Bogus Joint Liability Groups in Microfinance – Theory and Evidence from China

Alexander Karaivanov* Xiaochuan Xing† Yi Xue‡

March, 2016

Abstract

Survey data on clients of CFPAM, the leading microlender in China, indicate that nearly 70% of all surveyed joint-liability borrower groups are 'bogus' – that is, one person uses all loans given to group members in a single investment project while all other group members act as unproductive cosigners. This practice not only violates CFPAM rules but is also inconsistent with most of the theoretical literature on group lending, a basic tenet of which is that each borrower uses their own loan to implement their own investment project (what we call 'standard group'). We therefore extend the classic model of group lending under joint liability by explicitly allowing for both standard and bogus groups in a setting with the possibility of strategic default due to limited enforcement. The optimal choice between standard and bogus groups is endogenous and depends on the borrowers' characteristics (project productivity and probability of success). We analyze the optimal group loan contract (or menu of contracts) and show that bogus groups optimally arise when either the productivity differential between the projects in a group is high (in heterogeneous groups), or when the absolute level of project productivity is high (in homogeneous groups). Explicitly allowing for the possibility of bogus group formation not only helps the lender avoid losses which may occur if bogus groups are ignored, but also enhances the productive efficiency and borrower welfare in the economy. We test the model predictions with data from rural China and evaluate the welfare gains from implementing the optimal contract (or menu) relative to the benchmarks of: (a) lenders operating unaware of bogus groups or (b) lenders using a contract with terms that endogenously rule out bogus group formation.

JEL classification: C72; D82; G20; O12
Keywords: microfinance; group lending; bogus groups; optimal contracts; strategic default

*Department of Economics, Simon Fraser University, akaraiva@sfu.ca

†From School of Economics and Management at Tsinghua University.

‡From School of International Trade and Economics at the University of International Business and Economics, gratefully acknowledge financial support from the National Natural Science Foundation of China (NSFC), project No. 71471040.

1 Introduction

Group lending under joint liability has been the workhorse of microfinance since its origins in Mohammed Yunus' Grameen Bank in Bangladesh. It offers an elegant solution to the main problem of lending to the poor, namely their lack of physical collateral. Lacking collateral prevents borrowers from mitigating agency costs arising from information asymmetries or limited enforcement in the credit market, especially in developing countries.[1] As shown in both the theory and practice, group lending is able to create a substitute for physical collateral by using the joint liability clause and the social capital embedded in groups of poor borrowers to mitigate moral hazard, adverse selection, costly state verification, and/or enforcement problems.[2][3] In joint-liability group lending the lender provides a loan to each member of a group of borrowers with the requirement to invest the money in their own business project. All members are jointly responsible for the group's liability: if a member defaults, her liability must be repaid by the other members, otherwise the entire group can lose access to future credit (Chowdhury, 2005). In contrast to individual loans, this design allows lending to borrowers without collateral and at a relatively low interest rate due to the higher repayment rate resulting from risk-diversification, monitoring, and/or peer pressure within the group.

Survey data on 366 clients of the leading joint-liability microfinance lender in China, CFPAM (the China Foundation for Poverty Alleviation Microfinance) indicate that nearly 70% of the borrower groups in which they participate are what we call *bogus groups*, that is, one individual uses all loans given to the group members in a *single* investment project while the rest of the group act as unproductive cosigners.[4] This practice, known as *Lei Da Hu* in Chinese, is technically in violation of the loan covenant terms, however, CFPAM does not seem to take actions to eliminate it.

More importantly for the purposes of this paper, bogus groups are inconsistent with a basic tenet of group lending – both as designed in practice and as modeled in the economics literature – the requirement that each borrower invests in their own project (what we call a *standard group*). In this sense a bogus group resembles a larger-size individual loan – if the cosigners are poor, their nominal presence on the group loan offers no real protection to the lender. It is thus interesting to observe that the majority of CFPAM groups are bogus since according to the theory, group lending can reduce borrowing costs and increase the probability of obtaining future loans compared to individual lending. To the best of our knowledge ours is the first paper to formally analyze the issue of bogus microfinance groups both theoretically and empirically.

[1] See Armendariz de Aghion (1999), Ghatak and Guinnane (1999), Ahlin and Waters (2011), Ahlin (2012) among many others.

[2] See Mosley (1986), Udry (1990), Morduch (1999), Ghatak and Guinnane (1999), Ghatak (1999), Besley and Coate (1995), Ahlin and Waters (2011), etc.

[3] In addition to the risk diversification effect mentioned above, the group-lending benefits the poor people via the monitoring cost reduction effect since the group members has comparative advantages in information revealing relative to the lenders such as banks or moneylenders. Intuitively, the members in the group can monitor each other or verify the state of each other's investment projects with relatively low cost because usually they live geographically nearby (Ghatak and Guinnane, 1999; Karlan, 2005), they also have stronger enforcement power to make the members repay because they have close social links and can impose powerful social sanctions on members who default strategically (Besley and Coate, 1995; Armendariz de Aghion, 1999). Therefore the lender is willing to lend to the group even though they have no collateral since the members with these comparative advantages have incentive to monitor, verify and enforce each other when they have joint-responsibility for any of the members' default. In other words, the group members' comparative advantages on information revealing helps them access to the credit and improves the efficiency of the economy.

[4] The cosigners agree to participate either because the person who ends up using the money provides a monetary or social compensation to them or because they might want to invite her to cosign another loan in the future. Our model implicitly incorporates both these channels.

We build a model with a population of borrowers each of whom is endowed with a single investment project that can be of either high or low productivity. Each project is successful with a probability p and fails (yield zero) with probability $1-p$. The borrowers have no funds of their own to finance their projects, so they must borrow from a perfectly competitive microfinance lender.[5] Only group loans are available and are given to groups of two borrowers whose projects' outputs are i.i.d. across group members.[6] There is a limited commitment friction in the credit market – a borrower can default strategically on her loan if she finds it optimal to do so. This can be due either to unobservable project output or limited enforcement (both within and outside the group). After obtaining their loans, the group members can choose to either operate as a standard group or a bogus group. In both cases, all members bear joint responsibility for the group's total liability. If this joint liability clause is not fulfilled, all group members are excluded from access to future credit. The possibility of strategic default by an individual group member thus affects the other member's repay/default decision in equilibrium. Finally, in case of project failure (involuntary default) there is limited liability and the lender cannot collect anything from that borrower.

We first characterize the optimal choice between repayment and strategic default in a standard group. We show that, if the loan size is sufficiently small, there is a unique Nash equilibrium in which both borrowers repay in full (for themselves or for the whole group) whenever their project succeeds. We solve for the optimal standard group loan contract subject to the no-default and the lender's break-even constraints.

We then characterize bogus groups in the same setting and compare them to the standard groups. There are five key differences: (a) the expected future value of continued access to credit is larger in a standard group due to the larger repayment probability (risk sharing or diversification) when project returns are i.i.d. across borrowers but this is partially offset by the larger expected repayment arising from the need to cover for one's partner; (b) the loan interest rate is lower in a standard group for the same reasons as in (a); (c) strategic default incentives are stronger in a standard group since a defaulter would still receive the future value if their partner repays; (d) the optimal loan size is larger in a bogus group due to the less tight no-default constraint; and (e) expected output is higher in a bogus group with heterogeneous members since all funds are optimally invested in the highest-productivity project. Overall, items (a) and (b) favor standard groups while items (c)-(e) favor bogus groups. The trade-off between the various effects influences the overall outcome of the endogenous choice of group form by the borrowers and consequently the lender's optimal way of dealing with bogus groups.

We first show that a lender who ignores the possibility of bogus groups and simply offers the optimal standard group contract to all groups can suffer a monetary loss if it is optimal for some borrowers to form bogus groups which occurs if their projects' productivity differential is sufficiently large. The reason is that the lender relies on the risk diversification in a standard group to set a low interest rate but this diversification effect is absent in a bogus group in which all money is pooled into a single project.

We then analyze the optimal contracting problem of a lender who takes into account that bogus and standard groups would form endogenously for given contract terms but the group form decision is unobserved to the lender (ex-post moral hazard). In the case of observable project productivities, we show that the optimal loan contract (loan size and repayment) is tailored to the group composition and depends on the productivity values and the project success probability.

[5] We assume free entry by lenders which implies they make zero expected profits per loan (no cross-subsidization is possible).

[6] In the baseline setting we assume that the groups are formed before the project productivities are known to the borrowers (exogenous matching). We briefly discuss changing the timing to allow endogenous matching in the conclusions.

Specifically, for homogeneous groups in terms of project productivity, the trade-off is between effects (a)-(b) vs. (c)-(d) from the list above and we prove that it is optimal to offer a relatively large loan size and induce the formation of a bogus group whenever project productivity exceeds a certain threshold and to offer the optimal standard loan contract when the productivity parameter is low. Intuitively, groups with highly productive projects benefit more from the larger loan size supported in bogus groups while the low interest rate and risk sharing benefits dominate for lower productivity groups.

For heterogeneous groups (one high- and one low-productivity project) the optimal contract is more complex – providing a larger loan size and inducing bogus group is optimal when the high-productivity value is large enough in absolute sense or when the productivity differential between the two projects is sufficiently large to benefit from effect (e) in the list above. In contrast, when both projects have relatively low productivity and their productivity differential is small, it is optimal to offer a smaller loan size and induce a standard group. The loan size in the latter case is either the maximum size allowed by the no strategic default constraint, or even smaller depending on the parameters, to deter borrowers from benefiting from the low interest but forming a bogus group. Overall, we find that bogus groups are optimal when either project productivity is high or, in heterogeneous groups, when the productivity differential inside the group is sufficiently high. These are testable implications of the model that we take to the data.

We also show that our main findings remain largely unchanged when the project productivities are unobserved to the lender. In this case the lender offers a menu of two contracts which differ in their loan size and interest rate (repayment) – one contract designed for standard groups, with lower loan size and lower interest, and one contract designed for bogus groups, with larger size and higher interest. Once again, we prove that bogus groups are optimal when the productivity level is sufficiently large for homogeneous groups or when the productivity levels or productivity differential are large enough, in heterogeneous groups. One difference from the observed productivities case is that, if project productivity is unobserved, the lender not only has to take into account the ex-post moral hazard problem of a group taking a loan contract and deviating to the alternative group form but also the 'adverse selection' problem of a group mis-reporting its project composition. The free entry assumption helps us characterize the optimal contract menu since it implies that the lender can only offer at most two contracts – for a given interest rate, determined by the zero profit condition, the lender cannot use different loan sizes to screen across groups with different composition since all groups prefer larger loans.

Finally, we analyze the problem of a lender who, for exogenous reasons, wishes to design a group loan contract that rules out the endogenous formation of bogus groups.[7] Naturally, this results in a suboptimal outcome, associated with lower borrower welfare. The reason is that forcing all borrowers form standard groups is counter-productive when project productivity is large and the larger loan size that can be supported in a bogus group (effect d) or the additional output possible in a heterogeneous group (effect e) are left unexploited.

In the empirical part of the paper we use data from telephone interviews with 366 clients of CFPAM, the largest microlender in China. During the interview, we asked how the group loan is used. The responses indicate that in nearly 70% of the borrowing groups in our sample the entire loan is fully used by a single group member in her project while all other members play the role of nominal cosigners. The remainder of the groups are standard groups – in them each member uses their loan for their own investment project. Apart from the group form, bogus vs. standard, we have data on the loan size, required repayment, and intended loan use, as well as demographic

[7] For example, the lender may consider bogus groups as a violation of the legal aspects of the contract or the interest rate is regulated and the lender cannot break even if bogus groups exist, or the lender wants each borrower to invest in their own project to improve her skills, experience, etc.

information such as education, ethnic group, marital status, and age.

We perform regressions to study the determinants of the observed choice between bogus and standard group. We find that having chosen to form a bogus group is positively and statistically significantly associated with the loan size. We also intend to structurally estimate our model and evaluate the welfare gains from implementing the optimal contract (or menu) relative to the sub-optimal benchmarks of (a) lenders being unaware of bogus groups, or (b) lenders using a loan contract that endogenously excludes bogus groups. [*this part is in progress*]

Literature review [*to be added*]

2 The model

2.1 Basic setting

Consider an economy populated by two types of agents: lenders and borrowers. Each borrower is endowed with one investment project which has to be financed by taking a loan from a lender (microfinance institution) at time $t = 0$. The borrowers have no wealth, hence the entire initial investment L needed to implement an investment project must be financed by the loan. There are two types of investment projects: a 'conventional' project with productivity k_L and a 'high-return' project with productivity k_H, where $k_H \geq k_L > 0$. The projects' output generated at $t = 1$, is stochastic, as described below:

$$Y_i = \begin{cases} k_i L & \text{with probability } p, \\ 0 & \text{with probability } 1-p. \end{cases}$$

The parameter $p \in (0, 1)$ can be thus interpreted as the probability of a project being successful. The project output realizations are i.i.d. across borrowers.

Both the lenders and the borrowers are risk-neutral. The lenders have an opportunity cost of funds normalized to 1. We assume an environment with limited enforcement (the project return Y_i is non-verifiable) which gives rise to the possibility of strategic default – a borrower can report failure while in fact her project has succeeded. In addition, borrowers are subject to limited liability: if the project fails (yields zero), the borrower involuntarily defaults, in which case the lender cannot punish the borrower further. The loan terms must therefore be such that borrowers are given incentives to pay back when their project succeeds. For simplicity, the borrowers' outside option (if they do not invest in their project) is normalized to zero.

We assume that there is free entry in the micro-lending sector. This implies that any lender earns zero profits (breaks even) on each loan made and rules out cross-subsidization across borrowers. It is also assumed that the lender only provides a *group loan* contract. For example, as motivated by the empirical evidence from China we use, the lender could be an NGO with the mission to finance business projects of poor people. Specifically, as assumed in most of the group-lending literature, assume that the loan is issued to a group of two members. The group contract thus consists of two loans of size L each and a pre-specified total repayment amount R per loan (the gross interest rate is R/L). Since the borrowers have no wealth, requiring collateral is infeasible. Instead, the contract has a *joint liability* clause: each member is fully responsible for the total group obligation $2R$. That is, if the lender does not receive $2R$ at $t = 1$, either from a single borrower or from both combined, then both borrowers are cut off from access to credit in the future. Maintaining future access to credit has present discounted value $V > 0$ for a borrower.

The group of two borrowers can either operate as a "standard group" or as a "bogus group". In a standard group each member invests L into her *own* business project, as assumed in the group lending literature or as required in the loan terms of microlenders in practice. In contrast, in a bogus group the members invest the *total* loan amount, $2L$ into one of the two projects and share the total surplus. Normally, such behavior would be considered a violation of the loan covenants but, as motivated by the empirical evidence, such violations may be hard to verify in practice. In addition, we will show that lenders allowing bogus groups to form is in fact efficiency-enhancing.

The parameters k_i, k_j (the project productivities) are known to both borrowers in a group. In the baseline version of the model, we assume that the lender can also observe the project types. A possible interpretation is that, in practice, microfinance lenders acquire information on the projects the borrowers intend to implement before providing a loan. We also consider the case of k_i's that are unobservable to the lender in section 3.2 and show that our main conclusions and intuition remain largely unchanged.

2.2 Standard Groups Only

We start the analysis with the basic setting from the existing literature, in which bogus joint liability groups are exogenously ruled out and only standard groups exist. The timing is as follows:

- stage 0: Two borrowers form a group and then each is randomly endowed with an investment project with productivity k_i, $i \in \{H, L\}$;
- stage 1: The lender offers a group loan contract with terms (L, R);
- stage 2: Each borrower's investment is launched and her project output, Y_i is realized one period later; the project output is unobserved/non-verifiable to the lender.
- stage 3: Repayment or default decisions are made by each group member and all payoffs are realized (see below for details).

