Selection into and across Credit Contracts: Theory and Field Research

Christian Ahlin,[†] Vanderbilt University
Robert M. Townsend, University of Chicago
and Federal Reserve Bank of Chicago

revised July 2004

JEL No. D82, O16

Keywords: Empirical Contract Theory; Adverse Selection;
Mechanism Design; Joint Liability; Micro-credit

Abstract

Lenders may choose to encourage borrower side contracting using group, or co-signed, loans or discourage it using individual loans, so as to make relative performance comparisons. In this context wealth of the agents relative to outsiders, and wealth inequality among potential joint liability partners, are important factors determining the choice among loan contracts. In a related model of whether to borrow, higher covariance of returns mitigates an adverse selection effect. We test these models using relatively rich data gathered in field research in Thailand. The prevalence of joint liability contracts relative to individual contracts exhibits a U-shaped relationship with the wealth of the borrowing household and increases with the wealth dispersion. The likelihood of joint-liability borrowing increases the lower is the probability of project success, a direct affirmation of adverse selection. Higher correlation across projects makes joint liability borrowing more likely relative to all other alternatives. Strikingly, most of the results disappear if we do not condition the sample according to the dictates of the models, with selection into and across credit contracts.

[†]Corresponding author. Department of Economics, Vanderbilt University, VU Station B #351819, Nashville, TN 37235-1819. Phone: 615-322-2482. Email: c.ahlin@Vanderbilt.edu. We thank Eric Bond, William Collins, Bernard Salanie, and seminar participants at Chicago, London School of Economics, Stanford Institute for Theoretical Economics, University College London, Vanderbilt, and Yale, as well as the editor and two referees for helpful comments. Financial support from the National Institutes of Health and the National Science Foundation (including its Graduate Research Fellowship Program) is gratefully acknowledged. The views expressed in this article are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of Chicago or the Federal Reserve System. All errors are ours.

1 Introduction

Micro credit is viewed as a major tool to alleviate world poverty, but practitioners are polarized in a debate concerning the virtues of individual versus group, joint liability lending. For the most part, only one contract per lender is observed, e.g., group-based loans under the Grameen Bank in Bangladesh and individual loans under BRI in Indonesia. It is thus difficult to make progress in the debate. We do not see what would happen if the alternative contract were available, not to mention all the difficulties inherent in cross country comparisons. Here we take advantage of the menu of contracts offered by a dominant rural lender in a single country and some unusual data gathered in field research to make some headway on the issue. Potential borrowers from the BAAC[1] in Thailand decide whether to borrow at all, and if so, whether to borrow as individuals or in a group. That is, potential borrowers select into and across loan contracts. There is in fact much variation in these choices in the data.

Our use of the data is naturally enough dictated by various well-known models from the contract theory, mechanism design literature that debates the virtues of individual versus group lending. We exposit and test two related models that compare an individualistic, relative performance regime with a group, cooperative regime in the context of moral hazard in effort provision. The defining feature of each regime is whether or not the borrowers are able to side-contract. The basic advantage of the individualistic regime, where there is no side-contracting, is the opportunity for the lender to gage if one borrower has been diligent by looking to see if other individual borrowers are doing well. The group regime, where borrowers can side-contract, does not allow this kind of information extraction since borrowers can collude against the lender, coordinating in the choice of low mutual effort or reallocating internal resources (e.g., tunneling). Its potential advantage, however, is the enhanced ability to monitor and enforce intra-group agreements on actions and transfers (i.e. risk sharing). It is thus not obvious the kind of regime for which moral hazard considerations argue; the optimal regime may vary with the parameters of the model.

[1]BAAC is the Bank for Agriculture and Agricultural Cooperatives.

Indeed, Holmström and Milgrom (1990) show that one crucial determinant of the optimal regime is the degree of correlation between the random component of borrowers' returns. In the individualistic regime, the principal, or lender, can use the correlation to mitigate its imperfect information. It does so by rewarding or punishing based on borrower performance comparisons, which diminishes the risk cost of high-powered incentives. The higher the correlation, the more effectively the information problem is dealt with in this regime. In the group regime, however, rewards based on performance comparisons can be manipulated, so higher correlation does not carry the same benefits. In fact, this regime turns out to be more effective the lower is the correlation, since low correlation makes the ability of the borrowers to coordinate and share risk among themselves more valuable. As intuition suggests and Holmström and Milgrom (1990) show formally, it would be Pareto optimal for the borrowers to be acting non-cooperatively if technological correlation is high enough, and cooperatively if it is low enough. This is a testable implication.

Prescott and Townsend (2002) address similar questions but focus instead on wealth levels of borrowers relative to the lender and wealth dispersion across borrowers. They offer an extended version of the above-mentioned unobserved effort model, generalized in several directions, but without closed form solutions available. In simulations, they find that for sufficiently asymmetric Pareto weights on the two borrowers, interpretable as high wealth dispersion, the cooperative regime dominates. When the weights are more similar, the non-cooperative regime dominates. Apparently the cooperative regime is better at extracting effort from the low-weight borrower. Further, the cooperative regime dominates for high enough *or* low enough reservation payoff of the borrowers, interpretable as the borrowers' wealth, while an intermediate value makes the non-cooperative regime optimal. Again, both of these predictions are in principle testable.

These two models focus on conditions under which one borrowing regime dominates another. A third model, Ghatak (1999), makes predictions about whether a given borrower will choose to borrow in a group or take the best alternative, which may mean not borrowing at all. Some borrowers are inherently more risky than others. In a setting of limited, individual liability and lender ignorance of risk-type, risky borrowers are in effect subsidized by safe ones. This can cut out socially productive loans to safe borrowers and leave only the risky types to borrow.[2] In this setting, loans that make use of joint liability within an endogenously formed group can mitigate the adverse selection problem. They do so by inducing homogeneous matching among risk types and reducing the subsidy from safe to risky borrowers, drawing back into the market relatively safe borrowers. While group loans always lessen adverse selection relative to individual loans, they may not eliminate it. A testable implication of Ghatak (1999) is thus the existence of adverse selection, which some argue is negligible in practice due to lenders' ability to collect information and their offering of other incentives.[3]

We modify this model to introduce correlation in borrower returns and find that borrowers with higher correlation of returns are more likely to self-select into group contracts (a testable proposition). They do this because correlation raises their payoff of borrowing by lowering the chances of facing liability for their partner's loan.

We attempt to test these predictions from the various models using relatively rich data gathered in field research in Thailand. The BAAC is the primary institutional lender in rural Thailand : for example 64% of the institutional loans in our sample are from the BAAC. But it is not the only lender. In our data we have a complete enumeration of loans outstanding or repaid in the last twelve months, and for each loan the household respondent is asked what if any collateral was used (land title, cosigner, other). Among these are loans from village funds, money lenders, commercial banks, and the informal sector. The BAAC offers both individual loans and group, joint liability loans; other institutional lenders may offer only one or the other. We test the models on the smaller sample of BAAC loans and the larger sample of all individual and co-signed loans from institutional lenders, including but not restricted to the BAAC. In addition to loan data, we also have measures for wealth, wealth dispersion,

[2]This kind of adverse selection dates back at least to Stiglitz and Weiss (1981).

[3]See Robinson (2001) for example.

technological correlation, and the risk of the borrowers, as well as numerous controls.

We link the non-cooperative regime in the moral hazard theories with individual loans in the data, and the cooperative regime in these theories with joint liability (i.e. co-signed), group loans in the data. Since these models focus on comparing two regimes against each other but not against outside alternatives, we restrict the sample to those who have had a group or individual loan outstanding in the last twelve months and exclude those who did not borrow at all. By contrast, when testing the adverse selection models of the literature, we include the whole sample and focus on the decision whether to borrow under joint liability or to choose any alternative (not borrow at all or borrow under individual liability).[4]

Our association of the loan types in the theory with analogs in actual practice deserves some elaboration. BAAC policy dictates that to receive a group loan, one must form or join an official BAAC-registered borrowing group and enter into a joint liability agreement. A box denoting that the group is the collateral is checked off on the loan form, and sometimes a particular member in the group is named as a cosigner. In contrast, individual loans must be guaranteed by some form of collateral, usually land. Thus the link of group loans in the data to group loans in Ghatak (1999) is relatively clear. Liability of one borrower for another within the group is written into the borrowing contract explicitly. Cases do exist in which the BAAC required group members to pay for a delinquent member of the group. Further, liability does indeed appear to be limited since no explicit collateral is required.

The key element in Holmström and Milgrom (1990) and Prescott and Townsend (2002) is the re-allocation of risk. The BAAC has in place a risk-contingency system, as documented in Townsend and Yaron (2001), for example. Farmers experiencing force majeur events report their difficulties to the local branch. The loan can be rescheduled, and in some instances interest and even part of principal forgiven, as if an insurance indemnity had kicked in. Thus group and individual credit contracts are state contingent loan repayment agreements. The key distinction in Holmström and Milgrom (1990) and Prescott and Townsend (2002)

[4]The results are robust to various classifications of those borrowing under individual contracts – as borrowers, as non-borrowers, or as excluded from the sample.

between the group and individual regimes is whether the borrowers can side-contract. Of course in the data, group loans are not accompanied with a full legal apparatus that enables borrowers to enforce side contracts; neither do individual loan contracts stipulate that borrowers not coordinate on financial or production-related decisions.[5] However, the group loan format does appear to be a natural way of facilitating, even encouraging, side-contracting between borrowers. Group loans are accompanied by required group meetings and explicit group-related contingencies. Conversely, the individual loan format builds in an institutional neglect of borrower side-contracting capabilities. Further, while not explicitly in the loan contracts, performance comparisons seem to be made across individual borrowers as theory would predict. A credit officer visits the village (the farmer and neighboring households) to verify the adverse event. More generally the BAAC is well positioned to make relative performance comparisons in deciding what to do about repayment problems. Thus, while undoubtedly imperfect, there is a highly plausible link between group and individual loans in the data and the group and individual regimes in Holmström and Milgrom (1990) and Prescott and Townsend (2002).

It should be noted that some key assumptions are different across the models as written. Whether or not the models are irreconcilable at a deeper level is beyond the scope of this paper. A priori, however, each seems to have a reasonable chance at explaining the data at hand. Adverse selection is an obstacle thought to be pervasive in many contexts, and quite plausible in the Thai rural context. Moral hazard and the use of risk-contingencies in loan contracts, consistent with the practice of the BAAC, are also not unlikely here. Indeed, we find some support for both types of model.

As predicted by Prescott and Townsend (2002), the prevalence of joint liability contracts relative to individual contracts exhibits a U-shaped relationship with the wealth of the

[5] It seems reasonable to apply Holmström and Milgrom (1990) and Prescott and Townsend (2002) not only to situations where side-contracting can be explicitly forbidden or enabled by the principal, but also to address the principal's choice of whether or not to *encourage* side-contracting in cases where ruling it out or in explicitly would be difficult. Encouragement can take the form of providing infrastructure for communication and social interaction. Rai and Sjöström (2004) highlight this point in a model of strategic default, arguing that the key element of a group contract is its facilitation of cross-reporting.

borrowing pair and increases with the wealth dispersion. The latter result is especially robust. The results surface only when we limit the sample to that suggested by theory: households with either a group or an individual loan, but not both. The results do not appear to be due to a conventional collateral effect (in which higher wealth enables households to borrow without cosignors); we separate out and control for the subset of wealth that is most commonly used as collateral.

Contrary to the theory of Holmström and Milgrom (1990) (and presumably also Prescott and Townsend (2002)), we find no evidence that joint liability borrowing becomes less likely as covariance of output increases. This is true when restricting the sample only to those with group or individual loans. When we use the full sample, however, as Ghatak (1999) would dictate, we find some evidence for the opposite. That is, we find that higher correlation makes joint liability borrowing more likely relative to all outside options, just as our modification of Ghatak (1999) suggests.

Finally, we find direct evidence consistent with the prediction of adverse selection in the credit market, in that the likelihood of joint-liability borrowing increases the lower is the probability of project success. This result only appears when the full sample is used. Thus the more risky households are borrowing. Ausubel (1999) tests this proposition with data from credit card companies: those willing to select into higher interest rates are more likely to default. A parallel insurance literature also tests whether households paying for more complete coverage are more likely to experience the adverse, insured event. (See the literature review in Chiappori and Salonie (2000)). Here we take advantage of measurements of risk type also of those who choose *not* to borrow and offer a direct test of this prediction of the adverse selection model. As Chiappori, Jullien, Salanie, and Salanie (2002) emphasize it is difficult to determine whether accident probabilities increase for those purchasing more complete coverage because of adverse selection (ex ante selection) or moral hazard (ex post shirking). Here we further restrict our sample to those who have borrowed in the last year but are not currently doing so, and find that the adverse selection remains a force in the data.

Strikingly, as alluded to above, most of these results confirming the models' predictions disappear if we do not condition the sample according to the dictates of the models.

2 Theories and Implications

2.1 Moral Hazard and Technological Correlation

We focus initially on selection across contracts, given that borrowing is taking place. Correlation of output across agents is related by Holmström and Milgrom (1990), in the context of unobserved actions, to the optimality of an individualistic, relative performance regime vis a vis a cooperative regime. They find (in Proposition 5) that there exists a cutoff technological correlation coefficient, $\overline{\rho} > 0$. *For $\rho < \overline{\rho}$, the optimal contract employs a cooperative or group regime; for $\rho > \overline{\rho}$, the optimal contract employs a relative performance regime. Thus, one should expect to see more of the group regime among agents whose output is less correlated.*

Here we describe the model used in Holmström and Milgrom (1990) to illustrate this result.[6] There are two agents, or borrowers, indexed by i. Each produces output q_i as a function of his effort e_i and some random shock ε_i. One could think of output q_i as varying with loan size also, but the latter is regarded as fixed and dropped from the notation. Output is then additive in effort and the shock:

$$q_i = e_i + \varepsilon_i, \; i = 1, 2.$$

The ε_i's are distributed joint-normally with means zero and variance-covariance matrix

$$\mathbf{\Sigma} \equiv \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}.$$

[6]Some notation and terminology are changed for the sake of continuity within this paper. Other than that, the model here is exactly what is used for the result in Proposition 5 of Holmström and Milgrom (1990).

