Determinants of the Prevalence of Cooperative Borrowing

Christian Ahlin and Robert Townsend

June 2002

Abstract

Various theories shed light on the relative advantages of individualistic activities versus collective organizations. As is well known, the relative effectiveness depends critically on the covariance of output across agents. New work has also highlighted the wealth of the agents relative to the principal, as well as wealth inequality between agents, as important factors in determining the regime. Also important are the relative costs for the principal and the agent of monitoring an(other) agent. We attempt to test these theories by relating the prevalence of BAAC group borrowing, and group borrowing in general, across villages in Thailand. As predicted by theory, cooperative borrowing exhibits a U-shaped relationship with village wealth, and wealth inequality increases group regime prevalence. We find evidence that, contrary to theory, cooperative borrowing increases with covariance of output. Finally, as controls, high variability of income positively predicts cooperative borrowing and education levels negatively predict cooperative borrowing.

1 Theories and Implications

1.1 Technological Correlation

Correlation of output across agents is related by Holmström and Milgrom (1990), in the context of unobserved actions, to the optimality of a regime encouraging competition or cooperation. They find that there exists a cutoff technological correlation coefficient, call it $\overline{\rho}$. *For $\rho < \overline{\rho}$, the cost of implementing a given set of actions is lower under a cooperative regime; for $\rho > \overline{\rho}$, the competitive regime yields more surplus. Thus, one should expect to see more of the group regime among agents whose output is less correlated.*

Assume there are two agents, indexed by a and b respectively. Each produces output x_i as a function of his effort e_i and some random shock ε_i. In particular:

$$x_i = e_i + \varepsilon_i, \ \ i = a, b.$$

(In the following, any variable not otherwise defined, with suppressed subscript, indicates the vector of two elements, $[var_a, var_b]'$.) The epsilons are distributed joint-normally with

means zero and variance-covariance matrix Σ, where

$$\Sigma \equiv \begin{bmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{bmatrix}.$$

Since only the $x_i's$ are publicly observed, agent compensation must be in terms of them. Attention is restricted to linear compensation functions, R_i:

$$R_a(x;\alpha) = \alpha_0 + \alpha_a x_a + \alpha_b x_b; R_b(x;\beta) = \beta_0 + \beta_a x_a + \beta_b x_b.$$

Agents maximize expected utility of exponential functional form. This gives the well-known certainty equivalence utility function

$$CE_a(e;\alpha) = \alpha_0 + \alpha^T e - C_a(e_a) - (1/2) r_a [\alpha^T \Sigma \alpha]$$

for agent a, where $C_a(\cdot)$ is the cost of effort (strictly convex), r_a is his coefficient of absolute risk aversion, and $\alpha^T \Sigma \alpha = \alpha_a^2 \sigma_a^2 + \alpha_b^2 \sigma_b^2 + 2\alpha_a \alpha_b \sigma_{ab}$. An analogous expression holds for agent b. The principal is risk neutral and thus has certain equivalence utility of

$$CE_p(e;\alpha,\beta) = (\mathbf{1} - \alpha - \beta)^T e - \alpha_0 - \beta_0.$$

The principal maximizes his utility subject to agents' participation constraints and incentive compatibility constraints. Since the model exhibits transferable utility, this is the same as maximizing total surplus subject to the incentive compatibility constraints. Thus, the optimal contract chooses (α, β) to maximize

$$e_a + e_b - C_a(e_a) - C_b(e_b) - 1/2 r_a \alpha^T \Sigma \alpha - 1/2 r_b \beta^T \Sigma \beta$$
$$\text{subject to: } \alpha_a = C_a'(e_a) \text{ and } \beta_b = C_b'(e_b).$$

Total surplus is expected output less costs of effort and risk costs. The optimal solution satisfies

$$\alpha_a = \frac{1}{1 + r_a \sigma_a^2 (1-\rho^2) C_a''}, \ \alpha_b = -\alpha_a \sigma_{ab} / \sigma_b^2,$$

where $\rho \equiv \sigma_{ab}/\sigma_a \sigma_b$. Similar expressions hold for (β_a, β_b). The total risk cost under the contract (α, β) can be calculated to be:

$$(1/2) r_a \alpha_a^2 \sigma_a^2 (1-\rho^2) + (1/2) r_b \beta_b^2 \sigma_b^2 (1-\rho^2) \tag{1}$$

The preceding analysis has assumed that agents could not observe each other's actions or in fact contract with each other in any way. The risk sharing occurs because the correlation between shocks mitigates the lack of information on agents' efforts. In the extreme case, when $\rho = 1$, there is no risk cost in the optimal contract; rather full risk sharing obtains.

Next, the case is considered where the two agents *can cooperate*. In particular, they can contract with each other on actions and outcomes. This allows them to mutually reinsure each other and to enforce actions. The contracts they can write will be of the form

$$T(e,x) = \gamma_a x_a + \gamma_b x_b + t(e_a, e_b).$$

As above, contracts linear in output are considered. Now, given external incentives (α, β), the group chooses (e, t, γ) to maximize

$$\begin{gathered}(\alpha+\beta)^T e - C_a(e_a) - C_b(e_b) - 1/2 r_a(\alpha-\gamma)^T\Sigma(\alpha-\gamma) - 1/2 r_b(\beta+\gamma)^T\Sigma(\beta+\gamma) \\ \text{subject to: } t(\cdot) \text{ enforces e as a Nash equilibrium.}\end{gathered} \tag{2}$$

