The enforcement advantage of external monitoring: Lessons from an experiment with joint-liability groups in Burkina Faso

Wouter Gelade*and Catherine Guirkinger†

April 2018

Abstract

The theoretical literature sees cost-effective peer monitoring as a key to the success of joint-liability over individual-liability credit. Yet peer monitoring may also involve particular costs, such as the cost of denouncing wrongdoing, that are larger for group members than for the lender. Moreover, in practice, joint-liability credit often involves monitoring by the lender. To investigate the role of external monitoring and explore the nature of the costs associated with internal group monitoring, we conduct a field experiment in joint-liability credit groups in Burkina Faso. In the experiment, we randomly increase the intensity of external monitoring in credit groups. We find that external monitoring crowds out monitoring by the group leader, causes loan renewal decisions to be more severe and the handling of individual default to be more state-dependent. In addition, increased external monitoring decreases favoritism towards members of the leader's family. We argue that external monitoring reduces the costs of "pointing fingers" at moral-hazardous behaviors and of sanctioning.

*University of Namur, Center for Research in the Economics of Development, 8 rempart de la Vierge, Namur 5000, Belgium. Email: wouter.gelade@unamur.be

†Corresponding author. University of Namur, Center for Research in the Economics of Development, 8 rempart de la Vierge, Namur 5000, Belgium. Email: catherine.guirkinger@unamur.be

‡We aknowledge the funding of the United States Agency for International Development (USAID) under grant AID-OAA-L-12-00001 for the Feed the Future Innovation Lab for Assets and Market Access at UC Davis. The content of the paper is the sole responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government. Our gratitude also extends to seminar participants at Cornell University (United States). Special thanks are due to Jean-Marie Baland, Guilhem Cassan, Xavier Gine, William Pariente, Jean-Philippe Platteau, two anonymous referees and the journal editor for constructive comments.

1 Introduction

Throughout the world, microfinance institutions (MFI) continue to rely heavily on joint-liability lending: according to the best estimates available, joint-liability loans account for 42% of MFIs' portfolios in 2011, and it is not clear that this share will necessarily decrease as the sector continues to grow (deQuidt et. al. 2016a, deQuidt et. al. 2016b). In the economic literature, the success of joint-liability lending in serving borrowers who lack formal collateral, is largely attributed to the delegation of the monitoring efforts to credit group members. While this delegation effectively transfers the cost of monitoring to the group and is akin to a "tax" on borrowers (Gine and Karlan, 2014), it is typically considered efficient because group members can monitor at lower cost than the lender. The geographical and social proximity of group members give them an informational advantage over the lender and enable them to keep an eye on each other's economic activities without incurring great expenses (Stiglitz 1990, Banerjee et. al. 1994, Ghatak 1999, Magdajewicz, 2011, Cason et. al. 2012). It is surprising in this context to observe active external monitoring by lenders in joint-liability schemes (Chowdhury, 2005, Ghatak and Guinane, 1999).

In this paper we question the assumption that group members have an advantage in monitoring. We argue that it may be particularly difficult for them to use the information available to impose discipline or sanctions. Indeed, in order to be effective, internal monitoring requires that a group member who observes shirking acts upon this information and demands that the wrongdoer be sanctioned. Reporting wrongdoing or requiring punishment may however be more costly for group members than for the lender, as the former are more likely to suffer from a damaged relationship with the whistleblown person. In addition, it may be particularly difficult for an insider to discipline better-connected or more powerful individuals within the group. In other words, internal monitoring and sanctioning are prone to favoritism or cronyism. Therefore, information may be more efficiently and impartially used to discipline borrowers when it is collected by an outsider.

To explore this possibility empirically, we conducted a field experiment with joint-liability credit groups in Burkina Faso. Group members are cotton producers and they use credit to finance cotton production. The lender is a commercial bank which works closely with a cotton company. The company purchases the entire cotton output from the groups. Loans consist of a package of inputs delivered in-kind and they are repaid after the harvest, through a direct transfer from the cotton company to the bank. The output of a group thus serves as a collateral. It is extremely rare that a group defaults: the group's output is typically higher than the total amount owed by the members. However, an individual borrower's output may not be sufficient to cover his credit, in

which case the loss is de facto covered by his fellow group members. When this happens, the group has two options: asking the defaulting member to either contribute a larger amount for next year's repayment, or to make an immediate compensatory payment in cash. Decisions regarding the size of individual loans or the handling of individual default are taken during plenary group meetings.

In our randomized control trial, some randomly chosen credit groups received an intensive external monitoring treatment in the sense that, agents who work with the company and already monitor the groups were asked to make extra visits to borrowers' fields. These extra visits occurred after most inputs were applied, but before the cotton was harvested. During a field visit, agents typically measure the size of the field and observe the state of the cotton crop in order to detect input diversion, one of the leading causes of loan default. Group leaders regularly perform this type of visits as well, which they use to observe crops and measure fields. Because the monitoring technology is rather simple, the external agent can detect input diversion as easily as internal group leaders. In other words, groups leaders do not enjoy a large informational advantage when it comes to monitoring. This rather unique feature of the joint-liability loans under study helps us isolate the effect of the *identity* of the monitor on the consequences of monitoring. As mentioned above, we expect the change of monitor to affect the sanctioning regime.[1]

Our first result is that external monitoring fully crowds out monitoring by the group leader, implying that the experiment changes *who* is doing the monitoring. When we explore how the allocation of credit for the following campaign is affected, we find that in the treated groups, loan renewal is more stringent and the size of future loans is reduced. Furthermore, the effect is mainly observed for the members of the leader's family, suggesting that they are usually favored or less likely to be disciplined when monitoring is performed by the leader. Turning to the consequences of individual default in the treatment and the control groups, we find that the sanctioning regime changes: as external monitoring increases, individual defaulters are less often asked to pay back their loan immediately (by selling assets), unless they had shirked. Interestingly such leniency is not applied to the leader's family members. This is again consistent with the presence of favoritism in the internal monitoring regime. While there may be alternative interpretations of our set of results, we argue, with the help of qualitative evidence, that the main difference between internal and external monitoring is that the latter can be used more efficiently when it comes to "pointing fingers" at moral-hazardous behaviors, especially in the case of members of the leader's family. While we cannot formally draw

[1]While it is a priori hard to detect a change of sanctioning regime, we will explain later how the timing of our experiment inside the agricultural calendar offers a unique opportunity to observe the new sanctions.

welfare implications from our analysis, we argue that that external monitoring is likely to improve the efficiency of group management. Furthermore the cost of more intense external monitoring appears to be modest: a back-of-the-envelope calculation suggests that it represents an increase of the interest rate by 0.23 percentage points.

Our analysis relates to several strands of the literature. First, there is a very scarce literature on external monitoring in joint-liability credit groups. To the best of our knowledge, the only paper that explicitly addresses this question is Chowdhury (2005) who shows that external monitoring helps crowd-in cheaper peer monitoring. Strategic complementarity in monitoring efforts drives this result: a group member disciplined by monitoring has greater incentive to monitor his peers because more is at stake in case they default. A key assumption in Chowdhury's theoretical investigation is that the informational advantage of peers makes peer monitoring substantially cheaper than external monitoring. As mentioned above, in our specific setting, the informational advantage of internal monitors is not obvious and this helps us shed light on the potential costs of peer monitoring that are absent from Chowdhury's model (and typically overlooked in the literature), namely the costs of whistleblowing and cronyism.[2]

More directly related to our endeavor but situated in other contexts, a couple of papers have examined the effects of a change in the identity of the monitor on the consequences of monitoring. Vanderwalle (2017) shows that, in the case of self-help groups in India, the involvement of an external accountant instead of an internal accountant leads to a more equitable distribution of group benefits. Breza and Chandrasekhar (2015) conducted an experiment that aimed at stimulating savings in rural communities of India, where local monitors check saving balances. They show that the effectiveness of a monitor depends on her network position: when monitors are more central or closer to the saver, saving increases. This conclusion apparently is at odds with our own (closer monitors are less efficient), yet the two results may actually be the two sides of the same coin. In their setting, Breza and Chandrasekhar argue that the information collected by the monitor potentially translates into a *positive* signal for the community. In our case, however, monitors have to report *negative* information to the group and denouncing close acquaintances may be particularly costly. By implication, praising close acquaintances may be particularly beneficial.

Turning to the effect of social ties on the performance of credit groups, theory suggests that opposite effects are at play. On the one hand, social ties facilitate information

[2]In any case, our setting could not allow strategic complementarity to materialize because the specific timing of the intervention prevented large responses in terms of investment in cotton production and additional internal monitoring could not further reduce moral-hazard problems. Another notable difference with Chowdhury's setting is that we are considering monitoring by the leader and not strictly peer monitoring. To the best of our knowledge, there is no analysis of such two-tier monitoring structure (by an external agent and by a group leader) in the literature.

flows and provide an opportunity to use social sanctions. On the other hand, they may hinder sanctioning by raising its cost. In the context of microfinance, Karlan (2007) and Kassar (2007) find that the overall effect of social ties on repayment is positive. Wydick (1999) argues otherwise: group members with strong social ties may be less eager to enforce repayment. Hermes (2005) mentions that members of joint-liability credit groups may be reluctant to use pressure for fear of losing family or friends and Ahlin (2007) provides evidence that social ties negatively affect repayment rates. As for Ghatak and Guinnane (2009), they contend that joint-liability may not fully address moral-hazard when sanctions are too costly. Recent evidence suggests that this negative effect of social ties on loan performance is not restricted to developing country contexts. Thus, Haselmann, Schoenherr and Vig (2018) show that favoritism towards members of the same social network may result in a sub-optimal allocation of credit by commercial banks in Germany. They conducted a deep and detailed analysis of the allocation of credit between banks and firms whose officials are members of the same elite club. When the firm's CEO and the firm's banker are members of the same club branch, the bank appears more likely to finance the firm. Here again, a "reluctance to sanction" is at play: the result is driven by greater leniency in the renewal of loans to ailing firms. One contribution of our study is to show that external monitoring helps to suppress this bias in the context of joint-liability credit groups.

A last relevant strand of the literature examines the role of sanctioning in overcoming information asymmetries in experimental games. Thus Ostrom et al. (1992) and Gächter and Fehr (2000) point to the positive role of sanctioning on cooperation in public good games. However, both sanctioning and cooperation are reduced when counter-punishment are allowed for (Nikiforakis, 2008), or when groups include people prone to retaliation (Ones and Putterman, 2007).