To save on notation, we will say that a group has type ij if the productivities of its two members' projects are k_i and k_j where $i, j \in \{H, L\}$. There are three possible group types: HH, LL, HL. Without loss of generality assume $k_i \geq k_j$. We say that a group loan contract is *feasible* when each member's project generates enough output upon success to be able to cover repayment both when R (paying for oneself) and when $2R$ (paying for oneself and one's partner) is due. Feasibility requires $\min\{k_i L, k_j L\} \geq 2R$, or

$$R \leq \tfrac{1}{2} k_j L \tag{1}$$

In our baseline setting we assume that each borrower makes the decision to default or repay individually (non-cooperatively).[8] Appendix D considers the alternative assumption of two borrowers deciding on strategic default jointly. Similarly to Besley and Coate (1995), we model the borrowers' repayment decisions as a two-stage game. In the first stage, each borrower is asked to pay R by the lender and can either do so (repay) or report failure (default). Since the project outcome is non-verifiable by the lender, default could be either involuntary, if one's project fails, or strategic, if this is beneficial for a borrower. If both borrowers repay R or both default in this stage, their respective payoffs are realized (see below) and the game ends. The second stage is reached only if one borrower has repaid R in stage 1 while her partner has defaulted. In such case, the borrower who has repaid is asked by the lender to pay an additional amount R for her partner,

[8] To justify the non-cooperative nature of the game one can think that each borrower can only verify the output realization of her own project, as modeled in Armendariz de Aghion (1999) and others, or that sanctions within the group and period cannot be imposed.

as required by the joint liability clause. Again, the borrower decides to repay or default. We look for a subgame-perfect Nash equilibrium of the described game.

Start the analysis with stage 2 – the decision facing a borrower with a successful project who has already repaid R in stage 1. Note that it is never optimal to make a partial repayment (strictly between 0 or R) since since either defaulting (repaying zero) and forfeiting the entire future value V, or repaying in full and securing the future value V is a dominant strategy. The borrower will thus repay if her payoff, $-R+V$ is larger than that of defaulting, 0. Stage 2 repayment is hence optimal whenever

$$R \leq V \quad \text{(S2R)}$$

Suppose condition (S2R) holds, so any group member would repay if stage 2 is reached, and proceed by backward induction to stage 1. Conditional on project success, the borrowers play a simultaneous move game with normal form presented in Table 1. Only the payoffs of the row player i are listed; those of the column player j are symmetric around the diagonal. If one's project fails one simply announces default. Partial repayment is dominated by either defaulting or repaying in full (assume the lender considers partial repayment the same as default).

	Repay	Default
Repay	$k_iL - R - (1-p)R + V$	$k_iL - 2R + V$
Default	$k_iL + pV$	k_iL

Table 1: The normal form of the game in the repayment stage.

Conditional on borrower j playing Repay in stage 1, borrower i would also choose Repay when her project succeeds if her expected payoff from paying R is larger than her payoff from strategically defaulting (repaying zero), that is, if $k_iL - R - (1-p)R + V \geq k_iL + pV$, or

$$(2-p)R \leq (1-p)V \tag{2}$$

The repayment payoff reflects the fact that, under condition (S2R), a borrower playing Repay at stage 1 while her partner also plays Repay, will only be asked to repay an extra R at stage 2 with probability $1-p$ (if her partner's project fails).[9] Condition (2) implies the following constraint on R:

$$R \leq \tfrac{1-p}{2-p}V. \tag{3}$$

Similarly, conditional on member j choosing Default, member i would choose Repay if paying back $2R$ over the two stages and securing the future value V results in a larger payoff than default, that is, if $k_iL - 2R + V \geq k_iL$, or

$$R \leq \tfrac{1}{2}V. \tag{4}$$

Since condition (3) implies conditions (4) and (S2R), the profile (Repay, Repay) is the unique Nash equilibrium of the stage 1 game (Repay is at least weakly dominant strategy), and repayment is also optimal at stage 2, whenever $R \leq \frac{1-p}{2-p}V$. In general, all pure strategy Nash equilibria of the stage 1 game for given loan terms (L, R) are: (i) (Repay, Repay) if $R \leq \frac{1-p}{2-p}V$; (ii) (Repay, Default) or (Default, Repay) if $\frac{1-p}{2-p}V < R \leq \frac{1}{2}V$; and (iii) (Default, Default) if $R > \frac{1}{2}V$. Here, we focus on

[9]Here, consistent with the simultaneous move assumtion, we assume the borrower makes the stage 1 repayment decision without observing the partner's project outcome. The analysis would not change qualitatively if instead the borrower knew that her partner's project has failed, in which case the no default condition becomes $R \leq \frac{(1-p)V}{2}$.

the (Repay, Repay) equilibrium. In Appendix B we show that a lender could potentially offer a contract with $\frac{1-p}{2-p}V < R \leq \frac{1}{2}V$ inducing the (Repay, Default) outcome at stage 1, and break even by setting $L = pR$. However, we prove that such contract is always weakly dominated by inducing a bogus group when we allow the formation of bogus groups – see Section 3. Thus, our focus here on (Repay, Repay) equilibria is not restrictive and does not affect the main results of the paper.[10]

Given contract (L, R) satisfying conditions (1) and (3), the expected total payoff of a standard group of type ij equals

$$W_{ij}(L, R|S) = p(k_i + k_j)L - 2p(2-p)R + 2p(2-p)V, \tag{SEP}$$

where $|S$ indicates that the group form is standard.

Forming a standard joint-liability group increases the repayment probability from p (the project success probability) to $1-(1-p)^2 = p(2-p) > p$. This corresponds to the classic argument in favor of joint liability lending relative to individual liability lending. Due to limited liability, the lender receives the repayment $2R$ with probability $p(2-p)$ either from both or from a single borrower, and receives nothing otherwise. Hence, the lender's break-even condition is $2p(2-p)R - 2L \geq 0$, or

$$R \geq \tfrac{L}{p(2-p)}. \tag{5}$$

Given the free-entry assumption, the optimal standard group loan contract can be thus defined as the loan size and total repayment, (L_S, R_S) which maximize the expected joint payoff of the borrower group, that is, the pair (L_S, R_S) which solves:

$$\max_{L,R}\ W_{ij}(L, R|S) \tag{6}$$

$$\text{s.t. } (1), (3), \text{ and } (5).$$

It is clear that (5) must bind at optimum, that is $R = \frac{L}{p(2-p)}$ – if not, another lender can offer a contract with lower repayment R and still break-even. The two constraints (1) and (5) are then collectively equivalent to $\frac{L}{p(2-p)} = R \leq \frac{1}{2}k_j L$, which is non-empty only if $k_j \geq \frac{2}{p(2-p)}$. If $k_j < \frac{2}{p(2-p)}$, feasibility is violated and, when called upon, a k_j borrower would not be able to repay in case their project succeeded but their partner's project failed. We thus make the following assumption which ensures that the project productivities are large enough so that repayment is always feasible.[11]

Assumption 1:

$$k_L \geq \tfrac{2}{p(2-p)}$$

Note that Assumption 1 implies

$$pk_i > 1 \text{ for } i = L, H \tag{SE}$$

that is, both projects are socially efficient – the expected payoff per dollar invested in each strictly exceeds the opportunity cost of funds.

The following proposition characterizes the optimal standard group contract which solves problem (6).

[10] In addition, supporting the (Repay, Default) equilibrium may be hard in practice – it either means that one borrower consistently 'exploits' the other by free-riding on her repayment decision, or the borrowers take turns over time, which requires co-ordination and commitment at odds with the non-cooperative repayment decision assumption. See also Appendix D.

[11] Contrast with Besley and Coate (1995)

Proposition 1

(a) The optimal standard group loan contract that solves problem (6) is $\mathcal{S} \equiv (L_S, R_S)$ *with*

$$L_S = p(1-p)V \text{ and } R_S = \tfrac{1-p}{2-p}V.$$

(b) The optimal contract is the same as in part (a) when the borrowers' productivities are unobservable to the lender.

Proof: see Appendix

The intuition for part (b) is that the contract terms are set by the break-even and no-default conditions, neither of which depends on k_i and k_j since the group members' expected output does not affect their repayment incentives. The fact that the constraints are independent of the productivities also implies that the assumption that both borrowers receive the same loan size even though they may have different k's is not restrictive.

2.3 Allowing for bogus groups

A standard assumption in the literature on joint-liability lending, and also standard practice in microfinance, is that each borrower is expected to invest in her own business project. However, as motivated by the evidence from rural China reviewed in the introduction, suppose that lenders are unable to enforce or verify the requirement that each group member invests in her own project. Therefore, borrowers can choose to form either a "standard group" (described above), or a "bogus group", that is a group in which all loaned funds, $2L$ is invested into a *single* business project run by one of the borrowers. The other borrower is a cosigner in the eyes of the lender (joint liability still applies) but, by the limited liability assumption, she has no income or wealth to go after in case of her partner declaring default. Essentially, in this paper we extend the standard group-lending framework by allowing the endogenous choice of the group form – standard or bogus.

Allowing for endogenous choice between a standard or bogus group introduces a new step in the model's timing – given contract (L, R) the borrowers can decide to form a standard or bogus group. The rest of the timing remains the same as in Section 2.2.

In the analysis below we focus on maximizing the total group payoff, that is, borrowers would only choose to form a bogus group if it is jointly beneficial for them. Given our limited enforcement setting, one way to operationalize this idea is to assume, as in the literature, that the group members share some social capital which can be used to enforce a transfer T (which could possibly be non-monetary) between them, to ensure the co-signer obtains at least the same payoff as she would in a standard group, under repayment.[12] More precisely, assume that the borrower who invests $2L$ (the bogus group 'leader') makes the default/repay decision *individually*, based on his own payoff, as in the standard group case, but he does so with the understanding that the transfer T has to be made in either case. The repay/default trade-off is thus unaffected by the transfer. Appendix D considers the alternative case in which the borrowers make the repayment decision *jointly*.

Consider first the endogenous group form decision for given loan terms (L, R). If the borrowers form a bogus group, the two members first decide in which project to invest. It is clear that the borrowers would optimally invest all the loaned funds ($2L$) into the project with higher productivity. Next, let us look at the repay/default decision in a bogus group for given (L, R). As in Section 2.2, it is not optimal to repay partially. Repayment at stage 2 (reached if the leader repays in stage 1)

[12] It is easy to compute the required transfer amount and show that it is always feasible upon project success, both with repayment or strategic default. Alternatively, the transfer could be non-monetary.

is optimal as long as $R \leq V$ yielding a payoff of $V - R$. The cosigner has no decisions to make since her project is not funded and there is limited liability.

Assume $k_i \geq k_j$ without loss of generality, so all funds are invested in project i. Conditional on project success, the bogus group investor's payoff from repaying in Stage 1 is larger than the payoff from defaulting if

$$k_i(2L) - R - T + (V - R) \geq k_i(2L) - T$$

or,

$$R \leq \frac{V}{2} \tag{7}$$

This implies that, for given loan terms (L, R), the joint expected payoff, $W_{ij}(L, R|B)$ in a bogus group with project productivities k_i, k_j is:

$$W_{ij}(L, R|B) = \begin{cases} 2pk_iL - 2pR + 2pV & \text{if } R \leq \frac{V}{2} \text{ (Repay)} \\ 2pk_iL & \text{if } R > \frac{V}{2} \text{ (Default)} \end{cases} \tag{8}$$

When a bogus group is formed the lender is repaid $2R$ with probability p and zero otherwise. Hence, the lender's break-even condition on lending $2L$ in total is

$$pR \geq L \tag{9}$$

Free entry therefore implies $L = pR$.

2.3.1 Standard vs. bogus joint liability groups

There are thus five main differences between bogus and standard groups, as listed below. For fixed loan terms (L, R) we have:

(i) *insurance/risk-sharing*: in a standard group a person still receives the continuation value V if her project fails but her partner's project succeeds. The project output i.i.d. assumption is important here – the joint liability clause acts as a risk-sharing mechanism. The expected continuation value is $p(2-p)V$ in a standard group vs. pV in a bogus group. On the flip side, borrowers in a standard group repay more in expectation, $p(2-p)R$ vs. pR in a bogus group since standard group members cover for their partners. The net expected benefit, defined as the difference between the increased future value and the larger repayment, is positive if $R < V$ (see the discussion after Lemma 1 below).

(ii) *expected output:* is larger in a bogus group with heterogeneous borrowers than in a standard group with the same borrowers.

More generally, when the terms (L, R) could differ for standard and bogus groups:

(iii) *the interest rate:* is strictly lower in a standard joint liability group, $\frac{1}{p(2-p)}$ than in a bogus group, $\frac{1}{p}$. The reason is that the lender is paid with probability $1 - (1-p)^2$ in a standard group as two i.i.d. projects are funded vs. probability p in a bogus group. However, for the same nominal amount due R, This means that the "effective" interest rate is the same in both group forms.

(iv) *strategic default incentive*: is stronger in a standard group since, as shown in Section 2.2, if the loan size is large ($L > p(1-p)V$) a person has the incentive to strategically default and free ride on their partner's repayment knowing that she will be bailed out by their partner who wants to obtain the future value V. In a bogus group this free riding effect is absent since the partner has no income.[13]

(v) *loan size:* can be larger in bogus groups since the no-default condition (7) is weaker than the corresponding condition (3) for standard groups due to the free riding effect described in the previous point. Specifically, the break-even and no default conditions imply $L \leq p(1-p)V$ in standard groups and $L \leq \frac{pV}{2}$ in bogus groups (the latter is larger for $p > 1/2$).

Item (i), the net insurance benefit, favors standard groups whenever $R \leq V$ due to the risk-sharing from investing in two projects. The latter is also the reason for the lower standard group interest rate, item (iii). On the other hand items (ii), (iv) and (v) favor bogus groups by mitigating the strategic default incentives and/or allowing borrowers to benefit from larger loans and higher return projects.

2.3.2 Choice to form bogus group for given (L, R)

We proceed to formally characterize the trade-off between forming a standard vs. bogus joint liability group for given loan terms (L, R). We continue to focus on the (Repay, Repay) equilibrium for standard groups (that is, $R \leq \frac{1-p}{2-p}V$), which also implies repayment in bogus groups. See Appendix B for relaxing this. Since we hold the terms fixed, this comparison takes into account only items (i) and (ii) from the list in Section 2.3.1.

> **Lemma 1**: *Given group-lending loan contract (L, R), it is optimal for borrowers with projects with productivities k_i and k_j to form a bogus group instead of a standard group if and only if,*
>
> $$p(k_i - k_j)L > 2p(1-p)(V - R). \tag{10}$$
>
> **Proof:** see Appendix

Intuitively, forming a standard instead of a bogus group resembles buying an insurance – item (i) in the list in the previous section. For a borrower in a standard group, when her project succeeds (with probability p), she obtains the future value V as long as she repays. However, conditional on project success, she needs to repay in expectation $(2-p)R$ in a standard group ($2R$ in case of partner project failure and R in case of partner project success) instead of $2R/2 = R$ (per the same loan amount L) in a bogus group. The difference $(2-p)R - R = (1-p)R$ can be thought of as the insurance premium when a borrower's project succeeds (with probability p). If a borrower's project fails on the other hand, due to limited liability she pays nothing but would receive pV in expectation in a standard group as opposed to 0 in a bogus group. Thus, pV is the risk sharing benefit of being in a standard group when a borrower's project fails (with probability $1-p$). The ex-ante total net benefit (insurance value) from being in a standard group for both borrowers is hence

$$2(1-p)pV - 2p(1-p)R = 2p(1-p)(V-R) \tag{11}$$

which is positive for $V > R$ guaranteed by the no-default condition (3).

If the borrowers form a bogus group instead, they optimally invest all loaned funds ($2L$) into the higher productivity project, k_i. Total expected output thus increases by $p(k_i - k_j)L$ compared

[13] Note that deciding to strategically default in a bogus group does not require more "coordination" compared to the same decision in a standard group. In a standard group the joint liability (JL) clause means that if I decide to strategically default, the lender can go to the other member and collect $2R$ if he declares success. In a bogus group JL means that if I decide to default, the lender can still go to the other person but cannot collect anything due to limited liability.

to in a standard group. However, the borrowers forego the insurance benefit of being in a standard group described above.

Note that, for a homogeneous ii (HH or LL) groups, the LHS of (10) in Lemma 1 is zero and so this condition is never satisfied for such groups – forming a bogus group does not offer any benefit in terms of additional project return while still requires foregoing the risk-sharing value inherent in a standard joint liability group. Therefore condition (10) holds only for heterogeneous (HL) groups and only if the productivity difference $k_H - k_L$ is sufficiently large so that the extra output benefit of forming a bogus group overwhelms the loss of insurance (risk sharing) in a standard group.

2.3.3 Bogus groups – a problem?

The following Proposition uses Lemma 1 to show the consequences of bogus group endogenously forming if the lender offers the optimal standard group contract $\mathcal{S}$.

Proposition 2: *At the standard group loan contract $\mathcal{S} = (L_S, R_S)$, if*

$$\Delta \equiv k_H - k_L > \tfrac{2}{p(2-p)}, \tag{12}$$

then:
(a) all HL borrower pairs optimally form bogus groups
(b) all HL groups cause a loss to the lender.