The point is that a higher effort e_i makes higher output q_i more likely, but there is noise and the returns may be correlated. Note also that the projects may differ in risk σ_i^2; we control for variance of output in some empirical specifications (see section 3.3). Let $\boldsymbol{q}$ and $\boldsymbol{e}$ be the vectors $[q_1\ q_2]^T$ and $[e_1\ e_2]^T$, respectively.[7]

Since only the q_i's are publicly observed, borrower payoffs must be in terms of them. Attention is restricted to contracts giving agent i consumption c_i as a linear function of output:

$$c_i(\boldsymbol{q}; \kappa_{i0}, \boldsymbol{\kappa}_i) = \kappa_{i0} + \boldsymbol{\kappa}_i^T \boldsymbol{q},\ \ i = 1, 2,$$

where $\boldsymbol{\kappa}_i = [\kappa_{i1}\ \kappa_{i2}]^T$. Further let $\boldsymbol{\kappa} \equiv \{\kappa_{ij}\}$, for $i = 1, 2$ and $j = 0, 1, 2$; $\boldsymbol{\kappa}$ is just the collection of all the compensation parameters.

In this model, we consider preferences in which disutility of effort can be measured in consumption units by $C_i(\cdot)$, a strictly convex function. We also assume that agents maximize expected utility, where the utility function over consumption is exponential with coefficient of absolute risk aversion $r_i > 0$. Then, given effort choices e_1 and e_2, the certainty equivalent of contract $\boldsymbol{\kappa}$ has the following well-known form:

$$CE_i(\boldsymbol{e}; \boldsymbol{\kappa}) = \kappa_{i0} + \boldsymbol{\kappa}_i^T \boldsymbol{e} - C_i(e_i) - (1/2) r_i \boldsymbol{\kappa}_i^T \boldsymbol{\Sigma} \boldsymbol{\kappa}_i,$$

where $\boldsymbol{\kappa}_i^T \boldsymbol{\Sigma} \boldsymbol{\kappa}_i = \kappa_{i1}^2 \sigma_1^2 + \kappa_{i2}^2 \sigma_2^2 + 2\kappa_{i1}\kappa_{i2}\sigma_{12}$ is the variance of his compensation. Note that diversity in cost of effort and risk aversion is allowed, though we cannot control for these in the data beyond using education and demographic and occupational variables (see section 3.2). The lender is risk neutral and thus has certainty equivalence utility of

$$CE_p(\boldsymbol{e}; \boldsymbol{\kappa}) = e_1 + e_2 - (\boldsymbol{\kappa}_1 + \boldsymbol{\kappa}_2)^T \boldsymbol{e} - \kappa_{10} - \kappa_{20}.$$

Suppose first that the lender can deal with each borrower individually – the borrow-

[7] That is, $\boldsymbol{q}$ is a column vector and $\boldsymbol{q}^T$ a row vector; similarly for $\boldsymbol{e}$ and other vectors defined below.

ers do not observe each others' actions or outcomes and cannot side-contract. Then the lender sets both contracts to maximize his payoff subject to agents' participation constraints and incentive compatibility constraints. Since the model exhibits transferable utility, the optimal contract maximizes total surplus (the sum of all payoffs) subject to the incentive compatibility constraints only.[8] Thus, at the optimum, $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ solve

$$\max_{\boldsymbol{\kappa}_1,\boldsymbol{\kappa}_2} \quad e_1 + e_2 - C_1(e_1) - C_2(e_2) - 1/2 r_1 \boldsymbol{\kappa}_1^T \boldsymbol{\Sigma} \boldsymbol{\kappa}_1 - 1/2 r_2 \boldsymbol{\kappa}_2^T \boldsymbol{\Sigma} \boldsymbol{\kappa}_2$$

$$\text{subject to the first-order conditions for effort: } C_i'(e_i) = \kappa_{ii}, \; i = 1, 2.$$

Total surplus equals expected output (the first and second terms in the maximand) less costs of effort (the third and fourth terms) and risk costs (the fifth and sixth terms). The optimal contract satisfies

$$\kappa_{ii} = \frac{1}{1 + r_i \sigma_i^2 (1-\rho^2) C_i''}, \; \kappa_{ij} = -\kappa_{ii}\sigma_{12}/\sigma_j^2; \quad i = 1, 2, \; j = 1, 2, \; j \neq i$$

where $\rho \equiv \sigma_{12}/\sigma_1\sigma_2$ is the correlation coefficient. Note that the direct, own-production term κ_{ii} decreases in σ_i^2 and r_i as is natural with risk aversion. The cross term κ_{ij} varies inversely with the technological correlation σ_{12} and (in absolute value) with the risk of the other borrower σ_j^2. The risk cost, defined above, under this optimal contract can be calculated to be:

$$(1-\rho^2)\frac{r_1\kappa_{11}^2\sigma_1^2 + r_2\kappa_{22}^2\sigma_2^2}{2}. \tag{1}$$

[8]To see this, first note that a contract maximizing total surplus subject to incentive constraints, TSC, gives the same surplus as that of a contract maximizing total surplus subject to incentive constraints *and* agents' utility equalling their reservation level, TSCP. This is true because any contract can be transformed into one in which agents' participation constraints bind, without affecting incentive constraints or total surplus, through lump-sum utility transfers between lender and borrowers using the intercept terms κ_{10} and κ_{20}. Further, note that an optimal contract for the lender, OC, must give at least the same total surplus as TSCP. If it produced a lower surplus, i.e. sum of payoffs, the lender would be worse off under OC, since the borrowers can be no worse off than they are under TSCP (under which they receive their reservation utility level). But then OC would not be optimal, since TSCP gives a higher lender payoff and satisfies the same constraints. Of course, OC cannot give a greater total surplus than TSCP, since TSCP maximizes total surplus subject to the same constraints. Thus OC gives the same total surplus as TSCP, which gives the same total surplus as TSC. That is, any optimal contract maximizes total surplus subject only to incentive constraints.

If $\rho = 1$ then relative performance evaluation works perfectly well: all deviations in effort are perfectly detectable, and the lender offers full insurance. The risk sharing occurs because the correlation between shocks mitigates the principal's lack of information about the agents' efforts.

Suppose next that the two borrowers *can cooperate*, as is presumably easier within a joint liability group. Specifically, they do observe each others' actions and can commit to transfers with each other conditional on observed actions and outcomes.[9] (The principal still sees output only.) This allows them to mutually reinsure each other and to coordinate to an agreed upon set of actions. The side contracts they can write will be of the form

$$\tau(\boldsymbol{e}, \boldsymbol{q}) = \boldsymbol{\gamma}^T \boldsymbol{q} + t(e_1, e_2),$$

where $\tau(\boldsymbol{e}, \boldsymbol{q})$ gives the net transfer from agent 1 to agent 2 as a result of actions $\boldsymbol{e}$ and output realizations $\boldsymbol{q}$. The function $t(\cdot)$ allows the pair to enforce any set of actions as a Nash equilibrium. The mutual insurance agreements, for which coefficient $\boldsymbol{\gamma}$ denotes the vector $[\gamma_1 \ \gamma_2]^T$, are restricted to be linear in output, as above.

Holmström and Milgrom (1990) assume the pair will choose $t(\cdot)$ and $\boldsymbol{\gamma}$ to reach a Pareto optimal set of actions and transfers. Again, given transferable utility (which can be done within the group using $t(\cdot)$, for example), this implies the pair will maximize joint surplus. Given the external borrowing contract with its incentives $\boldsymbol{\kappa}$, the two borrowers thus choose $(\boldsymbol{e}, \boldsymbol{\gamma})$ to maximize[10]

$$(\boldsymbol{\kappa}_1 + \boldsymbol{\kappa}_2)^T \boldsymbol{e} - C_1(e_1) - C_2(e_2) - (r_1/2)\,(\boldsymbol{\kappa}_1 - \boldsymbol{\gamma})^T \boldsymbol{\Sigma} (\boldsymbol{\kappa}_1 - \boldsymbol{\gamma}) - (r_2/2)\,(\boldsymbol{\kappa}_2 + \boldsymbol{\gamma})^T \boldsymbol{\Sigma} (\boldsymbol{\kappa}_2 + \boldsymbol{\gamma}). \quad (2)$$

[9]Interestingly, Holmström and Milgrom (1990) show that as long as the pair can contract on actions, it makes no difference whether or not they can contract on outcomes. This is because the lender can accomplish the internal risk sharing through the contract even if they cannot, as discussed in footnote 11.

[10]For the sake of brevity, we suppress the incentive constraints that ensure $\boldsymbol{e}$ is optimal for each agent. In fact, any set of actions $\boldsymbol{e}$ can be enforced costlessly via a side contract that uses $t(\cdot)$ to punish deviations. For example, to enforce (e_x, e_y), the transfer function could be set $t(e_x, e_y) = 0$; $t(e', e_y) = B$ for $e' \neq e_x$; $t(e_x, e'') = -B$ for $e'' \neq e_y$; and $t(e', e'') = 0$ for $e' \neq e_x$ and $e'' \neq e_y$. For B large enough, (e_x, e_y) is a Nash equilibrium.

As before, the principal can be thought to be choosing $(\boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2)$ to maximize total surplus, constrained however by what the group is doing:

$$\max_{\boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2} \quad e_1 + e_2 - C_1(e_1) - C_2(e_2) - (r_1/2)\,(\boldsymbol{\kappa}_1 - \boldsymbol{\gamma})^T \boldsymbol{\Sigma}(\boldsymbol{\kappa}_1 - \boldsymbol{\gamma}) - (r_2/2)\,(\boldsymbol{\kappa}_2 + \boldsymbol{\gamma})^T \boldsymbol{\Sigma}(\boldsymbol{\kappa}_2 + \boldsymbol{\gamma})$$

$$\text{subject to: } (\boldsymbol{e}, \boldsymbol{\gamma}) \text{ maximizes (2) given } \boldsymbol{\kappa}.$$

Holmström and Milgrom (1990) prove (Proposition 3), using results of Wilson (1968), that the lender's optimal design coincides with that for a single ("syndicate") borrower whose effort cost function satisfies $C(e_1, e_2) = C_1(e_1) + C_2(e_2)$ and absolute risk aversion coefficient r satisfies $1/r = 1/r_1 + 1/r_2$. In other words, the risk tolerance coefficient (the inverse of r) of the syndicate borrower is greater than the risk tolerance of each individual borrower. This corresponds to a lower total risk cost to the pair, due to internal risk sharing.[11] [12] Now the principal is reduced from four degrees of freedom or parameters to two, since what matters is not $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ individually, but the sum $\boldsymbol{\kappa}_1 + \boldsymbol{\kappa}_2$. So, without loss of generality, assume $\kappa_{12} = \kappa_{21} = 0$.[13] Given $(\kappa_{11}, \kappa_{22})$, the total risk cost is

$$(1/2) r [\kappa_{11}^2 \sigma_1^2 + \kappa_{22}^2 \sigma_2^2 + 2\kappa_{11}\kappa_{22}\rho\sigma_1\sigma_2], \tag{3}$$

increasing in ρ.

Inspection of expression 2 also gives that here, as in the relative performance regime, actions will be chosen that equate κ_{ii} and $C_i'(e_i)$. Thus any pair of actions (e_1, e_2) must be

[11] This result obtains whether or not the borrowers can themselves contract on output realizations, as shown in Proposition 3 of Holmström and Milgrom (1990). If they can contract on output, this ability will lead them to set $\boldsymbol{\gamma}$ optimally to minimize the risk cost, without affecting incentives for actions. If the pair does not have this ability, the principal can build the same insurance into the contract without affecting incentives. In particular, if $(\boldsymbol{\kappa}_1^*, \boldsymbol{\kappa}_2^*)$ and $\boldsymbol{\gamma}^*$ are optimal in the case where the group can contract on output, then the contract parameters $(\boldsymbol{\kappa}_1^* - \boldsymbol{\gamma}^*, \boldsymbol{\kappa}_2^* + \boldsymbol{\gamma}^*)$ produce the same outcome in the case where the group cannot contract on output: the risk cost is the same, and the incentives are also the same since $\boldsymbol{\kappa}_1^* + \boldsymbol{\kappa}_2^* = \boldsymbol{\kappa}_1^* - \boldsymbol{\gamma}^* + \boldsymbol{\kappa}_2^* + \boldsymbol{\gamma}^*$.

[12] It is assumed that there is no mechanism available for the lender to elicit borrower revelation of each others actions (information that each borrower is now assumed to know about the other). The agents can side-contract, so any such mechanism would have to be collusion-proof.

[13] Fix an optimal contract $(\boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2)$ and optimal risk-sharing agreement $\boldsymbol{\gamma}$ and effort selection $\boldsymbol{e}$. Observe that $\boldsymbol{\kappa}_1' = [\kappa_{11} + \kappa_{21},\, 0]^T$, $\boldsymbol{\kappa}_2' = [0,\, \kappa_{12} + \kappa_{22}]^T$, $\boldsymbol{e}' = \boldsymbol{e}$, and $\boldsymbol{\gamma}' = [\gamma_1 + \kappa_{21},\, \gamma_2 - \kappa_{12}]^T$ give the same surplus and satisfy the same constraints. This contract must therefore also be optimal.

implemented by the same contract parameters $(\kappa_{11}, \kappa_{22})$ in both regimes. This enables us to determine easily which regime delivers higher total surplus when implementing a given set of actions. For a given set of actions, the only part of total surplus that varies by regime is the risk cost, given in expressions 1 and 3, respectively; and these risk costs are easily compared, for a given set of actions, since $(\kappa_{11}, \kappa_{22})$ are the same in both regimes.

It can be checked, using expressions 1 and 3, that at $\rho = 0$ the risk cost of implementing any set of actions $\boldsymbol{e}$ is lower under the cooperative regime.[14] It follows that at $\rho = 0$, the cooperative regime gives higher total surplus than the individualistic regime. Similarly, at $\rho = 1$ the risk cost of implementing any set of actions $\boldsymbol{e}$ is lower under the individualistic regime, and thus the individualistic regime gives higher total surplus. Further, as ρ increases, the cost of implementing every set of actions in the cooperative regime is increasing (see expression 3), which implies that maximized surplus under this regime is decreasing in ρ. The cost of implementing every set of actions in the individualistic regime is strictly decreasing in ρ (see expression 3), which implies that the maximized surplus under this regime is strictly increasing in ρ.