The remainder of the paper is organized as follows. Section 2 describes the data and provides institutional details regarding the functioning of the credit groups. Section 3 describes the experiment and the expected impacts. Section 4 presents the quantitative results which are discussed in Section 5 in the light of complementary qualitative insights. Section 6 concludes.

2 Data and institutional background

2.1 Data

First-hand data were collected for 890 cotton farmers in 40 villages in the area of Houndé in the South-West of Burkina Faso.[3] These farmers belong to 71 different joint-liability credit groups. In each group we interviewed 13 randomly chosen farmers, who were surveyed farmers in 2013, at the end of the agricultural season and following the harvest. We also conducted a systematic survey of all group leaders (71). The main purpose of this survey was to draw a list of all defaults (including of non-interviewed farmers) based on the group's book-keeping records. In addition to collecting general information about the functioning of the group, we asked detailed questions about the circumstances and consequences of each default. We also obtained administrative data by digitizing the books of the credit groups.[4] These books follow a template provided by the cotton company and include information on cotton production and credit for each group member. The main variable thus obtained that we use in the quantitative analysis is the "area planted to cotton". It is on the basis of this area that standard loan packages are supplied. Finally in 2015, we conducted long semi-structured interviews with five group leaders and five agents of the cotton company to inquire about their relative roles when it comes to sanctioning wrongdoing or contesting the loan demands of individual members. Figure 1 summarizes the timing of the quantitative surveys.

Table 1 presents descriptive statistics for the sample of farmers. Standards of living are low, and less than one fourth of farmers went to school. Households are formed of 8.5 individuals above 6, on average. They cultivate an average of 3.9 hectares of cotton and 4.5 hectares of cereals.

2.2 Cotton group structure

In Burkina Faso, cotton producers are organized in groups that receive joint-liability loans, pool their input purchases, and jointly sell their cotton to the local parastatal cotton company (a local monopsony). These groups consist of 8 to 79 producers (34 on average in our sample) living in the same village. In the area of study, some villages (32%) count only one group, while the majority have at most 3 groups (58%). Larger villages, however, can have multiple groups (up to 12 in our sample).

Groups are headed by a president who is typically a respected elder. In his day-to-day management tasks he is helped by a secretary who is often the most educated

[3]The data collection was part of a research project, funded by USAID, to evaluate a new insurance for these cotton farmers.

[4]For three groups, the book only contained information for one year because the registry was new and the old one was not accessible. In one case, the book had been lost and could not be digitized.

member of the group. The president and the secretary are the group's representatives vis-à-vis the cotton company. In charge of the group's administration, they supervise the admission of new members, control book-keeping and manage the group's loans. They organize at least two plenary meetings per year. During the first meeting they collect and discuss each member's credit demand for the following campaign while during the second meeting they distribute the proceeds from the sales of cotton and settle problems of individual default that arose during the current campaign.

2.3 Credit contracts

The bank supplying loans works in close relationship with the cotton company. The loans are disbursed exclusively in kind, in the form of cotton inputs delivered by the cotton company (seeds, fertilizers, herbicides and pesticides). The loan size is proportional to cotton area because the cotton company offers a standardized package of input per hectare.

The group loans are collateralized by the group's future cotton production. The joint-liability clause is strictly enforced : the cotton company pays the sale proceeds directly to the group's bank account after having duly subtracted the amount of debt outstanding.

The credit contract with the bank stipulates that group defaults are sanctioned by exclusion, which would imply that farmers are de facto forbidden to produce cotton (except if they are accepted in another group). Group defaults are in fact extremely rare and in practice not immediately sanctioned by exclusion.[5]

2.4 Individual default and moral hazard

While group defaults are rare, individual defaults are frequent. An individual farmer defaults when the value of the cotton he produced is lower than his share of the group loan.[6] In our sample 12.5% of farmers were found in this situation after the 2013 campaign (which was not particularly bad).

The high rate of individual default is, in part, related to the variability of cotton yields in the context of the Sudano-Sahelian climate, which is characterized by erratic rains and very variable levels of pest infestation. In fact, when queried about the causes

[5]Defaulting groups are often offered "a second chance" and are carefully monitored by the local agent of the cotton company. For instance, one defaulting group in the study area was temporarily denied a loan, after which it was required to follow a strict plan to reimburse the outstanding debt.

[6]Note that this definition implies that a farmer can default even when he uses another source of income to reimburse the loan. We define defaults in this way because any credit that is reimbursed using non-cotton revenue is still a concern for the group. As we discuss below, having to reimburse the credit immediately (using another source of income) is generally considered as a sanction and imposing this obligation can therefore create tensions in the group.

of defaults over the last year, group leaders reported that 13% of them were related to either an excess or a deficit of rain. (See Table 7 in the Appendix for an overview of the causes of default). However, the most common reasons for default, as reported to us, relate to factors under a farmer's control, namely labour and chemical inputs applications. The problem is that, because farmers do not have access to credit for their other crops (mainly cereals), they are keen to divert part of their chemical inputs (fertilizers, herbicides and pesticides) to these crops.[7] Another form of input diversion implies labour: 22% of default cases are attributed to a lack of labour effort, typically to the benefit of other crops. Overall, ex-ante moral hazard appears to be a pervasive feature among cotton producers.[8] By contrast, ex-post moral hazard appears to be very rare.[9]

In the case of individual default, the group's total cotton revenue covers the individual defaulter's deficit and it is the responsibility of the group (and in particular of its leader) to decide how the debt is settled. In most cases, the group's members who have covered the debt wait until the next campaign, when the defaulter's new cotton production is used to reimburse his debt to the group. Sometimes the group requests immediate repayment which the defaulter typically finances by selling foodgrain or animals (which is perceived as a severe sanction). In extreme cases, an individual can be permanently or temporarily excluded from the credit group.[10]

Interviews of leaders and group members suggest that the management of default is the most challenging aspect of cotton production. Enforcing repayment may be difficult and individual defaults generate tensions among group members who typically have close social ties. Interestingly, managers of the cotton company also indicate that individual defaults are a continuous source of concern for them, even when it does not cause group defaults. The reason is that individual defaults within joint-liability credit groups may discourage investment in cotton production because they decrease its profitability (and increase its variability), in particular for "good producers" who

[7]Inputs obtained through the cotton loan are also sometimes sold on the market, but this appears much less common than diverting them to other crops.

[8] Individual farmers may find it profitable to divert inputs away from cotton. However, both lenders and group leaders mention that they fight against input diversion (through monitoring in particular) because it raises the likelihood of loan defaults, which represent the most challenging aspect of credit group management (see below).

[9]There is very limited scope for ex-post moral hazard in this scheme, whereby farmers would fail to reimburse their loan despite having produced sufficient amounts of cotton. They could in principle do this by engaging in pirate sales, but this is made difficult by the fact that the cotton company is the only purchaser of cotton. This type of behavior was never mentioned by the many field actors we have interacted with.

[10]In less than 1% of the cases did the leader mention that the defaulter was excluded from the group. Furthermore there is about the same proportion (6%) of defaulters and non-defaulters who stopped cultivating cotton in 2014 in the control villages, confirming that very few defaulters were involuntarily excluded.

are never defaulting.

2.5 Information and monitoring

The literature on joint-liability credit suggests that group monitoring should limit opportunistic behavior such as ex-ante moral hazard. At first sight monitoring may appear particularly easy in the context of cotton groups. First, group members all produce cotton and thus have first-hand knowledge that helps them gauge the state of a crop and detect under-application of inputs. Second, farmers live close to each other and should therefore be able to visit each other's field at low cost. Direct peer-monitoring is however less widespread than expected. It appears difficult for a regular group member to make a courtesy visit to a peer's field, as those visits are perceived as intrusive. Our information on field visits in the control group in 2014 indicates that only 33% of farmers received a visit of any of their (non-leader) peers on their cotton field.[11]

In contrast, group leaders are actively engaged in monitoring: in the same sample 59% of sampled farmers have received the visit of a group leader. When visiting a field they always inspect the state of the crop and often measure its area. These two forms of monitoring allow to detect different forms of input diversion. Inspecting the field allows to detect under application of inputs on a given area while measuring can reveal a difference between the cotton area declared for loan demand and the planted area (which also frees up inputs to apply to other crops). During interviews leaders indicated that field visits help them deal with defaulting farmers since it enables them identify the cause of default. However, even visits by group leaders may give rise to frustrations. Some of them explained that their visit may offend the farmer who will interpret it as a sign that the leader does not trust him anymore. To avoid that suspicion, some leaders apply equal treatment by visiting every farmer in the group once in a campaign.

The cotton company also engages in direct monitoring of individual farmers. Local extension agents of the cotton company (ATC) not only meet with the group leaders and review with them the administrative data contained in the book but also visit and measure some farmers' fields during the cotton campaign. Overall, over the campaign, extension agents visit 37% of group members on average and they measure the visited fields of 24% of farmers, in line with the recommendation of the cotton company. These visits have several objectives. They enable agents to detect cases of credit diversion and to give advice regarding pest management. Managers of the cotton company perceive this intense monitoring as necessary, helping them to minimize internal group tensions susceptible of eventually harming cotton production. In fact, they are steadily increas-

[11] It would be interesting to investigate the effect of the intervention on peer monitoring rather than on leaders' monitoring only. Unfortunately we do not have information on peer monitoring in 2013, the year in which the intervention was implemented.

ing the level of monitoring and, in the near future, they intend to measure every cotton field every single year and to collect loan and repayment data for individual farmers, rather than only the group aggregates.