The principal then picks (α, β) to maximize

$$\begin{gathered}e_a + e_b - C_a(e_a) - C_b(e_b) - 1/2 r_a(\alpha-\gamma)^T\Sigma(\alpha-\gamma) - 1/2 r_b(\beta+\gamma)^T\Sigma(\beta+\gamma) \\ \text{subject to: } \exists\ t(\cdot) \text{ such that } (e, \gamma, t) \text{ solves (2) for } (\alpha, \beta).\end{gathered} \tag{3}$$

It is shown that the principal's optimal design coincides with that for a single ("syndicate") agent whose cost function $C\cdot$ satisfies $C(e_a, e_b) = C_a(e_a) + C_b(e_b)$ and absolute risk aversion r satisfies $1/r = 1/r_a + 1/r_b$. Thus the principal is reduced from four degrees of freedom to two, since what matters is not α and β individually, but $\alpha+\beta$. So, without loss of generality, assume $\alpha_b = \beta_a = 0$. Given (α_a, β_b), the total risk cost is

$$(1/2) r[\alpha_a^2\sigma_a^2 + \beta_b^2\sigma_b^2 + 2\alpha_a\beta_b\rho\sigma_a\sigma_b]. \tag{4}$$

To implement a given set of actions e, since the agents act as a syndicate, it is clear that α_a and β_b should be the same as under the non-cooperation regime when the same actions e are desired. Thus, using equations (1) and (4), we can compare across regimes the risk costs of implementing a given set of actions. Since risk costs are the only part of total surplus that varies across regime type, given the same actions being implemented, this comparison tells us which regime allows cheaper implementation of a set of actions.

It is easily checked that at $\rho = 0$, the cost of implementing any set of actions is lower under the cooperative regime. (Assume α_a and β_b are strictly positive.) Similarly, at $\rho = 1$ the cost of implementing any set of actions is lower under the non-cooperative regime. These two facts imply that the cooperative regime is preferred by the principal at $\rho = 0$ and the non-cooperative at $\rho = 1$. Now, both expressions in equations (1) and (4) are strictly monotonic in ρ: under cooperation, increasing, and under non-cooperation, decreasing. Thus the payoff to the principal is strictly increasing in ρ under non-cooperation and strictly decreasing under cooperation. The payoffs under each regime are clearly continuous in ρ. *This proves that there is a cutoff, $\overline{\rho} \in (0, 1)$, above which the non-cooperative regime dominates and below which the cooperative regime does.*

1.2 Levels and Distribution of Wealth

E. S. Prescott and Townsend (1999) offer a model in which technological correlation is not the only factor that affects regime optimality: wealth levels and distributions matter too. Their model is also in the context of unobserved effort. Analysis is done with a principal and two agents, but welfare theorem results are proved that link this setup to a competitive equilibrium in insurance markets where the agents are the buyers and sellers. In particular, the participation wealth level of the principal is identified with the amount of positive or negative wealth redistribution into a given type of economy.

Two types of cooperation are examined and compared with non-cooperation. We focus on the first, which is similar to that of section 1.1. The principal can choose a regime in which 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. Alternatively, the regime can disallow any transfers and coordination on actions.

Let there be two agents and two technologies, indexed by i and j, respectively. Call the effort agent i exerts on technology j, e_{ij}. Define e_i as the total effort exerted by agent i; $e_i \equiv e_{i1} + e_{i2}$. 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 $c \equiv (c_1, c_2)$, $q \equiv (q_1, q_2)$, $a \equiv (a_1, a_2)$ and $e_{i\bullet} \equiv (e_{i1}, e_{i2})$.

Each of c_i, q_j, and e_{ij} can take on a finite amount of values. Agent i maximizes utility $U_i(c_i) + V_i(T_i - e_i)$, defined over his consumption c_i and his leisure, which equals total time endowment T_i minus total effort. The principal's payoff is $W(q_1 + q_2 - c_1 - c_2)$. Technology is expressed as a probability mass function $p(q|a)$.

The principal chooses a mechanism for each regime. He chooses a mechanism that maximizes a weighted average of the agents' utilities subject to his own expected utility reaching at least $\overline{W}$. In the *non-cooperative case*, the mechanism is $\Pi(c, q, e_{1\bullet}, e_{2\bullet})$ which must satisfy probability measure constraints

$$\Pi(c, q, e_{1\bullet}, e_{2\bullet}) \geq 0, \quad \forall\ c, q, e_{1\bullet}, e_{2\bullet}$$

and

$$\sum_{c,q,e_{1\bullet},e_{2\bullet}} \Pi(c, q, e_{1\bullet}, e_{2\bullet}) = 1.$$

This mechanism should maximize

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

subject to:

$$\sum_{c,q,e_{1\bullet},e_{2\bullet}} \Pi(c, q, e_{1\bullet}, e_{2\bullet}) W(q_1 + q_2 - c_1 - c_2) \geq \overline{W},$$

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

$$\sum_{c,q,e_{2\bullet}} \Pi(c, q, e_{1\bullet}, e_{2\bullet})[U_1(c_1) + V_1(T_1 - e_1)] \geq$$

$$\sum_{c,q,e_{2\bullet}} \Pi(c, q, e_{1\bullet}, e_{2\bullet}) \frac{p(q|\hat{e}_{1\bullet} + e_{2\bullet})}{p(q|e_{1\bullet} + e_{2\bullet})} [U_1(c_1) + V_1(T_1 - \hat{e}_{1\bullet})], \ \forall e_{1\bullet}, \hat{e}_{1\bullet},$$

and an incentive constraint for agent b similar to the last inequality. The first constraint is a participation constraint for the principal, the second ensures that the mechanism assigns technologically feasible probabilities, and the last is an incentive constraint ensuring agent a abides by his recommended effort allocation.

