Bogus Joint Liability Groups in Microfinance*

Alexander Karaivanov†  Xiaochuan Xing  Yi Xue
Simon Fraser University  Yale University  UIBE
November 2018

Abstract

In a random sample of clients of CFPAM, the largest microlender in China, 73% of the observed joint-liability groups practice Lei Da Hu. That is, one person uses all group members’ loans in a single project. We call such borrower groups ‘bogus groups’. The Lei Da Hu practice violates CFPAM rules and is inconsistent with the key principle of group lending that each borrower should use their loan in a separate project (what we call ‘standard group’). We extend the theory of group lending by analyzing the endogenous formation and coexistence of standard and bogus groups and characterize the efficient lending terms. The chosen group form depends on the borrower projects’ productivities and probability of success. Bogus groups are formed by heterogeneous borrowers, when the gains from larger expected output exceed the foregone default-risk diversification. Accounting for bogus groups in their lending strategy can help MFIs avoid losses and raises productive efficiency and welfare.

JEL classification: C72; D82; G20; O12
Keywords: microfinance; group lending; Lei Da Hu; phantom borrowers; strategic default

*We thank audiences at the ThReD, SED and the Econometric Society meetings for many useful comments and suggestions. Karaivanov is grateful for financial support from the Social Sciences and Humanities Research Council (SSHRC), grant 435-2013-0698. Xue acknowledges financial support from the National Natural Science Foundation of China (NSFC), project 71471040.
†Corresponding author: Department of Economics, Simon Fraser University, akaraiva@sfu.ca


1 Introduction

Data from 353 randomly chosen clients of the leading Chinese microlender, the China Foundation for Poverty Alleviation – Microfinance (CFPAM) show that 73% of the surveyed joint-liability groups (59 out of 81) report practicing Lei Da Hu. That is, a single borrower uses the sum of all loans given to the group members while the rest of the group only act as cosigners. We call such borrower groups bogus groups.

Bogus groups violate a basic principle of group lending, both as implemented in microcredit practice and as modeled in the literature – the requirement that each borrower invest her loan in their own business project (what we call standard group). Many authors have emphasized the benefits of joint-liability group loans in microcredit, stemming from leveraging social capital and reducing default rates as the borrowers cover for each other in case of inability to repay. In contrast, in a bogus group all funds are invested in a single project or business and, if the cosigners are poor, then because of limited liability their nominal role as guarantors offers no protection to the lender. It is therefore surprising that a very large fraction of CFPAM groups report to be bogus since, according to the theory, group lending in its standard form can reduce transaction or agency costs and increase the probability of repayment and obtaining future credit. In addition, the Lei Da Hu practice violates the CFPAM loan covenant terms, although such infractions may be hard to verify in reality.

The existing theoretical literature on microcredit has focused almost exclusively on comparing group lending with individual lending. While bogus groups share a common characteristic with individual loans in the fact that, with limited liability, the probability of repayment depends on the success of a single investment, the two differ in several important aspects. First, bogus groups allow the borrowed funds to be invested in a project that is selected among all group members’ projects (e.g., the project with the highest expected return). Second, by pooling the members’ loans, bogus groups allow a larger scale of investment.

To the best of our knowledge this is the first paper to formally model and analyze Lei Da Hu and bogus groups in microfinance and compare them to standard joint-liability groups. We characterize the borrowers’ endogenous choice of group form, the possible coexistence of both group forms, and the lenders’ choice of loan terms in the presence of standard and bogus groups. In addition, we discuss individual loans in more detail in Section 5.3.

We highlight five key differences between standard and bogus groups. The trade-off among these factors shapes the endogenous group form choice by the borrowers and the lender’s offered loan terms.

(i) risk pooling / diversification – the expected value of continued access to credit is larger in a standard group, due to the lower default probability stemming from project-failure risk diversification;

\footnote{The cosigning members may agree to participate in a bogus group because the person who uses the funds may offer them monetary or other compensation or because they may want to ask this person to cosign a future loan.}

\footnote{See the literature review at the end of Section 1.}
(ii) **expected repayment** – the risk diversification effect (i) is partially offset by the larger expected repayment in a standard group, arising from the requirement to cover for other members;

(iii) **expected output** is larger in bogus groups with heterogeneous member projects, since all funds are optimally invested in the highest-productivity project;

(iv) **the interest rate** is lower in a standard group, since the repayment likelihood is larger as per item (i)

(v) **the loan size** is larger in a standard group due to the lower interest rate and lower default risk. Overall, items (i), (iv) and (v) favor standard groups while (ii) and (iii) favor bogus groups.

Specifically, we model borrowers, each of whom is endowed with a single investment project. The projects are heterogeneous in their productivity (return) which may be observable or unobservable to the lender. Each project either succeeds (yields positive output), with some positive probability, or fails (yields zero). The borrowers have no other funds and must borrow from a perfectly competitive microfinance sector. Only group loans are offered (as observed in the CFPAM data) and, for simplicity, they are made to groups of two borrowers (in Section 5 we study extensions allowing individual loans and more than two borrowers per group). The project outcomes are i.i.d. across the borrowers. Project output is non-verifiable by the lender. Thus, a borrower can strategically default (declare project failure) if s/he finds this optimal. In case of project failure we assume that the lender cannot collect anything from that borrower – there is limited liability.

After receiving their loans, the borrowers choose to operate either as a standard group (each person invests in her own project) or as a bogus group (one person uses all loaned funds). The group form is unobserved by the lender. There is joint liability – all borrowers bear full responsibility for the group’s total debt. If full repayment is not received by the lender, all group members are excluded from access to future credit. In our baseline model the group members make a joint decision whether to repay or not. We consider individual repayment decisions in Section 5.2.

We first analyze standard joint-liability groups, in which each borrower invest in their own project, and characterize the loan terms that maximize the group payoff subject to no strategic default. We then model bogus groups, in which all funds are invested in a single project, and compare them to standard groups. We show that a competitive lender who offers standard-group loan terms to everyone would suffer a loss when some borrowers (in equilibrium, those with heterogeneous projects) form bogus groups.

We then analyze the mechanism design problem of a lender who takes into account that bogus and standard groups form endogenously and can coexist for given loan terms, but who cannot observe the group form choice. In our baseline setting, assuming observable project productivities, the loan terms (size and repayment amount) and the chosen group form depend on the group composition. Specifically, for borrowers with homogeneous projects there is no expected output gain from forming a bogus group (item iii) and, hence, the constrained efficient (subject to no

---

3We assume free entry by lenders which implies that they make zero expected profits per loan in equilibrium (no cross-subsidization). An alternative interpretation could be non-profit / NGO lenders.

4We assume that the borrower groups are formed before the project productivities are known (exogenous matching). In Section 6 we briefly discuss the implications of allowing for endogenous matching.
default) loan size and interest rate are offered. In contrast, for borrower groups with heterogeneous projects, the offered loan terms depend on the productivity differential between the projects. If the differential is relatively small, the lender offers loan terms that induce a standard group, taking advantage of risk diversification. The loan size is, however, reduced relative to that offered to homogeneous groups, to prevent the borrowers to switch to bogus group. In contrast, for a large productivity differential, it is optimal to forego diversification and offer loan terms that induce a bogus group, that is, the lender encourages all funds to be invested in the highest-productivity project.

In Section 4.2 we show that our main results remain qualitatively unchanged when the project productivities are unobserved by the lender. In that setting, in addition to the moral hazard problem of unobserved group form choice, the lender must also address the adverse selection problem of groups possibly misreporting their project composition/type. As a result, the lender offers a menu of two loans which differ in their size and interest rate. Again, we show that bogus groups emerge when the productivity differential across the members’ projects is sufficiently large. For homogeneous borrowers, or those with small productivity differential, the optimal loan terms induce a standard group but the loan size is reduced to ensure incentive compatibility and truth-telling.

In Section 5 we consider four extensions of the baseline model. We first show how the analysis can be generalized to groups of more than two members (Section 5.1). In Section 5.2 we study the case of individual (non-cooperative) repayment decisions. Because of the joint liability clause, individual repay/default decisions introduce strategic interaction between the group members – a borrower may choose to default strategically if she knows that the other member would find it optimal to repay for both. We show that this free-riding effect tightens the no-strategic-default constraint in standard groups, relative to the baseline model with a joint repayment decision, and this results in a smaller feasible loan size. The strategic effect also implies that bogus groups could arise for homogeneous groups with high project productivities (see Appendix B for details). In a third extension (Section 5.3) we discuss allowing for individual loans and show that in our model they are dominated by group lending, that is, no group of borrowers would be better off by switching to individual loans. Finally, in Section 5.4 we study the problem of a lender who, for exogenous reasons, wants to prevent the formation of bogus groups.5 This results in lower borrower welfare since forcing all borrowers to form standard groups is inefficient when the project productivity differential is large and higher expected output can be achieved in a bogus group.

Related literature

There exists a large literature on joint liability microfinance and the comparison between group joint-liability lending and individual lending (Ghatak, 2000; Chowdhury, 2005; Gangopadhyay et al., 2005; Banerjee et al., 1994; Rai and Sjostrom, 2004, among many others). As shown in both theory and practice, the main advantage of group lending is that it can create a substitute for asset collateral by using joint liability and the social capital embedded in the borrowers’ networks.

5For example, a lender may deem bogus groups illegal or the interest rate may be regulated and the lender is unable to break even if bogus groups form. Alternatively, the lender (e.g., a non-profit MFI) may want each borrower to invest in their own project for non-monetary reasons, for example, to improve their skills, business experience, etc.
and relationships to mitigate moral hazard, adverse selection, costly state verification or debt enforcement problems (Mosley, 1986; Udry, 1990; Besley and Coate, 1995; Morduch, 1999; Ghatak, 1999; Ahlin and Waters, 2011; Ahlin, 2012). In comparison to individual liability loans, the joint liability design also allows lending at lower interest rates, due to the higher repayment rate enabled by risk diversification, peer selection, peer monitoring and peer enforcement within the group (Ghatak and Guinnane, 1999; Karlan, 2005; Besley and Coate, 1995; Armendariz de Aghion, 1999; Ghatak, 2000; Stiglitz, 1990). With limited enforcement, however, joint liability may, when a borrower is unable to pay for another member, decrease the repayment rate relative to individual lending (Besley and Coate, 1995). De Quidt et al. (2016) further show that individual lending can be welfare improving relative to joint liability when borrowers have sufficient social capital to sustain mutual insurance.

On the empirical side, Gine and Karlan (2014) find no significant difference in repayment rates between group and individual loans in the Philippines. Ahlin and Townsend (2007a) use data from Thai borrowing groups and find a U-shaped relationship between borrower’s wealth and the prevalence of joint-liability relative to individual loans. Ahlin and Townsend (2007b) further document that repayment rates are negatively affected by the joint liability rate and social ties and positively affected by the strength of local sanctions or correlated returns. De Quidt et al. (2017) explore the link between commercialization in microfinance and the perceived decline in joint liability lending. Using MIX Market data they find no evidence for such decline but document an increase in competition and shift from non-profit to for-profit lending which, as they show theoretically, can cause lenders to reduce the use of joint liability loans.

The above papers assume that each borrower invests in a separate project, often with uncorrelated returns across the borrowers (what we call standard groups). The authors mostly focus on the comparison and possible (dis-)advantages of joint liability loans relative to individual loans. In contrast, motivated by the evidence from China, we model and analyze the endogenous formation of bogus joint liability groups in which all loans are pooled and invested into a single project, as well as their coexistence alongside standard joint liability groups.

Further evidence on the importance of the Lei Da Hu (bogus group) practices that we study comes from Pakistan where, in a policy report, Burki (2009) observes that: “...often, the group leader had accessed more loans from an MFI than the MFI had record of by borrowing through a dummy or ghost borrower”. Also related is the qualitative study by Cieslik et al. (2015) on ‘unruly’ entrepreneurs in rural Burundi. The authors describe strategies used by poor entrepreneurs to bypass institutional microcredit rules. One strategy is “loan arrogation”, in which an entrepreneur asks another community member to obtain additional credit to be invested in the entrepreneur’s business. The authors argue that this allows for “...larger-scale investment, cementing social bonds and empowerment”, and more generally, that such illicit practices can be interpreted as value-creating acts.

---

6 We thank J-M. Baland for this reference.
2 Microfinance and Lei Da Hu in China

Microfinance was introduced in China in the early 1990s with the primary goal of alleviating rural poverty. Since 1996 the government has regarded microfinance as an effective channel of credit provision to the poor (Zhang et al., 2010). Non-profit and NGO microfinance institutions (MFIs) play a major role in this process. The main non-profit MFI in China, CFPAM (China Foundation for Poverty Alleviation – Microfinance) is the largest microlender by total issued loans and active members. CFPAM has been using joint-liability lending since its founding, offering only group loans until 2014. Recently most Chinese MFIs have introduced individual or other loan forms alongside group lending. This could be partly because of the Lei Da Hu phenomenon that we study and also reflects general trends worldwide, as microlenders expand their outreach beyond the extremely poor, who were traditionally targeted.

Microcredit in China is commonly regarded as constrained by financial regulation, lack of supply of wholesale funds, and credit risk. The latter is manifested in part as Lei Da Hu – bogus microfinance groups or ‘phantom’ borrowers. According to the Bank for International Settlements (BIS) (2010), “...the complexity and distinctive features of microlending, especially the decentralized lending process, raise important risk management issues for microfinance activities and institutions”. In particular, the BIS report mentions the problem of phantom borrowers, in whose presence the risk diversification and peer monitoring benefits of group lending are weakened or absent and the resulting increase in credit risk can cause MFIs to incur losses. Another policy report, by Planet Rating, a microfinance rating agency, concludes that CFPAM’s control functions regarding Lei Da Hu / phantom borrowers practices were “...not sufficiently formalized and existing control forms are not utilized” (Planet Rating, 2005). These reports, together with the additional evidence from Pakistan and Burundi discussed at the end of Section 1, suggest that Lei Da Hu or phantom borrowers can be an important factor affecting MFI’s credit risk and performance.

As further motivating evidence for the importance of Lei Da Hu / bogus groups, we use a confidential random sample of 81 CFPAM group loans issued in 2011 to 353 female borrowers in three of the poorest counties in Liaoning province in China – Beizhen, Xiuyan and Xingcheng. CFPAM advertises regularly in rural China and most people are aware of its microcredit program. Interested borrowers first form a group and then approach CFPAM. The CFPAM lending rules stipulate that each group must consist of 2 to 5 self-chosen members from the same village. There should be no more than one member from the same household per group and it is also not desirable for close relatives to be in the same group. After a group is created it elects a group leader from its members. If the group meets the basic requirements (each member has an existing business or business plan, understands the credit rules and needs a loan), CFPAM holds a training session explaining joint liability, group operations, the importance of group solidarity, monitoring, and meeting attendance by all members. Then each group member receives their first loan.

