MATCHING FOR CREDIT: RISK AND DIVERSIFICATION IN THAI MICROCREDIT GROUPS

CHRISTIAN AHLIN

BREAD WORKING PAPER No. 251
DECEMBER, 2009

BREAD

Working Paper

Bureau for Research and Economic Analysis of Development

Matching for Credit: Risk and Diversification in Thai Microcredit Groups

Christian Ahlin

December 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 rouse an otherwise dormant credit market. We extend the theory 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; structural multivariate analysis is carried out using Fox's (2008) matching maximum score estimator. We find evidence of a) homogeneous sorting by risk and b) risk anti-diversification within groups, though not along occupational lines. Thus there is evidence that group lending improves risk-pricing in this context and is part of the explanation of the rise in financial intermediation among the poor. However, the anti-diversification results reveal a potentially negative aspect of voluntary group formation and point to limitations of microcredit groups as risk-sharing mechanisms.

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

[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".

most promising recent advances in economic development.[2]

Despite its advances, questions about microcredit remain. An obvious one – does it work? – seems still to have eluded a definitive answer.[3] However, there seems to be a strong 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 gains from trade are being realized.

This leads to a different 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 focused on this question.

One candidate explanation is due to Ghatak (1999, 2000). The context is a standard adverse selection environment (Stiglitz and Weiss, 1981) in which there is no collateral (limited liability) and borrowers' distributions of project returns have identical means but vary in riskiness. In this environment, a lender that cannot observe risk offers all borrowers the same terms; its inability to price for risk results in effectively lower rates for risky borrowers, who fail more often, than for safe borrowers. Thus there is cross-subsidization of risky borrowers by safe borrowers, and this may cause safe borrowers to exit the market.

Ghatak 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. (Group lending contracts, popularized by the micro-credit movement, require borrowers to form official groups and to bear some liability for the loans of fellow group members.) The idea is as follows. First, borrowers voluntarily sort into groups that are homogeneous by risk. Second, given homogeneous

[2]The 2006 Nobel Peace Prize was awarded to Bangladesh's Muhammad Yunus and the Grameen Bank for pioneering the microcredit approach. Quoted on Armendariz and Morduch (2005), economist Timothy Besley calls microfinance "one of the most significant innovations in development policy of the past twenty-five years". Armendariz 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 and Morduch (2005) for a discussion of impact studies. Ahlin and Jiang (2008) explore the issue of long-run impact theoretically.

matching, the lender can use joint liability contracts to screen or pool borrowers 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 due to homogeneous matching, and thus the joint liability clause is less costly for them in expectation (conditional on success). 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) – as well as the growth of credit markets among the poor.

The lynchpin in this theory is the homogeneous risk-matching of the borrowing groups, which provides the effective discounts for safe borrowers. To our knowledge, however, matching patterns of micro-credit groups have yet to be empirically tested.[4] A main contribution of the current paper is to test directly for homogeneous risk-matching among borrowing groups in Thailand.

The paper also extends the theory on matching for credit to consider *correlated* risk, asking whether groups will form so as to diversify or anti-diversify group 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, since anti-diversification tends to limit the effectiveness of joint liability as a contracting tool.

We test empirically whether groups are homogeneous in both risk-type and risk-exposure.

[4]The literature seems to recognize this as an important open question. For example, it is the first one on the microfinance mechanisms empirical research agenda Morduch (1999, p. 1586) lays out: "Is there evidence of assortative matching through group lending as postulated by Ghatak (1999)?"

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 use a new approach. For each village and variable, we calculate a variance decomposition, rank correlation, and/or chi-squared test statistic to assess homogeneity of sorting. We then put these calculations in perspective with a permutation test. That is, we first repeat the calculations for all possible groupings of the village borrowers into groups of the observed sizes; the result from the observed grouping is then mapped 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 random with respect to the variable in question, 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 far from perfect homogeneity, the matching pattern can reject random matching in the direction of homogeneity. We also find evidence for anti-diversification within groups. While random matching based on agricultural occupation cannot be rejected, groups appear anti-diversified in terms of clustering of bad income years and income shocks.

We turn next to a multivariate analysis using 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 some support for homogeneous matching along both risk-type and correlated risk dimensions.

In sum, Ghatak's theory receives support from the data: the degree of risk-type homogeneity is significantly greater than random matching would predict. Apparently, group lending is successfully embedding a non-negligible discount for safe borrowers via matching behavior; this may partly explain how group lending and microcredit have successfully awakened previously dormant credit markets. However, results on anti-diversification caution about a potentially negative aspect of voluntary group formation; they also suggest that joint liability contracts can contain their own disincentives for using microcredit groups to share risk.

It should be noted that we do not decisively 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 4.3). 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 (section 4.1), as well as the results (section 4.2) and a discussion of causality (section 4.3). Section 5 presents the multivariate estimation. Section 6 concludes. Figures and proofs are in the appendix.

2 Theory

2.1 Baseline model and results

Risk-neutral agents are each endowed with no capital and one project. Each project requires one unit of capital and has expected value $\overline{Y}$. Agents and their projects differ in risk, indexed by $p \in \mathcal{P}$, where $\mathcal{P} = [\underline{p}, \overline{p}]$ and $0 < \underline{p} < \overline{p} < 1$. The project of an agent of type p yields gross returns of Y_p ("succeeds") with probability p and yields gross returns of 0 ("fails") with probability $1 - p$. This implies that $p \cdot Y_p = \overline{Y}$, for all $p \in \mathcal{P}$. The higher p, the lower the agent's risk.

Agents' types are observable to other agents, but not to the outside lender. In this context, uncollateralized individual loan contracts can be inefficient. They bear an interest rate based on the average risk of a borrowing pool, a rate at which safe borrowers may find it unprofitable to borrow.[5] Thus, the lending market can (partially) collapse – even when all projects are efficient – excluding all but the riskiest borrowers due to a failure to price for risk.

In this context, group lending can increase efficiency by improving risk-pricing. 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

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

assuming the borrowers' returns are uncorrelated. Note that

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

That is, risk-types are complements in the group payoff function and homogeneous matching by risk is the stable outcome when there is a continuum of agents, as Ghatak has shown. The intuition is that having a more reliable partner is worth more to safe borrowers, since a

[5]For evidence of this behavior in the Thai context, see Ahlin and Townsend (2007b).

borrower is "on the hook" for his partner iff he succeeds.

In order to compare to a standard individual loan contract, where the payoff is $\overline{Y} - p_i r$ and the interest rate does not vary by risk-type, one can rewrite the borrower's payoff under a group contract (equation 1) as

$$\pi_{ij} = \overline{Y} - p_i \tilde{r},$$

where

$$\tilde{r} \equiv r + q(1 - p_j).$$

Here $\tilde{r}$ is interpretable as the *effective* interest rate under group lending.[6] It includes both the direct interest rate r and expected bailout payment $q(1 - p_j)$.