In summary, holding risk aversion and other parameters constant, the payoff to the lender is strictly increasing in ρ under relative performance and decreasing in ρ under cooperation; and at $\rho = 0$, the cooperative regime dominates, while at $\rho = 1$, the relative performance regime dominates. *This proves that there is a cutoff,* $\overline{\rho} \in (0, 1)$*, above which the individualistic regime dominates and below which the cooperative regime does.* The intuition is that when correlation is high, the scope for internal risk-sharing is low, while the lender is able to offer significant insurance through relative performance comparisons; and vice versa for low correlation.[15]

[14] This follows from the fact that $r = r_1 r_2/(r_1 + r_2) < r_1, r_2$.

[15] Holmström and Milgrom (1990) and this exposition restrict attention to non-negative ρ. For ρ in the negative range, the comparison is likely to reverse again at some point in the neighborhood of $\rho = -1$. To see this, note from expression 1 that a high correlation *in absolute value* is what makes the relative performance regime effective. In fact, for $\rho = -1, 1$, borrowers bear no risk even under first-best incentives ($\kappa_{ii} = 1$). But note from expression 3 that the cooperative regime is *also* getting better at bearing risk as $\rho \to -1$. Clearly more negative correlation increases the scope for internal risk-sharing. But note that at $\rho = -1$ and $\kappa_{ii} = 1$, the risk cost in the cooperative regime equals $1/2r(\sigma_1 - \sigma_2)^2$. Thus unless both agents have the

2.2 Moral Hazard and Wealth Distribution

E. S. Prescott and Townsend (2002) offer an extended version of Holmström and Milgrom (1990) in which wealth levels and distribution also affect regime optimality. The model still features unobserved, costly effort and analyzes a Pareto problem with a principal and two agents. Non-transferable utility, however, will imply that the optimal regime varies as one moves along the Pareto frontier.

Cooperation or joint liability is again compared with non-cooperative or individual loans. The two regimes are compared in terms of their respective Pareto frontiers. In the cooperative regime, agents can internally enforce any set of actions and commit to a set of internal Pareto weights, according to which they will divide effort and consumption. In the individualistic regime, transfers and coordination on actions are disallowed.

There are two agents and two technologies, indexed by i and j, respectively. The effort agent i exerts on technology j is e_{ij}. A special case (as in Holmström and Milgrom (1990) and used in the simulations here) would be $e_{ij} = 0$ for $i \neq j$, so agent i works only his own project. Total effort exerted by agent i is $e_i \equiv e_{i1} + e_{i2}$. Define a_j as the total effort exerted on technology j; $a_j \equiv e_{1j} + e_{2j}$. Let q_j be the output from technology j. Finally, define vectors $\boldsymbol{c} \equiv [c_1, c_2]^T$, $\boldsymbol{q} \equiv [q_1, q_2]^T$, $\boldsymbol{a} \equiv [a_1, a_2]^T$ and $\boldsymbol{e_{i\cdot}} \equiv [e_{i1}, e_{i2}]^T$.

Agent i maximizes utility $U_i(c_i) + V_i(T_i - e_i)$, defined over consumption c_i and leisure, which equals total time endowment T_i minus total effort. As in Holmström and Milgrom (1990), varying degrees of risk aversion and work disutility could be allowed. In contrast to Holmström and Milgrom (1990), here utility is not generally transferable and preferences of the agents do not aggregate. Thus the analysis does not reduce to maximizing total surplus subject to incentive constraints. The principal's payoff is $W(q_1 + q_2 - c_1 - c_2)$, a special case

same variance of output, the group still bears positive risk. This is because the group still faces aggregate risk under perfect negative correlation, unless both members have the same variance.

So in general, for ρ in the neighborhood of negative one, the relative performance regime begins to dominate again. We restrict attention to the result of Holmström and Milgrom (1990) on $\rho \in [0, 1]$. In the context of mainly agricultural, rural borrowers, it is hard to imagine significant negative correlation between individual farmers.

of which is linear as in Holmström and Milgrom (1990) and as we assume in the exposition below. Technology is expressed as a probability mass function $p(\boldsymbol{q}|\boldsymbol{a})$, and again, this can be parameterized and allows heterogeneity.[16]

The Pareto problem to determine the information-constrained optimal allocation for each regime consists of maximizing the lender's expected utility, subject to the λ-weighted average of the borrowers' expected utilities reaching at least $\overline{G}$, and the appropriate incentive constraints.[17] Suppose for simplicity that each of c_i, q_j, and e_{ij} can take on a finite number of values. In the *individualistic case*, the credit contract induces individuals' effort $\boldsymbol{e_{i\cdot}}$ and specifies individualized consumption amounts $\boldsymbol{c}$ as a function of output $\boldsymbol{q}$.[18] Linearity in compensation is not imposed however, and further, randomness in consumption is allowed. Even the contract itself can be randomized ex ante, inducing a random choice of effort. The more general notation is thus a lottery over all objects: $\Pi(\boldsymbol{c}, \boldsymbol{q}, \boldsymbol{e_{1\cdot}}, \boldsymbol{e_{2\cdot}})$. However, the optimal contract may often involve non-random effort recommendations and consumption allocations, dependent on $\boldsymbol{q}$ alone.

The optimal contract is found by maximizing

$$\sum_{\boldsymbol{c},\boldsymbol{q},\boldsymbol{e_{1\cdot}},\boldsymbol{e_{2\cdot}}} \Pi(\boldsymbol{c}, \boldsymbol{q}, \boldsymbol{e_{1\cdot}}, \boldsymbol{e_{2\cdot}}) \ (q_1 + q_2 - c_1 - c_2)$$

subject to:

$$\sum_{\boldsymbol{c},\boldsymbol{q},\boldsymbol{e_{1\cdot}},\boldsymbol{e_{2\cdot}}} \Pi(\boldsymbol{c}, \boldsymbol{q}, \boldsymbol{e_{1\cdot}}, \boldsymbol{e_{2\cdot}}) \sum_i \lambda_i [U_i(c_i) + V_i(T_i - e_i)] \geq \overline{G},$$

$$\sum_{\boldsymbol{c}} \Pi(\boldsymbol{c}, \overline{\boldsymbol{q}}, \overline{\boldsymbol{e}}_{\boldsymbol{1}\cdot}, \overline{\boldsymbol{e}}_{\boldsymbol{2}\cdot}) = p(\overline{\boldsymbol{q}}|\overline{\boldsymbol{e}}_{\boldsymbol{1}\cdot} + \overline{\boldsymbol{e}}_{\boldsymbol{2}\cdot}) \sum_{\boldsymbol{c},\boldsymbol{q}} \Pi(\boldsymbol{c}, \boldsymbol{q}, \overline{\boldsymbol{e}}_{\boldsymbol{1}\cdot}, \overline{\boldsymbol{e}}_{\boldsymbol{2}\cdot}), \quad \forall \overline{\boldsymbol{q}}, \overline{\boldsymbol{e}}_{\boldsymbol{1}\cdot}, \overline{\boldsymbol{e}}_{\boldsymbol{2}\cdot},$$

[16] The uncertainty, including the correlation between borrower returns, is imbedded in the function $p(\boldsymbol{q}|\boldsymbol{a})$. It need not take the form of joint normality as in Holmström and Milgrom (1990).

[17] The exposition of the paper itself maximizes the (weighted) sum of borrower utilities subject to a participation constraint for the lender. Here we reverse it to parallel the exposition of Holmström and Milgrom (1990). Either formulation of the problem enables one to trace out the Pareto frontier.

[18] As in Holmström and Milgrom (1990), loan size could be made to vary, but is regarded as fixed and dropped from the notation.

and incentive constraints

$$\sum_{\boldsymbol{c},\boldsymbol{q},\boldsymbol{e_2.}} \Pi(\boldsymbol{c},\boldsymbol{q},\boldsymbol{e_1.},\boldsymbol{e_2.})[U_1(c_1)+V_1(T_1-e_1)] \geq$$

$$\sum_{\boldsymbol{c},\boldsymbol{q},\boldsymbol{e_2.}} \Pi(\boldsymbol{c},\boldsymbol{q},\boldsymbol{e_1.},\boldsymbol{e_2.}) \frac{p(\boldsymbol{q}|\boldsymbol{\hat{e}_1.}+\boldsymbol{e_2.})}{p(\boldsymbol{q}|\boldsymbol{e_1.}+\boldsymbol{e_2.})}[U_1(c_1)+V_1(T_1-\hat{e}_1)],\ \forall \boldsymbol{e_1.},\boldsymbol{\hat{e}_1.}$$

and

$$\sum_{\boldsymbol{c},\boldsymbol{q},\boldsymbol{e_1.}} \Pi(\boldsymbol{c},\boldsymbol{q},\boldsymbol{e_1.},\boldsymbol{e_2.})[U_2(c_2)+V_2(T_2-e_2)] \geq$$

$$\sum_{\boldsymbol{c},\boldsymbol{q},\boldsymbol{e_1.}} \Pi(\boldsymbol{c},\boldsymbol{q},\boldsymbol{e_1.},\boldsymbol{e_2.}) \frac{p(\boldsymbol{q}|\boldsymbol{e_1.}+\boldsymbol{\hat{e}_2.})}{p(\boldsymbol{q}|\boldsymbol{e_1.}+\boldsymbol{e_2.})}[U_2(c_2)+V_2(T_2-\hat{e}_2)],\ \forall \boldsymbol{e_2.},\boldsymbol{\hat{e}_2.}.$$

The first constraint ensures a given (λ_i-weighted) amount of utility for the pair. Note that as $\overline{G}$ increases, in effect the wealth of each borrower increases. The second ensures that the mechanism assigns technologically feasible probabilities. The last two are incentive constraints ensuring agents 1 and 2 each abide by their respective recommended effort allocations. Of course the contract must also satisfy probability measure constraints

$$\Pi(\boldsymbol{c},\boldsymbol{q},\boldsymbol{e_1.},\boldsymbol{e_2.}) \geq 0, \quad \forall\ \boldsymbol{c},\boldsymbol{q},\boldsymbol{e_1.},\boldsymbol{e_2.}$$

and

$$\sum_{\boldsymbol{c},\boldsymbol{q},\boldsymbol{e_1.},\boldsymbol{e_2.}} \Pi(\boldsymbol{c},\boldsymbol{q},\boldsymbol{e_1.},\boldsymbol{e_2.}) = 1.$$

In the *cooperative, joint liability case*, the borrowers are able to side contract. Thus they can internally enforce any set of (Pareto optimal) actions, conditional on the borrowing contract offered by the lender. They are also free to share risk internally by committing to ex post transfers as in Holmström and Milgrom (1990). This means that the contract cannot specify the c_i's separately, only total group consumption, $c_g \equiv c_1 + c_2$ as a function of $\boldsymbol{q}$. Similarly, it cannot specify the individual effort choices ($\boldsymbol{e_1.}$, $\boldsymbol{e_2.}$) separately, only the total effort on each technology, the a_j's. Given preferences that are separable in consumption and leisure, the actual consumption of agent i can be expressed as a function of total consumption and the internal Pareto weights, $c_i(c_g, \mu)$, where μ is the weight on agent 1, say. Similarly, effort of agent i can be expressed as a function of total effort on each technology and the internal Pareto weights, $e_i(\boldsymbol{a}, \mu)$.[19]

The joint liability contract here takes the form $\Pi(c_g, \boldsymbol{q}, \boldsymbol{a}, \mu)$. It should also maximize

$$\sum_{c_g, \boldsymbol{q}, \boldsymbol{a}, \mu} \Pi(c_g, \boldsymbol{q}, \boldsymbol{a}, \mu)\ (q_1 + q_2 - c_g)$$

subject to:

$$\sum_{c_g, \boldsymbol{q}, \boldsymbol{a}, \mu} \Pi(c_g, \boldsymbol{q}, \boldsymbol{a}, \mu) \sum_i \lambda_i [U_i\left(c_i(c_g, \mu)\right) + V_i\left(T_i - e_i(\boldsymbol{a}, \mu)\right)] \geq \overline{G},$$

$$\sum_{c_g} \Pi(c_g, \overline{\boldsymbol{q}}, \overline{\boldsymbol{a}}, \overline{\mu}) = p(\overline{\boldsymbol{q}}|\overline{\boldsymbol{a}}) \sum_{c_g, \boldsymbol{q}} \Pi(c_g, \boldsymbol{q}, \overline{\boldsymbol{a}}, \overline{\mu}), \quad \forall \overline{\boldsymbol{q}}, \overline{\boldsymbol{a}}, \overline{\mu},$$

and

$$\sum_{c_g, \boldsymbol{q}} \Pi(c_g, \boldsymbol{q}, \boldsymbol{a}, \mu) \sum_i \mu_i [U_i\left(c_i(c_g, \mu)\right) + V_i\left(T_i - e_i(\boldsymbol{a}, \mu)\right)] \geq$$
$$\sum_{c_g, \boldsymbol{q}} \Pi(c_g, \boldsymbol{q}, \boldsymbol{a}, \mu) \frac{p(\boldsymbol{q}|\hat{\boldsymbol{a}})}{p(\boldsymbol{q}|\boldsymbol{a})} \sum_i \mu_i [U_i\left(c_i(c_g, \mu)\right) + V_i\left(T_i - e_i(\hat{\boldsymbol{a}}, \mu)\right)] \quad \forall \boldsymbol{a}, \hat{\boldsymbol{a}}, \mu.$$

There is a single group incentive constraint in this case: only joint deviations in effort are considered since the group can internally commit to a set of actions. Finally, the contract must satisfy analogous probability measure constraints to those defined above.

