Group lending and individual lending with strategic default

Bharat Bhole *, Sean Ogden

Department of Economics, Rochester Institute of Technology, 92 Lomb Memorial Drive, Rochester, NY 14623-5604, United States

Abstract

Papers that compare group lending and individual lending in the presence of strategic default suggest that unless group members can impose costly social sanctions on one another, or unless the bank uses cross-reporting mechanisms group lending may do worse than individual lending. In this paper, we show that if, (1) the amount that a successful borrower owes for his defaulting partner is optimally determined, and (2) the penalty is allowed to vary across group members, then even in the absence of any social sanctions or cross-reporting, (1) expected borrower welfare is strictly higher with group lending when both group lending and individual lending are feasible and (2) group lending is feasible for a greater range of opportunity cost of capital. These results are robust to collusion between borrowers.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Conventional lending to the poor has traditionally been considered infeasible as a result of the riskiness of loans that are not secured with adequate collateral. In developing countries, this risk is exacerbated by the lack of sound legal infrastructure and credit scoring mechanisms. In such circumstances joint liability institutions, which lend to groups of people and where the entire group is considered responsible for default by any one member, have had some success in lending to the poor. The results, however, have been mixed. As Conning (1996) states (referring to Pitt and Khandker, 1996), “...group lending programs have been quite successfully implemented in the Cameroon, Malawi, South Korea, Malaysia and Bangladesh but similar schemes have had problems in India, Egypt, Venezuela, Kenya and Lesotho.” Karlan and Giné (2007) note that in the new system of Grameen Bank (dubbed Grameen II by the founder Muhammad Yunus) one of the changes has been to relax the joint liability aspect. The reward/punishment of a borrower now depends on both her performance and group performance as opposed to only the group performance.2

Focusing on the problem of strategic default, Besley and Coate (1995) and Armendariz de Aghion (1999) suggest a possible explanation for the mixed performance of group lending. Besley and Coate (1995) argue that unless social sanctions are sufficiently strong, group lending may encourage default by members who would have repaid under individual lending. On the other hand, if social sanctions are sufficiently strong, group lending can improve repayment rates by encouraging borrowers to help each other. Similarly, Armendariz de Aghion (1999) in her model with costly peer monitoring and social sanctions finds that if peer monitoring is not very costly and social sanctions are sufficiently strong, group lending can improve repayment rates by encouraging borrowers to help each other. Similarly, Armendariz de Aghion (1999) in her model with costly peer monitoring and social sanctions finds that if peer monitoring is not very costly and social sanctions are sufficiently strong, group lending can improve repayment rates by encouraging borrowers to help each other.

2 To quote Dowla and Barua (2006, p. 76), “Grameen II has also altered the loan ceiling policy. Instead of a common loan ceiling for all borrowers of a branch each borrower now [emphasis added] has his or her own loan ceiling. This ceiling is customized based on the performance of the borrower, her group, her center and the amount of deposits in her savings accounts.”

3 The intuition is the following: In a group of say two borrowers A and B, once A knows that B is going to default he compares the cost of defaulting (penalty imposed by the bank + social sanctions) with the entire group liability (which he is now responsible for) in deciding whether to repay. In some cases his individual liability is smaller than the penalty imposed by the bank, which is in turn smaller than the total group liability. In such a case, if social sanctions are not sufficiently strong, then a borrower who would not have defaulted under individual contract does so under the group contract.

* Corresponding author. Tel.: +1 585 475 7954; fax: +1 585 475 2510.
E-mail address: bharat.bhole@rit.edu (B. Bhole).

1 To quote Basu (2006, p. 31) referring to the situation in rural India, “Government has not been able to develop and enforce a legal and regulatory framework conducive to rural finance, so that contract design, contract renegotiation and contract enforcement remain weak, making it even more difficult for financiers to provide borrowers with the right incentives for the payment.”
sanctions are sufficiently large then the bank’s profit is higher with group lending; otherwise, individual lending does better.\textsuperscript{4} These papers suggest that a bank whose clients are unable to impose strong social sanctions on one another may be better off choosing individual liability over joint liability, at least when dealing with the problem of strategic default.\textsuperscript{5} In this paper we demonstrate that this is not necessarily true. In particular, we suggest a form of group lending contract in which the amount that a successful borrower owes for his defaulting partner is optimally determined and the penalty is allowed to vary across group members. We show that this flexible group lending contract has two advantages over individual lending. First, expected borrower welfare is higher with group lending for any given opportunity cost of capital for which both group lending and individual lending are feasible. Second, the range of opportunity cost of capital for which lending is feasible is also greater under group lending. Further, we show that these results are robust to what we refer to as weak collusion between borrowers.\textsuperscript{6} Our suggestion of greater flexibility in group lending is analogous to the relaxed form of group lending implemented in Grameen II, which involves differential treatment of group members who repay their loans and those who do not.

In Besley and Coate (1995) and Armendariz de Aghion (1999) group lending performs worse than individual lending due to certain restrictions placed on the group lending contract. In the former, for example, if one member defaults, the other member has to pay the full amount, $2R, owed by the group to avoid penalty ($R$ is the amount owed by each member). The amount that a member owes for his defaulting partner is not optimally determined. Armendariz de Aghion (1999) relaxes this restriction and solves for the optimal amount that a borrower owes for a defaulting partner. However, the penalty (which takes the form of cancellation of future financing) is restricted to be the same across members who meet their obligations and those who default, i.e., either all group members have access to future financing or every member’s access is cancelled.\textsuperscript{7} We refer to group lending contracts with these restrictions on payments and penalties as simple group lending contracts. We show that once these restrictions are relaxed, then the resulting flexible group lending contract performs unambiguously better than individual lending even in the absence of any social sanctions.\textsuperscript{8}

As in our paper, Rai and Sjostrom (2004) also relax the restrictions inherent in simple group lending contracts and find that group lending always does at least as well as individual lending. They design a clever cross-reporting mechanism, where each borrower’s payoff depends not only on the payments made by the borrower and his partner but also on the messages sent by unsuccessful borrowers. An important feature of their mechanism involves giving an unsuccessful borrower, say $i$, the ability to threaten her successful partner, say $j$, by saying “Borrower $j$, help me repay, or I will tell the bank that you refused to help me out and they will impose a harsh punishment on you (but not on me).”\textsuperscript{9} As Armendariz de Aghion and Morduch (2007) point out “...[the mechanism] works well on paper in a specific theoretical context”. However, in reality, microfinance institutions such as Grameen Bank and Spandana\textsuperscript{10} do not make use of cross-reporting; especially not in the manner suggested in Rai and Sjostrom (2004). They do not have in place any rewards or punishments that rely on messages sent by unsuccessful borrowers about other borrowers. For example, while Grameen II has joint liability aspects to it, such as the one mentioned in footnote 2, Grameen II members are not obliged to pay a member’s loan when the member fails to do so (see Rutherford et al., 2004). Rutherford et al. (2004) note that Dr. Yunus, the founder of Grameen Bank, emphasized that no member is ever required to pay for others.\textsuperscript{11} If a borrower is not required to pay for others, an unsuccessful borrower cannot threaten a successful borrower in the manner suggested in the Rai and Sjostrom mechanism.\textsuperscript{12}

A possible reason why a Rai–Sjostrom type mechanism may not be used in reality is the fear that it would create tension amongst borrowers. Referring to the Rai–Sjostrom mechanism, Armendariz de Aghion and Morduch (2007, p. 134) note “One fear is that formalizing such a system [of cross-reporting] can create tensions amongst individual borrowers or a strong incentive for them to collude.” Further, Rutherford et al. (2004) note “With Grameen II, and its emphasis on tension-free microcredit, it appears that Grameen is making even stronger efforts to ensure that strong forms of joint liability are a thing of the past.” Karlan and Giné (2007) also mention that excessive tension among members is often responsible for voluntary dropout. A Rai–Sjostrom type mechanism, which not only involves strong forms of group-liability but relies on borrowers threatening each other, is likely to cause even more tension among borrowers. Another possible problem with the mechanism is that it may discourage a borrower from sharing information about factors that have made her successful, with other borrowers. This is because by doing so she reveals that she is successful, which may be used against her.

Given these potential problems with the Rai and Sjostrom mechanism and the fact that there does not seem to be any practical experience with it, a bank that is interested in using microfinance to help the poor may have reservations about adopting it. If such a bank also finds lack of adequate social collateral among its potential clients, it may feel individual liability is superior to joint liability. Our paper shows that this is not the case. While the Rai and Sjostrom mechanism may be necessary for efficiency, we demonstrate that flexible group lending performs better than

\textsuperscript{4} In fact, in Aghion’s model, if we set both peer monitoring costs and social sanctions equal to zero then group lending always does worse.

\textsuperscript{5} Even in the absence of social sanctions, joint liability may mitigate costs arising due to other problems such as adverse selection (Chhatk, 2000; Armendariz De Aghion and Gollier, 2000; Laffont, 2003) and moral hazard (Stiglitz, 1990; Varian, 1990; Laffont and Rey, 2003).

\textsuperscript{6} Weak collusion refers to a situation where the borrowers know each other well enough to collude on their repayment amounts by making instantaneous transfers amongst themselves, when it is mutually beneficial to do so. They do not, however, trust each other enough to lend money in return for a promise of future repayment.

\textsuperscript{7} When peer-monitoring costs and social sanctions are both zero in Armendariz de Aghion (1999), group lending also does worse than individual lending because of the particular timing of the game. It requires that borrowers make their payment decisions before they learn the other borrower’s output. If a borrower is allowed to modify his decisions once he learns the other borrower’s output then performance under group lending can be improved making it the same as that under individual lending.

\textsuperscript{8} Note we are not saying that social sanctions do not improve the performance of group lending. It is certainly possible that their presence further improves the performance of group lending.

\textsuperscript{9} See Rai and Sjostrom (2004, p. 218).

\textsuperscript{10} Spandana is one of the largest microfinance institutions in India. As of April 2009, it operated in 9 states, had over 2.5 million clients and had outstanding loans of over $350 million US dollars (see http://www.spandanaindia.com/updateofoperations.htm).

\textsuperscript{11} See also the video entitled “Delinquency Management Process, Spandana, India” posted on the website http://www.indiamicrofinance.com/2009/03/microfinance-delinquency-management.html. It shows the procedures for loan recovery from defaulting borrowers at the microfinance institution Spandana (see footnote 10). The bank first tries its best to recover the loan from the defaulting borrower. If the bank is unsuccessful in doing so, it tries to recover the loan from other group members. However, at no point does it ask the defaulting borrower any information about the other successful borrowers. Hence, the question of the successful borrower’s penalty depending on messages of the unsuccessful borrower does not arise.

\textsuperscript{12} Rutherford notes that some Grameen Bank officials have not embraced the new rules completely. They refuse to end the group meeting until the complete group amount is repaid. However, even in this case, pressure is put on those who are known to the bank to have cash on the day of the meeting. This includes people who are due to get a loan that day and who have large deposits to make. Neither the defaulting nor the successful borrowers are rewarded or penalized based on any information provided by the former. In fact, no information about the other borrowers is obtained from those who are unable to pay.
individual lending even without such a mechanism and without social sanctions.

Another important difference between our paper and Armendariz de Aghion (1999) and Rai and Sjostrom (2004) is that in these papers the maximum penalty that can be imposed on borrowers is exogenously given; it is not affected by the contractual terms. In our paper the maximum penalty is endogenously determined. We consider an infinite period model where the penalty is in the form of denial of future loans. As a result, the maximum penalty that can be imposed is the continuation payoff from access to future loans, which is endogenously determined. This maximum level of penalty determines the maximum level of repayment that can be extracted from a borrower. The maximum repayment amount, in turn, determines the maximum opportunity cost of capital for which lending is feasible. This allows us to meaningfully compare individual lending and group lending not only with respect to borrower welfare when both are feasible, but also with respect to the maximum opportunity cost of capital for which lending is feasible.

The rest of the paper is organized as follows: In Section 2 we present our basic model. In Sections 2.1 and 2.2 we derive the optimal individual lending contract and the optimal simple group lending contract respectively. Then we compare these two. As in Besley and Coate (1995), we find that for some parameter values individual lending is better while for others simple group lending is better (although instead of repayment rates we compare these on the basis of borrower welfare). Section 3 studies the optimal flexible group lending contract. In Section 3.1 we consider the case where borrowers cannot collude; in Section 3.2 borrowers are allowed to collude on their repayment amounts. We show that in each case welfare is unambiguously higher with the flexible group lending contract than with the individual lending contract. We also compare welfare under flexible group lending with that under the Rai–Sjostrom mechanism in these sections. In Section 4 we present concluding remarks. The proofs of all propositions are in Appendix A. Appendix B presents technical details involving comparison with the Rai–Sjostrom mechanism.