Note that a borrower's effective interest rate declines in his partner's risk-type, p_j. What homogeneous matching gives is that one's effective interest rate declines in one's *own* risk-type, p_i. Safer borrowers have safer partners, and thus can expect fewer bailout payments (conditional on success); in other words, safe borrowers face a lower *effective* interest rate under joint liability, as they would under full information. In this sense, group lending harnesses social information to vary the interest rate by risk-type and improve risk-pricing.

All this is true even under unsophisticated pooling contracts, where the lender simply offers all comers a standard joint liability contract. Whether the lender knows it or not, if matching is homogeneous, the contract embeds discounts for safe borrowers and can draw more of them into the market. In this sense, unsophisticated group lending can be responsible for reviving a lending market.

We next briefly discuss three variants on the basic model. First, assume a finite population of borrowers rather than a continuum. Though homogeneous matching will generally not be possible, it can be shown that groups will be rank-ordered by risk-type in equilibrium. That is, the two riskiest will pair together, the next two riskiest will pair together,

[6]Of course, the direct interest rates (r) need not be the same in group and individual contracts – even holding borrowing pool fixed, r can be lower in a group contract because the bailout payments q also provide the bank revenue.

and so on. Given rank-ordered matching, group lending has qualitatively similar risk-pricing advantages over individual lending. For example, the safest borrower is paired with the next safest borrower, rather than with someone identical, and so receives a similar (if slightly smaller) discount in effective interest rate.

Second, consider removing the assumption that borrowers know each other's risk types, so that random matching results instead of homogeneous matching. All borrowers would then face the same expected effective interest rate based on matching with the average risk-type in the borrowing pool.[7] With no variation in ex ante effective interest rate across risk-type, group lending would lose its risk-pricing advantage over individual lending in this context and could not draw safe borrowers back into the market.[8]

Third, continue to assume borrower ignorance of each other's risk, but assume they match on characteristics other than risk – e.g. proximity or friendship – that are themselves predictive of risk. Or, one could assume borrowers do know each others' risk but face custom-based or other constraints on matching. In either case, one might observe "relatively" homogeneous risk-matching. Interestingly, group lending would then still tend to embed an effective discount for safe borrowers, for the reasons discussed. If borrowers grasped their payoffs, and to the extent that groups formed risk-homogeneously, then group lending could still be a force for expanding the lending market.

In summary, within-group risk homogeneity could be expected in a number of settings. Group lending could in each case improve risk-pricing and facilitate more efficient lending.

2.2 Heterogeneous matching over risk

Several authors have made the point that a different form of joint liability can reverse the matching pattern. Specifically, Sadoulet (1999) and Guttman (2008)[9] consider dynamic contracts where liability for one's partner carries the threat of being denied future loans. In

[7]See Ahlin and Townsend (2002, section 5.4.7) for more formal analysis.

[8]In a slightly different context, Armendariz and Gollier (2000) show how joint liability can raise efficiency even with random matching.

[9]See also Chiappori and Reny (2006) for heterogeneous matching in a different context.

the spirit of these models, we modify the previous section's model by ignoring saving and assuming that there are potentially two periods of loans. If the lender receives nothing from the group in the first period, it denies both borrowers a second loan with probability d. Here, d can be thought of as **dynamic liability**, while q represents **direct liability**.

A borrower of type p_i matching with a borrower of type p_j now has expected payoff

$$\pi_{ij} = [\overline{Y} - rp_i - qp_i(1-p_j)] \cdot [2 - d(1-p_i)(1-p_j)]. \tag{3}$$

This is the same as in payoff equation 1 except for the additional bracketed term which reflects the fact that there will be two loans unless both borrowers fail in the first period (probability $(1-p_i)(1-p_j)$) and the lender responds by denying a second loan (probability d). The cross-partial of the group payoff in this case is

$$\frac{\partial^2(\pi_{ij}+\pi_{ji})}{\partial p_i \partial p_j} = -2d[\overline{Y} - (p_i+p_j-1)r] + 4q + 2dq[(2p_i-1)(1-p_j) + (2p_j-1)(1-p_i)]. \tag{4}$$

For any $q > 0$, if d is small enough the cross-partial is positive (close to $4q$). Homogeneous matching would result.

For any $d > 0$ and q small enough, however, the cross-partial is close to

$$-2d[\overline{Y} - (p_i+p_j-1)r],$$

which must be negative for any pair of agents that finds borrowing worthwhile. Thus, with direct liability (q) negligible relative to dynamic liability (d), risk-types are substitutes. The intuition here is that having a more reliable partner is worth more to risky borrowers, for several reasons: they more often need their partner to be able to bail them out in order to get a second loan, and they value a second loan more due to cross-subsidization from safe to risky borrowers.

With risk-types as substitutes, the matching would be heterogeneous.[10] That is, the safest borrower would match with the riskiest, the next safest with the next riskiest, and so on onion-style.[11]

From a contracting perspective, it is not clear why dynamic liability would be used more heavily than direct liability. Indeed, Sadoulet (1999) and Guttman (2008) take the contract form as given and focus on matching. Regardless, the point here is simply that a dynamic operationalization of joint liability in this context could lead to very different, heterogeneous matching patterns.

A second well-known type of contract – a relative performance contract – could also lead to heterogeneous matching when applied in this context. The basic joint liability contract discussed in the previous section (payoff equation 1) can be interpreted as a relative performance contract if $q < 0$, since then a borrower's payment is *lower* when his partner fails. With this contract, the cross-partial (equation 2) would be negative and heterogeneous matching would result.

Again, the point is simply that some common contracting forms lead to heterogeneous matching in this environment. If lenders for some reason have a misspecified model in mind or operate under other constraints,[12] these contract forms may be observed in practice even if not strictly optimal.

2.3 Matching over degree and source of risk

This section points out a potential pitfall of relying on voluntary matching. We add to the baseline model (with direct liability only) the possibility for correlated risk. Given the agricultural setting of many micro-lenders, including the one in our data, this is a potentially important modification. However, it is little analyzed in the group lending literature, and

[10]This is true of the first period. The final period would involve risk-type complementarity and the desire to re-match homogeneously for continuing borrowers. The cited papers use an infinite horizon; we use a two-period framework for simplicity.

[11]Heterogeneous matching is more complicated with n-person groups when $n > 2$; see Ahlin (2009).

[12]The lender in our data is a government development bank focused on agriculture and arguably operating under significant political constraints.

to our knowledge not at all in the context of group formation.

Given two borrowers i and j with unconditional probabilities of success p_i and p_j, respectively, the joint output distribution can be written uniquely as:

	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 $(1-p_i)$	$(1-p_i)p_j - \epsilon_{ij}$	$(1-p_i)(1-p_j) + \epsilon_{ij}$

(5)

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 correlations which will ensure that $\epsilon_{ij} = \epsilon > 0$ for any two borrowers exposed to the same shock, and $\epsilon_{ij} = 0$ for all other pairings.