7 In 2015 CFPAM issued 324,228 loans with total value 4.13 billion RMB and served 306,101 borrowers.
8 See for example: http://www.rong360.com/gl/2015/11/24/82.097.html. Grameen Bank – China has kept group lending as its main lending practice but borrower groups are now formed using an online platform, instead of the traditional way of grouping people closely familiar with each other.
process typically takes a week.

Our random sample of CFPAM loans includes loan identification number, starting date, size, duration, required monthly payment, interest rate and proposed loan use (by the applicant). There is also basic demographic information about the borrowers, including age, gender, ethnicity, education and marital status. According to CFPAM loan officers, some of the self-reported information is unverifiable, e.g., the proposed loan use or a borrower’s education. In addition to the administrative CFPAM data, each borrower in the sample was called by phone at randomly chosen time during working hours by the survey team and asked: “Does your borrower group practice Lei Da Hu?” The survey team explained the meaning of Lei Da Hu if a borrower was unaware. In 59 of the 81 groups (73%) at least one member reports that their group practices Lei Da Hu; in 54 groups all borrowers report so; and in 56 groups the majority (50% or more) of the members report so.

Unfortunately, the CFPAM data have serious limitations – the small sample size and available variables do not allow formal statistical analysis. There is almost no variation in the interest rate and repayment schedule: 79 of the 81 loans have the same interest rate (13.5%) and only 2 loans have a different interest rate (12% and 16%). Only 1 of the 81 loans (the outlier with 16% interest) has a different total required number of repayments (4, versus 10 for the rest). The loan size and corresponding monthly required repayment vary across the borrowers but do not differ in a statistically significant way with the reported group form (bogus or standard). Sample statistics are reported in Table A1 in Appendix C.

The main patterns observed in the CFPAM data are:

(i) bogus groups (self-reported Lei Da Hu practice) constitute a large fraction of the random sample and coexist with standard groups.

(ii) being in a bogus group is not associated with the observed borrower characteristics: age, marital status, education (% with college is the only exception), ethnicity and reported intended loan use; although it could be associated with other unobserved by us factors, such as project quality or heterogeneity.

(iii) CFPAM offered similar loan terms to all groups.

Facts (i) and (ii) motivate our theoretical model of bogus groups and their coexistence with standard groups. The observed homogeneity of loan terms in the sample (fact iii) is consistent the hypothesis that CFPAM was unaware or ignored bogus groups and offered standardized loan terms to all borrowers. Our earlier discussion about losses incurred by Chinese MFIs from phantom borrowers / Lei Da Hu lends indirect support for this hypothesis. Unfortunately, we do not have direct data on default rates or subsequent survey rounds to test this formally.

---

9All borrowers were also asked and confirmed their awareness of the joint liability clause and knowing the other group members personally.
10We did try multiple econometric specifications.
3 Model

3.1 Setting

Consider an economy populated by lenders and borrowers. Each borrower has a single investment project requiring initial investment $L > 0$. The borrowers have no wealth. Hence, the entire initial investment $L > 0$ needed to implement the investment project must be financed at time $t = 0$ by taking a loan from a lender (microfinance institution). There are two types of projects: a ‘conventional’ project with productivity $k_L$ and a ‘high-return’ project with productivity $k_H$, where $k_H > k_L > 0$. The project output, $Y_i$ for $i = L, H$ is generated at $t = 1$ and is stochastic:

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

The parameter $p \in (0, 1)$ is interpreted as the probability of a project being successful. The project output realizations are assumed i.i.d. across borrowers. The borrowers’ outside option (if they do not invest) is normalized to zero.

Both the lenders and the borrowers are risk-neutral. The lenders have an opportunity cost of funds normalized to 1. Because of information or enforcement frictions, the project return $Y_i$ is non-verifiable by the lender. This allows the possibility of strategic default – a borrower can report failure while in fact her investment project has succeeded. In addition, the borrowers have limited liability: if the investment fails (yields zero), the borrower defaults involuntarily and the lender does receive anything. The loan terms must therefore be chosen so that borrowers are given incentive to pay back when their projects succeed.

We assume free entry in the micro-lending sector. This implies that any lender earns zero profits on each loan and rules out cross-subsidization across loans. We also assume that only group loans are provided, as in our motivating evidence from China. Allowing for individual loans is discussed as an extension, in Section 5.4. As in most of the theory literature on microcredit, suppose groups consist of two members (we relax this in Section 5.1). The group lending contract thus consists of two loans, each with size $L$ and required repayment $R$ where $R \geq L$ (the gross interest rate is $R/L$). Since the borrowers have no wealth, requiring collateral is infeasible. Instead, the group loan has a joint liability clause – each member is fully responsible for the entire group obligation $2R$. If the lender does not receive $2R$ at $t = 1$, from one or both borrowers combined, then both borrowers are cut off from access to future credit.

To keep the analysis tractable, suppose that the borrowers’ project returns $k_i$ are i.i.d. over time, e.g., each borrower draws a new project per credit cycle. That is, we can think of the borrower groups as being randomly ‘reset’ in any future period $t > 1$, that we do not model explicitly. The i.i.d. assumption allows us to focus on a single credit cycle (two periods, $t = 0$ and $t = 1$ only), since it implies that the ex-post continuation value of future credit access is the same for all borrowers who are not cut off from credit because of default.\footnote{Our emphasis in this paper is on the heterogeneity in project quality within borrower groups. In reality, the...} We call this ex-post continuation (future-credit...}
access) value \( V \), where \( V > 0 \), and treat it as given hereafter. Note, however, that we do allow the \textit{ex-ante} (expected) continuation value of future credit access to differ across borrower groups when their endogenous probability of repayment is different.

Each borrower group can operate as either a \textit{standard group} or as a \textit{bogus group}. In a standard group each member invests \( L \) into her \textit{own} business project, as assumed in the literature and as required by the MFI loan terms in practice. In contrast, in a bogus group the members invest the total loan amount \( 2L \) into \textit{one} of the two projects.

The project productivities \( k_i \) and \( k_j \) are known to both borrowers in a group but may be observed or unobserved by the lender (we study both cases). A possible interpretation of the observed productivity case is that most microlenders require detailed information about the investments that borrowers intend to implement before providing a loan. To simplify the notation, we will say that a group ‘has type \( ij \)’ if the productivities of its members’ projects are \( k_i \) and \( k_j \) where \( i, j \in \{H, L\} \). There are three possible group types, \( ij \in \{HH, LL, HL\} \). Without loss of generality assume that \( k_i \geq k_j \).

As our baseline setting we assume that the loan repayment decision is made \textit{jointly} by the group members. The members default if their expected joint payoff exceeds the payoff from repaying and keeping access to future credit. This assumption can be interpreted as the members being able to observe each other’s project outputs or as sharing sufficient social capital to enforce social penalties in case of uncoordinated strategic default. In Section 5.2 we also consider the alternative scenario of individual (non-cooperative) default decisions, which opens the possibility for strategic interaction and free riding on a partner’s joint-liability obligation.

\textbf{Model timing}

- Stage 0: a group of two borrowers is formed; then each borrower draws an investment project with productivity \( k_i, i \in \{H, L\};^{12} \)
- Stage 1: the lender offers a group loan with terms \( (L, R) \);
- Stage 2: the borrowers choose to operate as a standard or bogus group (unobserved to the lender);
- Stage 3: each borrower’s investment is launched and the project output \( Y_i \) is realized one period later; \( Y_i \) is non-verifiable by the lender;
- Stage 4: repayment / default decision is made jointly by the group members;
- Stage 5: all payoffs are realized (see below for details).

\footnote{continuation value may depend on various factors such as age, education, family status, business experience, etc. but we abstract from those.}

\footnote{This timing rules out endogenous matching between borrowers (see Section 6 for more discussion).}
3.2 Standard Groups

We start the analysis with the basic setting from the literature in which bogus joint liability groups are exogenously ruled out and only standard groups exist. Assume for now that the group member’s project productivities $k_i$ and $k_j$ with $k_i \geq k_j$ are observed/known by the lender (the case of unobservable productivities will be discussed later on).

We say that a group loan is feasible if each member’s project generates sufficient output upon success to be able to cover both $R$ (paying for oneself) and $2R$ (paying for oneself and one’s partner). Without this condition, joint liability loans are disadvantageous relative to individual loans. Feasibility requires \( \min\{k_iL, k_jL\} \geq 2R \), or

\[
R \leq \frac{1}{2}k_jL.
\]

We make the following assumption which, as shown below, ensures that the project returns are sufficiently large so that repayment is always feasible.\(^{13}\)

**Assumption 1 (feasibility)**

\[
k_H \geq k_L \geq \frac{2}{p(2-p)}
\]

Assumption 1 implies

\[ pk_i > 1 \text{ for } i = L, H \quad (\text{SE}) \]

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

In the repayment stage (Stage 4 in the timing) the borrowers choose between repaying the entire group liability $2R$ or repaying zero (default). Note that it is never optimal to make a partial repayment (an amount between 0 or $2R$) since, because of the joint liability clause, either defaulting (repaying zero) and forfeiting the continuation value $2V$, or repaying in full and securing $2V$ is dominant strategy. Thus, to maximize the joint group payoff, the borrowers repay if and only if\(^{14}\)

\[
R \leq V
\]

Lending to a standard group increases the loan repayment probability from $p$ (the project success probability) to $1 - (1 - p)^2 = p(2 - p)$. This reflects the classic argument for joint-liability lending as compared to individual lending. For loan terms $(L, R)$ satisfying (1) and (2), the expected payoff of a standard group of type $ij$ equals

\[
W_{ij}^{S}(L, R) = p(k_i + k_j)L - 2p(2-p)R + 2p(2-p)V.
\]

The first term in (SEP) is the expected output in a standard group, the second term is the expected

\(^{13}\)Contrast with Besley and Coate (1995) where a borrower may be unable to repay both loans.

\(^{14}\)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 $k_iL - 2R + 2V \geq k_iL$. 
repayment and the third term, \(2p(2 - p)V\) is the expected continuation value of access to future credit.

The lender receives \(2R\) with probability \(p(2 - p)\) and zero otherwise (limited liability). Hence, the lender’s break-even condition is \(2p(2 - p)R - 2L \geq 0\), or

\[
R \geq \frac{L}{p(2-p)}.
\]  

(3)

The optimal loan terms for standard group \(ij \in \{LL, LH, HH\}\) solve

\[
\max_{L,R} W_{ij}^S (L, R) \\
\text{s.t. (2) and (3)}
\]

With free entry, the zero-profits condition (3) must bind at optimum, that is \(R = \frac{L}{p(2-p)}\). If not, another lender can offer a loan with lower repayment \(R\) and still break-even. The feasibility condition (1) is then equivalent to \(\frac{L}{p(2-p)} = R \leq \frac{1}{2} k_j L\), which is satisfied given Assumption 1.

**Proposition 1**

(a) The loan size and repayment terms solving problem (SP) are \((L_S, R_S)\) with

\[
L_S = p(2 - p)V \text{ and } R_S = V.
\]

(b) The same loan terms \((L_S, R_S)\) would be offered when the project productivities are unobservable to the lender.

**Proof:** see Appendix A.

The reason for the result in (b) is that the loan terms \((L_S, R_S)\) are determined by the break-even and no-default conditions, neither of which depends on the project productivities \(k_i\) and \(k_j\). This also implies that our assumption of both borrowers receiving the same loan size even though they may have different \(k\)'s is not restrictive.

### 3.3 Bogus groups

A standard assumption in the theoretical 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, motivated by the evidence reviewed in the introduction, suppose that lenders are unable to enforce or verify the requirement that each group member invest in her own project, hence bogus groups may exist.

In a bogus group all loaned funds \(2L\) are invested into a single project run by one of the members. The other borrower is a cosigner in the eyes of the lender (joint liability formally applies) but, because of limited liability, this co-signing borrower has no income or wealth for the lender to
go after in case of declared default. As with the repayment decision, we assume that the borrowers form a bogus group if and only if this is jointly beneficial for them. A possible interpretation is the presence of social capital that allows within-group enforcement.

For any given loan terms \((L, R)\) a bogus group invests all funds \(2L\) into the higher-productivity \((k_i)\) project. As in Section 3.2, it is not optimal to repay partially. Conditional on project success, the group’s joint payoff from repaying is larger than the payoff from default if \(k_i(2L) - 2R + 2V \geq k_i(2L)\) or, equivalently

\[
R \leq V
\]

which is the same no strategic default condition as (2) for standard groups.

The above implies that, for given loan terms \((L, R)\) satisfying the no-default condition \(R \leq V\), the expected payoff of a bogus group with project productivities \(k_i, k_j\) is:

\[
W_{ij}^B(L, R) = 2pk_iL - 2pR + 2pV. \quad \text{(BEP)}
\]

The first term in (BEP) is the expected output in a bogus group, the second term is the expected repayment and the third term, \(2pV\) is the expected continuation value of access to future credit. Note that the expected continuation value differs from its counterpart in a standard group, \(2p(2 - p)V\) in (SEP).

For \(R \leq V\) the lender is repaid \(2R\) with probability \(p\) and zero otherwise. Hence, the lender’s break-even condition for lending \(2L\) to a bogus group is \(pR \geq L\) and thus free entry implies \(L = pR\). Combined with the no-default condition \(R \leq V\), the payoff-maximizing loan terms for a bogus group (as if known as such by the lender and taken in isolation) are therefore

\[
L_B = pV \quad \text{and} \quad R_B = V. \quad \text{(4)}
\]

### 3.3.1 Standard vs. bogus groups – comparison

There are five main differences between bogus and standard groups. First, for given loan terms \((L, R)\), we have:

(i) **risk-pooling / diversification**: in a standard group a borrower receives the continuation value \(V\) when her own project fails but her partner’s project succeeds. The i.i.d. project returns assumption is important for this, as the joint liability clause ensures that in such scenario the group is not cut off from future credit. Thus, the expected continuation value per member is \(p(2 - p)V\) in a standard group vs. \(pV\) in a bogus group.

