is used, $C_{i,j}^G = \phi_{occ}(\vec{occ}_i \cdot \vec{occ}_j)$ (i.e. the dot product of the occupational vectors of borrowers i and j). The ϕ parameters are assumed strictly positive. Thus, correlatedness is proxied by similarity in income shocks, bad income years, and/or (agricultural) occupations.

Incorporating $C_{i,j}^G$ – for concreteness, proxied here using worst_year – and notation similar to the above into the sampled group payoff function 15 gives

$$\begin{aligned}\Pi_{\mathcal{S}^G} &= q\sum_{i\in\mathcal{S}^G} p_i^G \overline{p}_{\mathcal{S}_{-i}^G}^G + q\sum_{i\in\mathcal{S}^G}\overline{C}_{i,\mathcal{S}_{-i}^G}^G \\ &= \sum_{i\in\mathcal{S}^G}\left(\beta_1 p_i^G\overline{p}_{\mathcal{S}_{-i}^G}^G + \beta_2\overline{1\{worst\_year_i^G = worst\_year_{\mathcal{S}_{-i}^G}^G\}}\right),\end{aligned} \tag{16}$$

where $\beta_1 = q$, $\beta_2 = q\phi_{wst}$.[51] Parameter q can thus be identified in sign but not magnitude; hence, β_1 is normalized to $+1$ or -1 in estimation.

The main test of the Ghatak (1999) theory and our extension is whether all β's are positive. The model assumes that $q > 0$, which underlies complementarity of types in the payoff function and hence drives homogeneous matching. A positive estimate of β_1 is thus direct evidence for this complementarity, while a negative estimate would imply substitutability of types and that more heterogeneous patterns of matching are being observed. Regarding

[51] The term multiplying β_2 is simply the fraction of other same-group, sampled borrowers naming the same worst_year as borrower i.

$\beta_2(= q\phi_{wst})$, since ϕ_{wst} (and ϕ_{shk} and ϕ_{occ}) are restricted to be positive and since $q > 0$ is assumed, the model requires a positive estimate for β_2. A negative estimate would contradict the model; in particular, it would suggest that sorting is more consistent with payoffs that value diversification (less correlatedness) rather than anti-diversification, in contrast to our theory.

Risk types p_i^G, p_j^G are measured by *prob*, discussed in section 3. Correlatedness is proxied by various subsets of the three measures discussed, worst_year, shock, and occupation. If there are V villages indexed by v, and each village v has two (sampled) groups, L_v and M_v, the estimator comes from

$$max_{\beta_1 \in \{-1,1\}, \beta_2(,\beta_3)} \sum_{v=1}^{V} \sum_{L_v', M_v'} 1\{\Pi_{L_v} + \Pi_{M_v} > \Pi_{L_v'} + \Pi_{M_v'}\},$$

where[52] the alternate groupings L_v' and M_v' come from all k-for-k borrower swaps, as discussed above, and there are three parameters when two proxies for correlatedness are included.

We also estimate based on a slightly different objective function, where the score is the sum of all villages' *shares* of correct inequalities rather than *numbers* of correct inequalities; that is, the indicator function for village v is normalized by the total number of inequalities for village v. This weights each village equally in its contribution to the estimation and provides a more similar basis of comparison with the univariate KS results, where each village counts as a single draw from a distribution.[53]

Maximization is carried out using the genetic algorithm routine in Matlab. Results from

[52] The estimator uses a strict inequality though theory requires only a weak one. Given a continuous distribution of match-specific error terms introduced to support the estimator, equalities can be ignored with probability one.

[53] The approach here and the univariate approach have similarities and differences. Both essentially compare the observed grouping to unobserved alternatives. The univariate approach proceeds by putting a metric on this comparison (e.g. variance decomposition) and testing against a random-matching benchmark. This approach could be applied in the multivariate setting *if* we knew the relative importance of the multiple dimensions; in equation 16 above, this is equivalent to knowing β_2. In this case, we could calculate sorting percentiles for each village's grouping based on the full, multi-variate payoff function (up to scale), and then proceed to the KS test. However, we do not know β_2, and it is the matching maximum score estimator that provides a way to estimate it.