2. The model

Consider a setting with a benevolent bank and many risk neutral cashless microentrepreneurs. Each entrepreneur is interested in carrying out a project that requires an investment of $l$ per period. It is assumed that entrepreneurs do not accrue savings and hence any entrepreneur needs to borrow $l$ from the bank in every period that she wants to carry out the project. The output $\theta$ per project is identically and independently distributed across entrepreneurs and periods: $\theta \in (0, H)$ with $Pr(\theta = H) = v$ and $Pr(\theta = 0) = (1 - v)$, i.e., either the project succeeds and yields $H$ or it fails and the output is zero. In any period in which an entrepreneur does not invest in the project she receives a utility normalized to zero.

Let $H$ denote the bank's opportunity cost of capital for a project financed in any period. It is assumed that $vH > I > 0$, i.e., investment in the project is socially desirable. As in Rai and Sjostrom (2004), we assume that the bank is a non-profit organization or is operated by a benevolent government and is therefore interested in maximizing expected lifetime utility of borrowers subject to itself breaking even. Borrowers are interested in maximizing their own utility. Let $\delta \in (0,1)$ denote their discount factor.

It is assumed that borrowers can costlessly observe each other's output realization, but the bank cannot observe whether a borrower has succeeded or failed. This implies that when a borrower defaults on a payment, the bank cannot ascertain whether it was a strategic default or whether it was due to project failure. Also, it is assumed that if a borrower claims that she is unable to repay, the bank cannot forcefully extract any payment from her. Therefore, the only way borrowers can be induced to repay their loans is through the threat of cancellation of future financing. As in Armendariz de Aghion (1999) and Tedeschi (2006), we assume that this threat takes the form of a probability with which a borrower is excluded from future financing; henceforth referred to as probability of cancellation. The precise nature of the bank's problem (the variables to be solved for and the constraints faced) depends on whether we are considering individual lending, simple group lending or flexible group lending and on whether borrowers collude or not. Below we briefly explain each of these and describe the variables to be solved for in each case. The timing, objective function and the constraints are explained in detail in the relevant subsections where we solve for the optimal contracts.

In individual lending (IL) the payment owed by a borrower is independent of the performance and repayment decisions of the other borrowers. Also, the probability with which future financing is cancelled depends only on the borrower's own repayment decision. Let $R_i$ denote the repayment amount; and let $q$ (respectively, $p$) denote the probability of cancellation when the borrower repays (respectively, does not repay) $R_i$. We assume that once the relationship is terminated, the bank does not lend to that borrower again.

In simple group lending (SGL), borrowers have to apply for loans in groups comprising two members. Each borrower gets a loan of $\frac{l}{2}$ and has to repay $R_G$ to the bank. If a member defaults on his loan then his partner owes $2R_G$ to the bank, i.e., the complete liability of a defaulting member is transferred to the other member. If the bank receives a total payment of $2R_G$, the group then future financing is cancelled for both members with probability $q$. If the total payment received falls short of $2R_G$, future financing is cancelled (again for both members) with probability $p$.

As in simple group lending, flexible group lending (FGL) requires borrowers to apply for loans in groups comprising two members. Also, as in SGL, the bank stipulates a payment of $R_G$ per borrower. If this is received, future financing is cancelled for each borrower with probability $q$. However, now if a member defaults on his loan, his partner owes an amount $R_G - R_F$ to the bank which is not restricted to be $2R_G$; it is optimally determined by the bank. When one borrower defaults and the other borrower repays $\bar{R}$, the bank excludes the defaulting (respectively, solvent) borrower from future financing with probability $p$ (respectively, $q$). That is, the penalty is allowed to differ across the defaulting and the solvent borrowers. If the total payment falls short of $\bar{R}$ then both borrowers lose access to future financing with probability $p$. In each case (individual lending, simple group lending and flexible group lending) we restrict attention to stationary contracts, i.e., the bank's choice variables are assumed to be independent of the period and the past history.

2.1. Individual lending

The timing in individual lending is as follows. Each period consists of three stages, $s = 0, 1, 2$. At $s = 0$, the bank offers a loan contract $(l, R_i)$.
max \( q, p \) to the borrower. At \( s=1 \), the borrower invests \( I \) and realizes output \( \theta \in [0, H] \); finally, at \( s=2 \), she makes her repayment decision. If the repayment amount is less than \( R_I \) the bank excludes the borrower from future financing with probability \( p \); otherwise, financing is cancelled with probability \( q \). If the bank decides to continue lending to the borrower, these stages are repeated until the borrower gets her reservation utility, zero, in that and all subsequent periods.

Let \( M \) denote the borrower’s expected lifetime utility from any period onwards in which she gets the loan.\(^{17}\) Given that the borrower discounts future payoffs at the rate \( \delta \), the bank’s problem is to:

\[
\begin{align*}
\max & \quad M \quad q, p, R_I \\
\text{s.t.} & \quad R_I \leq (p - q)\delta M \\
& \quad R_I \leq H \\
& \quad v_R I \geq I \\
& \quad 0 \leq q \leq 1 \quad \text{and} \quad 0 \leq p \leq 1.
\end{align*}
\]

Constraint (1) ensures that a borrower has an incentive to repay her loan when successful. The LHS is the amount that the borrower has to repay. The RHS is the expected increase in punishment.\(^{18}\) If the bank does not renew the contract, the borrower gets her reservation utility, zero. Setting \( I < H \) is feasible only if \( I \) is less than the maximum probability \( q \) that the borrower will repay.\(^{19}\) Setting \( I \) equal to \( H \) implies constraint (2) and hence (3) (along with (1)) imply constraint (5) since \( q = 1 \) and \( p \leq 1 \) is feasible.

Lending is feasible only if this maximum amount, \( v_R I \geq I / \nu \) (see constraint (3)), or only if \( I \leq v_R H \). Hence, given other things, the higher is \( \nu \), or \( \delta \), or \( H \), the larger is the maximum opportunity cost of capital for which lending is feasible. The intuition is straightforward. The larger is any of these parameters, the more attractive is future refinancing and therefore more potent is the threat of termination of the lender–borrower relationship, enabling the bank to extract a higher repayment in the current period.

Note that while the above proposition is stated and proven for the case where there is no collusion between the borrowers, the result remains the same even if we allow for weak collusion. Since any borrower’s payoff is completely independent of the other borrower’s performance and there are no messages, there is no incentive for them to collude on their repayment amount by making instant transfers among themselves.

2.2. Simple group lending

We now solve for the simple group lending contract. The timing is as follows: each period consists of three stages, \( s=0–2 \). At \( s=0 \), the bank offers a loan contract \((l, R_c, q, p)\) to each borrower. At \( s=1 \), each borrower invests \( I \) and realizes output \( \theta \in [0, H] \). At \( s=2 \), the borrowers simultaneously make their repayment decisions. If the total repayment amount does exceed \( 2R_c \), the group is excluded from future financing with probability \( q \). If the total repayment, \( R \), is less than \( 2R_c \), the bank asks the borrower who has repaid the higher amount whether he wants to pay \( 2R_c - R \).\(^{20}\) This borrower decides whether to make up for the shortfall. If she chooses to do so, the bank excludes the group from future financing with probability \( q \); if not, i.e., if she only pays additional amount \( A - 2R_c - R \), then the bank sends a request for \( 2R_c - R - A \) to the other borrower. If the other borrower complies (resp. does not comply) the bank cancels future financing with probability \( q \) (resp. \( p \)).\(^{20}\)

To ensure that each borrower paying \( R_c \) when both are successful is a subgame perfect equilibrium, the bank needs to satisfy:

\[
R_c \leq (p - q)\delta M
\]

where, as before, \( M \) is the expected lifetime utility of a borrower any period onwards in which she gets financing. To induce a successful borrower to pay \( 2R_c \) when her partner is unsuccessful, it is necessary that\(^{21}\)

\[
2R_c \leq (p - q)\delta M.
\]

Clearly, constraint (6) implies constraint (5) and hence (6) (along with the described timing) also ensures that each borrower paying \( R_c \) when both are successful is a subgame perfect equilibrium. Constraint (5) implies that a borrower’s expected lifetime utility (exploiting the

\[^{17}\] Given that we have normalized the reservation utility to zero, the borrower’s expected lifetime utility from any period onwards in which she does not get the loan, is zero.

\[^{18}\] This can be easily seen by differentiating the optimal value of \( p^* \) with respect to \( I \). We get, \( dL(\delta(1-\delta)/(\theta^2H-\delta^2)) = \delta(1-\delta)\nu^2H/(\theta^2H-\delta^2) > 0 \).

\[^{19}\] If each borrower has contributed the same amount then one borrower is chosen randomly.

\[^{20}\] Sending a payment reminder and giving the borrowers an opportunity to repay if they have not done so rules out asymmetric and/or multiple subgame perfect equilibria. For example, if both constraints (5) and (6) are satisfied and if the bank terminates future financing with probability \( p \) whenever total repayment amount falls short of \( 2R_c \), then \( 2R_c=0 \) is a subgame perfect equilibrium in state \((H,H)\). In such cases, asymmetric equilibria is the only equilibrium in Basley and Coate (1995) in some cases where both borrowers are successful. By inclusion of reminders we ensure that \((R_c, R_c)\) is the unique equilibrium in this state. Similarity, if only constraint (5) is satisfied then without reminders both \((R_c, R_c)\) and \((0,0)\) are subgame-perfect equilibria when both borrowers are successful. Payment reminder from the bank rules out the latter equilibrium and ensures uniqueness.

\[^{21}\] When a borrower’s partner is unsuccessful, she needs to pay \( 2R_c \) to be considered solvent.
stationary structure of the model) is: \( M = v^2(2H - R_H + (1 - q)\delta M) + v(1 - v)(2 - 2R_C + (1 - q)\delta M) + (1 - v)v(1 - q)\delta M + (1 - v)^2(1 - p)\delta M \).\(^{22}\) Using this to solve for \( M \) we can write the bank's problem as:

\[
\max_{R_C, q, p} M = vH - (2v - v^2)R_C \\
\text{s.t. } R_C \leq \frac{1}{2}(p - q)\delta M \\
2R_C \leq H \\
(2v - v^2)R_C \geq I \\
0 \leq q \leq 1 \quad \text{and} \quad 0 \leq p \leq 1.
\]

Constraint (8) is the incentive constraint obtained from (6) by solving for \( R_C \). Constraint (9) ensures that the stipulated repayment amount for a successful borrower does not exceed her output. Finally, (10) is the bank's break-even constraint. It says that the expected repayment amount per borrower should be at least as large as \( I \).

**Proposition 2.** If \( \delta \) and \( v \) are such that \( 2 - v\delta(4 - v) \geq 0 \), then simple group lending (SGL) is feasible for any \( I \leq (2 - v)\delta H/(2 - 2v)\delta v \). Otherwise, SGL is feasible only if \( I \leq (2 - v)\delta H/2. \) For any feasible \( I \), repayment amount per borrower, \( R_{SC} = 1/(2 - v)\delta H \), and the probability of cancellation of future financing when the group is solvent is \( q_{SC} = 0 \), and when it defaults is \( p_{SC} = 2(1 - \delta)/(2 - 2v)\delta H - \delta(1 + (1 - v)^2)I \). Each borrower's expected lifetime utility is \( U_{SC} = -(vH - I)/(1 - \delta + \delta(1 - v)^2p_{SC}) \).

As in the case of individual lending, if the group meets its obligation then the contract is renewed with certainty. If not, then future financing is cancelled for both borrowers with a positive probability (\( p_{SC} = 0 \)) that is set just large enough to induce the borrowers to pay up to \( 2R_{SC} \) when successful. This probability increases as the required repayment increases. Note that the results obtained in this proposition remain valid even if we allow for weak collusion between the borrowers. Since under SGL, the total repayment amount in state \((\theta_b, \theta_b)\) is the same as the total repayment amount in state \((\theta_h, \theta_h)\), there is no incentive for the borrowers to collude in state \((\theta_b, \theta_b)\). And, in state \((\theta_h, 0)\) there is no possibility of collusion since the unsuccessful borrower has no resources.

**Propositions 1 and 2** can be used to compare individual lending and SGL along two dimensions: (i) Feasibility — finding which of these two is feasible for a greater range of \( I \); and (ii) Borrower welfare — when both are feasible, finding which one yields a greater expected lifetime utility for the borrower. Doing so, gives us the following proposition (see also Figs. 1 and 2).