(ii) **expected repayment**: the 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. For given loan terms \((L, R)\) the difference between the increased continuation value in item (i) and the larger expected repayment (item ii) is non-negative for \(R \leq V\) (see also the discussion below).

(iii) **expected output** is larger in a bogus group with heterogeneous borrowers \((2pk_HL)\) than in a standard group with the same borrowers, \(p(k_H + k_L)L\).
We characterize the interplay of items (i), (ii) and (iii) for given \((L,R)\) in the Lemma below.

**Lemma 1:** For given loan terms \((L,R)\) satisfying the no-default condition \(R \leq V\), borrowers with projects productivities \(k_i\) and \(k_j\) with \(k_i \geq k_j\) prefer to form a bogus group if and only if

\[
p(k_i - \frac{k_i + k_j}{2})L > [1 - (1 - p)^2 - p](V - R)
\]

and prefer to form a standard group otherwise.

**Proof:** see Appendix A

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

(iv) *the interest rate* \(\left(\frac{R}{L}, \text{gross}\right)\) at which the lender breaks even is strictly lower in a standard group, \(\frac{1}{p(2-p)}\), than in a bogus group, \(\frac{1}{p}\). The reason, as in item (i), is that the lender is repaid with probability \(1 - (1 - p)^2\) in a standard group in which two i.i.d. projects are funded vs. repaid with probability \(p\) in a bogus group.

(v) *the loan size* can be larger in standard groups due to the lower interest rate and more likely repayment. Specifically, the break-even and no-default conditions imply \(L \leq p(2-p)V\) in standard groups vs. \(L \leq pV\) in bogus groups.

**Lemma 1** implies that, all else equal, larger loan size \(L\), or larger repayment amount \(R\) or larger productivity differential \(k_i - k_j\) make bogus groups more preferred. The left hand side of (5) is the gain in expected output per member from forming a bogus group in comparison to a standard group, that is, the expected gain from investing \(2L\) at the high return \(k_i\) instead of investing \(L\) at return \(k_i\) and \(L\) at return \(k_j\). This corresponds to item (iii). Conversely, the right hand side of (5) is the net gain per member from forming standard group instead of bogus group thereby raising the probability of repaying and obtaining the continuation value \(V\) from \(p\) to \(1 - (1 - p)^2\). This corresponds to items (i) and (ii) in the list.

The right hand side of (5) is always non-negative since \(R \leq V\) by the no-default condition \(10\) and the left hand side is also non-negative since \(k_i \geq k_j\) by assumption. For homogeneous \((HH\) or \(LL)\) borrower pairs, the left hand side is zero – forming a bogus group does not offer any benefit in additional project return while it foregoes the diversification benefit of a standard group. Note also that when \(R = V\), which is the maximum repayment amount satisfying the no-default condition \(2\), homogeneous pairs (with \(k_i = k_j\)) are indifferent between the two group forms. Hereafter, we assume that when indifferent borrowers choose to form a standard group. This can be justified by adding a small exogenous cost (e.g., detection risk) of operating as bogus group.

To sum up, item (i) favors standard groups due to the diversification effect from investing in two different projects, which is also the reason for the lower standard-group interest rate (item iv). In addition, larger loan size (item v) is feasible in standard groups. Conversely, items (ii) and (iii) favor bogus groups, by allowing the borrowers to benefit from investing a larger amount in the high-return project and from a smaller expected repayment.
3.3.2 Bogus groups and lender’s loss

We now use Lemma 1 to show that a lender would lose money if (s)he ignored the possibility of bogus groups.

**Proposition 2:** If all borrowers are offered loan terms \((L_S, R_S) = (p(2-p)V, V)\) then,

(a) all heterogeneous (HL) borrower pairs form bogus groups

(b) all homogeneous (HH or LL) borrower pairs are indifferent between forming bogus or standard group

(c) bogus groups cause a loss to the lender.

**Proof:** see Appendix A.

Intuitively, remember that the right hand side of condition (5) in Lemma 1 corresponds to the net benefit from forming a standard vs. bogus group for given loan terms \((L, R)\). At the given loan terms \((L_S, R_S)\) with \(R_S = V\), this benefit is zero. Thus, only the increased output effect (item iii) remains and hence all heterogeneous pairs prefer to form bogus groups while homogeneous pairs are indifferent. All HL borrower pairs benefit from the larger expected output in a bogus group but repay less often than required for the lender to break even at the gross interest rate \(\frac{R_S}{L_S}\). Hence, the lender loses money on any such loans. The implication is that endogenous bogus groups formation must be addressed by the lender by designing appropriate loan terms.

4 Endogenous group form

We now analyze the decision problem of lenders who design loan terms taking into account the borrowers’ hidden action to form bogus or standard group. We first study the simpler setting in which the borrowers’ project productivities \(k_i\) and \(k_j\) are observable by the lender. We interpret this as the lender knowing or monitoring the intended loan use: agriculture, retail, etc. In Section 4.2 we relax this assumption and study an alternative setting in which lenders do not observe the productivities \(k_i\) and \(k_j\).

4.1 Loan terms

By the free-entry assumption, facing a group of type \(ij\) the lender would offer the loan terms that maximize the group payoff subject to the break-even constraint. When the project returns \(k_i\) and \(k_j\) are observable, the lender rationally expects the optimal group form choice (standard or bogus) for any \((L, R)\), and thus offers loan terms inducing the payoff-maximizing group form. The lender’s zero-profit condition and no-strategic-default constraint depend on the induced group form. Below we suppress the indexation by \(ij\) for notational simplicity, e.g., we write \(W^S(L, R)\) instead of \(W^S_{ij}(L, R)\), etc.

Given \(k_i, k_j\) and \(p\) the group loan design problem can be written as:
Problem OP

\[
\max_{L,R,\tau \in (0,1)} \tau W^S(L, R) + (1 - \tau)W^B(L, R)
\]

s.t. \( \tau W^S(L, R) + (1 - \tau)W^B(L, R) \geq \tau W^B(L, R) + (1 - \tau)W^S(L, R) \) (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 expected payoffs of a standard group, \( W^S(L, R) \), and a bogus group, \( W^B(L, R) \), are as defined in (SEP) and (BEP) above, omitting the subscript \( ij \).

We use the indicator \( \tau \in \{0, 1\} \) to write in a compact way the choice of the larger of the two payoffs, \( W^S(L, R) \) and \( W^B(L, R) \), corresponding to the two group forms, standard and bogus. Constraint (IC) is an incentive constraint which requires that, if the lender chooses to offer loan terms \((L, R)\) inducing a standard group \((\tau = 1)\), then the borrowers must indeed prefer to form a standard group, i.e., \( W^S(L, R) \geq W^B(L, R) \). Similarly, the opposite must hold if inducing a bogus group \((\tau = 0)\) is payoff-maximizing. The no-default and zero-profit constraints were derived in Section 3. The next proposition characterizes the loan terms and group form for any possible values of the parameters \( k_L, k_H \) and \( p \).

**Proposition 3:** Suppose the project productivities \( k_i \) and \( k_j \) are observable. The loan terms solving Problem OP for any group type \( ij \in \{HH, LL, HL\} \) and any \( p, k_L, k_H \) satisfying Assumption 1 are:

<table>
<thead>
<tr>
<th>group type, ( ij )</th>
<th>loan terms</th>
<th>group form</th>
<th>for</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. homogeneous, LL or HH ( k_i = k_j \in {k_L, k_H} )</td>
<td>((L_S, R_S))</td>
<td>standard</td>
<td>any ( k_L, k_H, p )</td>
</tr>
<tr>
<td>2. heterogeneous, HL ( k_i = k_H, k_j = k_L )</td>
<td>((L_E, R_E)), ((L_B, R_B))</td>
<td>standard, bogus</td>
<td>( k_H \in [k_L, \bar{k}] ), ( k_H &gt; \bar{k} )</td>
</tr>
</tbody>
</table>

where \( L_E \equiv \frac{p(2-p)\tilde{V}}{1 + \frac{2-p}{2-p}(k_H - k_L)} \), \( R_E = \frac{L_E}{p(2-p)} \) and \( \bar{k} \) is the larger root of \( W^S(L_E, R_E) = W^B(L_B, R_B) \) written as quadratic equation in \( k_H \).\(^{15}\)

**Proof:** see Appendix A.

Intuitively, homogeneous borrower pairs \((LL \text{ or } HH)\) always prefer a standard group, since they benefit from the associated reduced risk of default (diversification) and the larger loan size. The benefit from the lower interest rate \( \left( \frac{1}{p(2-p)} < \frac{1}{p} \right) \) is offset by the larger expected repayment. For homogeneous borrowers there is also no increase in expected output from forming a bogus group. Hence, homogeneous borrower pairs have no incentive problem and always choose a standard group and receive the maximum feasible (subject to no default) loan size \( L_S \).

\(^{15}\)The threshold \( \bar{k} \) depends on the values of \( k_L \) and \( p \) – see the proposition proof for details.
In contrast, heterogeneous borrowers face a trade-off between the larger expected output that can be obtained in a bogus group vs. the larger expected payoff from risk diversification in a standard group. Therefore, when the productivity differential is sufficiently large, for \( k_H > \hat{k} \) which is equivalent to \( W^B(L_B, R_B) > W^S(L_E, R_E) \), it is optimal to induce a bogus group and offer the largest feasible loan size, \( L_B \). Note also that, since \( \hat{k} > k_L \) as shown in the proof of Proposition 3, the condition \( k_H > \hat{k} \) is more stringent than the condition \( k_H > k_L \) for forming a bogus group in Proposition 2 – the lender optimally responds to the possibility of bogus groups by raising the interest rate, which in turn reduces the bogus group’s expected payoff.

Conversely, when the project productivity differential within a group is relatively small, for \( k_H \in [k_L, \hat{k}] \), which is equivalent to \( W^S(L_E, R_E) = W^B(L_B, R_B) \), the risk diversification benefit from inducing a standard group outweighs the loss of larger potential output. However, a smaller loan size, \( L_E < L_S \) has to be offered, to deter the heterogeneous borrowers from forming a bogus group.

Comparing the three possible equilibrium loan sizes, it is easy to show that \( L_S > L_B \) and \( L_S > L_E \). That is, the model predicts that borrowers in homogeneous groups (\( HH \) and \( LL \)) receive larger loans than borrowers in heterogeneous groups (\( HL \)). The reason is that the former borrowers have no benefit from forming a bogus group and so default less often.

Comparing the loan terms for standard vs. bogus groups, we have \( L_S > L_B \), however, either of \( L_B \) and \( L_E \) can be larger than the other, depending on the parameter values. Thus, from observing group form alone, without knowing the within-group project quality composition, we cannot draw a definite conclusion about the relationship between loan size and group form. The repayments satisfy \( R_S = R_B > R_E \); that is, they are either the same across bogus and standard groups or larger in bogus groups.

Finally, the above results show that efficiency is enhanced and lender losses avoided when the endogenous group form (bogus or standard) is taken into account by the lender and the appropriate loan terms are designed. The conclusion is that Lei Da Hu or bogus groups should not be viewed as money-losing illicit practice but instead as an opportunity that MFIs can incorporate in their lending strategies.

### 4.2 Unobserved productivity

#### 4.2.1 Preliminaries

Suppose now that the borrowers’ project productivities \( k_i, k_j \in \{L, H\} \) are unobserved by the lender. The lender knows the values \( k_L \) and \( k_H \), that is, the returns of a high- and a low-productivity project. Assume also that the lender knows what fraction (mass) of agents, \( q_{ij} \) belong to \( HH \), \( LL \) and \( HL \) borrower pairs.\(^{16}\) As we show below, our main results and intuition from Section 4.1 in which productivity was assumed observable generalize with small modifications.

Note first that the loan terms in Proposition 3 are no longer feasible, since the lender cannot

\(^{16}\)For example, if project productivities are i.i.d. and \( k_L \) and \( k_H \) are drawn with equal probability after the group is formed, the mass of \( HH \) and \( LL \) groups would be 1/4 while the mass of \( HL \) groups would be 1/2.
observe the group members’ productivities $k_i$ and $k_j$ and hence cannot offer terms contingent on the group composition $ij$. Second, when $k_i, k_j$ are unobserved by the lender, the borrowers have incentive to report the $ij$ value which maximizes their joint payoff, which may differ from the true group composition. For example, the borrowers in an $HL$ pair that would be offered loan terms $(L_E, R_E)$ in Proposition 3 can misreport the group type as $HH$, receive loan $(L_S, R_S)$ and form a bogus group, which would cause a loss to the lender.

Unlike in Proposition 3 where the lender customizes the loan terms for each observed borrower pair $(k_i, k_j)$, when the $k$’s are unobservable the lender must design a loan menu such that no borrower pair has incentive to misreport its type ($LL$, $HH$ or $HL$) and also, after choosing a loan from the menu, no pair has incentive to choose group form (bogus or standard) different from the form that the lender wants to induce. This is a mechanism design problem which is is more complex than a standard screening (adverse selection) problem since the lender faces both unobserved types (the $k_i, k_j$) and an unobserved action (the ex-post moral hazard in group form choice).

### 4.2.2 The loan menu

First, note that, because of free entry, the lender’s break-even condition for each possible group form implies that the gross interest rate $\frac{L}{L}$ must equal either $\frac{1}{p^{(2-p)}}$ (if a standard group is being induced) or $\frac{1}{p}$ (if bogus group is being induced). Any lower interest rate would result in a loss to the lender while any higher interest rate would be competed away.

Second, remember that for any loan terms, either $(L, \frac{L}{p})$ or $(L, \frac{L}{p^{(2-p)}})$, the group payoff is strictly increasing in the loan size $L$, regardless of the borrowers’ productivities. This implies that, within each of the two possible loan terms defined by the interest rate, $(L, \frac{L}{p})$ or $(L, \frac{L}{p^{(2-p)}})$, only a unique 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 is no way for the lender to screen across the $ij$ borrower pairs by varying the loan size, only by the interest rate. If two distinct loan sizes were offered both carrying the same interest rate, then any borrowers would prefer 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 loan terms, $\mathcal{N} \equiv (L_N, R_N)$ and $\mathcal{M} \equiv (L_M, R_M)$, inducing respectively a standard group and a bogus group. The terms $\mathcal{N}$ and $\mathcal{M}$ are chosen such that:

(a) each group $ij$ which selects loan $\mathcal{N}$ chooses to be standard and each group $ij$ which selects $\mathcal{M}$ chooses to be bogus;

(b) no borrower pair defaults strategically;

(c) lenders break even within each loan type, $\mathcal{N}$ or $\mathcal{M}$ (no cross-subsidization because of free entry);

(d) the total borrowers’ payoff is maximized.