Table 1 — Maximum Score Matching Estimation

Variable		Number	Share	Number	Share	Number	Share	Number	Share
Prob	Est.	+1	+1	+1	+1	+1	+1	+1	+1
	p-val.				Superconsistent				
Worst_Year	Est.	0.40	0.32*					0.32	0.33
	p-val.	0.18	0.06					0.27	0.31
Income Shock	Est.			-0.46**	0.0037				
	p-val.			0.02	0.30				
Occupation	Est.					-0.062	-0.0029	-0.051	-0.025
	p-val.					0.46	0.47	0.34	0.51
Number of Inequalities		3620		3620		3620		3620	
Number of Villages			32		32		32		32
Maximized Objective Fn.		2348	19.3	2208	18.4	2205	18.3	2382	19.8
Percent Correct		65%	60%	61%	58%	61%	57%	66%	62%

Note: Each column corresponds to a different estimation; differences arise from the objective function used (noted atop each column) and the proxies for correlated risk. P-values are from one-sided tests for a negative (positive) true parameter if the point estimate is positive (negative). They are constructed using subsampling methods on 200 subsamples, each containing 24 distinct villages. Significance at 5% and 10% levels is denoted by ** and *, respectively.

eight estimations that alternately use the two objective functions combined with four sets of proxies for correlated risk are reported in Table 1. The point estimates are based on the 32 villages with sufficient data, and the corresponding 3620 total inequalities. As suggested by Fox (2008), inference is carried out by subsampling.[54]

We find that the estimated coefficient on risk-type, measured by Prob, is consistently positive. Thus, even when controlling for correlated risk measures, including occupational similarity,[55] unconditional risk has positive explanatory power for group formation. It also supports the model, since a positive estimate implies *complementarity* of risk-types in the payoff function, which is the basis for homogeneous matching and the Ghatak (1999) theory.

[54] He notes that the bootstrap is proved inconsistent by Abrevaya and Huang (2005) for a class of estimators that converge at rate $\sqrt[3]{n}$, which almost certainly includes the matching maximum score estimator. Thus, for each estimation, we create 200 subsamples containing 24 villages' data, by randomly sampling without replacement from the 32 villages. Estimation is carried out for each subsample. Operating under the assumption of $\sqrt[3]{n}$-convergence, one can apply the distribution of $\left(\frac{24}{32}\right)^{1/3}(\hat{\beta}_{24,i} - \hat{\beta}_{32})$ to $(\hat{\beta}_{32} - \beta_0)$ to construct confidence intervals, where $i \in \{1, ..., 200\}$ corresponds to the subsamples, $\hat{\beta}_{24,i}$ are the subsample estimates, $\hat{\beta}_{32}$ is the full-sample estimate, and β_0 is the true parameter. See Politis et al. (1999, Section 2.2).

[55] This reduces the likelihood that the risk-homogeneity results are being driven by borrowers sorting by occupation in order to keep monitoring costs low.

The correlated risk results seem less clear. The estimates for occupation are not statistically different from zero; this is no different from the univariate results, which were discussed in section 4.2. One estimate for income shock is actually significantly negative, which would contradict the theory, but the estimate using share rather than number is slightly positive. Given the relative instability of the shock estimate, and since neither shock nor occupation have as much explanatory power as worst_year (see last two rows of Table 1), worst_year seems the preferred proxy for correlated risk.

The results using worst_year are more salient, if somewhat weak statistically. The estimates are relatively stable in the $0.32 - 0.40$ range, and a negative coefficient can be rejected at the 10% level in the share case and a 20%-level in the number case. When combined with occupation, the estimates do not change but the significance levels drop. We interpret these results are mildly supportive of both aspects of the sorting theory; they suggest that group payoffs are higher with greater homogeneity on both unconditional risk and correlated risk dimensions. Though the other two proxies are not as supportive of the theory, the case can be made – based on the results of Table 1 and caveats in section 3.2 – that they are inferior proxies for correlated risk.