As has been pointed out, the setup of Prescott and Townsend (2002) is similar to that of Holmström and Milgrom (1990) of section 2.1, but more general in several dimensions. This generality necessitates computation in order to characterize the solution, and in particular, how the optimal regime varies with $\overline{G}$ and the λ_i's. (We note that the optimal regime is

[19] If effort is not reallocable across technologies, that is agent 1 is tied to technology 1 and similarly for agent 2, then $e_i(\boldsymbol{a}, \mu) = a_i$. If effort is reallocable, then $e_i(\boldsymbol{a}, \mu) = e_i(e_g, \mu)$, where $e_g = a_1 + a_2$.

likely to vary with correlation in the same way as in Holmström and Milgrom (1990), but this is not the focus.) Attention in the simulations was restricted to the case in which each agent must work his own technology (that is, $e_{ij} \equiv 0, i \neq j$) and there is some positive correlation across technologies. Each agent also was given symmetric utility, endowment, and technological parameters.

The simulations indicate that the cooperative regime dominates for highly asymmetric λ_i's, no matter what $\overline{G}$ is. The cooperative regime is relatively better at extracting wealth from the low-weight agent and giving it to the high-weight agent. The low-weight agent in the group can be forced to work hard without incentive problems, due to the commitment ability within the group. For more symmetric λ_i's, the relative performance regime can dominate, but this is highly dependent on $\overline{G}$. For $\overline{G}$ low enough, it strictly dominates nowhere; the cooperative regime is always optimal. As $\overline{G}$ increases, the (symmetric) range around $\lambda_1 = 1/2$,[20] for which the non-cooperative regime dominates increases up to a certain point, making individual loans more likely. For further increases in $\overline{G}$, as the group becomes wealthier, the (symmetric) range around $\lambda_1 = 1/2$ for which the relative performance regime dominates diminishes, making the group regime attractive over a wider range.[21]

In a decentralization of the model, $\overline{G}$ would vary in the same direction as the wealth of the borrowers. Prescott and Townsend (2002) show that in a decentralized competitive equilibrium with a continuum of borrowing pairs for each type (distinguished by preferences, technology, and the λ_i-weights), lenders break even on the contract Π and $\overline{G}$ is such that the expected surplus is zero for each group type. In Thailand the BAAC branch does not necessarily break even. It can make losses covered by a government subsidy (or make

[20] We normalize the sum $\lambda_1 + \lambda_2$ to one.

[21] The intuition for the group regime's wider range of dominance at high or low levels of $\overline{G}$ seems to be as follows. For $\overline{G}$ high enough, the agents are not required to exert much effort. Both regimes are equally good when providing incentives for effort is not critical and redistribution is unimportant. Thus both are equally good at symmetric λ_i's, whereas for asymmetric λ_i's, the group regime dominates due to its advantage at redistributing wealth among the agents. For $\overline{G}$ low, the key is extracting wealth from the group. It appears the group regime is better at this because of internal commitment, while the relative performance regime requires more high-powered incentives, i.e consumption rewards, to induce high effort. It is unclear how robust this latter result is.

profits). Similarly, the theory can allow taxes and transfers to each group type as in the second welfare theorem. More to the point, as $\overline{G}$ goes up and the expected surplus to the lender goes down, consumption and leisure go up, as if wealth of the borrowing group were going up. Consumption and wealth are correlated in the data.

Likewise the asymmetry of the λ_i's would reflect wealth dispersion. As λ_1 increases relative to λ_2, there is a tendency to assign higher mean consumption and lower mean effort to household 1 relative to household 2. The λ_i's within the group might be thought of as determined ex ante in a bargaining problem, determined by reservation (autarky) utility, hence captured by wealth.[22]

Under these interpretations, the simulations thus suggest that *we should expect to see the cooperative regime occurring more frequently at higher levels of wealth inequality within a group. Controlling for inequality, we should expect to see the cooperative regime varying negatively with wealth when wealth is low and positively when it is high. This is a U-shaped relationship of the cooperative regime prevalence with group wealth.*

2.3 Adverse Selection and Technological Correlation

Ghatak (1999) provides an adverse selection model of joint liability borrowing. Here the internal information is on the risk-type. The key choices of the agent in this model are the partner to join with in borrowing, and whether or not to borrow at all. Joint liability is shown to take advantage of the information borrowers have about each other in a way that enables decreasing the interest rate and drawing in safer borrowers, thus mitigating the adverse selection problem.

Regarding the choice of the borrowing partner of the agent, Ghatak (1999) shows that agents form groups that are homogeneous in risk-type. Regarding the choice about whether or not to borrow, the model delivers a cutoff risk-type, $\hat{p}$, such that all borrowers riskier than

[22]In a standard competitive equilibrium, exposited in Mas-Colell et al. (1995) and especially Negishi (1960), there is a positive monotone relationship between the λ-weight in the Pareto problem and the wealth on the right hand side of the Arrow-Debreu budget constraint.

$\hat{p}$ borrow, and others do not. The latter prediction is our focus here.

A continuum of borrowers is assumed, each one with an exogenously given project indexed by a risk-type, p. This is the key heterogeneity. There is a density $g(p) > 0$ of borrowers at each type $p \in [\underline{p}, 1]$, where $\underline{p}$ is a parameter of the model. The project of a borrower of type p yields output $q(p)$ with probability p, independent of other borrowers' returns, and gives zero output otherwise. Ghatak (1999) also assumes that

$$pq(p) = E, \quad \forall p \in [\underline{p}, 1].$$

Hence the borrowers do not differ in expected output. The only heterogeneity is in the second moment of the distribution. A lower p implies higher variance, i.e. higher risk.

Thus the distribution of output is simpler here – independently distributed according to the binomial distribution – than in Holmström and Milgrom (1990) and Prescott and Townsend (2002), where there is correlation, for example. A further distinction is that in Ghatak (1999) the adverse selection is with respect to the risk but not the mean, while in Holmström and Milgrom (1990) and Prescott and Townsend (2002), the moral hazard affects the mean but not the risk structure (variance or correlation) of the output distribution.

Borrowers are taken to be risk-neutral, unlike the previous models. It is assumed that one unit each of capital and labor are required to carry out the project. Each potential borrower is endowed with one unit of labor, which could be used by itself to produce $\underline{u}$, and no capital. Capital is available from a lender who offers borrowers take-it-or-leave-it contracts. Again, loan size could be introduced into the current notation, but as in the other models, we suppress it here.

The lender is assumed to observe output imperfectly. In particular, it sees only whether each project has succeeded ($q > 0$) or failed ($q = 0$), but not the precise amount of output (q). Thus the lender can contract on the binary outcome of success or failure, but not on the precise amount of output or risk-type, which would be optimal if these were observable.[23] This assumption of imperfectly observed output in Ghatak (1999) distinguishes it from Holmström and Milgrom (1990) and Prescott and Townsend (2002), where the lender observes output precisely. However, it may be that the results of Holmström and Milgrom (1990) and Prescott and Townsend (2002) would also hold up under a case of imperfect output observability, e.g. if the lender could only classify the output level into two or three categories. We do not verify this here, but do note that substantial yet imperfect output observability is likely the case in our data.

Ghatak (1999) focuses the analysis on groups of size two. A borrower of type p who pairs with one of type p' has expected payoff of

$$E - pr - p(1 - p')l, \tag{4}$$

where $r > 0$ is the interest plus principal, paid whenever the borrower succeeds, and $l > 0$ the joint liability payment, paid whenever the borrower succeeds and his partner does not. Limited liability means that nothing can be extracted from a borrower who fails. In a sense this is equivalent to a participation constraint that must hold in all states of the world, rather than one that must hold only in expectation, as in Holmström and Milgrom (1990) and Prescott and Townsend (2002).

By inspection, the cross-partial of payoff 4 with respect to p and p' is $l > 0$. Intuitively, safe borrowers benefit relatively more from safe partners. This is because safe borrowers succeed more often, and are thus in the position of having a delinquent partner affect their payoff more often. Thus borrowers match homogeneously in risk-type, as Ghatak (1999) proves formally.

Homogeneous matching means the payoff of a borrower of type p will be the following in

[23]Clearly, if type were observable, the lender could tailor the interest rate to each borrower's likelihood of success and eliminate the adverse selection problem. If output $q(p)$ were observable, this would be equivalent to knowing the borrower's type in this model, since $p = E/q(p)$ (and E would presumably be known if there were many borrowers). This would then reduce to the case where type is observable and the lender can perfectly price discriminate.

equilibrium:

$$E - pr - p(1-p)l.$$

Note that the derivative of the payoff with respect to p is $-[r + l(1-2p)]$. As long as $l \leq r$, which we assume, this derivative is strictly negative for $p \in [\underline{p}, 1)$. Thus, the safer an agent's type, the lower his payoff from borrowing and undertaking the project. Since all agents have an outside option that pays $\underline{u}$, agents will borrow if and only if the payoff from borrowing is greater than $\underline{u}$. Given that the payoff of borrowing is declining in p, *there exists a cutoff risk-type, call it $\hat{p}$, such that borrowers of type $p > \hat{p}$ will not find it optimal to borrow, and borrowers of type $p < \hat{p}$ will borrow.*[24] [25] This prediction of adverse selection in a limited liability credit market is general and can be found in Theorem 1 of Stiglitz and Weiss (1981).

Next we extend the model to consider correlation of output between borrowers, in the spirit of Holmström and Milgrom (1990).[26] There is a unique way to introduce correlation that preserves the individual probabilities of success, for a given pair of borrowers, (p, p'). It is:

	2 Succeeds (p')	2 Fails $(1-p')$
1 Succeeds (p)	$p\,p' + \epsilon$	$p(1-p') - \epsilon$
1 Fails $(1-p)$	$(1-p)p' - \epsilon$	$(1-p)(1-p') + \epsilon$

By inspection, each row and column adds to the correct individual probability of success or failure. Of course $\epsilon = 0$ is the zero-correlation case, while $\epsilon > 0$ introduces positive correlation.[27] The above is the unique joint distribution, *given a pair* (p, p') of individual probabilities. In general, ϵ may depend on (p, p'), so we write $\epsilon(p, p')$.

[24] The level of $\hat{p}$ depends on other characteristics of an agent, such as his outside option $\underline{u}$ and expected return E. We control for these using observable demographic, capital and occupational choice variables, as discussed in sections 3.1 and 3.2.

[25] Ghatak (1999) provides conditions under which joint liability will completely eliminate adverse selection (so that $\hat{p} = 1$) and under which it will reduce but not eliminate the adverse selection problem (so that $\hat{p} < 1$). Our test will only be valid for an interior cutoff (i.e. $\hat{p} < 1$) and thus could in theory fail if joint liability has eliminated the adverse selection problem.

[26] The same modification to Ghatak (1999) is made in Ahlin and Townsend (2002). There the focus is on how the repayment rate is expected to vary with the level of correlation, while here we look at how correlation affects the likelihood of borrowing.

[27] As in Holmström and Milgrom (1990), attention is restricted here to positive correlation.

No general result is available. We thus proceed by imposing some symmetry in the correlation across all project pairings. First we may assume $\epsilon(p, p')$ to be constant:[28]

$$\epsilon(p, p') = \epsilon, \ \forall p, p'. \tag{5}$$

The second assumption,

$$\epsilon(p, p') = \tilde{\rho} * \min\{p'(1-p), p(1-p')\}, \tag{6}$$

gives any homogeneous group the same *correlation coefficient* over project returns, equal to $\tilde{\rho}$, as a straightforward calculation shows.[29] For non-homogeneous groups, that is, those for whom $p \neq p'$, $\tilde{\rho}$ is not the correlation coefficient, but something closely related: it is the correlation, expressed as a fraction of the maximum correlation possible given individual probabilities of p and p'.[30] When $p = p'$, the maximum possible correlation coefficient is one, so $\tilde{\rho}$ equals the correlation coefficient. This formulation is the unique way of affecting each potential group's (appropriately normalized) correlation coefficient symmetrically. It is therefore perhaps the closest analog to the setup of Holmström and Milgrom (1990), in

[28]Note that ϵ is restricted in value by the need to keep each cell of the distribution matrix no less than zero and no greater than one. For the off-diagonal cells of the matrix to be positive, we need ϵ to be less than $p(1-p')$ and $p'(1-p)$. As p or p' approaches one, ϵ is forced to zero, which is not interesting. To avoid this problem, we can modify the support of risk-types innocuously to exclude those with probability of success near 1: $[\underline{p}, \overline{p}]$ for some $\overline{p} < 1$. This ensures that all borrowers face at least some uncertainty. One can verify that under this modified support, $\epsilon \leq \underline{p}(1-\overline{p})$ is necessary (and sufficient if $\epsilon \geq 0$) to ensure no cell in the distribution matrix of every possible pairing (p, p') exceeds one or falls below zero.

[29]It is clear that under this second assumption, any $\tilde{\rho} \in [0, 1]$ is allowable, that is, keeps each cell in the distribution matrix of every possible pairing (p, p') from exceeding one or falling below zero.

[30]Note that for non-homogeneous groups, it is a theoretical impossibility to have perfect correlation; two binomial variables with different individual probabilities of success can never be perfectly correlated. The further apart are the two probabilities, the lower the maximum correlation possible. One can calculate that the maximum correlation coefficient across two projects (p, p'), call it $\overline{\rho}(p, p')$, is equal to

$$\overline{\rho}(p, p') = \min\{\sqrt{\frac{p(1-p')}{p'(1-p)}}, \sqrt{\frac{p'(1-p)}{p(1-p')}}\}.$$

Further, the correlation coefficient corresponding to the $\epsilon(p, p')$ of assumption 6 can be shown to equal $\tilde{\rho} * \overline{\rho}(p, p')$. Thus $\tilde{\rho}$ has the interpretation, described in the text, as the correlation expressed as a fraction of the maximum possible correlation, and it can vary freely from zero to one. Of course, it is clear from the formula that $\overline{\rho}(p, p) = 1$, so for homogeneous groups, $\tilde{\rho}$ is just the correlation coefficient itself.

which the correlation coefficient is fixed independent of the group's effort choices, (e_1, e_2).