Proof: see Appendix

Proposition 2 demonstrates the problem a lender would have from not taking into account the possibility of endogenous formation of bogus groups – she would no longer break even using contract (L_S, R_S) from Section 2.2. The intuition is clear – for given (L, R), HL pairs are the only pairs which can benefit from the higher expected output in a bogus group. This happens if the productivity differential Δ is large enough to offset the loss of expected future value relative to forming a standard group.

Proposition 2 also implies that if condition (12) does not hold (k_H and k_L are relatively close), then offering contract (L_S, R_S) would not cause a loss to the lender since all groups will be standard. However, offering this contract to all borrowers may not be optimal if in fact allowing bogus groups to form is welfare-improving due to the benefit from investing larger amounts of funds in the higher return project (items ii, iv and v in the list in Section 2.3.1). In the following section we thus proceed to address the question of what is the optimal loan contract when the lender recognizes that bogus groups can form. We discuss both the scenario when condition (12) holds, which may be more relevant in practice for lenders who are losing money by failing to acknowledge bogus groups and also the scenario when it does not hold – lenders would not lose money by offering contract (L_S, R_S) but could be leaving money on the table by failing to capitalize on the productivity benefits which bogus groups offer.

3 Optimal loan terms with endogenous group form

Suppose the lenders acknowledge the possibility and allow the existence of bogus groups and would like to design the optimal loan contract taking into account the borrowers' hidden action (the choice to form bogus or standard group). We start with the case in which the borrowers' productivities are assumed observable to the lender. For example, we can think of the lender knowing the intended use of the loans, e.g., agriculture, retail, etc. Then, in Section 3.2 we relax this assumption and assume that the lender does not observe k_i, k_j.

3.1 Optimal loan terms – observed productivity

The lender designs an optimal contract contingent on the (observed) group type ij. The timing is as follows:

- stage 0: two borrowers form a group and then each borrower is endowed with a project with productivity k_i, $i \in \{H, L\}$, which is observable to the lender;
- stage 1: given the productivities k_i, k_j the lender offers a contract (L, R)
- stage 2: the borrowers choose to form either a standard or bogus group;
- stage 3: investment is launched and output is realized;
- stage 4: repayment decisions are made by the borrowers, as explained in Section 2; all payoffs are realized.

By the free entry assumption, facing a group of productivity type ij the lender would offer the loan contract which maximizes the group's total payoff subject to the lender breaking even. When k_i and k_j are observed, the lender knows the optimal group form choice (standard or bogus) for any (L, R) and thus offers the lending contract which induces the payoff-maximizing choice.[14] Basically, the lender uses the contract terms (loan size and interest rate) to induce either a standard or bogus group, whichever is total payoff maximizing for the borrowers. Of course, the lender's zero-profit condition and the no-default constraint depend on the group form induced via the loan terms. Below we suppress the indexation by ij for notational simplicity.

Given k_i, k_j and p the optimal contracting problem can be written as:

$$\max_{L,R,\tau\in\{0,1\}} \tau W(L,R|S) + (1-\tau)W(L,R|B) \tag{OP}$$

$$\text{s.t. } \tau W(L,R|S) + (1-\tau)W(L,R|B) \geq \tau W(L,R|B) + (1-\tau)W(L,R|S) \quad \text{(IC)}$$

$$R \leq \tau\frac{(1-p)V}{2-p} + (1-\tau)\frac{V}{2} \quad \text{(no default)}$$

$$R = \tau\frac{L}{p(2-p)} + (1-\tau)\frac{L}{p} \quad \text{(zero profit)}$$

where the total payoffs for a standard group, $W(L, R|S)$ and bogus group, $W(L, R|B)$ are:[15]

$$W(L,R|S) = \begin{cases} p(k_i+k_j)L - 2p(2-p)R + 2p(2-p)V & \text{if } R \leq \frac{1-p}{2-p}V \text{ (Repay,Repay)} \\ p(k_i+k_j)L - 2pR + 2pV & \text{if } \frac{1-p}{2-p}V < R \leq \frac{V}{2} \text{ (Repay,Default)} \\ p(k_i+k_j)L & \text{if } R > \frac{V}{2} \text{ (Default, Default)} \end{cases} \tag{13}$$

and

$$W(L,R|B) = \begin{cases} 2pk_iL - 2pR + 2pV, & \text{if } R \leq \frac{V}{2} \text{ (Repay)} \\ 2pk_iL, & \text{if } R > \frac{V}{2} \text{ (Default)} \end{cases} \tag{14}$$

The group form indicator variable τ allows a compact way of writing the problem of picking the larger of the two maximized (over L, R) total payoffs, $W(L, R|S)$ and $W(L, R|B)$.

[14] This implies that even if the group form were assumed enforceable, the lender would offer the exact same contract described here. Of course, if the lender had monopoly power the enforceability of group form may matter.

[15] We exhibit the total payoff under all possible NE for standard groups even though (Repay, Default) and (Default, Default) will not occur at the optimal contract. The reason is the hidden action of borrowers choosing the group form which requires evaluating deviations from the prescribed behaviour.

Start by observing that for, any k_i, k_j, p, the (Repay, Default) equilibrium in a standard group is always weakly dominated by the (Repay) equilibrium in a bogus group. The reason is that the interest rate and the upper bound on loan size are the same (see Section 2.3 and Appendix B), however, a weakly larger expected output can be achieved in the (Repay) bogus group outcome relative to the (Repay, Default) standard group outcome. Hence, it is not optimal for a lender to induce a (Repay, Default) equilibrium in a standard group.[16] We hence write the no-default and zero-profit constraints in problem (OP) only for the (Repay, Repay) equilibrium, whenever it is optimal to induce a standard group. As before, the zero profit constraint must hold at equality due to the free-entry assumption – if the lender made positive profit, then L can be increased or R reduced to increase the objective function.

Constraint (IC) guarantees that if, for example, the lender finds it optimal to choose (L, R) to induce a standard group (choose $\tau^* = 1$), then the borrowers must indeed prefer to form a standard group, i.e., $W(L, R|S) \geq W(L, R|B)$ must hold, and similarly, the opposite inequality must hold if inducing a bogus group is optimal for the given k_i, k_j, p.

Proposition 3: *Suppose Assumption 1 holds, that is $k_L \geq k_{\min} \equiv \frac{2}{p(2-p)}$, and the borrowers' productivities k_i and k_j are observed by the lender. The optimal loan terms (L^*, R^*) for a (k_i, k_j) group are:*
(a) for homogeneous, ii (HH or LL) groups: if $p(2p-1)k_i > 1$ then $L^* = \frac{pV}{2} \equiv L_B$, $R^* = \frac{V}{2} \equiv R_B$ and the group is bogus ($\tau^* = 0$) while if $p(2p-1)k_i \leq 1$ then $L^* = p(1-p)V = L_S$, $R^* = \frac{(1-p)V}{2-p} = R_S$ and the group is standard ($\tau^* = 1$).
(b) for heterogeneous, ij (HL) groups: depending on the parameter values,[17] either $L^* = L_B$ and $R^* = R_B$ and the group is bogus ($\tau^* = 0$) or $L^* = \min\{L_S, L_E\}$, $R^* = \frac{L^*}{p(2-p)}$ and the group is standard ($\tau^* = 1$), where $L_E \equiv \frac{p(1-p)V}{\frac{1-p}{2-p} + \frac{p}{2}(k_H - k_L)}$.

The following Corollary exhibits the exact mapping between the model parameters (k_i, k_j, p), the loan terms (L^*, R^*), and the group form.

Corollary 1 to Proposition 3: *Suppose Assumption 1 holds, that is $k_L \geq k_{\min} \equiv \frac{2}{p(2-p)}$, and let L_B, R_B, L_S, R_S, and L_E be as defined in Propositions 1 and 3. Let also $R_E \equiv \frac{L_E}{p(2-p)}$, $\hat{k} \equiv \frac{2-3p}{p(2-p)(2p-1)}$, $\tilde{k} \equiv \frac{1}{p(2p-1)}$, and the functions $d(k) \equiv \frac{2}{p(2-p)} + k$, $g(k) \equiv \frac{1}{p^2} + \frac{1-p}{p}k$, and $f(k)$ be as defined in the proof of Proposition 3. Defining $\mathcal{S} \equiv (L_S, R_S)$, $\mathcal{B} \equiv (L_B, R_B)$, and $\mathcal{E} = (L_E, R_E)$, the optimal loan contract for every possible group composition, LL, HH or HL, and all possible parameter configurations are exhibited in the Table below:*

[16] Obviously, setting (L, R) to induce (Default, Default) in a standard group or (Default) in a bogus group is not compatible with the break-even condition, so it is not optimal either.

[17] The exact conditions are exhibited in the Corrolary below.

Parameter conditions	**Optimal loan and group form**		
	LL groups	HH groups	HL groups
(a) high productivities or high p: $p(2p-1)k_L > 1$	$\mathcal{B}$, *bogus*	$\mathcal{B}$, *bogus*	$\mathcal{B}$, *bogus*
(b) high k_H *and low* k_L*:* $p(2p-1)k_H > 1 \geq p(2p-1)k_L$	$\mathcal{S}$, *standard*	$\mathcal{B}$, *bogus*	$\mathcal{B}$, *bogus*
(c) low productivities or low p: $p(2p-1)k_H \leq 1 \wedge$			
– *(i)* $k_H \in [k_L, \min\{g(k_L), d(k_L)\}]$	$\mathcal{S}$, *standard*	$\mathcal{S}$, *standard*	$\mathcal{S}$, *standard*
– *(ii)* $k_H \in (d(k_L), f(k_L)] \wedge \{(k_L < \hat{k} \wedge p \in (\frac{1}{2}, \frac{4}{7})) \vee p \leq \frac{1}{2}\}$	$\mathcal{S}$, *standard*	$\mathcal{S}$, *standard*	$\mathcal{E}$, *standard*
– *(iii-1)* $k_H > g(k_L) \wedge k_L \geq \hat{k}$	$\mathcal{S}$, *standard*	$\mathcal{S}$, *standard*	$\mathcal{B}$, *bogus*
– *(iii-2)* $k_H > f(k_L) \wedge \{(k_L < \hat{k} \wedge p \in (\frac{1}{2}, \frac{4}{7})) \vee p \leq \frac{1}{2}\}$	$\mathcal{S}$, *standard*	$\mathcal{S}$, *standard*	$\mathcal{B}$, *bogus*

Proof: (see Appendix)

[*Figure 1 here*]

3.1.1 Discussion

1. We see from Corrolary 1 and its proof that in homogeneous pairs bogus groups arise when either the project success probability, p is very high ($p > \frac{4}{5}$) or for intermediate success probability ($p \in (\frac{1}{2}, \frac{4}{5}]$) and high project productivity, $k_i > \tilde{k}$. Intuitively, the gains from forming a bogus group are largest when the risk of failure is relatively low (high p) and/or when the project productivity is large. Standard groups are optimal otherwise.

2. For heterogeneous pairs, contract $\mathcal{B}$ is offered and bogus groups optimally form when: (i) p is very high ($p > \frac{4}{5}$), or (ii) the productivities are large enough, or (iii) sufficiently large productivity difference between the projects (when $k_H > \tilde{k} \geq k_L$ or $k_H > f(k_L) > k_L$ or $k_H > g(k_L) > k_L$). Scenarios (i) and (ii) map into part (a) in the Table, while (iii) maps into parts (b) and (c-iii). In part (b) what is crucial is that the high-productivity project is sufficiently profitable ($k_H > \tilde{k}$) – the value of k_L does not matter since all funds are invested in the k_H project. Scenario (iii) reflects the borrowers' benefit from the extra output that can be generated in a bogus group compared to in a standard group, per dollar lent. This effect obviously does not arise in homogeneous groups (see part a).

Conversely, if the project productivities k_H and k_L are close to each other (the case $k_H \in [k_L, \min\{g(k_L), d(k_L)\}]$ in line (c-i) of the Table), then offering contrast $\mathcal{S}$ inducing a standard group is optimal. We also see that there arises an intermediate case with optimal contract $\mathcal{E}$ and loan size, $L_E < L_S$ (line c-ii). Contract $\mathcal{E}$ also induces a standard group and trades off the smaller loan size with the benefit of a lower interest rate.[18] The reason for the reduced loan size is the need to respect the incentive compatibility constraint arising from the unobservability of the group form decision. This case can only arise for relatively low success probability, $p < 4/7$ and an intermediate productivity differential, $f(k_L) \geq k_H > d(k_L) > k_L$.

3. The optimal contract inducing a bogus group is always (L_B, R_B) with terms independent of the productivities k_i, k_j. The intuition is that the loan size L_B is determined by the no-default constraint (in which k_i, k_j enter on both the benefit and cost sides) while the interest rate (R_B/L_B) is determined by the lender's break-even condition. The same applies for the (L_S, R_S) contract. The only contract with terms dependent on the productivity values k_i, k_j is the reduced-loan-size contract $\mathcal{E}$. Intuitively, to deter an HL group from switching to bogus it is necessary to know the

[18] For these parameters we have: $W(L_S, R_S|B) > W(L_S, R_S|S) > W(L_E, R_E|S) \geq W(L_B, R_B|B)$ so the reduced loan size is used to satisfy the (IC) and the lender's break even constraint.

exact output gain from changing the group form which depends on the difference between k_H and k_L.

4. The analysis shows that borrower welfare and productive efficiency in the economy are raised if the possibility of endogenously arising bogus groups is acknowledged by the lender and such groups are offered the appropriate loan contract, as opposed to 'ignoring' bogus groups and offering contract (L_S, R_S) only. This result is true both when the lender would lose money (Proposition 2) but also more generally, as shown in Proposition 3 and Corollary 1, by benefitting from the larger loans (if $p > 1/2$) or additional output (for HL pairs) that obtain in bogus groups.

3.2 Optimal loan terms – unobserved productivity

In this section we assume that the borrowers' project productivities $k_i, k_j \in \{L, H\}$ are unobserved by the lender. The lender does know the values k_L and k_H, that is the return of a high- vs. a low-productivity project. Assume also that the lender knows what fraction (mass) of agents, q_{ij} belong to HH, LL and HL groups.[19] For example, the lender may only know a broad description of the project but not the actual ability of the borrower who implements it, etc. As we show below, the main results and intuition from Section 3.1 where productivity was assumed observable generalize only with minor modifications. We continue to assume that group members freely choose to form either standard or bogus groups after observing the contract(s) the lender offers.

Observe first that clearly the lending contract from Proposition 3 is no longer feasible since the lender cannot observe the group members' productivities k_i and k_j and hence cannot offer contracts contingent on ij as described in Corollary 1. The reason is that when k_i, k_j are unobserved, the borrowers have the incentive to report the group composition ij that would maximize their payoff, which may not necessarily be the true composition. For example, the borrowers in an HL group who would be offered contract (L_E, R_E) in Proposition 3 could lie that they are both type k_H, receive contract (L_S, R_S) and then form a bogus group which would result in a loss to the lender.

Unlike in Proposition 3 where the lender could simply offer the (single) best contract to each possible borrower pair, with unobservable $k's$ the lender's problem is therefore to offer a *menu of contracts* designed such that no borrower pair has an incentive to misreport its productivity type (LL, HH or HL) and, in addition, no group has incentive to choose a group form (bogus vs. standard) different from the form the contract is designed to induce. Note that the resulting mechanism design problem is more complex than a standard screening or adverse selection problem since the lender faces both unobserved types (the k_i, k_j pairing) and an unobserved action (ex-post moral hazard).

We start the analysis by noticing that, because of the free-entry assumption, the lender's break-even condition for each of the two possible group forms implies that the interest rate on any loan contract (L, R) cannot be different from either $\frac{1}{p(2-p)}$ (if standard group is induced) or $\frac{1}{p}$ (if bogus group). Any lower interest rate would result in a loss to the lender while any higher interest rate will be competed away.

Second, remember from the previous analysis that, for any lending terms $(L, \frac{L}{p})$ or $(L, \frac{L}{p(2-p)})$, the joint payoff of any group is strictly increasing in the loan size L regardless of the productivities of its members.[20] This implies that, within each contract type defined in terms of its interest rate, $(L, \frac{L}{p})$ or $(L, \frac{L}{p(2-p)})$, only one loan size can be offered, namely the size that maximizes the borrowers' joint payoff subject to the no default and break-even constraints. In other words, there

[19] For example, if the project productivities are i.i.d. after the group is formed and both k_L and k_H can be drawn with equal probability we would have that the mass of HH and LL groups is 1/4 while the mass of HL groups is 1/2.