In the *cooperative case*, the group is free to allocate consumption internally. This means that the principal cannot specify the c_i's separately, only total group consumption, $c_g (= c_1 + c_2)$. Similarly, he cannot specify the $e_{i\bullet}$'s separately, only 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, $c_i(c_g, \mu)$. Similarly, effort of agent i can be expressed as a function of total effort and the pareto weights, $e_i(e_g, \mu)$, where $e_g \equiv a_1 + a_2$.

The principal chooses a mechanism $\Pi(c_g, q, a, \mu)$ satisfying similar probability measure constraints as above. The mechanism should maximize

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

subject to:

$$\sum_{c,q,a,\mu} \Pi(c, q, a, \mu) W(q_1 + q_2 - c_g) \geq \overline{W},$$

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

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

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

For obvious reasons, joint deviations in effort are considered in the incentive compatibility constraint.

The cooperative regime and non-cooperative regime have different implications for individual and group consumption. In particular, in the cooperative regime full insurance occurs, meaning that an individual's consumption is fully determined by group consumption (plus his pareto weight.) In the non-cooperative regime, however, an individual's consumption will in general vary with the disturbance to the technology into which he is putting comparatively more effort. *This implies that tests of full consumption insurance can be used to define the exact boundaries of the group regime.*

The two regimes can be compared at different values for $\overline{W}$ and λ in terms of which can deliver a higher value for the objective function. Simulations were used to analyze the case in which each agent is restricted to working his own technology (that is, $e_{ij} \equiv 0, i \neq j$) [in the non-cooperative case only?] and there is some correlation across technologies. Each agent also has symmetric utility, endowment, and technological parameters. The simulations indicate that the cooperative regime dominates for asymmetric λ's, no matter what $\overline{W}$ is. The reason is because it is relatively better at extracting wealth from the low-weight agent and giving it to the high-weight. For more symmetric λ's, the non-cooperative regime can dominate, but this is highly dependent on $\overline{W}$. For $\overline{W}$ low enough, it strictly dominates nowhere. As $\overline{W}$ increases, the (symmetric) range around $\lambda_1 = 1/2$, for which the non-cooperative regime dominates, increases up to a certain point and then declines. These simulations indicate *we should expect to see the cooperative regime occurring more frequently at higher levels of inequality. 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 wealth.*

1.3 Relative Costs of Monitoring

E. S. Prescott (1997) considers a simple model, based on one of delegated monitoring by Diamond (1984, 1996), in which lenders decide which type of contracts to employ with borrowers. They can opt for individual lending with either liquidation or direct monitoring, or group lending. The choice is seen to depend on the lender's cost of direct monitoring and the borrowers' cost of monitoring a (fellow) borrower.

Consider an economy with many lenders and borrowers. Each borrower has access to a project requiring 1 unit of capital and producing 1 unit of output with probability 0.2 and 1.4 units with probability 0.8. The output is private information. Borrowers have no capital.

Lenders have capital and access to a technology that produces 1.05 units of output per unit invested. They are risk-neutral, so they will lend to borrowers if the expected return is at least 1.05 per unit lent. Lenders can monitor an individual borrower at cost K; that is, by paying K they can observe output and thus contract on it. They can also liquidate projects, resulting in net payoffs of zero for borrower(s) and lender. Lenders can also lend to pairs of borrowers, who pay k_g to monitor each other. Monitoring makes their output internal common knowledge and they are seen to act as a unit, maximizing total expected payoffs and sharing equally.

Thus a lender has three choices: individual lending with liquidation, individual lending with monitoring, and group lending (with liquidation). Let R be the expected payoff to the lender (assume $R = 1.05$), U be the expected utility of the borrower, L be the liquidation costs, and M be the monitoring costs. Let F be the payoff required under a debt contract.

Under individual lending with liquidation, it must be that $F > 1$ since the lender must earn 1.05. Thus the borrower can pay F only when output is 1.4. So $F = 1.05/0.8 = 1.3125$ ensures that the lender's expected payoff is 1.05, as required. The borrower's utility is then $0.8(1.4 - 1.3125) = 0.07$. Liquidation costs are 0.2, since output of 1 is destroyed with probability 0.2. There is no monitoring. Thus, under this system, $R = 1.05$, $U = 0.07$, $M = 0$, and $L = 0.2$.

Under individual lending with monitoring, the lender can make the contract conditional on output. Since both parties are risk-neutral, a continuum of contracts provides the lender with payoff 1.05. Consider the one where the borrower pays the lender everything (1 unit) in the low-output state. Then the payoff F in the high output state must satisfy $1.05 = (0.2)1 + (0.8)F - K$ to ensure the lender expected payoff of 1.05. The borrower's expected payoff is $(0.8)(1.4 - F) = 1.12 - (0.8)F$. Given the previous equation for F, this means $U = .27 - K$. We also have $M = K$, $L = 0$, and $R = 1.05$ as usual.

Under group lending, the lender will liquidate output of both borrowers unless they pay F. As long as $F < 2.4$, the group will pay in all cases except when both realize low output. This means that F must satisfy $2.10 = [1 - (0.2)^2]F$ so that the lender receives 1.05 per unit of capital. Thus $F = 2.1875$. Each borrower has expected utility of $0.64(2.8 - F)/2 + 0.32(2.4 - F)/2 - K_g$. Substituting for F, we get $U = 0.23 - K_g$. Also, per person, $R = 1.05$, $M = K_g$, and $L = 0.04$.