We assume here that the borrowers know their own correlation and factor it into their payoffs, but that the lender does not observe borrower correlation. Thus the contract parameters are independent of the correlation. This is a substantial difference from Holmström and Milgrom (1990) in which correlation is a key determinant of the contract. Thus, any related empirical test of the two models' correlation predictions may be partly an implicit test of lender rationality with respect to correlation. On the other hand, one might argue that in Holmström and Milgrom (1990) all that matters is that the borrowers know their own correlation, as assumed here. This is because the contracts derived there maximize total surplus. Imagine a lender offering a menu of contracts, each element of which earns it some baseline expected payoff. The borrowers will then choose the surplus-maximizing contract based on their own knowledge of correlation, since this also maximizes their own surplus. The lender does not need to know all the borrower details, only that the menu of contracts earns it a baseline payoff and offers the borrower sensible choices. These it may know in practice through history, borrower feedback, and experimentation, rather than direct observation of correlation. Thus the prediction of Holmström and Milgrom (1990) may be compatible with an imperfectly rational lender.

The first step in examining correlation is to determine how it affects the equilibrium homogeneous matching result, if at all. The forms of correlation expressed in 5 and 6 do not, as can be shown by verifying that the gain to a risky borrower from switching to a safe partner is less than the loss to a safe borrower from switching to a risky partner (for $\tilde{\rho}$ non-negative, see Ahlin and Townsend 2002). Given that homogeneous matching still obtains, the borrower payoff under correlation now becomes

$$E - pr - [p(1-p) - \epsilon(p,p)]l.$$

The indifference equation that defines the cutoff type $\hat{p}$[31] under assumption 5 becomes

$$E - \hat{p}r - [\hat{p}(1-\hat{p}) - \epsilon]l = \underline{u}, \tag{7}$$

and under assumption 6,

$$E - \hat{p}r - \hat{p}(1-\hat{p})(1-\tilde{\rho})l = \underline{u}. \tag{8}$$

By inspection, the left-hand sides of equations 7 and 8 are increasing in ϵ and $\tilde{\rho}$, respectively. That is, higher correlation raises the payoff of the indifferent borrower. Clearly, for higher ϵ, $\hat{p}$ will also have to be higher (since the borrowing payoff is still decreasing in p) in order to bring down the payoff and keep it equal to $\underline{u}$; the same is true for higher $\tilde{\rho}$.[32] Thus $\hat{p}$ is increasing in the degree of correlation. A borrower of risk-type p will therefore choose to borrow only if correlation is strong enough to raise $\hat{p}$ above p. That is, *there is a level of correlation above which the agent will choose to borrow and below which he will choose not to borrow.*[33]

This result comes from the property of the model that a positive payoff only occurs when the borrower is successful. Correlation shifts the weight in this state of the world toward the sub-state where the borrower's partner is successful, away from the sub-state where the borrower's partner fails. Thus it raises the payoff to joint liability borrowing and draws in more borrowers.

[31] This is true for interior $\hat{p}$.

[32] In other words, total differentiation of equations 7 and 8, respectively, shows that $d\hat{p}/d\epsilon > 0$ and $d\hat{p}/d\tilde{\rho} > 0$.

[33] As with the other prediction of Ghatak (1999), this one holds for a non-corner solution. There may be some risk-types that borrow no matter what the degree of correlation and some that do not borrow regardless of the level of correlation.

3 Empirical Results

3.1 Empirical Strategy

The predictions of Holmström and Milgrom (1990) and Prescott and Townsend (2002) can be tested by relating the likelihood of the cooperative, group regime – vis a vis the relative performance, individual regime – to correlation and wealth distribution variables. In particular, Holmström and Milgrom (1990) predict the group regime to be less likely the higher is the correlation between agents' output. Prescott and Townsend (2002) predict the group regime to be more likely the more disparate the wealth levels within the potential group, and the more extreme the wealth level of the potential group relative to the lender. In other words, there is a U-shaped relationship of group wealth with the likelihood of the group regime.

To measure existence of the cooperative and relative performance regimes, we focus on borrowing contracts. The rural credit market is a likely candidate for the existence of moral hazard, due to lack of collateral of many poor borrowers and unobservability of effort. There is also significant risk to be borne in these predominantly agricultural settings. Thus the tradeoff between group and individual contracts analyzed in theory is likely to be present in our data.

In the theory, the key difference between the individual and group regimes is whether the agents can collude and enforce internal agreements. In practice this ability seems hard to measure, though it may be proxied well by variables measuring sharing and cooperative activities in other spheres of life (see Ahlin and Townsend 2002). The point of the theory, however, is that *the contract itself* may build in encouragement or facilitation of such co-operation, or leave it conspicuously absent. In other words, the theory relies on the lender and/or borrowers being able to turn on or off the borrowers' ability to side-contract. It seems natural to interpret the formation of groups as such a facilitation of side-contracting. Groups have meetings together, are under a designated internal leader, and engage in significant internal monitoring. In our data, the average borrowing group (from the BAAC) has 70% of its members communicating with the group leader at least twice per week and 35% of its members at least six times per week.

Individual loans, on the other hand, can be thought of as the relative performance regime, though of course in reality the principal cannot enforce complete lack of cooperation between agents. He can, however, abstain from encouraging it. Further, the primary lender in our data, the BAAC, does seem to base its stringency toward one borrower in part on the performance of other borrowers, as discussed in the introduction. It is at least reasonable then to associate the relative performance regime in the theory with individual loans in the data.

Tying these two regimes to individual and group loans, respectively, we face the difficulty of drawing the boundary lines of the group empirically. In the theory, contracts always involve two agents. In practice, joint liability groups (of the BAAC) typically have between five and fifteen members, a small minority having between fifteen and thirty. Further, we do not observe in our data the group with whom each borrowing household is matched. It is even more difficult to determine the set of borrowers with whom an individual borrower is being compared in a relative performance regime; certainly there is no reason to expect that the comparison group contains just one other borrower, as in the theory. Our simplifying assumption in both cases is that the *village* forms the boundary from which the group can be drawn and among which comparisons are made for individual borrowers. The village is the smallest unit into which our data classify each borrower. It makes sense as a boundary for explicit borrowing groups: in our data, over 65% of such groups (of the BAAC) have all their membership from a single village, and over 90% have more than half. Since correlation is likely higher within a village than across villages, it is probably also the best choice of a relative performance comparison group.[34]

[34]One might argue under logic related to Holmström and Milgrom (1990) that forming cooperative groups out of borrowers from different villages would be optimal, since this would minimize correlation. However, the ability of such borrowers to monitor and enforce internal agreements would presumably be low, due to geographic separation and lack of contact. Thus a group contract between borrowers from different villages

To summarize, the data we use are household-level observations that show whether the household is borrowing and under which regime, but do not allow us to determine explicitly group membership (whether the group is a comparison group under relative performance, or an explicit borrowing group under the cooperative regime). To measure[35] wealth dispersion within the group, then, we use a measure of distance between the borrowing household's wealth and the village mean wealth. The assumption is that the further the household is away from the village mean, the higher the likelihood that it is in a group with asymmetric wealth.[36] We measure correlation within the group by the degree of clustering of superlative years (for household income) of this household with other households in the village. Again, the assumption is that the more this household's good and bad years coincide with others in the village, the more likely the case that it is in a group with high correlation. Finally, to proxy the wealth of the group, we use different combinations of the wealth of the household itself and the village average wealth of borrowing households. Each of these two components has its advantage. The wealth of the household itself is certainly a component of group wealth, since we know for certain this household belongs to the borrowing group in question. However, the wealth of the remainder of the group is a more significant component of total group wealth, since groups contain multiple households; but the village average wealth of borrowers measures wealth of the remainder of the group with error.

Ultimately, we use these measures of group wealth, wealth dispersion, and correlation, all at the household level, to explain whether the household has a group loan. Since Holmström and Milgrom (1990) and Prescott and Townsend (2002) do not address agents involved in both kinds of regimes or in neither, we exclude from the regression households that either have both group and individual loans or that have no loan. The results are detailed in section 3.3.

Ghatak (1999) predicts that risk-type and correlation are key determinants of whether a

may be ruled out as highly costly a priori.

[35]Variable construction is discussed in more detail in section 3.2.

[36]We provide evidence that groups are not segregating by wealth levels in section 3.3. The evidence involves decomposition of land inequality into between-group and within-group components.

household will borrow. For correlation, we use the same household-level measure described in the preceding paragraphs, reflecting the coincidence of the household's good and bad years with other households in the village. For risk-type, our measure applies the theory directly to predictions the household makes about its future income. This measure is detailed in section 3.2. These predictions are derived holding other characteristics, such as outside, non-borrowing economic opportunities, expected income from borrowing, and interest rate r, constant. We therefore control for expected income using the household's assessment of next year's income, and for economic opportunities using measures of wealth, occupation, landholdings, and occupation. Controlling for r is of less concern in our data, since BAAC national policy at the time of the survey dictated charging a fixed interest rate of 9% on loans under sixty thousand Thai baht; the group loan ceiling was below this.

The key decision, then, is what to do with households having individual loans, since Ghatak (1999) does not leave room for uncollateralized households taking individual loans.[37] One option is to exclude them, just as households that did not fit into Holmström and Milgrom (1990) and Prescott and Townsend (2002) were excluded in the related empirical specifications. Another option is to include them and group them with households having group loans. The rationale for this would be that the key to the adverse selection result is not the form of liability, individual or joint, but rather whether liability is limited.[38] Thus if the individual loans are limited liability, they too should exhibit adverse selection. In fact it is likely they are limited liability to some degree; the collateral goes unseized not infrequently and provisions are sometimes made. However, even in this case, it may be that these loans are significantly closer to full liability than are joint liability loans. This would argue for their inclusion with non-borrowing households in the regression, since full liability makes risk-type unimportant. Further, the correlation result relies critically on both joint

[37]Ghatak (1999) shows that joint liability contracts always produce more social surplus than individual liability contracts, though they may not be Pareto optimal in that they can hurt the risky borrowers in equilibrium. These results are all in the context of limited liability. Thus Ghatak (1999) does not rule out individual loans that are fully collateralized.

[38]Stiglitz and Weiss (1981) the same adverse selection result in the context of *individual*, limited liability.

and limited liability within groups, so this latter specification is clearly preferred on those grounds. Fortunately, the key results are insensitive to the differences in these specifications.

3.2 Description of Variables

Our data are taken from a detailed survey of households in Thailand conducted in 1997, the Townsend Thai data. The survey covers two contrasting regions in Thailand. The central region is relatively close to Bangkok and enjoys a degree of industrialization as well as fertile land for farming. The northeast region is poorer and semi-arid. There is significant wealth variation both within and across regions. In all 192 villages were surveyed. In virtually all of these villages, fifteen households were surveyed, giving a total sample size of 2875. Table 1 gives a summary of all dependent and independent variables used, reporting means for all but squared terms and standard deviations for all but squared terms and dummy variables.

[Table 1 here]

The dependent variables that we use are dummies reflecting whether the household has taken out a group-guaranteed loan from a lending institution in the past year. There are two versions of this variable. The first, BAGPLOAN, restricts attention to group-guaranteed loans from the BAAC. As noted, this government institution is the primary institutional lender in rural Thailand: for example, 64% of institutional loans in our sample are from the BAAC. The BAAC offers both individual loans, which must be guaranteed by some form of collateral, usually land, and joint liability loans. To receive the latter, one must form or join an official BAAC-registered borrowing group and enter into a joint liability arrangement. BAGPLOAN equals one if the household has had outstanding a loan from the BAAC in the past year and lists the collateral for this loan as either none, a single guarantor, or multiple guarantors. About 23% of the household sample has such a loan.

The second version of the dependent variable is GRUPLOAN, which incorporates group-guaranteed loans from the BAAC *and other institutions*. These others are typically smaller institutions such as agricultural cooperatives and often village-based ones such as production credit groups (PCGs), but they also include commercial banks.[39] Using this broader definition increases the proportion of the sample that qualifies as having a group-guaranteed loan to about 30%.[40] However, the institutions incorporated are diverse in size and practice, which makes isolating contracts that are clearly group contracts more imprecise. We report specifications using both BAGPLOAN and GRUPLOAN.

Two analogous variables measure whether the household has an individual loan contract from a lending institution. BAIDLOAN and INDLOAN correspond to BAGPLOAN and GRUPLOAN, respectively, in the lenders they cover. The criterion for a loan counting as an individual loan are that the collateral used was land, savings, current or future crops, and other collateral such as house or boat.[41] BAIDLOAN is positive for about 13% of the population, INDLOAN for about 22%. Neither of these will be used directly in regressions, but will at times be used to limit the sample to only those households having secured either an individual or a group loan.

Key independent variables from Prescott and Townsend (2002) include WEALTH and its square, WEALTHSQ, measured in millions of 1997 Thai baht and millions of baht-squared, respectively. These count up all types of household wealth, including land, household assets, business assets, agricultural assets, and net financial assets, including savings less debt. The part of this wealth attributable to landholdings to which the household has a legal title is called TITLE, and its square is TITLESQ. These measures are helpful as proxies for the value of available collateral, since titled land is the predominate form of collateral, at least

[39] Specifically, they include PCGs, commercial banks, agricultural cooperatives, village funds, rice banks, as well as any other cooperatives, institutions, or programs that also lend funds.

[40] For all but the BAAC, we tighten the criterion for a loan from a given institution to qualify as group-guaranteed. It must list as collateral a single guarantor or multiple guarantors; listing the loan as requiring no collateral leads us to count the loan as indeterminate. The reason for our distinguishing between lenders in this way is our better knowledge of BAAC policy. They do not give loans to anyone without collateral unless they belong to a borrowing group, so those who list no collateral requirement almost certainly are referring to a group loan.

[41] A few loans for which the collateral was a relative's land or something undecipherable were counted as indeterminate. Loans effectively collateralized through some contract interlinkage, such as by future wages where the loan is from an employer, were categorized as neither group nor individual.

with the BAAC. Recall that collateral plays no role in the theory.

To measure wealth dispersion WLTHDSPR, which corresponds to the λ_i's of Prescott and Townsend (2002), we use the following function of household wealth and village average wealth, call it $\overline{WEALTH}$:

$$WLTHDSPR = |1 - \frac{WEALTH}{\overline{WEALTH}}|^{1/2}.$$

This is similar to a simple distance function, $|\overline{WEALTH} - WEALTH|$. The differences are that we divide by $\overline{WEALTH}$, which makes it a scale-free measure,[42] and we take a square root which dampens the effect of WEALTH's long right tail. To check robustness relative to functional form, we will also use a nonparametric regression technique on $\frac{WEALTH}{\overline{WEALTH}}$ itself.