The contracting problem with unobservable productivities is:
Problem UP

\[
\max_{L_N, R_N, L_M, R_M} \sum_{ij} q_{ij} W_{ij}(L_N, R_N, L_M, R_M)
\]

subject to

\[
R_M \leq V \tag{6}
\]
\[
R_M = \frac{L_M}{p} \tag{7}
\]
\[
R_N \leq V \tag{8}
\]
\[
R_N = \frac{L_N}{p(2-p)} \tag{9}
\]

\[
W_{ij}(L_N, R_N, L_M, R_M) \geq \max \{W_{ij}^B(L_N, R_N), W_{ij}^S(L_M, R_M)\}, \forall ij \in \{HH, HL, LL\} \tag{IC}
\]

where \(W_{ij}(L_N, R_N, L_M, R_M) \equiv \max \{W_{ij}^S(L_N, R_N), W_{ij}^B(L_M, R_M)\}\) and the group payoffs \(W_{ij}^S(L, R)\) and \(W_{ij}^B(L, R)\) are defined in (SEP) and (BEP).

The first four constraints are the no-default and break-even constraints for each loan in the menu. The last constraint ensures that any group \(ij\) would choose its intended contract (\(N\) or \(M\)) and group form (standard or bogus) that maximizes its payoff. Selecting the alternative contract or deviating to the alternative group form, or both, is not optimal.

**Lemma 2:** Constraint (6) binds at the solution to Problem UP, that is, \(R_M = V\) and \(L_M = pV\).

**Proof:** see Appendix A.

Lemma 2 shows that loan \(M\), designed for bogus groups, always has the maximum feasible loan size that precludes strategic default and hence \(M = (L_B, R_B)\), the same bogus group loan terms as in Section 3.3 (observable productivities). The intuition is that, for any given \((L, R)\), the marginal payoff increase from raising the loan size is always larger for a bogus group compared to in a standard group; hence the incentive constraint (IC) is (weakly) relaxed by increasing the loan size \(L_M\) to its maximum possible value.

**Proposition 4:** Suppose the project productivities \(k_i\) and \(k_j\) are unobserved by the lender. Let \(E \equiv (L_E, R_E)\), \(B \equiv (L_B, R_B)\) and \(\bar{k}\) be as defined in Proposition 3. Define also \(L_F = \frac{p k_H - 1}{p - 1} p V\), \(R_F = \frac{L_F}{p(2-p)}\) and \(F \equiv (L_F, R_F)\). Then, for any \(k_L, k_H\) and \(p\) satisfying Assumption 1, the loan menu and the chosen loan terms and group form solving Problem UP are:

<table>
<thead>
<tr>
<th>parameter condition: (M, N = LL \text{ and } HH) pairs</th>
<th>chosen loan terms and group form</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (k_H \in [k_L, \bar{k}])</td>
<td>(E, B)</td>
</tr>
<tr>
<td>2. (k_H &gt; \bar{k})</td>
<td>(F, B)</td>
</tr>
</tbody>
</table>
Proof: See Appendix A.

As proved in Lemma 2, the loan terms $M$ inducing a bogus group feature the maximum feasible loan size $L_B$. In contrast, the loan size $L_N$ designed to induce a standard group is reduced from the maximum feasible no-default size $L_S$ to either $L_E$ or $L_F$. This is necessary to prevent the borrowers from selecting loan $N$, benefitting from its low interest rate $\frac{1}{p(2-p)}$ and choosing to form a bogus group. Comparing the loan terms across standard and bogus groups, the model predicts that the repayment $R_B$ should be strictly higher in a bogus groups than in a standard group, $R_F$ or $R_E$ (since $R_B = R_S = V$ and $R_S > R_E, R_S > R_F$). On the other hand, the relative magnitudes of the loan sizes $L_B$ vs. $L_E$ or $L_F$ depend on the model parameters.

The asymmetric information about the project productivities naturally reduces efficiency relative to the observable productivities case (Section 4.1) – comparing with Proposition 3, the reduction in the loan size for standard groups (from $L_S$ to $L_E$ or $L_F$) now holds for all possible borrower pairs (even though the IC constraint is slack for homogeneous pairs), since the loan terms $N$ cannot depend on the unobserved group composition. For any given parameters, the offered loan size is (weakly) smaller than the loan size in the observable productivity setting.

As in Proposition 3, bogus groups endogenously emerge when the relative heterogeneity in project productivity is relatively high ($k_H > \bar{k}$), while standard groups are formed otherwise. The main intuition and trade-off between the diversification benefit in standard groups vs. the expected output gain in bogus group discussed earlier still hold when project productivity is unobserved by the lender. As in Section 4.1, when the endogenous formation of bogus groups is taken into account by the lender and the borrowers are offered the optimal loan menu $M, N$ to choose from, efficiency is still enhanced relative to Proposition 2 and lender’s losses avoided.

5 Extensions

5.1 $n$-member groups

Our analysis can be extended to $n$-member groups where $n > 2$. As in Section 3, it is easy to show that with joint repayment decisions both bogus and standard groups repay when $R \leq V$. Denote the repayment probability of a standard group by $p_S$ where $p_S \geq p$. The exact value of $p_S$ depends on the feasibility constraint, that is, how much extra liability a member is able to cover when her partner(s) fail(s). For example, suppose that a borrower can only cover up to $mR$ in total liability when other members fail, where $m \leq n$. In our baseline model we assumed $m = n = 2$, and so $p_S = 1 - (1 - p)^2 = p(2 - p)$. In general, the expected joint payoffs in a standard and a bogus groups are:

$WS(L, R) = p \sum_{i=1}^{n} k_i L + np_S(V - R)$ and $WB(L, R) = npk_{max} L + np(V - R)$,
where $k_{\text{max}}$ is the largest productivity value in the group and superscripts $S$ and $B$ denote standard and bogus group as before. For given loan terms $(L, R)$ a group would choose to be bogus if

$$npk_{\text{max}}L + np(V - R) \geq p \sum_{i} k_i L + np_S(V - R)$$

$$\Leftrightarrow p(k_{\text{max}} - \frac{1}{n} \sum_{i=1}^{n} k_i) L \geq (p_S - p)(V - R)$$

The left hand side reflects the expected output gain from forming a bogus group and raising the average project productivity from $\frac{1}{n} \sum_{i=1}^{n} k_i$ to $k_{\text{max}}$. The right hand side corresponds to the net gain from forming a standard group and raising the probability of repayment and obtaining the future value $V$ from $p$ to $p_S$. Note that for $n = 2$ and $p_S = 1 - (1 - p)^2$ we obtain Lemma 1. The rest of our previous analysis can be extended for $n > 2$ in a similar way, with the complication of additional parametric cases and algebra – full details are available upon request.

### 5.2 Individual repayment decisions

We next explore an alternative model setting in which the group members make the decision to repay or default individually (non-cooperatively) instead of jointly (as assumed in Sections 3 and 4). Similarly to Besley and Coate (1995), we then model the borrowers’ repayment decisions as a two-stage game. In the first stage, each borrower is asked by the lender to repay $R$ and can either do that or report default. If both borrowers repay $R$ or if both default in the first stage, their respective payoffs (described below) are realized 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 that case, the borrower who repaid is asked by the lender to pay additional $R$ for her partner, as stipulated by the joint liability clause. Again, the borrower chooses to repay or default. We solve for the subgame-perfect Nash equilibrium of the described game.

Start with stage 2 – the decision facing a borrower with successful project who has repaid $R$ in stage 1. It is never optimal to make a partial repayment (strictly between 0 or $R$) since either defaulting (repaying zero) and forfeiting the continuation value $V$, or repaying in full and securing the value $V$ is dominant strategy (this assumes that the lender considers incomplete repayment a default). The borrower thus repays in stage 2 if her payoff, $-R + V$ is larger than the payoff of defaulting, 0 or:

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

Suppose (S2R) holds, so that each 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 for the row player $i$ are listed; the payoffs for the column player $j$ are symmetric across the diagonal. If a borrower’s project fails

\[\text{For example, each borrower may only be able to verify her own project outcome, as in Armendariz de Aghion (1999), or within-group sanctions cannot be imposed.}\]
she announces default. As before, partial repayment is dominated by either defaulting or repaying in full.

<table>
<thead>
<tr>
<th>Table 1: Individual repayment – stage 1 game</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repay</td>
</tr>
<tr>
<td>Default</td>
</tr>
<tr>
<td>Repay</td>
</tr>
<tr>
<td>$k_i L - R - (1 - p) R + V$</td>
</tr>
<tr>
<td>$k_i L - 2 R + V$</td>
</tr>
<tr>
<td>Default</td>
</tr>
<tr>
<td>$k_i L + p V$</td>
</tr>
<tr>
<td>$k_i L$</td>
</tr>
</tbody>
</table>

The repayment payoff in Table 1 reflects the fact that, assuming (S2R) holds, a borrower playing Repay in stage 1 while her partner also plays Repay, would only be asked to repay extra $R$ in stage 2 with probability $1 - p$ (if her partner’s project fails).\(^\text{18}\) Conditional on $j$ playing Repay in stage 1, borrower $i$ would also choose Repay when her project succeeds if her expected payoff from repaying $R$ is larger than her payoff from strategically defaulting (repaying zero), that is, if

$$k_i L - R - (1 - p) R + V \geq k_i L + p V,$$

or

$$R \leq \frac{1 - p}{2 - p} V.$$  \(\text{(10)}\)

Similarly, conditional on borrower $j$ choosing Default, borrower $i$ would choose Repay if paying back $2R$ over the two stages and securing the continuation value $V$ results in a larger payoff than declaring default, that is, if

$$k_i L - 2 R + V \geq k_i L,$$

or

$$R \leq \frac{1}{2} V.$$  \(\text{(11)}\)

Since $\frac{1 - p}{2 - p} < \frac{1}{2}$, condition (10) implies conditions (11) and (S2R). Thus, for $R \leq \frac{1 - p}{2 - p} V$ (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.\(^\text{19}\) Comparing constraint (10) with constraint (2) in Section 3 (the case of joint repayment decision) we see that assuming non-cooperative repayment decisions restricts the maximum feasible repayment (and hence the loan size) for which no strategic default can be supported. The intuition is that without coordination each borrower can free-ride on the repayment of the other group member, which increases the incentive for strategic default compared to the joint-decision setting in Sections 3 and 4.

The analysis in Sections 3 and 4 can be re-done in a relatively straightforward way for the model with individual repayment decisions. We discuss only the main results here and refer the reader to Appendix B.2 for all details and formal proofs.

We first show that, with standard groups only (as in Proposition 1), individual repayment decisions imply strictly smaller loan size, $L_{S'} < L_S$ (and, correspondingly, smaller repayment $R_{S'}$).

\(^\text{18}\)Here, consistent with the non-cooperative assumption, we assume the borrower makes her stage 1 repayment decision without observing her 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 $R \leq \frac{(1-p)V}{2}$.

\(^\text{19}\)See Appendix B.1 for further discussion.
and group payoff). The reason is the additional incentive for strategic default from being able to free ride on the repayment of the other member. Specifically, \( L^* = p(1 - p)V \) and \( R^* = \frac{1 - 2p}{1 - p}V \) – see Proposition B-1 in Appendix B. The interest rate is the same as in Proposition 1 since it is determined by the break-even (free entry) condition.

We then compare standard and bogus groups and show that the basic intuition from Section 3 still holds. Standard groups offer a diversification benefit and lower interest rate while expected output is larger in bogus groups. However, there is an additional dimension to the comparison, since the incentive for strategic default in a standard group is now stronger because of the free-riding possibility – see the discussion surrounding condition (10). In a bogus group this free-riding effect is absent, since the inactive partner has no income and cannot repay.\(^{20}\) The additional free-riding effect is present in both heterogeneous and homogeneous pairs and results in an additional benefit from forming a bogus group.

As in Proposition 2, when bogus groups can form endogenously, offering standard loan terms to all groups can cause a loss to the lender (see Proposition B-2 in Appendix B). This happens for borrowers with sufficiently heterogeneous projects.

We also characterize the constrained-optimal loan terms when the lender takes into account the endogenous choice of group form (see Proposition B-3 in Appendix B, for the observable \( k \)’s setting). Compared to Proposition 3, there exist more cases dependent on the parameter values for \( p, k_L \) and \( k_H \). The basic intuition – that more heterogeneity in the productivities \( k_i \) and \( k_j \) makes bogus groups more likely – still holds. What is different from the joint-decision setting, however, is that with individual repayment decisions bogus groups may also be formed by homogeneous pairs. This happens either when the success probability \( p \) or when the project productivities are sufficiently high. Intuitively, the gains from forming a bogus group are the largest when the risk of failure is relatively low (large \( p \)) and/or when project productivity is large, so that the benefits from the reduced free riding incentive and larger loan size in a bogus group outweigh the loss of risk diversification. Standard groups are optimal otherwise (see Proposition B-3 in Appendix B).

For heterogeneous borrower pairs, either offering loan terms inducing bogus group is optimal, which happens when the productivity differential or level is sufficiently high enough; or loan terms inducing standard group are optimal, when the project productivities are relatively low and close to each other. As in Proposition 3, the loan terms inducing a bogus group do not depend on the values \( k_i, k_j \). The only loan terms that depend on the productivity values \( k_i, k_j \) are \( E \) since, to deter an \( HL \) group from switching to bogus, it is necessary to take into account the output gain from changing the group form, which depends on \( k_H \) and \( k_L \). Finally, as in Section 4, borrower welfare and productive efficiency are improved when the endogenous formation of bogus groups is incorporated in the lender’s strategy and the borrowers are offered appropriate (choice of) loan terms.

\(^{20}\)Note that deciding to strategically default in a bogus group does not require additional coordination compared to in a standard group. In a standard group the joint liability (JL) clause means that if member 1 decides to strategically default, the lender can go to member 2 and collect \( 2R \) if 2 declares success. In a bogus group JL implies that if 1 decides to default, the lender can still go to member 2 but cannot collect anything due to limited liability.
5.3 Individual loans

So far, as motivated by the CFPAM data, we only considered group joint-liability loans and studied the endogenous choice of borrowers to form bogus or standard groups. To clarify further this paper’s contribution to the literature we also explore allowing individual loans in our setting. We show that when payoffs are transferable within a group then any borrower pair has no incentive to switch to individual loans.

