Microcredit and Financial Distress ☆

Joeri Smits, Isabel Günther

Chair of Development Economics, ETH Zurich, Zurich, Switzerland

Abstract

This paper studies the effect of microcredit uptake on household financial distress. Drawing on quasi-experimental household panel survey data collected in urban Uganda, coupled with bank administrative data, we find that on average, microcredit uptake increases household financial distress. These average impacts, however, conceal fundamental heterogeneity in treatment effects for different subpopulations. Households whose members' financial literacy skills are low, and those with more volatile income, lose most on average. These households may make suboptimal borrowing decisions and are especially hard hit by a double trigger (adverse income shock and a failed investment). For households with a ROSCA participant, we fail to reject the null of no distress response to microcredit uptake, suggesting a complementary, consumption smoothing role for ROSCA's to formal credit.

Keywords: microcredit, financial distress, effect heterogeneity, Uganda

1. Introduction

Whereas many poverty-alleviation policies have focused on reducing credit constraints, concerns have increasingly been voiced about potential overborrowing by the poor (e.g. Roodman (2012); Angelucci et al. (2015); Fafchamps

☆This research was supported by the Financial Cooperation Independent Evaluation Unit of the KfW Development Bank. The authors acknowledge project funding from the Independent Evaluation Unit of KfW. We thank Eva Terberger, Franziska Spörri, Thomas Gietzen, Martin Brown, Vincent Somville, Stefan Klonner, Furio Rosati, Ethan Ligon, Marcel Fafchamps, Adán L. Martinez Cruz, Daniel Rozas, and numerous seminar and conference participants for helpful comments. Any errors are of course, our own.

(2013); Schicks (2013)). The aim of this paper is to analyze impacts of microcredit uptake on urban Ugandans households financial distress, and to identify sources of heterogeneity in these impacts. Household financial distress is an outcome not previously included in impact evaluations of microcredit. The risk of overborrowing by households in developing countries is increasingly being recognized as a problem, fuelled by crises in the microfinance sectors of Andhra Pradesh (India) and Bolivia, among others (CSFI, 2012, 2014). Household overborrowing and overindebtedness may also have contributed to the US subprime crisis of 2008 that led to the subsequent global financial crisis. Even in the absence of such crises, overindebtedness of households is a concern for various reasons. On the borrower side, overindebtedness and the associated household financial distress constitutes a welfare loss for households. School dropout has long term consequences and so may distress asset sales in the presence of multiple equilibria due to asset-based poverty traps (Barrett & Carter, 2013; Carter & Barrett, 2006). On the lender side, overborrowing increases credit risk. Despite its critical role in understanding impact pathways of microcredit, surprisingly little research has been conducted on financial distress in the context of microcredit, and to the best of our knowledge, no microcredit impact evaluation has included financial distress in its outcome measures.

We develop a toy model of a household with multiple volatile income streams and which may or may not have access to informal insurance. In the model, financial distress is defined as the shortfall of income with respect to a subsistence level, and can thus be thought of a continuous variable bounded below by zero. The model combines elements from (i) intertemporal choice models with human capital investment choices and (ii) informal insurance, to derive predictions for our empirical analysis. The novelty of the model lies in it considering the borrowing and investment choices of households and their consequences in an environment with multiple stochastic income streams, which is common to most households in developing countries (but is not typically part of intertemporal choice models). The model clarifies two points. First, depending on the household initial conditions in terms of income level, credit uptake may increase financial distress in expectation. Second, the model shows how this effect will vary across households with different (a) degrees of income volatility; (b) access to informal insurance, and (c) financial literacy levels.