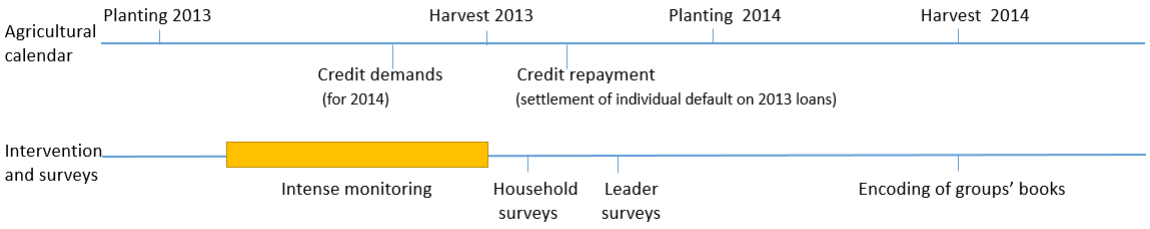
In our area of study, agents and group leaders are generally on good terms. Leaders and group members show a lot of respect for the agent. While the agent is not a member of the rural community he works with, he lives in a nearby village and leaders often go and talk to him about practical matters or problems related to the management of their group. Revealingly, both the agents and the leaders speak of themselves as "jointly" managing the group. For instance, some agents mentioned that they choose which farmers to visit together with the leaders of the group. This cooperation is facilitated by the fact that their incentives are well aligned: both see it as their goal to avoid defaults and conflicts in the group. The agent additionally tries to maximize cotton production (which determines the company's profits), which could induce him to get tougher regarding input diversion. Yet, agents and leaders agree that limited amounts of input diversion are acceptable and that fighting against excessive moral hazard is the priority.[12]

2.6 Timing of the cotton campaign

A peculiar aspect of the credit contracts is that loan demands precede the reimbursement of the loans taken during the previous campaign. In other words, decisions about loan size for the next campaign are taken before the cotton of the current campaign is harvested. Specifically a plenary group meeting takes place a couple of months before harvest and during this meeting, each farmer of the group announces the number of hectares of cotton that he intends to plant. Since loan size per hectare is largely fixed by the cotton company, the area planted to cotton determines it. During the meeting, the group has to formally agree with each farmer's demand and the aggregate demand is then transmitted to the cotton company and the bank. In some cases, the loan size may be adjusted later but the cotton company strongly discourages adjustments because they complicate the logistic of in-kind disbursements. As a result, in practice, the subsequent loan disbursement is not contingent upon repayment. The above timing implies that decisions over future loan size are taken before individual defaults are established for the current campaign. The timeline in Figure 1 illustrates the timing of the events.

[12]The incentives of the extension agent are thus also quite similar to the incentives of a typical credit officer. Both care about loan repayment and about increasing the size of loans when possible. The difference is that the agent also has an interest in increasing expected yields beyond levels at which default becomes impossible.

Figure 1: Timeline of agricultural activities, the intervention and surveys



3 Experimental design and expected outcomes

Experimental design

The experiment took place in 2013 and consisted of increasing the level of external monitoring in a random sample of credit groups.[13] Randomization occurred at the village level and we selected 20 monitored villages and 20 control villages. In a monitored village all interviewed groups were monitored while in a control village no group was. We randomly selected the groups to be interviewed, and the sample consists of 31 monitored groups and 40 control groups.[14]

The monitoring treatment involved an increase in the number of field visits by the extension agent in charge of the group. As discussed above, extension agents are always involved in group monitoring so that the effect of our intervention was to increase the intensity of this external monitoring. Specifically, in treatment groups the agent was instructed to visit the group every 10 days for three months, from the middle to the end of the campaign. The intervention came as a surprise, both for agents and for farmers. At every visit, the agent had to visit the fields of two farmers of the group and to record some information (status of the crop, status of the field, recommendations given). He could freely choose which farmers to visit, but needed to visit different farmers every time. In control groups, no such instructions were given.

As has already been stressed, an important aspect of the intervention is that it

[13]Some information relative to the 2013 campaign - such as the settlement of default - was collected in 2014.

[14]When a village hosted one or two groups, all groups were interviewed. In "big villages" (holding at least 3 groups), three groups were chosen at random in control villages and two groups were chosen at random in monitored villages. This explains why there are fewer monitored groups than control groups. This also implies that treatment is only random conditional on the type of village (big or small). In other words, we have a stratified random sample. We control in all tests and regressions for this stratification variable (whether the group belongs to a village that hosts three or more groups). Additionally, since we sample 13 farmers in each group, the sampling probability of an individual depends on the size of the group. For this reason, we also control for group size when using the household survey sample (See Solon et. al. (2015) for a discussion).

started half-way during the agricultural campaign when many input decisions had already been made (cotton was planted and fertilizer applications had largely taken place). Because farmers and agents were not informed of the intervention in advance, the experiment did not affect decisions taken before it started. Actual visits then happened until the end of the campaign (Figure 1).

The sample appears well-balanced across treatment and control areas. Table 2 reports differences for key pre-determined variables for the sample of surveyed households (upper-panel) and for all group members (lower-panel). The only significant differences are whether the household lives in a village with five or more groups, whether the household owns a motorbike and whether the household was self-sufficient in cereal production for the last 3 years, that is, produced enough cereals for its own consumption every year. In all regressions, we control for these systematic differences.[15]

Expected impacts of the intervention

As detailed above, the monitoring technology is rather simple: both the agent and the group leader measure the size of cotton fields and inspect the state of the crop to detect input diversion. This implies that the identity of the monitor visiting a given farmer should not change the probability to detect moral hazard. It may however affect the way this information is used in group meetings where individual default is sanctioned or credit is allocated. Indeed we believe that the main difference between external and internal monitoring lays in the cost of using the collected information to constrain or sanction wrongdoers: an internal monitor is likely to be more reluctant to point finger at specific members of his group than an external monitor. Under this assumption, external monitoring may change the sanctioning regime and the allocation of credit. We highlight these effects below. In Section 5, we discuss alternative explanations for our set of results.

The first outcome of interest is the impact of an increase in external monitoring on internal monitoring by the group leader. Because both monitors collect the same type of information during their field visit, we expect a crowding-out effect on the leader's field visits. The experiment's immediate outcome would then be a change in the identity of the monitor.

What are the expected consequences of a change in the identity of the monitor? If different monitors face different cost of acting upon the information they collect - for

[15]Note that the variable "Self-sufficient in cereal production" is missing for 38 households. This is because these are newly formed households who have just started their agricultural production and for whom this variable is undefined. We still include these households in the analysis and systematically control for the variable "Self-sufficient in cereal production" to be missing, that is, for whether households are newly formed.

example different costs of sanctioning - a change in monitor may affect the punishment of shirking behavior. Isolating the effect of the monitor's identity on the sanctioning regime is a priori very difficult, as group members would respond to the change in monitor by adapting their behavior. If for example, thanks to the new sanctioning regime, external monitoring succeeds in deterring moral hazard, we would not observe the new sanctions at equilibrium. The specific setting of our experiment offers a possibility to overcome this difficulty: because external monitoring started after input allocation, it should have no effect on moral hazard and output. We can therefore explore the consequences of a change in the identity of the monitor on future loan allocation and on the sanctioning of default.

The first step, before turning to these latter two outcomes, is thus to verify that the intervention did not affect production decisions. Towards this purpose, we estimate the impact of the intervention on various measures, including cotton yield, individual default, the risk of default and, when default occurred, the cause of default. After establishing that the intervention did not modified farmer's productive behavior, we examine the impact of external monitoring on loan size for the next campaign, keeping in mind that decisions over future loan size take place before the repayment of current loans but after more intensive external monitoring.

If the external monitor voices his suspicions of input diversion more easily than the group leader, wrongdoers may face new constraints in their access to credit: the group would sanction their behavior to reduce its exposure to individual default. In addition to this supply effect, the new monitoring regime may also affect demands for loans, albeit in an ambiguous manner. A decrease in individual default reduces the cost of borrowing for non-defaulters, which may stimulate demand. On the other hand, farmers who would otherwise divert part of their loan for other uses may demand less credit as a result of external monitoring. For these farmers, the supply and demand effects go in the same direction and external monitoring should reduce future loan size. In other words, farmers whose potential diversion plans are more constrained as a result of the intervention should experience a larger decrease in loan size. To explore this possibility empirically, we estimate the heterogeneous impact of the intervention on members of the leader's family. The rationale is that the leader's cost of reporting wrongdoing may be especially large when members of his family are involved. Therefore, the negative effect of external monitoring on loan size is expected to be especially important for these members.

Finally, we explore the effect of the experiment on the sanctioning of default. We first provide further evidence that default cases are not different across treatment and control groups. This is necessary to attribute any change in the handling of default to a change in the sanctioning regime (rather than to a selection effect). What are the expected

impacts of external monitoring on the sanctioning regime? If wrongdoing is more easily established when the information is collected by the external agent, individual default caused by shirking should be more harshly punished. In contrast, default caused by external events (beyond the farmer's control) may be treated with more leniency. This is because when cases of wrongdoing are systematically exposed in public, the simple fact of not exposing other cases implies that the responsibility of the defaulted may be disclaimed. In short, we may expect an increase in the state-dependency of sanctioning. Empirically we thus explore the impact of the intervention on sanctioning, distinguishing between defaults that are likely to have arisen from shirking and other defaults. We also examine whether the impact is different for members of the leader's family, expecting a greater strengthening of sanctions for them.

4 Results

All results are estimated using ordinary least square regressions with standard errors clustered at the village level and controls for unbalanced characteristics.[16]

Field visits and measurements: group leaders are crowded out

Table 3 presents the impacts of the intervention on the probability that a farmer's field is visited or measured by the agent (columns (1) and (2)), by the group's leader (columns (3) and (4)) or by either one or the other (columns (5) and (6)). First, in line with the experimental design, agents' visits are more frequent in treatment villages: the share of farmers who received a visit of the extension agent over the campaign increased by 14.6 percentage points as a result of the intervention. Agents also increased field measurements by 12.8 percentage points.

Second, leaders substantially reduced their visits and field measurements, in almost the same proportion as external agents increased their monitoring: leaders reduced their visits by 15.3 percentage points and their measurements by 13.4 percentage points. External monitoring thus seems to have entirely substituted for internal monitoring by group leaders. In fact, there is no change in the overall probability to receive a field visit or measurement whether by the agent or the leader.

While the overall probability to be monitored has not changed, the intervention may have changed the composition of the pool of farmers who are visited. To explore this possibility, in Table 4 we compare the characteristics of monitored farmers in control and in treatment villages. Columns (5) and (6) reveal that farmers monitored by either

[16]Given the limited number of clusters, we have also done all these regressions using wild bootstrap with the same level of clustering. This does not affect the results.

the agent or the leader have similar characteristics in the two types of villages.

Taken together, these results indicate that the intervention did not increase the probability of a field visit or measurement (by either a leader or an agent) nor did it affect the type of farmers visited (by either a leader or an agent). Monitoring is however more likely to have been undertaken by the agent than by the leader. In other words, the intervention experimentally shocked who is doing the monitoring. We now investigate how this substitution of internal for external monitoring changed production and credit outcomes.