[20] This statement also includes the scenario when in a standard group a larger loan size would result in a switch from the (Repay, Repay) to the (Repay, Default) equilibrium.

is no way to screen across the ij borrower pairs by loan size, only by the interest rate. If two distinct loan sizes were offered both carrying the same interest rate, any borrower pair prefers to take the larger loan.

These two observations imply that when the borrowers' productivities are unobservable, the lender can offer a menu consisting of at most two different contracts, $\mathcal{N} \equiv (L_N, R_N)$ and $\mathcal{M} \equiv (L_M, R_M)$ which induce a standard group and a bogus group respectively. The contracts $\mathcal{N}$ and $\mathcal{M}$ must be such that:

(a) any group ij that selects contract $\mathcal{N}$ chooses to be standard and any group ij that selects contract $\mathcal{M}$ chooses to be bogus;

(b) no borrower defaults strategically;

(c) the lender breaks even for each contract $\mathcal{N}$ and $\mathcal{M}$ (no cross-subsidization); and

(d) the ex-ante expected total borrower payoff is maximized.

$$\max_{L_N, R_N, L_M, R_M} \sum_{ij} q_{ij} W_{ij}(L_N, R_N, L_M, R_M) \tag{UPP}$$

subject to

$$R_M \leq \tfrac{V}{2} \tag{15}$$

$$R_M = \tfrac{L_M}{p} \tag{16}$$

$$R_N \leq \tfrac{1-p}{2-p} V \tag{17}$$

$$R_N = \tfrac{L_N}{p(2-p)} \tag{18}$$

$$\bar{W}_{ij}(L_N, R_N, L_M, R_M) \geq \max\{W_{ij}(L_N, R_N|B), W_{ij}(L_M, R_M|S)\}, \ \forall ij \in \{HH, HL, LL\} \tag{IC}$$

where $\bar{W}_{ij}(L_N, R_N, L_M, R_M) \equiv \max\{W_{ij}(L_N, R_N|S), W_{ij}(L_M, R_M|B)\}$ and the joint payoff function $W_{ij}(\cdot, \cdot|S)$ and $W_{ij}(\cdot, \cdot|B)$ are as defined in (13) and (14) respectively.

The first four constraints are the no-default and break-even constraints for each contract in the menu. The final constraint ensures that any group ij would select the contract, $\mathcal{N}$ or $\mathcal{M}$ and the group form that maximizes its payoff. Selecting the alternative contract, or deviating to the alternative group form, or both, is not optimal.

As in Section 3.1, note that contract $\mathcal{N}$, designed for standard groups, would never induce the (Repay, Default) equilibrium, that is, choose $\frac{1-p}{2-p}V < R_N \leq \frac{V}{2}$. The reason is that in the latter case the interest rate would be equal to that of contract $\mathcal{M}$ using the break-even constraint. Hence, for such R_N the functional form of $W_{ij}(\cdot, \cdot|S)$ is the same as that of $W_{ij}(\cdot, \cdot|B)$ and the no-default constraint is the same as that for contract $\mathcal{M}$. This implies that any such contract is weakly dominated by contract $\mathcal{M}$. That is why, in problem (UPP) above, we only search for the optimal contract $\mathcal{N}$ in the (Repay, Repay) range defined by constraint (17).

Substituting in for R_N and R_M from the break-even constraints, problem (UPP) can be simplified as:

$$\max_{L_N, L_M} \sum_{ij} q_{ij} W_{ij} \left(L_N, \tfrac{L_N}{p(2-p)}, L_M, \tfrac{L_M}{p} \right) \tag{UPP'}$$

subject to

$$L_M \leq \tfrac{pV}{2} \tag{19}$$

$$L_N \leq p(1-p)V \tag{20}$$

$$\max\left\{W_{ij}\left(L_N, \frac{L_N}{p(2-p)}|S\right), W_{ij}\left(L_M, \frac{L_M}{p}|B\right)\right\} \geq \tag{IC}$$
$$\geq \max\left\{W_{ij}\left(L_N, \frac{L_N}{p(2-p)}|B\right), W_{ij}\left(L_M, \frac{L_M}{p}|S\right)\right\} \textit{ for all } ij \in \{HH, HL, LL\}$$

and where

$$W_{ij}\left(L_N, \frac{L_N}{p(2-p)}|S\right) = (p(k_i + k_j) - 2)L_N + 2p(2-p)V \tag{21}$$

$$W_{ij}\left(L_N, \frac{L_N}{p(2-p)}|B\right) = 2\left(pk_i - \frac{1}{2-p}\right)L_N + 2pV \tag{22}$$

$$W_{ij}\left(L_M, \frac{L_M}{p}|B\right) = 2(pk_i - 1)L_M + 2pV \tag{23}$$

$$W_{ij}\left(L_M, \frac{L_M}{p}|S\right) = \begin{cases}(p(k_i + k_j) - 2(2-p))L_M + 2p(2-p)V & \text{if } L_M \leq \frac{p(1-p)}{2-p}V \text{ (Repay,Repay)} \\ (p(k_i + k_j) - 2)L_M + 2pV & \text{if } L_M \in (\frac{p(1-p)}{2-p}V, \frac{pV}{2}] \text{ (Repay,Default)}\end{cases} \tag{24}$$

Lemma 2: *Constraint (19) is binding at the solution to problem (UPP'), that is,* $L_M = \frac{pV}{2}$.

Proof: see Appendix

Lemma 2 implies that the optimal contract designed for bogus groups always features the maximum possible loan size, $\frac{pV}{2}$ that ensures repayment. The $\mathcal{M}$ contract indended for bogus groups in the unobserved productivity setting is hence the same as contract (L_B, R_B) from Section 3.1. The intuition is that the marginal payoff increase for bogus groups is always larger, for the same contract, than that in standard groups, hence the incentive-compatibility and selection constraint (IC) are (weakly) relaxed by raising L_M to its maximum possible value.

> **Proposition 4**: *Suppose Assumption 1 holds and the borrowers' productivities k_i and k_j are unobserved by the lender. Let L_B, R_B, L_S, L_E be as defined in Propositions 1 and 3. The loan menu which maximizes the borrowers' ex-ante expected payoff consists of two loan contracts with different loan size and interest rate: $\mathcal{M} \equiv (L_M^*, R_M^*)$ and $\mathcal{N} \equiv (L_N^*, R_N^*)$. Contract $\mathcal{M}$ has terms $L_M^* = L_B$ and $R_M^* = R_B$ for any k_H, k_L, p. Contract $\mathcal{N}$ has terms $R_N^* = \frac{L_N^*}{p(2-p)}$ and $L_N^* = L_S$, or $L_N^* = L_E < L_S$, or $L_N^* = L_F < L_S$ depending on the exact values of k_H, k_L, p (see below) where $L_F \equiv \frac{pk_H - 1}{pk_H - \frac{1}{2-p}} pV$. Any borrowers who select contract $\mathcal{N}$ form a standard group and any borrowers who select contract $\mathcal{M}$ form a bogus group.*

Proof: see Appendix

Proposition 4 mirrors the results obtained in the observable productivity setting in Proposition 3 and Corollary 1. As in the previous results, contract $\mathcal{M}$ designed for bogus groups always has the maximum possible loan size, L_B while the loan size L_N designed for standard groups is sometimes reduced to L_E or L_F to prevent heterogeneous groups from selecting this contract, benefitting from its low interest rate, and choosing to operate as bogus. The case when L_N is reduced from L_S to L_E arises for the same reason as in Proposition 3 – to prevent the HL group from choosing contract $\mathcal{N}$ intended for standard groups but operate as bogus. The additional case with loan size L_F where $L_F < L_S$ arises due to the need to deter HL groups from selecting contract $\mathcal{N}$ and operating as bogus. This was not an issue in Proposition 3 and Corollary 1 since the group composition is known to the lender and HL groups are only offered contract (L_B, R_B) in the corresponding parametric case.

Figure 2 illustrates the possible scenarios depending on the model parameters.

[*Figure 2 here*]

In general, for a borrower pair with different types of projects (HL), a bogus group is more likely to arise if the relative heterogeneity of productivity is high; while for a borrower pair with same types of projects (LL or HH), a bogus group is more likely to arise if the absolute level of productivity is high. The basic intuition of the trade-off among three factors: (i) the risk sharing benefit, (ii) the strategic interaction (free riding) cost and (iii) the productivity increment in bogus group, still apply when the each borrower's productivity is unobservable for the lender: for ii group, the strategic interaction cost in the repayment stage may overwhelm the risk sharing benefit in standard group when the absolute level of productivity k_i is sufficiently high so that operating as bogus group under contract $\mathcal{M}$ is preferred. For HL groups, besides the first two trade-offs, when the relative heterogeneity in productivity is sufficiently high, the extra benefit from increased output further strengthens their incentive to operate as bogus group under contract $\mathcal{M}$.

Similar to our findings in the case of observable project types, borrowers' welfare is enhanced when the presence of bogus groups is taken into account and they are offered an appropriate contract. Thus, bogus groups are not a 'problem' costing lenders money but instead represent an opportunity that lenders should take into account and incorporate in designing the optimal loan contract.

4 Excluding bogus groups

Suppose that the lender is aware of the possibility of bogus groups but, for some exogenous reason, would like to design a group loan contract that eliminates the incentive for all borrower pairs to operate as a bogus group.[21] Focus on the case of observed productivities k_i and k_j (see Appendix C for the unobservable productivities case).

The lender wants to design an 'excluding' contract $(L^{\#}, R^{\#})$ so that: (i) the borrower pair ij chooses to form a standard group; (ii) each group member repays in full when her own project succeeds; (iii) the lender breaks even and (iv) the group's joint payoff is maximized.

From the previous analysis we know that contract $(L^{\#}, R^{\#})$ must satisfy the no default constrain for standard groups (condition (3)) and that bogus group are not optimal at $(L^{\#}, R^{\#})$. From Lemma 1, the latter requires that,

$$(k_i - k_j)L^{\#} \leq 2(1-p)(V - R^{\#}) \tag{25}$$

Therefore, the payoff-maximizing excluding contract $(L^{\#}, R^{\#})$ solves:

$$\max_{L,R} W_{ij}(L, R|S) = p(k_i + k_j)L - 2p(2-p)R + 2p(2-p)V$$

$$\text{s.t. (1), (3), (25), and } R = \frac{L}{p(2-p)}$$

Proposition 5: *Suppose the lender wants to exclude bogus groups and borrowers' productivities are observable. Let the quantities L_S, R_S, L_E and R_E are as defined in Propositions 1-3. Then:*

(i) the payoff-maximizing excluding contract for ii groups is (L_S, R_S);

(ii) the payoff-maximizing excluding contract for HL groups is (L_E, R_E) with $L_E < L_S$ if $k_H - k_L > \frac{2}{p(2-p)}$ and (L_S, R_S) if $k_H - k_L \leq \frac{2}{p(2-p)}$.

[21] For example, as discussed in the introduction, the lender may be intolerant to the rule violation in operating as a bogus group; or the interest rate is regulated and the lender cannot break even if bogus groups exist; or the lender wants to ensure that every borrower engages in their own project to improve her long run skills and experience, or other social or poverty alleviation reasons.

Proof: see Appendix.

Proposition 5 shows that the optimal contract for standard groups from Proposition 1 (L_S, R_S) is optimally offered to *ii* groups because under this contract, such groups has no incentive to be bogus group. The same holds for HL groups if the productivity differential is relatively small – the gain in extra output is not worth losing the future value. In the remaining case, $k_H - k_L > \frac{2}{p(2-p)}$ the HL group is deterred to switch to bogus form by reducing the loan size to L_E. The interest rate in contract (L_E, R_E) is the same as in contract (L_S, R_S) since in both cases all groups are standard in equilibrium and hence the lender's break-even condition is the same per dollar lent.

Comparing with the optimal contracts in Section 3.1, we see that by excluding bogus groups borrowers' welfare is reduced (compare to Proposition 3 and Corollary 1). The excluding contract fails to capitalize on the larger loan size L_B if project productivity is high and also reduces instead of increases the loan size if the productivity differential is high. These distortions are even larger if the productivities are unobserved (see Appendix C).

5 Empirical Analysis

5.1 Institutional background

According to the annual report by the People's Bank of China, by the end of the third quarter of 2013, there were 7,398 microfinance institutions in China with total issued loans approximating 161.2 billion RMB.[22] Among all microcredit institutions, CFPAM (the China Foundation for Poor Alleviation Microfinance) is ranked 1st in both total issued loans and active members.[23]

In its group lending program CFPAM employs trained loan officers who introduce the regulations and expected program costs and benefits to potential members. Later, these officers also assume responsibility for training and monitoring borrower groups. CFPAM's rules for group formation include:

1. A group must consist of 2 to 5 self-chosen members.
2. All group members must be from the same village.
3. There shall be no more than one member from the same household in a group. It is also not desirable for close relatives to be in the same group.
4. After each group is formed, it elects its own leader among the members.

CFPAM advertises regularly in the rural areas of China and most people in those areas are aware of its lending program. The clients first form a group and then approach CFPAM to apply for loans. If the group meets certain basic criteria (each member has an existing business or a business plan, understands the rules, and requires a loan), CFPAM provides a training session explaining the concept of joint liability, group operations, the importance of group solidarity and monitoring of loan repayment, use and meetings attendance by all members. Then each group member receives her first loan. The whole process typically takes a week.

[22] See the microfinance annual report by People's Bank of China at www.pbc.gov.cn.

[23] During 2013, CFPAM issued 182,397 loans, with the total amount 1.87 billion RMB serving 174,577 clients (borrowers).

5.2 Data

We obtained proprietary information regarding detailed contract terms of group lending contracts from CFPAM.[24] These contracts are for the borrowers served between February, 2011 and November, 2011 in three of the poorest counties in Liaoning province – Beizhen, Xiuyan and Xingcheng. The contract terms include: contract number, which indicates the sequential order of the group of clients served in 2011, the starting date of the loan, and the proposed usage of the loan. We also obtained demographic information about the borrowers in each group. This information includes age, gender, ethnicity, education, and marital status.

Due to regulation constraints imposed by CFPAM, we randomly choose 80 group lending contracts, with 366 borrowers out of all group lending contracts provided by CFPAM. We called each borrower in a random sequence at randomly selected time slots during working hours for two weeks. We interviewed each borrower by asking the following questions in a random sequence and customized the interview conditional on feedbacks from the borrowers:

1. *The group type (GT) question:* is this group a bogus group (*Lei Da Hu*) or not? We explain what is meant by bogus group if the borrower is not aware.

2. *The joint liability (JL) question:* Are you aware of the joint liability clause in group lending?

3. *The group membership (GM) question:* Do you know the other group member personally?

We then construct three indicator variables as follows. The group type dummy, "BOGUS", equals one if *all* group members answered "yes" when asked the GT question and zero otherwise.[25] Similarly, the "AWARENESS" ("KNOW") dummies equal one if the borrower answered yes when (s)he was asked the JL (GM) question and zero otherwise.

In addition, the data contain six categories of proposed loan usage, five education categories, and three geographical dummies (see Appendix E for details). We provide summary statistics in Table 2. Perhaps the most important number is that 69% of all sampled groups are bogus. This suggests that in CFPAM's group lending practice, bogus groups are the rule rather than the exception. As argued in the theory section of the paper, ignoring the existence of bogus groups can lead to an inferior outcome in terms of efficiency and borrower welfare via causing a loss to the lender and/or sub-optimal contract design (e.g., too small loan size).

5.3 Loan terms

As Table 1 indicates, there is not much difference in the loan repayment amount, the loan size, and interest rate across the 80 group loan contracts in our sample. The monthly repayment amount is determined by the total number of payments to be made and the loan size issued to each borrower. In our sample, only 1 out of 80 contract has a different total number of payments (4); all other contracts have 10 payments in total. Another important contract term is the interest rate. In our sample, only 2 out of all 80 contracts have a different interest rate, 12% and 16% (the latter is the same outlier group with 4 payments); all 78 remaining contracts have the same interest rate of 13.5%. Overall, this evidence suggests that the group loan contract is very standardized and there are only very minor differences in the number of payments and the interest rate across the groups.

The contract term that is observed to differ is the loan size since even within the same group different borrowers may be lent different amounts. For instance, in contract No. 681 issued to a

[24] Until 2014 CFPAM issued loans only on group basis.

[25] This definition thus provides a lower bound on the actual number of bogus groups.