The main contribution of this paper however, is empirical: we exploit quasi-experimental household panel survey data collected in urban Uganda to estimate the impact of microcredit uptake on household financial distress. Since the subsistence consumption level will depend on household characteristics and given measurement error challenges with income and consumption data, in the empirical part of this paper we proxy for financial distress by eliciting households' experiencing of discrete 'distress events' among which: the inability to meet health expenses when falling sick, distress assets sales, distress school dropout, cutbacks in food consumption. Working with a Ugandan microfinance deposit taking institution (MDI)[1], At baseline, we observe active microcredit borrowers as well as a control group consisting of loan applicants (for the analysis we exclude those whose applications were rejected). A random subsample of the households is re-interviewed around 13 months later to create a panel dataset. On average, microcredit uptake increases financial distress in the urban Ugandan sample, as predicted. While our quasi-experimental design does not match the rigor of an RCT, a battery of tests and sensitivity analyses, aided by administrative data from the MDI, addresses possible concerns with respect to potential biases of the estimates, which are shown to be robust.

Going beyond averages, fundamental effect heterogeneity is identified, primarily by financial literacy levels, access to informal insurance, and income volatility. We interpret the finding with respect to ROSCA membership as being due to participation in ROSCAs providing access to emergency liquidity to smooth consumption risk against adverse income shocks. For households that are more favorably endowed in one or more of these dimensions - income volatility, financial literacy skills and participation in ROSCAs - we are not able to reject the null of no effect of microcredit uptake on financial distress. The

[1]Explanation of MDI

findings with regard to financial literacy levels suggest a role perhaps for an elicitation of numeracy skills to be part of the credit screening process.

Our findings contribute to two strands of literature. First, our findings contribute to the strand of the household finance literature that analyzes overborrowing and financial distress. Research on household financial distress and overindebtedness in the context of microcredit is scant, nothwitstanding its critical role in understanding impact pathways of microcredit. An exception is Schicks (2013), who explicitly tried to infer the burden debt places by asking respondents about 'sacrifices' respondents had to make in order to meet repayment obligations. In her study, 531 microborrowers in Ghana were asked about a list of 'sacrifices' they had to make in the last 6 months to meet their loan repayment obligation(s) (see Subsection 3.1 for a detailed list). She then asked her sample respondents to classify each of these 'sacrifices' into acceptable and unacceptable sacrifices. A household was deemed overindebted if (a) it experiences a sacrifice that indicates structural problems (an asset seizure, loan recycling or selling/pawning assets) and (b) it makes unacceptable sacrifices repeatedly[2]. While such approach is subjective in nature, its novelty and strength are that it looks at the problem from the client perspective by trying to directly infer the burden of debt across various dimensions of borrowers' lives in a quantitative manner[3]. On the other hand, some borrowers may meet their repayment obligations but at the cost of their households needing to sell productive assets, cut back on education or even miss medical treatments. There are however two, related, concerns with this approach. First, by 'framing' or 'priming' the respondent on their repayment obligations, confirmation bias may result. Second, Schicks only interviewed microcredit borrowers, so the attribu-

[2]Either >3 unacceptable sacrifices, or ≥ 1 unacceptable sacrifice made >3 times.

[3]Supply side measures, such as repayment delays and delinquency, may reflect debt-induced hardship to some extent, but do not provide the whole picture: delinquent borrowers may include those facing a short-term liquidity shortfall without necessarily facing a structurally unsustainable debt burden. Moreover, such figures also capture willful delinquency and default (which is potentially more important for non-collateralized loans and in weak legal systems

tion of the observed distress to microcredit can not be ascertained (which is implicit in that approach). We address the first of these issues by avoiding to frame the - what we refer to as - 'distress events' as repayment struggles, instead framing them more broadly (and calling the resulting index that is used as an outcome measure 'financial distress index').