In summary, the evidence from these tests is consistent with the theory in which the incentives are to match homogeneously along lines of both risk-type and risk exposure. Controlling for correlated risk, matching appears homogeneous by risk-type. There is also some evidence that controlling for risk-type, anti-diversification of risk raises group payoffs and thus becomes a goal in group formation.

6 Conclusion

In the context of direct group liability lending and unobserved risk type, theory (Ghatak, 1999) suggests that borrowers will sort homogeneously by risk; this embeds an effective discount for safe borrowers and improves efficiency. However, theory also suggests that different forms of group liability may lead to heterogeneous risk-matching, and that borrowers may sort to anti-diversify risk and thereby to minimize potential liability for their fellow group members. While the first kind of sorting works in favor of efficiency, the second and third kind may work against it by limiting the lender's ability to use group lending effectively.

We test these matching predictions using data from Thai borrowing groups. The first and third predictions are upheld to some degree. Direct comparisons to completely random matching using Kolmogorov-Smirnov tests give evidence that groups are more homogeneous than random in unconditional risk and in types of risk exposure. Multivariate tests using Fox's (2008) matching maximum score estimator give some confirmation that the payoff to similarity is positive, in both dimensions. Thus voluntary sorting appears to be resulting in a discount to safe borrowers, but also may be limiting effective liability via anti-diversification.

These results provide direct evidence on a mechanism by which group lending can improve efficiency in micro-lending markets. They add to our understanding of how innovations in lending have been able to extend finance to the world's poor and why group lending has been such a popular lending mechanism in the microcredit movement.

The anti-diversification results of this paper shed light on the common view of microcredit groups mainly as risk-sharing mechanisms. While the lender may prefer as diversified a group as possible – because this causes joint liability to bind more frequently and, thus, enhances the safe-borrower discount – groups themselves have (private) incentives to purposefully form groups that cannot share risk well. Thus, via sorting incentives, group contracts may limit the effectiveness of microcredit groups as risk-sharing mechanisms.

From a policy standpoint these results show that voluntary sorting by borrowers may also have its downside. Sorting to anti-diversify can work against the lender's interests and, in equilibrium, the borrowers'. Applying the results narrowly, lenders may want to intervene to promote risk diversification within groups – for example, requiring occupational diversity – provided their intervention will not prevent homogeneous risk matching. More generally, lenders may wish to step in with respect to group composition on certain dimensions while leaving other dimensions to the borrowers' discretion. Accomplishing this in a precise fashion may be impossible; however, attempts to do so may be some of the profitable future steps of lender experimentation in extending credit markets to the poor.

References

[1] Jason Abrevaya and Jian Huang. On the bootstrap of the maximum score estimator. *Econometrica*, 73(4):1175–1204, July 2005.

[2] Christian Ahlin and Neville Jiang. Can micro-credit bring development? *Journal of Development Economics*, 86(1):1–21, April 2008.

[3] Christian Ahlin and Robert M. Townsend. Using repayment data to test across theories of joint liability lending. *Economic Journal*, 117(517):F11–F51, February 2007a.

[4] Christian Ahlin and Robert M. Townsend. Selection into and across contracts: Theory and field research. *Journal of Econometrics*, 136(2):665–698, February 2007b.

[5] Beatriz Armendariz de Aghion and Jonathan J. Morduch. *The Economics of Microfinance*. MIT Press, Cambridge MA, 2005.

[6] Eric Bellman. Invisible hand: Entrepreneur gets big banks to back very small loans. *Wall Street Journal*, page A1, May 15, 2006.

[7] Pierre-Andre Chiappori and Philip J. Reny. Matching to share risk. Working Paper, May 2006.

[8] Morris H. DeGroot. *Probability and Statistics*. Addison Wesley, 2nd edition, 1986.

[9] Thomas A. Dowling and Douglas R. Shier. Special sequences. In Kenneth H. Rosen, John G. Michaels, Jonathan L. Gross, Jerrold W. Grossman, and Douglas R. Shier, editors, *Handbook of Discrete and Combinatorial Mathematics*. CRC Press, 2000.