Technological correlation ρ is of course difficult to measure. Our proxy focuses on correlation of good and bad years of a household with the rest of their village. Each household was asked which of the past five years was best and which was worst, in terms of household income. (The wording of the question does seem to evoke income gross of any informal risk sharing, thus increasing our confidence that it is capturing technological variation in income.) SAMEBEST, SAMEWRST, and SAMEITHR, reflect the percent of the other village households' responses that coincide with this household's responses regarding the best year, worst year, and about either year, respectively.[43] We use each of these measures. The theories we test do not give clear direction on which should be more helpful.

The final key variable to measure is the risk-type of the borrower. We do so using subjective income assessments, taking Ghatak (1999) quite literally. Specifically, each household was 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

[42] In other words, WLTHDSPR does not change if the wealth of everyone in the village goes up or down by the same multiplicative factor. Since power utility is used in the computational results, and since the ratio of marginal utilities under power is scale-free, WLTHDSPR also should be scale-free.

[43] For example, if a household lists the same best year as 4 of 12 other village households and the same worst year as 7 of 14 other village respondents, then $SAMEBEST = 4/12$, $SAMEWRST = 7/14$, and $SAMEITHR = 11/26$. (Some households answer one question but not the other.)

be (Ex). We assume the income distribution is binomial over the high and low states, as in Ghatak (1999). The probability of success, PROBHI, is then calculated to be

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

using the fact that $PROBHI * Hi + (1 - PROBHI) * Lo = Ex$.

Several control variables are included, suggested by theory, institutional details, or practicality. We include next year's expected income (Ex), called EXINCOME, in our regressions. This is necessary because Ghatak (1999) assumes expected returns to be the same across all borrowers, while it of course varies in our data. We also include a comprehensive measure of income over the past twelve months, INCOME.[44]

The household head's gender and years of education are controlled for in MALEH (male coded as one) and EDYEARH, respectively. We also include a dummy variable, NRTHEAST, equaling one for the 50% of households in the poorer Northeast region. These are also important controls because Ghatak (1999) assumes the same outside (non-borrowing) option $\underline{u}$ for all borrowers, while outside options clearly vary in our data.

Finally, we control for eligibility and desire for a loan. Households were asked whether or not they engage in agricultural activity, and if so, whether or not they would like to expand their operations. From these questions we derive two dummy variables, AGRYES and AGRNO. AGRYES (AGRNO) equals one for the 46% (33%) of households that engage in agricultural activity and would (would not) like to expand their activity. The remaining 21% of households for whom neither equals one are those who do not engage in agricultural activity. It is crucial to control for occupation, since the BAAC and several other institutional lenders targeted agricultural activities exclusively at the time of the survey. One further proxy for desirability of a loan is the dummy variable OWNSBSNS, which equals one for the

[44] In some footnoted specifications we also control for the variance of next year's income, using the subjective income assessments described in the preceding paragraphs.

21% of households that own a business.[45]

3.3 Logit Results

We use simple logits since the dependent variables, whether the household has a group loan, are dummies. We report twelve specifications, varying across the three measures of correlation, the two versions of the dependent variable, and the sample chosen, whether restricted or unrestricted. More specifically, the restricted sample includes only those households that have either a group-guaranteed BAAC (or institutional) loan, or an individual BAAC (or institutional) loan. That is, it includes only households for whom $BAGPLOAN + BAIDLOAN = 1$ (or $GRUPLOAN + INDLOAN = 1$). The rationale for examining this restricted sample is that Holmström and Milgrom (1990) and Prescott and Townsend (2002) focus on the choice *between* cooperative and relative performance regimes given that one must be chosen. Thus the restricted sample regressions, though smaller in amount of data, are arguably more appropriate tests of Holmström and Milgrom (1990) and Prescott and Townsend (2002).

The logit results are reported in tables II and III, for the restricted and entire samples, respectively. It should also be noted that in each regression, we exclude households who have not answered the questions about both best and worst years or whose village does not contain at least four responses to each of these questions.[46]

[Table 2 here]

Table 2 is the preferred specification for testing Holmström and Milgrom (1990) and

[45] It is purely coincidental that the same percent of households own businesses and do not engage in agriculture: there are households of whom both are true, and of whom neither are true.

[46] SAMEITHR is valid in about 66% of households, after dropping households from villages where less than eight total responses to the two questions exist. SAMEBEST is valid in about 71% of households and SAMEWRST in about 80% of households, after dropping households from villages where less than four total responses to the respective question exist. For the regressions we report, we restrict each specification to the 66% of households that have valid information for both SAMEBEST and SAMEWRST. This ensures the sample is not changing as we use different measures of correlation. We have run the regressions using all the data available in each specification, and the results become somewhat different, but not enough to change our qualitative conclusions.

Prescott and Townsend (2002). In all specifications, wealth levels exhibit a U-shaped relationship with the group contract. These coefficients are all significant at the 5% significance level, except that under BCGPLOAN the coefficient on WEALTHSQ drops to the 10% significance level and under GRUPLOAN the coefficient on WEALTH drops to the 10% level. These baseline specifications give strong evidence for the wealth *level* predictions of Prescott and Townsend (2002).

The trough of the U-shaped relationship is around 6.4 million baht for BAGPLOAN and 8.4 million baht for GRUPLOAN. This leaves 5-6% of the sample on the upward sloping segment of the relationship. The U-shape remains significant under exclusion of households with wealth greater than 20 million (1.4% of the data).[47] We check a cubic relationship for robustness and find a very similar picture: coefficients on WEALTH and WEALTHSQ of the same sign as in Table 2 and significant at 1%, and the coefficient on WEALTHCB negative and significant at 5% or 10% depending on specification. Further, the relationship is declining to a trough between 5 and 6 million baht, then increasing to a peak at just over 21 million, and declining from there (for just over 1% of the data). Thus the U-shape over virtually all the data appears robust.

To allow the relationship with wealth to take a more flexible form, we also employ an alternative, partially linear model. This model is simpler in discarding the logistic transform of the logit model, but allows one or more variables to follow any smooth functional form. Specifically, we assume a partially linear model in which the decision to borrow D is some smooth function of WEALTH added to a linear function of the remaining variables, X_{-j}:

$$D = \beta_{-j}' X_{-j} + k(WEALTH) + \epsilon,$$

where β_{-j} is a vector of coefficients and k is a continuous function. Robinson (1988) and Yatchew (1998), among others, have proposed estimators for this model. We use Yatchew's

[47]It does disappear as this cutoff decreases to 10 million (3.4% of the data); but this is not surprising since we are ignoring a significant fraction of households above the turning point.

technique here. The coefficients β_{-j} are estimated first by ordering the observations according to WEALTH, differencing across nearby observations (we use optimal fifth-order differencing, described in Yatchew 1998), and regressing (the differenced) X_{-j} on (the differenced) D. This produces an estimate $\hat{\beta}_{-j}$. We then form residuals $D - \hat{\beta}'_{-j} X_{-j}$. Finally, we carry out a non-parametric regression[48] of these residuals (which are clearly not restricted to equal 0 or 1) on WEALTH. Standard errors at 90% confidence are calculated using the bootstrap method, that is, recreating 1000 samples from the original sample by sampling with replacement, and calculating an estimate of function k (evaluated at each unique value w of the *original* sample) using the same technique on each sample. For each value w, the confidence interval is the 51st and 950th smallest value of $k(w)$ from these samples. Since our main concern is with the shape (slope) of the functions, we normalize the residuals in each estimate to have mean zero.[49]

The results using both dependent variables (and SAMEITHR as the measure of correlation) are pictured in Figure 1, which depicts the relationships for wealth below twenty million baht.[50] The U-shape is evident in the function estimates over almost all the data, though the upward-sloping part is less tightly estimated due to sparser data. At a minimum, it does appear that the logit U-shape results are not being driven by imposed quadratic or cubic curvature.[51]

[48]We use a local linear regression similar to Lowess (see for example Cleveland 1979 and Fan 1992). For each unique value of WEALTH, w, we calculate a fitted value of the dependent variable (the residuals) from a weighted least squares regression on a "nearby" subset of the total sample. Thus the choices are weights for the regression and a bandwidth which determines the subsample. The bandwidth $h(w)$ is set for each unique value w to ensure inclusion of the 80% of the data whose values w_i are closest to w. (If there are clusters of observations at the boundary of the bandwidth with the same value for the independent variable, all are included. Thus potentially more than 80% of the sample is used.) The weighting function is the tri-cube weighting function:

$$w_i = (1 - (\frac{|w - w_i|}{1.0001\ h(w)})^3)^3.$$

This function places more weight on observations w_i located more closely to w.

[49]Note that the Yatchew procedure identifies the function up to a constant. De-meaning the residuals is then essentially a normalization of each bootstrapped estimate with respect to the constants. Without this normalization, bootstrap error bands can get large merely because the constants are varying.

[50]That is, the graph (not the actual procedure) leaves out the part of the curve from wealth of twenty million baht to one hundred million baht, to focus on the key part of the relationship. There are only five to ten data points about twenty million, so the error bands are quite wide above there.

[51]To provide a comparison of the two models, logit and partially linear, we compare the implied marginal

[Figure 1 here]

There may be other explanations for such a relationship. Chief among these could be a story revolving around collateralizable wealth: since individual loans require collateral, we would expect them to be more prevalent at higher wealth levels, at the expense of group loans. This story, however, does not immediately make clear why the relationship should turn up again to deliver a U-shape. Further, we separate out and control for the part of wealth that is most commonly used and accepted as collateral, TITLE and TITLESQ, and still find the U-shaped relationship between total wealth and being in a group contract. Thus the collateral story does not seem to be driving the results.[52]

One might also think that poor households borrow in groups, moderately wealthy households borrow as individuals, and the most wealthy households take out both kinds of loans since their demand for credit is higher. This would produce a U-shape relationship between wealth and having a group loan. But note that this explanation is ruled out since the sample includes households with one or the other kind of loan, but not both.

Of course, one caveat regarding these results is that we do not have group-level wealth, only wealth of households. Another approach would be to proxy group wealth for a given household as the average between that household's wealth and the average household wealth level in the village (or among villagers who borrow). The inverted-U shape loses significance in the majority of specifications under this approach. This need not be due to a deficiency of the model, but rather an imperfect proxy for group wealth. We continue to view the evidence presented in Table 2 as suggestive evidence for the model.

effects for all of the variables except wealth. In the partially linear model, the marginal effect for a variable X_k is just the estimated coefficient $\hat{\beta}_k$. For the logit model, it is the product of the estimated coefficient and the logistic density of the index $\beta' X$, conventionally evaluated at the means of the variables. For all variables but two (TITLE and TITLESQ) the signs of the marginal effects are the same across the two models. The magnitudes are generally of the same order of magnitude also.

[52]In fact, excluding TITLE and TITLESQ from the regression gives a much more linearly negative relationship between wealth and borrowing (i.e. the positive coefficients on WEALTHSQ drop in significance). Evidently, in that specification, WEALTH measures both collateral and wealth in the sense of Prescott and Townsend (2002), and the more uniformly negative (or at least concave) effect of collateral is keeping the relationship with WEALTH from turning up to form a U.

A strong and statistically significant result for Prescott and Townsend (2002) is that group borrowing is more likely the higher is WLTHDSPR, that is, for those with wealth further away from the village average. Evidently the group contract is used more heavily for borrowers with asymmetric wealth. Thus the wealth inequality predictions of Prescott and Townsend (2002) are also borne out. This result is even more robust than that of the U-shape: in every specification we report in the preceding paragraphs, it remains significant.

We also estimate the relationship using the partially linear model with the Yatchew estimator described above. We do so using the simple measure $\frac{WEALTH}{\overline{WEALTH}}$, that is household wealth relative to the village mean of household wealth. This allows us to remain agnostic about the functional form of the wealth dispersion variable. The results using both dependent variables (and SAMEITHR as the measure of correlation) are pictured in Figure 2. They are strikingly close to the theoretical prediction in that the chance of having a group loan is minimized at the point of symmetry, that is when relative wealth equals one. (For reference, vertical hash lines are drawn through the curves where relative wealth equals one.) The probability of having a group loan increases the further from one is relative wealth in either direction. The effect is stronger on the margin for low relative wealth, but clearly seems present for high relative wealth as well, at least in the point estimates.

[Figure 2 here]

One might worry that the U-shape and the wealth dispersion are actually the same result: high- and low-wealth borrowers are more likely to take group loans. But recall that the U-shape relationship is the same for the whole sample, while the dispersion relationship is *relative* to village average wealth, and thus different in each village. Thus two poor households of equal wealth have the same boost to their chance of borrowing in a group due to the U-shape wealth level result, but a poor household from the richer village has an even higher boost due to the wealth dispersion result (since it is likely to be in a more asymmetric group). Of course, this interpretation relies on significant variation in village average wealth.

This we have: the coefficient of variation of village average wealth is around two.[53]

One might also worry that the empirical result on relative wealth need not imply high wealth dispersion *within* groups. For example, it could be that relatively high-wealth villagers form their own group and quite low-wealth villagers their own group. If so, groups could be homogeneous in wealth (contrary to Prescott and Townsend (2002)) but households further from the village mean would still be more likely to borrow. We cannot directly check for wealth dispersion within groups from the household data, since we do not know who is paired with whom. However, also part of the Townsend Thai data is a different survey, of BAAC borrowing groups rather than households. It contains information on each group member's landholdings, which we can use to proxy wealth. Further, of the 262 groups interviewed, 200 are one of two groups interviewed from their village. Thus we have 100 villages with two groups each, and we can decompose land inequality of village group-borrowers into between-group and within-group inequality. Results from this exercise support Prescott and Townsend (2002). We rank villages based on the percent of such inequality attributable to between-group inequality. The median village has about 25-27% of inequality attributable to between-group inequality using the gini or coefficient of variation, and only 10-12% using Theil's second measure or Foster-Shneyerov's $q = 2$ measure.[54] Even the most extreme village has only 47-51% of inequality explained by between-group inequality using the gini or coefficient of variation, and 52-54% using Theil's second measure or Foster-Shneyerov's $q = 2$

[53] One might attribute the wealth dispersion result to rich people joining with their servants or poor neighbors merely to increase their loan size (by ex post appropriating the loan their servants receive). This seems unlikely. Most group-guaranteed loans are below 50,000 baht, while the mean individual loan is 87,000 and about 10% are over 200,000. Thus applying individually seems to be a better way of obtaining a large loan. Further, a group loan is only given to an individual who can show that he is undertaking some agricultural project and has access to cultivated land (whether rented or owned), so above-mentioned scheme would be difficult to implement.