Ex ante moral hazard: production and defaulting are unchanged

Let us now look at the impact of the intervention on farmers' production and defaulting. As mentioned above, the monitoring intervention was designed as a surprise so that farmer's production behavior (or ex-ante moral-hazard) should not have been affected. Table 5 confirms this expectation: both cotton yields and the probability that a farmer defaults - his cotton revenue is smaller than his debt - are very similar in treatment and in control samples (columns 1 and 2). Likewise, the same proportion of farmers is "at risk of default" in both samples (column 3). A farmer is "at risk of default" if his cotton output barely enabled him to reimburse his loan or, more formally, if a one standard deviation reduction in his yield would have led to default.[17] Finally, defaults have the same probability to be caused by factors under a farmer's control (column 4). To construct the dependent variable of this last regression, we use the leaders' descriptions of the causes of default for all cases of defaults (318), and we designate as shirking the instances where default might have been avoided if the farmer had invested more time, energy or inputs into cotton production. [18]A note of caution is needed here: given the lack of precision of our estimation we cannot rule out small changes in behavior. For instance, the 95% confidence interval of the effect on cotton yields ranges from -74 to 87 kg/ha. With an average cotton yield of 736 kg/ha, we can only reject a reduction or an increase of more than 10% in the yields. What we can conclude, however, is that the intervention did not have a large effect on production behavior (or moral-hazard) during the year of the intervention.

While the intervention has not affected the propensity of default, it may have

[17]The measure of standard deviations is based on estimates of farmers' individual yield distribution, assuming farmers have different expected yields but the same variance in yields.

[18]Two remarks are in order. Because this classification may appear arbitrary in some instances, we constructed three measures of shirking that are increasingly stringent: 70% of defaults are considered to be the result of shirking with the first definition, 42% with the second and 34% with the third. The results remain similar whichever the definition used. Table 7 in the Appendix presents the frequency of each cause and their classification using the three definitions. Because the leaders themselves report the causes of default, a legitimate question is whether their ability to detect shirking was not directly affected by the intervention. We take up this quesiton in the last sub-section (page 18).

changed the type of farmers who defaulted. Table 9 and 10 in Appendix suggest that this is not the case. When we compare the characteristics of farmers who defaulted during the year of the intervention across the treatment and control groups (Table 9), we find no significant difference. The same conclusion holds when we restrict the comparison to default attributed to shirking: farmers who defaulted for this reason share similar characteristics in control and in treatment villages (Table 10). This provides additional support to the claim that farmers' production behavior has not been affected by the intervention.

Loan renewal: loans are smaller with external monitoring

We now explore whether more intense external monitoring changed group members' access to credit. As interest rates and maturity are fixed by the lender and are the same for all group members, we focus on loan size. Recall that loan size is decided during a plenary group meeting that takes place before harvest and thus before the repayment of the previous loan (Figure 1). As explained earlier, the effect of the intervention on loan size is a priori ambiguous, yet, we expect a larger negative effect for members of the leader's family.

Table 6 reports the impacts of the intervention on different measures of credit outcomes for all group members (book-keeping sample). The dependent variable in columns (1) and (2) is a binary variable indicating whether the farm's loan size decreased during the post-intervention year. It is constructed by comparing the area under cotton in 2013 (which corresponds to the area planted when the intervention occurred) and the same area in 2014 (which has been negotiated towards the end of the 2013 campaign, but before harvest). In columns (3) and (4) we investigate the impact of the intervention on the change in loan size, proxied by the change in the cotton area (per hectare loan size is standardized).[19] Finally, columns (5) and (6) report the impact of the intervention on a binary variable indicating whether the farmer stopped planting cotton in 2014.

While columns (1), (3) and (5) of Table 6 report the average treatment effect, columns (2), (4) and (6) distinguish between group members who belong to the leader's family and other members. In practice, the variable "leader kin" takes value one if the farmer has the same surname as the group's leader. Sharing the surname of the group's leader indicates that the member is likely to belong to the same patri-kin as the leader but does not enable to detect family links through the mother or the wife of the farmer. However, in this strongly patriarchal society, a man's primary responsibilities concern

[19]We choose to use the area declared for cotton in the book of the group rather than loan size expressed in monetary terms because the latter variable is less systematically recorded in the books and more noisy. Cotton area for credit is also the information transmitted to the cotton company by the group.

members of his paternal family.[20]

Compared to those in the control groups, farmers in the treatment groups are 12 percentage points more likely to experience a decrease in their credit between 2013 and 2014 (column 1). Using a continuous measure of cotton area also indicates a larger decrease in loan size in treatment groups, albeit not significant. The interaction with "leader kin" suggests that these effects are mostly driven by a reduction in credit for members of the leader's family. For the subsample of farmers from the leader's kin, the increase in external monitoring translates into a 19 percentage point increase in the probability to experience a significant decrease in credit size, and into a decrease of 0.7 hectares in the area under cotton. Results on the probability to stop cultivating cotton are not statistically significant (columns 5 and 6), which is not surprising given the small incidence of this event. Yet the sign of the coefficients suggests that the intervention increased the probability to terminate cotton production, for members of the leader's family in particular.

Overall, our results indicate that the change in the monitor's identity caused a reduction in the size of new loans, especially for members of the leader's family.

Sanctioning of individual defaults with external monitoring

If external and internal monitors face different costs of denouncing wrongdoing, the intervention may affect how cases of individual default are settled within the group. As described earlier, joint-liability is strictly enforced by the lender who directly subtracts the group's total debt from the sale proceeds. It then falls under the responsibility of the leader to settle cases of individual default.

Table 8 reports the impact of the intervention on the probability for a defaulter to be asked to return his debt immediately. The variable "Paid on spot" takes on value one if the defaulter had to reimburse right after (or at) the group meeting during which sale proceeds were distributed and defaults settled. We use two subsamples of defaulters: the sample of all farmers whose cotton income was smaller than their debt in the control and treatment groups (318 farmers), and the subsample of them who were surveyed (118 farmers). The advantage of using the second subsample is that we can construct an indicator of the farmers' closeness to other groups members. In particular we build a Hamiltonian index that captures the (genetic) distance between a farmer and the group members liable for his debt.[21] The premise is that, much like in

[20]There is no heterogeneous impact of the program (for members the leader's family) on the probability to be monitored or on current yields and default (Table 11 in appendix). This confirms that there is no change in the composition of the pool of monitored farmers (only the identity of the monitor changes). Furthermore, the intervention did not have a differential impact on the moral-hazardous behavior of members of the leader's family during the year of the intervention.

[21]This information was elicited by first asking the respondent whether specific members had been

the case of the leader's kin, social norms or peculiar incentives may come to play when sanctioning connected members is at stake.

The results reported in column (1) indicate that the intervention decreased by 16.4 percentage points the probability that a defaulter is asked to immediately reimburse his debt. Furthermore, while in control groups members of the leader's kin were systematically less likely to have to return their debt on the spot, they were treated like any other group members when the external agents provided intense monitoring: the coefficient on the interaction between "leader's kin" and "monitored" nearly exactly compensates the negative coefficient on "leader's kin" (column 2). Much in the same line, defaulters who share stronger blood ties with farmers liable for their debt were less likely to reimburse on the spot in control groups. By contrast, they appear to have been treated like any other defaulters in the treatment groups (column (4)).

As expected, an interesting contrast emerges when we investigate the correlation between immediate repayment and the causes of default (Column 3). With more intense external monitoring, sanctioning is more state-dependent, that is, it depends on whether the default arises from factors under the farmer's control.[22] Inside externally monitored groups, reduction in the immediate repayment requirement is concentrated on cases of default that do not involve shirking: such cases are now 32 percentage points less likely to be sanctioned by the above requirement. This effect disappears for default cases attributed to shirking (the coefficient on the interaction compensates the coefficient on "Monitored").[23] In short, while external monitoring is associated with a more lenient handling of individual default cases, it leads to more severe punishment of defaulters accused of shirking (compared to other defaulters in treatment villages). The aggregate effect is thus an increase in the state-dependency of sanctioning.

Threats to identification

Before turning to potential explanations for our results, we discuss the potential threats to identification, particularly in the analysis of individual default. First, and most

designated to pay back his debt in case of default. If the answer was positive, the respondent was asked about his relationship with each of these specific members. If the answer was negative, the whole group would be asked to cover his debt. Three group members were then randomly chosen in the list of members (by the computer used for the survey) and the respondent was asked about his relationship with each of them. In practice fathers, sons and brothers receive a weight of 0.5; grand-father, grand-son, uncle and nephew receive a weight of 0.25; cousin and other family members receive a weight of 0.125.

[22] A possible problem with using a measure like "Shirked" is ex-post rationalization: we may worry that the reported cause of default depends on whether or not the farmer was asked to pay on the spot. However this would be a confounding factor only if ex-post rationalization is itself affected by the intervention, which is not very likely.

[23] The results hold for the other classification of default, but are not statistically significant (not shown).

importantly, we worry that leaders' reports of the causes for default may be less reliable in treatment groups precisely because they did less monitoring. This concern applies even if the likelihood to report shirking is similar in treatment and in control groups. We mentioned above that leaders and agents tightly collaborate and we will provide additional qualitative evidence in Section 5 that suggests that they share information about individual members' performances. Additional quantitative evidence from our leaders' surveys also suggests that the level of information of leaders has not suffered from their reduced involvement in direct monitoring. Specifically we systematically asked leaders, for each case of default, whether they had anticipated the default before the cotton harvest. This data indicates that the likelihood that the leader expected default to occur is not significantly different across treatment and control groups (and, if anything, larger in treatment group). Similarly, the likelihood that the leaders admitted to have tried to avoid the problem is not different across treatment and control groups (Table 14 in Appendix). Finally, reports by external agents confirm that they shared with the leaders the information gathered during their visit. We thus asked external monitors to fill in, right after their field visit, a form where they were asked to report who accompanied them during their visit and whether they discussed its outcome with someone immediately after it took place. In as many as 54% of the cases, agents were either accompanied by a group leader, or talked to a leader immediately after the field visit. Unfortunately, the agents were not asked to report whether they talked about the field visit at a later date (when they next saw the leader), but the high incidence of immediate discussions is by itself a strong indicator that information gathered by the external monitor is often reported to the group leaders. Overall, therefore, we are rather confident that the intervention did not impair the leaders' information about the causes of default.

Second, sample selection may be an issue since our regressions only involve farmers who defaulted. Defaulters could be different in treatment and control samples if the intervention affected moral hazardous behavior. We have argued above that the intervention came largely as a surprise and left little time to change behavior. Additionally, the effects on moral hazard show point estimates that are very close to zero (Table 5). In Table 9 in the Appendix we also compare the characteristics of defaulters in control and treatment and show that they are very similar. We thus find no evidence for a differential selection between control and treatment, and we interpret our results as indicating that external monitoring affects the sanctioning regime applied in credit groups.[24]

Another concern relates to the state-dependency of sanctioning: sanctioning of farm-

[24]Since we are interested in sanctioning following default, this is the natural subsample. We cannot rule out, however, that non-defaulters would have been affected differently, had they defaulted.

ers varies depending upon the (presumed) cause of default. Here, there are two potential problems. First, the treatment could have affected the selection of who defaults because of misbehavior (or who is classified as such). Again, moral hazard does not seem to be affected by the treatment and, additionally, shirkers are quite similar in control and treatment (see Table 10 in the Appendix). The second problem is that, while monitoring is randomized, shirking is not. Heterogeneous effects could therefore be generated by any variable that (1) is correlated with shirking and (2) causes heterogeneous effects in sanctioning as a result of the treatment.[25] While there are many reasons to believe that shirking defaulters are systematically different from other defaulters, Table 12 in the Appendix suggests that the two groups are very similar. If we are willing to accept that the observed similarity of these different subgroups implies similarity on unobservables, our results point to a causal effect of the intervention on the state-dependency of sanctioning.