Assume there are two i.i.d. aggregate shocks, A and B. Each equals 1 or -1 with equal probability. Every agent is assumed to be exposed to risk from either shock A or shock B, or neither. 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. The project of an A-risk agent is independent of shock B: $p_i|B = p_i$. The success of a "B-risk" agent of type p_i depends on the realization of shock B but not shock A in the exactly analogous way: $p_i|B = p_i + \gamma B$, $p_i|A = p_i$. The remaining "N-risk" agents are exposed to neither aggregate shock: $p_i|A = p_i|B = p_i$.

With these assumptions, the ϵ_{ij} of equation 5 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[13]

$$\epsilon_{ij} = \epsilon \equiv \gamma^2. \tag{6}$$

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.[14]

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 borrower exposure-type. For two borrowers i and j, let $\kappa_{i,j}$ equal 1 if $s_i = s_j = A$ or $s_i = s_j = B$, and 0 otherwise. Then the payoff of borrower i when matched with borrower j is now

$$\pi_{ij} = \overline{Y} - rp_i - q[p_i(1-p_j) - \epsilon\kappa_{i,j}] = \overline{Y} - rp_i - qp_i(1-p_j) + q\epsilon\kappa_{i,j}\ . \tag{7}$$

The last term ($q\epsilon$) represents a payoff boost from matching with a partner exposed to the same risk. This is because positive correlation of returns in the group lowers chances of having to bail out one's partner.

In this context, the following can be shown:

Proposition 1. *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, groups match homogeneously in risk-type and exposure-type; they contain either all A-risk, all B-risk, or all N-risk borrowers. The intuition for exposure-type homogeneity is simple: borrowers choose to anti-diversify their groups so as to lower their chances of facing liability for their partners.

[13] With probability 1/2, the shock to which both are exposed 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 $p_i p_j + \gamma^2$.

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

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. 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, especially the dimension that has greater payoff salience.

Homogeneous borrower sorting along the correlated risk dimension appears to work against efficient lending. In a finite population, it may divert borrowers from rank-ordering by risk-type, which is the basis for group lending's efficiency gains in this context. More importantly, correlated risk lowers the *effective* rate of joint liability. In the extreme case of perfect correlation, for example, the effective rate of joint liability is 0 regardless of how the bank sets q, since when one borrower fails, they both 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.[15] Thus, some dimensions of voluntary matching may not work in favor of efficiency.

Though joint liability groups are sometimes thought of as risk-sharing groups, the result of Proposition 1 casts them in a different light. In particular, joint liability contracts can reward anti-diversification, and thus limit the usefulness of the group for risk-sharing. However, it need not imply that group lending is bad for risk-sharing: households may share risk with other households regardless of whether they are in the same borrowing group.

On the other hand, if joint liability is primarily the dynamic liability of section 2.2, the contract would reward formation of diversified groups; diversification would raise chances of partner bailouts that can extend the borrowing relationship. In this sense, risk-type and exposure-type heterogeneity would be consistent with predominantly dynamic liability, while homogeneity in both dimensions is consistent with predominantly direct liability.

[15] The lender could in principle use a higher q to neutralize higher correlated risk. However, this is not possible since, as Gangopadhyay et al. (2005) argue, q is bounded above by r and is optimally set at its upper bound when risk is uncorrelated.

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, groups drawn from the same pool of borrowers will be compared.

3.1 Data description and environment

The data come from the Townsend Thai survey data. In May 1997, a cross section of l92 villages was surveyed, covering four provinces from two contrasting regions of Thailand, both with large agricultural sectors. In each village as many borrowing groups of the Bank for Agriculture and Agricultural Cooperatives (BAAC) as 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.[16]

Hence, we turn to a resurvey, conducted in April and May 2000. The resurvey data were collected from a random subset of the same villages, stratified at the sub-district (tambon) level. 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. Relative to the baseline survey, the resurvey data have two decisive advantages. First, individual group members respond to questions *on their own behalf*, up to five per group and on average 4.5. Second, several resurvey questions were designed 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

[16]The main concern is that when data on all borrowers in a group are provided by one person, the responses can exaggerate within-group homogeneity. Unreported results from this dataset validate this concern.

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[17] 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 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 dominates in practice, if either.

Groups typically have between five and fifteen members; about 15% are larger. 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.[18] 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.

[17] 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.

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

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.[19] 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, Hi and Lo, and that Ex represents the mean, 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.[20] Based on the same projected income distribution, this works out to be

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

This is simply the geometric average of the percentage deviations from the mean for good and bad years.

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 a *vector* with ten entries, each giving the fraction of total household revenue accounted for by one agricultural category.[21]

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

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

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

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 binary measure of occupation/risk-exposure because some degree of within-household occupational heterogeneity is common in our data.[22] 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.

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.

Third, we calculate a direct measure of a household income **shock**. Under certain assumptions, it captures the percent deviation of this year's income from its expected value. The household's current income, Inc, comes from a detailed compilation of realized business and farm revenues and expenses, along with wage and other income sources, for the just-completed year. 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 each household's income is 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 each household's income exhibits i.i.d. fluctuations around a growth path, where the growth rate is common across households, the measure captures the income shock up to a constant. However, if household incomes are growing at different rates, then $Shock$ captures not only the income shock but differential growth rates of income.

[22] Occupational similarity between two borrowers is then the dot product of their respective vectors. The dot product would give the $\kappa_{i,j}$ of section 2.3 if we had only two categories and all vector entries had to be binary. 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.

4 Univariate Methodology and Results

In this section, we assess homogeneity/heterogeneity of sorting in one dimension at a time. Section 5 extends the analysis to consider multi-dimensional sorting.

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 in a nonparametric statistical test across all villages of whether groupings tend to be homogeneous or heterogeneous. Section 4.1.3 proposes alternative, model-driven metrics of sorting that can be used as the basis for these nonparametric tests for homogeneous/heterogeneous sorting.

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, e.g. risk-type. One way to measure within-group homogeneity is to calculate a *variance decomposition* of $X = (x_1, ..., x_{l+m})$ into between-group and within-group components. The between-group variance is maximized in a rank-ordered grouping, so a larger between-group component can be taken as evidence of homogeneous sorting. To illustrate, consider a village with 2 groups of size 4, with risk-types[23] $X = (1, 2, 4, 5, 6, 7, 8, 9)$. Compare the following borrower grouping: $L = (2, 5, 6, 8)$ and $M = (1, 4, 7, 9)$, with an alternative grouping: $L' = (1, 2, 5, 6)$ and $M' = (4, 7, 8, 9)$. Of the overall variance, 0% in the first grouping and 44% in the second grouping is attributed to between-group differences. The higher value reflects the more homogeneous sorting of the second grouping, and the lower value the more mixed first grouping.

[23]For brevity, all risk-types are multiplied by 10; the true data are $(0.1, 0.2, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)$.