**Proposition 3.** (a) When both individual lending and SGL are feasible, borrower welfare is higher with simple group lending than with individual lending. (b) If \( v \) is sufficiently large (in particular, \( v \) is such that \( 6v - v^2 - 4 \geq 0 \)) then individual lending is feasible for a greater range of opportunity cost of capital. (c) If \( v \) is such that \( 6v - v^2 - 4 \geq 0 \) then we have the following: for \( \delta \in (1/2 - v), (2 - v)/(2\delta), \) SGL is feasible for a greater range of \( I \). Finally, for \( \delta \in ((2 - v)/(2\delta), 1) \), again individual lending is feasible for a greater range of \( I \).

\(^{22}\) With probability \( v^2 \) both borrowers are successful, each pays \( R_c \), the group is solvent and future financing is renewed with probability \( (1 - q) \); with probability \( v(1 - v) \) the borrower succeeds but her partner fails, she pays \( 2R_c \), again the group is considered solvent and future financing is renewed with probability \( (1 - q) \); with probability \( (1 - v)^2 \) both borrowers fail, the group defaults and the contract is cancelled only with probability \( (1 - p) \).

The reason SGL yields a greater borrower welfare than individual lending when both are feasible is the following: since the bank is interested in breaking even, the expected payment of each borrower is the same, \( I \), in both SGL and IL. However, under IL a borrower's contract is cancelled whenever she fails, whereas in SGL the borrower's contract is cancelled only when both fail. Since the likelihood of both borrowers failing at the same time is smaller than the borrower alone failing, financing is cancelled less often under SGL resulting in a higher expected payoff.

We also see in the above proposition that when \( v \) is not very large, SGL is feasible for a greater range of \( I \) for intermediate values of \( v \) but not for extreme values of \( \delta \). The intuition is the following. The maximum \( I \) that is feasible under either mechanism is obtained when \( p \) in that mechanism is one. Suppose that this is the case for both SGL and IL. Now note that the maximum payment required under SGL, \( 2R_{SC} \), is greater than that required under individual lending, \( R_i \) (because in some states in SGL a borrower has to pay both for herself and her partner, whereas in IL she only pays for herself). With \( p_{SC} = p^H = 1 \), the former is constrained by \( M_{SC} \) and the latter is constrained by \( M_{IL} \) where \( M_{SC} \) (resp., \( M_{IL} \)) is the expected lifetime utility of each borrower under SGL (resp. individual lending) if the contract is renewed in the next period. When \( \delta \) is small the difference \( M_{SC} - M_{IL} \) is small. Consequently, the maximum payment that can be extracted under SGL is not much larger than that under individual lending, i.e., the difference between \( 2R_{SC} \) and \( R_{IL} \) is small. This makes the overall expected payment smaller under SGL.\(^{24}\) As \( \delta \) increases, the difference \( \delta(M_{SC} - M_{IL}) \) increases enabling the maximum payment to be sufficiently high under SGL such that the overall expected payment under SGL can be higher. After a point, however, as \( \delta \) becomes sufficiently large the limited liability constraint under SGL binds \( 2R_{SC} \) and \( R_{IL} \) cannot be increased further. But \( R_{IL} \) can be increased, which results in lending being feasible for a greater range of \( I \) under IL.

Our finding here that simple group lending can either improve or worsen borrower welfare is similar to that of Besley and Coate (1995) who find that group lending can improve or worsen repayment rates. This conclusion is reached in Besley and Coate and in this section in our paper due to the following two restrictions on the group lending.

---

23 Strictly speaking, conditional on both borrowers failing, the contract is cancelled with a higher probability under SGL than under IL (i.e., \( p_{SC}^H > p^H \)). However, it is still not high enough for the overall cancellation probability to be higher under SGL.

24 In state \( (H, H) \) the bank receives \( 2R_c \) (resp. \( 2R_h \)) under simple group lending (resp. individual lending); and in state \( (H, 0) \) the bank receives \( 2R_c \) (resp. \( R_h \)) under simple group lending (resp. individual lending). When \( 2R_c \) is only slightly larger than \( R_h \), the advantage under SGL in state \( (H, H) \) is small, but the disadvantage in state \( (H, 0) \) is large \( 2R_c \) (is substantially smaller than \( 2R_h \)) resulting in an overall smaller expected payment under SGL.
contract: (i) if one borrower defaults, the entire debt obligation of the defaulting borrower is transferred to the other borrower (the amount owed by the other borrower is not optimally determined); and, (ii) the probability with which financing is cancelled is identical for both the borrowers; the bank cannot differentiate between a borrower who meets her obligation (including the amount owed for the other borrower) and the borrower who does not repay. In the following section we show that when these restrictions are relaxed, the resulting flexible group lending does strictly better than individual lending even for those parameter values for which we found individual lending to be superior.

3. Flexible group lending

To keep the exposition simple we assume that borrowers either pay an amount specified in the loan contract or they pay nothing. The timing is as follows: At stage \( s = 0 \), the bank offers a contract \((i, R_C, \hat{R}, q, \hat{q}, \hat{p}, p)\) with \( R = (1 + \eta)R_C \) to each borrower, where \( \eta \in [0, 1] \) (or \( \hat{R} \in [R_C, 2R_C] \)).\(^{25}\) At \( s = 1 \), each borrower invests \( l \) and realizes output \( \theta \in [0, H] \). At \( s = 2 \), the borrowers simultaneously make their repayment decision. Let \((R_i, \hat{R}_i)\) denote the amounts paid by borrowers \( i \) and \( j \) respectively, where \( i, j \in \{1, 2\} \) and \( i \neq j \). At \( s = 3 \), depending on these repayment amounts, the bank makes its decision in the following manner:

- If \((R_i \geq R_C, R_j \geq R_C)\) financing is cancelled for each borrower with probability \( q \);
- If \((\hat{R} = R_C, \hat{R} = 0)\) then the bank asks borrower \( j \) whether she wants to pay \( R_C \). If she complies, future financing is cancelled with probability \( q \) for each borrower. If not, then future financing is cancelled with probability \( p \) for each borrower;
- If \((R_i = \hat{R}, R_j = 0)\) future financing for borrower \( i \) (resp. \( j \)) is cancelled with probability \( \hat{q} \) (resp. \( \hat{p} \));
- If \((R_i = 0, R_j = 0)\) or for any other combination of \((R_i, R_j)\) not mentioned above, future financing is cancelled for each borrower with probability \( p \).

The bank’s problem is to determine \((R_C, \hat{R}, q, \hat{q}, \hat{p}, p)\) such that borrowers pay \((R_C, \hat{R}_C)\) in state \((H, H)\) and \((\hat{R}, 0)\) in state \((H, 0)\) (or \((0, \hat{R})\) in state \((0, H)\)), the bank breaks even, and the borrowers’ expected lifetime utility is maximized. Note that since the bank’s response when it receives payment \((R_C, 0)\) is different from its response when it receives \((\hat{R}, 0)\), we must have \(\hat{R} = R_C\) (or \(\eta > 0\)). The constraints faced by the bank depend on whether borrowers collude or not and the form of collusion. We first consider the case of no collusion, followed by the case of weak collusion. In weak collusion borrowers are able to collude on their payment decisions by making instant transfers but they do not trust each other enough for the successful borrower to lend money to the unsuccessful one.

3.1. Flexible group lending with no collusion

When borrowers do not collude with each other, the bank’s problem is:

\[
\max_{R_C, \hat{R} \in [0, \hat{R}]} M
\]

s.t. \( H - R_C + \hat{q}(1 - q)M \geq H + \hat{q}(1 - p)M \)

\( H - \hat{R} + q(1 - \hat{q})M \geq H - \hat{R} + q(1 - \hat{q})M \)

\( \hat{R} = (1 + \eta)R_C \leq H \)

\( v^2R_C + v(1 - v)\hat{R} \geq l \)

\( R_C - \hat{R} \leq 2R_C \)

\( H - 2R_C + q(1 - q)M \geq H - \hat{R} + q(1 - \hat{q})M \)

\( 0 \leq q \leq 1, 0 \leq p \leq 1, 0 \leq \hat{q} \leq 1 \) and \( 0 \leq \hat{p} \leq 1 \).

Constraints (13) and (14) are similar to constraints (5) and (6). The former ensures that when both borrowers are successful then each pays \( R_C \) and the latter ensures that when one borrower fails, the other (successful) borrower pays \( \hat{R} \). Constraints (13) and (15) rule out payment of \((\hat{R}, 0)\) as an equilibrium when both borrowers are successful.\(^{26}\) Constraint (16) ensures that borrowers are able to make the required payments when successful; and constraint (17) ensures that the bank breaks even. Constraint (18) has been explained above. Finally, constraint (19) ensures that the successful borrower does not pay \(2R_C\) instead of \(\hat{R}\) in state \((H, 0)\).\(^{27}\) Solving this problem we get the following result.

**Proposition 4.** With no threat of collusion between borrowers, lending is feasible for any \( I < v^2\delta H/(1 - \delta v + \delta v^2) \). For any feasible \( I \), probability of cancellation of future financing, \( q^\infty = \hat{q}^\infty = p^\infty = 0 \) and

\[
p^\infty = \frac{I(1 - \hat{q})(1 + \eta)}{(1 + \eta - \eta v^2)\delta H - (1 - v)(1 + \eta) + v^2\delta I}
\]

where \( \eta \) is sufficiently small. Payment \( R_C \) induced per borrower when both are successful is \( R_C^\infty = I(1 + \eta)/v(1 + (1 - v)\eta) \); and payment induced from the successful borrower (when only one borrower is successful) is \( R_C^\infty = I(1 + \eta)/v(1 + (1 - v)\eta) \). The borrower’s expected lifetime utility is \( M_{B^\infty}^C = (vH - I)/(1 - \hat{q} + \delta(1 + \eta)^2)\).

\(^{25} \) The assumption that \( \hat{R} \in (R_C, 2R_C) \) is without any loss of generality. Even if we do not restrict \( R \) to be in this range, the optimal value ends up being in this range. Showing this formally however requires extra steps which make the proofs longer and more difficult to read without adding anything. Also, restricting \( \eta \) to \([0, 1]\) lends it the natural interpretation as the fraction of defaulting borrower’s debt that a successful borrower owes to the bank if she is to be considered solvent.

\(^{26} \) A borrower knows that if she pays \( R_C \) and her partner pays anything less than \( R_C \), the bank will give the latter a chance to repay the difference. Constraint (13) ensures that a successful partner will repay the difference. Hence, by paying \( R_C \), the borrower gets \( H - R_C = (1 - \hat{q})\delta M \), whereas by paying \( \hat{R} \) she gets \( H - \hat{R} + \hat{q}(1 - \hat{q})M \). Constraint (15) ensures that the borrower is better off paying \( R_C \).

\(^{27} \) Having \( R < 2R_C \) is not sufficient to guarantee this because \( q \) and \( \hat{q} \) are not known yet. Even with \( R < 2R_C \), if \( q < \hat{q} \), the borrower may be induced to pay \( 2R_C \) instead of \( R \).
One feature of this proposition is noteworthy. When only one borrower succeeds, the bank does not induce this borrower to pay the entire debt owed by the defaulting borrower ($\tilde{R} < 2R_c$), i.e., the bank does not induce mutual insurance. In fact, inducing mutual insurance (by setting a larger $\eta$) reduces both, each borrower’s expected lifetime utility and the range of investment costs for which lending is feasible.\(^\text{28}\) To see the intuition consider the following: when a borrower has to pay a greater fraction of the loan owed by her partner, her repayment amount, $\tilde{R}$, increases. To ensure that this incentive compatible requires an increase in the expected penalty of the successful borrower when she does not pay, which requires an increase in the probability with which future financing is cancelled when repayment is $(0,0)$. With overall expected payment in any period held the same (it is equal to $I$) this increase in probability reduces a borrower’s expected lifetime utility. Also, the fact that a higher probability of cancellation is required to support any given $I$ means that as $I$ increases, the probability of cancellation hits the bound of one for a smaller $I$ than when $\eta$ was lower, thereby decreasing the range of investment cost for which lending is feasible.