[10] Marcel Fafchamps and Flore Gubert. The formation of risk sharing networks. *Journal of Development Economics*, 83(2):326–350, July 2007.

[11] Jeremy T. Fox. Estimating matching games with transfers. Working Paper, September 2008.

[12] Shubhashis Gangopadhyay, Maitreesh Ghatak, and Robert Lensink. On joint liability lending and the peer selection effect. *Economic Journal*, 115(506):1005–1015, October 2005.

[13] Maitreesh Ghatak. Group lending, local information and peer selection. *Journal of Development Economics*, 60(1):27–50, October 1999.

[14] Maitreesh Ghatak. Screening by the company you keep: Joint liability lending and the peer selection effect. *Economic Journal*, 110(465):601–631, July 2000.

[15] Maitreesh Ghatak and Timothy W. Guinnane. The economics of lending with joint liability: Theory and practice. *Journal of Development Economics*, 60(1):195–228, October 1999.

[16] Jean Dickinson Gibbons and Subhabrata Chakraborti. *Nonparametric Statistical Inference*. Marcel Dekker, New York, 4th edition, 2003.

[17] Joel M. Guttman. Assortative matching, adverse selection, and group lending. *Journal of Development Economics*, 87(1):51–56, August 2008.

[18] Joseph Kaboski and Robert M. Townsend. Borrowing and lending in semi-urban and rural Thailand. Mimeo, 1998.

[19] Patrick Legros and Andrew F. Newman. Monotone matching in perfect and imperfect worlds. *Review of Economic Studies*, 69(4):925–942, October 2002.

[20] Jonathan Morduch. The microfinance promise. *Journal of Economic Literature*, 37(4):1569–1614, December 1999.

[21] Dimitris N. Politis, Joseph P. Romano, and Michael Wolf. *Subsampling*. Springer, 1999.

[22] Loic Sadoulet. Equilibrium risk-matching in group lending. Working paper, June 1999.

[23] Joseph E. Stiglitz and Andrew Weiss. Credit rationing in markets with imperfect information. *American Economic Review*, 71(3):393–410, June 1981.

Appendix

Proof of Proposition 1. Note from equation 4 that Π_G is twice continuously differentiable in the p_i^G's and that $\partial^2\Pi_G/\partial p_i^G \partial p_j^G = 2q > 0$; it follows that Π_G is strictly supermodular.[56] Also, Π_G is symmetric in the p_i^G's (i.e. all that matters is the list of group members, not the order).

Assume not all equilibrium groups are rank-ordered. Then two groups exist, L and M, such that $\overline{p}^L > \underline{p}^M$ and $\overline{p}^M > \underline{p}^L$. Let L' and M' be (re-)orderings of L and M satisfying $p_1^{L'} \leq p_2^{L'} \leq \ldots \leq p_n^{L'}$ and $p_1^{M'} \geq p_2^{M'} \geq \ldots \geq p_n^{M'}$. The fact that L and M are not rank-ordered implies that $p_1^{L'} < p_1^{M'}$ and that $p_n^{M'} < p_n^{L'}$; hence $L' \ngeq M'$ and $M' \ngeq L'$. Finally, let $L'' = L' \wedge M'$ and $M'' = L' \vee M'$. Then

$$\Pi_{L''} + \Pi_{M''} > \Pi_{L'} + \Pi_{M'} = \Pi_L + \Pi_M,$$

where the inequality follows from strict supermodularity of Π_G and the equality from symmetry of Π_G. Since L'' and M'' represent an alternative, feasible grouping of the $2n$ borrowers that produces higher total payoffs, this contradicts L and M being equilibrium groups. ■

Proof of Corollary 1. A strictly positive measure of rank-ordered groups with strictly positive width (i.e. difference between the highest and lowest risk-type in the group) cannot be drawn from the bounded interval $[\underline{p}, \overline{p}]$. ■