An alternative sorting metric that uses only the data's ordinality is a rank correlation. In particular, one can calculate Kendall's tau_b[24] between the data X and a group index variable, Y, where for example $y_1 = ... = y_l = 1$ and $y_{l+1} = ... = y_{l+m} = 2$. Using the groupings 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%.[25] Again, the higher value reflects more homogeneous sorting, and the lower value more mixing; Kendall's tau_b is maximized under rank-ordering.

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.[26] 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 (shocks), compare the following grouping: $L = (A, A, B, B)$ and $M = (A, A, B, B)$, with an alternative: $L' = (A, A, A, B)$ and $M' = (A, B, B, B)$. The chi-squared test statistic for the first grouping is 0 and for the second grouping is 2.[27] Again, the higher value reflects more homogeneous sorting; the chi-squared statistic is maximized under group homogeneity.

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 the relevant matching universe for group formation is the village – a reasonable assumption since villages are relatively small

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

[25]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).

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

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

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. 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 with 2 groups of size 4, with risk-types $X = (1, 2, 4, 5, 6, 7, 8, 9)$. There are $\binom{8}{4}/2 = 35$ groupings of these eight borrowers into two groups of size four. Compared to the grouping $L = (2, 5, 6, 8)$ and $M = (1, 4, 7, 9)$, 32 groupings register higher between-group variance while 3 (including the grouping itself) register exactly the same, i.e. zero between-group variance. Thus this grouping 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 ties and the relatively small number of groupings. Compared to the grouping $L' = (1, 2, 5, 6)$ and $M' = (4, 7, 8, 9)$, 31 groupings have lower, 2 have the same, and 2 have higher between-group variance. This grouping's sorting percentile range is thus $[88.6, 94.3]$. Similarly, applying this permutation test to the Kendall's tau$_b$ rank correlation metric gives a slightly wider sorting percentile range to the first grouping, $[0, 11.4]$, and the same sorting percentile range to the second grouping, $[88.6, 94.3]$.

The same approach can be used with the chi-squared test statistic.[28] There are 17 combinations with a larger chi-squared test statistic and 18 combinations tied with grouping $L = (A, A, B, B)$ and $M = (A, A, B, B)$. This grouping's percentile range is then $[0, 51.4]$. Compared to grouping $L' = (A, A, A, B)$ and $M' = (A, B, B, B)$, 18 combinations have less, 1 combination has greater, and 16 combinations have the same chi-squared test statistic.

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

This grouping'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 used, while a lower sorting percentile reflects more heterogeneous sorting. One can then interpret villages with percentiles above the 95th as exhibiting homogeneous 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 corresponding to a null hypothesis. An advantage of this approach is that it is non-parametric and requires no distributional assumptions.

Our null hypothesis is that matching is random with respect to the given variable. The rationale is the same as the one underpinning the t-statistic in a linear regression, where the null hypothesis is also that the variable has no explanatory power.

The distribution of sorting percentiles that reflects random matching is the uniform on $[0, 1]$. The idea is as follows. Consider 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 2. *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 observed 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.[29]

[29]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.

4.1.3 The test with structural sorting measures

The sorting metrics discussed so far – 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.[30] So, we turn next to alternative metrics derived directly from the theory. The theory predicts that any grouping we observe must maximize the sum of group payoffs; otherwise, the groups could re-organize with everyone being made better off. 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 seen, leads to homogeneous matching under direct liability).

Consider first the baseline model with uncorrelated risk and observed groups $L = \{i, j\}$ and $M = \{i', j'\}$ in a village. Let group payoff functions be Π_L and Π_M, where $\Pi_L = \pi_{ij} + \pi_{ji}$ and $\Pi_M = \pi_{i'j'} + \pi_{j'i'}$. Then we would like to use $\Pi_L + \Pi_M$ as the measure of sorting, since this is what must be maximized by the observed grouping in the theory. Using equation 1,

$$\Pi_L + \Pi_M = 4\overline{Y} - (r + q)(p_i + p_j + p_{i'} + p_{j'}) + q(p_i p_j + p_j p_i + p_{i'} p_{j'} + p_{j'} p_{i'}).$$

Note that only the interaction terms (the last parenthetical) may differ across groupings of the 4 borrowers. 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 these terms.[31] Letting p_{-k} be

[30] What is clear is that they are maximized under a homogeneous/rank-ordered grouping. Less clear is how they rank other groupings on the homogeneity/heterogeneity scale. This matters because when groups have more than 2 members and types are substitutes, only a relatively small fraction of groupings can be ruled out without knowledge of the exact payoff function and type distribution (see Ahlin, 2009). However, while substitutability by itself may rule out few groupings, knowledge of the production function and the types generally leaves only one optimal grouping.

[31] In other words, results using the entire group payoff function are the same for any $\overline{Y}$, r, and $q > 0$.

the risk-type of borrower k's partner, they can be written

$$\sum_{k \in L} p_k p_{-k} + \sum_{k \in M} p_k p_{-k}. \tag{8}$$

Taking the theory to data is complicated, however, by the fact that the borrowing groups in the data are not pairs, but typically involve 5-15 members. Further, we do not typically have the entire group's data, primarily because a maximum of five group members are sampled. Our strategy will be to proxy for p_{-k} in the above expression (8) using the average risk-type of the other sampled group members.

Specifically, let group G be a set of grouped borrowers, $\mathcal{S}^G$ be the sampled subset of group G, and $\overline{p}_{-k}^{\mathcal{S}^G}$ be the average risk-type in the sampled subset of group G excluding borrower k. The following is our sample estimate of the above payoff expression 8:

$$\sum_{k \in \mathcal{S}^L} p_k \overline{p}_{-k}^{\mathcal{S}^L} + \sum_{k \in \mathcal{S}^M} p_k \overline{p}_{-k}^{\mathcal{S}^M}. \tag{9}$$

This estimate is simply the sum, over all sampled village borrowers, of the borrower's risk-type multiplied by the average risk-type of other, same-group sampled borrowers.[32] To illustrate, grouping $L = (2, 5, 6, 8)$ and $M = (1, 4, 7, 9)$ has sum of group payoffs of 202,[33] compared to 235 for more homogeneously matched grouping 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.

Consider also the contract under correlated risk. Letting $\kappa_{k,-k}$ be the dummy equaling 1

[32]This form of the payoff function can be justified by the n-person group contract suggested in Ghatak (1999), in which each borrower who succeeds owes q per fellow unsuccessful borrower.