Finally, our interpretation of the heterogeneous effects obtained for the members of the leader's family raises a similar concern: if these individuals are systematically different from other group members, these differences may confound the effects of the variable "Leader kin" in our regressions. Table 13 indeed suggests that in our household sample, members of the leader's family are younger, less likely to own a bed or a television, more likely to live in a house with solid walls and to cultivate a larger area (both in cotton and cereals). As a result we cannot claim causality for the results on the "Leader kin" variable, but only conclude that the intervention affected farmers sharing similar characteristics with the leader's family in a different manner from other farmers.

5 Discussion

Taken together our results show that external monitoring crowds out internal monitoring by the group leader. This substitution is not neutral and modifies both credit renewal decisions and the sanctioning of delinquent loans. A more intense external monitoring leads to a firmer handling of renewal decisions, especially those of members of the leader's family (or members who share similar characteristics with members of the leader's family). Simultaneously, more leniency is displayed in case of defaulting be-

[25]Consider this illustrative example. Suppose that educated farmers are more likely both to divert inputs and to challenge the information collected by the group leader in their field (which is used to decide of the sanction). Assume that they cannot challenge the information collected by the external agent so that their sanctioning regime changes as a result of the intervention. In that case, the interaction terms in Table 8 would capture the heterogeneous effect of education and not (strictly) that of shirking. In that case, we would need to control for the interaction between education and the treatment. When we add interactions between the treatment and various control variables in the subsample of farmers who took part in the household survey, the point estimates on the interaction between shirking and monitoring remain similar (results not shown).

yond the control of the member. In addition, the handling of delinquent loans becomes less biased in favor of members of the leader's family.

These results are in line with our main hypothesis regarding the cost of denouncing wrongdoing (see Section 3). In this section, we discuss whether alternative mechanisms may account for our set of result and we provide additional descriptive evidence in favour of our preferred mechanism. We then turn to the question of whether a substitution of external monitoring for leaders' monitoring enables a more efficient management of credit groups.

Alternative mechanism

We have argued that the group leader and the agent collect the same type of information during their field visits. However, part of our results would hold under the alternative assumption that the agent has a greater expertise and collects more precise information. In particular, if agents are better at detecting input diversion or predicting yields, they may trigger a more state-dependent sanctioning of default. Agents' monitoring may then also have a negative impact on future loan size because they better identify excessive demands for loans by individual farmers. This said, the strong monitoring skills of the agent cannot directly account for the heterogeneous effect observed for members of the leader's kin. In addition, during interviews, agents indicated that they do not get more information from a field visit than the leader would. They believe that the group leaders, who are experienced cotton farmers themselves, are just as competent at detecting signs of input diversion or lack of labour application. Also, group leaders have the advantage of knowing the historical record of individual farmers while the rapid rotation of extension agents makes it very difficult for them to acquire this knowledge. In the light of this evidence, we do not believe that, alone, the hypothesis of different monitoring skills can provide a satisfactory explanation for our results.[26]

Additional evidence in support of the enforcement hypothesis

If the extension agent does not learn more from his visit to a given field than the group leader does, the fact that the agent, instead of the group leader, made the visit makes a difference. Our preferred explanation for this difference is that the information collected by the agent is less costly to use when it comes to deciding future loan sizes and sanctioning detected cases of default. Our leaders' survey and our semi-structured interviews with leaders and agents lend support to this explanation.

[26]The possibility that the intervention affected the composition of the group of farmers who are monitored (by either the leader or the external agent) was discussed in the results section. While we cannot formally rule this possibility out, we find no indication of such a selection effect.

When queried about the reasons invoked to oppose an excessive loan demand, leaders systematically mentioned information collected during field visits. According to them, it is difficult to ask a farmer to reduce his cotton area or to request the immediate repayment of a delinquent loan. To be effective, such moves must typically be based on evidence of the lack of diligence of the farmer, which only a field visit by a leader or an agent can provide. This is confirmed by the result of the leaders' survey. Thus, 77% of all leaders said that their field visit is helpful when they have to deal with "moral hazardous" behavior. Interestingly this number further increases to 88% for a visit by the extension agent. One external agent explained that his intervention is particularly effective when "powerful" group members have to be disciplined: if a group leader thinks it necessary to oppose the credit demand of a powerful member, he informs the agent who collects information and prepares arguments to speak against this person during the group's plenary meeting. In contrast, according to both agents and leaders, ordinary members do not directly use their knowledge of individual cases to require more severe treatment regarding loan demands or sanctions. The reason is that ordinary members "do not want to risk to make an enemy." If they wish to denounce shirking, they address themselves privately to the group's leader, suggesting that he visits the field of the farmer suspected.

Even external agents sometimes use intricate strategies to avoid that monitored farmers suspect the leader to have sent them to their field. For instance, one agent described the case of a leader who suspected the group's secretary to massively divert inputs. After having been informed by the leader, the agent asked the secretary to accompany him on several visits. During one of these visits, he asked to visit the fields of the leader and the group's accountant, which "helped justify his demand to see his field as well". The secretary did not take him to his own field, but to the cotton field of another member. The agent acted as if he had not noticed. This field was however much smaller than the area officially declared by the secretary. Two weeks later, the agent told the secretary he had decided to measure all the cotton fields of the group. After measuring the field alleged to belong to the secretary, the agent told the latter that the area was much smaller than it was supposed to be, indicating input diversion. He informed the secretary that he would report this fact to the group's leader who then used the information to request an immediate repayment of the delinquent loan (the secretary had to sell an ox to pay back the loan). The agent said: "Coming from me, the information could be used to sanction the secretary". This anecdote corroborates the claim that information gathered by the agent is "easier to use" to ensure effective enforcement of the rule.[27]

[27]In exceptional cases, the agent is even present during the second plenary meeting where defaults are settled. One agent thus mentioned that leaders sometimes ask him to intervene directly to sanction

A last piece of qualitative evidence supporting the idea that sanctioning is difficult to enforce internally comes from the answers that agents and leaders gave to the open question: "What are the characteristics of a good leader?" All ten respondents who participated in our in-depth interviews in 2015 first mentioned "being tough", or "being able to sanction." This again suggests that while sanctioning is necessary, it is not an easy task for a leader.

In short, our interviews suggest that external monitoring helps when it comes to discussing loan size and to sanctioning. There are two possible explanations for more effective enforcement when information is collected by an external agent. Either the agent has more legitimacy than the leader and the information he collects is perceived as less biased. Or, it is less costly for the agent to "point fingers" and for the leader to be content with repeating the agent's allegation of wrongdoing than to take direct responsibility for it. These two explanations are not mutually exclusive and they are likely to be both valid. Together, they also help understand the heterogeneous effects of the intervention on members of the leader's family. Indeed, the leader's bias is likely to be towards favoring members of his kin, while the cost of "pointing fingers" is no doubt higher when members of one's family are concerned. Overall, our analysis of credit groups in Burkina Faso suggests that in contexts of closely linked members, a key obstacle in overcoming ex-ante moral hazard lies in the enforcement of the rules rather than in getting evidence of rule violations. State-dependent sanctioning and tight screening of new loan demands appear to be facilitated by an external monitor.

Welfare implications

What do our results suggest regarding the performance of joint-liability credit groups?

While we cannot provide a definite answer to this question, both our quantitative and qualitative evidence suggest that the group is managed more efficiently when the external agent monitors more intensively. First, the fact that members of the leader's family are less systematically favored in their access to credit point to a fairer and more efficient management. Second, greater leniency in sanctioning is also compatible with greater efficiency. This is because less stringent requirement of immediate repayment only occurs in case of "legitimate default" and not when borrowers shirked. Rather than being the result of a more lax attitude on the part of the leader, the new sanctioning regime improves the insurance value of the loan: farmers who experienced a negative shock and default for reasons beyond their control are less likely to be sanctioned. This implies that risk-sharing across group members de facto increases, which might increase

a member. When doing so he needs to be prudent and motivate the sanction on objective grounds because otherwise "there will be trouble in the village the moment I leave". A field visit provides the objective information that he needs.

the willingness to make profitable but risky investments.[28]

Quantifying these benefits would require a longer-term intervention to measure its impacts on moral-hazard and risk-taking. This is outside the scope of the current analysis. The costs, on the other hand, can be estimated and are relatively small: implementing the intervention would require an increase in the interest rate of about 0.23 percentage points.[29] Costs are low because the loans are big (they finance the entire cotton production of the group) and because the intervention is small, requiring only one visit to some group members. Therefore, it is highly likely that, in our case, the benefits outweigh the costs.

6 Conclusion

We have argued that a substitution of external for internal monitoring in joint liability credit groups facilitates group discipline, suggesting that external monitoring is more efficient than internal monitoring in our setting. While it is impossible to do a cost-benefit analysis that compares these forms of monitoring, our evidence indicates that external monitoring in joint-liability groups can be extremely useful to help enforce rules and sanctions. This helps to understand why it is frequently, and increasingly, used in our context as well as in other credit groups.

A legitimate question is whether joint-liability with external monitoring has any advantage over individual liability: if external monitoring, which is costly, is required, even in joint-liability groups, the cost advantage of the latter seems to vanish. What we argue is that such conclusion is unwarranted because it bypasses the issue of sanctioning detected wrongdoing. To be effective, sanctioning of default behavior requires a combination of external monitoring (to detect and expose wrongdoing) and group level punishment (to implement sanctions). Joint-liability in combination with external monitoring may thus be needed for group members to be in a position to adequately use social sanctions to enforce good behavior. This is all the more important as high levels of social capital are likely to be positively correlated with high costs of applying punishment: it is precisely when social sanctions are feasible that their application may

[28]This insurance aspect of joint-liability, and its impact on risk-taking has been studied by Fischer (2013) and Gine (2010) theoretically and through the use of experimental games. Their work indicates that risk-sharing through joint-liability does not necessarily benefit all members (relatively risk-averse borrowers may, for example, suffer from greater risk-taking by their peers).