Suppose a borrower with productivity $k_i \in \{k_{L}, k_{H}\}$ is given an individual loan $L$ which requires gross repayment $R$. The borrower’s payoff upon strategic default is $k_i L$, the payoff upon repayment is $k_i L - R + V$ and hence the no strategic default constraint is $R \leq V$ which is the same as the condition for group loans derived in Section 3.

The lender’s break-even condition is $L = pR$, which is the same as that in a bogus group. The borrower’s expected payoff given $R \leq V$ is

$$W^I(L, R) = pk_i L - pR + pV$$

The above results imply that the borrower’s payoff maximizing individual loan has terms

$$L_I = pV \quad \text{and} \quad R_I = V,$$

which are identical to the bogus group loan terms $(L_B, R_B)$ from Section 3. Feasibility of repayment requires $pk_i \geq 1$ which is satisfied by Assumption 1. Consistent with the previous analysis, assume transferable utility within the group, so that the borrowers are concerned with maximizing their joint expected payoff.

Take a pair of borrowers with productivities $k_i, k_j$ where $k_i = \max\{k_i, k_j\}$. If both receive individual loans with terms $(L_I, R_I) = (pV, V)$, then the pair’s total payoff is

$$\pi^I(k_i, k_j) = p^2 (k_i + k_j)V,$$

which equals $2p^2 k_i V$ for a homogeneous pair $(k_i, k_i)$. In contrast, by Proposition 3, a homogeneous borrower pair $(k_i, k_i)$ borrowing at the standard group loan terms $(L_S, R_S)$ has joint payoff,

$$\pi^S(k_i, k_i) = 2p^2 k_i (2 - p)V,$$

which is clearly larger than the pair’s total payoff with individual loans, $2p^2 k_i V$. The intuition is that joint liability allows risk diversification (members covering for each other), which reduces the risk of default (and losing the continuation value $V$) and allows for lower interest rate and larger loan size, while keeping the total repayment the same as in an individual loan. Hence, homogeneous borrower pairs have no incentive to use individual loans.

Consider now a heterogeneous borrower pair $(k_i, k_j)$, with $k_i > k_j$. If the pair forms a bogus
group, its total payoff at loan terms \((L_B, R_B)\) is:

\[
\pi^B(k_i, k_j) = 2p^2k_iV
\]

which is strictly larger than \(\pi^I(k_i, k_j)\). Again, the two borrowers have no incentive to use individual loans (transferable utility is important here). Finally, suppose a heterogeneous group faces the loan terms \((L_E, R_E)\) in the Proposition 3 case \(k_H \leq \bar{k}\) and call its joint payoff \(\pi^E(k_i, k_j)\). We then know (see the proof of Proposition 3) that the group’s joint payoff satisfies \(W^S(L_E, R_E) \geq W^B(L_B, R_B)\), which implies \(\pi^E(k_i, k_j) \geq \pi^B(k_i, k_j)\) and since \(\pi^B(k_i, k_j) > \pi^I(k_i, k_j)\) as shown before, the borrowers are once again better off with their group loan \(E\).

In sum, we conclude that, starting at the equilibrium described in Proposition 3, if individual loans were made available at the best possible terms for the borrowers \((L_I, R_I) = (pV, V)\) and payoffs are transferable within the group, any borrower group would have no incentive to switch to individual loans.

Intuitively, in our model group lending dominates individual lending for two reasons. First, in a standard group, group loans allow reduced risk of default and larger loan size. Second, in a bogus group, group lending allows the borrowers to benefit from investing a larger amount into the higher productivity project.

5.4 Preventing bogus groups

Suppose that the lender is aware of the possibility of Lei Da Hu but, for exogenous reasons, would like to design a group loan contract that eliminates the incentive for all borrowers to form bogus groups.\(^\text{21}\) We only discuss the case of observed productivities \(k_i\) and \(k_j\) (the unobservable productivities case is available upon request).

The lender wants to design loan terms \((L^#, R^#)\) such that: (i) borrower pair \(ij\) for any \(ij \in \{LL, HH, HL\}\) forms a standard group; (ii) there is no strategic default; (iii) the lender breaks even; and (iv) the borrowers’ total payoff is maximized.

As in Section 3, the terms \((L^#, R^#)\) must satisfy the feasibility constraint, the break-even constraint and the no-default constraint for standard groups ((2)). In addition, forming a bogus group should not be optimal for any borrowers at \((L^#, R^#)\). By Lemma 1, the latter requires,

\[
(k_i - k_j)L^# \leq 2(1 - p)(V - R^#)
\]

Therefore, the payoff-maximizing loan terms \((L^#, R^#)\) that prevent the formation of bogus groups solve:

\[
\max_{L, R} W^S_{ij}(L, R) = p(k_i + k_j)L - 2p(2 - p)R + 2p(2 - p)V
\]

s.t. (1), (2), (12), and \(R = \frac{L}{p(2-p)}\) (break even)

\(^{21}\)A list of possible reasons was given in footnote 4.
Proposition 5: Suppose the lender wants to prevent bogus groups and project productivities are observable. Let $L_S, R_S, L_E$, and $R_E$ be as defined in Propositions 1 and 3. Then:

(i) the loan terms for homogeneous pairs (HH or LL) are $(L_S, R_S)$

(ii) the loan terms for heterogeneous pairs (HL) are $(L_E, R_E)$.

Proof: see Appendix A.

Proposition 5 shows that the payoff-maximizing standard-group loan $(L_S, R_S)$ from Proposition 1 can be safely offered to homogeneous pairs since these borrowers would not form a bogus group. In contrast, heterogeneous borrower pairs always have to be deterred from forming a bogus group by reducing the loan size to $L_E$ for any values of $p, k_L$ and $k_H$ satisfying Assumption 1.

Comparing with the constrained-optimal loan terms derived in Section 4.1, we see that preventing the formation of bogus groups reduces overall borrower welfare (compare Proposition 5 with Proposition 3). Specifically, the terms $(L^+, R^+)$ fail to capitalize on the larger borrower surplus that can be generated if an HL group is offered loan $(L_B, R_B)$ when the productivity differential between the borrower projects is large, the case $k_H > \bar{k}$.

6 Conclusions

We study group lending by explicitly modeling ‘bogus’ microfinance groups, that is, groups in which one borrower invests all members’ loans into a single project, a practice called Lei Da Hu in China. We model the endogenous formation of bogus groups and their coexistence with ‘standard’ borrower groups, in which each member invests their loan in a separate project.

We highlight two main economic mechanisms which determine the offered loan terms and the endogenously chosen group form. The first mechanism is the risk diversification benefit of a standard group – the probability that the borrowers can repay their loans and obtain the continuation value of future credit is strictly higher compared to in a bogus group. This allows lenders to offer lower interest rate and larger loan size to standard groups. The second mechanism is the larger expected output (return) in a bogus group with heterogeneous investment projects, since a bogus group always invests all loaned funds into the project with the highest productivity among all member’s projects. The trade-off between these two mechanisms underpins our theoretical results. The main take-away is that bogus groups arise when the productivity differential between the group members’ projects is sufficiently large, to benefit from the larger expected output. In contrast, the risk diversification benefit prevails among homogeneous borrowers who form standard groups.

An important conclusion from our analysis is that the practice of Lei Da Hu should not be viewed as an undesirable phenomenon which microlenders must eradicate but, instead, as an optimal response by the borrowers which can increase credit allocation efficiency, provided that the lenders design appropriate loan terms or menus. Bogus groups would only cause losses for lenders which

\[22\text{Remember from the proof of Proposition 3 that } k_H > \bar{k} \text{ is equivalent to } W^B(L_B, R_B) > W^S(L_E, R_E).\]
are unaware or ignore their existence. Related to this, one of the ingredients in our model is the moral hazard friction between the borrowers and the lender regarding the group form, that is, how the members’ loans are used after disbursement. If this friction were to be relaxed, then the inefficiency in loan size (e.g., the case \((L_E, R_E)\) in Proposition 3) would be removed. However, our conclusion that bogus groups can be efficiency-improving in heterogeneous groups (through the increase in output) stands.

To an extent, our setting parallels the decision problem of an investor who chooses whether to invest a large amount into a single (riskier) asset vs. smaller amounts into multiple assets. However, this parallel is incomplete, since we go beyond individual portfolio choice and in addition model the strategic interaction between a lender and a group of borrowers who are jointly liable, and also the strategic interaction within the group of borrowers, in one of the extensions.

A common issue with group lending in its standard form is that it may be difficult for entrepreneurs to find partners with ready-to-go investment projects or business ideas with whom to form a borrowing group. We show that bogus groups offer a possible solution to this problem, by allowing all funds to be invested in a single project while preserving or enabling future access to credit (conditional on project success) to all group members.

A natural question is why don’t MFIs avoid the need to tackle Lei Da Hu by using sequential lending within the group, that is, by waiting for a repayment to be received before the next loan is disbursed. A significant downside of this approach is that it involves having productive borrowers wait and available funds not being utilized which could be very inefficient. In addition, the benefit from default risk diversification through the joint liability clause is also lost in sequential lending when the cosigners have no other source of wealth.

The assumption of risk neutrality matters for our results regarding the trade-off between forming a standard vs. a bogus group. If the borrowers and/or lenders were risk averse, then there would be an additional insurance benefit from forming a standard group and diversifying the risk of project failure.

One final issue which we have not addressed thus far is endogenous matching. The model timing can be modified so that the project productivities are drawn first and then the borrowers sort into groups, rationally expecting the loan terms. In that case, it is easy to show that, if only standard groups can be formed, any equilibrium matching pattern is optimal since two \(HL\) groups achieve the same total surplus as one \(HH\) and one \(LL\) group. With endogenous unobserved group form and observable productivities \(k_i, k_j\) (as in Section 4.1), the lenders would offer the loan terms described in Proposition 3 to whichever groups they end up facing after the matching stage. That is, we simply need to check whether positive (PAM) or negative (NAM) assortative matching pattern is ex-ante optimal. It turns out that NAM arises if the productivity differential \(k_H - k_L\) is large while PAM holds otherwise.\(^{23}\) Allowing for endogenous matching is more complex if the lenders could not observe the project productivities, as in Section 4.2. The reason is that the loan menu described in Proposition 4 cannot be taken as given, since a lender may wish to offer a different

\(^{23}\)Proofs of these statements are available from the authors.
menu if he knew, for example, that all groups which he would face in equilibrium would be type \(HL\) (the reason is that the IC constraint is affected). On the other hand, the equilibrium group composition / matching pattern would depend on the loan menu that the borrowers expect. Thus is a hard problem, potentially with multiple equilibria, that we leave for future research.

**References**


Appendix A – Proofs

Proof of Proposition 1

(a) Substituting \( R = \frac{L}{p(2-p)} \) (shown in the main text) into the no-strategic-default constraint (2) transforms it to

\[ L \leq p(2-p)V \]  \hspace{1cm} (13)
Therefore, problem (SP) is equivalent to

$$\max_L (p(k_i + k_j) - 2)L + 2p(2-p)V \text{ s.t. (13)}$$

The objective, (14) is strictly increasing in $L$, thus (13) binds at optimum. Hence, the loan terms solving problem (SP) are $L_S = p(2-p)V$ and $R_S = \frac{L_S}{p(2-p)} = V$.

(b) Note that the loan terms $(L_S, R_S)$ do not depend on $k_i$ and $k_j$. Hence, if $k_i$ and $k_j$ are unobservable by the lender, Proposition 1(a) still holds. To show this formally, re-write the objective as weighted sum of $ij$’s total expected payoffs, with weights equal to the population shares of HH, LL and HL groups. Since the no-default and break-even constraints do not depend on $k_i$ and $k_j$, the result in part (a) obtains.

**Proof of Lemma 1**

An $ij \in \{LL, LH, HH\}$ borrower pair prefers to form a bogus group instead of a standard group for given loan terms $(L, R)$ if and only if $W^B_{ij}(L, R) > W^S_{ij}(L, R)$ which, using (SEP) and (BEP), is equivalent to condition (5).

**Proof of Proposition 2**

(a)-(b) These results follow directly from Lemma 1. At $R = R_S = V$ the r.h.s. of condition ((5)) is zero. For heterogeneous borrower pairs ($k_i > k_j$) the l.h.s. of condition ((5)) is positive, hence they prefer a bogus group. For homogeneous pairs ($k_i = k_j$), the l.h.s. of ((5)) is zero and so they are indifferent between forming bogus or standard group.

(c) By part (a), all heterogenous pairs prefer a bogus group. The lender’s expected profit from lending to a bogus group at terms $(L_S, R_S)$ is

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

therefore bogus HL groups cause a loss to the lender.

**Proof or Proposition 3**

We start with the following observation:

**Result 1**: For $R \leq V$, $W^S(L, \frac{L}{p}) \leq W^B(L, \frac{L}{p})$ for homogeneous group ($k_i = k_j$) if and only if $L \geq pV$.

Proof: using (SEP) and (BEP), $2pk_iL - 2(2-p)L + 2p(2-p)V \leq 2pk_iL - 2L + 2pV$ is equivalent to $pV \leq L$.

Call $(L^*, R^*)$ the loan terms solving **Problem OP**.

**Homogeneous groups**
Suppose the lender faces a homogeneous borrower pair (LL or HH) and wants to induce a bogus group \((\tau = 0)\). Then \(R = \frac{L}{p}\) by the zero-profit constraint. Result 1 then implies that the incentive constraint (IC) is incompatible with the no-default constraint \((R \leq V, \text{i.e., } L \leq pV)\) unless \(L = pV\) and \(R = V\), in which case the borrowers are indifferent between the two group forms.

Now suppose the lender wants to induce a standard group \((\tau = 1)\). The zero-profit condition implies \(R = \frac{L}{p(2-p)}\), which substituted into the no-default constraint yields \(L \leq p(2-p)V\). Constraint (IC) is satisfied for any \(R\) satisfying the no-default condition \(R \leq V\) (see Lemma 1). The objective \(W^S(L, \frac{L}{p(2-p)})\) equals \((2p^2 - 2p)L + 2p(2-p)V\), which is strictly increasing in \(L\). Consequently, the no-default constraint written in terms of the loan size, \(L \leq p(2-p)V\) must bind at the optimum, implying \(L^* = L_S = p(2-p)V\) and \(R^* = R_S = V\).