[33]This comes from $2 * 19/3 + 5 * 16/3 + ... + 7 * 14/3 + 9 * 12/3$.

if borrower k faces the same risk-exposure as his partner and using payoff function 7 gives

$$\begin{aligned}\Pi_L + \Pi_M =& 4\overline{Y} - (r+q)(p_i + p_j + p_{i'} + p_{j'}) \\ &+ q\left(\sum_{k\in L} p_k p_{-k} + \sum_{k\in M} p_k p_{-k}\right) + q\epsilon\left(\sum_{k\in L}\kappa_{k,-k} + \sum_{k\in M}\kappa_{k,-k}\right).\end{aligned} \tag{10}$$

There are two types of interaction terms in the groups' payoff function, involving unconditional risk-type and risk exposure-type (the first and second parentheticals on the second line, respectively). 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. Following the above techniques and defining $\overline{\kappa}^{\mathcal{S}^G}_{k,-k}$ as the average correlatedness dummy of borrower k in group G with other sampled group G members, our estimator for sum of group payoffs due to correlated risk is

$$\sum_{k\in\mathcal{S}^L}\overline{\kappa}^{\mathcal{S}^L}_{k,-k} + \sum_{k\in\mathcal{S}^M}\overline{\kappa}^{\mathcal{S}^M}_{k,-k}.$$

This estimate is simply the sum, over all sampled village borrowers, of the fraction of other, same-group sampled borrowers exposed to the same risk. Again, compare grouping $L = (A,A,B,B)$ and $M = (A,A,B,B)$ with grouping $L' = (A,A,A,B)$ and $M' = (A,B,B,B)$. The correlation-related payoffs sum to 2.67 in the first grouping and to 4 in the second, more anti-diversified (homogeneous) grouping.[34]

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 $\kappa_{i,j}$, the correlatedness dummy. In the case of worst_year (see section 3 for descriptions of these measures), we proxy $\kappa_{i,j}$ simply by $1\{worst\_year^G_i = worst\_year^G_j\}$; that is, if the two

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

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 $\kappa_{i,j}$ by $e^{-|shock_i^G - shock_j^G|}$; 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, $\kappa_{i,j}$ 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 high income, **prob**, is the focus of our empirical tests for homogeneous risk-matching. The sample CDFs of village sorting percentile ranges for *prob* based on Kendall's tau_b and the structural sorting metric, respectively, are graphed in Figure 1.[35] Based on 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 dashed 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.[36] These results point to risk-matching that, while not rank-ordered, is statistically distinguishable from random matching in the direction of homogeneity.

The results using the structural sorting metric, which uses the model's specific payoff function (see section 4.1.3), are quite similar to the atheoretic results. The mean (median) sorting percentile is 56% (61%) and heterogeneous matching is rejected at the 5% level.

A second measure of risk is the **coefficient of variation** of projected income. This

[35]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.

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

measure is a proxy and cannot be substituted directly into the structural sorting metric. Hence, 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: 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 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. From the standpoint of the theories of section 2, this suggests that direct liability is a non-negligible aspect of the group contracts, and that safe borrowers are receiving somewhat lower implicit borrowing rates because they tend to have safer partners.

Sorting by exposure-type. We next examine diversification within groups. Consider the **worst_year** measure. Since this is a categorical variable, Figure 3 reports results using the chi-squared test statistic and the structural sorting metric. 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.[37] 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 not too different from random based on occupation, where occupation is measured by shares of total 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, unobserved traits.

4.3 Discussion of Univariate Results

The univariate tests suggest that group composition is homogeneous along both risk-type and exposure-type dimensions – far from perfectly, but moreso than under random matching. Of course, the evidence is not proof of causality running from risk amount or exposure type to sorting behavior. For example, it may 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, if the goal is to assess whether the Ghatak (1999) model is an empirically

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

plausible (partial) explanation of group lending's popularity and ability 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 this discount, which prices risk more accurately, that allows group lending 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 model. Conversely, 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.

Of course, establishing causal links between risk and group formation would be ideal for testing other aspects of the Ghatak model and extensions. It would also make the main argument more robust if the risk-homogeneity is achieved purposefully rather than as an unintended consequence. Still, as argued above, we believe the results presented are first-order informative about the ability and limitations of the theory to explain the rise of group lending and micro-credit.

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. Hence, 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).[38] This estimator can be used with an atheoretic or structural approach; here we take a structural approach and derive the specification from the exact payoff function of the model.[39]

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 derives moment inequality conditions from 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 $\tilde{L}$ and $\tilde{M}$ 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 $\tilde{L}$ and $\tilde{M}$ 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

$$\Pi_L(\phi) + \Pi_M(\phi) \geq \Pi_{\tilde{L}}(\phi) + \Pi_{\tilde{M}}(\phi). \qquad (11)$$

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

[39]A reduced-form estimation that included more controls than we use could also be interesting. However, this is not as attractive in part since our dataset lacks data on social networks.

The matching maximum score estimator chooses parameters ϕ that maximize the score, i.e. the number of inequalities of the form 11 that are true, where each inequality corresponds to a different unobserved grouping $\tilde{L}, \tilde{M}$.

Our set of unobserved groupings, and thus inequalities, comes from all k-for-k borrower swaps across two groups in the same village.[40] 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.[41]

Consider the model's expression for $\Pi_L + \Pi_M$ from section 2.3, reproduced from equation 10 here:

$$\begin{aligned}\Pi_L + \Pi_M =& 4\overline{Y} - (r+q)(p_i + p_j + p_{i'} + p_{j'}) \\ &+ q\left(\sum_{k\in L} p_k p_{-k} + \sum_{k\in M} p_k p_{-k}\right) + q\epsilon\left(\sum_{k\in L} \kappa_{k,-k} + \sum_{k\in M} \kappa_{k,-k}\right).\end{aligned}$$

Note that all terms in the group payoff function that do not involve interactions between borrower characteristics drop out of inequality 11, since they appear identically on both sides;[42] hence, we can ignore the non-interaction terms.

We proceed as in section 4.1.3. Since groups contain more than 2 members and since our data contain a subset of each group (up to 5 members), we use a sample analog expression for the payoff function. Again, let G be defined as a set of grouped borrowers, $\mathcal{S}^G$ as the sampled subset of group G, k as a sampled group-G borrower, $\overline{p}_{-k}^{\mathcal{S}^G}$ as the average risk-type in the sampled subset of group G excluding borrower k, and $\overline{\kappa}_{k,-k}^{\mathcal{S}^G}$ as the average correlatedness dummy of borrower k with other sampled group-G borrowers. Then, the sample analog to

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

[41] 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.