[54] See Foster and Shneyerov (2000), who characterize the class of inequality measures I_q for which total inequality of a sample always exactly equals inequality of the between-population plus inequality of the within-population. (The between-population sets the income of each group to the group mean; the within-population scales the incomes within each group by a group-specific scalar so that all groups end with the same mean.) Theil's second measure (which equals I_1) and I_2 both satisfy this property. The gini and coefficient of variation do not. For the latter two, we calculate the percent contribution of between-group inequality as $I(B)/[I(B) + I(W)]$, where B and W are the between-population and the within-population, respectively.

measure. According to land, then, the groups do not appear to be sorting homogeneously; rather, the usual case is substantially more variation within groups than across groups.

With respect to technological correlation, none of the measures shows up significantly, except SAMEBEST at the 90% level in the specification with GRUPLOAN as dependent variable. However, it shows up positively, counter to what Holmström and Milgrom (1990) predict. Thus there is some evidence against, or at least no evidence for, the prediction of Holmström and Milgrom (1990) (and presumably also Prescott and Townsend (2002)) regarding correlation. It is plausible that the lack of result is due to a very noisy measure of correlation, but less so since we do get some results with the same variable when testing Ghatak (1999).[55]

None of the models incorporates the size of loan, desired or actual, though in theory this may be an important factor in determining the type of contract.[56] It is important in the data also, as the BAAC has a de jure loan limit of fifty thousand baht for group loans but can give much higher individual loans. We therefore add loan size into the above specifications: first without sample restrictions, second excluding all households with loans greater than one hundred thousand, and third excluding all households with loans greater than fifty thousand. The third sample allows comparison over the range of loans that can be observed under either contract. The basic results are that loan size is a negative and significant predictor of the group regime, at the 1% level always except under BAGPLOAN in the third sample, when it is at the 10% level. WLTHDSPR continues to be positive and significant at the 5% or 10% level. The U-shape on wealth continues to be seen in the coefficients on WEALTH and

[55] One might also wonder if this variable is endogenous, since the group contract may explicitly encourage internal risk-sharing and thus may induce high correlation within the group, at least for consumption and income net of risk-sharing. However, our measure makes this unlikely. It asks for the best or worst year in terms of a Thai word best translated as "income" or "earnings", and we believe this would generally be interpreted as income gross of risk-sharing transfers.

Further, we run correlations between two household indices for sharing rice, helping with free labor, and helping with money - plausible proxies for informal risk-sharing activity - with our three measures of household correlatedness. We thus compute six correlations, only two of which are significantly different from zero. Point estimates of the correlation are typically around .02 and always below .08. It appears safe to interpret our correlation measure as applying to income gross of informal risk-sharing transfers.

[56] For a modification of Ghatak (1999) that incorporates endogenous loan size, see Ahlin and Townsend (2002).

WEALTHSQ, but only in a minority of cases at conventional significance levels (most often the significance is 20-30%). Correlation remains an insignificant predictor. In short, the slightly weakened significance of the U-shape relationship with wealth is the only change in results when controlling for loan.

[Table 3 here]

The predictions of Ghatak (1999) are best tested in the full sample regressions, seen in Table 3. This is because the selection there is between a group contract and some unspecified outside option (see discussion in section 3.1 and below for robustness checks). Correlation shows up as significant when GRUPLOAN is the dependent variable. Then it shows up positively and significantly at the 15% level using SAMEBEST and SAMEWRST, and at the 5% level using SAMEITHR. This is some evidence that households highly correlated with the rest of the village are more likely to borrow with a group contract. It provides support for the result on correlation of our modification of Ghatak (1999), that higher correlation makes joint liability borrowing more attractive and thus draws in more borrowers.

The coefficient on PROBHI is always negative and significant. This is striking evidence for adverse selection in the joint liability credit market, as Ghatak (1999) predicts. The more likely a borrower is to succeed, the less likely he is to borrow in a joint liability group. The result is all the more remarkable because of the unique, literal interpretation we use to measure risk-type.[57]

Note that we have not identified adverse selection vis a vis moral hazard. A moral hazard interpretation of the data could be that risk was endogenous to having a group loan; that is, having the loan was causing the borrower to operate with more risk, rather than vice versa. In order to identify a pure adverse selection effect, we make use of the fact that some of the loans in our data have already been repaid (usually within the last few months).

[57] If the environment is such that every household is either endowed with a risky project or not endowed with any project, then our result could be spurious: both taking a loan and bearing risk could be the result of being endowed with a project. Our results do not seem to be driven by such a phenomenon, since we control for ownership of a business, agricultural occupation, and the desire to expand one's agricultural operation.

Under adverse selection, those who took a loan but have already repaid it would still be forecasting low probabilities of success, since they are inherently more risky. Under moral hazard, the incentives for risk-taking from having a limited liability loan vanish when the loan is repaid, so those who have already repaid their loan should look no different from those who never had a loan. We run the logits again after eliminating all households with a group loan that has not yet been repaid. Unfortunately, this leaves less than 20% of the households who had group loans. In all specifications, the coefficient actually increases noticeably in magnitude (i.e. becomes more negative), but so do the standard errors. In the three specifications using BAGPLOAN, PROBHI remains significant, twice at the 10% level and once at the 15% level. In the specifications using GRUPLOAN, the estimates drop just beyond conventional significance levels (15-30%). This may be expected from the significant drop in sample size. We interpret these results as suggestive that adverse selection *specifically* is occurring in this credit market.[58]

The previous test provides evidence for adverse selection, but it does not rule out that moral hazard is also occurring. One can approach the moral hazard question in a similar way.[59] Among those who took a loan, moral hazard suggests that those who have already repaid should look less risky than those with a loan still outstanding. Again, the reason is that the incentives for risk-taking disappear once the loan is repaid. We therefore limit the sample to those who did take out a loan (those for whom BAGPLOAN or GRUPLOAN, depending on the specification, equals one) and predict PROBHI using the same covariates as above and a dummy for whether the loan is already repaid. Again, the skewed incentives leading to moral hazard should not be an issue for those whose loan is repaid, and thus the coefficient on the dummy should be positive. The coefficient is negative, though insignificant, in all six specifications, using both OLS and tobit. The data here thus provide no evidence

[58]It is unlikely that our lender observes the information contained in PROBHI or something similar. Regardless, the lender was not using information to tailor the contract (except for loan size), but was offering a standard contract to whoever wanted it.

[59]We thank Eric Bond for suggesting the following test.

that moral hazard is occurring.[60]

One might argue that when testing adverse selection, individual loans should not be included in the outside option ($D = 0$, where $D = 1$ represents having a loan) but as part of those having a loan ($D = 1$). It is only limited liability that is needed for adverse selection, irrespective of whether liability is individual or joint. If the individual loans, though technically collateralized, are in fact limited liability to some degree, they should also exhibit adverse selection. Accordingly, we rerun the specifications in Table 3 coding households that have individual loans as $D = 1$. We also skirt the issue by running them again excluding all households that have an individual loan. The results in both cases change little,[61] or if anything become sharper. These results may suggest that the individual loans exhibit limited liability, and thus are subject to adverse selection, but both to a lesser degree than group loans.[62] Thus, whether they are included with group loans or the outside option, we expect the adverse selection result.

Finally, one might want to vary the conditioning sets more faithfully by model. To this end, we run the same logits but for Holmström and Milgrom (1990) omitting the income variables, PROBHI, and the wealth variables except TITLE and TITLESQ; for Prescott and Townsend (2002) omitting the correlation variables, the income variables, and PROBHI; and for Ghatak (1999) omitting the wealth variables except TITLE and TITLESQ. We leave the remaining controls in for each specification. The titled land variables are always included, as

[60]Note that moral hazard here in the sense of Stiglitz and Weiss (1981) and Ghatak (1999) refers to effects on the riskiness, not mean, of the distribution. This test is therefore not directly related to the moral hazard of Holmström and Milgrom (1990) and Prescott and Townsend (2002).

[61]Specifically, the coefficient on PROBHI is negative and significant at 10% in every specification when individual loans are excluded. It is negative and significant at 5% in the first and third specifications and 10% in the remaining four when individual loans are included with group loans.

Oddly enough, SAMEWRST becomes an even stronger predictor of borrowing when individual loans are included with group loans. The result we show in section 2.3 cannot explain this since it relies on joint liability specifically, not just limited liability. One interpretation we do not explore here is that if all loans are insurance contracts bought at some price, those with the least ability to ensure themselves (i.e. high correlation of bad years) would be the ones buying them. (The frameworks of Holmström and Milgrom (1990) and Prescott and Townsend (2002) hint at this, though they restrict analysis to the choice of contract *given* a contract is bought.)

[62]If this were true, we would expect the magnitude of the coefficients on PROBHI to be highest when individual loans are excluded altogether, since the comparison would be starkest in that case. Indeed, this is what we find.

controls for collateral. The income variables, INCOME and EXINCOME are excluded from Holmström and Milgrom (1990) and Prescott and Townsend (2002) since the (expected) income level is endogenous in the models. The results of each of these are not appreciably different from what is reported. To test Holmström and Milgrom (1990) more faithfully one might also want to include the variance of technology, as this is an exogenous parameter held constant in the comparison performed in Holmström and Milgrom (1990). We experimented with the inclusion of a measure of the variance,[63] and the results did not change. Further, the variance was not a significant predictor (nor have we explored what the model would predict about its implications). Thus the results reported in Tables II and III are highly similar to those obtained from more faithful specifications.

4 Conclusion

The results on wealth and contract selection are interesting in several respects. Indirectly, they give plausibility to the interpretation of group lending as a way for the lender to promote side-contracting between borrowers. This has not always been the primary view of group lending, but it does appear to have some empirical validity.

More directly, these results suggest that wealth distribution, as emphasized by Prescott and Townsend (2002), can be a key determinant of the optimal contract. Specifically here, a wealth level further away from the village average, or further away from the population cutoff of 6-8 million baht, makes choice of a group loan over an individual loan more likely. From a theoretical point of view, moving away from a transferable utility framework seems warranted. From a policy point of view, one might be led to wonder whether group lending accomplishes its goal of reducing inequality, or whether, at least locally, it thrives on inequality and perhaps perpetuates it through unequal intra-group allocation of consumption and labor.

[63] The variance was measured using the subjective income assessments described in section 3.2. In particular, again assuming the binomial distribution on income over Hi and Lo, and calculating the probability of success using Ex, one can also calculate the variance as $(Hi - Ex)(Ex - Lo)$.

The absence of corroboration in the data for the correlation result of Holmström and Milgrom (1990) is puzzling. There is a range of possible explanations. First, the restricted sample size used to test Holmström and Milgrom (1990) is less than half of the full sample size used to test Ghatak (1999). Second, the result is proved under very specific functional form assumptions, which may not be satisfied. On the other hand, the intuition appears likely to be valid more generally. Third, it may be that correlation is unobservable or unused by the lender. But as discussed in section 2.3, one might still expect the patterns in the data that we look for even if only borrowers observe correlation. Finally, it is possible that the correlation measure is too noisy. This is less likely since the measure does fairly well in the Ghatak (1999) specification.

The results confirming Ghatak (1999) on adverse selection and correlation indirectly support the more common interpretation of group lending as joint liability. Though we do not view this interpretation as necessarily contradictory with the other, we know of no attempt to analyze their interaction. This attempt could lead to fruitful research.

At any rate, adverse selection of risky borrowers seems to be an issue not only in the ivory tower and formal insurance companies, but also in the rice patties. In terms of policy, at least partial screening by risk information that is available is an obvious first step in improving the extension of credit and the allocation of funds. Since the survey, the BAAC has in fact begun offering different interest rates depending on credit history. Still, there may remain room for more efficient screening by offering a wider array of contracts with different combinations of interest rate and joint liability (as in Ghatak (2000)). On the theoretical front, introducing loan size into such an adverse selection model could also prove fruitful.

Ahlin, C. and R.M. Townsend, 2002, Using Repayment Data to Test Across Theories of Joint Liability Lending, Vanderbilt University Department of Economics Working Paper.

Ausubel, L.M., 1999, Adverse Selection in the Credit Card Market, University of Maryland Department of Economics Working Paper.

Chiappori, P.-A. and B. Salanie, 2003, Testing Contract Theory: A Survey of Some Recent Work, in: M. Dewatripont, L.P. Hansen, and S. Turnovsky, eds., Advances in Economic Theory and Econometrics: Theory and Applications, Eighth World Congress, Volume I (Cambridge University Press, Cambridge UK) 115-149.

Cleveland, W.S., 1979, Robust locally weighted regression and smoothing scatterplots, Journal of the American Statistical Association 74, 829-836.

Fan, J., 1992, Design-Adaptive Nonparametric Regression, Journal of the American Statistical Association 87, 998-1004.

Foster, J.E. and A.A. Shneyerov, 2000, Path Independent Inequality Measures, Journal of Economic Theory 91, 199-222.

Ghatak, M., 1999, Group Lending, Local Information and Peer Selection, Journal of Development Economics 60, 27-50.

Ghatak, M., 2000, Screening by the Company You Keep: Joint Liability Lending and the Peer Selection Effect, The Economic Journal 110, 601-631.

Holmström, B. and P. Milgrom, 1990, Regulating Trade Among Agents, Journal of Institutional and Theoretical Economics 146, 85-105.

Mas-Colell, A., M.D. Whinston, and J.R. Green, 1995, Microeconomic Theory (Oxford University Press, New York).