Second, our findings fit into the literature on microcredit impacts. A series of Randomized Controlled Trial (RCT) studies have investigated the average impacts of microcredit supply expansion in a variety of settings. A meta-analysis of seven randomized trials found that the general impact of microcredit access on on investment in self-employed activities is likely small, but positive (Meager, 2015). A less rigorous summary of results of six RCTs by Banerjee, noted that they did not provide clear evidence of strong effects on higher-level outcomes (household income, consumption, education, health) (Banerjee et al., 2015b). One plausible explanation for the lack of effects is that the often female-owned, small businesses that the households gaining access to microcredit invest in have low marginal product of capital (Crépon et al., 2015). This would explain why these studies often find no significant impact of microcredit access on business profits or income from self-employment activities on average, although several do find an impact on profits for pre-existing businesses or for businesses at the top end of the distribution of profits (Angelucci et al., 2015; Banerjee et al., 2015a; De Mel et al., 2008). An alternative or complementary explanation for the lack of average effects is that microcredit may have opposing effects (in- or decreasing the expected solvency position and welfare of a household) under different household initial conditions, and that cancellation of positive and negative effects within a pooled sample population may result in a lack of treatment effects[4]. It seems plausible that the expected solvency position of a household may respond positively or negatively to microcredit expansion, in a context of

[4]In other words, in contrast to interventions such as improved sanitation or a school meal program, the assumption of *monotone treatment response* seems implausible for the case of microcredit.

low and varying financial literacy levels. And even if for all of those who take up microcredit, the effect would be to improve the *expected* solvency position, the actual solvency status may not, given income shocks and investment risk. Noisy outcome measures in levels (as opposed to events), such as income, profits or expenditure, compound power challenges in randomized trials of microcredit arising from small differences in take-up rates between treated and control(s) (areas) due to the availability of substitutes as well as individual treatment effects of opposite sign. Our discrete-events based outcome may have more power to detect gainers and losers from microcredit. Whereas in the current study, we are not able to identify households who gain from taking up microcredit, this may in part be due to power challenge and be a feature of the particular setting. The effect heterogeneity identified in this study at least suggests the possibility that this may explain the prevalence of null findings that characterize some RCT evaluations of microcredit expansion.

The rest of this paper is organized as follows. The next section develops a stylized theoretical model highlighting several factors that may drive heterogeneity in the effects of microcredit, with a focus on overborrowing and financial distress. Section 3 describes the data and methods used to test the model's predictions. Section 4 reports the estimation results on the mean impacts of microcredit uptake in urban Uganda, with Section 5 describing a range of sensitivity analyses conducted to check the robustness of those estimates. Section 6 reports on results regarding microcredit treatment effect heterogeneity. Finally, Section 7 discusses the results in their relation to the existing literature and concludes.

2. A model of household (over-)borrowing under uncertainty

We develop a simple stylized three-period model of a household making borrowing decisions, while facing various sources of uncertainty. The purpose is to evaluate borrowing decisions and their consequences to derive predictions for the empirical analysis. The base case (Subsection 2.1) assumes unlimited access to emergency credit, unlimited liability, and zero correlation between income volatility and the returns to investment. In Subsection 2.2, the household is assumed not to have access to emergency credit. In Subsection 2.3, we consider other extensions, including limited liability of loan contracts and financial literacy skills.

2.1. The basic model: access to emergency credit

There are four dates {0,0.5,1,2}. In period 0 ($t = 0$), the household is endowed with disposable wealth $W = 0$. Income in period 1, y_1, is a stochastic variable with two possible realizations: $y_1 = y_{high}$ with probability p and $y_1 = y_{low} = 0$ with probability $(1-p)$. For simplicity, formal loans are assumed to be exclusively used for investment. A fixed size formal loan of size L can be taken up in period 0 to enable an investment of the same size I, i.e. $L = I$. The returns to investment are stochastic too: $r = r_{high}$ with probability q (the investment project is succesful) and $r = r_{low} = 0$ with probability $(1-q)$ (the investment fails). The returns to investment r and period 1 income y_1 are revealed in period $t = 0.5$. The returns are defined so as to include the investment itself, so for instance $r = 1$ for a break-even investment (i.e., for an investment with 'zero returns').