The reason the bank sets $\tilde{R} = R_c$ (but as close as $R_c$ as possible), is to learn about the defaulting borrower’s outcome when one borrower pays and the other does not. Through the payment of the successful borrower (whether she pays $R_c$ or $\tilde{R}$) the bank can ascertain whether the other borrower’s default is strategic or due to his project failure. This enables the bank to avoid penalizing the latter when he is truly unable to pay ($p=0$ but $\tilde{b}=0$). Hence, as far as the repayment amounts are concerned, the optimal IL contract is very similar to the optimal IL contract. However, the probability of renewal is different. Under no collusion, with FGL, a borrower who fails but has a successful partner gets his contract renewed with probability $\eta^{P2} > 0$ regardless of whether the other borrower is successful or fails.

Comparing Propositions 1 and 4 it can be seen that when there is no threat of collusion, the optimal group lending contract does better than individual lending for all parameter values. First, consider feasibility. Group lending is feasible for any $I < v^2\delta H / (1 - \delta v + \delta v^2)$; whereas individual lending is feasible for any $I < v^2\delta H$. Comparing the two terms on the right hand side, it is clear that group lending is feasible for a greater range of $I$. Second, consider a borrower’s welfare when both individual and group lending are feasible. Comparing her expected lifetime utility under individual lending, $M_{\text{IL}}$, with her utility under individual lending, $M_{\text{IL}}$, we get that the former is larger provided $(1 - V)\eta^{NC} < \eta^{P2}$. Using the equilibrium values of these probabilities, $M_{\text{IL}}$ is larger than $M_{\text{IL}}$ provided

$$
(1 - V) \frac{I(1 - \delta)(1 + \eta)}{(\delta^2(1 + \eta) - \delta v^2 H)(1 - (1 - V) + \delta v^2 I)} < (1 - \delta) \frac{I}{(\delta^2 H - \delta I)}
$$

As $\eta \rightarrow 0$, the left hand side of the this inequality goes to $(1 - V)$ $(1 - \delta) I / (\delta^2 H - \delta I)$, since this last term is smaller than $(1 - \delta) I / (\delta^2 H - \delta I)$, we know that by choosing $\eta$ small enough the bank can attain higher borrower welfare under group lending. Essentially, group lending does better because it enables the bank to avoid penalizing a group member when she fails but her partner succeeds, because the successful member reveals the failed member’s true output through her payment choice. In individual lending a borrower is penalized with a positive probability whenever she fails. Summarizing this discussion, we have:

**Proposition 5.** In the absence of collusion between borrowers, the flexible group lending contract yields higher welfare than individual lending for any $I$ for which both are feasible. Moreover, the flexible group lending contract is feasible for a greater range of the opportunity cost of capital.

### 3.1.1. Comparison with the Rai–Sjostrom mechanism

**Rai and Sjostrom** (2004) have shown that if the bank uses cross-reporting (and, hence, does not rely on repayment amounts alone to ascertain borrowers’ success or failure) then it can achieve the first best in the absence of collusion. It is interesting to know what proportion of the borrower welfare in the Rai–Sjostrom mechanism is secured by flexible group lending. Henceforth, we refer to this proportion as the efficiency ratio. Let $W_{\text{NC}}$ (resp. $W_{\text{ES}}$) denote the expected welfare of a borrower in any given period in the FGL (resp. Rai–Sjostrom) mechanism. Then, in the absence of collusion between borrowers, the efficiency ratio is given by $\text{ER}\text{NC} = (v H - (1 - v + v^2)I) / (v H - v I)$.\(^\text{29}\) It can be checked that this ratio increases with $v$. The intuition is the following: any difference in welfare between FGL and the Rai–Sjostrom mechanism occurs only in state $(0,0)$.\(^\text{30}\) Greater is $v$, the smaller is the likelihood that state $(0,0)$ is realized. To get an idea about the size of this ratio, in Table 1 we present some numerical values corresponding to different parameter combinations. Specifically, we choose three different values of $v$, $(v = 0.85, 0.9$ and $0.95)$, $\delta = 0.9$, $I = 1$ and $H$ equal to the minimum value required for lending to be feasible.\(^\text{31}\) This requires $H$ to be slightly higher than $(1 - \delta v + \delta v^2 I) / \delta v^2 I$; we choose $H$ to be 0.01 more than this quantity. We find that when $v = 0.85$, FGL secures 84% of the welfare attained under the Rai–Sjostrom mechanism. When $v = 0.95$, the efficiency ratio is about 98%. This suggests that if borrowers cannot collude and $v$ is sufficiently large, then the optimal FGL mechanism comes quite close to the efficient Rai–Sjostrom mechanism.

### 3.2. Flexible group lending with weak collusion

Borrowers may have a strong incentive to collude especially when the difference between $\tilde{R}$ and $R_c$ is small. Even if mutual trust is not sufficiently strong for the borrowers to lend money to each other or accept a promise of future payment, they will most likely know each other well enough to collude on their payments to the bank by making instant transfers among themselves when it is mutually beneficial to do so. For example, with the contract described in Proposition 4 the borrowers can obtain an extra $2R_{\text{NC}}^\text{WC} - \tilde{R}_{\text{NC}}$ for the group by colluding in state $(H, H)$ and paying $(\tilde{R}_{\text{NC}}, 0)$ instead of $(R^\text{ES}, R^\text{NC})$. While the contract of the previous subsection is not robust to such collusion, we show here that the bank can nevertheless design a collusion-proof group lending contract which performs better than individual lending.\(^\text{32}\)

\(^\text{28}\) See Appendix B for a derivation.

\(^\text{30}\) The repayment amount plus the punishment of a borrower is the same in other states.

\(^\text{31}\) Note that the efficiency ratio is increasing in $H$; hence, choosing higher values for $H$ will result in higher efficiency ratios than the ones we find.

\(^\text{32}\) The notion of weak-collusion referred to above is very similar to the idea of interim side contracting in **Rai and Sjostrom** (2004); the only difference being that in **Rai and Sjostrom** (2004) in addition to the repayment amounts, the borrowers can also collude on messages. By assumption there are no messages in our model and hence the borrowers can collude only on the repayment amounts.

---

**Table 1**

<table>
<thead>
<tr>
<th>$v$</th>
<th>ER\text{NC}</th>
<th>ER\text{ES}</th>
<th>ER\text{WC}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.85</td>
<td>84.0%</td>
<td>96.7%</td>
<td>94.8%</td>
</tr>
<tr>
<td>0.90</td>
<td>92.3%</td>
<td>98.6%</td>
<td>98.1%</td>
</tr>
<tr>
<td>0.95</td>
<td>98.0%</td>
<td>99.7%</td>
<td>99.6%</td>
</tr>
</tbody>
</table>

\(^\text{4}\) $\delta = 0.9$, $I = 1$ and $H$ is chosen to be the minimum value for which lending is feasible.

\[^\text{29}\] This follows from $dP^{\text{NC}} / d\eta > 0$ (which implies $d[M^{\text{NC}}] / d\eta > 0$) and $d[\eta^2 I (v - \eta + 1) / (1 - \delta v (1 - \eta + \eta v (1 - \delta v))) / d\eta > 0$, where $(\delta^2 H (v - \eta + 1) / (1 - \delta v (1 - \eta + \eta v (1 - \delta v))))$ is the maximum feasible $I$ for any given $\eta$.\)
As above, let \( (I, R_C, \tilde{R}, q, \tilde{q}, \tilde{p}, p) \) be the contract offered by the bank. To ensure that borrowers do not collude in state \((H, H)\) to pay \((\tilde{R}, 0)\) instead of \((R_C, R_C)\), in addition to constraints \((13)-(20)\) we need:

\[
2H + 2(1-q)\tilde{q}M - 2R_C \geq 2H + ((1 - \tilde{q}) + (1 - \tilde{p}))\tilde{q}M - \tilde{R} \quad (21)
\]

The LHS is the group's aggregate payoff when the borrowers pay \((R_C, R_C)\) and the RHS is their payoff when they pay \((\tilde{R}, 0)\). Constraint \((21)\) ensures that gain from collusion is smaller than the loss suffered, thereby preventing any mutually beneficial exchange.\(^33\) The other constraints and the timing remain the same as in the no collusion case. Hence, the bank's problem is:

\[
\max_{\tilde{R}, \tilde{q}, \tilde{p}, p} M \quad \text{s.t.} \quad (13)-(21).
\]

Solving this problem we get:

**Proposition 6.** If borrowers can collude on their repayment decision and if,

i. \( \tilde{q} \leq 2/\sqrt{4 - v} \)

ii. \( I \leq (2-v)\tilde{v}\delta H/(2 - (2-v)\tilde{v}) \) the flexible group lending contract is the same as the optimal SGL contract; ii.\( I < 1/(2-v) \) and \( I \) in the interval \( ((2-v)\tilde{v}\delta H/(2 - (2-v)\tilde{v}), \tilde{v}^2\delta H) \); \( q^m = \tilde{q}^m = 0 \);

\[
\tilde{p}^{wc} = \frac{2(1-v)^2\delta H - (2 - (2-v)\tilde{v})\tilde{v}\delta I}{(2-v)\tilde{v}^2\delta H - v\delta I} \quad (22)
\]

\( \tilde{p} > 2/(4-v) \) then,

i. \( I \leq (2-v)\tilde{v}^2\delta H/2 \) the optimal contract is the same as the simple group lending contract; ii. \( I \geq (1-v)/(2-v)\tilde{v}\delta H/2 \); \( p^{wc} = (1 - \tilde{q})\tilde{v}\delta H/\tilde{v}\delta H \) and \( p^{wc} = (1 - \tilde{q})/(2 - (2-v)\tilde{v})\tilde{v}\delta H/\tilde{v}\delta H \) \( \delta \leq 1 \) (i.e., if \( \tilde{v} \) is sufficiently small) and \( I \leq (1-v)/(2-v)\tilde{v}\delta H/2 \); ii. \( \tilde{p}^{wc} = 1 \) and \( \tilde{p} = \tilde{p}^{wc} \).

In both parts (A) and (B) the repayment amounts are

\[
\tilde{p}^{wc} = \frac{(2 - (2-v)\tilde{v})\tilde{v}\delta H}{(2-v)\tilde{v}^2\delta H - v\delta I} \quad \text{and}
\]

\[
\tilde{p} = \frac{2\tilde{v}\delta H}{2\tilde{v}^2\delta H + (1 - \tilde{v})\delta p^{wc}}
\]

and the expected borrower welfare is \( M = (vH - I)/(1 - \tilde{v} + \tilde{v}(1-v)\tilde{p}^{wc} + (1-v)\tilde{p}^{wc}) \).

To discourage borrowers from colluding in state \((H, H)\) and from behaving as they would in state \((H, 0)\), the bank has to make the latter state less desirable from the group's viewpoint. It can do so by sufficiently increasing \( \tilde{R} \) and/or the cancellation probability \( \tilde{p} \). The proposition says that if lending is feasible with the simple group lending type mechanism, which involves \( \tilde{R} = 2R_C \) and \( \tilde{p} = 0 \), then that is the optimal choice. To see why this is so, that is, to see why the bank cannot do better by decreasing \( \tilde{R} \) from \( 2R_C \) and increasing \( \tilde{p} \), note the following. The real advantage from decreasing \( \tilde{R} \) is that it allows for a decrease in \( p \). A decrease in \( \tilde{R} \) without any decrease in \( p \) worsens borrower welfare.\(^34\) The question then becomes whether, for any

---

\(^{33}\) Weak collusion can also be prevented if instead \( H = (\tilde{R} - R_C) + (\tilde{q} - q)\tilde{q}M \) i.e., output is smaller than a borrower's loss who pays \( \tilde{R} \) (because it simply means that this borrower cannot be sufficiently compensated by the other borrower). It turns out, however, that if \( \tilde{p} \leq \tilde{q} \) then unless \((21)\) is satisfied, it is not possible to satisfy this constraint. Hence, it is sufficient to consider only constraint \((21)\).

\(^{34}\) Any decrease in \( \tilde{R} \) requires an increase in \( R_C \) so that the bank breaks even, and an increase in \( \tilde{p} \) so that the borrowers do not collude in state \((H, H)\). This implies that when \( \tilde{R} \) is decreased and \( \tilde{p} \) is increased, expected payment per borrower remains the same but her future financing is cancelled with a greater probability. This decreases the borrower's expected lifetime utility.
increase in \( v \) means that the likelihood of success is higher, and hence, the expected utility of the borrowers from the renewal of contract is higher for any given \( \delta \). This allows the bank to ask for higher payments without violating the incentive compatibility for any given \( \delta \). But then this also means that the limited liability constraint binds for smaller values of \( \delta \).

We can use the results in Propositions 1 and 6 to compare IL and FGL under weak collusion. Doing so, gives us the following proposition (see also Figs. 1 and 2). Note that, as noted in Section 2.1, the outcome under IL remains the same as described in Proposition 1, even when we allow for weak collusion.