Proof of Proposition 2. If not, then there are equilibrium groups L and M such that M rank-wise dominates L, say, and L and M are not nearly rank-wise identical. Let L' and M' be (re-)orderings of the types in L and M, respectively, satisfying $p_1^{L'} \leq p_2^{L'} \leq \ldots \leq p_n^{L'}$ and $p_1^{M'} \leq p_2^{M'} \leq \ldots \leq p_n^{M'}$. Since M rank-wise dominates L, $p_i^{L'} \leq p_i^{M'}$ for all $i \in \mathcal{N}$; further, since L and M are not nearly rank-wise identical, there at least two strict inequalities, say for $j, k \in \mathcal{N}$. Let L'' and M'' be groups formed from L' and M' by interchanging $p_j^{L'}$ and $p_j^{M'}$: $L'' = (p_1^{L'}, \ldots, p_{j-1}^{L'}, p_j^{M'}, p_{j+1}^{L'}, \ldots, p_n^{L'})$ and $M'' = (p_1^{M'}, \ldots, p_{j-1}^{M'}, p_j^{L'}, p_{j+1}^{M'}, \ldots, p_n^{M'})$. Note that rank-wise dominance implies $L' = L'' \wedge M''$ and $M' = L'' \vee M''$. Also, $M'' \ngeq L''$ (since $p_j^{L'} < p_j^{M'}$) and $L'' \ngeq M''$ (since $p_k^{L'} < p_k^{M'}$). Thus

$$\Pi_{L''} + \Pi_{M''} > \Pi_{L'} + \Pi_{M'} = \Pi_L + \Pi_M,$$

where the inequality follows from strict submodularity of Π_G (which follows from the smoothness and strictly negative cross-partial of Π_G) and the equality from symmetry of Π_G. Since L'' and M'' represent an alternative, feasible grouping of the $2n$ borrowers that produces higher total payoffs, this contradicts L and M being equilibrium groups. ■

Proof of Corollary 2. Fix any equilibrium group L. Since there are no mass points in the distribution of types, the maximum measure of groups that can be formed that are

[56] A function $f : D \subseteq \Re^n \to \Re$ is supermodular if for any $x, y \in D$, $f(x \wedge y) + f(x \vee y) \geq f(x) + f(y)$, where $x \wedge y$ and $x \vee y$ denote the component-wise minimum and maximum of x and y, respectively. It is *strictly* supermodular if the inequality is strict for any $x, y \in D$ such that $x \ngeq y$ and $y \ngeq x$. The function f is (strictly) submodular if $-f$ is (strictly) supermodular.

nearly rank-wise identical to L is zero. The result then follows from proposition 2. ■

Proof of Corollary 3. This follows from proposition 2, since no two groups can be nearly rank-wise identical when all borrowers have distinct types. ■

Proof of Corollary 4. Let $\mathcal{G}$ be the set of equilibrium groups, $\underline{p} = \sup_{G \in \mathcal{G}} \underline{p}^G$, and $\overline{p} = \inf_{G \in \mathcal{G}} \overline{p}^G$. If $\overline{p} < \underline{p}$, then there exist equilibrium groups L and M such that $\overline{p}^L < \underline{p}^M$. But then M rank-wise dominates L without being nearly rank-wise identical, contradicting proposition 2. So, $\underline{p} \leq \overline{p}$ and for any $\tilde{p} \in [\underline{p}, \overline{p}]$, each equilibrium group G has $\underline{p}^G \leq \underline{p} \leq \tilde{p}$ and $\overline{p}^G \geq \overline{p} \geq \tilde{p}$. ■

Proof of Proposition 3. For group size n and population size $2n$, the total number of groupings is $\binom{2n}{n}/2$. The division by 2 is because group labels are irrelevant and hence $\binom{2n}{n}$ counts each grouping twice. It suffices to show that the total number of groupings in which one group rank-wise dominates the other is $\binom{2n}{n}/(n+1)$.