The income y_2 in period 2 is deterministic (and known ex-ante). With i_L denoting the interest rate (the interest rate is fixed and known ex-ante) on the formal loan L, we define $i = 1 + i_L$ (and refer hereinafter to i as the 'interest rate'). Assume the principal and simple interest to be repaid in two installments of equal size: the first one at $t = 1$ and the second at $t = 2$. Figure 1 visualizes the timing of the determination of the choice and stochastic variables.

Figure 1: Time line of the determination of choice and stochastic variables.

Let $\bar{c}_1, \bar{c}_2$ denote the expected (from the viewpoint of $t = 0$) consumption in period 1 respectively period 2. Suppose that the household can bring forward period 2 consumption through (interest-free, hence informal) emergency credit b. The three-period model can now be represented as:

$$\begin{array}{ll} \text{t=0:} & L = I \\ \text{t=1:} & \bar{c_1} = p \cdot y_{high} + \frac{1}{2}(q \cdot r_{high} - i) \cdot L + b \\ \text{t=2:} & \bar{c_2} = y_2 + \frac{1}{2}(q \cdot r_{high} - i) \cdot L - b \end{array}$$

which can be summarized as expected life-time consumption (denoted c_{1+1}) equaling

$$\bar{c}_{1+2} = \bar{c_1} + \bar{c_2} = p \cdot y_{high} + y_2 + (q \cdot r_{high} - i) \cdot L \tag{1}$$

Let $c_{min} \geq 0$ denote a subsistence, or minimum acceptable, level of consumption. If $c_1 < c_{min}$, the household takes up emergency credit, with the amount borrowed equaling the shortfall in consumption:

$$\begin{array}{ll} b = c_{min} - (y_1 + \frac{1}{2} \cdot (r - i) \cdot L) & \text{if } c_1 = y_1 + \frac{1}{2} \cdot (r - i) \cdot L < c_{min} \\ b = 0 & \text{if } c_1 = y_1 + \frac{1}{2} \cdot (r - i) \cdot L \geq c_{min} \end{array}$$

For households without a formal loan:

$$\begin{array}{ll} b = c_{min} - y_1 & \text{if } c_1 = y_1 < c_{min} \\ b = 0 & \text{if } c_1 = y_1 \geq c_{min} \end{array}$$

Ex-ante, the household is solvent in expectation if

$$\bar{c}_{1+2} \equiv \bar{c_1} + \bar{c_2} = p \cdot y_{high} + y_2 + (q \cdot r_{high} - i) \cdot L \geq 2 \cdot c_{min} \tag{2}$$

The states of the world in period 1, their probabilities and corresponding realized life-time consumption levels are tabulated in Table 1.

State of the world	Probability	Realized life-time consumption: with loan	without loan
y_{high}, r_{high}	$p \cdot q$	$y_{high} + y_2 + (r_{high} - i) \cdot I$	$y_{high} + y_2$
y_{high}, r_{low}	$p \cdot (1-q)$	$y_{high} + y_2 - i \cdot I$	$y_{high} + y_2$
y_{low}, r_{high}	$(1-p) \cdot (q)$	$y_2 + (r_{high} - i) \cdot I$	y_2
y_{low}, r_{low}	$(1-p) \cdot (1-q)$	$y_2 - i \cdot I$	y_2

Table 1: States of the world and the corresponding realized life-time consumption.

By computing expectations based on Table 1, it follows that ex-ante, the effect of formal borrowing on consumption is positive in expectation if

$$(q \cdot r_{high} - i) \cdot L > 0 \tag{3}$$

We refer to the expression on the left-hand side of equation as the expected net return to formal borrowing (ERB); a rational agent will only borrow if the inequality (3) holds true. We assume throughout this Section that the ERB are positive.