[29]We do not have precise information on the exact costs involved, but we can provide some rough estimates. Agents need to visit each group 8 times. Visiting the farmers (by motorbike) requires 1 liter of gas (750 CFA). The daily wage, including maintenance of the motorbike, is about 5000 CFA and an agent can visit 2 groups per day. The total cost involved in implementing the monitoring intervention for one group is thus about 26 000 CFA (or 40 Euro). By contrast, the average loan size of the groups in our sample is about 11.3 million CFA. The cost of the intervention thus represents an increase of 0.23 percentage points of the interest rate.

be prohibitively costly.

If our explanation is correct, our results also imply that simply increasing information can not always solve the incentive problems arising from information asymmetry. Even if group leaders were perfectly informed about the actions of the members, that would not solve moral hazard problems insofar as they cannot (or do not want to) act upon this information. The intervention of an external monitor, by avoiding the obstacles faced by indigenous leaders, facilitates effective application of the rules in the event of detected rule violations.

References

Banerjee, A. V., T. Besley, and T. W. Guinnane (1994). Thy neighbor's keeper: The design of a credit cooperative with theory and a test. *The Quarterly Journal of Economics 109*(2), 491–515.

Breza, E. and A. G. Chandrasekhar (2015). Social networks, reputation and commitment: Evidence from a savings monitors experiment. Technical report, National Bureau of Economic Research.

Cason, T. N., L. Gangadharan, and P. Maitra (2012). Moral hazard and peer monitoring in a laboratory microfinance experiment. *Journal of Economic Behavior & Organization 82*(1), 192 – 209.

Chowdhury, P. R. (2005). Group-lending: Sequential financing, lender monitoring and joint liability. *Journal of Development Economics 77*(2), 415–439.

De Quidt, J., T. R. Fetzer, and M. Ghatak (2016a). Commercialization and the decline of joint liability microcredit. *mimeo*.

De Quidt, J., T. R. Fetzer, and M. Ghatak (2016b). Market structure and borrower welfare in microfinance. *mimeo*.

Fischer, G. (2013). Contract structure, risk-sharing, and investment choice. *Econometrica 81*(3), 883–939.

Ghatak, M. (1999). Group lending, local information and peer selection. *Journal of Development Economics 60*(1), 27–50.

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

Giné, X., P. Jakiela, D. Karlan, and J. Morduch (2010). Microfinance games. *American Economic Journal: Applied Economics 2*(3), 60–95.

Giné, X., D. Karlan, and M. Ngatia (2014). Social networks, financial literacy and index insurance. In M. Lundberg and F. Mulaj (Eds.), *Enhancing financial capability and behaviour in low- and middle-income countries*. World Bank, Washington, DC.

Haselmann, R. F., D. Schoenherr, and V. Vig (2018). Rent-seeking in elite networks. *Journal of Political Economy*.

Madajewicz, M. (2011). Joint liability versus individual liability in credit contracts. *Journal of Economic Behavior & Organization 77*(2), 107–123.

Schreiner, M. (2012). Une carte de score de pauvreté au Burkina Faso. *Technical Report*.

Solon, G., S. J. Haider, and J. M. Wooldridge (2015). What are we weighting for? *Journal of Human Resources 50*(2), 301–316.

Stiglitz, J. E. (1990). Peer monitoring and credit markets. *World Bank Economic Review 4*(3), 351–66.

Main tables

Table 1: Descriptive statistics

	Mean	Sd	Median	N
Characteristics household head and Household				
Age household head	43.61573	12.85161	42	890
Nr. years of education household head	1.204494	2.493362	0	890
Household head has any education	.2404494	.4275967	0	890
Nr. years head manages household	15.5748	11.5828	12	889
Nr. years household head belong to credit group	10.05293	6.35288	10	888
Household size (above 6 years old)	8.505618	5.251651	7	890
Consumption indicator (Progress out of Poverty Index, PPI)	36.66854	12.53419	36	890
Self-sufficient in cereal production (last 3 years)	.6115023	.487695	1	852
Average Hamiltonian index with group members	.2124084	.1795223	.1875	889
Agricultural production				
Cotton area cultivated 2013 (ha)	3.865618	3.405493	3	890
Cotton yield 2013 (kg/ha)	826.7994	344.363	800	890
Average cotton yield (2008-2012)	876.7581	255.4646	859.9	851
Cultivated GM cotton 2013	.5505618	.4977166	1	890
Cereal area cultivated 2013 (ha)	4.512277	3.197328	4	887
Credit for cotton production				
Had a loan for cotton 2013	.994382	.0747844	1	890
Total value of cotton production (in CFA)	808247.4	903720.4	520877.5	890
Cotton area declared for credit (ha)	4.049045	3.682221	3	890
Total credit for cotton production (in CFA)	399963	396861.6	275000	877
Defaulted (individual default)	.1254276	.3313921	0	877
At risk of default (default if yield decreased by 1 sd)	.2827822	.4506084	0	877
Monitoring in credit group				
Agent visited field	.5112613	.5001549	1	888
Agent measured field	.3617978	.4807909	0	890
Leader visited field	.4208543	.4940066	0	796
Leader measured field	.1559748	.3630599	0	795
Leader or agent visited field	.7166247	.4509211	1	794
Leader or agent measured field	.4402516	.4967298	0	795

Data source: Household surveys.

Notes: Sample sizes differ across variables. Out of the 890 farmers interviewed, 38 were in newly formed households and could not report historical cereal and cotton yields, 13 did not know their credit amount (which is used to calculate defaults), although they did know their credit area. Information about leader monitoring was obtained from the leaders, who were only asked detailed questions for a random subsample of 796 farmers. All other missing values are due to individual farmers not knowing the reply (or not wanting to reply) to a particular question

Table 2: Table of balance using pre-determined characteristics

	Control	sd	Diff: T - C	se	N
HOUSEHOLD SAMPLE					
Characteristics household					
Age household head	44.57	(12.72)	-0.439	(1.361)	890
Nr. years of education household head	1.320	(2.379)	0.236	(0.244)	890
Household head has any education	0.269	(0.413)	0.0505	(0.0481)	890
Nr. years head manages household	17.69	(11.86)	-0.243	(0.863)	889
Nr. years head is part of credit group	10.13	(6.512)	-0.393	(0.640)	888
Household size (above 6 years old)	9.107	(5.422)	-0.828	(0.519)	890
Household wealth and food security					
Household has television	0.255	(0.517)	0.0295	(0.0437)	890
Household has bed	0.757	(1.458)	-0.0692	(0.135)	890
Household has moto	0.774	(1.195)	-0.193**	(0.0905)	890
Household has house with solid walls	0.150	(0.354)	0.0346	(0.0493)	890
Consumption indicator (PPI index)	36.74	(12.30)	-0.377	(1.487)	890
Self-sufficient in cereal (last 3 years)	0.553	(0.477)	-0.0929*	(0.0493)	852
Any meal reduced (last year)	0.181	(0.393)	0.0243	(0.0314)	890
Any meal skipped (last year)	0.114	(0.291)	0.0187	(0.0244)	890
Agricultural production					
Cotton area cultivated 2013	4.494	(3.505)	-0.222	(0.490)	890
Average cotton yield (2008-2012)	855.0	(260.9)	-21.57	(32.06)	851
Cultivated GM cotton 2013	0.663	(0.498)	0.007	(0.106)	890
Cereal area cultivated 2013	4.882	(3.126)	-0.116	(0.348)	887
Credit for cotton (household survey)					
Cotton area for credit	4.474	(3.764)	-0.249	(0.531)	890
Total credit for cotton production (in CFA)	448870.4	(424258.6)	-50154.1	(58923.5)	877
Five or more groups in village	0.013	(0)	0.159*	(0.0839)	890
ADMINISTRATIVE DATA					
Number of producers in group	41.66	(15.38)	6.804	(6.042)	2022
Same kin as leader (same family name)	0.356	(0.493)	0.024	(0.092)	2022
Cotton area declared for credit 2013	3.569	(3.655)	-0.037	(0.482)	2022
Cotton area declared for credit 2012	3.787	(3.886)	0.012	(0.536)	1346
Individual default 2012	0.118	(0.329)	0.015	(0.027)	1345

Data source: Household surveys and administrative data (bookeeping records).
Notes: The first column reports the constant in a regression where the variable of interest is regressed on the treatment ("Monitored") and the stratification variables (see footnote 14). The third column reports the coefficient on "Monitored" from the same regression. Standard errors are clustered at the village level. Sample sizes differ across variables. Out of the 890 farmers interviewed, 38 were in newly formed households and could not report historical cereal and cotton yields. 13 did not know their credit amount, which is also used to calculate defaults, although they did know their credit area. All other missing are due to individual farmers not knowing the reply (or not wanting to reply) to a particular question. The "Consumption Indicator" is computed using ten indicators and has been shown to be an appropriate indicator of poverty in Burkina Faso (Schreiner 2012). The average of the index correspond to a poverty likelihood of 17% (using the national poverty line).

Table 3: Impact of the experiment on external and internal monitoring

	(1) Agent visit	(2) Agent measurement	(3) Leader Visit	(4) Leader measurement	(5) Agent or leader visit	(6) Agent or leader measurement
Monitored	0.146**	0.128**	-0.153	-0.134*	0.0198	0.0250
	(0.0597)	(0.0608)	(0.0968)	(0.0662)	(0.0574)	(0.0751)
Constant	0.526***	0.396***	0.279**	0.149*	0.643***	0.435***
	(0.0813)	(0.0742)	(0.124)	(0.0869)	(0.0612)	(0.0914)
N	888	890	796	795	794	795

Data source: Household surveys and leader surveys for the construction of the dependent variables "leader visit" and "leader measurement".