It is easy to compare the regimes by comparing the utility of a borrower under each. In particular, as long as $K_g < \min\{K - 0.04, .16\}$, the group lending regime is preferred to either of the individual schemes. *The prediction is that the lower are group costs of monitoring and the higher are the lending institution's costs of monitoring, the more likely it is that the*

group lending scheme will be observed.

2 Empirical Results

2.1 Description of Variables

Our data are taken from a survey of households in Thailand conducted in 1997. 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. Within regions, sampling was stratified based on ecological zones, to ensure variation in good and bad years. This stratification affected choices of sub-counties, or *tambons*. Within tambons, choices of villages and of households within villages were random. There were 192 villages surveyed. In virtually all villages, fifteen households were surveyed, giving a total sample size of 2875. Our analysis will be on the village level.

To measure existence of the cooperative regime, we focus on borrowing schemes. The credit market is one in which moral hazard is prevalent due to lack of collateral of many poor borrowers. Thus the lending institution faces the tradeoff of offering cooperative or non-cooperative contracts to borrowers. In particular, joint liability lending schemes, in which groups borrow and each individual is held liable for all loans in the group, can be thought of as cooperative. This is because a significant amount of insurance comes about within the group as members ensure that other members can repay their loans. Individual loans can be thought of as the non-cooperative regime, though in reality the principal cannot enforce complete lack of cooperation between agents. He can, however, abstain from encouraging it.

The first measure of the group regime we consider, BAACGLON, is the percentage of respondents in a given village who are currently borrowing from the BAAC (Bank for Agriculture and Agricultural Cooperatives) under a joint liability lending scheme. We restrict initial attention to this institutional lender because it is the primary supplier of joint liability loans in the areas surveyed. The BAAC targets agricultural workers almost exclusively. It offers both individual loans, which much 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, as described above.

We next consider the same measure *normalized* by the total percent in the village borrowing from the BAAC. This is called BAACGVST. This is because the first measure captures not only *prevalence of the group regime* in borrowing from the BAAC, but also the *total amount of borrowing from the BAAC*. By normalizing in the way described above, we capture the percent of BAAC loans in the village that use the group regime instead of the individual one.

The final measure we look at, CPVSNONC, includes lending institutions other than the BAAC. This is because, while the BAAC is the main supplier of group-guaranteed loans, this is not always the case for individual loans. If we restrict attention to BAAC loans only, we are often seriously underestimating prevalence of individual loans. This measurement error is likely correlated with unobservables, namely whatever influences supply of individual

Table I - Summary of Variables

Variable	Description
	DEPENDENT:
BAACGLON	Fraction of village with a group-guaranteed loan from the BAAC
BAACGVST	Village fraction of BAAC loans that are group-guaranteed
CPVSNONC	Village fraction of loans from lending institutions that are guaranteed by one or more individuals besides the borrower
	INDEPENDENT:
	Correlation:
GOODCORR	Village measure of coincidence of economically 'good' years across villagers
BADCORR	Village measure of coincidence of economically 'bad' years across villagers
	Wealth:
WEALTH	Village average wealth
WEALTHSQ	WEALTH squared
COEFVARW	Coefficient of variation for wealth in the village
	Monitoring Costs:
BCOOPPCT	Percent in tambon naming this village best in the tambon for cooperation among villagers
SHAREREL	Index for sharing among relatives in the tambon
SHARENON	Index for sharing among non-relatives in the tambon
NRTHEAST	Dummy = 1 if the village is in the northeast region
	Control:
AGPCT	Percent of village engaging in agricultural activity
LAND	Average landholdings (in rai) of villagers
EDCATION	Village average highest amount of education attained by any household member
VARIBLTY	Village average coefficient of variation for next year's expected income

loans in a given region. Thus the second measure we consider does not accurately capture relative prevalence of the group regime. In an effort to create a broader measure of group and individual loans, we consider loans from every medium or large institution in the survey: the BAAC, PCGs (Production Cooperative Group), commercial banks, agricultural cooperatives, village funds, and rice banks. We categorize a loan as a group loan if it is guaranteed by one or more people other than the borrower. It is considered an individual loan if it uses some other form of collateral, for example, land, savings, or future crops. (Note that this does not exhaust all village loans, since we only consider the institutions listed and only loans for which the type of collateral was clear.) For each village, our measure of relative prevalence of the group regime is just the percent with group loans divided by the sum of the percent with group loans and the percent with individual loans.

Table I gives a summary of all dependent and independent variables used. Our baseline

explanatory variables fit into four categories. Several reflect the *costs of monitoring*, for the borrowers and the lending institution. BCOOPPCT gives the percentage of respondents in this village's tambon (four villages) naming this village as having the best cooperation among villagers. Cooperation among villagers is likely a good indication of low borrower costs of monitoring. SHAREREL and SHARENON are measures of sharing among relatives and non-relatives, respectively, in the same tambon. These also seem to indicate existence of tight social networks that could cheapen monitoring. Measures of sharing are first computed for each responding household in the village, then SHAREREL and SHARENON are the village averages of these measures. The household measure for sharing among relatives equals the sum of four dummies. The dummies equal one if, in the past year, there has been sharing of rice, helping with money, helping with free labor, and free use of farm equipment, respectively, among this household and its relatives in the tambon. The household measure for sharing among non-relatives is computed analogously. NRTHEAST is a dummy equalling one if the village is in the northeast region. This is included in the baseline model when the dependent variable is restricted to the BAAC, since the BAAC is headquartered much closer to the central region. Thus, costs of institutional monitoring are probably higher in the less readily accessible northeast region.