The focus of the current model is household financial distress, which, for operational reasons, we define in the model as lifetime consumption falling below subsistence level, i.e. $c_{1+2} < 2 \cdot c_{min}$. To make predictions about the household's solvency status, assumptions are needed regarding the level of income in its relation to the subsistence level. Fixing the ratio between the loan size and stochastic income to, $y_{high} = 2 \cdot i \cdot I$ and $q \cdot y_{high} = \alpha \cdot i \cdot I$, allows us to separate the analysis according to household income level. We consider the effects of microcredit for three income brackets, depicted in Figure 2. For all income brackets, $0 < \alpha < 1$. Households for whom the income level is such that $y_2 - i \cdot I < 2 \cdot c_{min}$ are never in financial distress, and households whose income is such that $y_{high} + y_2 > 2 \cdot c_{min}$ are always in financial distress, see Figure 2.

1) Not so poor: $y_2 - \alpha \cdot i \cdot I = 2 \cdot c_{min}$

The only state of the world in which the agent is insolvent is y_{low}, r_{low}, when borrowing. Whereas solvency is understood here as a binary concept,

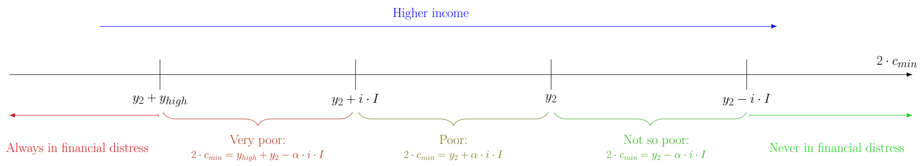

Figure 2: Income line.

the concept of financial distress captures as well the extent of the shortfall in consumption below $2 \cdot c_{min}$, akin to the concept of a poverty gap, or the Foster-Greer Thorbecke poverty with exponent $\alpha \leq 1$. Insolvency is relatively unlikely in this scenario, only occuring when borrowing and having simultaneously an adverse income shock ($y = y_{low} = 0$) and a low return to investment ($r = r_{low} = 0$). The consumption shortfall then equals $(1 - \alpha) \cdot I$, hence the expected value of the increase in the consumption shortfall when borrowing ($L > 0$) vis-a-vis not borrowing ($L = 0$), is

$$(1 - p)(1 - q)(1 - \alpha) \cdot I \tag{4}$$

Hence, when keeping the ratio of loan size to expected income fixed, the level of expected income does not play a role in the expected increase in distress (i.e. consumption shortfall) due to microcredit, only its volatility. To see this, keep α, i, I fixed in the relation $p \cdot y_{high} = \alpha \cdot i \cdot I$. Since y_{high} does not feature in equation (4), an increase in income volatility, i.e., a higher p, leads to an increase in the expected distress effect of borrowing.

The two other income brackets consider households for which the initial conditions are such that income is closer to subsistence level. To save space, and because the conclusions of the other two income brackets are similar to the one presented here, we derive the consumption shortfall expressions with and without microcredit for the other income brackets in Appendix A.

In all three cases (income brackets), the uptake of microcredit increases expected financial distress.

Proposition 1. *The uptake of microcredit ($L > 0$ rather than $L = 0$) in-*

creases financial distress (i.e. the expected value of life-time consumption with respect to $2 \cdot c_{min}$*) in expectation for any non-trivial income level, i.e. whenever* $2 \cdot c_{min} \in (y_2 - \alpha \cdot i \cdot I, y_{high} + y_2 - i \cdot I)$.

The theoretical model also gives predictions regarding two sources of heterogeneity in the effect of microcredit uptake: income level and income volatility. We summarize the model's conclusions with respect to the heterogeneity of the effect of microcredit for those who take it up:

Proposition 2. *a) There is no case where an increase in income volatility, parametrized by* $(1-p)$ *(for a given level of expected income* $p \cdot y_{high}$*), reduces expected financial distress, but there are income brackets (income brackets (1),(3) above) wherein income volatility raises the expected increase in financial distress due to microcredit uptake.*

b) Within some of the income brackts ((1),(3) above), lower expected household income raises the expected increase in financial distress due to microcredit uptake. Since income level and income volatility inter the distress expressions multiplicatively (see equation (4) *and* (15)*, a combination of low and volatile income especially increases the microcredit-distress response.*