**Proposition 7.** Even when borrowers can collude on their payment decision, flexible group lending results in a higher borrower welfare than individual lending for any I for which lending is feasible under both. If \( \delta > 1/(2−v) \) then lending is also feasible for a greater range of I under flexible group lending. For \( \delta < 1/(2−v) \), the range of I for which lending is feasible is identical under both.

The intuition is the same as before. For any feasible I, a borrower’s expected payment is the same in both group lending and individual lending, but the overall likelihood of cancellation is smaller under the former.

**3.2.1. Comparison with the Rai–Sjostrom mechanism**

We now compare borrower welfare under FGL with that under the (efficient) Rai–Sjostrom mechanism in the presence of weak collusion. Note that this comparison is reasonable only for parameter combinations corresponding to parts A(i), B(i) and B(ii) of Proposition 6. In the case of other parameter combinations (specifically, those corresponding to parts A(ii) and B(iii)), the welfare is also smaller under FGL due to restrictions on the punishment that follow from endogenous determination of the maximum penalty in our framework. Since we are mainly interested in the efficiency loss that results from lack of cross-reporting, it seems reasonable to restrict the comparison to those cases in which any welfare loss is solely due to lack of cross-reporting.

Denote by \( W^{WRC} \) (resp. \( W^{RS} \)) the expected welfare of a borrower in any given period under the FGL (resp. Rai–Sjostrom) mechanism. Let \( ER^{WC} = W^{WRC}/W^{RS} \) denote the efficiency ratio.\(^{37}\) When the parameter values are such that the optimal FGL contract is the same as the SGL contract (cases A(i) and B(i) in Proposition 6), this efficiency ratio is

\[
ER^{WC}_S = \frac{(2−v)v^2H−(2−2v+v^2)I}{(2−v)v^2H−I}.
\]

When the parameter values are those corresponding to case B(ii), the efficiency ratio is

\[
ER^{WC}_B = \frac{vH(2−v)I}{vH(1−v+v^2)−I}.
\]

\(^{37}\) For example, consider a situation with \( \delta > 1/(2−v) \) and I slightly greater than \( (2−v)v^2H/[(2−2v)v\delta] \). This corresponds to part A(ii) in Proposition 6. In this situation, for lending to be feasible, we need to impose punishment on the defaulting borrower in state \( (H,0) \) (i.e., \( p = \delta > 0 \)). We have to do this in order to increase \( \delta \) but without increasing \( \delta \). The reason is that \( \delta \) cannot be increased because the punishment when both borrowers default cannot be increased any further (that is, because \( p = 1 \)). If this limit on maximum penalty did not exist, the bank could have done better with FGL by increasing the punishment in state \( (0,0) \) and increasing both \( \delta \) and \( \delta \) (rather than punishing the defaulting borrower in state \( (H,0) \)). In Rai and Sjostrom (2004), maximum penalty is exogenous and it is assumed to be larger than the output of the borrower. Consequently, in their paper the bank does not face this restriction on penalty.

It can be checked that in each case the efficiency ratio increases with \( v \). The intuition is the same as that given in the case of ‘no collusion’.

To get an idea about the size of these ratios, in Table 1 we present some numerical values for different parameter combinations. As in the case of ‘no collusion’, we choose three different values of \( v \) (specifically, \( v = 0.85, 0.9 \) and \( 0.95 \)), \( \delta = 0.9, I = 1 \), and \( H \) equal to the minimum value required for lending to be feasible. Accordingly, when we compare the FGL contract in part B(i) (resp. B(ii)) of Proposition 6 with the Rai–Sjostrom mechanism, we choose \( H = 2I/(2−v) \).\(^{38}\) We find that in each of these cases the efficiency ratio is approximately 95% or higher. That is, in these cases the optimal FGL contract attains more than 95% of the welfare under the Rai–Sjostrom mechanism.

**4. Concluding remarks**

The papers that compare group lending and individual lending in the presence of strategic default (Besley and Coate, 1995; Armendariz de Aghion, 1999 and Rai and Sjostrom, 2004) seem to suggest that unless group members can impose sufficiently strong social sanctions on their strategically defaulting partners, or unless the bank uses cross-reporting mechanisms, group lending can perform worse than individual lending (it can yield lower repayment rates, smaller bank profit or smaller borrower welfare). In this paper, we identify the reasons why group lending performs worse than individual lending (when it does) in Besley and Coate (1995) and Armendariz de Aghion (1999). We show that when certain restrictions on the group lending contract, as modelled in these papers, are relaxed then group lending yields a higher welfare than individual lending even in the absence of any social sanctions or cross-reporting. In particular, we demonstrate that if the amount that a group member owes for her defaulting partner is optimally determined and if the reward/penalty is allowed to vary from one member to another then (i) group lending yields higher borrower welfare when both GL and IL are feasible, and (ii) the range of opportunity cost of capital for which lending is feasible is also greater under group lending. We show that these results are robust to collusion against the bank.

Further, in this paper the maximum penalty is endogenously determined. This allows for comparison of individual lending and group lending not only with respect to borrower welfare, but also with respect to the maximum opportunity cost of capital for which investment is feasible.

Our findings suggest a possible explanation for some of the observed changes in the Grameen Bank of Bangladesh, namely dissimilar treatment of solvent and defaulting group members. In the new system the terms of the future loan of a member depend not only on group performance but also on her individual performance (see footnote 2). Also, by showing that group lending performs better than individual lending even in the absence of social sanctions, our paper suggests that success of microfinance in countries like Bangladesh and Guatemala (see Wydick, 1999) may be replicated in societies where social connectedness and the ability to impose social sanctions is low.

It should be noted that in this paper we only consider stationary dynamic contracts. An anonymous referee of this paper pointed out that allowing for non-stationary contracts along with savings can result in a higher borrower welfare than that attained under stationary contracts. This is an interesting avenue for future work as it may explain the savings requirements in Grameen Bank II.

\(^{38}\) Note that choice of parameters \( v \) and \( \delta \) rules out case A(i). Hence, in Table 1 we are essentially comparing the performance of FGL contract in cases B(i) and B(ii) with that of the Rai–Sjostrom mechanism.
Acknowledgements

The first author is extremely grateful to Jean-Jacques Laffont for suggesting this problem to him. We would like to thank Amit Batabyal, Sunita Surana and Mike Vernarelli for their comments. We are especially grateful to Dilip Mookherjee and three anonymous referees for their invaluable suggestions that we believe significantly improved the paper. We are responsible for any remaining errors.

Appendix A

Proof of Proposition 1. The bank’s objective is to max $M$ subject to constraints (1)–(4), where $M = \frac{vH - vR_0}{1 - \delta + \delta q + \delta(1 - v)p}$. Substituting this expression for $M$ in constraint (1), the bank’s problem can be written as:

$$\max_{q, \delta, R_0} M = \frac{vH - vR_0}{1 - \delta + \delta q + \delta(1 - v)p} \quad (25)$$

s.t. $R_0 \geq 0$, $R_0 \leq H$, $0 \leq q \leq 1$ and $0 \leq p \leq 1$.

We solve this problem in three steps:

Step 1: Note first that in the optimal solution $q$ must be zero. Suppose not, i.e., suppose $q > 0$. From the objective function it is clear that given other things a small decrease in $q$ will increase borrower’s expected lifetime utility $M$. Also, decreasing $q$ does not violate any constraint, implying that $q > 0$ cannot be optimal.

Step 2: In the optimal solution constraint (26) must hold as an equality. Differentiating RHS (26) with respect to $p$, we get:

$$\frac{\partial M}{\partial p} = \frac{vH\left\{1 - \delta + \delta q + \delta(1 - v)p\right\} - vR_0\delta(1 - v)}{(1 - \delta + \delta q + \delta(1 - v)p)^2} = 0$$

From (34), it is clear that $M$ decreases as $R_0$ increases. Hence, subject to the constraints, the bank would like to set $R_0$ as low as possible. Constraint (34) gives the minimum $R_0$ required for breaking even, $R_0 = \frac{-vH}{\delta(1 - v)}$. As long as this $R_0$ is below the upper limit expressed in constraints (35) and (36) lending is feasible, otherwise lending is not feasible. It can be checked that if $2 - v\delta(4 - v) > 0$ then satisfaction of constraint (35) implies satisfaction of (36) and hence only the former needs to be taken into account. Opposite is the case when $2 - v\delta(4 - v) < 0$.

Step 3: Using equality of constraint (26) and $q = 0$, in any optimal solution $p = \frac{1 - \delta}{\delta(1 - v) - \delta q - \delta(1 - v)p}$. This $p$ is feasible only if it is less than one, i.e., only if $\frac{1 - \delta}{\delta(1 - v) - \delta q - \delta(1 - v)p} < 1$, or $R_0 \leq vH$. Substituting for $p$ in the objective function, taking account of the limit on $R_0$ and for the moment ignoring constraint (28), we can rewrite the bank’s problem as:

$$\max_{R_0} M = \frac{vH - R_0}{1 - \delta} \quad (30)$$

s.t. $l \leq vR_0$, $R_0 \leq vH$. (31)

From (30) it is clear that value of $M$ falls as $R_0$ increases. Hence, the bank would like to set $R_0$ as small as possible, i.e., in any optimal solution constraint (31) must hold with equality. This implies $R_0 = \frac{l}{v}$.

This solution is admissible only provided that $R_0$ so defined is within the limit imposed by constraint (32), or in other words only as long as $\frac{l}{v} \leq vH$, or $l \leq v^2H$. This gives us the maximum feasible level of $l$. For any feasible $l$, substituting $R_0 = \frac{l}{v}$ in the expression for $p$, we get $p = \frac{1 - \delta}{v - vH}$. It is clear that constraint (28) is satisfied whenever constraint (32) is satisfied and hence we do not need to worry about it.

Proof of Proposition 2. The optimal SGL contract is a solution to the problem:

$$\max_{R_C} M = \frac{vH - (1 + (1 - v)^2)R_C}{1 - \delta} \quad (33)$$

s.t. $l \leq (2 - v)vR_C$, $R_C \leq \frac{vH}{2(\frac{1}{\delta} + 1 + (1 - v)^2 + \delta)}$.

Substituting the expression for $p$ in (7) the bank’s problem can be rewritten as:

$$\max_{R_C} M = \frac{vH - (1 + (1 - v)^2)R_C}{1 - \delta} \quad (33)$$

s.t. $l \leq (2 - v)vR_C$.

From (33) it is clear that $M$ decreases as $R_C$ increases. Hence, subject to the constraints, the bank would like to set $R_C$ as low as possible. Constraint (34) gives the minimum $R_C$ required for breaking even, $R_C = \frac{vH}{2(\frac{1}{\delta} + 1 + (1 - v)^2 + \delta)}$. As long as this $R_C$ is below the upper limit expressed in constraints (35) and (36) lending is feasible, otherwise lending is not feasible. It can be checked that if $2 - v\delta(4 - v) > 0$ then satisfaction of constraint (35) implies satisfaction of (36) and hence only the former needs to be taken into account. Opposite is the case when $2 - v\delta(4 - v) < 0$.

Part a: When both individual lending and SGL are feasible, the borrower’s expected lifetime utility under SGL is $M_{SGL} = (vH - l)/(1 - \delta + \delta(1 - v)^2p_S)$ and under individual lending it is $M_{IL} = (vH - l)/(1 - \delta + \delta(1 - v)^2p^H)$. From these expressions, SGL is superior (inferior) to individual lending if $\delta(1 - v)^2p_S^H < (>) \delta(1 - v)^2p^H$. Subtracting $\delta(1 - v)^2p_S^H$ from $\delta(1 - v)^2p^H$ we get, sign $\{\delta(1 - v)^2p_S^H - \delta(1 - v)^2p^H\} = \text{sign}[l - vH]$, which is negative.

Parts b and c: First we compare feasibility of SGL and IL when $2 - v\delta(4 - v) > 0$ and then we consider the case $2 - v\delta(4 - v) < 0$.

(i) $2 - v\delta(4 - v) > 0$.