Notes: Standard errors, clustered at the village level, in parentheses. Regressions control for the stratification variables as well as for unbalanced characteristics: Whether the household has a moto, is self-sufficient in cereal production (last 3 years), is newly formed and belongs to a village hosting five or more groups. There are fewer observations about leader monitoring because this information comes from the leader interview, who were asked only about a random subset of interviewed farmers. Other differences in sample sizes are because of respondents not knowing the answer. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Characteristics of individuals visited by leader or external agents in treatment and control

	Agent visit		Leader visit		Agent or leader visit	
	(1) C	(2) T-C (se)	(3) C	(4) T-C (se)	(5) C	(6) T-C (se)
Characteristics household						
Age household head	46.16	-1.43 (1.57)	46.56	-0.65 (1.99)	45.76	-0.74 (1.49)
Nr. years of education household head	1.69	-0.10 (0.29)	1.23	0.02 (0.21)	1.49	0.21 (0.23)
Household head has any education	0.29	0.01 (0.05)	0.23	0.04 (0.05)	0.27	0.06 (0.05)
Nr. years head manages household	18.22	-1.24 (1.24)	18.92	0.12 (1.28)	18.13	-0.61 (0.99)
Nr. years head is part of credit group	11.67	-0.69 (0.74)	10.02	-0.09 (0.72)	10.45	-0.10 (0.67)
Household size (above 6 years old)	10.95	-0.74 (0.63)	10.24	-1.73*** (0.58)	9.90	-0.75 (0.60)
Same kin as leader (same family name)	0.12	-0.03 (0.11)	0.31	-0.04 (0.08)	0.16	0.01 (0.09)
Household wealth and food security						
Household has television	0.35	-0.05 (0.05)	0.29	0.02 (0.06)	0.30	0.03 (0.05)
Household has bed	1.11	-0.12 (0.17)	1.00	-0.17 (0.19)	0.91	-0.10 (0.14)
Household has moto	0.98	-0.24* (0.12)	0.89	-0.24* (0.12)	0.86	-0.20* (0.10)
Household has house with solid walls	0.21	-0.01 (0.05)	0.18	0.03 (0.08)	0.18	0.02 (0.05)
Consumption indicator (PPI index)	38.94	-1.80 (1.66)	38.41	-0.47 (2.01)	37.84	-0.41 (1.55)
Self-sufficient in cereal (last 3 years)	0.58	-0.10* (0.06)	0.60	-0.03 (0.08)	0.55	-0.06 (0.06)
Any meal reduced (last year)	0.09	0.02 (0.04)	0.25	-0.04 (0.05)	0.19	-0.00 (0.04)
Any meal skipped (last year)	0.07	-0.01 (0.02)	0.12	-0.02 (0.03)	0.12	-0.01 (0.02)
Agricultural production						
Cotton area cultivated 2013	6.24	-0.94 (0.60)	4.32	-0.18 (0.60)	4.96	-0.16 (0.58)
Average cotton yield (2008-2012)	872.90	-48.51 (40.23)	919.47	-27.45 (43.54)	881.95	-28.89 (37.31)
Cultivated GM cotton 2013	0.82	-0.11 (0.12)	0.75	0.03 (0.11)	0.74	-0.02 (0.11)
Cereal area cultivated 2013	6.00	-0.35 (0.44)	4.55	-0.04 (0.36)	5.04	0.04 (0.41)
Credit for cotton						
Cotton area declared for credit	5.87	-0.84 (0.66)	4.64	-0.41 (0.68)	4.89	-0.17 (0.63)
Total credit for cotton production (in CFA)	606042	-126854* (68104)	486985	-45111 (76298)	500746	-50784 (68736)

Data source: Household surveys and leader surveys for the construction of the dependent variables "leader visit" and "leader measurement".

Notes: The columns "C" report the constant in a regression where the variable of interest is regressed on "visit" and the stratification variables. The columns "T-C" reports the coefficient on "visit" from the same regression. Standard errors are clustered at the village level.

Table 5: Impact of external monitoring on (current) yields and defaults

	(1)	(2)	(3)	(4)
	Yield	Default	At risk of default	Shirked
Monitored	2.518	-0.0275	0.000896	0.0266
	(39.72)	(0.0287)	(0.0475)	(0.0992)
Constant	749.4***	0.186***	0.409***	0.700***
	(56.22)	(0.0467)	(0.0695)	(0.0839)
N	890	877	877	319

Data source: Household surveys and leader surveys (in particular for the construction of the dependent variable "Shirked", available for the 319 defaulting farmers).
Notes: "At risk of default" = 1 if a reduction in yield of one standard deviation would have caused default. "Shirked" = 1 if the default was due to a lack of farmer's diligence (see page 15 for a discussion). Standard errors, clustered at the village level, in parentheses. Regressions control for the stratification variables as well as for unbalanced characteristics: whether the household has a moto, is self-sufficient in cereal production (last 3 years), is newly formed and belongs to a village hosting five or more groups. There are fewer observations for defaults than for yields because some farmers did not know their exact credit amount.

Table 6: Impact of external monitoring on (future) credit sizes

	(1)	(2)	(3)	(4)	(5)	(6)
	Decrease (binary)	Decrease (binary)	Change (in ha)	Change (in ha)	Stop (binary)	Stop (binary)
Monitored	0.117**	0.0375	-0.276	0.0146	0.0480	0.00000292
	(0.0498)	(0.0593)	(0.197)	(0.204)	(0.0499)	(0.0467)
Leader kin		-0.113***		0.273**		-0.0794**
		(0.0302)		(0.134)		(0.0294)
Monitored * Leader kin		0.190***		-0.682**		0.115
		(0.0610)		(0.253)		(0.0942)
Constant	0.394***	0.442***	-0.372**	-0.499***	0.235***	0.268***
	(0.0495)	(0.0494)	(0.164)	(0.162)	(0.0360)	(0.0391)
N	2022	2022	2022	2022	2022	2022

Data source: Administrative data (bookkeeping records).
Notes: "Decrease" =1 if the farmer's cotton area decreased following the intervention, "Change" is the difference between the cotton area in 2014 and in 2013, "Stop" =1 if the farmer stopped cultivating cotton in 2014, "Leader kin" =1 if member has the same family name as the leader of the group. Standard errors, clustered at the village level, in parentheses. Regressions control for the stratification variable.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Comparison of causes of default in treatment and control and definition of own fault variables

	Control	sd	Diff: T - C	se	N	Shirked	Shirked (alt def 1)	Shirked (alt def 2)
Did not use enough inputs	0.106	(0.264)	0.102	(0.0753)	310	X	X	X
Did not work enough	0.235	(0.382)	0.0218	(0.0611)	310	X	X	X
Area cultivated too small	0.0101	(0.162)	-0.0299	(0.0207)	310	X	X	X
Bad inputs	0.0114	(0.103)	-0.0106	(0.00709)	310	X	X	
Parasites	0.00548	(0.145)	-0.0163	(0.0187)	310	X	X	
Planted late	0.0128	(0.145)	-0.0149	(0.0201)	310	X	X	
Weeds	0.0548	(0.215)	-0.0468**	(0.0186)	310	X	X	
Bad land	0.0697	(0.177)	0.0483	(0.0347)	310	X		
Not enough labour available	0.194	(0.403)	-0.0181	(0.101)	310	X		
Other	0.00915	(0.177)	-0.00398	(0.0323)	310	X		
Flooding or drought	0.133	(0.425)	0.0215	(0.107)	310			
Lost livestock	0.0900	(0.236)	-0.0122	(0.0422)	310			
Sickness	0.0641	(0.226)	-0.0514***	(0.0182)	310			
Fire	0.00422	(0.0731)	0.0107	(0.0114)	310			

Data source: Leader surveys.

Notes: The column "Control" reports the constant in a regression where the variable of interest is regressed on the treatment ("Monitored") and the stratification variable. The columns "T-C" reports the coefficient on "Monitored" from the same regression. Standard errors are clustered at the village level.

Table 8: Impact of external monitoring on the sanctioning of default

	(1)	(2)	(3)	(4)
	Paid on spot	Paid on spot	Paid on spot	Paid on spot
Monitored	-0.164**	-0.234**	-0.322***	-0.379***
	(0.0802)	(0.0902)	(0.113)	(0.108)
Leader kin		-0.177**		
		(0.0667)		
Monitored*Leader kin		0.188*		
		(0.111)		
Shirked			-0.141	
			(0.122)	
Monitored*Shirked			0.248*	
			(0.128)	
Hamiltonian				-0.713***
				(0.212)
Monitored*Hamiltonian				0.762**
				(0.308)
Constant	0.428***	0.496***	0.511***	0.408**
	(0.102)	(0.111)	(0.144)	(0.161)
N	318	318	318	118

Data Source: Leader surveys and, for the variable "Hamiltonian", household surveys.
Notes: "Paid on stop" is defined for all defaulting households (318), "Hamiltonian" only for defaulting households also belonging to the household sample (118). "Hamiltonian" is an index of genetic distance between the defaulting farmer and the group members liable for his debt (see page 17 for a discussion). "Shirked" = 1 if the default was due to a lack of farmer's diligence (see page 15 for a discussion). "Leader kin" =1 if member has the same family name as the leader of the group. Standard errors, clustered at the village level, in parentheses. Regressions control for the stratification variables as well as for unbalanced characteristics: whether the household has a moto, is self-sufficient in cereal production (last 3 years), is newly formed and belongs to a village hosting five or more groups. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix

Table 9: Comparison of defaulters in treatment and control (household survey sample)

	Control	sd	Diff: T - C	se	N
Characteristics household head and households					
Age household head	46.57	(12.32)	-3.716**	(1.720)	119
Nr. years of education household head	2.143	(2.275)	0.557	(0.550)	119
Household head has any education	0.329	(0.336)	0.102	(0.0822)	119
Nr. years head manages household	19.63	(11.14)	-0.798	(1.684)	119
Nr. years head belongs to credit group	8.915	(6.212)	0.0979	(1.386)	118
Household size (above 6 years old)	7.054	(4.188)	-0.928	(1.023)	119
Same kin as leader (same family name)	-0.0231	(0.475)	-0.179	(0.133)	119
Wealth and food security HH					
Household has television	-0.0137	(0.458)	0.00473	(0.0929)	119
Household has bed	0.145	(1.044)	0.0766	(0.201)	119
Household has moto	0.309	(0.840)	-0.108	(0.184)	119
Household has house with solid walls	0.0839	(0.177)	0.0955	(0.0628)	119
Consumption indicator (PPI index)	29.93	(11.69)	-0.562	(2.735)	119
Self-sufficient in cereal (last 3 years)	0.199	(0.504)	-0.0412	(0.0848)	115
Any meal reduced (last year)	0.457	(0.429)	0.0936	(0.0783)	119
Any meal skipped (last year)	0.220	(0.272)	0.0932	(0.0588)	119
Agricultural production					
Cotton area cultivated 2013	2.269	(1.877)	0.481	(0.671)	119
Average cotton yield (2008-2012)	843.6	(255.1)	-19.72	(47.43)	111
Cultivated GM cotton 2013	0.400	(0.502)	-0.0587	(0.112)	119
Cereal area cultivated 2013	3.275	(2.279)	0.406	(0.651)	119
Credit for cotton production					
Cotton area declared for credit	2.334	(2.955)	0.0981	(0.837)	119
Total credit for cotton production (in CFA)	275661.8	(377866.7)	-39180.9	(103571.7)	114