Table 2: Summary Statistics

Variable	Variable Definition	Obs	Mean	Std. Dev.	Min	Max
bogus	group type dummy	366	0.69	0.21	0	1
mpayment	monthly payment (in RMB)	366	828.6	192.3	50.7	908
loansize	loan amount (in RMB)	366	7194	1774	500	8000
duration	number of payments in total	366	9.93	0.62	4	10
ir	interest rate	366	13.5%	0.32%	12%	16%
age	age	366	43.8	9.68	21	64
married	marital status dummy	366	0.94	0.24	0	1
AFAF	industry dummy	366	0.80	0.40	0	1
manufacture	industry dummy	366	0.06	0.23	0	1
service	industry dummy	366	0.02	0.15	0	1
wholesale	industry dummy	366	0.08	0.27	0	1
transport	industry dummy	366	0.02	0.14	0	1
housing	industry dummy	366	0.02	0.15	0	1
below	education dummy	366	0.01	0.10	0	1
primary	education dummy	366	0.27	0.44	0	1
junior	education dummy	366	0.69	0.46	0	1
highschool	education dummy	366	0.03	0.17	0	1
college	education dummy	366	0.01	0.09	0	1
beizhen	county dummy	366	0.54	0.50	0	1
xiuyan	county dummy	366	0.22	0.41	0	1
xingcheng	county dummy	366	0.25	0.43	0	1
Han	the majority of Chinese	366	0.29	0.46	0	1
Manchu	one of the minorities of Chinese	366	0.70	0.46	0	1
Mongols	one of the minorities of Chinese	366	0.01	0.07	0	1

group of four, two group members obtain 500 RMB each, while the other two obtain 1,000 RMB and 8,000 RMB.

The microlender can only determine the contract terms conditional on observable characteristics of the borrowers at the individual or group level. According to the application process regulated by CFPAM, the observable characteristics of the borrowers are age, gender, marital status, education, and proposed usage of the loan. In addition, CFPAM might take into account the ethnicity of a borrower.[26]

In Table 3 we regress loan size (the only term that we observe to differ across contracts) on various borrower characteristics after excluding the outlier group with 4 payments and 16% interest rate. We find that the borrowers' characteristics, including age, marital status, and proposed usage of the loan are not statistically significantly associated with the loan size.

To sum up, the data suggest that contract terms are basically identical for all borrower groups conditional on their observable characteristics. In the eyes of our model, this could indicate one of two possibilities: first, the lender ignores the existence of bogus groups and provides the optimal standard group lending contract to all groups. The second possibility is that the micro-lender does acknowledge the existence of bogus groups but the parameters are such that all groups select or are optimally given the same contract. This can happen in the model when either contract (L_B, R_B) or contract (L_S, R_S) are optimal for all groups (see Corollaries 1 and 2). However, this scenario is contradicted by the data since approximately 30% of all groups are standard (unless they mis-report). We hypothesize that it is more likely that CFPAM ignores or is unaware of the existence of bogus groups, and provides a contract assuming standard group form, which, as this paper shows, may lead to losses or be suboptimal given the borrowers' project returns.

5.4 Determinants of the choice between bogus and standard group (*very preliminary*)

Next, we examine empirically the group form choice (bogus vs. standard) made by the borrowers in our sample. The evidence on contract uniformity and lack of conditioning on borrower characteristics presented above suggests that it is unlikely that an endogeneity problem exists from the lender optimizing the contract terms in expectation of the equilibrium choices by the borrowers and simultaneously the borrowers choosing a group form given the offered contract.[27]

In the model one of motivation for a group to choose to be bogus is the larger loan size, that is, a larger loan size is positively associated with a group's choice to be bogus. We investigate the determinants of group form choice (bogus = 1 vs. standard = 0) using a logit regression. Specifically, we regress the "bogus" dummy on loan characteristics such as the loan size and the loan monthly payment (which is a function of the interest rate, the total number of payments, and the loan size) and the characteristics of the borrowers.[28] The results are reported in table 4.

As demonstrated in Table 4, the loan size is significantly related to the group form choice – for larger loan size (for any reason), the group is more likely to be bogus. The magnitude remains stable around 5.2 in all the regression specification, which indicates that the relationship between loan size and bogus group choice is robust to controlling for any other variable we consider. The

[26] In general, borrowers of the minorities can be treated more favorably or unfavorably.

[27] It could be the case that, the micro-lender knows whether a person is a frequent borrower or a new borrower and offer different contracts, e.g., due to reduced information asymmetry. Due to data limitations, we can not identify whether a borrower is frequent or new. However, ignoring such factors would not bias our results as long as being frequent or new borrowers is uncorrelated with the group form choice or it only affects the group form choice via the same channel.

[28] We cannot use the interest rate and the number of total payments directly in the regression since the variation in these two variables is too small which results in a high correlation with the constant term.

Table 3: Determinants of loan size

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	lloansize	lloansize	lloansize	lloansize	lloansize	lloansize	lloansize	lloansize
AFAF		-0.02	-0.04					-0.10
		(0.15)	(0.15)					(0.15)
manufacture		0.06	0.05					0.01
		(0.18)	(0.18)					(0.18)
service		0.18	0.16					0.09
		(0.21)	(0.21)					(0.21)
wholesale		0.17	0.14					0.10
		(0.17)	(0.17)					(0.17)
transportation		0.18	0.17					0.16
		(0.22)	(0.22)					(0.22)
lage	0.04		0.08		0.10		0.03	0.14
	(0.09)		(0.10)		(0.09)		(0.09)	(0.10)
married	-0.14		-0.12		-0.12		-0.13	-0.09
	(0.09)		(0.10)		(0.09)		(0.09)	(0.09)
below				-0.64**	-0.65**			-0.50
				(0.32)	(0.32)			(0.33)
primary				-0.17	-0.18			-0.09
				(0.25)	(0.25)			(0.25)
junior				-0.03	-0.04			0.06
				(0.25)	(0.25)			(0.25)
highschool				0.06	0.05			0.12
				(0.28)	(0.27)			(0.28)
manchu						-0.08	-0.07	-0.07
						(0.05)	(0.05)	(0.05)
mongols						0.11	0.13	0.20
						(0.31)	(0.31)	(0.30)
Constant	8.82***	8.81***	8.64***	8.89***	8.63***	8.88***	8.89***	8.46***
	(0.36)	(0.15)	(0.39)	(0.24)	(0.44)	(0.04)	(0.37)	(0.46)
Observations	366	366	366	366	366	366	366	366
R-squared	0.01	0.02	0.03	0.04	0.05	0.01	0.01	0.08

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Determinants of bogus vs. standard group form

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	bogus	bogus	bogus	bogus	bogus	bogus	bogus	bogus
lmpayment	-4.78**				-4.75**	-4.85**	-4.44**	-4.60**
	(2.14)				(2.10)	(2.12)	(2.00)	(2.06)
lloansize	5.41**				5.38**	5.53**	5.12**	5.28**
	(2.18)				(2.13)	(2.16)	(2.04)	(2.10)
lage				-0.47	-0.53	-0.60	-0.60	-0.66
				(0.48)	(0.50)	(0.52)	(0.54)	(0.54)
married				-0.22	-0.077	0.04	0.08	0.15
				(0.49)	(0.50)	(0.51)	(0.51)	(0.51)
AFAF		-1.16				-0.98	-0.97	-0.99
		(1.08)				(1.08)	(1.09)	(1.09)
manufacture		-1.54				-1.56	-1.41	-1.39
		(1.16)				(1.17)	(1.18)	(1.18)
service		0.00				-0.10	-0.09	-0.13
		(1.51)				(1.52)	(1.52)	(1.52)
wholesale		-0.76				-0.82	-0.65	-0.62
		(1.15)				(1.16)	(1.17)	(1.17)
transportation		-3.74**				-3.82**	-3.82**	-3.78**
		(1.52)				(1.52)	(1.52)	(1.52)
below			13.80				15.16	15.89
			(574.1)				(716.9)	(894.2)
primary			14.47				15.00	15.68
			(574.1)				(716.9)	(894.2)
junior			14.68				14.97	15.62
			(574.1)				(716.9)	(894.2)
highschool			14.78				14.98	15.58
			(574.1)				(716.9)	(894.2)
manchu								-0.26
								(0.28)
mongols								-1.30
								(1.45)
Constant	-15.10***	1.95*	-13.80	2.76	-12.97**	-12.42**	-26.64	-27.25
	(5.41)	(1.07)	(574.1)	(1.87)	(5.68)	(5.82)	(716.9)	(894.3)
Observations	366	366	366	366	366	366	366	366

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

controls such as age, marital status, education, and the proposed usage of the project are not statistically significant, with the sole exception of the transport dummy. To gauge the economic significance of the coefficient on loan size note that for a one standard deviation from the mean (when the loan size is adjusted from 7,194 to 8,968) the probability of observing a bogus group increases from 65.4% to 94.7%, ceteris paribus.

Our model also suggests that groups with more productive investment projects and/or with higher probability of success (safe projects) are more likely to be bogus. [to be completed]

6 Conclusion

In this paper, we study optimal group-lending contracts by considering the possibility of bogus group formation in a limited enforcement setting. There are three key mechanisms that influence the contract design and group form in our model. The first is the risk-sharing benefit in a standard group: the probability that the group repays the total amount due and obtain the continuation value of future credit is higher than in a bogus group. The second mechanism is the strategic interaction cost in a standard group: each member makes the repayment decision independently and has the opportunity to strategically default while free riding on her partner's repayment. As a result the maximum loan size for a standard group must be reduced so that the members do not default strategically. In contrast, only one investment project is implemented in a bogus group and hence there is no strategic interaction between borrowers, thus the loan size of bogus loan contract is always larger.

The third mechanism is the possibility of higher expected output that exists in a bogus group with borrowers of heterogeneous productivities – the reason is that the bogus group always invests the entire loaned amount into the member project with the highest productivity.

The trade-offs among these three factors underpin our main theoretical findings. The main take-away is that bogus groups arise optimally when the project productivities are sufficiently high (to benefit from the larger feasible loan size) or when the productivity differential between projects within the group is sufficiently large (to benefit from both the larger loan size and extra output). These benefits are outweighed by the risk-sharing advantage of standard groups for less productive projects and/or homogeneous groups.

An important lesson from our theory is that bogus groups are not an undesirable phenomenon that microlenders should try to eradicate but an optimal response by borrowers that can increases efficiency and the allocation of credit if lenders react adequately by designing an optimal contract (menu). Bogus groups can only cause losses to lenders that are unaware or purposefully ignore their existence.

A key assumption in our model is the information or enforcement friction (both within the groups and between the borrowers and the lender) causing the possibility of strategic default decided at the individual level. This assumption is at the heart of our results that standard groups obtain smaller loans at the optimum. In Appendix D we study how changing this assumption to a collective default decision within the group affects the basic trade-off by eliminating the second of the three main mechanisms listed above. The results simplify since in such a setting only heterogeneous groups have an incentive to be bogus.

Another issue we did not discuss due to the timing we assume is the possibility of endogenous matching between borrowers of different productivities. One could consider changing the timing so that the productivities are drawn first and then agents sort into groups expecting the optimal loan contract(s). In that case, it is easy to see that if only standard groups can be formed, any equilibrium matching pattern is optimal since two HL groups have the same joint surplus as one HH and one LL group. In contrast, if endogenous bogus group formation is allowed and the borrower productivities k_i, k_j are observable, then the lender would simply offer the optimal contract specified in Proposition 3 and Corollary 1 to whatever groups he is facing and we simply need to check whether PAM or NAM pattern is optimal. It turns out that NAM would be optimal if the differential $k_H - k_L$ is large and PAM is optimal otherwise. Allowing endogenous matching is complex if the lender does not observe the borrower productivities. The reason is that we cannot just take the contract menu from Proposition 4 and Corollary 2 as given since the lender may want to offer a different menu if he knew, for example, that all groups that he would face will be HL (the IC constraint will be affected). On the other hand, the equilibrium group composition would depend on the contract menu that will be offered. This is a hard problem, potentially with multiple equilibria that remains for future research.

A Appendix A – Proofs

A.1 Proof of Proposition 1

(a) Since the objective function (6) is decreasing in R, the lender's break-even constraint must be binding (zero profits), i.e. $R = \frac{L}{p(2-p)}$. Therefore, the problem can be re-written as

$$\max_{L} \ (p(k_i + k_j) - 2)L + 2p(2-p)V, \tag{26}$$

$$\text{s.t. } L \leq p(1-p)V. \tag{27}$$

Since the objective function is strictly increasing in L, constraint (27) must bind at the optimum. Thus the contract $\mathcal{S} = \{L_S, R_S\}$ described in Proposition 1 is the optimal contract.

(b) Notice that contract $\mathcal{S}$ is independent of k_i and k_j. Therefore, even if k_i and k_j are unobserved by the lender, Proposition 1(a) still holds. To show this formally, we can re-write the objective as the weighted sum (with weights equal to the population shares of HH, LL and HL groups) of ij's total expected payoffs. Since the no-default and break-even constraint do not depend on k_i, k_j the results from part (a) remain intact.

A.2 Proof of Lemma 1

An ij borrower pair will form a bogus group instead of a standard group for given (L, R) if and only if $W_{ij}(L, R|B) \geq W_{ij}(L, R|S)$ which is equivalent to condition (10) using (SEP) and (8).

A.3 Proof of Proposition 2

Given contract (L_S, R_S), condition (10) is equivalent to $L_S > \frac{p(2-p)V}{1+\frac{p(2-p)}{2(1-p)}(k_i - k_j)}$, or

$$k_i > k_j + \tfrac{2}{p(2-p)}.$$

If $i = j$, the above is impossible to hold. Therefore, given the standard group loan contract $\{L_S, R_S\}$, bogus group will arise only if $i = H$ and $j = L$ and (12) holds. The lender's profit from such group is:

$$2pR_S - 2L_S = 2\left(p\tfrac{L_S}{p(2-p)} - L_S\right) = -\tfrac{2(1-p)}{2-p}L_S < 0,$$

so bogus HL groups cause loss to the lender.

A.4 Proof of Proposition 3

Remark 1: Comparing the payoffs $W(L, R|S)$ and $W(L, R|B)$ from (13) and (14), Lemma 1 shows that, for a given contract (L, R) with $R \leq \frac{(1-p)V}{2-p}$, an ij group would optimally choose to operate as bogus group if and only if,

$$p(k_i - k_j)L > 2p(1-p)(V - R) \tag{BC}$$

and operate as standard group otherwise.

Remark 2: It is easy to see from the expressions for $W(L, R|S)$ and $W(L, R|B)$ that forming a bogus group is always optimal for any group ij facing contract (L, R) with $R > \frac{(1-p)V}{2-p}$.

(a) Suppose the lender faces a homogeneous, ii group and wants to induce a standard group (choose $\tau = 1$). The zero profit conditon implies $R = \frac{L}{p(2-p)}$, which substituted into the no default constraint gives $L \leq p(1-p)V$. Constraint (IC) is satisfied for any R satisfying the no default constraint $R \leq \frac{(1-p)V}{2-p}$ since then inequality (BC) does not hold for $k_i = k_j$ (see Remark 1). The objective $W(L, \frac{L}{p(2-p)}|S)$ equals $(2pk_i - 2)L + 2p(2-p)V$ which is strictly increasing in L. Consequently, the no default constraint, $L \leq p(1-p)V$ must bind at the optimum, implying $L^* = L_S = p(1-p)V$ and $R^* = R_S = \frac{(1-p)V}{2-p}$.

Suppose now the lender wants to induce a bogus group (choose $\tau = 0$). The zero profit constraint implies $L = pR$, which substituted into the no default constraint gives $L \leq \frac{pV}{2}$. The objective $W(L, \frac{L}{p}|B)$ equals $(2pk_i - 2)L + 2pV$ which is strictly increasing in L implying that, as long as (IC) is not violated, it is optimal to have the no default constraint bind and so $L^* = \frac{pV}{2} \equiv L_B$ and $R^* = \frac{V}{2} \equiv R_B$. Constraint (IC) is indeed not violated at (L_B, R_B) since forming a bogus group is optimal for $R_B = \frac{V}{2} > \frac{(1-p)V}{2-p}$ (see Remark 2 above).

Finally, to decide whether choosing $\tau = 0$ or $\tau = 1$ is optimal, the lender must compare the group payoff from inducing a standard group at contract (L_S, R_S) vs. bogus group at (L_B, R_B). Setting $\tau^* = 1$, $L^* = L_S$ and $R^* = R_S$ is optimal when $W(L_S, R_S|S) \geq W(L_B, R_B|B)$ which is equivalent to:

$$(2pk_i - 2)L_S + 2(2-p)pV \geq (2pk_i - 2)L_B + 2pV \text{ or,}$$
$$pk_i(1-p)V \geq (pk_i - 1)\frac{V}{2} \text{ or}$$
$$p(2p-1)k_i \leq 1$$

and setting $\tau^* = 0$, $L^* = L_B$ and $R^* = R_B$ is optimal otherwise.