In this case SGL is feasible if and only if $l \leq (2 - v)v^2M_{SGL}/(1 - \delta + (1 + (1 - v)^2)\delta)$ and IL is feasible if and only if $l \leq v^2H$. Subtracting the latter from the former, we have $((2 - v)v^2\delta(1 - \delta)/2 - (1 - \delta) + (1 + (1 - v)^2)\delta) - v^2H = \delta((2 - v)\delta - (1 + (1 - v)^2)\delta)$ which is positive when $\delta > 1/(2 - v)$ and negative when $\delta < 1/(2 - v)$. Since $0 \leq 2 - v\delta(4 - v)$ (the case we are studying).
it is possible to have \( \delta > 1/(2-v) \) only if \((1/(2-v) < 2/(v(4-v)),\) or only if \(6v-v^2 - 4 < 0.\) Hence, if \(6v-v^2 - 4 > 0\) then \(\delta \leq 2/(v(4-v))<1/(2-v)\) and individual lending is feasible for a greater range of \(v.\) If \(6v-v^2 - 4 < 0\) then individual lending is feasible for a greater range of \(v\) for \(\delta < 1/(2-v)\) and SGL is feasible for a greater range of \(v\) when \(\delta \in (1/(2-v), 2/(v(4-v))).\)

(ii) \(2-v \delta (4-v) < 0.\)

In this case SGL is feasible if and only if \(I \leq (2-v) vH/2\) and as above IL is feasible if and only if \(I \leq \sqrt{\delta} H.\) Subtracting the latter from the former, we have that SGL is feasible for a greater range of \(I\) if and only if \(\delta < (2-v)/2v.\) When \(v\) is sufficiently small, in particular when \(v\) is such that \(6v-v^2 - 4 < 0,\) we have \(2/(v(4-v)) < (2-v)/2v.\) In this case SGL is feasible for a greater range of \(I\) as long as \(\delta \in (2/(v(4-v)),(2-v)/2v)\) and IL is feasible for a greater range of \(I\) for \(\delta > (2-v)/2v.\) When \(v\) is such that \(6v-v^2 - 4 > 0\) then \(2/(v(4-v)) > (2-v)/2v,\) implying \(\delta > (2-v)/2v,\) which implies that IL is feasible for a greater range of \(I.\)

**Proof of Proposition 4.** The bank’s problem is to maximize \(M\) subject to constraints (13)–(20). Provided these constraints are satisfied, we can write the borrower’s expected lifetime utility as: 

\[
M = \mathbb{E}\left(\mathcal{L} - \mathcal{R} - \delta \left(1 - q|\delta\right) M\right) + v (1-q) \left((1-H) - R_G + \delta \left(1-q\right) M\right) + \delta \left(1 - \delta \right) \left(1 - p\right) M, \quad \text{or} \quad M = \frac{vH - \delta \mathcal{R}_G - (1-v)\hat{R}}{1-\delta} + (1-v)\hat{p} + (1-v)^2\hat{q} + (1-v)^2\hat{p}(v(4-v) - \delta) (1-q) M
\]

Using this and simplifying the constraints we can rewrite the bank’s problem as:

\[
\max_{\mathcal{R}_G, \mathcal{R}, \mathcal{R}_C, \mathcal{R}_C} M = \frac{vH - \delta \mathcal{R}_G - (1-v)\hat{R}}{1-\delta} + (1-v)\hat{p} + (1-v)^2\hat{q} + (1-v)^2\hat{p}(v(4-v) - \delta) (1-q) M
\]

s.t. \(\mathcal{R}_G \leq \delta(p - q) M\)

\[
\hat{R} \leq \delta(p - q) M
\]

\[
\mathcal{R}_C \leq \hat{R} + \delta(q - \hat{q}) M
\]

\[
\hat{R} \leq H
\]

\[
I \leq v^2 \mathcal{R}_C + \mathcal{R}(1-v)\hat{R}
\]

\[
\mathcal{R}_C \leq \hat{R} \leq 2\mathcal{R}_C
\]

\[
\hat{R} \leq 2\mathcal{R}_C + \delta(q - \hat{q}) M
\]

\[
0 \leq q \leq 1; \ 0 \leq p \leq 1; \ 0 \leq \hat{q} \leq 1; \ 0 \leq \hat{p} \leq 1.
\]

We solve the problem ignoring constraints (41) and (44) and show later that they are satisfied.

**Step 1:** In any optimal solution, \(q, \hat{q}, \text{ and } \hat{p}\) are all zero. Also, we can ignore constraint (40). Suppose \(q=0.\) From the problem above it is clear that decreasing \(q\) by a small quantity will increase \(M\) and will not violate any constraint (remember we are ignoring constraint (44) for the time being). This implies that any \(q>0\) cannot be optimal. Similarly, if \(\hat{q} \text{ and } \hat{p}\) are greater than zero they can be reduced without violating any constraint and reducing them increases the borrower’s expected utility. We need not worry about violating constraint (40) when reducing \(\hat{q}\) because if \(q=0\) and constraint (43) is satisfied then constraint (40) must be satisfied too. Hence, we can ignore constraint (40).

**Step 2:** Using \(q=\hat{q}=0,\) in any optimal solution constraint (39) must bind. Note also that constraints (39) and (43) imply constraint (38) and hence the latter can be ignored. If constraint (39) does not bind then we can reduce \(p\) until this constraint binds. Doing so increases \(M\) and does not violate any other constraint. Since \(\hat{R} > \mathcal{R}_C,\) we need not worry about violating constraint (38) as long as (39) is satisfied.

**Step 3:** We use equality of constraint (39) and the expression for \(M\) to solve for \(p.\) This gives \(p = \frac{\mathcal{R}_G - \mathcal{R}_C}{vH - \delta \mathcal{R}_G - (1-v)\hat{R}}\) Substituting this \(p\) in the objective function and using steps 1–2 above we can rewrite the problem as:

\[
\max_{\mathcal{R}_C, \mathcal{R}} M = \frac{vH - \delta \mathcal{R}_G - (1-v)\hat{R}}{1-\delta} + (1-v)\hat{p} + (1-v)^2\hat{q} + (1-v)^2\hat{p}(v(4-v) - \delta) (1-q) M
\]

s.t. \(I \leq v^2 \mathcal{R}_C + (1-v)\hat{R}\)

\[
\mathcal{R}_C \leq \hat{R} \leq 2\mathcal{R}_C
\]

**Step 4:** If \(I\) is positive then in any optimal solution, constraint (47) must hold as an equality. This is straightforward. Suppose we have a solution with \(\mathcal{R}_C \text{ and } \hat{R}\) such that constraint (47) is slack. Then we can reduce either \(\mathcal{R}_C \text{ or } \hat{R}\) (or maybe both) by a small amount without violating constraints (47) and (48). This increases a borrower’s expected utility implying that the solution we started with could not have been optimal.

**Step 5:** We prove the following lemma in this step. It shows that the bank can choose \(\hat{R}\) arbitrarily close to \(\mathcal{R}_C\) (alternatively, it can choose \(\eta\) arbitrarily close to zero). This implies that constraint (47) can be satisfied.

**Lemma 1.** Any pair \((\mathcal{R}_C, \hat{R})\) that satisfies constraints (47) and (48) can be replaced by an alternative pair \((\mathcal{R}_C^*, \hat{R}^*)\) with \(\hat{R}^* < \hat{R}\) which also satisfies these constraints and yields higher expected utility to each borrower.

**Proof.** Suppose we have a pair \((\mathcal{R}_C, \hat{R})\) which satisfies constraints (47) and (48). Define \(\hat{R} = \hat{R} - \delta \epsilon \) and \(\mathcal{R}_C = \mathcal{R}_C + \frac{\epsilon}{v(4-v)}\) where \(\epsilon > 0\) is chosen such that \(\frac{\epsilon}{v(4-v)} < 1\). It is clear that pair \((\mathcal{R}_C^*, \hat{R}^*)\) satisfies constraints (47) and (48). Now we show that \((\mathcal{R}_C^*, \hat{R}^*)\) yields higher borrower utility than \((\mathcal{R}_C, \hat{R})\). Define \(\hat{R}^* = \hat{R}\) be a borrower’s expected utility is \(\frac{1}{\mathcal{R}_C^*} \), whereas with \((\mathcal{R}_C, \hat{R})\) it is \(\frac{1}{\mathcal{R}_C}\). Note that the term \(v^2 \mathcal{R}_C + (1-v)\hat{R} - (1-v)\hat{R} = v^2 \mathcal{R}_C + (1-v)\hat{R} - (1-v)\hat{R} = v^2 \mathcal{R}_C + (1-v)\hat{R} - (1-v)\hat{R} = v^2 \mathcal{R}_C + (1-v)\hat{R} - (1-v)\hat{R}\) which implies \(\frac{1}{\mathcal{R}_C^*} > \frac{1}{\mathcal{R}_C}\).

**Step 6:** Lending is feasible for any \(l \leq \frac{v^2 i}{\delta + \epsilon}\) and for any feasible \(l, \mathcal{R}_C = \frac{l}{\delta + \epsilon} - \frac{\epsilon}{\delta + \epsilon} v^2 \hat{R}\). Suppose \(\eta\) is chosen to be sufficiently small. For any \(\eta\) chosen, \(\hat{R} = \mathcal{R}_C + \frac{\eta}{\delta + \epsilon}\). Substituting this in constraint (47) in the expression for \(p\) obtained in step 3 we get that for any \(I\) and \(\eta,\) required \(p = \frac{\epsilon}{v(4-v)} \) which is chosen to be sufficiently small. This choice is possible only if \(I \leq \frac{v^2 (1+\eta)}{\delta+\epsilon} \) or only if \(I \leq \frac{v^2 (1+\eta)}{\delta+\epsilon} \leq 1\), or only if \(I \leq \frac{v^2 (1+\eta)}{\delta+\epsilon} \leq 1\).
This is the maximum feasible $l$ for any given $\eta$. But since $\eta$ can be chosen as small as necessary (but not zero), substituting $\eta=0$ in $v\frac{\partial G}{\partial v} + \frac{\partial G}{\partial p}$, we get that any $l$ is feasible.

**Step 7:** Constraints (41) and (44) that we ignored so far are satisfied by the solution obtained.

Since in the above solution $q = \tilde{q} = 0$, constraint (44) reduces to $\tilde{R} \leq 2R_c$, which is clearly satisfied since it is implied by constraint (48) which was explicitly taken into account. Constraint (41) requires that $\tilde{R} = R_c(1+\eta) \leq H$, i.e., it requires that for any feasible $l$, $\frac{v^2}{1-\theta} + \frac{v^2}{1-\eta} \leq H$ (using $R_c = \frac{v^2}{1-\theta} + \frac{v^2}{1-\eta}$). Since $\frac{v^2}{1-\theta} + \frac{v^2}{1-\eta}$ is increasing in $l$ if we can show that the inequality is satisfied for the maximum feasible $l$ then it follows that it is satisfied for any feasible $l$.

Substituting $\frac{v^2}{1-\theta} + \frac{v^2}{1-\eta}$ in place of $l$, satisfaction of (41) requires $\frac{v^2}{1-\theta} + \frac{v^2}{1-\eta} \leq H$, or $(v + v^3 - 2v^3\delta + v\eta(1-v)(1-2v\delta) - v^3\delta) \geq 0$. The second term can be negative, however, the smaller is $\eta$ the smaller is its magnitude. Since the bank can choose $\eta$ as small as necessary (see step 5 above), it can ensure that $(v + v^3 - 2v^3\delta + v\eta(1-v)(1-2v\delta) - v^3\delta) \geq 0$, that is constraint (41) is satisfied for any feasible $l$.

**Proof of Proposition 5.** The proof is straightforward given the discussion in the paragraph preceding Proposition 5.

**Proof of Proposition 6.** (Part A: $\delta \leq 2/\sqrt{4-v}$)

The bank’s problem is to maximize $M$ subject to constraints (13)–(21). Provided these constraints are satisfied then using the stationary structure of the problem (as in the proof of Proposition 4) we get, $M = \max_{\delta, \phi \leq \hat{\delta}} \frac{\mu H - \frac{v^2}{1-\delta} - (1-v)\hat{\delta}}{1-\delta + \theta\left(\frac{v(1-v)\hat{\delta}}{1-\delta} + (1-v)^2\hat{\delta}\right)}$ (59)

s.t. $R_c \leq p\delta M$ (50)

$\hat{R} \leq (p + \hat{q})\delta M$ (51)

$R_c \leq \hat{R} + \hat{q}\delta M$ (52)

$\hat{R} \leq H$ (53)

$I \leq \frac{v^2}{1-\delta} + (1-v)\hat{\delta}$ (54)

$R_c \geq \hat{R} \leq 2R_c$ (55)

$\hat{R} \geq 2R_c - (\hat{q} + \hat{p})\delta M$ (56)

$0 \leq \hat{p} \leq 1; 0 \leq \hat{q} \leq 1; 0 \leq \hat{p} \leq 1$. (57)

Note that constraint (52) is implied by constraint (55) and hence can be ignored.