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

the (relevant part of the) payoff function is

$$\Pi_L + \Pi_M = q\left(\sum_{k\in\mathcal{S}^L} p_k\overline{p}_{-k}^{\mathcal{S}^L} + \sum_{k\in\mathcal{S}^M} p_k\overline{p}_{-k}^{\mathcal{S}^M}\right) + q\epsilon\left(\sum_{k\in\mathcal{S}^L} \overline{\kappa}_{k,-k}^{\mathcal{S}^L} + \sum_{k\in\mathcal{S}^M} \overline{\kappa}_{k,-k}^{\mathcal{S}^M}\right) . \qquad (12)$$

It is this expression that we use for group payoffs in the inequality 11.[43]

Given data on borrower probabilities of success (p_i^G's) and correlatedness ($\kappa_{i,j}$'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 $\tilde{\beta}/q$. This approach, however, requires data that can capture the existence of correlation ($\kappa_{i,j}$) as distinct from the intensiveness (ϵ) of correlation. That is, to identify ϵ, $\kappa_{i,j}$ 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 $\kappa_{i,j}$, we focus on the overall correlation between borrowers i and j, call it $C_{i,j} \equiv \epsilon\kappa_{i,j}$. $C_{i,j}$ is proxied in different ways, depending on the variable used (see section 4.1.3). When worst_year is used, $C_{i,j} = \phi_{wst}\, 1\{worst\_year_i = worst\_year_j\}$. When income shock is used, $C_{i,j} = \phi_{shk}\, e^{-|shock_i - shock_j|}$. When occupation is used, $C_{i,j} = \phi_{occ}(\vec{occ}_i \cdot \vec{occ}_j)$ (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 $C_{i,j}$ – for concreteness, proxied here using worst_year – and notation similar

[43]That is, interaction terms involving sampled borrowers are used to estimate the group payoff function. Similarly, the counterfactual groups are formed via k-for-k borrower swaps across sampled subsets of the groups. In general, sorting optimality conditions (here, inequality 11) need not hold for random *subsets* of groups. However, if types are complements, they do, since if two groups are rank-ordered, so are any two subsets of the two groups.

to the above into the sampled groups' payoff function 12 gives

$$\Pi_L + \Pi_M = \beta_1 \left(\sum_{k \in \mathcal{S}^L} p_k \overline{p}_{-k}^{\mathcal{S}^L} + \sum_{k \in \mathcal{S}^M} p_k \overline{p}_{-k}^{\mathcal{S}^M} \right) + \beta_2 \left(\sum_{k \in \mathcal{S}^L} \overline{1\{worst\_year_k = worst\_year_{-k}\}} + \sum_{k \in \mathcal{S}^M} \overline{1\{worst\_year_k = worst\_year_{-k}\}} \right), \tag{13}$$

where $\beta_1 = q$, $\beta_2 = q\phi_{wst}$, and the terms in the second-line sums represent the fraction of other, same-group sampled borrowers naming the same worst_year as borrower k. 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 suggest substitutability of types and that more heterogeneous patterns of matching are being observed. Regarding $\beta_2 (= q\phi_{wst})$, 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 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

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

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.

where[44] 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.[45]

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

[44]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.

[45]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 13 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.

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. Inference is carried out by subsampling.[46]

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, 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 group lending's improved risk-pricing.

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

[46] Fox (2008) 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, 2.2).

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, controlling for correlated risk, matching appears homogeneous by risk-type. There is also suggestive 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 joint liability lending and unobserved risk type, theory suggests that borrowers will sort homogeneously by risk; this embeds an effective discount for safe borrowers and improves efficiency. However, theory also suggests 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 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, and find supporting evidence. Direct comparisons to 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 cast doubt 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.

The paper leaves some open questions for future work to address. First, the risk and correlation measures used here could be improved upon. Future work with income histories and/or more detailed elicitations of future income distributions could perhaps push the analysis further, including in a more quantitative direction. Second, it would be ideal for sorting tests to use measures of risk that pre-date group formation, to distinguish sorting patterns from within-group conformity or imitation that occur after group formation. Third, richer datasets that include data on social networks, physical distances, etc., could potentially be used to identify whether risk-homogeneity and anti-diversification are purposeful or are by-products of other matching considerations. They could also help quantify and pinpoint matching frictions in these environments.

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. Matching under transferable utility when group size is greater than two. Working Paper, December 2009.

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

[4] Christian Ahlin and Robert M. Townsend. Using repayment data to test across theories of joint liability lending. Vanderbilt University Department of Economics Working Paper 02-W27, December 2002.

[5] 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.

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

[7] Beatriz Armendariz and Christian Gollier. Peer group formation in an adverse selection model. *Economic Journal*, 110(465):632–643, July 2000.

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

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

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

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

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

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

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

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

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

[17] 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.

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

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

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

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

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

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

[24] 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. Consider a set of equilibrium groups. There are six sets into which all groups can be partitioned: AA, BB, NN, AB, AN, BN, where the set names denote the risks faced by the two borrowers. For example, set AA contains all groups with two A-risks and set BN contains all groups with one B-risk and one N-risk.

The cross-partial of group payoff functions with respect to p_i and p_j is still given by equation 2. This implies that the baseline result of homogeneous matching in almost every group holds in any set of groups within which $\kappa_{i,j}$ is fixed for all possible pairings of borrowers within the set – in particular, AA, BB, and NN.

We show next that the sets AB, AN, and BN have zero measure in equilibrium. Consider AB, for example. Risk-type complementarity implies rank-ordering within risk exposure type: the safest A-risk matches with the safest B-risk, and so on. That is, if (i, j) and (i', j') are groups and borrowers i, i' (j, j') are A-risk (B-risk), then one of the following pairs of statements must hold: $p_i \geq p_{i'}$ and $p_j \geq p_{j'}$, or $p_{i'} \geq p_i$ and $p_{j'} \geq p_j$. Otherwise, the grouping (i, j') and (i', j) would raise surplus by increasing payoffs from risk-type complementarity without altering the nature of the exposure-type matching.

Given this fact and if set AB has positive measure, then for any $\delta > 0$, there must exist two groups (i, j) and (i', j') with $|p_i - p_{i'}| < \delta$ and $|p_j - p_{j'}| < \delta$. Fix $\delta = \sqrt{\epsilon/4}$ and two such groups. We will show that with risk-types so close, the gains from anti-diversification (matching A with A, B with B) outweigh any losses from decreased risk-type similarity.

Without loss of generality, let (i, j) be the safer group (i.e. $p_i \geq p_{i'}$ and $p_j \geq p_{j'}$) and borrower j be the safest in that group ($p_i \leq p_j$). Using equation 7, the sum of both groups' payoffs can be written

$$4\overline{Y} - (r + q)(p_i + p_j + p_{i'} + p_{j'}) + 2q(p_i p_j + p_{i'} p_{j'}) \ ,$$

since no borrowers are exposed to the same shocks. An (i, i') and (j, j') grouping would instead give rise to

$$4\overline{Y} - (r + q)(p_i + p_j + p_{i'} + p_{j'}) + 2q(p_i p_{i'} + p_j p_{j'}) + 4q\epsilon \ ;$$

the last term represents the savings from matching with borrowers exposed to the same shock. This alternative grouping gives higher total surplus and thus represents a contradiction if

$$4q\epsilon - 2q(p_i - p_{j'})(p_j - p_{i'}) > 0.$$

Note that $p_j \geq p_i \geq p_{i'}$, the first inequality by construction and the second by rank-ordering; so, the last parenthetical is non-negative. If $p_i \leq p_{j'}$, the first parenthetical is non-positive and the inequality holds.