(b) Now suppose the lender faces a heterogeneous group (k_L, k_H) and considers inducing standard group (choose $\tau = 1$). As before, the zero profit constraint implies $R = \frac{L}{p(2-p)}$, which substituted into the no default constraint implies $L \leq p(1-p)V = L_S$. However, unlike in the homogeneous group case, now constraint (IC) is not automatically satisfied for any R satisfying the no default condition. Instead, in order for (IC) to hold, that is $W(L, R|S) \geq W(L, R|B)$, using (BC), we need:

$$p(k_H - k_L)L \leq 2p(1-p)(V - \tfrac{L}{p(2-p)}) \text{ or,}$$
$$L \leq \frac{p(1-p)V}{\frac{1-p}{2-p} + \frac{p}{2}(k_H - k_L)} = L_E$$

Therefore, the optimal contract in this case is $L_N^* \equiv \min\{L_E, L_S\}$ and $R_N^* \equiv \frac{L_N^*}{p(2-p)}$.

Suppose now the lender wants to induce a bogus group (choose $\tau^* = 0$). The analysis is analogous to that for homogeneous groups. Constraint (IC) is still satisfied at $L^* = \frac{pV}{2}$ and $R^* = \frac{V}{2}$. Therefore, the optimal contract in this case is $L^* = L_B$ and $R^* = R_B$, as before.

To decide whether choosing $\tau^* = 0$ or $\tau^* = 1$ is optimal, compare the group payoff from operating as standard group with contract (L_N, R_N) vs. operating as bogus group with contract (L_B, R_B). There are two cases depending on whether L_E or L_S is smaller. Note that $L_E \geq L_S$ is equivalent to

$$\frac{p(1-p)V}{\frac{1-p}{2-p}+\frac{p}{2}(k_H-k_L)} \geq p(1-p)V \text{ or,}$$

$$\frac{2}{p(2-p)} \geq k_H - k_L \qquad \text{(CC1)}$$

which is the converse of the condition in Proposition 2 about when bogus groups arise at contract (L_S, R_S).

Suppose first (CC1) holds, that is

$$k_H \leq d(k_L) \equiv k_L + \frac{2}{p(2-p)}$$

and hence $L_N = L_S$. In such case, setting $\tau^* = 1$, $L^* = L_S$ and $R^* = R_S$ is optimal if $W(L_S, R_S|S) \geq W(L_B, R_B|B)$ which is equivalent to:

$$(p(k_L + k_H) - 2)L_S + 2(2-p)pV \geq (2pk_H - 2)L_B + 2pV \text{ or,} \qquad \text{(CC2)}$$

$$k_H \leq g(k_L) \equiv \frac{1}{p^2} + \frac{1-p}{p}k_L$$

and setting $\tau^* = 0$, $L^* = L_B$ and $R^* = R_B$ is optimal otherwise.

If (CC1) does not hold, that is $k_H > d(k_L)$, then $L_N = L_E < L_S$ and $R_N = R_E \equiv \frac{L_E}{p(2-p)}$. Choosing $\tau^* = 1$, $L^* = L_N$ and $R^* = R_N$ is optimal if $W(L_E, R_E|S) \geq W(L_B, R_B|B)$ which is equivalent to

$$(p(k_L + k_H) - 2)\frac{p(1-p)V}{\frac{1-p}{2-p}+\frac{p}{2}(k_H-k_L)} + 2(2-p)pV \geq (2pk_H - 2)\frac{pV}{2} + 2pV \text{ or } f(k_L) \geq k_H \qquad \text{(CC3)}$$

where[29]

$$f(k) \equiv \frac{1}{2}\left(k + c + \sqrt{(k+c)^2 - \frac{8(1-p)}{p^2(2-p)} - \frac{4k}{p}}\right) \text{ and } c \equiv \frac{4p^2-11p+8}{p(2-p)}.$$

Setting $\tau^* = 0$, $L^* = L_B$ and $R^* = R_B$ is optimal otherwise, for $f(k_L) < k_H$.

Remark 3. Note that, since $L_E < L_S$, then $k_H > d(k_L)$ implies $W(L_S, R_S|S) > W(L_E, R_E|S)$, which means that condition (CC3) is tighter than (implies) condition (CC2).

In sum, we obtain:

(a) if $k_H \leq d(k_L)$ then it is optimal to offer $\mathcal{S}$ if $k_H \leq g(k_L)$ and offer $\mathcal{B}$ if $k_H > g(k_L)$ and

(b) if $k_H > d(k_L)$ then it is optimal to offer $\mathcal{E}$ if $k_H \leq f(k_L)$ and offer $\mathcal{B}$ if $k_H > f(k_L)$.

Corrolary 1 shows the exact parameter configurations for which these cases occur.

A.5 Proof of Corollary 1

The results for homogeneous groups (the columns labeled LL and HH) depend only on whether $p(2p-1)k_i > 1$ or $p(2p-1)k_i \leq 1$ and are implied directly by Proposition 3(a). Hence, we only discuss heterogeneous groups below. Parts (a)–(c) in the proof below refer to the corresponding

[29] One can show that the smaller root of the quadratic equation of which $f(k)$ is the larger root is strictly smaller than k. Thus, for $k_H \geq k_L$ inequality (CC3) is satisfied iff $k_H \leq f(k_L)$.

cases of the table in the corollary statement. Call $\hat{k} = \frac{2-3p}{p(2-p)(2p-1)}$ and $\tilde{k} = \frac{1}{p(2p-1)}$ defined for $p > 1/2$. The following auxiliary results can be easily shown.

Result 1. Suppose $p \leq 1/2$. Then, $(2p-1)\hat{k} \leq 0$, $d(k) < g(k)$ and $d(k) < f(k)$ for any $k \geq 0$.

Proof: The first two statements are easy to check directly. To prove that $d(k) < f(k)$, rewrite $k + c$ in the definition of $f(k)$ as $d(k) + n$ where $n \equiv \frac{3}{p} - 4$. Then, $d(k) < f(k)$ is equivalent to:

$$d(k) < \frac{1}{2}\left(d(k) + n + \sqrt{(d(k)+n)^2 - \frac{4k}{p} - \frac{8(1-p)}{p^2(2-p)}}\right)$$

or $d(k) - n < \sqrt{(d(k)+n)^2 - \frac{4k}{p} - \frac{8(1-p)}{p^2(2-p)}}$ which is equivalent to $\frac{k}{p} + \frac{2(1-p)}{p^2(2-p)} < d(k)n = (k_{\min} + k)(\frac{3}{p} - 4)$, or, $2k(2p-1) < k_{\min}(2-3p)$ which is true for $p \leq 1/2$.

Result 2. Suppose $p > 1/2$. Then $d(\hat{k}) = g(\hat{k}) = f(\hat{k})$.

Proof: $d(\hat{k}) = g(\hat{k})$ is verified directly. To show that $d(\hat{k}) = f(\hat{k})$ follow the proof of Result 1 above and use $k_{\min} = \frac{2}{p(2-p)}$ at the last step.

Result 3. Suppose $p > 1/2$. Then $\hat{k} \leq k_{\min}$ if $p \geq 4/7$ and $\hat{k} > k_{\min}$ if $p \in (1/2, 4/7)$.

Result 4. Suppose $p > 1/2$. Then, if $k \geq \hat{k}$ we have $d(k) \geq f(k)$ and $d(k) \geq g(k)$ while if $k < \hat{k}$ we have $d(k) < f(k)$ and $d(k) < g(k)$.

Result 5. Suppose $p > 1/2$. Then, $k > \tilde{k} \Longleftrightarrow k > g(k)$.

Result 6. Suppose $p > 1/2$. Then, $\tilde{k} > \hat{k}$ and $\tilde{k} < k_{\min} \Leftrightarrow p > 4/5$.

(a) The condition $p(2p-1)k_L > 1$ (equivalent to $k_L > \tilde{k}$) can only hold if $p > 1/2$. Also, note that if $p > 4/5$ then $\tilde{k} < k_{\min}$ by Result 6 and hence $p(2p-1)k_L > 1$ is satisfied for any $k_L \geq k_{\min}$. Using Result 5, $k_L > \tilde{k}$ implies $k_H \geq k_L > g(k_L)$, and hence it is optimal to offer contract $\mathcal{B}$ for any such k_H, k_L.

(b) As explained in part (a), this case is impossible for $p > 4/5$. It is also impossible for $p \leq 1/2$ because $p(2p-1)k_H > 1$ cannot hold. Supposing $p \in (\frac{1}{2}, \frac{4}{5}]$, the inequality $k_H > \tilde{k}$ implies (Result 5) $k_H > g(k_H) \geq g(k_L)$, that is, $k_H > g(k_L)$ for any such k_H, k_L. Note also that, in the case $k_H > d(k_L)$, the inequality $k_H > g(k_L)$ (the negation of inequality CC2) implies $k_H > f(k_L)$ which is the negation of the stricter inequality CC3. Using Proposition 3(b), in either case it is then optimal to offer contract $\mathcal{B}$ to HL groups for any such k_H, k_L.

(c) The following sub-cases depending on the value of p are possible:

(ci) Suppose $p \leq 1/2$, in which case $p(2p-1)k_H \leq 1$ is satisfied for any $k_H \geq k_{\min}$. Then, from Result 1 and since $d(k) > k$ for all $k > 0$, by Proposition 3(b), it is optimal to offer contract $\mathcal{S}$ if $k_H \in [k_L, d(k_L)]$, $\mathcal{E}$ if $k_H \in (d(k_L), f(k_L)]$, and $\mathcal{B}$ if $k_H > f(k_L)$. These correspond to lines (i), (ii) and (iii-2) in the Table.

(cii) Suppose $p \in (\frac{1}{2}, \frac{4}{7})$. Then, by Result 3, $\hat{k} > k_{\min}$. If $k_L \in [k_{\min}, \hat{k})$, then from Result 4, $d(k_L) < f(k_L)$ and $d(k_L) < g(k_L)$ and so, using Proposition 3(b), it is optimal to offer contract $\mathcal{S}$ if $k_H \in [k_L, d(k_L)]$, $\mathcal{E}$ if $k_H \in (d(k_L), f(k_L)]$, and contract $\mathcal{B}$ if $k_H > f(k_L)$. This maps to lines (i), (ii) and (iii-2) in the Table.

Alternatively, if $k_L \geq \hat{k}$, then $d(k_L) \geq f(k_L)$ and $d(k_L) \geq g(k_L)$ by Result 4, and so it is optimal to offer $\mathcal{S}$ for $k_H \in [k_L, g(k_L)]$ and offer $\mathcal{B}$ for $k_H > g(k_L)$. These correspond to lines (i) and (iii-1) in the Table.

(ciii) Suppose $p \in [\frac{4}{7}, \frac{4}{5}]$. By Result 3, $\hat{k} \leq k_{\min}$ and so $k_L \geq \hat{k}$. Thus, by Result 4, $d(k_L) \geq f(k_L)$ and $d(k_L) \geq g(k_L)$. If $k_L > \tilde{k}$, then by Result 5, $k_L > g(k_L)$ and so it is always true that $k_H > g(k_L)$. Hence, it is always optimal to offer contract $\mathcal{B}$ (line iii-1 in Table). If instead $k_L \in [k_{\min}, \tilde{k}]$ then, by Result 5, $g(k_L) \geq k_L$ and hence it is optimal to offer contract $\mathcal{S}$ for $k_H \in [k_L, g(k_L)]$ and $\mathcal{B}$ for $k_H > g(k_L)$, corresponding to lines (i) and (iii-1) in the Table.

Finally, remember that if $p > 4/5$ only case (a) is possible since $p(2p-1)k_L > 1$ for any $k_L \geq k_{\min}$.$\square$

A.6 Proof of Lemma 2

Before proving Lemma 2, we show the following result.

> **Lemma A1**: *Suppose $L_M < \frac{pV}{2}$. Then, respecting constraints (19), (20) and (IC), we can increase L_M holding L_N constant so that constraint (IC) still holds.*

Consider a local increase of L_M to $L_M + \varepsilon$ with $\varepsilon > 0$. Given constraints (19), (20) and expressions (23) and (24) we have that $W\left(L_M, \frac{L_M}{p}|S\right) \geq W\left(L_M, \frac{L_M}{p}|B\right)$ if and only if $L_M \leq L_M^* \equiv \min\left\{\frac{2p(1-p)V}{p(k_i-k_j)+2(1-p)}, \frac{p(1-p)V}{2-p}\right\}$.[30] When $L_M < L_M^*$, we have

$$\frac{dW\left(L_M, \frac{L_M}{p}|B\right)}{dL_M} = 2(pk_i - 1) > p(k_i + k_j) - 2(2-p) = \frac{dW\left(L_M, \frac{L_M}{p}|S\right)}{dL_M}$$

Therefore, by increasing L_M by a small amount (so that $L_M + \varepsilon < L_M^*$) while holding L_N constant, the LHS of constraint (IC) increases by (weakly) more than its RHS, so the constraint still holds.

Suppose now $L_M \in [L_M^*, \frac{pV}{2})$. If the RHS of (IC) equals $W\left(L_M, \frac{L_M}{p}|S\right)$ the result follows from above. Suppose the RHS of (IC) equals $W\left(L_N, \frac{L_N}{p(2-p)}|B\right)$ instead. Then, constraint (IC) is equivalent to

$$\max\left\{W\left(L_N, \frac{L_N}{p(2-p)}|S\right), W\left(L_M, \frac{L_M}{p}|B\right)\right\} \geq W\left(L_N, \frac{L_N}{p(2-p)}|B\right) \tag{28}$$

The expression $W\left(L_M, \frac{L_M}{p}|B\right)$ is increasing in L_M. Hence, by increasing L_M while holding L_N constant, the LHS of (28) weakly increases while its RHS remains constant, that is, both the condition $L_M \geq L_M^*$ and constraint (28) still hold, which means that (IC) still holds.$\square$

Lemma A1 implies that if $L_M < \frac{pV}{2}$ we can increase L_M while holding L_N constant, so that the constraint (IC) still holds. However, since $W\left(L_M, \frac{L_M}{p}|B\right)$ is strictly increasing in L_M, as a result the objective function weakly increases in L_M, therefore setting $L_M = \frac{pV}{2} = L_B$ is optimal and so the optimal contract $\mathcal{M}$ intended for bogus groups equals (L_B, R_B) for any parameter values.

A.7 Proof of Proposition 4

We start by showing some preliminary results that allow us to simplify the IC constraint, (IC) in Problem (UPP').

A. Preliminaries

First, using Lemma 2 from where $L_M = pV/2$, the expressions (23) and (24) imply

$$W_{ij}(\mathcal{M}|B) = W_{ij}(L_M, \tfrac{L_M}{p}|B) = (pk_i + 1)pV \geq$$
$$\geq (\tfrac{p(k_i+k_j)}{2} + 1)pV = W_{ij}(L_M, \tfrac{L_M}{p}|S) = W_{ij}(\mathcal{M}|S)$$

[30] Note that $W\left(L_M, \frac{L_M}{p}|S\right) \geq W\left(L_M, \frac{L_M}{p}|B\right)$ is only possible when $L_M \leq \frac{p(1-p)}{2-p}$ in (24); otherwise forming a bogus group always (weakly) dominates for the same loan size. Also, $\frac{p(1-p)V}{2-p} < \frac{pV}{2}$ so the no default constraint is not violated.

Also,

$$W_{ij}(\mathcal{N}|S) \equiv W_{ij}\left(L_N, \tfrac{L_N}{p(2-p)}|S\right) = (p(k_i+k_j)-2)L_N + 2p(2-p)V$$

$$W_{ij}(\mathcal{N}|B) \equiv W_{ij}\left(L_N, \tfrac{L_N}{p(2-p)}|B\right) = 2(pk_i - \tfrac{1}{2-p})L_N + 2pV$$

From Lemma 1,

$$W_{ij}(\mathcal{N}|S) \geq W_{ij}(\mathcal{N}|B) \Leftrightarrow L_N \leq L_1(ij) \equiv \frac{p(1-p)V}{\frac{1-p}{2-p}+\frac{p}{2}(k_i-k_j)}, \tag{29}$$

and it is easy to show that,

$$W_{ij}(\mathcal{M}|B) \geq W_{ij}(\mathcal{N}|B) \Leftrightarrow L_N \leq L_2(i) \equiv \frac{pk_i-1}{2pk_i-\frac{2}{2-p}}pV \tag{30}$$

where $k_i \geq k_j$ as assumed earlier.