Let us explain this restatement of the problem. From steps 1–2 we know that constraints (51) and (56) must hold as equalities in an optimal solution. Substituting the value of $M$ (from (49)) in these two constraints, we get two equations which we use to solve for $R_c$ and $\hat{R}$. We get, $R_c = \frac{(p + \hat{p})\xi}{\mu H}$ and $\hat{R} = \frac{2\xi}{\mu H}$ by substituting these values in the objective function, we get, $M = \frac{\mu H}{2\xi - (2 - v\eta)\phi + (1 + (1-v)^2)\phi}$. Similarly, substituting for $R_c$ and $\hat{R}$ in constraints (50) and (54) gives $\hat{p} \leq p$ and $I \leq \frac{(2 - v\eta)\phi + (1 + (1-v)^2)\phi}{\mu H}$ respectively. Constraint (62) is obtained when we substitute for $\hat{R}$ in (53). Given the expressions for $R_c$ and $\hat{R}$, the second inequality of constraint (55), $\hat{R} \leq 2R_c$, is satisfied for all $\hat{q}$, $\hat{p}$ and $p$. The first inequality requires $p - \hat{p} \geq 2\delta q$ which is stated as constraint (61).

We ignore constraint (62) in steps 4–6 below. In step 7 we find the conditions under which this ignored constraint is satisfied and hence the conditions under which the solution obtained in steps 1–6 definitely holds.

**Step 4:** We can without any loss of generality set $\hat{q} = 0$.

This is straightforward. Decreasing $\hat{q}$ does not violate any constraint and has no effect on the objective function.

**Step 5:** In the optimal solution to the above problem we cannot have both $p < 1$ and $\hat{p} > 0$, i.e., the bank will not increase $\hat{p}$ from zero.
unless \( p \) equals one. This implies that if lending is feasible with \( p \leq 1 \) the bank will choose \( \hat{p} = 0 \).

Maximizing (58) is the same as maximizing \(-2V\nu\hat{p}\delta(1 + (1 - v^2)\delta)p\). Using this, the Lagrangian for the bank's problem is given as (we have ignored constraint \( \hat{p} \leq p \)):

\[
L = -(2 - v)\nu\hat{p} - (1 + (1 - v^2)\delta)p + \lambda \left[ \frac{(2 - v)p + \nu\hat{p}}{2(1 - \delta)} (2 - v)\nu\hat{p} + (1 + (1 - v^2)\delta)p - l \right].
\]

The first order conditions with respect to \( \hat{p} \) and \( p \) are:

\[
\text{FOC}(\hat{p}) : -\delta(2 - v)
+ \lambda \left[ \frac{2v^2\delta H(1 + (1 + p(1 - v))\delta)}{D^2} \right] \geq 0 \text{ if } \hat{p} = 0
+ \lambda \left[ \frac{2v^2\delta H(1 + (1 + p(1 - v))\delta)}{D^2} \right] \geq 0 \text{ if } \hat{p} = 1
\]

\[
\text{FOC}(p) : -(1 + (1 - v^2))\delta
+ \lambda \left[ \frac{2v^2\delta H(1 + (1 + p(1 - v))\delta)}{D^2} \right] \geq 0 \text{ if } p = 0
+ \lambda \left[ \frac{2v^2\delta H(1 + (1 + p(1 - v))\delta)}{D^2} \right] \geq 0 \text{ if } p = 1
\]

where, \( D = 2(1 - \delta) + (2 - v)\nu\hat{p} + (1 + (1 - v^2)\delta)p \). From these first order conditions, as long as \( p \in (0,1) \) we must have \( \lambda = \frac{2v^2\delta H(2 - v(1 - \delta) + (1 + p(1 - v))\delta)}{D(1 + (1 - v^2))\delta} \).

Substituting this value of \( \lambda \) in LHS of the first order condition (64), we get

\[
\text{LHS(64)} = -(2 - v)\nu\hat{p} + \frac{2v^2\delta H(2 - v(1 - \delta) + (1 + p(1 - v))\delta)}{D(1 + (1 - v^2))\delta}.
\]

Step 7: In the above steps we ignored constraint (62). This constraint is satisfied as long as \( 2 - \nu(4 - \nu) \geq 0 \).

Differentiating LHS(62) with respect to \( p \) (using \( \hat{q} = 0 \)), we get

\[
\frac{\partial \text{LHS}(62)}{\partial p} = \frac{2v^2\delta H(2 - v(1 - \delta) + (1 + p(1 - v))\delta)}{D(1 + (1 - v^2))}\delta > 0.
\]

Also, it is straightforward to see that LHS(62) is decreasing in \( \hat{p} \). Hence, LHS(62) takes its maximum value when \( p = 1 \) and \( \hat{p} = 0 \), which is

\[
\frac{2v^2\delta H(2 - v(1 - \delta) + (1 + p(1 - v))\delta)}{D(1 + (1 - v^2))}\delta > 0.
\]

This is less than or equal to one as long as

\[
2 - \nu(4 - \nu) \geq 0.
\]

Hence, as long as \( 2 - \nu(4 - \nu) \geq 0 \) the constraint is satisfied for any feasible \( I \) in the above obtained solution.

Proof of Proposition 6. (Part B: \( \delta > 2/\nu(4 - \nu) \))

Step 1: We first identify the \( I \) values for which the solution obtained in part A above goes through even when \( 2 - \nu(4 - \nu) < 0 \).

Remember that for \( I \leq \frac{2(1 - \delta) + (2 - v)\nu\hat{p} + (1 + (1 - v^2)\delta)p}{2(1 - \delta) + (2 - v)\nu\hat{p} + (1 + (1 - v^2)\delta)p} \leq \frac{2}{2 - v} \).

In part A, we had \( \hat{p} = \hat{q} = 0 \). Also, note that the LHS of constraint (62) is increasing in \( p \). Solving for \( p \) for which LHS(62) equals one (when \( \hat{p} = 0 \) and \( \hat{q} = 0 \)), we get

\[
p = \frac{2(1 - \delta) + (2 - v)\nu\hat{p} + (1 + (1 - v^2)\delta)p}{2(1 - \delta) + (2 - v)\nu\hat{p} + (1 + (1 - v^2)\delta)p}.
\]

For any smaller \( p \), LHS(62) is less than one and hence constraint (62) is satisfied. Substituting this \( p \) value in the RHS(60) we get

\[
\text{RHS}(60) = \frac{2v^2\delta H(2 - v(1 - \delta) + (1 + p(1 - v))\delta)}{D(1 + (1 - v^2))}\delta > 0.
\]

Hence, for any \( I \leq \frac{2}{2 - v} \), the solution obtained above satisfies the limited liability constraint even when \( 2 - \nu(4 - \nu) < 0 \) and hence is valid. We need to resolve the problem for \( I > \frac{2}{2 - v} \), which we do now.

Step 2: The bank's problem is to:

max 58

s.t. (59) (63) (67)

where, unlike part A, we cannot ignore constraint (62) anymore. Note first that in the optimal solution to this problem \( \hat{q} \) must be 0. Suppose not. That is suppose we have an optimal solution with \( \hat{q} > 0 \). If (62) is slack, then reduce \( \hat{q} \) until constraint (62) binds or until \( \hat{q} \) becomes zero. Doing so does not violate any other constraint; it in fact increases RHS(60) which allows for a small decrease in \( p \), which in turn increases the borrower's expected utility. This contradicts the fact that our solution with \( \hat{q} = 0 \) when (62) is slack was optimal.

If constraint (62) is binding and \( \hat{q} = 0 \) then do the following: decrease \( \hat{q} \) and \( p \) by small amount \( \epsilon \) and increase \( \hat{p} \) by \( \frac{(2 - v)\nu\hat{p} + (1 + (1 - v^2)\delta)p}{2(1 - \delta)} \).

This keeps the LHS of (62) unchanged. We show that these changes increase the RHS (60), making constraint (60) slack, which in turn allows a further decrease in \( p \) and corresponding increase in the borrower's expected utility. Hence, whenever \( \hat{q} = 0 \) we can increase borrower's utility by decreasing \( \hat{q} \) slightly (and carrying out other changes mentioned above), which contradicts the assumption
The objective is to maximize $\hat{q}$ subject to (69) and (70). Let us solve

Step 3: With $\hat{q}$, constraint (61) is implied by constraint (59) and hence can be ignored. Also, ignoring constraint (59) for now (we will show later that if it is satisfied) we can write the bank's problem as:

$$\max_{\hat{p}} \quad 2\hat{v}_{H}$$
$$\text{s.t.} \quad I \leq \frac{(2-v)\hat{p} + \hat{v}_{H}}{(2-\hat{v})\hat{p}}, \quad \hat{p} \leq \frac{(2-v)\hat{p} + \hat{v}_{H}}{(2-\hat{v})\hat{p}}, \quad \hat{p} \leq \frac{2\hat{v}_{H}}{(2-\hat{v})\hat{p}} \leq 1$$

Step 4: Note first that in any solution to this problem constraint (70) must bind: Suppose that is not the case for some $I > (2-v)\hat{p}$. That is suppose that there is a solution (for which (70) is slack). We know that if $\hat{p} = 0$, the required $p$ for any such $I$ is greater than $\frac{2(1-\hat{v})}{(4\hat{v} - \hat{v}_{H} - 2\hat{v})}$ (from RHS(69)) and for any such $p$ the LHS of (70) exceeds one. Hence, the only way constraint (70) can be slack for $I > (2-v)\hat{p}$ is if $\hat{p} > 0$. Now if $\hat{p} > 0$ and constraint (70) is slack then we can reduce $\hat{p}$ slightly without violating constraint (70). This increases RHS (69) and the borrower's expected utility, contradicting the fact that we had an optimal solution.

Step 5: Given step 4 above we can use (70) to solve for $\hat{p}$ in terms of $p$, which gives $\hat{p} = \frac{2(1-\hat{v}) + (2-4v + \hat{v}_{H})}{(2-v)\hat{p}}$. This is feasible only if it does not exceed one. Hence, any $I$ is feasible only if it satisfies $\frac{(2-v)\hat{p} + \hat{v}_{H}}{(2-\hat{v})\hat{p}} \leq \frac{2\hat{v}_{H}}{(2-\hat{v})\hat{p}} \leq 1$; solving for $I$ we get that the maximum feasible $I$ is $I = \frac{(2-v)\hat{p} + \hat{v}_{H}}{(2-\hat{v})\hat{p}}$. Subtracting the optimal $\hat{p}$ from $p$ we get, $p - \hat{p} = \frac{2(1-\hat{v}) - (2-4v + \hat{v}_{H})}{(2-v)\hat{p}} > 0$ for any $I \leq \frac{(2-v)\hat{p} + \hat{v}_{H}}{(2-\hat{v})\hat{p}}$, implying that constraint (59) which we ignored, is satisfied.

Case 2. When $\hat{v}_{H} < 1/(2-v)$

The statement of the problem remains the same as in Case 1 above. The objective is to maximize (68) subject to (69) and (70). Let us solve this problem.

Step 7: For any $I \leq \frac{2\hat{v}_{H}}{(2-v)\hat{p}}$ constraint (70) must bind in the optimal solution: Substituting $\hat{p} = \frac{2(1-\hat{v}) + (2-4v + \hat{v}_{H})}{(2-v)\hat{p}}$ in LHS of constraint (70), we get that LHS(70) = 1 for any $p$. That is, for any given $p$, as long as $\hat{p} = \frac{2(1-\hat{v}) + (2-4v + \hat{v}_{H})}{(2-v)\hat{p}}$, this constraint holds as an equality. Now suppose we have an optimal solution for which $I = \frac{2\hat{v}_{H}}{(2-v)\hat{p}}$ and $p$ are so chosen that constraint (70) is slack. Then it must be the case that, $\hat{p} = \frac{2(1-\hat{v}) + (2-4v + \hat{v}_{H})}{(2-v)\hat{p}}$. Decrease $\hat{p}$ by a small amount $\varepsilon$. To ensure that break-even constraint is still satisfied, this requires an increase in $p$ of $\frac{(2-v)\hat{p}}{(2-\hat{v})\hat{p}} \leq \frac{(2-v)\hat{p}}{(2-\hat{v})\hat{p}} - \frac{(2-v)\hat{p}}{(2-\hat{v})\hat{p}} = \varepsilon$. These changes increase the LHS of constraint (70), however when it is slack there always exists a small enough $\varepsilon$ for which we can carry out the above changes without violating this constraint. Substituting these changes in the objective function we can see that the denominator of the objective function falls by $\frac{2(1-\hat{v}) + (2-4v + \hat{v}_{H})}{(2-v)\hat{p}} - \frac{(2-v)\hat{p}}{(2-\hat{v})\hat{p}}$, i.e., the borrower's expected utility increases. This contradicts the fact that the solution in which constraint (70) is slack (when $I = \frac{2\hat{v}_{H}}{(2-v)\hat{p}}$) is optimal.