To determine whether inducing standard or bogus group is optimal, compare the respective payoffs

\[
W^S(p(2-p)V, V) = 2kp^2(2-p)V \quad \text{and} \quad W^B(pV, V) = 2kp^2V.
\]

Clearly, \(W^S(p(2-p)V, V)\) is larger for any \(k_i > 0\). Hence, setting \((L^*, R^*) = (L_S, R_S)\) which induces a standard group is always optimal for homogeneous borrower pairs.

**Heterogeneous groups**

Suppose first the lender wants to induce a standard group \((\tau = 1)\). As before, the zero-profit constraint implies \(R = \frac{L}{p(2-p)}\), which substituted into the no-default constraint implies \(L \leq p(2-p)V = L_S\). Unlike in the homogeneous case above, constraint (IC) is no longer automatically satisfied for any \(R\) satisfying the no-default condition. To satisfy (IC), i.e., \(W^S(L, R) \geq W^B(L, R)\) we need, using Lemma 1:

\[
(k_H - k_L)L \leq 2(1-p)(V - \frac{L}{p(2-p)}) \quad \text{or,} \quad L \leq \frac{p(2-p)V}{1 + \frac{p(2-p)}{2(1-p)}(k_H - k_L)} \equiv L_E
\]

It is easy to verify that \(L_S \geq L_E\) whenever \(k_H \geq k_L\). Therefore, the loan terms in this case are \(L^* = L_E\) and \(R^* = \frac{L_E}{p(2-p)}\).

Suppose instead the lender wants to induce a bogus group \((\tau = 0)\). The zero-profit constraint implies \(R = \frac{L}{p}\) and the no-default condition is \(R \leq V\). By Lemma 1, constraint (IC) is clearly satisfied for \(R = V\) and \(L = pV\). Therefore, the offered loan terms are \((L^*, R^*) = (L_B, R_B)\).

To determine whether inducing standard or bogus group is optimal, compare the payoff of a standard group with loan terms \((L_E, R_E)\) with that of a bogus group with terms \((L_B, R_B)\). Choosing \(L^* = L_E\) and \(R^* = R_E\) which induces a standard group is optimal if \(W^S(L_E, R_E) \geq W^B(L_B, R_B)\), that is:

\[
(p(k_L + k_H) - 2) \frac{p(2-p)V}{1 + \frac{p(2-p)}{2(1-p)}(k_H - k_L)} + 2(2-p)pV \geq (2pk_H - 2)pV + 2pV \quad \text{(CC3)}
\]

or equivalently, \(f(k_L) \geq k_H\)
where
\[ f(k) \equiv \frac{1}{2} \left( k + c + \sqrt{(k + c)^2 - \frac{4k}{p}} \right) \quad \text{and} \quad c \equiv \frac{2p^2 - 5p + 4}{p(2-p)}. \]

Choosing \( L^* = L_B \) and \( R^* = R_B \) (inducing a bogus group) is optimal otherwise, for \( f(k_L) < k_H. \) \(^{24}\) Call \( \tilde{k} = f(k_L) \) as defined above.

The table in the statement of Proposition 3 summarize the results proven above.

Proof of Lemma 2
Substituting for \( R_N \) and \( R_M \) from the break-even constraints, Problem UP simplifies to:

\[
\max_{L_N,L_M} \sum_{ij} q_{ij} W_{ij} \left( L_N, \frac{L_N}{p(2-p)}, L_M, \frac{L_M}{p} \right) \quad \text{(UP')} \]

subject to

\[
L_M \leq pV \quad \text{(15)}
\]
\[
L_N \leq p(2-p)V \quad \text{(16)}
\]

\[
\max \left\{ W^S_{ij}(L_N, \frac{L_N}{p(2-p)}), W^B_{ij}(L_M, \frac{L_M}{p}) \right\} \geq \max \left\{ W^B_{ij}(L_N, \frac{L_N}{p(2-p)}), W^S_{ij}(L_M, \frac{L_M}{p}) \right\} \quad \text{for all } ij \in \{HH, HL, LL\} \quad \text{(IC)}
\]

where, using (SEP) and (BEP),

\[
W^S_{ij}(L_N, \frac{L_N}{p(2-p)}) = (p(k_i + k_j) - 2)L_N + 2p(2-p)V \quad \text{(17)}
\]
\[
W^B_{ij}(L_N, \frac{L_N}{p(2-p)}) = 2(pk_i - \frac{1}{p})L_N + 2pV \quad \text{(18)}
\]
\[
W^B_{ij}(L_M, \frac{L_M}{p}) = 2(pk_i - 1)L_M + 2pV \quad \text{(19)}
\]
\[
W^S_{ij}(L_M, \frac{L_M}{p}) = (p(k_i + k_j) - 2(2-p))L_M + 2p(2-p)V \quad \text{(20)}
\]

We will use the following auxiliary result to Prove Lemma 2 and Proposition 4 below.

**Lemma A1:** Suppose \( L_M < pV. \) Then, respecting the no-default constraints (15) and (16), one can increase \( L_M \) to \( L_M + \varepsilon \) for some \( \varepsilon > 0 \) while holding \( L_N \) constant, such that the incentive constraint (IC) remains satisfied.

Consider a local increase of \( L_M \) to \( L_M + \varepsilon \) with \( \varepsilon > 0. \) Given constraints (15), (16) and expressions (19) and (20) we have that \( W^S(L_M, \frac{L_M}{p}) \geq W^B(L_M, \frac{L_M}{p}) \) if and only if \( L_M \leq L^*_M \equiv \frac{pV}{1 + \frac{pV}{2(1-p)(k_i-k_j)}. \text{ Note also that } L^*_M \leq pV, \text{ with equality if } k_i = k_j.\)

\(^{24}\) Re-write \( W^S(L_E, R_E) = W^B(L_B, R_B) \) as a quadratic equation in terms of \( k_H. \) It is easy to show that its larger root equals \( f(k_L) \) while its smaller root is strictly smaller than \( k_L. \) Thus, for \( k_H \geq k_L \) inequality (CC3) is satisfied if and only if \( k_H \leq f(k_L). \)
Suppose first, $L_M < L_M^*$. We then have
\[
\frac{dW^B(L_M, \frac{L_M}{p})}{dL_M} = 2(pk_i - 1) > p(k_i + k_j) - 2(2 - p) = \frac{dW^S(L_M, \frac{L_M}{p})}{dL_M}
\]
Therefore, by increasing $L_M$ by a small amount (so that $L_M + \varepsilon < L_M^*$) while holding $L_N$ constant, the l.h.s. of constraint (IC) increases by more than its r.h.s., hence the constraint remains satisfied.

Suppose now $L_M \in [L_M^*, pV]$ (if possible). If the r.h.s. of (IC) equals $W^S(L_M, \frac{L_M}{p})$ the result follows from above. Suppose the r.h.s. of (IC) equals $W^B(L_N, \frac{L_N}{p(2-p)})$ instead. Then, constraint (IC) is equivalent to
\[
\max \left\{ W^S(L_N, \frac{L_N}{p(2-p)}), W^B(L_M, \frac{L_M}{p}) \right\} \geq W^B(L_N, \frac{L_N}{p(2-p)})
\]
(21)
The payoff $W^B(L_M, \frac{L_M}{p})$ is increasing in $L_M$. Hence, by increasing $L_M$ while holding $L_N$ constant, the l.h.s. of (21) weakly increases while its r.h.s. remains constant. That is, both the condition $L_M \geq L_M^*$ and constraint (21) still hold, which implies that (IC) remains satisfied. □

Lemma A1 shows that, for any $L_M < pV$, we can increase $L_M$ locally while holding $L_N$ constant, so that the incentive constraint (IC) still holds. Since $W^B(L_M, \frac{L_M}{p})$ is strictly increasing in $L_M$, the objective function in problem (UP') weakly increases in $L_M$. Therefore, setting $L_M = pV = L_B$ is optimal and hence the loan terms $M$ designed to induce bogus groups and solving problem (UP') are $(L_B, R_B)$ for any parameter values.

**Proof of Proposition 4**

We first establish several preliminary results that allow simplifying the IC constraint, (IC) in Problem (UP') defined in the proof of Lemma 2.

**A. Preliminaries**

Using $L_M = pV$ from Lemma 2, the expressions (19) and (20) imply
\[
W^B_i(M) \equiv W^B_i(L_M, \frac{L_M}{p}) = 2k_ip^2V \geq
\]
\[
\geq (k_i + k_j)p^2V = W^S_{ij}(L_M, \frac{L_M}{p}) = W^S_{ij}(M).
\]

Using (17) and (18),
\[
W^S_{ij}(N) \geq W^B_{ij}(N) \iff L_N \leq L_1(ij) \equiv \frac{p(2-p)V}{1 + \frac{p(2-p)}{2(1-p)}(k_i - k_j)},
\]
(22)
and, by Lemma 2,
\[
W^B_{ij}(M) \geq W^B_{ij}(N) \iff L_N \leq L_2(i) \equiv \frac{pk_i - 1}{2 - p} - pV
\]
(23)
where $k_i \geq k_j$ as assumed earlier.

Any given borrower pair $ij \in \{LL, HH, HL\}$ chooses the loan terms that result in a higher
joint payoff. The pair would choose loan $\mathcal{M}$ if

$$W_{ij}^S(\mathcal{N}) < W_{ij}^B(\mathcal{M}) \Leftrightarrow L_N < L_3(ij) \equiv \frac{2pk_i - 4 + 2p}{p(k_i + k_j)} - \frac{1}{2} p V$$

(CHM)

and $\mathcal{N}$ otherwise.

If loan terms $\mathcal{M}$ are payoff-maximizing, i.e., $W_{ij}^B(\mathcal{M}) > W_{ij}^S(\mathcal{N})$ for pair $ij$ (that is, $L_N < L_3(ij)$ by (CHM)) then the l.h.s. of the IC constraint (IC) equals $W_{ij}^B(\mathcal{M})$. We showed above that $W_{ij}^B(\mathcal{M}) \geq W_{ij}^S(\mathcal{M})$ always holds. Hence, to satisfy the IC constraint in this case we need to ensure that $W_{ij}^B(\mathcal{M}) \geq W_{ij}^B(\mathcal{N})$ or, equivalently, $L_N \leq L_2(i)$.

Alternatively, if loan terms $\mathcal{N}$ are payoff-maximizing, i.e., $W_{ij}^S(\mathcal{N}) \geq W_{ij}^B(\mathcal{M})$ for pair $ij$ (that is $L_N \geq L_3(ij)$ by (CHM)), then the l.h.s. of the IC constraint (IC) equals $W_{ij}^S(\mathcal{N})$. Since $W_{ij}^B(\mathcal{M}) \geq W_{ij}^S(\mathcal{M})$ for any $ij$, this implies that $W_{ij}^S(\mathcal{N}) \geq W_{ij}^S(\mathcal{M})$ holds in this case. Thus, to satisfy the IC constraint we need to ensure that $W_{ij}^S(\mathcal{N}) \geq W_{ij}^B(\mathcal{N})$, or equivalently $L_N \leq L_1(ij)$.

Intuitively, in both cases constraint (IC) requires that the lender ensure that the borrowers have no incentive to take the low-interest loan $\mathcal{N}$ but operate as bogus group.

**B. Homogeneous groups**

We show that the incentive constraint (IC) is always satisfied for homogeneous $ii$ groups for $L_M = pV$ and any $L_N \leq p(2 - p)V$. Indeed, from (22) we have $L_1(ii) = p(2 - p)V$. The no-default constraint (16) thus ensures $L_N \leq L_1(ii)$ and hence (IC) is always satisfied when loan terms $\mathcal{N}$ are payoff-maximizing. It is also easy to verify that $L_3(ii) = \frac{2pk_i - 4 + 2p}{p(k_i + k_j)} - \frac{1}{2} p V = L_2(i)$ for any $k_i$ and $p$ and hence, whenever loan terms $\mathcal{M}$ are payoff-maximizing (when $L_N < L_3(ii)$), the incentive constraint is also automatically satisfied for any $L_N \leq p(2 - p)V$. Overall, we conclude that homogeneous $ii$ groups choose loan terms $\mathcal{N}$ whenever $L_N \geq L_3(ii)$ or

$$L_N \geq \frac{pk_i - 2 + p}{pk_i - 1} p V$$

(24)

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

**C. Heterogeneous groups**

From Lemma 2 we know that $L_M^* = pV = L_B$, hence the (bogus-group) loan terms $\mathcal{M}$ solving (UP') are always $(L_B, R_B)$. Remember that $L_S = p(2 - p)V$ and $L_E = \frac{p(2 - p)V}{1 + \frac{p(2 - p)}{2l_1 - p}(k_H - k_L)}$. Define

$$L_F \equiv L_2(H) = \frac{pk_H - 1}{pk_H - \frac{1}{2} p} p V.$$

There are two possible cases, depending on whether loan terms $\mathcal{N}$ or $\mathcal{M}$ are optimally chosen by an $ij$ group. If $\mathcal{M}$ are payoff-maximizing (that is, $W_{ij}^B(\mathcal{M}) > W_{ij}^S(\mathcal{N})$), which happens when $L_N < L_3(\mathcal{HL})$, then the lender must also ensure that $L_N \leq L_2(H) = L_F$ to satisfy (IC) and prevent bogus groups from picking loan terms $\mathcal{N}$ and causing a loss. If, alternatively, loan terms $\mathcal{N}$ are payoff-maximizing, which happens when $L_N \geq L_3(\mathcal{HL})$, then, to satisfy (IC), the lender must ensure that $L_N \leq L_1(\mathcal{HL}) = L_E$. In both cases we also need $L_N \leq p(2 - p) V = L_S$ from
the no-default constraint. The objective function of problem (UP') is weakly increasing in $L_N$, hence the optimal loan size $L_N$ is the largest possible value satisfying both the (IC) and no-default constraints. Overall, this requires $L_N \leq \min\{L_S, L_E\}$ when loan terms $N$ are payoff-maximizing and $L_N \leq \min\{L_S, L_F\}$ when loan terms $M$ are payoff-maximizing.

We know that $L_S \geq L_E$ whenever $k_H \geq k_L$, which is always true. It is also easy to show that $L_S > L_F$ for any $p \in (0,1)$ and $k_H > 0$. This implies that constraint (IC) always binds for $HL$ pairs. Hence, for such pairs $L_N^*$ must equal either $L_E$ or $L_F$, both of which were shown to be smaller than $L_S$. Overall, this implies that $L_N^* = L_E$ when loan terms $N$ are payoff-maximizing and $L_N^* = L_F$ when loan terms $M$ are payoff-maximizing.