Any given group ij chooses the contract that yields higher joint payoff, so it will choose contract $\mathcal{M}$ if

$$W_{ij}(\mathcal{N}|S) < W_{ij}(\mathcal{M}|B) \Leftrightarrow L_N < L_3(ij) \equiv \frac{pk_i-3+2p}{p(k_i+k_j)-2}pV \tag{CHM}$$

and contract $\mathcal{N}$ otherwise.

If $W_{ij}(\mathcal{M}|B) > W_{ij}(\mathcal{N}|S)$ for a given group ij (that is, $L_N < L_3(ij)$) then the LHS of the IC constraint (IC) is $W_{ij}(\mathcal{M}|B)$. We showed above that $W_{ij}(\mathcal{M}|B) \geq W_{ij}(\mathcal{M}|S)$ always holds. Hence, to satisfy the IC constraint we need to ensure that $W_{ij}(\mathcal{M}|B) \geq W_{ij}(\mathcal{N}|B)$ or, $L_N \leq L_2(i)$.

Alternatively, if $W_{ij}(\mathcal{N}|S) \geq W_{ij}(\mathcal{M}|B)$ for a given ij (that is $L_N \geq L_3(ij)$), then the LHS of the IC constraint (IC) is $W_{ij}(\mathcal{N}|S)$. Since $W_{ij}(\mathcal{M}|B) \geq W_{ij}(\mathcal{M}|S)$ we then know that $W_{ij}(\mathcal{N}|S) \geq W_{ij}(\mathcal{M}|S)$ holds in this case. Thus, to satisfy the IC constraint we need to ensure that $W_{ij}(\mathcal{N}|S) \geq W_{ij}(\mathcal{N}|B)$, or equivalently $L_N \leq L_1(ij)$. Intuitively, in both cases we need to ensure that the group has no incentive to take the low-interest loan contract $\mathcal{N}$ but operate as bogus.

B. Homogeneous groups

Result 1: *Constraint (IC) is satisfied for homogeneous ii groups at* $L_M = pV/2$ *and any* $L_N \leq p(1-p)V$.

Proof: For an ii group we have $L_1(ii) = p(2-p)V > p(1-p)V$. The no-default constraint (20) then ensures $L_N < L_1(ii)$ and hence (IC) is always satisfied when contract $\mathcal{N}$ is payoff-maximizing for the ii group. It is also a matter of simple algebra to verify that $L_3(ii) = \frac{pk_i-3+2p}{2pk_i-2}pV < L_2(i)$ and hence, whenever contract $\mathcal{M}$ is payoff-maximizing, the (IC) constraint for homogeneous groups is also automatically satisfied for any $L_N \leq p(1-p)V$.

Overall, we conclude that homogeneous ii groups choose contract $\mathcal{M}$ whenever $L_N < L_3(ii)$ or

$$L_N < \frac{pk_i-3+2p}{pk_i-1}\frac{pV}{2} \tag{31}$$

and choose contract $\mathcal{N}$ otherwise. The value of L_N solving problem (UPP') is determined in part C below.

C. Heterogeneous groups

For ij groups, we know from the proof of Proposition 3 that $L_E = L_1(HL) = \frac{p(1-p)V}{\frac{1-p}{2-p}+\frac{p}{2}(k_H-k_L)}$. Similarly, call $L_F \equiv L_2(H) = \frac{pk_H-1}{2pk_H-\frac{2}{2-p}}pV$. There are two cases depending on whether contract $\mathcal{N}$ or $\mathcal{M}$ would be optimally chosen by an ij group. If contract $\mathcal{M}$ is payoff-maximizing (that is, $W_{ij}(\mathcal{M}|B) > W_{ij}(\mathcal{N}|S)$), which happens when $L_N < L_3(HL)$, then the lender needs to also ensure that $L_N \leq L_F$ to satisfy (IC) and prevent bogus groups from deviating to $\mathcal{N}$ causing a loss. If, alternatively, contract $\mathcal{N}$ is optimally chosen by the ij group, which happens when $L_N \geq L_3(HL)$, then the lender must ensure $L_N \leq L_E$ to satisfy (IC). In both cases we also need

$L_N \leq p(1-p)V = L_S$ from the no-default constraint. Notice that the objective function is weakly increasing in L_N, so the optimal loan size will be the largest possible value satisfying the (IC) and no-default constraints. Clearly, setting $L_N^* = L_S$ (the maximum loan size for a standard group compatible with no default) is optimal whenever (IC) is satisfied. The only reason that the loan size L_N^* is reduced to L_E or L_F is when constraint (IC) is violated at $L_N = L_S$ for ij groups.Overall, these results imply

Table M	
loan size, $L_N^* =$	**condition**
1. L_S	$L_S < L_3(HL)$ and $L_S \leq L_F$
2. L_F	$L_F < L_3(HL)$ and $L_F < L_S$
3. L_S	$L_S \geq L_3(HL)$ and $L_S \leq L_E$
4. L_E	$L_E \geq L_3(HL)$ and $L_E < L_S$

The exact conditions for each case to occur depend on the parameters k_L, k_H and p (see Figure 2 for an illustration). Since the lender cannot observe the group composition, the reduced loan size (L_E or L_F) applies to *all groups* (HH, LL and HL) even though for homogeneous groups the (IC) constraint is slack, as shown in part B. above.

A.8 Proof of Proposition 5

Condition (25) for ii groups is always true, and thus redundant. Thus, the optimal excluding contract for homogeneous groups is therefore (L_S, R_S) – the same as in Proposition 1.

For HL groups, using the lender's break-even constraint, $p(2-p)R = L$, condition (25) becomes

$$(k_H - k_L)L + \tfrac{2(1-p)L}{p(2-p)} \leq 2(1-p)V \text{ or}$$
$$L \leq L_E$$

Thus, if $L_E < L_S$ which is equivalent to $k_H - k_L > \frac{2}{p(2-p)}$, since the lender's objective function is increasing in L, the optimal excluding contract for HL groups has $L^{\#} = L_E$ and $R^{\#} = \frac{L_E}{p(2-p)}$ while if the productivity differential is small, $k_H - k_L \leq \frac{2}{p(2-p)}$ (or $L_S \leq L_E$) we have $(L^{\#}, R^{\#}) = (L_S, R_S)$ since constraint (25) is slack.

B Appendix B – (Repay, Default) equilibrium

In Section 2.2 we showed that a (Repay, Default) Nash equilibrium would arise in a standard borrower group if $\frac{1-p}{2-p}V < R \leq \frac{V}{2}$. Here we show how Proposition 1 is modified if such equilibrium is allowed.

Start by noticing that the lender's break-even condition in a (Repay, Default) equilibrium is $pR = L$. The no default constraint then implies $L \leq \frac{pV}{2}$. The lender would thus offer either the optimal contract inducing the (Repay, Repay) – contract (L_S, R_S), or the payoff-maximizing contract $(L_H, R_H) = (\frac{pV}{2}, \frac{V}{2})$ inducing the (Repay, Default) equilibrium, whichever of the two maximizes the group's total payoff for the given parameters.

Proposition 1B: *Suppose Assumption 1 holds. The optimal standard group loan contract is* (L_S, R_S) *if* $p(p - \frac{1}{2})(k_i + k_j) \leq 1$ *and* $(L_H, R_H) = (\frac{pV}{2}, \frac{V}{2})$ *otherwise.*

For example, Proposition 1B implies that $p \leq 1/2$ is a sufficient condition for (L_S, R_S) to be optimal when the (Repay, Default) equilibrium is feasible. In contrast, when $p > 1/2$ and $k_i + k_j$ is large enough, it pays off to the lender to induce the (Repay, Default) outcome and benefit from the larger loan size possible.

Note, however, that in a (Repay, Default) equilibrium, switching from standard to bogus group does not change the repayment probability (it equals p in both cases) and the interest rate $(1/p)$ and also does not change the expected future value (equal to $2pV$ in both cases). However, forming a bogus group increases the expected joint output from $p(k_i + k_j)L$ to $2pk_iL$ while supporting the same maximum loan size $\frac{pV}{2}$. Hence, it is always (weakly) better for the borrowers to form a bogus group in the (Repay, Default) equilibrium. This implies that our focus on the (Repay, Repay) equilibrium in Section 2.2 is not restrictive when bogus groups can be formed since the maximum group payoff in the (Repay, Default) equilibrium is always weakly dominated by the bogus group payoff at contract (L_B, R_B).

C Appendix C – Excluding contract for unobservable k_i

In this section we analyze the optimal contract in which the lender, for exogenous reasons wants to rule out bogus groups. The lender also does not observe borrowers productivities. Therefore, the lender cannot condition the contract on k_i, k_j and all groups must receive the same contract. As in Section 3.2, holding the interest rate fixed from the break-even condition, the lender cannot screen by using different loan sizes since all groups prefer larger loans. The optimal excluding contract is such that: (i) any borrower group will be standard, (ii) each group member will optimally repay when her own project succeeds (iii) the lender breaks even and (iv) the ex-ante expected borrower joint payoff is maximized.

As in Sections 3.2 and 4 we have that the excluding contract $\{L^{\#}, R^{\#}\}$ solves:

$$\max_{L,R} \sum_{ij} q_{ij} W_{ij}(L, R|S) = \sum_{ij} q_{ij}[p(k_i + k_j)L - 2p(2-p)R + 2p(2-p)V]$$

$$\text{s.t. } (1), (3), \text{ and } R = \frac{L}{p(2-p)}$$

and

$$(k_i - k_j)L \leq 2(1-p)(V - R), \ \forall ij \in \{HH, HL, LL\} \quad (32)$$

where constraint (32) must hold for all types of groups. As in Section 2.3, the above constraint is binding for HL groups if $k_H - k_L > \frac{2}{p(2-p)}$ and slack otherwise. We therefore obtain,

Proposition 5B: *Suppose the lender wants to exclude bogus groups and does not observe borrowers' productivities. Let the quantities L_S, R_S, L_E and R_E are as defined in Propositions 1-3. Then:*

(i) the payoff-maximizing excluding contract is (L_S, R_S) if $k_H - k_L \leq \frac{2}{p(2-p)}$.

(ii) the payoff-maximizing excluding contract is (L_E, R_E) with $L_E < L_S$ if $k_H - k_L > \frac{2}{p(2-p)}$.

In this case there is an additional welfare loss compared to that in Proposition 5 coming from restricting the loan size to HH and LL groups if the productivity differential is large – while only borrowers in HL groups would choose to operate as bogus, all borrowers obtain inefficiently smaller loans.

D Appendix D – Joint Strategic Default Decision

In the previous sections, we assume that in a standard group each borrower makes their repayment decision individually, opening the possibility for strategic interaction and free riding incentives on the other member's repayment decision. We saw that this information or commitment friction leads to a tight restriction on loan size in standard groups and becomes an important factor influencing the contract design problem for the lender and group form choice. We now extend our basic setting by assuming that the repayment decision in a standard group is made collectively by both members, not by each individual independently. This can be interpreted as the members being able to observe each other's project output and/or enforce social or other penalties on uncoordinated strategic default.

As in Section 2.2, we start with the basic case where due to an exogenous reason, the formation of bogus groups is not possible. In the repayment stage the group makes a binary choice between repaying the entire group liability $2R$ or repaying zero. Similarly to the previous analysis, repaying an amount strictly between 0 and $2R$ is not optimal. At the repayment stage, the borrowers maximize the joint group payoff, thus they choose to repay in full if and only if[31]

$$R \leq V \tag{33}$$

Compare constraint (33) with constraint (3) from Section 2.2 – the collective strategic default decision relaxes the maximum loan size at which a repaying standard group can be supported to the level of the maximum loan size for a bogus group. The lender's break even condition remains

$$R = \tfrac{L}{p(2-p)}. \tag{34}$$

Hence, substituting for R into the no-default constraint (33), the problem of finding the optimal standard group contract is

$$\max_{L} W_{ij}\left(L, \tfrac{L}{p(2-p)}|S\right) = (p(k_i + k_j) - 1)L + p(2-p)V$$

$$\text{s.t. } L \leq p(2-p)V \tag{35}$$

Since the objective function is increasing in L, constraint (35) is binding at the optimum, thus we obtain:

> **Proposition D1**: *When bogus group formation is impossible and group members make the repayment decision collectively, the optimal standard group loan contract is* $\mathcal{S}' = (L_{S'}, R_{S'})$ *with* $L_{S'} \equiv p(2-p)V$ *and* $R_{S'} \equiv V$.

The joint payoff for an ij group under contract $\mathcal{S}'$ is

$$W_{ij}(L_{S'}, R_{S'}|S) = p^2(2-p)(k_i + k_j)V > W_{ij}(L_S, R_S|S).$$

Notice also that contract $\mathcal{S}'$ (like $\mathcal{S}$) is independent of k_i and k_j, therefore, it does not matter whether k_i and k_j are observable for the lender or not when bogus group formation is exogenously ruled out.

[31] If both projects succeed, the borrowers repay if $(k_i + k_j)L - 2R + 2V \geq (k_i + k_j)L$. Similarly, if only project i succeeds, the borrowers repay if $k_iL - 2R + 2V \geq k_iL$.

D.1 Standard vs. bogus groups

Compared to the list in Section 2.3.1, the net insurance value effect (item i) cancels out since $R = V$. Standard groups now receive larger loans $L_{S'} > L_B = pV$ and a lower interest rate, so the former items (iii) and (v) go in their favor too. The strategic default (free riding) effect (item iv) is no longer present when the default decision is made jointly. Only the additional expected-ouput item (ii) remains in favor of bogus groups. Hence the basic trade-off is now between the larger and lower-interest rate loans in standard groups vs. larger output in heterogeneous groups. Clearly, homogeneous groups have no incentive to be bogus in the collective default decision setting, as we formally show below.

Following Section 2.3.2, Lemma 1 implies that if the borrowers make the default decision collectively, at contract $\mathcal{S}'$ with $R_{S'} = V$ we have:

Proposition D2: *If borrowers make repayment decision collectively in a standard group and loan contract $\mathcal{S}' = (L_{S'}, R_{S'})$ is offered, any group with*

$$k_i - k_j > 0, \tag{36}$$

or equivalently, any HL group, will optimally operate as bogus group and consequently incur a loss to the lender.

Intuitively, remember that the RHS of condition (10) in Lemma 1 is the total ex-ante net benefit from forming a standard vs. bogus group for the same given contract (L, R). At contract $\mathcal{S}'$, since $R_{S'} = V$, this net value corresponding to the difference between the increased expected future value and the larger expected repayment in a standard vs. bogus group is zero. Thus, only the increased productivity effect (ii) remains. The "bogus group problem" now affects all heterogeneous groups, not only those with sufficiently high productivity differential (compare with Proposition 2).