Without loss of generality, consider $p_1 < p_2 < ... < p_{2n}$. Any grouping of these $2n$ borrowers can be expressed uniquely in a $2 \times n$ matrix as follows: each group is placed on a single row in increasing order, with the group containing p_1 in the first row.[57] Clearly, any such grouping exhibits rank-wise dominance of one group over the other iff the matrix is monotone increasing going down each column. Thus, the number of rank-wise dominance groupings is equal to the number of ways to construct a $2 \times n$ matrix of $2n$ ordered numbers that is monotone within each row and column. This is the nth Catalan number: $\binom{2n}{n}/(n+1)$. (See Dowling and Shier, 2000, pp. 145-147, especially Example 11.) ■

Proof of Proposition 4. Let n_A^G and n_B^G be the numbers of borrowers in group G exposed to shocks A and B, respectively. Then group payoffs 11 can be written

$$\Pi_G = nE - [r + q(n-1)] \sum_{i \in \mathcal{N}} p_i^G + q \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}_{-i}} p_i^G p_j^G + q\epsilon \sum_{s \in \{A,B\}} n_s^G (n_s^G - 1). \qquad (17)$$

Further, the set of equilibrium groups can be finitely partitioned into subsets in which all groups have the same number of borrowers exposed to shocks A and B, n_A and n_B, respectively. Fix any such subset that is not homogeneous by risk-exposure: i.e. fix n_A and n_B such that $n_A < n$, $n_B < n$, and $n_A + n_B > 0$. We argue that this subset of groups is of measure zero in any equilibrium.

Assume not. By arguments used in proposition 1, borrowers are rank-ordered by risk-type *within exposure-type*. That is, fix any two groups L and M in this subset: either the minimum risk-type among A-risks in L is greater than the maximum risk-type among A-risks in M, or the minimum risk-type among A-risks in M is greater than the maximum risk-type among A-risks in L; and similarly for B-risks and N-risks. If this were not so, borrowers could be exchanged across groups within an exposure-type (e.g. A-risk for A-risk) to increase payoffs from risk-type complementarities, without affecting payoffs from correlated risk.

Further, it must be that the ordering of risk-types between two groups is the same for

[57] Without the normalization that p_1 goes in the first row, each grouping has two such matrix expressions.

all exposure-types. For example, if group L has higher A-risk risk-types than group M, it must also have the higher B-risk and/or N-risk risk-types. If not, it is easy to show that swapping all borrowers within one exposure-type across the two groups to ensure that one group has dominant risk-types within all exposure-types would raise total payoffs: it would increase payoffs from risk-type complementarities, without affecting payoffs from correlated risk.

Given rank-ordering of groups within exposure-type, with the same direction across exposure-type, it must be that for any $\delta > 0$, there exist two groups in this subset such that the difference between any two A-risk, B-risk, and N-risk risk-types, respectively, in the two groups is less than δ. If this were not so, the number of groups in the subset would be finite (no more than $3/\delta$ for some $\delta > 0$), and the partition would have measure zero. Hence, fix $\delta = \sqrt{\epsilon/[8(n-1)]}$ and consider two groups L and M with differences in risk-type, within risk-exposure, strictly less than δ.

Consider switching the lowest risk-type from the dominant group, L say, with the borrower from M with the highest risk-type subject to having different exposure-type than the lowest L borrower. That is, switch the ith borrower from L such that $i = argmin_k\ p_k^L$ with the jth borrower from M such that $j = argmax_{k\ s.t.\ s_k^M \neq s_i^L}\ p_k^M$. One can show that the loss in total payoffs due to changes in risk-type complementarities (the third term in equation 17) in this re-grouping is no greater than

$$2q(p_i^L - p_j^M)\left(\sum_{x\in\mathcal{N}_{-i}} p_x^L - \sum_{y\in\mathcal{N}_{-j}} p_y^M\right). \tag{18}$$

We argue that the second parenthesized term is non-negative. One can form $(n-2)$ pairs of differences, $p_\cdot^L - p_\cdot^M$, each corresponding to two borrowers that share the same exposure-type; by rank-ordering within exposure-type, these differences are all non-negative. The remaining difference involves borrowers x and y, say, of groups L and M, respectively, with $s_y^M = s_i^L$. This difference is also non-negative since $p_x^L \geq p_i^L \geq p_y^M$, where the first inequality is by construction and the second is due to rank-ordering within exposure-type.