2.2. No access to emergency credit

Assume now that the household has no access to emergency credit in period 1. Then future income and investment gains cannot be brought forward, and the household is illiquid but solvent in period 1 if

$$y_1 + y_2 + (r - i) \cdot I \geq 2 \cdot c_{min} \tag{5}$$

but

$$y_1 + \frac{1}{2}(r - i) \cdot I < c_{min} \tag{6}$$

Hence, without access to emergency credit, the household is equally likely to be insolvent as in Case 1 (unlimited access to emergency credit), but it is now more likely to be illiquid in period 1, as period 2 consumption cannot be brought forward through emergency borrowing if needed.

Proposition 3. *Access to emergency credit does not moderate the effect of the uptake of microcredit on household insolvency, but it does increase the expected response of illiquidity to microcredit uptake.*

Whereas both illiquidity and insolvency can be due to a failed investment and/or a negative income shock, income volatility (independent of the loan-induced investment) plays a relatively larger role for illiquidity, whereas investment returns and loan size play a larger role for insolvency.[5] However, in the real-world, illiquidity may have dynamic effects on future solvency status in the presence of multiple equilibria due to for instance asset-based poverty traps (Barrett & Carter, 2013; Carter & Barrett, 2006). Moreover, empirically, it's hard to distinguish between illiquidity and insolvency of a household.

To anticipate the empirical part of this paper, we have data on informal loans households may have outstanding, but we lack variation in their *access* to informal or emergency credit. Controlling for informal loans is not an option: the model suggests that the likelihood of informal borrowing is affected by the uptake a formal loan (a microcredit), as its uptake affects the likelihood of the households' period 1 consumption falling below subsistence level. Including informal loans a household has outstanding as control variable would therefore introduce post-treatment bias.

We do however, have data on whether at least one household member joins (and is member of) a Rotating Savings and Credit Association (ROSCA). Typically, in Sub-Saharan Africa, the order in which members receive the pot of money is not random, but determined endogenously: either agreed on ex-ante

[5]To see this, multiply equation 7 by 2.

by the members, determined by a bidding process, or determined by the ROSCA leader (Baland et al., 2016). Klonner (2003) showed how risk-averse participants in a bidding ROSCA can insure themselves against idiosyncratic risks by smoothing income shocks. Fang et al. (2015) expands on Klonner's earlier work by showing that as long as formal credit markets are imperfect, formal credit and ROSCAs act as complements in consumption smoothing. In an empirical study of Indian ROSCAs with concurrent bidding, Calomiris & Rajaraman (1998) also show that the actual amount received by the winner is subject to variation through the bidding process in a manner consistent with insurance. Baland et al. (2016) do not assume a particular allocation mechanism and ask how people position themselves within ROSCAs which allow for changes in the allotment order. They cite and provide ethnographic evidence that participants in ROSCAs in various countries in Sub-Saharan Africa (Uganda was not included) often prefer not to take an early turn for the pot of money from the ROSCAs but to wait and even go last, whereby waiting provides an insurance function. The model of Baland et al. (2016) predicts that waiting will always be preferred if the likelihood of a shock occurring is large enough.

Combining Proposition 3 with the aforementioned theorectical results on the insurance function of ROSCAs, we will test Proposition of by studying heterogeneity in the effect of microcredit uptake across subsamples of household that participate and households that do not participate in ROSCAs.

2.3. Extensions: Limited liability and behavioral biases

The decision whether or not to apply for a loan was not modeled so far, because all respondents in our sample have selected into a loan: either they already borrow at baseline or they applied for credit at baseline and we could ascertain that their application was not rejected. Under the identifying assumption stated in the Data and Methods Section, the borrowing decision does not have to be modeled to obtain predictions on the average treatment effect on the treated. For an evaluation of treatment effect heterogeneity however, the borrowing decision plays a potentially important role.