We proceed to solve for the optimal loan contract for any group ij assuming that borrower productivities are observable. Analogously to Section 3.1, the lender's problem can be written as:

$$\max_{L,R,\tau\in\{0,1\}} \ \tau W(L,R|S) + (1-\tau)W(L,R|B)$$
$$\text{s.t. } \tau W(L,R|S) + (1-\tau)W(L,R|B) \geq \tau W(L,R|B) + (1-\tau)W(L,R|S) \quad \text{(IC)}$$
$$R \leq V \quad \text{(no default)}$$
$$R = \tau \frac{L}{p(2-p)} + (1-\tau)\frac{L}{p} \quad \text{(zero profit)}$$

where the total payoffs for a standard group, $W(L,R|S)$ and bogus group, $W(L,R|B)$ are now:

$$W(L,R|S) = \begin{cases} p(k_i + k_j)L - 2p(2-p)R + 2p(2-p)V, & \text{if } R \leq V \text{ (Repay)} \\ p(k_i + k_j)L, & \text{if } R > V \text{ (Default)} \end{cases}$$

and

$$W(L,R|B) = \begin{cases} 2pk_iL - 2pR + 2pV, & \text{if } R \leq V \text{ (Repay)} \\ 2pk_iL, & \text{if } R > V \text{ (Default)} \end{cases}$$

Proposition D3: *Suppose Assumption 1 holds, the borrowers make the strategic default decision jointly, and the borrowers' productivity parameters k_i and k_j are observed by the lender. Let L_E, R_E, L_B, R_B and the function $f(k_L)$ be as defined in the proof of Proposition 3. Define also $L_{S'} = p(2-p)V$, $R_{S'} = V$ and $\mathcal{S}' = (L_{S'}, R_{S'})$, $\mathcal{E} = (L_E, R_E)$*

and $\mathcal{B} = (L_B, R_B)$. Then, the optimal contract depending on the group composition and k_L, k_H and p is as follows:[32]

	Optimal contract and group form		
Parametric conditions:	*LL groups*	*HH groups*	*HL groups*
1. $k_H \in [k_L, f(k_L)]$	$\mathcal{S}'$, standard	$\mathcal{S}'$, standard	$\mathcal{E}$, standard
2. $k_H > f(k_L)$	$\mathcal{S}'$, standard	$\mathcal{S}'$, standard	$\mathcal{B}$, bogus

Proof of Proposition D3:
Remark D1: $W_{ii}(L, \frac{L}{p}|S) \leq W_{ii}(L, \frac{L}{p}|B)$ *for homogeneous groups if and only if* $L \geq (1-p)V$.
Proof: Indeed, $(2pk_i-(2-p))L+2p(2-p)V \leq 2(pk_i-1)L+2pV$ is equivalent to $p(1-p)V \leq pL$.□

- *Homogeneous groups*

Suppose the lender faces a homogeneous group and wants to induce a standard group ($\tau^* = 1$). As in the proof of Proposition 3, using Proposition D2 implies that constraint (IC) is satisfied. Thus, we obtain from the no-default and zero-profit constraints, $L^* = p(2-p)V = L_{S'}$ and $R^* = V = R_{S'}$. Suppose now the lender wants to induce a bogus group ($\tau^* = 0$) and hence $R = \frac{L}{p}$. Remark D1 implies that, if $p < 1/2$ then constraint (IC) is incompatible with the no-default constraint, $L \leq pV$. If $p \geq 1/2$ then, since the objective is increasing in L, setting $L^* = pV$ is optimal. To decide whether $\tau^* = 1$ or $\tau^* = 0$ (only valid choice if $p \geq 1/2$) is optimal, compare:

$$W(p(2-p)V, V|S) = 2p^2 k_i(2-p)V$$

with

$$W(pV, V|B) = 2p^2 k_i V$$

Clearly, the former is always larger, thus setting $(L^*, R^*) = \mathcal{S}'$ is always optimal for homogeneous groups.

- *Heterogeneous groups*

Suppose the lender wants to induce a standard group ($\tau^* = 1$). As in the proof of Proposition 3, Lemma 1 implies that constraint (IC) is satisfied if $L^* \leq L_E$. It is easy to show that $L_{S'} \geq L_E$ whenever $k_H \geq k_L$. Therefore, the optimal contract in this case is $L^* \equiv L_E$ and $R^* \equiv \frac{L_E}{p(2-p)}$. Suppose now the lender wants to induce a bogus group, $\tau^* = 0$. From Lemma 1, constraint (IC) is satisfied at $L^* = pV$ and $R^* = V$. Therefore, the optimal contract in this case is $L^* = L_B$ and $R^* = R_B$.

To decide whether choosing $\tau^* = 0$ or $\tau^* = 1$ is optimal, compare the total payoff from operating as standard group with contract (L_E, R_E) vs. operating as bogus group with contract (L_B, R_B). From the proof of Proposition 3 we know that $W(L_E, R_E|S) \geq W(L_B, R_B|B)$ whenever $k_H \leq f(k_L)$.■

D.2 Discussion

Intuitively, standard groups are always better for homogeneous borrower pairs since they benefit from the higher expected future value. The larger loan size, $L_{S'} > L_B$ and lower interest rate are offset by the larger expected repayment. There is no increased output (item ii). Unlike in

[32] In the table, "standard" corresponds to $\tau^* = 1$ while "bogus" corresponds to $\tau^* = 0$ in the solution to the contracting problem.

Proposition 3 this result does not depend on the absolute productivity level since R^* is the same for both standard and bogus groups.

For heterogeneous groups, the trade-off is between the increased output and the higher future value. Therefore, when the productivity differential is sufficiently large (when $k_H > f(k_L)$) it is optimal to set $L^* = L_B$ and induce a bogus group. When the productivity differential is relatively small the benefits from a standard group outweigh the loss of extra output, but as in Proposition 3 and Corollary 1 an inefficiently small loan size, L_E must be offered to deter heterogeneous borrower pairs to benefit from the lower interest rate $\frac{R_E}{L_E}$ and form a bogus group.

D.3 Unobserved productivities

The problem is very similar to problem (UPP') with the only difference that the no default constraint for a standard group is $L_N \leq p(2-p)V$ instead of (20) and that

$$W_{ij}\left(L_M, \tfrac{L_M}{p}|S\right) = \begin{cases} (p(k_i+k_j) - 2(2-p))L_M + 2p(2-p)V, & L_M \leq p(2-p)V \text{ (Repay,Repay)} \\ p(k_i+k_j)L_M, & L_M > p(2-p)V \text{ (Default,Default)} \end{cases}$$

Following the previous analysis, it is easy to check that the proof of Lemma 2 still applies and so $L_M^* = pV$ as in the individual default decision setting. The proof of Proposition 4 also remains virtually the same as before but using $L_{S'} = p(2-p)V$ instead of L_S. This means that the (IC) constraint is always satisfied for homogeneous groups.

Proposition D4: *Suppose Assumption 1 holds, the borrowers make the strategic default decision jointly, and the borrowers' productivity parameters k_i and k_j are unobservable to the lender. Let $L_E, R_E, L_B, R_B, L_F, R_F$ and the function $f(k_L)$ be as defined in Propositions 3 and 4. Let also $L_{S'} = p(2-p)V$, $R_{S'} = V$ and call $\mathcal{S}' = (L_{S'}, R_{S'})$, $\mathcal{E} = (L_E, R_E)$, $\mathcal{F} = (L_F, R_F)$ and $\mathcal{B} = (L_B, R_B)$. Then, the optimal contract menu and chosen group form depending on k_L, k_H and p is as follows:*[33]

	Optimal menu	**Selected contract and group form**		
Parametric conditions:		*LL groups*	*HH groups*	*HL groups*
1. $k_H \in [k_L, f(k_L)]$	$\mathcal{E}$,$\mathcal{B}$	$\mathcal{E}$, standard	$\mathcal{E}$, standard	$\mathcal{E}$, standard
2. $k_H > f(k_L)$	$\mathcal{F}$,$\mathcal{B}$	$\mathcal{F}$, standard	$\mathcal{F}$, standard	$\mathcal{B}$, bogus

Proof of Proposition D4: From the above discussion, we know that $L_M^* = L_B$ so contract $\mathcal{M}$ is always (L_B, R_B). As in Proposition D3 we have $L_{S'} \geq L_E$ whenever $k_H \geq k_L$ which is always true. It is also easy to show that $L_{S'} > L_F$ holds for any $p \in (0,1)$. Then, following part C of the Proposition 4 proof, this implies that the (IC) will always bind for HL groups. In addition, as shown in the proof of Corollary 2,

$$L_3(HL) \geq L_E \iff k_H \geq f(k_L)$$

and it is also easy to show that[34]

$$L_3(HL) \geq L_F \iff k_H \geq f(k_L)$$

[33] In the table, "standard" corresponds to $\tau^* = 1$ while "bogus" corresponds to $\tau^* = 0$ in the solution to the contracting problem.

[34] Obviously $L_E \neq L_F$, however, it turns out that, for any given k_L, the expressions $L_3(HL)$, L_F and L_E taken as functions of k_H cross at the same point (at which $k_H = f(k_L)$) and thus $L_3(HL)$ is larger than both L_F and L_E whenever $k_H \geq f(k_L)$.

If $k_H \in [k_L, f(k_L)]$ then, since $L_E \geq L_3(HL)$ and $L_E \leq L_{S'}$ setting $L_N^* = L_E$ is optimal and HL groups select contract $\mathcal{N}$ and operate as standard. Similarly, as in the proof of Corollary 2(d), homogeneous groups also select contract $\mathcal{N}$ (here (L_E, R_E)) in this case and operate as standard. If $k_H > f(k_L)$ we have $L_3(HL) > L_F$ and since also $L_F < L_{S'}$ this implies that $L_N^* = L_F$ in this case. As in the proof of Corollary 2(d), we again obtain that homogeneous groups optimally select contract $\mathcal{N}$. ■

The results parallel those in Propositions 4 – with unobservable productivities all groups receive the reduced loan size contract $\mathcal{N}$ (here equal to either E or F) to deter them from switching to bogus, hence efficiency is reduced relative to the observable productivities case in Proposition D3. Also, as in Proposition D3, if the borrowers decide jointly on strategic default, only heterogeneous groups with sufficiently large productivity differential would optimally be bogus.

E Appendix E – Data Details

There are six category of proposed usage of the loan reported when borrowers apply for a loan, namely, agriculture/ forestry/ animal husbandry/ fishery (AFAF), production and processing industry (manufacture), transport industry (transport), wholesale and retail trade (wholesale), social services (service), and rural housing reconstruction (housing). We construct indicator variables for the proposed loan usage accordingly. For example, the variable $AFAF$ equals one if the proposed loan usage is in the agriculture/forestry/animal husbandry/fishery industry and zero otherwise. Similarly, for the variables $Manufacture$, $Transport$, $Wholesale$, $Service$ and $Housing$.

There are also five education levels reported by the borrowers in their loan applications: college or higher (*college*), high school (*highschool*), junior school (*junior*), primary school (*primary*), and below primary school (*below*). As with the loan use we construct indicator variables for each education level that equal to one if the individual reports that level and zero otherwise.

We also construct three county dummies: *beizhen*, *xiuyan* and *xingcheng*, which equal 1 if the borrower belongs to the respective county and zero otherwise. In addition, we construct three ethnicity dummies: Han, $Manchu$, $Mongols$ which equal one if the borrower belongs to the respective ethnicity and zero otherwise. A marital status dummy, *married* is also created equal to one if the borrower is married and zero otherwise. We do not use gender dummies since all borrowers in our sample are women (in comparison, the 2011 CFPAM annual report reports that 91% of its borrowers all over China are women).

In terms of the proposed usage of the loan, our sample composition is very similar to that reported by CFPAM in its annual report country-wide.[35] This suggests that our sample is not significantly different from the universe of all contracts issued by CFPAM and can be viewed as a reasonable representative sample. The majority of the borrowers in our rural sample have either primary school or junior school level of education while less than 1% of the borrowers (3 out of 366) in our sample obtain college or higher education. Around 50% of the sample borrowers come from Beizhen, while 25% each are from Xiuyan and Xingcheng. 70% of the borrowers are Manchu, nearly 30% are Han while less than 1% are Mongol.

[35] In 2011 CFPAM reports that the fractions of total loans invested in AFAF, manufacture, service, wholesale, transport and housing industries were: 67.84%, 2.98% , 4.98%, 12.63%, 6.36%, and 0.13%. In our sample, the corresponding fractions are 80%, 5.5%, 2.2%, 8.1%, 2%, and 2% accordingly. See "Microfinance annual report 2011" at http://www.cfpamf.org.cn.

References

Ahlin, C. (2012). Group lending under adverse selection with local information: The role of group size. *Mimeo*. 1

Ahlin, C. and Waters, B. (2011). Dynamic lending vs. group lending in the ghatak/stiglitz/weiss adverse selection model. *Mimeo*. 1, 2

Armendariz de Aghion, B. (1999). On the design of a credit agreement with peer monitoring. *Journal of Development Economics*, 60(1):79–104. 1, 3, 8

Besley, T. and Coate, S. (1995). Group lending, repayment incentives and social collateral. *Journal of Development Economics*, 46:1–18. 3, 10

Chowdhury, P. R. (2005). Group-lending: Sequential financing, lender monitoring and joint liability. *Journal of Development Economics*, 77:415–439. 1

Ghatak, M. and Guinnane, T. (1999). The economics of lending with joint liability: theory and practice. *Journal of development economics*, 60(1):195–228. 1, 3

Karlan, D. S. (2005). Social connection and group banking. Working Paper. 3

Morduch, J. (1999). The microfinance promise. *Journal of Economic Literature*, 37:1569–1614. 2

Mosley, P. (1986). Risk, insurance and small farm credit in developing countries: A policy proposal. *Public Administration and Development*, 6(3):309–319. 2

Udry, C. (1990). Credit markets in northern nigeria: Credit as insurance in a rural economy. *The World Bank Economic Review*, 4(3):251–269. 2

Figure 1: The equilibrium contracts and group forms under different parameter values when productivity is observable

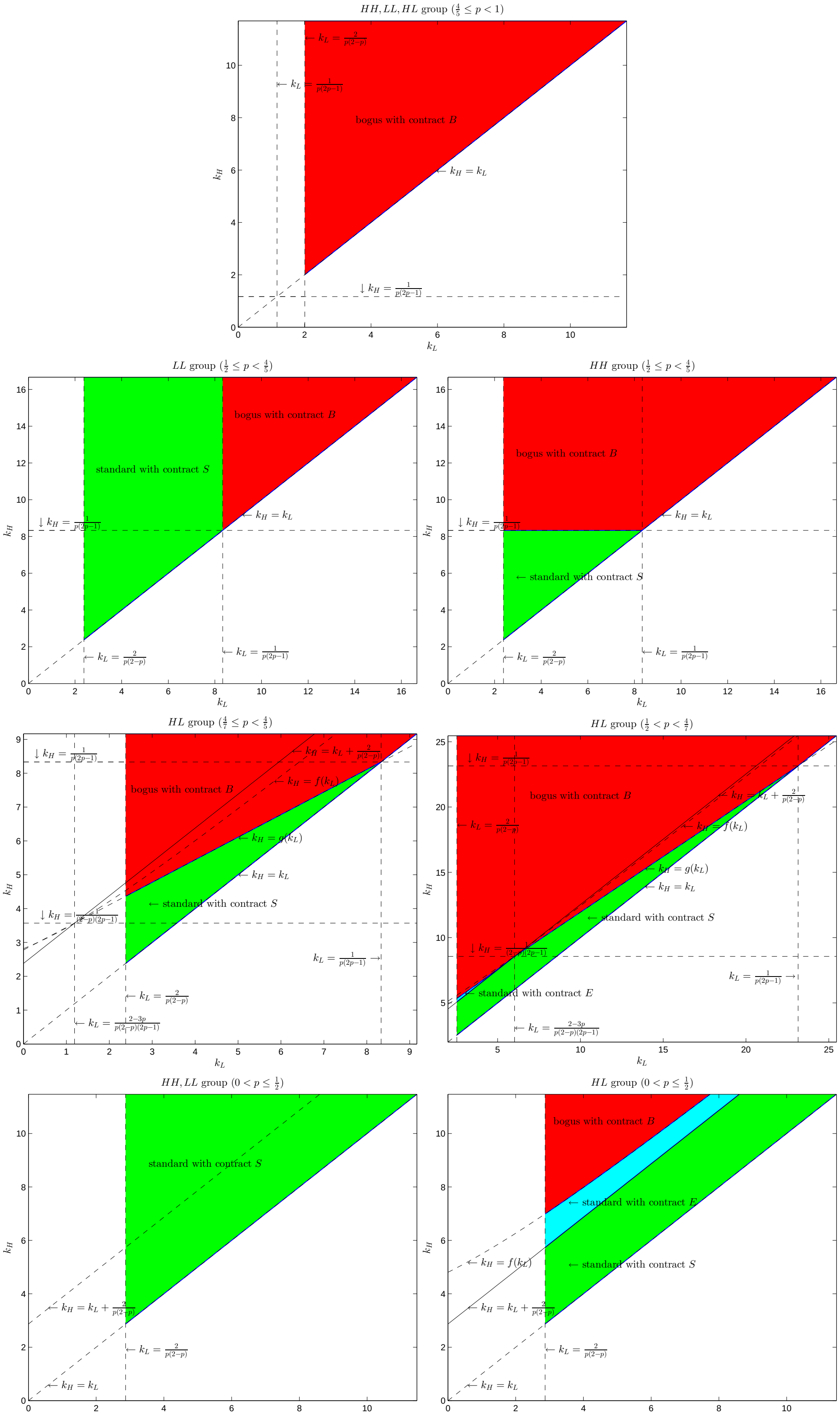

Figure 2: The equilibrium contracts and group forms under different parameter values when productivity is unobservable (The figure for HH, LL and HL group when $4/5 \leq p < 1$ is the same with the top panel of figure 1, i.e., all of the groups select contract $\mathcal{M}$ (i.e. $\mathcal{B}$) and operate as bogus groups.)