Consider two cases. First, if $p_i^L \leq p_j^M$, then the loss due to changes in risk-type complementarities (expression 18) is non-positive. Second, if $p_i^L > p_j^M$, then by rank-ordering within exposure-type and construction of i and j, $p_i^L \geq p_y^M$ for any $y \in \mathcal{N}$; further, $p_x^L \geq p_y^M$ for all $(x, y) \in \mathcal{N}^2$. This implies that no group-L borrower has risk-type greater than $p_i^L + \delta$ and no group-M borrower has risk-type less than $p_i^L - \delta$ (otherwise some borrower would have a risk-type more than δ away from the risk-type of *any* borrower in the other group). Hence, $p_x^L - p_y^M \leq 2\delta$ for all $(x, y) \in \mathcal{N}^2$. The loss (expression 18) is then bounded by

$$2q \cdot 2\delta[2\delta(n-1)] = 8q(n-1)\delta^2 = q\epsilon.$$

Combining the two cases, this is the upper bound for loss in total payoffs due to risk-type complementarities.

The gain in total payoffs due to increased intra-group correlated risk (the fourth term in equation 17) when switching an A-risk and a B-risk is

$$q\epsilon \sum_{s\in\{A,B\}} [(n_s+1)n_s + (n_s-1)(n_s-2) - 2n_s(n_s-1)] = 4q\epsilon,$$

and, similarly, when switching an N-risk and either an A-risk or a B-risk is $2q\epsilon$. Since this gain from anti-diversification is greater than the potential loss from decreased risk-type homogeneity, capped at $q\epsilon$ as shown above, the proposed re-grouping strictly increases total payoffs. This contradicts L and M being equilibrium groups. Thus this subset has measure zero.

Hence, almost every group is homogeneous in exposure-type: all A-risk, all B-risk, or all N-risk. Arguments used in proposition 1 and corollary 1 establish that almost every A-risk, B-risk, and N-risk group is homogeneous in risk-type. ■

Proof of Proposition 5. Let there be N groupings and $K \leq N$ unique values that arise when the given sorting metric is applied to the N groupings, with values $v_1 < v_2 < ... < v_K$. (Ties involve $K < N$.) Let n_i be the number of combinations that give rise to value v_i and N_i be the number of combinations that give rise to any value $v \leq v_i$, with $N_0 \equiv 0$; then $N_i = \sum_{k=1}^{i} n_k$ and $N_K = N$. If sorting is random, each of the N combinations of borrowers is equally likely to obtain. With probability $\pi_i \equiv n_i/N$ the realized combination will result in value v_i, leading to calculated sorting percentile range $[\frac{N_{i-1}}{N}, \frac{N_i}{N}]$.

We show next that the CDF of sorting percentiles is uniform, i.e. $F(z) = z$. Fix $z \in [0,1]$. There exists some $i \in \{1, 2, \ldots, K\}$ such that $z \in [\frac{N_{i-1}}{N}, \frac{N_i}{N}]$. Then the probability that a village's sorting percentile is less than z, i.e. $F(z)$, is the probability that its grouping leads to any value strictly less than v_i plus the probability that its grouping leads to value v_i *and* its sorting percentile picked from the uniform on $[\frac{N_{i-1}}{N}, \frac{N_i}{N}]$ is below z:

$$F(z) = \sum_{k=1}^{i-1} \pi_k + \pi_i \int_{\frac{N_{i-1}}{N}}^{z} \frac{1}{\frac{N_i - N_{i-1}}{N}}\, dz = \sum_{k=1}^{i-1} \frac{n_k}{N} + \frac{n_i}{N}\frac{N}{n_i}(z - \frac{N_{i-1}}{N}) = z,$$

where the definitions of the π_i's and the N_i's have been used in the simplification. ■