Consider instead $p_{j'} < p_i$. Since $p_j \geq p_i$ and since $p_{j'} \geq p_j - \delta$, by hypothesis, then $p_{j'} \geq p_i - \delta$; thus, $p_{j'} \in [p_i - \delta, p_i)$, so the first parenthetical is bounded above by δ. Regarding the second parenthetical, note that $p_{i'} \in [p_i - \delta, p_i]$, by hypothesis and rank-ordering. Also, it must be that $p_j \leq p_i + \delta$; otherwise, $p_{j'} \geq p_j - \delta$ could not be true (given $p_{j'} < p_i$). Thus the second parenthetical is bounded above by 2δ. Using these upper bounds minimizes the left-hand side of the above inequality, which becomes

$$4q\epsilon - 2q\delta(2\delta) = 4q\epsilon - 4q\delta^2 = 4q\epsilon - q\epsilon = 3q\epsilon > 0.$$

Thus AB cannot have positive measure in equilibrium.

Similarly, AN and BN cannot. The only difference in the argument is that the borrower swap outlined above would add $2q\epsilon$ rather than $4q\epsilon$ to sum of the groups' payoffs – this would still raise total surplus. ■

Proof of Proposition 2. 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. ■

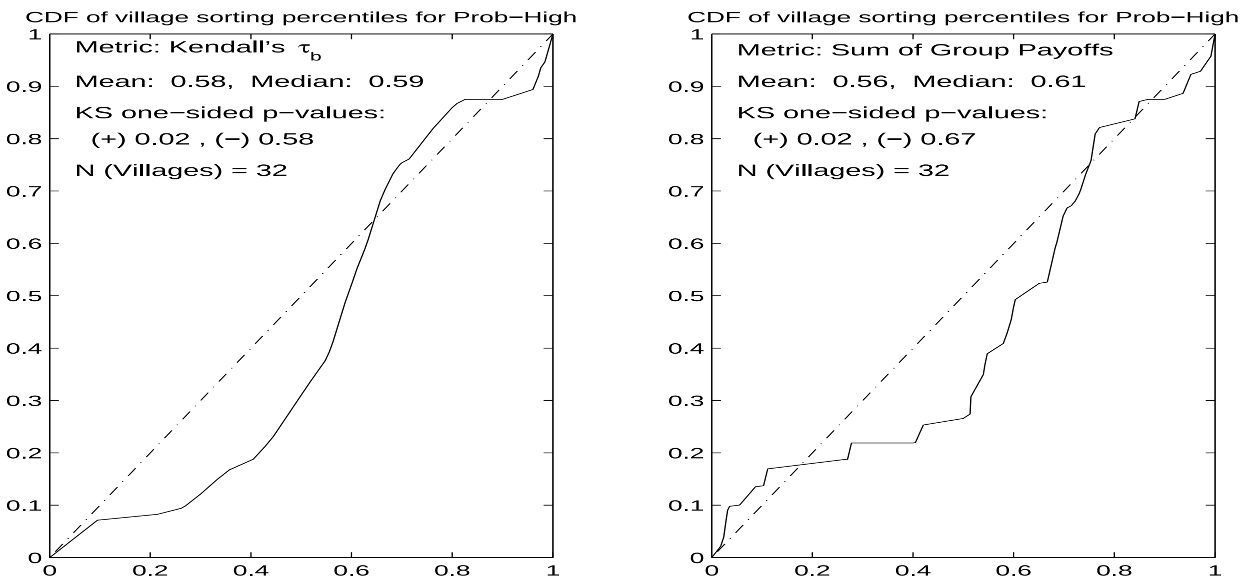

Figure 1: Solid: Sample CDF of villages' sorting percentiles based on rank correlation Kendall's tau$_b$ (left panel) and the sum of group payoff functions (right panel) for *prob*, the probability of realizing high income. Dashed: Uniform CDF.

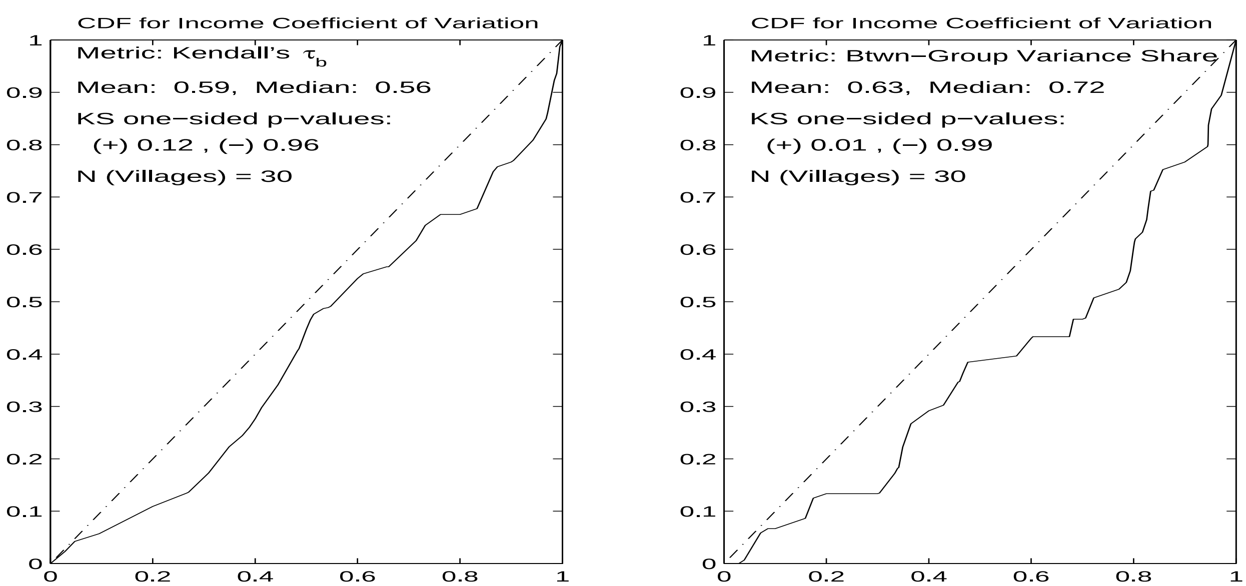

Figure 2: Solid: Sample CDF of villages' sorting percentiles based on rank correlation Kendall's tau$_b$ (left panel) and variance decomposition (right panel) of the coefficient of variation for income (standard deviation / mean). Dashed: Uniform CDF.