We use two variables to measure *correlation of output* in the village. Each villager was asked to indicate which of the past five years was best for household income, and which was worst. From these questions we construct two measures of correlation in the village, GOODCORR and BADCORR, in the following way. Each variable is constructed as the percent of pairwise matchings of village respondents in which the same year is mentioned. The number of pairwise matchings if there are N respondents is $(N-1)+(N-2)+...+2+1 = N(N-1)/2$.[1] If N_i is the number of year-i respondents in the village, then the number of pairwise matchings in which both respondent indicated year i is $(N_i-1)+(N_i-2)+...+2+1 = N_i(N_i-1)/2$.[2] Since the question is restricted to the previous five years, BADCORR works out to be

$$\frac{N_1(N_1-1)+N_2(N_2-1)+N_3(N_3-1)+N_4(N_4-1)+N_5(N_5-1)}{N(N-1)},$$

where N_i is the number of respondents indicating year i as worst for income. GOODCORR is exactly analogous.

Wealth and wealth distribution are captured in three variables. WEALTH is the average household wealth in the village. WEALTHSQ is its square. We include the squared term because of the non-monotonic relationship predicted by the theory. COEFVARW is the coefficient of variation for wealth in the village. It is intended to capture the degree of inequality.

Two *control variables* are included in both baseline models. AGPCT is the percent of respondents in the village engaging in agricultural activity. This is needed because the

[1] To see this, one could imagine N people in a line. The first goes down the line, matching with the N-1 remaining people, then exits. The second repeats this, matching with the remaining N-2 people, then exits. This continues until the second to last person matches with the last person.

[2] Again, one could picture the line of N respondents, but this time sorted so that all the year-i respondents come last. Then the first $N - N_i$ people exit with no year-i matches, and we are left with a line of N_i people all matching on year i.

BAAC lends almost exclusively to agricultural workers. Since the BAAC is a chief supplier of group loans, we must also include it even when considering additional institutions. LAND is the average landholdings of villagers interviewed. It is included because, in the context of credit, existence of collateral (such as land) is another factor that determines the optimality of individual or group loans. Two other control variables beyond the baseline model are considered. EDCATION gives the village average of highest years of education attained among members of a household. VARIBLTY is the village average of household coefficients of variation for expected income next year. Household respondents were asked how much they will earn if next year is a good year (Hi), how much if bad (Lo), and how much do they expect to earn (Ex). We assume the distribution of income is over these three mass points, Hi, Lo, and Ex, and that the probabilities of each are a, q-a, and 1-q, respectively. Given that the expected value must be Ex, we know $aHi + (q-a)Lo + (1-q)Ex = Ex$. This gives that $a = q(Ex - Lo)/(Hi - Lo)$. The variance σ^2 is, using the definition, $a(Hi - Ex)^2 + (q-a)(Ex - Lo)^2$. Substituting in our expression for a in terms of q, the variance simplifies to

$$\sigma^2 = q(Hi - Ex)(Ex - Lo).$$

Dividing by $(Ex)^2$ and taking the square root, we calculate the individual's coefficient of variation to be:

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

It is evident that the coefficient of variation varies positively with q. (Recall that q is the probability mass on the extreme outcomes, Hi and Lo.) Fortunately, q enters multiplicatively. This implies that, as long as q is strictly positive, choice of q scales a regression coefficient but does not affect hypothesis tests where the null hypothesis is a zero coefficient. We use $q = 1$.

2.2 Regression Results

We use OLS regressions to analyze the baseline models and slight variations on them. These regressions are reported in tables II, III, and IV for dependent variables BAACGLON, BAACGVST, and CPVSNONC, respectively.

Table II contains some interesting results. It is surprising that high correlation of bad years is very significantly and positively related to the existence of the group regime. Theory predicts that high correlation should result in greater use of the non-cooperative regime as the principal makes optimal use of the information provided by correlation. The result may not be surprising, though, in the case of the BAAC. They seem to offer choice of contracts to the borrowers, allowing them to decide whether to form groups or borrow individually. Given the choice remains with the borrower, it may not be surprising that areas of high correlation in bad years would see more group-guaranteed loans, because high correlation would tend to cut down on joint liability payments. In short, when the choice of regime is left to the borrower, he may self-select into the suboptimal regime. Correlation of good years has a positive, insignificant coefficient.

Wealth exhibits the predicted U-shaped pattern, but the coefficients are not significantly different from zero in most cases. The exception is when the dummy variable for the poorer northeast region is excluded, in which case the WEALTH variable becomes significant at the

Table II - BAACGLON as dependent variable

BAACGLON is the percent of the village with a group-guaranteed loan from the BAAC.
|Tstats| in parentheses; significance at 80, 90, 95 and 99% denoted by *, **, ***, and ****, respectively.