Negishi, T., 1960, Welfare Economics and the Existence of an Equilibrium for a Competitive Economy, Metroeconomica 12, 3-8.

Prescott, E.S. and R.M. Townsend, 2002, Collective Organizations versus Relative Performance Contracts: Inequality, Risk Sharing, and Moral Hazard, Journal of Economic Theory 103, 282-310.

Rai, A.S. and T. Sjöström, 2004, Is Grameen Lending Efficient? Repayment Incentives and Insurance in Village Economies, Review of Economic Studies 71, 217-234.

Robinson, M.S., 2001, The Microfinance Revolution: Sustainable Finance for the Poor (World Bank, Washington, D.C.).

Robinson, P.M., 1988, Root-N-Consistent Semiparametric Regression, Econometrica 56, 931-954.

Stiglitz, J.E. and A. Weiss, 1981, Credit Rationing in Markets with Imperfect Information, American Economic Review 71, 393-410.

Townsend, R.M. and J. Yaron, 2001, The credit risk-contingency system of an Asian development bank, Federal Reserve Bank of Chicago Economic Perspectives XXV, 31-48.

Wilson, Robert, 1968, The Theory of Syndicates, Econometrica 36, 119-132.

Yatchew, A., 1998, Nonparametric Regression Techniques in Economics, Journal of Economic Literature 36, 669-721.

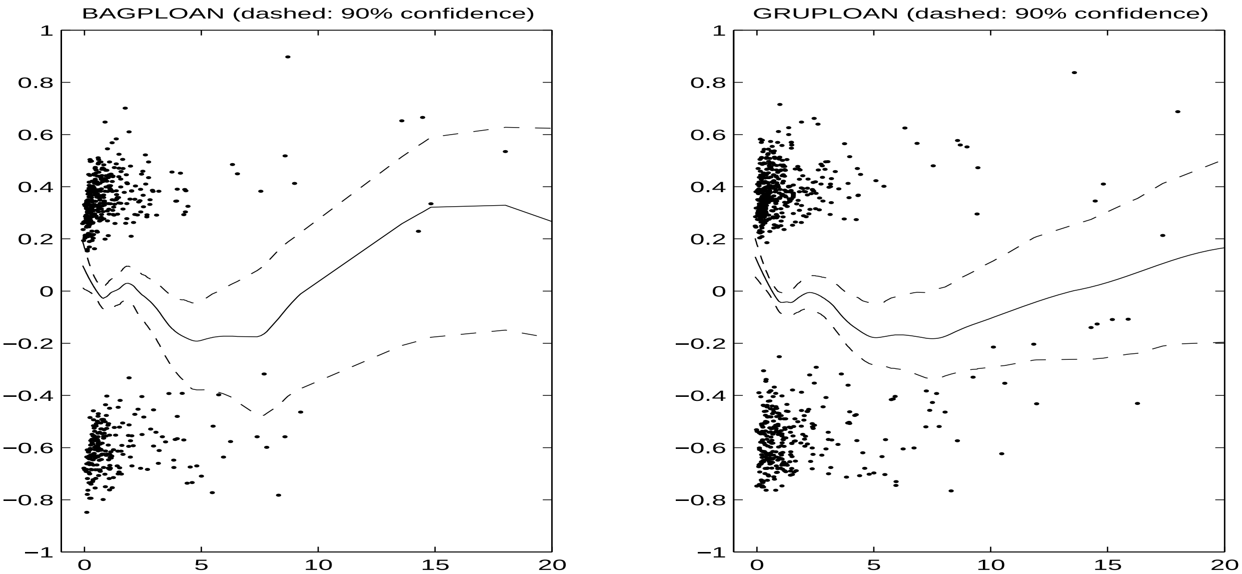

Figure 1: Wealth as a predictor of having a group loan.

Table 1 - Summary of Variables

Variable	Description	Mean	(σ)
	DEPENDENT:		
BAGPLOAN	Has the household obtained a BAAC group-guaranteed loan in the past year?	0.22	
GRUPLOAN	Has the household obtained an institutional group-guaranteed loan in the past year?	0.29	
BAIDLOAN	Has the household obtained a BAAC individual loan in the past year?	0.14	
INDLOAN	Has the household obtained an institutional individual loan in the past year?	0.23	
	INDEPENDENT:		
	Correlation:		
SAMEBEST	Percent of other village responses agreeing with this one on best year	0.30	(0.23)
SAMEWRST	Percent of other village responses agreeing with this one on worst year	0.28	(0.22)
SAMEITHR	Percent of other village responses agreeing with this one on superlative year	0.29	(0.17)
	Wealth:		
WEALTH	Wealth of the household, in million baht	1.8	(5.8)
WEALTHSQ	Wealth squared, in million baht-squared		
WLTHDSPR	Measure of distance from household wealth to village mean wealth	0.81	(0.39)
TITLE	Wealth attributable to titled land, in million baht	0.94	(5.0)
TITLESQ	Wealth attributable to titled land squared, in million baht-squared		
	Risk Type:		
PROBHI	Probability of realizing high income next year	0.57	(0.26)
	Control:		
INCOME	Gross household income from past year, in million 1997 Thai baht	0.15	(2.2)
EXINCOME	Expected household income for next year, in million baht	0.11	(0.30)
NRTHEAST	Dummy = 1 if the village is in the northeast region	0.50	
AGRYES	Does household engage in agriculture and want to expand operations?	0.46	
AGRNO	Does household engage in agriculture and not want to expand operations?	0.33	
OWNSBSNS	Does household own a business?	0.21	
LANDOWND	Average landholdings (in rai) of household	17.2	(30.4)
EDYEARH	Years of education attained by head of household	4.1	(2.6)
MALEH	Is head of household male?	0.77	

Table 2 - Restricted Samples

The samples are restricted only to those borrowers having either a group-guaranteed loan, or an individual loan.

Standard errors in parentheses; significance at 5, 10, and 15% denoted by ***, **, and *, respectively.

	BAGPLOAN			GRUPLOAN		
WEALTH	-.417 (.169)***	-.418 (.169)***	-.418 (.169)***	-.174 (.096)**	-.173 (.097)**	-.175 (.097)**
WEALTHSQ	3.24E-8 (1.73E-8)**	3.28E-8 (1.72E-8)**	3.26E-8 (1.73E-8)**	1.03E-8 (4.91E-9)***	1.03E-8 (4.98E-9)***	1.02E-8 (4.94E-9)***
WLTHDSPR	.696 (.253)***	.686 (.254)***	.695 (.254)***	.505 (.217)***	.500 (.217)***	.509 (.217)***
TITLE	.213 (.164)	.219 (.165)	.216 (.165)	-.053 (.090)	-.043 (.090)	-.046 (.090)
TITLESQ	-3.14E-8 (1.70E-8)**	-3.21E-8 (1.70E-8)**	-3.17E-8 (1.70E-8)**	-8.14E-9 (4.63E-9)**	-8.19E-9 (4.71E-9)**	-8.10E-9 (4.66E-9)**
PROBHI	.023 (.353)	.017 (.354)	.029 (.353)	-.139 (.308)	-.124 (.308)	-.123 (.308)
SAMEBEST	.176 (.401)			.549 (.350)*		
SAMEWRST		-.190 (.407)			-.059 (.367)	
SAMEITHR			.009 (.513)			.482 (.471)
INCOME	.312 (.877)	.335 (.882)	.310 (.876)	.075 (.738)	.082 (.735)	.037 (.727)
EXINCOME	-.732 (1.10)	-.771 (1.11)	-.734 (1.10)	-.090 (.943)	-.102 (.938)	-.042 (.928)
NRTHEAST	-.368 (.214)**	-.344 (.213)*	-.355 (.214)**	-.223 (.176)	-.191 (.176)	-.222 (.177)
AGRYES	-.538 (.522)	-.537 (.522)	-.530 (.522)	.003 (.354)	.040 (.353)	.021 (.353)
AGRNO	-.804 (.531)*	-.802 (.532)*	-.798 (.531)*	-.415 (.361)	-.387 (.360)	-.403 (.361)
OWNSBSNS	-.116 (.225)	-.125 (.226)	-.117 (.226)	-.512 (.189)***	-.521 (.189)***	-.508 (.189)***
LANDOWND	-7.28E-3 (4.12E-3)**	-7.21E-3 (4.12E-3)**	-7.25E-3 (4.11E-3)**	-6.21E-3 (3.45E-3)**	-6.22E-3 (3.45E-3)**	-6.27E-3 (3.44E-3)**
EDYEARH	.030 (.042)	.029 (.042)	.030 (.042)	.010 (.033)	.007 (.033)	.010 (.033)
MALEH	.437 (.243)**	.446 (.243)**	.445 (.243)**	.628 (.208)***	.644 (.208)***	.632 (.208)***
N	573	573	573	736	736	736

Table 3 -Full Samples

The full sample is used in each of these regressions.

Standard errors in parentheses; significance at 5, 10, and 15% denoted by ***, **, and *, respectively.

	BAGPLOAN			GRUPLOAN		
WEALTH	-.077 (.078)	-.077 (.078)	-.077 (.078)	.016 (.067)	.016 (.068)	.015 (.068)
WEALTHSQ	1.08E-9 (2.81E-9)	1.04E-9 (2.84E-9)	1.08E-9 (2.81E-9)	-1.37E-9 (2.85E-9)	-1.48E-9 (2.96E-9)	-1.43E-9 (2.91E-9)
WLTHDSPR	.119 (.176)	.130 (.176)	.123 (.176)	-.098 (.164)	-.078 (.165)	-.084 (.165)
TITLE	.017 (.110)	.014 (.110)	.014 (.110)	-.125 (.071)**	-.124 (.071)**	-.127 (.071)**
TITLESQ	-3.30E-9 (6.06E-9)	-3.09E-9 (6.06E-9)	-3.16E-9 (6.04E-9)	2.21E-9 (2.80E-9)	2.30E-9 (2.90E-9)	2.29E-9 (2.84E-9)
PROBHI	-.460 (.235)***	-.431 (.236)**	-.448 (.235)**	-.386 (.218)**	-.344 (.219)*	-.353 (.218)*
SAMEBEST	-.008 (.269)			.370 (.251)*		
SAMEWRST		.351 (.289)			.408 (.272)*	
SAMEITHR			.331 (.367)			.784 (.346)***
INCOME	.611 (.559)	.605 (.562)	.603 (.560)	.049 (.508)	.042 (.513)	.035 (.511)
EXINCOME	-.578 (.672)	-.565 (.674)	-.563 (.672)	.066 (.591)	.075 (.596)	.087 (.594)
NRTHEAST	.618 (.133)***	.603 (.133)***	.600 (.134)***	.317 (.122)***	.326 (.121)***	.301 (.122)***
AGRYES	1.03 (.244)***	1.03 (.244)***	1.02 (.244)***	1.16 (.214)***	1.17 (.213)***	1.15 (.214)***
AGRNO	.761 (.251)***	.753 (.251)***	.752 (.252)***	.737 (.221)***	.735 (.221)***	.725 (.221)***
OWNBSNS	.280 (.152)**	.294 (.153)**	.287 (.153)**	.079 (.141)	.093 (.142)	.093 (.142)
LANDOWND	4.21E-3 (2.68E-3)*	4.13E-3 (2.67E-3)*	4.19E-3 (2.67E-3)*	4.17E-3 (2.36E-3)**	4.01E-3 (2.36E-3)**	4.09E-3 (2.35E-3)**
EDYEARH	.037 (.024)*	.038 (.024)*	.038 (.024)*	.056 (.023)***	.057 (.023)***	.057 (.023)***
MALEH	.510 (.168)***	.508 (.168)***	.504 (.168)***	.468 (.152)***	.476 (.152)***	.467 (.152)***
N	1666	1666	1666	1602	1602	1602

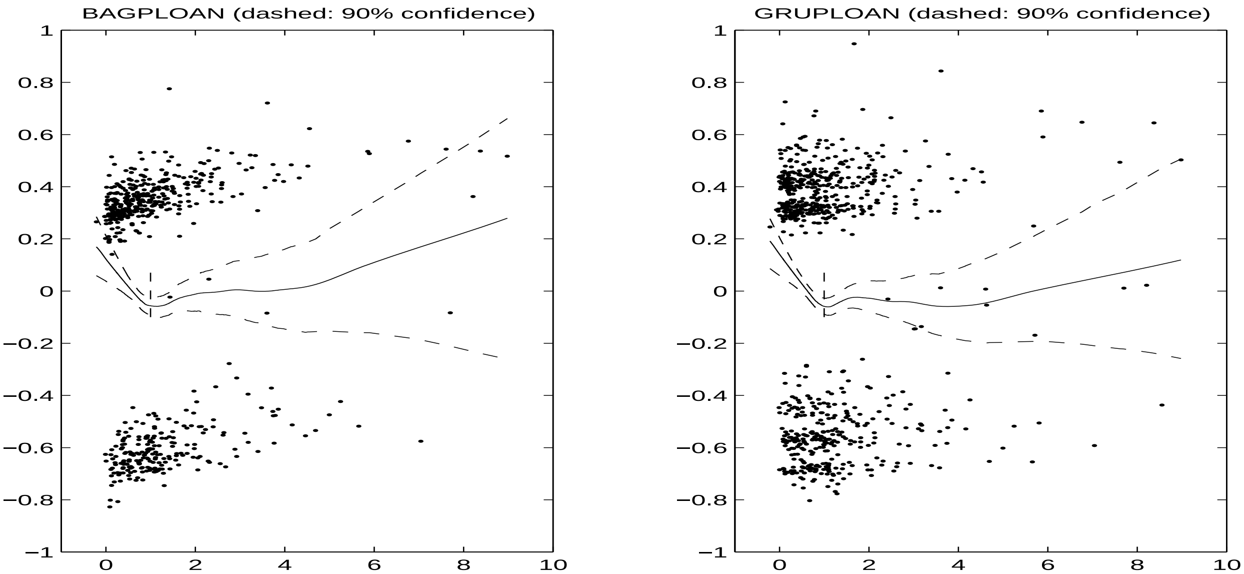

Figure 2: Relative wealth – household wealth divided by village mean – as a predictor of having a group loan.