To complete the analysis, we determine the (standard-group) loan size $L_N^*$ depending on the model parameters $p, k_H, k_L$. First, it is easy to show that,

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

where $f(k)$ was defined in the proof of Proposition 3. Second, we can show that

$$L_3(HL) > L_F \iff k_H > f(k_L).$$

The above results imply that for $k_H \in [k_L, f(k_L)]$ the solution to problem (UP') is $L_N^* = L_E$. This is so, since for such $k_H$ and $k_L$ we have $L_E \geq L_3(HL), L_F \geq L_3(ML)$ and (IC) requires $L_N^* \leq L_F$. In this case $HL$ borrower pairs select loan terms $N = (L_E, \frac{L_E}{p(2-p)})$ and form standard groups. Alternatively, when $k_H > f(k_L)$ the above results imply $L_E < L_3(ML)$ and $L_F < L_3(ML)$. Hence, $L_N < L_3(ML)$ and thus loan terms $M$ are payoff-maximizing. Constraint (IC) then implies $L_N^* = L_F$ and $HL$ pairs select terms $M = (L_B, R_B)$ (see Lemma 2) and form bogus groups.

Going back to homogeneous borrower pairs, note first that $L_3(LL) < L_3(HH) \leq L_3(ML)$ for any $p < 1$ and $k_H \geq k_L$. The first (strict) inequality follows from the fact that $L_3(ii)$ (see (CHM)) is strictly increasing in $k_i$. The second (weak) inequality is verified directly, with equality holding only for $k_H = k_L$. This implies that, when $k_H \in [k_L, f(k_L)]$ and so $L_N^* = L_E$, we have $L_E \geq L_3(ii)$ for both $i = L$ and $H$ and hence, using (24), homogeneous groups choose loan $N$ and a form a standard group. In the alternative case, when $k_H > f(k_L)$, we showed above that $L_N^* = L_F$. Using the definitions of $L_F$ and $L_3(ij)$ it is easy to verify directly that $L_F > L_3(HH)$ for any $k_H > 0$ and $p \in (0,1)$. Therefore, $L_F > L_3(LL)$ and so once again, by (24), homogeneous groups choose loan terms $N$ and a form standard groups. This completes the characterization of the loan menu, loan terms and group forms solving problem (UP').

---

25 Obviously $L_E \neq L_F$ in general, however, for any given $k_L$, the expressions $L_3(ML), L_F$ and $L_E$ taken as functions of $k_H$ cross at the same point (at $k_H = f(k_L)$) and thus $L_3(ML)$ is larger than both $L_F$ and $L_E$ whenever $k_H \geq f(k_L)$ and smaller otherwise.
Proof of Proposition 5

Condition (12) is always satisfied for homogeneous \((HH\ or\ LL)\) pairs, given that \(R \leq V\) by the no-default constraint. This implies that the loan terms for homogeneous groups are \((L_S, R_S)\) – the same terms as in Proposition 1.

For \(HL\) pairs, using the lender’s break-even constraint, \(p(2 - p)R = L\), condition (12) is equivalent to
\[
(k_H - k_L)L + \frac{2(1-p)L}{p(2-p)} \leq 2(1-p)V, \text{ or } L \leq L_E
\]
where \(L_E = \frac{p(2-p)V}{1 + \frac{p(2-p)}{2(1-p)}(k_H - k_L)}\) as defined earlier. Since \(L_E < L_S\) for \(k_H > k_L\) and since the lender’s objective function is increasing in \(L\), the payoff-maximizing loan terms for heterogeneous \((HL)\) groups are \(L^# = L_E\) and \(R^# = \frac{L_E}{p(2-p)}\).

Appendix B – Individual repayment decisions

B.1 (Repay, Default) equilibrium

In general, all pure strategy Nash equilibria of the stage 1 game described in Section 5.2 for given loan terms \((L, R)\) are: (i) (Repay, Repay) if \(R \leq \frac{1-2p}{2-p}V\); (ii) (Repay, Default) or (Default, Repay) if \(\frac{1-2p}{2-p}V < R \leq \frac{1}{2}V\); and (iii) (Default, Default) if \(R > \frac{1}{2}V\). In the main text we focus on the (Repay, Repay) equilibrium. A lender could potentially offer a loan with \(\frac{1-2p}{2-p}V < R \leq \frac{1}{2}V\) inducing the (Repay, Default) outcome it stage 1 and break even by setting \(L = pR\). However, we show that it is never optimal for the lender to choose loan terms that induce this equilibrium for standard groups. Indeed, note that in a (Repay, Default) equilibrium, switching from standard to bogus group does not affect the repayment probability (it is \(p\) in both cases), the interest rate \((1/p\) in both), and the expected continuation value \((2pV\) in both). However, forming a bogus group raises expected output from \(p(k_i + k_j)L\) to \(2pk_iL\), while supporting the same maximum loan size \(\frac{pk_i}{2}\). Therefore, it is always (weakly) better for the borrowers to form a bogus group if facing loan terms inducing (Repay, Default) equilibrium. This implies that our focus on the (Repay, Repay) equilibrium in Section 5.2 is not restrictive when bogus groups can be formed since the maximum group payoff in (Repay, Default) equilibrium is always weakly dominated by the bogus group payoff at loan terms \((L_B, R_B)\).\(^{26}\)

B.2 Results

Given loan terms \((L, R)\) satisfying feasibility (1) and the no strategic default condition (10), the expected total payoff of standard group of type \(ij \in \{HH, LL, HL\}\) equals
\[
\hat{W}_{ij}^S(L, R) = p(k_i + k_j)L - 2p(2-p)R + 2p(2-p)V, \quad \text{(SEP)}
\]
\(^{26}\)Supporting the (Repay, Default) equilibrium may be hard in practice – it either means that one borrower is allowed to consistently free ride on the repayment decision of the other, or the borrowers take turns over time which requires co-ordination and commitment at odds with the non-cooperative setting assumption.
where we use $\bar{W}$ to indicate all payoffs in the individual default decision setting.

The lender’s break even condition remains as in (3),

$$R = \frac{L}{p(2 - p)}.$$

**Standard groups**

As in Section 3, we first characterize the loan terms if bogus groups are assumed away exogenously.

**Proposition B-1**

(a) With standard groups only and individual default decisions, the optimal loan terms are $S^0 \equiv (L_{S^0}, R_{S^0})$ with

$$L_{S^0} = p(1 - p)V \text{ and } R_{S^0} = \frac{1 - p}{2 - p}V.$$

(b) The loan terms $S^0$ remain optimal if the borrowers’ productivities are unobservable to the lender.

The proof is isomorphic to that of Proposition 1 and hence omitted.

**Bogus groups**

To introduce bogus groups in the individual-default setting, we follow the microfinance literature and assume that the group members share social capital which can be used to enforce a (possibly non-monetary) transfer $T$ between them in order to ensure that the bogus group cosigner (‘ghost member’) obtains at least the same payoff as she would in a standard group under repayment.\(^{27}\)

More precisely, assume that the bogus group leader (the member who invests $2L$) makes a default/repay decision individually, based on his own payoff as the borrowers in a standard group, however, he does so with the understanding that a transfer $T$ must be made in either case. The repay/default trade-off is thus unaffected by the transfer. The cosigner has no decisions to make since her project is not funded and there is limited liability (she has no other wealth or income).

If the borrowers consider forming a bogus group, they optimally invest all funds ($2L$) into the higher productivity project. Assuming $k_i \geq k_j$ without loss of generality, all funds are invested in project $i$. Consider the bogus group leader’s repay/default decision for given loan terms $(L, R)$. As before, it is not optimal to repay partially. Repayment in stage 2 (reached if the leader repays in stage 1) is optimal as long as

$$R \leq V,$$

\(^{27}\)It is easy to compute the required transfer amount and show that it is always feasible upon project success, both under repayment and strategic default. Alternatively, the transfer can be non-monetary.
yielding a payoff of $V - R$. Conditional on project success, the bogus group leader’s payoff from repaying in stage 1 is larger than her payoff from defaulting if

$$k_i(2L) - R - T + (V - R) \geq k_i(2L) - T \quad \text{or} \quad R \leq V/2$$

The no-strategic-default condition over both stages 1 and 2 is then:

$$R \leq \frac{V}{2} \quad (25)$$

This implies that, for given loan terms $(L, R)$, the joint expected payoff $ar{W}_{ij}^B(L, R)$ of a bogus group with productivities $k_i, k_j$ is:

$$\bar{W}_{ij}^B(L, R) = 2pk_iL - 2pR + 2pV \quad (26)$$

Free entry implies $L = pR$ as in Section 3. Combined with the no-default condition (25), the payoff-maximizing loan terms for a bogus group (as if known and taken in isolation) are

$$L_{B'} \equiv \frac{pV}{2} \quad \text{and} \quad R_{B'} \equiv \frac{V}{2}. \quad (27)$$

Comparing bogus and standard groups, the basic intuition from Section 3.3 still holds, as discussed in Section 5.2 in the main text.

It is easy to show that Lemma 1 still applies. The following Proposition is the counterpart of Proposition 2 and shows the consequence of endogenous bogus group formation in the lender were to offer terms $S'$ to all borrowers.

**Proposition B-2:** If loan $S' = (L_{S'}, R_{S'})$ defined in Proposition B-1 is offered to all borrowers, then:

(a) if

$$\Delta \equiv k_H - k_L > \frac{2}{p(2-p)}:$$

heterogeneous (HL) borrower pairs form bogus groups; otherwise HL pairs form standard groups;

(b) all homogeneous (HH or LL) borrower pairs form standard groups;

(c) bogus groups cause a loss to the lender.

**Proof:** see Appendix B.3.

The previous intuition (see Proposition 2) carries over – at loan terms $S'$ designed for standard groups only, bogus groups are strictly optimal only for heterogeneous borrower pairs. An additional condition is now needed – the productivity differential across the two project in an HL pair must be large enough for the borrowers to form a bogus group. If condition (28) does not hold (that is, for $k_H$ and $k_L$ relatively close), then offering loan terms $(L_{S'}, R_{S'})$ would not cause a loss to the lender since all groups will be standard. However, as we show below, offering these terms to all borrowers may not be optimal since allowing (inducing) bogus groups can be efficiency-improving as larger amount of funds can be invested in the higher return project.
Endogenous group form and loan terms

Here we only analyze the case of observable project productivities. The analysis of the unobservable productivity case is similar to that in Section 4.2 and is available upon request.

**Proposition B-3:** Suppose repayment decisions are individual and the productivities \( k_i \) and \( k_j \) are observed by the lender. The optimal loan terms \((L^*, R^*)\) for an \( ij \in \{HH, LL, HL\} \) group are:

(a) homogeneous, \( k_i = k_j \) (HH or LL) groups: if \( p(2p-1)k_i > 1 \) then \( L^* = \frac{pW}{2} \equiv L_{B'} \), \( R^* = \frac{V}{2} \equiv R_{B'} \) and the group is bogus \((\tau^* = 0)\) while if \( p(2p-1)k_i \leq 1 \) then \( L^* = p(1-p)V = L_{S'}, R^* = \frac{(1-p)W}{2-p} = R_{S'} \) and the group is standard \((\tau^* = 1)\).

(b) heterogeneous, \( k_i > k_j \) (HL) groups: depending on the parameter values,\(^{28}\) either \( L^* = L_{B'} \) and \( R^* = R_{B'} \) and the group is bogus or \( L^* = \min\{L_{S'}, L_{E}\}, R^* = \frac{L^*}{p(2-p)} \) and the group is standard, where \( L_E \) is as defined in Proposition 3.

**Proof:** see Appendix B.3.

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

**Corollary B.** Let \( R_E = \frac{L_E}{p(2-p)}, \tilde{k} = \frac{2-3p}{p(2-p)(2-p-1)}, \tilde{k} = \frac{1}{p(2-p-1)}, d(k) = \frac{2}{p(2-p)} + k, g(k) = \frac{1}{p} + \frac{1-p}{p} k \) and \( k^l \) be the larger root of \( W^S(L_E, R_E) = W^B(L_{B'}, R_{B'}) \) written as quadratic equation in \( k_H \).\(^{29}\) Calling \( S' \equiv (L_{S'}, R_{S'}), B' \equiv (L_{B'}, R_{B'}) \) and \( E = (L_{E}, R_{E}) \), the loan terms for all possible group types and parameter values are:

<table>
<thead>
<tr>
<th>Parameter conditions</th>
<th>Loan terms and group form</th>
</tr>
</thead>
<tbody>
<tr>
<td>((a)) high productivities or high ( p ): ( p(2p-1)k_L &gt; 1 )</td>
<td>( B', ) bogus</td>
</tr>
<tr>
<td>((b)) high ( k_H ) and low ( k_L ): ( p(2p-1)k_H &gt; 1 \geq p(2p-1)k_L )</td>
<td>( S', ) standard</td>
</tr>
<tr>
<td>((c)) low productivities or low ( p ): ( p(2p-1)k_H \leq 1 ) and</td>
<td></td>
</tr>
<tr>
<td>i) ( k_H \in [k_L, \min{g(k_L), d(k_L)}] )</td>
<td>( S', ) standard</td>
</tr>
<tr>
<td>ii) ( k_H \in (d(k_L), k^l] \wedge {(k_L &lt; \tilde{k} \wedge p \in (\frac{1}{2}, \frac{p}{2}) } \lor p \leq \frac{1}{2} )</td>
<td>( S', ) standard</td>
</tr>
<tr>
<td>iii-1) ( k_H &gt; g(k_L) \land k_L \geq \tilde{k} )</td>
<td>( S', ) standard</td>
</tr>
<tr>
<td>iii-2) ( k_H &gt; k^l \lor {(k_L &lt; \tilde{k} \wedge p \in (\frac{1}{2}, \frac{p}{2}) } \lor p \leq \frac{1}{2} )</td>
<td>( S', ) standard</td>
</tr>
</tbody>
</table>

**Proof:** see Appendix B.3.

Discussion of these results is provided in Section 5.2.

---

\(^{28}\)The exact conditions are shown in Corollary B.

\(^{29}\)The threshold \( k^l \) depends on \( k_L \) and \( p \) – see the proof of Proposition B-3 for details.
B.3 Proofs (online appendix)

Proof of Proposition B-2:
(a) Consider a heterogeneous (HL) borrower pair, that is, \( k_i = k_H \) and \( k_j = k_L \). Substituting in the loan terms \((L^S, R^S)\) derived in Proposition B-1, it is easy to verify that condition (5) from Lemma 1 implies that bogus groups are payoff-maximizing if and only if inequality (28) in part (a) holds. Otherwise, Lemma 1 implies that forming standard group is optimal.