Independent Variables	Baseline	−(NRTHEAST)	+(EDCATION)	+(VARIBLTY)
GOODCORR	9.36E-3	3.41E-2	1.48E-2	2.66E-3
	(0.12)	(0.44)	(0.19)	(0.04)
BADCORR	.217	.233	.219	.204
	(2.79)****	(2.94)****	(2.81)****	(2.59)***
WEALTH	-4.56E-9	-1.68E-8	-9.32E-9	-2.51E-9
	(0.35)	(1.34)*	(0.68)	(0.19)
WEALTHSQ	1.21E-17	5.78E-16	2.04E-16	-8.13E-17
	(0.01)	(0.70)	(0.24)	(0.10)
COEFVARW	-1.20E-2	-1.88E-2	-7.52E-3	-1.61E-2
	(0.46)	(0.71)	(0.29)	(0.61)
BCOOPPCT	.239	.229	.228	.245
	(2.53)***	(2.38)***	(2.40)***	(2.59)***
SHAREREL	-.0304	-.0523	-.0358	-.0325
	(0.80)	(1.37)*	(0.93)	(0.85)
SHARENON	.0198	.0695	.0192	.0160
	(0.53)	(2.03)***	(0.51)	(0.42)
NRTHEAST	.092		.096	.093
	(2.85)****		(2.96)****	(2.89)****
AGPCT	.191	.244	.212	.168
	(2.76)****	(3.60)****	(2.97)****	(2.31)***
LAND	-1.79E-4	-7.52E-4	-8.16E-5	-3.62E-4
	(0.17)	(0.71)	(0.08)	(0.34)
EDCATION			.0115	
			(1.15)	
VARIBLTY				.142
				(1.00)
R^2	.286	.254	.292	.290
Adjusted R^2	.243	.213	.244	.243

80% level. Wealth dispersion exhibits an unexpected sign, but is not significantly different from zero.

The costs of monitoring variables perform well. The measure of village cooperation, which proxies for low costs of monitoring, is significantly and positively related to existence of the group regime. Likewise, sharing among non-relatives is a positive predictor, but only significantly when the NRTHEAST dummy is excluded. Sharing among relatives is a negative predictor of group borrowing, though only marginally significant when the NRTHEAST dummy is excluded. Possibly this reflects the fact that strong family risk-sharing networks are a substitute for institutionally encouraged ones. The NRTHEAST coefficient appears to confirm the fact that high costs of monitoring by the institution make the group regime more likely. This is because the BAAC is headquartered much closer to the central region and the northeast provinces are less accessible. However, evidence from the second set of regressions, in Table III, belies this interpretation. In Table II, what is being picked up by NRTHEAST is the overall prevalence of BAAC borrowing in the northeast region, not group borrowing per se. Since less borrowing options exist in the northeast region, the BAAC is more heavily relied upon there.

The percent of agricultural workers in the village is also a strong predictor of group borrowing from the BAAC. This is not surprising since they target agricultural workers. LAND's coefficient has the expected sign, since land collateral makes individual borrowing more possible, but is not significantly different from zero.

The second set of regressions, in Table III, has as its dependent variable BAACGVST, the percent of BAAC loans in the village that are group-guaranteed. There are a few differences between this and the first regression, some of them substantial. The main explanation for differences is likely to be that some variables influence not just relative amount of group borrowing, but absolute amount. Both of these effects are likely to be picked up in the first regression, while only the former should be seen in the second regression.

Coefficients on the correlation variables remain positive, but become less significantly different from zero in the case of bad-year correlation. The U-shaped wealth relationship becomes fairly significant in all regressions. In addition, the coefficient on wealth dispersion becomes positive, though remaining insignificant. The rankings of village cooperation and the measures of sharing retain their signs but become less significant, except for sharing among non-relatives. This variable is now significant only when the NRTHEAST dummy is included, in contrast to the first regression results. The explanation is clearly that, without the NRTHEAST dummy, this sharing variable picks up some of the negative effect of the northeast region (since sharing is higher there) and its coefficient thus becomes less positive.

The NRTHEAST dummy changes sign and becomes fairly insignificant. Obviously, there is more BAAC borrowing of either type in the northeast, and this fact was picked up by the dummy in the first regression. Nevertheless, the sign is surprising; one would expect group monitoring to be relatively cheaper in the northeast region. AGPCT becomes less important, not surprisingly, since its main effect was on the magnitude of any kind of BAAC borrowing. Land becomes more significant, as expected. The reason for its negative sign is that more land makes the individual loan more possible, since land can be used as collateral. The assumption that land is a positive predictor of total BAAC borrowing and a negative predictor of the group regime explains the difference in significance between the first and second set of regressions.

Table III - BAACGVST as dependent variable

BAACGVST is the fraction of village loans from the BAAC that are group- guaranteed.
|Tstats| in parentheses; significance at 80, 90, 95 and 99% denoted by *, **, ***, and ****, respectively.

Independent Variables	Baseline	−(NRTHEAST)	+(EDCATION)	+(VARIBLTY)
GOODCORR	.117	.0971	.112	.0797
	(0.75)	(0.63)	(0.71)	(0.51)
BADCORR	.215	.210	.212	.148
	(1.37)*	(1.34)*	(1.35)*	(0.93)
WEALTH	-4.61E-8	-3.58E-8	-4.30E-8	-3.46E-8
	(1.53)*	(1.30)*	(1.35)*	(1.13)
WEALTHSQ	3.11E-15	2.62E-15	2.99E-15	2.55E-15
	(1.68)**	(1.49)*	(1.57)*	(1.37)*
COEFVARW	.0441	.0483	.0414	.0301
	(0.79)	(0.63)	(0.73)	(0.54)
BCOOPPCT	.0976	.106	.106	.149
	(0.46)	(0.51)	(0.50)	(0.71)
SHAREREL	-.0528	-.0405	-.0492	-.0613
	(0.65)	(0.51)	(0.60)	(0.76)
SHARENON	.129	.0925	.131	.113
	(1.50)*	(1.24)	(1.51)*	(1.32)*
NRTHEAST	-.0611		-.0641	-.0527
	(0.85)		(0.88)	(0.73)
AGPCT	.169	.126	.160	.0970
	(1.00)	(0.78)	(0.92)	(0.56)
LAND	-3.78E-3	-3.39E-3	-3.79E-3	-4.41E-3
	(1.69)**	(1.55)*	(1.69)**	(1.97)**
EDCATION			-7.03E-3	
			(0.32)	
VARIBLTY				.588
				(1.92)**
R^2	.073	.068	.073	.095
Adjusted R^2	.005	.007	-.001	.023

Finally, VARIBLTY is a strong predictor of the group regime. The theory we have covered has no prediction in either direction. Absent any endogeneity problem,[3] this result can be evidence that the group regime is turned to as a forum for better provision of mutual insurance, and moreso when the need is greater.