A useful and plausible way to think about the borrowing decision, is to assume that the agent will take up the microcredit only if life-time consumption exceeds the subsistence constraint in expectation:

$$\bar{c}_{1+2} = \bar{c_1} + \bar{c_2} = p \cdot y_{high} + y_2 + (q \cdot r_{high} - i) \cdot L \geq 2 \cdot c_{min} \tag{7}$$

We have so far assumed that there is unlimited liability of loan contract enforcement. However, one cause of overborrowing may be limited liability. In the case of limited liability, the household takes into account only the scenario that $y_1 = y_{high}, r = r_{high}$, and anticipates to default strategically in case of other states of the world. The household will then borrow even if

$$\bar{c}_{1+2} = \bar{c_1} + \bar{c_2} = p \cdot y_{high} + y_2 + (q \cdot r_{high} - i) \cdot L < 2 \cdot c_{min} \tag{8}$$

as long as

$$\bar{c}_{1+2} = \bar{c_1} + \bar{c_2} = y_{high} + y_2 + (r_{high} - i) \cdot L \geq 2 \cdot c_{min} \tag{9}$$

Compared to the case of unlimited liability, now households with a smaller difference $q \cdot r_{high} - i$[6] and/or with more volatile income, will select as well into formal credit. Therefore, formal borrowing is (more) likely to increase financial distress in expectation than when liability is unlimited. Liability of microcredit loan contract enforcement in the real world is not binary, of course, so what matters is the degree of liability, and the perceived cost (both pecuniary as well as non-pecuniary) of delinquency and default by the potential borrower ex-ante.

Another reason for overborrowing or costly borrowing may be a lack of numeracy and financial literacy skills. To be able to calculate if the expected net returns to borrowing are non-negative in equation (7), basic numeracy and financial literacy skills are required. Stango & Zinman (2009) in the US, and Mantilla et al. (2014) in urban Colombia, that overborrowing increases in the *interest rate bias*, defined as the difference between the perceived interest rate

[6] This expression may even be negative while equation (9) holds true.

(by the borrower) and the actual interest rate. For the expected financial distress expressions for all three income brackets above, i shows up in way that a higher i leads to more distress. In sum,

Proposition 4. *The lower the financial literacy skills level of the household, the more adverse the effect of the uptake of microcredit on household socio-economic status.*

Finally a note on the relationship between investment returns and income volatility. Typically, loan-induced investment will increase risk, i.e., fatten the tails of the counterfactual distribution of socio-economic outcomes (vis-a-vis a non-borrowing scenario), for two otherwise identical households. However, there are investments that reduce the risk in the returns to capital and/or reduce income volatility, think for instance of loan-induced diversification of the business, or irrigation technology that reduces sensitivity to rainfall shocks. In many other cases, however, investment returns and income volatility are positively correlated in at least two ways. First, through covariate shocks that hit not only generate an adverse income shock to an individual but also to the community (and thereby demand for products and services of the borrower), which negatively affects investment returns. Second, income shocks and investment returns are likely to be temporally correlated: an unexpected health shock reduces income, and may also reduce (hard to substitute) labor input in the critical, early stages of an investment. Such positive correlation between income shocks and investment returns would further increase both up- and downside risks of credit. It would aggravate the increase in expected financial distress due to credit uptake for overborrowing agents.

3. Data and methods

3.1. Data structure

For data collection, we collaborated with a Ugandan Microfinance Deposit-Taking Institution ('the MDI' hereinafter) in order to have enough formal microcredit clients in our sample. Loan officers directed us to their borrowing clients. Which clients were visited depended on the schedule of the loan officer(s) within a month; little systematic bias is expected here since loan officers visit every group according to a pre-determined repayment frequency. Some bias may have arisen from only interviewing those present at the group repayment meeting, though efforts were made to track down and interview those group members not showing up at the group meeting. Loan officers also pointed us to the places of residence of first-time applicants for loans. Households were sampled in all 5 divisions of Kampala, as well as in the Wakiso, Luweero and Mukono districts surrounding Kampala. The final baseline sample consists of 855 respondents (and their households) interviewed between September 2013 and March 2014. Of the 855 respondent houshoelds, 124 were applicants at the MDI and the remaining 731 microcredit borrowers.[7]