Data source: Household surveys

Notes: The first column reports the constant in a regression where the variable of interest is regressed on the treatment ("Monitored") and the stratification variables (see footnote 14). The third column reports the coefficient on "Monitored" from the same regression. Standard errors are clustered at the village level. See the note below Table 1 for an explanation of the missing variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Comparison of people defaulting because of shirking in treatment and control (household survey sample)

	Control	sd	Diff: T - C	se	N
Characteristics household head and households					
Age household head	49.40	(12.99)	-5.617**	(2.505)	75
Nr. years of education household head	1.558	(1.767)	1.000	(0.592)	75
Household head has any education	0.222	(0.273)	0.190*	(0.0945)	75
Nr. years head manages household	22.01	(12.10)	-1.869	(2.347)	75
Nr. years head belongs to credit group	9.609	(6.189)	-1.153	(1.756)	74
Household size (above 6 years old)	6.876	(4.759)	-1.098	(1.177)	75
Same kin as leader (same family name)	-0.0630	(0.471)	-0.176	(0.134)	75
Wealth and food security HH					
Household has television	0.0600	(0.563)	-0.0202	(0.118)	75
Household has bed	0.296	(0.847)	0.199	(0.235)	75
Household has moto	0.399	(0.887)	-0.225	(0.212)	75
Household has house with solid walls	0.0413	(0.162)	0.0547	(0.0514)	75
Consumption indicator (PPI index)	32.11	(11.72)	-1.268	(3.246)	75
Self-sufficient in cereal (last 3 years)	0.259	(0.500)	-0.0905	(0.118)	73
Any meal reduced (last year)	0.454	(0.446)	0.0862	(0.111)	75
Any meal skipped (last year)	0.131	(0.273)	0.109*	(0.0626)	75
Agricultural production					
Cotton area cultivated 2013	1.827	(2.063)	0.180	(0.781)	75
Average cotton yield (2008-2012)	855.7	(234.0)	-3.912	(41.91)	71
Cultivated GM cotton 2013	0.396	(0.500)	-0.0438	(0.134)	75
Cereal area cultivated 2013	3.218	(2.530)	-0.0761	(0.727)	75
Credit for cotton production					
Cotton area declared for credit	2.330	(3.538)	-0.532	(1.074)	75
Total credit for cotton production (in CFA)	191802.5	(429868.0)	-80170.6	(124678.1)	73

Data source: Leader's survey for the identification of defaults caused by Shirking (see Table 7 for an exact definition) and household surveys for individual characteristics.

Notes: The first column reports the constant in a regression where the variable of interest is regressed on the treatment ("Monitored") and the stratification variables (see footnote 14). The third column reports the coefficient on "Monitored" from the same regression. Standard errors are clustered at the village level. See the note below Table 1 for an explanation of the missing variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Monitoring, yields and default: heterogeneous impacts for leader's kins

	(1) Agent or leader visit	(2) Agent or leader measurement	(3) Yield	(4) Default	(5) At risk of default	(6) Shirked
Monitored	0.029	0.013	-4.297	-0.009	0.013	0.053
	(0.048)	(0.080)	(46.08)	(0.030)	(0.053)	(0.126)
Leader kin	-0.064	-0.018	-42.14	0.092**	0.104*	0.090
	(0.102)	(0.073)	(44.78)	(0.036)	(0.051)	(0.085)
Monitored * Leader kin	-0.0314	0.066	27.27	-0.077	-0.045	-0.073
	(0.114)	(0.090)	(58.16)	(0.044)	(0.076)	(0.151)
Constant	0.668***	0.469***	758.3***	0.163***	0.390***	0.667***
	(0.061)	(0.089)	(60.13)	(0.046)	(0.070)	(0.100)
N	794	795	890	877	877	319

Data source: Household surveys and leader surveys (in particular for the construction of the dependent variable "Shirked", available for the 319 defaulting farmers).
Notes: "At risk of default" = 1 if a reduction in yield of one standard deviation would have caused default. "Shirked" = 1 if the default was due to a lack of farmer's diligence (see page 15 for a discussion). "Leader kin" =1 if member has the same family name as the leader of the group. Standard errors, clustered at the village level, in parentheses. Regressions control for the stratification variables as well as for unbalanced characteristics: whether the household has a moto, is self-sufficient in cereal production (last 3 years), is newly formed and belongs to a village hosting five or more groups. There are fewer observations for defaults than for yields because some farmers did not know their exact credit amount.

Table 12: Characteristics of farmers who shirked (among defaulters)

	Control	sd	Diff: T - C	se	N
Characteristics household head and households					
Age household head	46.97	(10.61)	-1.807	(2.172)	119
Nr. years of education household head	2.515	(2.715)	-0.192	(0.421)	119
Household head has any education	0.379	(0.370)	-0.0164	(0.0724)	119
Nr. years head manages household	19.55	(8.986)	-0.209	(1.830)	119
Nr. years head belongs to credit group	8.995	(6.560)	-0.0491	(1.222)	118
Household size (above 6 years old)	6.558	(4.123)	0.189	(0.773)	119
Same kin as leader (same family name)	-0.0467	(0.471)	-0.0409	(0.0875)	119
Wealth and food security HH					
Household has television	-0.137	(0.255)	0.134*	(0.0691)	119
Household has bed	0.251	(1.087)	-0.0852	(0.166)	119
Household has moto	0.120	(0.685)	0.163	(0.129)	119
Household has house with solid walls	0.181	(0.321)	-0.0694	(0.0657)	119
Consumption indicator (PPI index)	26.52	(11.54)	3.454	(2.682)	119
Self-sufficient in cereal (last 3 years)	0.491	(0.457)	-0.309***	(0.106)	115
Any meal reduced (last year)	0.443	(0.424)	0.0504	(0.0907)	119
Any meal skipped (last year)	0.231	(0.291)	0.0225	(0.0565)	119
Agricultural production					
Cotton area cultivated 2013	2.704	(2.882)	-0.289	(0.548)	119
Average cotton yield (2008-2012)	836.9	(264.7)	-0.877	(50.79)	111
Cultivated GM cotton 2013	0.450	(0.505)	-0.0760	(0.0924)	119
Cereal area cultivated 2013	3.594	(3.172)	-0.192	(0.591)	119
Credit for cotton production					
Cotton area declared for credit	2.400	(2.644)	-0.0347	(0.692)	119
Total credit for cotton production (in CFA)	314314.7	(314010.6)	-55298.3	(82731.6)	114

Data source: Household surveys and leader surveys (in particular for the construction of the variable "Shirked").
Notes:The first column reports the constant in a regression where the variable of interest is regressed on "Shirked" and the stratification variables (see footnote 14). The third column reports the coefficient on "Shirked" from the same regression. Standard errors are clustered at the village level. See the note below Table 1 for an explanation of the missing variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Characteristics of farmers who belong to the leader's family

	Control	sd	Diff: T - C	se	N
Characteristics household head and households					
Age household head	45.56	(13.23)	-1.696*	(0.876)	890
Nr. years of education household head	1.340	(2.463)	0.140	(0.289)	890
Household head has any education	0.274	(0.424)	0.0280	(0.0483)	890
Nr. years head manages household	18.18	(12.05)	-1.423*	(0.836)	889
Nr. years head belongs to credit group	10.01	(6.376)	0.559	(0.604)	888
Household size (above 6 years old)	9.042	(5.270)	0.438	(0.489)	890
Same kin as leader (same family name)	0	(0)	1.000***	(2.82e-17)	890
Wealth and food security HH					
Household has television	0.294	(0.555)	-0.0728*	(0.0389)	890
Household has bed	0.770	(1.424)	-0.273**	(0.104)	890
Household has moto	0.763	(1.100)	-0.0255	(0.0869)	890
Household has house with solid walls	0.152	(0.354)	0.0660**	(0.0290)	890
Consumption indicator (PPI index)	36.70	(12.68)	-0.841	(1.555)	890
Self-sufficient in cereal (last 3 years)	0.506	(0.491)	0.0618	(0.0556)	852
Any meal reduced (last year)	0.175	(0.403)	-0.0112	(0.0342)	890
Any meal skipped (last year)	0.0940	(0.295)	0.0192	(0.0252)	890
Agricultural production					
Cotton area cultivated 2013	4.276	(3.341)	1.123**	(0.427)	890
Average cotton yield (2008-2012)	849.7	(250.9)	41.78	(37.19)	851
Cultivated GM cotton 2013	0.657	(0.500)	0.205**	(0.0758)	890
Cereal area cultivated 2013	4.768	(3.041)	0.813**	(0.337)	887
Credit for cotton production					
Cotton area declared for credit	4.224	(3.538)	1.166**	(0.500)	890
Total credit for cotton production (in CFA)	408232.8	(358984.8)	174125.1***	(60667.7)	877
Monitoring by leader and agent					
Agent visited field	0.654	(0.500)	0.0121	(0.0493)	888
Agent measured field	0.525	(0.480)	0.0247	(0.0416)	890
Leader visited field	0.296	(0.498)	-0.132	(0.105)	796
Leader measured field	0.106	(0.377)	-0.0707	(0.0688)	795
Leader or agent visited field	0.752	(0.443)	-0.0697	(0.0621)	794
Leader or agent measured field	0.520	(0.497)	0.0136	(0.0534)	795

Data source: Household surveys.

Notes: The first column reports the constant in a regression where the variable of interest is regressed on "Leader kin" and the stratification variables (see footnote 14). The third column reports the coefficient on "Leader kin" from the same regression. Standard errors are clustered at the village level. See the note below Table 1 for an explanation of the missing variables. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Leader's information and intervention

	(1) Anticipated default (before harvest)	(2) Anticipated default (before harvest)	(3) Anticipated default (before harvest)	(4) Did something to avoid default
Monitored	0.113	0.114	0.0179	-0.0156
	(0.0682)	(0.0763)	(0.0904)	(0.117)
Leader kin		0.0107		
		(0.0482)		
Monitored*Leader kin		-0.00116		
		(0.0881)		
Shirking			0.0490	
			(0.0396)	
Monitored*Shirking			0.155	
			(0.123)	
Constant	0.0889***	0.0851**	0.0422	0.420**
	(0.0326)	(0.0409)	(0.0394)	(0.169)
N	319	319	319	53

Data Source: Leader surveys.

Notes: "Anticipated default" = 1 if the leader knew the default would occur before cotton was harvested. Only when they knew did we ask whether they did something to avoid default (column 4). Standard errors, clustered at the village level, in parentheses. Regressions control for the stratification variables as well as for whether the household belongs to a village hosting five or more groups.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$