In our third set of regressions, reported in Table IV, the dependent variable CPVSNONC incorporates borrowing not just from the BAAC but from all lending institutions. The distinguishing characteristic between cooperative and non-cooperative loans is taken to be that people other than the borrower cosign the former while some other type of collateral is used in the latter. CPVSNONC is thus the fraction of both types of loans that are considered cooperative. It is similar to BAACGVST, but in some ways should be a better indicator for reasons discussed above, namely that the BAAC individual loans are a very poor measure of total individual borrowing, in the central region especially.

The results are strongest for the wealth and wealth distribution variables. The U-shaped relationship with wealth is strongly significant, especially in the baseline model. In addition, wealth dispersion measured by COEFVARW has the correct positive sign and is significant at the 80% level.

In a strange twist, the correlation of good years becomes a significant (positive) predictor of cooperative borrowing, while correlation of bad years becomes negative and insignificant. It is hard to imagine what might be going on, although it seems that institutions besides the BAAC avoid offering the group regime to villages with high correlation of bad years, unlike the BAAC. This could be the force making the coefficient on bad-year correlation negative in this set of regressions.

The cooperation variables remain largely as before, with sharing among relatives a negative predictor and sharing among non-relatives a positive predictor of the group regime. The cooperation ranking variable inexplicably becomes negative and insignificant. The dummy for the northeast region is not significantly different from zero.

Of the control variables, the prevalence of agricultural workers remains a positive and significant explanatory variable. However, when education and variability levels are controlled for, it loses significance. Land takes on a positive sign, but is insignificant. It could be that land is less readily accepted as collateral by other lending institutions than by the BAAC. This would mean we no longer expect a negative sign. Finally, both education levels and variability are strong predictors. Higher education levels lead to lower prevalence of jointly liable loans. Higher variability of income leads to higher prevalence of the cooperative regime.

[3] In some models, for example Stiglitz 1990, joint liability increases income variability; it acts as a mean-preserving spread on income. In this case, village income variability would itself be a function of the prevalence of joint liability borrowing. However, this is not likely to be a problem here, for two reasons. First, the percent of village respondents with joint liability loans is low (on average, 20%), so joint liability borrowing is at best a minor determinant of village income variability. Second, the joint liability mechanism of the BAAC seems rarely to be direct liability in the sense of enforcing repayment for another or seizing of assets, as in Stiglitz 1990. Rather, it mostly appears to be in the form of delay or refusal of future loans or slight raising of interest rates. Arguably, neither of these affect future income as dramatically.

Table IV - CPVSNONC as dependent variable

CPVSNONC is the fraction of village loans from institutions that are guaranteed by one or more individuals other than the borrower.

|Tstats| in parentheses; significance at 85, 90, and 95% denoted by *, **, and ***, respectively.

Independent Variables	Baseline	+(NRTHEAST)	+(EDCATION)	+(VARIBLTY)
GOODCORR	.231	.233	.203	.202
	(1.67)**	(1.67)**	(1.47)*	(1.46)*
BADCORR	-.0160	-.0143	-.0303	-.0662
	(0.11)	(0.10)	(0.22)	(0.46)
WEALTH	-5.75E-8	-5.89E-8	-4.33E-8	-5.11E-8
	(2.59)***	(2.48)***	(1.83)**	(2.29)***
WEALTHSQ	2.44E-15	2.51E-15	1.87E-15	2.16E-15
	(1.66)**	(1.65)*	(1.25)	(1.47)*
COEFVARW	.0771	.0765	.0663	.0632
	(1.64)*	(1.62)*	(1.41)	(1.34)
BCOOPPCT	-.103	-.104	-.0778	-.0857
	(0.59)	(0.59)	(0.45)	(0.49)
SHAREREL	-.0877	-.0901	-.0699	-.0989
	(1.27)	(1.28)	(1.01)	(1.44)
SHARENON	.0822	.0878	.0817	.0765
	(1.29)	(1.22)	(1.29)	(1.21)
NRTHEAST		-.0101		
		(0.17)		
AGPCT	.209	.215	.146	.137
	(1.73)**	(1.71)**	(1.15)	(1.08)
LAND	1.58E-3	1.51E-3	1.54E-3	1.07E-3
	(0.82)	(0.76)	(0.80)	(0.55)
EDCATION			-.0315	
			(1.71)**	
VARIBLTY				.476
				(1.80)**
R^2	.137	.138	.152	.153
Adjusted R^2	.089	.084	.098	.100

Table V - CPVSNONC as dependent variable

CPVSNONC is the fraction of village loans from institutions that are guaranteed by one or more individuals other than the borrower.

|Tstats| in parentheses; significance at 85, 90, and 95% denoted by *, **, and ***, respectively.