Proof of Proposition 7. Case $\hat{v}_{H} \leq 2/v(4-v)$

Since whenever $I \leq \frac{(2-v)\hat{v}_{H}}{(2-\hat{v})\hat{p}} (2(1-\hat{v}) + (1 + (1-v)^2)\hat{p})$ or $\hat{v}_{H} \leq (2-v)/(2-v)$ the flexible group lending contract is identical to the simple group lending contract, SGL, we already know (from Proposition 3) that flexible group lending does better than individual lending for parameter values satisfying these conditions. We now show that flexible group lending results in a higher expected utility per borrower than individual lending even when $\hat{v}_{H} < 1/(2-v)$ and $I \leq \frac{(2-v)\hat{v}_{H}}{(2-\hat{v})\hat{p}} (2(1-\hat{v}) + (1 + (1-v)^2)\hat{p})$. Comparing $M_{p+}$ with $M_{\hat{p}}$ we get that a borrower's expected utility is higher under flexible group lending provided $\hat{v}_{H} < 1/(2-v)$ and $(1 + (1-v)^2)\hat{p} > \hat{v}_{H}$. Substituting the values of $\hat{v}_{H}$ and $\hat{p}$, we get that flexible group lending results in a higher expected utility $\frac{(1 + (1-v)^2)\hat{p} - \hat{v}_{H}}{(2-v)\hat{p}} > \frac{(2-v)\hat{v}_{H}}{(2-\hat{v})\hat{p}} > \frac{(2-v)\hat{v}_{H}}{(2-\hat{v})\hat{p}}$. When $I \leq \frac{(2-v)\hat{v}_{H}}{(2-\hat{v})\hat{p}}$, it can be easily checked that this is the case when $I \leq \frac{(2-v)\hat{p}}{(2-\hat{v})\hat{p}}$ and $\hat{v}_{H} < 1/(2-v)$.

Proof of Proposition 7. Case $\hat{v}_{H} > 2/v(4-v)$

We first show that lending is feasible for a greater range of investment cost under flexible group lending than under individual lending. Under individual lending investment is feasible only for $I \leq \hat{v}_{H}$, whereas under FGL it is feasible for any $I \leq \frac{(2-v)\hat{v}_{H}}{(2-\hat{v})\hat{p}}$. Subtracting $\hat{v}_{H}$ from $\frac{(2-v)\hat{v}_{H}}{(2-\hat{v})\hat{p}}$, we get, $\frac{(2-v)\hat{v}_{H}}{(2-\hat{v})\hat{p}} - \frac{(2-v)\hat{v}_{H}}{(2-\hat{v})\hat{p}} = \frac{(2-v)\hat{v}_{H}}{(2-\hat{v})\hat{p}}$.

41 For the case we are considering, $\hat{v} > 1/(2-v)$, RHS(69) is decreasing in $\hat{p}$. 

42 By choosing a small enough $\varepsilon$ we can ensure that it is possible to increase $p$ by the required amount for any $\hat{p} < 1$. If $\hat{p} > 1$ and $\hat{p} > 1 + (2-4v + \hat{v}_{H})$, then constraint (70) will be slack for any $I \leq \frac{(2-v)\hat{p}}{(2-\hat{v})\hat{p}}$ and hence $\hat{p}$ can be decreased by a small $\varepsilon$ without violating this constraint.

43 It can be shown that as $\varepsilon \to 0$, the change in LHS of constraint (70) goes to zero.
Differentiating the term \((2\delta - 1 - 2v\delta^2 + v^2\delta^2)\) in the numerator with respect to \(\delta\), we get \(\frac{d}{d\delta}(2\delta - 1 - 2v\delta^2 + v^2\delta^2) = 2 - 4v\delta + 2v^2\delta\). Note that \(2 - 4v\delta + 2v^2\delta\) is falling in \(v\) (for any \(\delta > 0\)). Substituting \(v = 1\) in it, we get \(2 - 4v\delta + 2v^2\delta = 2 - 2\delta > 0\) for all \(\delta\). Hence, the derivative of \((2\delta - 1 - 2v\delta^2 + v^2\delta^2)\) with respect to \(\delta\) is positive for all \(v\) and \(\delta\). This implies \((2\delta - 1 - 2v\delta^2 + v^2\delta^2)\) is increasing in \(\delta\), for all \(v\) and \(\delta\). Evaluating it at the smallest possible value of \(\delta\) (which is \(1/2\) for the case we are considering), we get \((2\delta - 1 - 2v\delta^2 + v^2\delta^2) = 0\) for \(\delta = \frac{1}{2}\). Hence, for any \(\delta\) above that it should be positive, implying \(\frac{2\delta - 1 - 2v\delta^2 + v^2\delta^2}{\delta - 1} > v^2\delta H\), i.e., the range of investment costs for which lending is feasible is greater under group lending.

Now we show that for any \(I\) for which lending is feasible under both group lending and individual lending, the borrower’s expected utility is higher under the former; for any \(I \leq (2 - v)\nu H/2\), the FGL is the same as the SGL and hence as shown in the proof of Proposition 3, the borrower’s expected utility is higher in this case. For any \(I \equiv \frac{2(1 - \delta)}{\delta - 1} (\nu H^2 + \nu^2 H)\), substituting \(\hat{\delta} = \frac{2(1 - \delta)}{\nu H^2 + \nu^2 H}\) and \(p = \frac{\delta - 1}{\delta - 1}\) in the objective function (68) we have \(M_{\text{GL}} = \frac{1}{\delta - 1}\left(\frac{\delta - 1}{\delta - 1}\right)\). Comparing this utility to the borrower’s utility under individual lending \(M_{\text{GL}} = \frac{1}{\delta - 1}\left(\frac{\delta - 1}{\delta - 1}\right)\), we get \(M_{\text{GL}} - M_{\text{GL}} = \frac{1}{\delta - 1}\left(\frac{\delta - 1}{\delta - 1}\right) = \frac{1}{\delta - 1}(\delta - 1) > 0\). That is, the group lending contract gives the borrowers a higher expected utility than the individual lending contract.

**Case 2.** \(\delta \leq 1/(2 - v)\)

When \(\delta = 1/(2 - v)\), both FGL and IL are feasible for the same range of \(I\), which is, \(I \leq \nu H^2\). We show that any \(I\) for which lending is feasible under both, group lending yields a higher borrower utility than individual lending. Consider the range \(I \equiv \frac{2(1 - \delta)}{\delta - 1} (\nu H^2 + \nu^2 H)\), and the range \(I \equiv \frac{2(1 - \delta)}{\delta - 1} (\nu H^2 + \nu^2 H)\) separately. When \(I \equiv \frac{2(1 - \delta)}{\delta - 1} (\nu H^2 + \nu^2 H)\), we have \(p = \frac{\delta - 1}{\delta - 1}\) and \(\hat{\delta} = \frac{2(1 - \delta)}{\nu H^2 + \nu^2 H}\), which is the same as in Case 1 above; we have shown there that group lending yields higher expected utility. When \(I \equiv \frac{2(1 - \delta)}{\delta - 1} (\nu H^2 + \nu^2 H)\), we have \(p = 1\) and \(\hat{\delta} = \frac{1}{\nu H^2 + \nu^2 H}\). Substituting these in the objective function we get \(M_{\text{GL}} = \frac{1}{\delta - 1}\left(\frac{\delta - 1}{\delta - 1}\right)\). Subtracting from this a borrower’s utility under individual lending we get \(M_{\text{GL}} - M_{\text{GL}} = \frac{2(1 - \delta)}{\delta - 1} (\nu H^2 + \nu^2 H)\). From Proposition 4, the repayment amount plus the punishment imposed on a borrower in the optimal FGL mechanism can be made arbitrarily close to: \(I/\nu\) in state \((H, H)\); \(I/\nu\) in state \((H, 0)\) when the borrower is successful; zero in state \((0, 0)\) if the borrower is unsuccessful; and \(I/\nu\) in state \((0, 0)\).\(^{44}\) Hence, welfare under FGL can be made arbitrarily close to:

\[
W_{\text{WC}} = \left(v^2 + v(1 - v) + (1 - v)^2\right)\left(H - \frac{1}{\nu}\right). \tag{73}
\]

Using Eqs. (72) and (73) we get the efficiency ratio in the case of no collusion.

**Case 2. Weak collusion between borrowers – optimal FGL is the same as SGL**

Consider the case when parameter values are such that the optimal FGL mechanism is the same as the SGL mechanism. In this case, the optimal contract under the Rai–Sjostrom mechanism is described in Proposition 3 of their paper, with \(\tilde{R} = \tilde{r}/(p(h, h) + 2p(0, h))\) (see Rai and Sjostrom (2004, pp. 222 and 227)). Therefore, the repayment amount plus the punishment imposed on a borrower in their mechanism (in terms of the notation in our paper) is: \(I/(2v - v^2)\) in state \((H, H); 2I/(2v - v^2)\) in state \((H, 0)\) when the borrower is successful; zero in state \((0, H)\) when the borrower is unsuccessful; and \(I/(2v - v^2)\) in state \((0, 0)\). Using this, a borrower’s welfare is

\[
W_{\text{RS}} = \left(\frac{v^2}{2v - v^2} + v(1 - v)\right)\left(H - \frac{2I}{2v - v^2}\right) + (1 - v)^2\left(-\frac{2I}{2v - v^2}\right). \tag{74}
\]

The optimal FGL mechanism is that given in Proposition 2. It can be checked that in all states except state \((0, 0)\), the sum of the repayment amount and punishment of a borrower is the same as that in the Rai–Sjostrom mechanism. In state \((0, 0)\), a borrower’s punishment is \(p\delta^2\delta^2M_{\text{GL}} = 2I/(2v - v^2)\). Using this, a borrower’s welfare under FGL is

\[
W_{\text{WC}} = \left(H - \frac{I}{2v - v^2}\right) + v(1 - v)\left(H - \frac{2I}{2v - v^2}\right) \tag{75}
\]

Using Eqs. (74) and (75) we get the efficiency ratio, \(ER_{\text{WC}}\), in Eq. (23).

**Case 3. Weak collusion between borrowers – optimal FGL is that in part B(ii) of Proposition 6**

Consider the case when parameter values are those corresponding to part B(ii) of Proposition 6. In this case the limited liability constraint binds and the optimal contract under the Rai–Sjostrom mechanism is that described in Proposition 3 of their paper, with \(\tilde{R} = r)/(p(h, h) + p(0, 0))\) (see Rai and Sjostrom (2004, pp. 222 and 227)). This implies a repayment amount plus punishment that equals: \((1 - (1 - v)H)/v^2\) in state \((H, H); H\) in state \((H, 0)\) when the borrower is successful; zero in state \((2v - v^2)/v^2\) when the borrower is unsuccessful; and \((1 - (1 - v)H)/v^2\) in state \((0, 0)\). Using this, a borrower’s welfare under the Rai–Sjostrom mechanism is

\[
W_{\text{RS}} = \left(H - \frac{I}{v^2} - \frac{v(1 - v)H}{v^2}\right) + (1 - v)^2\left(-\frac{2I}{v^2}\right). \tag{76}
\]

\(^{44}\) This is obtained by substituting \(\eta = 0\) in the expressions for \(R_c\), \(\tilde{R}\) and \(p\) in Proposition 4, and using \(\tilde{\delta}^2\delta^2M_{\text{GL}}\) to calculate the punishment in state \((0, 0)\).
The optimal FGL mechanism in this case is that given in Proposition 6, part B(ii). It can be checked that in all states except state (0, 0), the sum of the repayment amount and punishment of a borrower in FGL is the same as that in the Rai–Sjostrom mechanism. In state (0, 0), a borrower’s punishment is $\delta p^{MC} = H$. Using this, a borrower’s welfare under the Rai–Sjostrom mechanism is

$$W^{WC} = v^2 \left( H - \frac{(1 - v)\nu H}{v^2} \right) + (1 - v)v \left( \frac{(2 - v)\nu H - 2\nu}{v^2} \right)$$

Using Eqs. (76) and (77) we get the efficiency ratio, $ER^{MC}$, in Eq. (24).

References