(b) Using Lemma 1, since for homogeneous pairs \((k_i = k_j)\) the l.h.s. of (5) is zero while its r.h.s. is positive at \((L^S, R^S)\), it is always payoff-maximizing for a homogeneous pair to form a standard group.

(c) Shown analogously to Proposition 2(c).

Proof of Proposition B-3:
As in Section 4.1, the lender chooses loan terms \((L, R)\), contingent on the observed pair type \(ij\) (omitted to save on notation), to maximize the group payoff subject to incentive compatibility, no-default and break-even constraints:

\[
\max_{L, R, \tau \in \{0, 1\}} \tau \bar{W}^S(L, R) + (1 - \tau)\bar{W}^B(L, R) \tag{OP}
\]

s.t. \(\tau \bar{W}^S(L, R) + (1 - \tau)\bar{W}^B(L, R) \geq \tau \bar{W}^B(L, R) + (1 - \tau)\bar{W}^S(L, R) \tag{IC}
\]

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

\[R = \frac{L}{(1-p)} + (1 - \tau)\frac{L}{p} \text{ (zero profit)} \]

and where the payoffs for a standard group, \(\bar{W}^S(L, R)\) and bogus group, \(\bar{W}^B(L, R)\) are:\(^{30}\)

\[\bar{W}^S(L, R) = p(k_i + k_j)L - 2p(2 - p)R + 2p(2 - p)V \tag{29}\]

and

\[\bar{W}^B(L, R) = 2pk_iL - 2pR + 2pV \tag{30}\]

Above we used the result 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 – see Appendix B.1.\(^{31}\) Hence, we write the no-default and zero-profit constraints in problem (OP) only for the (Repay, Repay) equilibrium whenever the lender wants to induce a standard group \((\tau = 1)\). 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) has the same interpretation as in Section 4.1.

To prove Proposition B-3 we will use two auxiliary results.

---

\(^{30}\)We exhibit the total payoff under all possible Nash equilibria for standard groups even though the (Repay, Default) and (Default, Default) equilibria cannot obtain at optimum. The reason is that the hidden action of borrowers choosing the group form requires evaluating deviations from the prescribed behaviour.

\(^{31}\)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.
Remark 1: Comparing the payoffs $\tilde{W}^S(L, R)$ and $\tilde{W}^B(L, R)$ from (29) and (30) for given loan terms $(L, R)$ with $R \leq \frac{(1-p)V}{2-p}$, an $ij$ pair would optimally choose form bogus group if and only if,

$$(k_i - k_j)L > 2(1-p)(V - R)$$

(BC)

and form standard group otherwise.

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

(a) Homogeneous groups

Suppose the lender faces a homogeneous borrower pair $(ii$ with $i = L$ or $H$) and wants to induce a standard group (choose $\tau = 1$). The zero profit constraint implies $R = \frac{L}{p(2-p)}$, which substituted into the no-default constraint yields $L \leq p(1-p)V$. Constraint (IC) is satisfied for any $R$ satisfying the no-default constraint $R \leq \frac{(1-p)V}{2-p}$ since condition (BC) does not hold for $k_i = k_j$ (see Remark 1). The objective $\tilde{W}^S(L, \frac{L}{p(2-p)})$ 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 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 $\tilde{W}^B(L, \frac{L}{p})$ equals $(2pk_i - 2)L + 2pV$ which is strictly increasing in $L$. Hence, as long as (IC) is not violated, it is optimal to have the no-default constraint bind and so $L^* = \frac{pV}{2} = L_{B'}$ and $R^* = \frac{V}{2} = R_{B'}$. Constraint (IC) is indeed not violated at $(L_{B'}, R_{B'})$ since forming a bogus group is optimal at $R = R_{B'} = \frac{V}{2} > \frac{(1-p)V}{2-p}$ (see Remark 2).

To decide whether choosing $\tau = 0$ or $\tau = 1$ is optimal, the lender compares the payoffs from inducing a standard group with loan terms $(L_{S'}, R_{S'})$ vs. bogus group with loan terms $(L_{B'}, R_{B'})$. Choosing $\tau = 1$, $L^* = L_{S'}$ and $R^* = R_{S'}$ is optimal when $\tilde{W}^S(L_{S'}, R_{S'}) \geq \tilde{W}^B(L_{B'}, R_{B'})$ which is equivalent to:

$$(2pk_i - 2)L_S + 2(2-p)pV \geq (2pk_i - 2)L_B + 2pV \text{ or,}$$

$$p(2p - 1)k_i \leq 1.$$  

Choosing $\tau = 0$, $L^* = L_{B'}$ and $R^* = R_{B'}$ is optimal otherwise.

(b) Heterogeneous groups

Suppose the lender faces a heterogeneous $(k_H, k_L)$ borrower pair and considers inducing standard group ($\tau = 1$). The zero profit constraint implies $R = \frac{L}{p(2-p)}$, which substituted into the no-default constraint implies $L \leq p(1-p)V = L_{S'}$. Unlike for homogeneous pairs, constraint (IC) is not automatically satisfied for any $R$ satisfying the no-default condition. Indeed, for (IC) to
hold, that is $W^S(L, R) \geq W^B(L, R)$, using (BC) we need,

$$p(k_H - k_L)L \leq 2p(1 - p)(V - \frac{L}{p(2-p)})$$

or,

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

Therefore, the optimal loan terms in this case are $L_N \equiv \min\{L_E, L_{S'}\}$ and $R_N \equiv \frac{L^*}{p(2-p)}$.

Suppose instead the lender wants to induce a bogus group ($\tau = 0$). The analysis is analogous to that for homogeneous pairs. Constraint (IC) is satisfied at $L = \frac{pL}{2}$ and $R = \frac{V}{2}$. Therefore, the optimal loan terms in this case are $L^* = L_{B'}$ and $R^* = R_{B'}$, as before.

To choose $\tau = 0$ or $\tau = 1$ the lender compares the group payoff from inducing a standard group with loan terms ($L_N, R_N$) vs. a bogus group with loan terms ($L_{B'}, R_{B'}$). There are two cases, depending on whether $L_E$ or $L_{S'}$ is smaller. $L_E \geq L_{S'}$ is equivalent to

$$\frac{p(1-p)V}{\frac{pL}{2} + \frac{p}{2}(k_H - k_L)} \geq \frac{p(1-p)V}{L_E} \iff \frac{2}{p(2-p)} \geq k_H - k_L$$

which is the converse of the condition in Proposition B.2.

Suppose first that condition (CC1) is satisfied, that is

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

and hence $L_N = L_{S'} \leq L_E$. Then, choosing $\tau = 1$, $L^* = L_{S'}$ and $R^* = R_{S'}$ is optimal if $W^S(L_{S'}, R_{S'}) \geq W^B(L_{B'}, R_{B'})$ which is equivalent to

$$(p(k_L + k_H) - 2)L_S + 2(2-p)pV \geq (2pk_H - 2)L_B + 2pV$$

or,

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

and $\tau = 0$, $L^* = L_{B'}$ and $R^* = R_{B'}$ is optimal otherwise.

If inequality (CC1) does not hold, that is if $k_H > d(k_L)$, 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 $W^S(L_E, R_E) \geq W^B(L_{B'}, R_{B'})$ which is equivalent to

$$(p(k_L + k_H) - 2)\frac{p(1-p)V}{\frac{pL}{2} + \frac{p}{2}(k_H - k_L)} + 2(2-p)pV \geq (2pk_H - 2)\frac{pV}{2} + 2pV$$

or $\tilde{f}(k) \geq k_H$

(CC3)

where

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

and $c \equiv \frac{4p^2 - 11p + 8}{p^2(2-p)}$.

Choosing $\tau = 0$, $L^* = L_{B'}$ and $R^* = R_{B'}$ is optimal otherwise, for $\tilde{f}(k_L) < k_H$.

Remark 3. Note that, since $L_E < L_{S'}$, then $k_H > d(k_L)$ implies $W^S(L_{S'}, R_{S'}) > W^S(L_E, R_E)$, which means that condition (CC3) is tighter than (implies) condition (CC2).

\footnote{One can show that the smaller root of the quadratic equation of which $\tilde{f}(k)$ is the larger root is strictly smaller than $k$. Thus, for $k_H \geq k_L$ inequality (CC3) is satisfied iff $k_H \leq \tilde{f}(k_L)$.}
In sum, we obtain:

(a) if $k_H \leq d(k_L)$ then it is optimal to offer loan terms $S'$ if $k_H \leq g(k_L)$ and $B'$ if $k_H > g(k_L)$

(b) if $k_H > d(k_L)$ then it is optimal to offer loan terms $E$ if $k_H \leq \tilde{f}(k_L)$ and $B'$ if $k_H > \tilde{f}(k_L)$

Proof of Corollary B

The results for homogeneous groups (Table columns 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 B-3(a). Hence, we only discuss heterogeneous groups below. Call $\hat{k} \equiv \frac{2-3p}{p(2-p)(2p-1)}$ and $\tilde{k} \equiv \frac{1}{p(2p-1)}$, defined for $p > 1/2$. The following results are easily shown (all functions are defined in the proof of Proposition B-3).

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

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

$$d(k) < \frac{1}{2} \left( d(k) + n \right) + \sqrt{d(k) + n}$$

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

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

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

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

Result 4. Suppose $p > 1/2$. If $k \geq \hat{k}$, then $d(k) \geq \tilde{f}(k)$ and $d(k) \geq g(k)$. If $k < \hat{k}$, then $d(k) < \tilde{f}(k)$ and $d(k) < g(k)$.

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

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

We proceed with the proof of Corollary B. Parts (a)--(c) below refer to the corresponding Table lines in the corollary statement.

(a) The condition $p(2p-1)k_L > 1$ (equivalent to $k_L > \hat{k}$) can only hold for $p > 1/2$. Note also that if $p > 4/5$ then $\hat{k} < k_{\text{min}}$ by Result 6 and hence $p(2p-1)k_L > 1$ is satisfied for any $k_L \geq k_{\text{min}}$. Using Result 5, $k_L > \hat{k}$ implies $k_H \geq k_L > g(k_L)$, and hence, by (CC2), it is optimal to offer loan terms $B'$ for any such $k_H, k_L$.

(b) As explained in (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. Assuming $p \in (1/2, 4/5]$, the inequality $k_H > \hat{k}$ implies (by 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 CC2) implies $k_H > \tilde{f}(k_L)$, which is the negation of the stricter inequality (CC3). Using Proposition B-3(b), in either of these cases it is optimal to offer loan terms $B'$ to HL pairs for any such $k_H, k_L$.

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

(c-i) Suppose $p \leq 1/2$, in which case $p(2p-1)k_H \leq 1$ holds for any $k_H \geq k_{\text{min}}$. Then, by
Result 1 and since \( d(k) > k \) for all \( k > 0 \), using Proposition B-3(b), the lender offers loan terms \( S' \) if \( k_H \in [k_L, d(k_L)] \), loan terms \( E \) if \( k_H \in (d(k_L), \bar{f}(k_L)] \), and loan terms \( B' \) if \( k_H > \bar{f}(k_L) \). Calling \( \bar{k}^{-1} \equiv \bar{f}(k_L) \), these cases correspond to lines (i), (ii) and (iii-2) in part (c) of the Table.

(c-ii) Suppose \( p \in \left( \frac{1}{2}, \frac{4}{7} \right) \). Then, by Result 3, \( \hat{k} > k_{\min} \). If \( k_L \in [k_{\min}, \hat{k}] \), then from Result 4, \( d(k_L) < \bar{f}(k_L) \) and \( d(k_L) < g(k_L) \). Hence, using Proposition B-3(b), the lender offers loan terms \( S' \) if \( k_H \in [k_L, d(k_L)] \), loan terms \( E \) if \( k_H \in (d(k_L), \bar{f}(k_L)] \), and loan terms \( B' \) if \( k_H > \bar{f}(k_L) \). These cases map to lines (i), (ii) and (iii-2) in part (c) of the Table. Alternatively, if \( k_L \geq \hat{k} \), then \( d(k_L) \geq \bar{f}(k_L) \) and \( d(k_L) \geq g(k_L) \) by Result 4, and so it is optimal to offer \( S' \) for \( k_H \in [k_L, g(k_L)] \) and offer \( B' \) for \( k_H > g(k_L) \). These correspond to lines (i) and (iii-1) in the Table.

(c-iii) Suppose \( p \in \left[ \frac{4}{7}, \frac{2}{5} \right] \). By Result 3, \( \bar{k} \leq k_{\min} \) and so \( k_L \geq \hat{k} \). Thus, by Result 4, \( d(k_L) \geq \bar{f}(k_L) \) and \( d(k_L) \geq g(k_L) \). If \( k_L > \hat{k} \), then, by Result 5, \( k_L > g(k_L) \) and so \( k_H > g(k_L) \). Hence, the lender offers loan terms \( B' \) (line iii-1 in the Table). If, instead, \( k_L \in [k_{\min}, \hat{k}] \) then, by Result 5, \( g(k_L) \geq k_L \) and the lender offers loan terms \( S' \) for \( k_H \in [k_L, g(k_L)] \) and loan terms \( B' \) for \( k_H > g(k_L) \), corresponding to lines (i) and (iii-1) in the Table.

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

### Appendix C

<table>
<thead>
<tr>
<th>Table A1: Sample statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>group size</td>
</tr>
<tr>
<td>loan size</td>
</tr>
<tr>
<td>required monthly repayment</td>
</tr>
<tr>
<td>age</td>
</tr>
<tr>
<td>% married</td>
</tr>
<tr>
<td>% intended loan use = agriculture</td>
</tr>
<tr>
<td>% intended loan use = consumption</td>
</tr>
<tr>
<td>% intended loan use = business#</td>
</tr>
<tr>
<td>% without education</td>
</tr>
<tr>
<td>% with primary school</td>
</tr>
<tr>
<td>% with junior school</td>
</tr>
<tr>
<td>% with high school</td>
</tr>
<tr>
<td>% with college or above</td>
</tr>
<tr>
<td>% Manchu</td>
</tr>
<tr>
<td>% Mongol</td>
</tr>
<tr>
<td>sample size</td>
</tr>
</tbody>
</table>

Note: #business use = manufacturing, services, wholesale or transportation.