Independent Variables	Baseline	+(NRTHEAST)	+(EDCATION)	+(VARIBLTY)
GOODCORR	.225	.227	.198	.196
	(1.63)*	(1.63)*	(1.44)	(1.43)
BADCORR	-.0142	-.0127	-.0293	-.0655
	(0.10)	(0.09)	(0.21)	(0.46)
WEALTH	-5.79E-8	-5.91E-8	-4.32E-8	-5.14E-8
	(2.62)***	(2.49)***	(1.84)**	(2.31)***
WEALTHSQ	2.44E-15	2.49E-15	1.85E-15	2.15E-15
	(1.66)**	(1.64)*	(1.24)	(1.46)*
COEFVARW	.0798	.0793	.0680	.0653
	(1.71)**	(1.69)**	(1.45)*	(1.39)
SHAREREL	-.0913	-.0934	-.0722	-.102
	(1.33)	(1.33)	(1.04)	(1.49)*
SHARENON	.0805	.0853	.0804	.0751
	(1.27)	(1.19)	(1.28)	(1.19)
NRTHEAST		-.0086		
		(0.14)		
AGPCT	.215	.219	.148	.141
	(1.78)**	(1.75)**	(1.18)	(1.11)
LAND	1.54E-3	1.48E-3	1.51E-3	1.03E-3
	(0.79)	(0.74)	(0.78)	(0.53)
EDCATION			-.0322	
			(1.76)**	
VARIBLTY				.484
				(1.83)**
R^2	.136	.136	.151	.152
Adjusted R^2	.092	.087	.103	.104

Table VI - CPVSNONC as dependent variable

CPVSNONC is the fraction of village loans from institutions that are guaranteed by one or more individuals other than the borrower.

|Tstats| in parentheses; significance at 85, 90, and 95% denoted by *, **, and ***, respectively.

Independent Variables	Baseline	-(BCOOPPCT)	+(EDCATION)	+(VARIBLTY)
GOODCORR	.249	.242	.225	.214
	(1.83)**	(1.78)**	(1.63)*	(1.57)*
BADCORR	-.0456	-.0439	-.0536	-.0948
	(0.33)	(0.31)	(0.38)	(0.67)
WEALTH	-1.18E-7	-1.17E-7	-9.53E-8	-1.10E-7
	(3.07)***	(3.05)***	(2.29)***	(2.88)***
WEALTHSQ	2.89E-15	2.84E-15	2.31E-15	2.66E-15
	(1.94)**	(1.91)**	(1.49)*	(1.79)**
STDDEVW	3.60E-8	3.56E-8	3.01E-8	3.45E-8
	(2.24)***	(2.22)***	(1.81)**	(2.16)**
SHAREREL	-.0603	-.0663	-.0506	-.0714
	(.86)	(.95)	(.72)	(1.02)
SHARENON	.0503	.0476	.0475	.0546
	(.80)	(.75)	(.86)	(.76)
BCOOPPCT	-.142		-.116	-.118
	(.82)		(.67)	(.68)
AGPCT	.221	.227	.165	.147
	(1.83)**	(1.89)**	(1.30)	(1.17)
LAND	1.35E-3	1.28E-3	1.35E-3	8.27E-4
	(0.70)	(0.66)	(0.70)	(0.43)
EDCATION			-.0262	
			(1.39)	
VARIBLTY				.507
				(1.95)**
R^2	.148	.145	.158	.166
Adjusted R^2	.100	.102	.105	.114

3 Conclusions

There is obvious difficulty in measuring prevalence of a cooperative, full risk-sharing, full-commitment borrowing setup compared with an individualistic setup where transfers and collusion are banned. In reality, there must exist a continuum of shades of these institutional setups.

However, difficulties notwithstanding, the empirical exercises we undertake contain some striking results. *In particular, the U-shaped relationship with wealth and the positive correlation with wealth inequality in the third set of regressions seem to be strong confirmation of the results of E. S. Prescott and Townsend (1999) on wealth and organizational choice.* These results are also born out in the second set of regressions, though with less significance due to the poorer measure of cooperative regime prevalence we use there.

Another interesting result is the significant positive effect on the cooperative regime, seen in all three sets of regressions, by a measure of correlation of output. This is, on the face of it, counter to the predictions of Holmström and Milgrom 1990. They show that higher correlation increases the amount of information a more individualistic regime can convey, making this setup more likely to be optimal to the principal. However, the explanation could be that in the data we examine, the agent, not the principal, is the one choosing the mechanism under which to borrow. For areas with high correlation, group-guaranteed loans may be more attractive, since joint liability payments are less likely: if one borrower fails, all are likely to fail. This begs the question of why the BAAC would primarily lend to these types. The answer could be political. It also may have to do with cross-regional insurance that the BAAC provides in the case of observable natural disasters such as floods.

Measures of cost of monitoring perform unreliably, particularly the measure of cooperation among villagers. *The sharing variables paint a consistent picture across all three sets of regressions. It appears that sharing among non-relatives encourages cooperative borrowing, while sharing among relatives discourages it.* Nothing in the theory sheds light on why this may be the case. One might hazard a guess that family risk-sharing networks need less of an external, institutional structure to thrive and are substitutes to non-family risk-sharing networks.[4] Thus non-family risk-sharing is enhanced by the organizational structure provided by joint liability lending institutions, while family risk-sharing is not. There could also be a relationship here to the theory involving output correlation. If much economic activity (including sharing) is along family lines, then possibly the correlation is higher and the individual setup is optimal; the opposite would be true for activity organized along non-family lines. The underlying assumption here is that blood relationships have more correlated shocks, for example illness.

Variability of income also seems to be a positive predictor of the cooperative regime. It is as if the cooperative regime is a better way to bring about insurance, which is more valuable the more risk is faced.

[4] This explanation assumes that non-cooperative loan arrangements are less capable as risk-sharing mechanisms than cooperative ones, an assumption not justified from the theory. However, typical individual loan contracts are also not as dependent on others' output as the theory would predict.