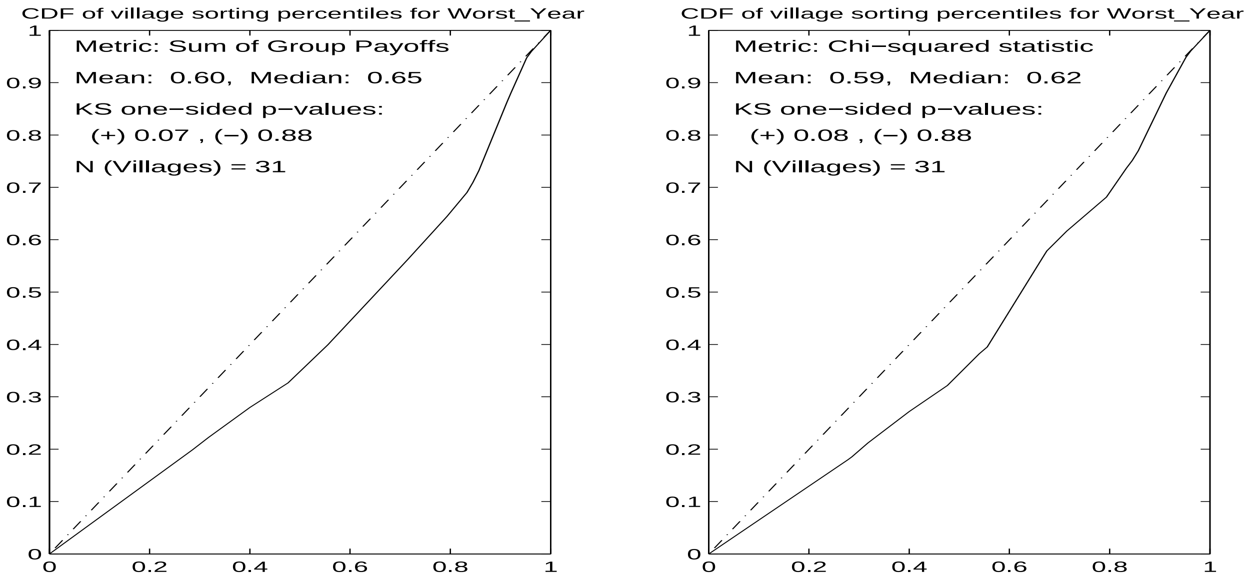

Figure 3: Solid: Sample CDF of villages' sorting percentiles based on the sum of group payoff functions (left panel) and the chi-squared statistic (right panel) of the worst_year for income. Dashed: Uniform CDF.

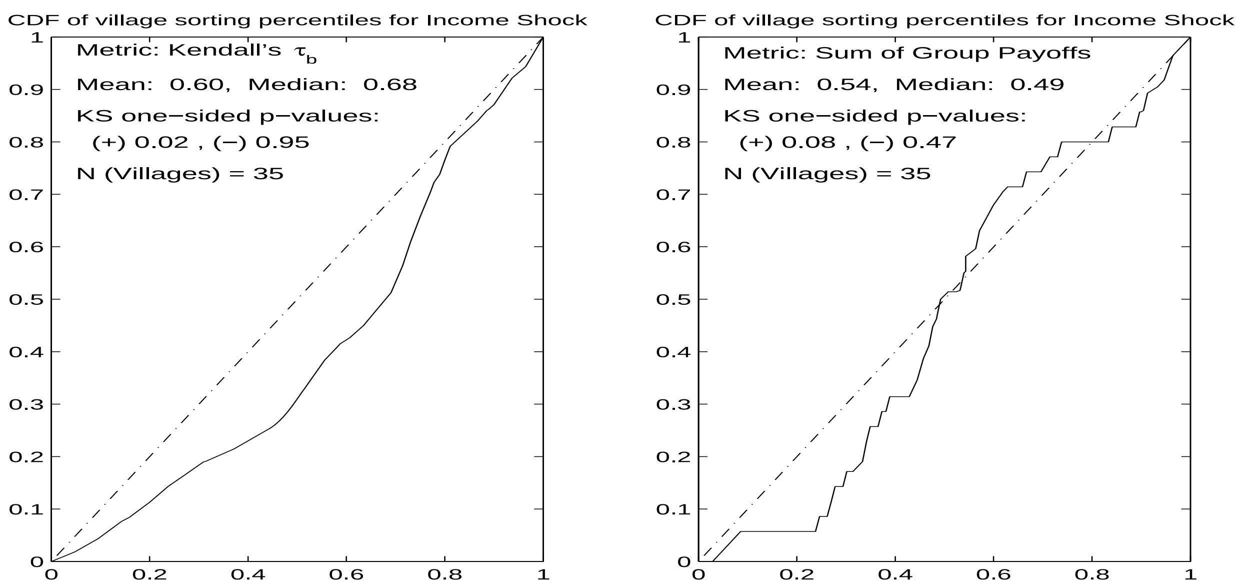

Figure 4: Solid: Sample CDF of villages' sorting percentiles based on rank correlation Kendall's tau_b (left panel) and the sum of group payoff functions (right panel) for income shocks. Dashed: Uniform CDF.

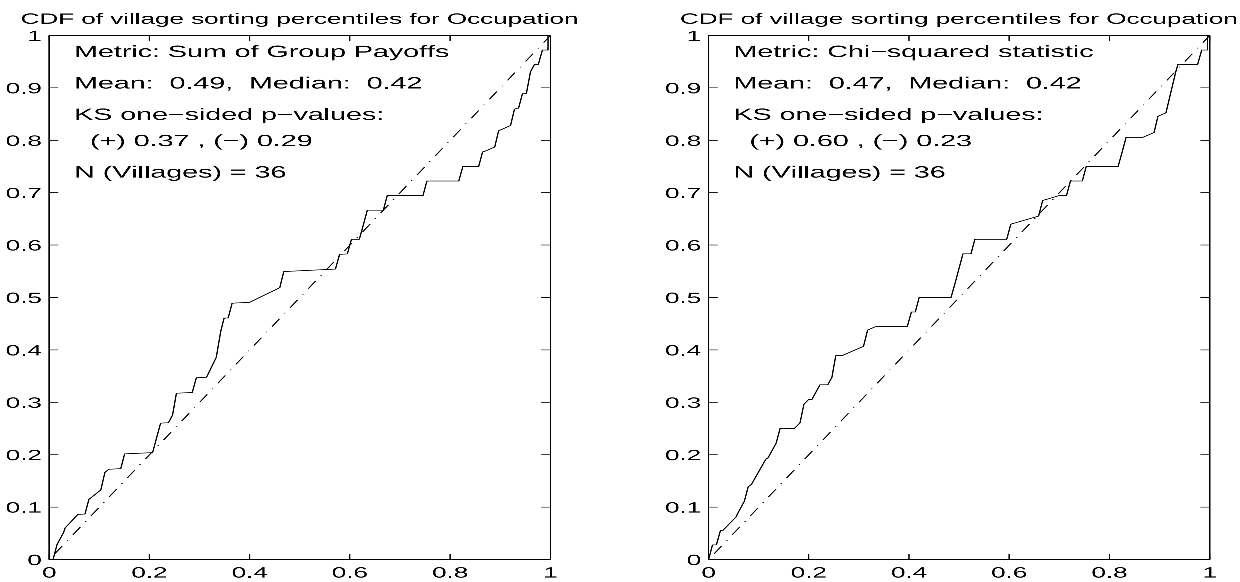

Figure 5: Solid: Sample CDF of villages' sorting percentiles based on the sum of group payoff functions (left panel) and the chi-squared statistic (right panel) for occupation. Dashed: Uniform CDF.