A follow-up survey was conducted between October 2014 and March 2015 on a random subsample of respondents who were either formal borrowers or applicants for formal credit at baseline. Of the 320 households in the second wave, 222 were already formal borrower at baseline and 98 were applicant at baseline and had taken up their loan by the time of the follow-up survey. The main objective of our study is not to obtain unbiased estimates of the incidence of financial distress in a larger reference population, but rather to infer the average treatment effect on the treated of microcredit uptake on financial distress. We therefore follow (Solon et al., 2015) and do not use sampling weights[8].

3.2. Outcome measures

As touched upon in the Introduction section, our outcome measures focus on financial distress, which captures the symptomatic events of household dropping

[7]We excluded the 144-124=20 baseline applicants whom we could not identify in the MIS data, and whose loan applications were therefore, likely rejected.

[8]That is, we do not use sampling weights to adjust for the non-random sampling scheme. We will, later on, include sampling weights to achieve covariate balance between treated and control units, using entropy balancing.

below subsistence level. It is thus a more general concept then overindebtedness, which refers to distress that is debt-induced or for which a high debt burden prevents a transition to a state of less distress.

Inspired by the approach of Schicks (2013), the following questions were asked to elicit respondents' level of financial struggling[9]

(1) *assetless*: During the last 6 months, did your household have to sell any of its assets (i.e. land, motocycle, etc.) to meet other payment obligations?

(2) *eatless*: During the last month, did you eat less or of less quality than usual because of lack of money?

(3) *schoolless*: During the last 6 months, did you have to take out your children from school because of lacking funds?

(4) *runoutofmoney*: How often during the last 6 months, did your household

[9] In Schicks (2013), interviewers asked each respondent about the following list of sacrifices in relation to their loan obligations (% of her sample respondents making the sacrifice that find it unacceptable):

- Reduce food quantity/quality (cut down eating) (73 %)
- Reduce education (e.g. taking children out of school) (80 %)
- Work more than usual (e.g. take additional paid labour or work longer hours) (32 %)
- Postpone important expenses (e.g. for health, housing, business assets etc.) (33 %)
- Deplete your financial savings (e.g. money in the house or in a savings account) (38 %)
- Borrow anew to repay (take an additional loan) (85 %)
- Sell or pawn assets (e.g. jewellery, cattle, productive or household assets) (90 %)
- Seizure of assets (MFI takes property by force to make up for missed payments) (100 %)
- Use family/friends' support to repay (e.g. monetary contribution or other help) (72 %)
- Suffer from shame or insults (also gossip about you/exclusion from a contract) (100 %)
- Feel threatened/harassed by peers/family/loan officer (100 %)
- Suffer psychological stress in your marriage (80 %)

run out of money from previous revenues before the next revenues arrived (e.g. wages)? Choose one response: every month, every other month, twice, once, never, don't know.

(5) *healthless*: During the last month, was your household unable to pay for medicine/visiting the doctor because of lack of money?

(6) *openbill*: Do you (or any other member of your household) currently have any unpaid bills - open balance - and where? Do you have any open bill outstanding?

In contrast to (Schicks, 2013), none of the questions makes a direct link to debt or borrowing, to avoid priming effects (i.e. confirmation bias). In question 3, we mention 'payment obligations' to try to capture those asset sales that were induced by financial distress of the household (as opposed to sales due to reduced need for the asset or because a better substitute asset has just been acquired). We restricted ourselves to these six distress events (out of the 12 used by Schicks) because it is possible to ask them in a relatively objective way without having to 'frame' the question in terms of debt servicing. An additional reason for the exclusion of some of Schicks (2013)'s questions is that those capture illiquidity to a greater extent than that they capture insolvency, especially the question related to 'deplete financial savings'. Seizure and sales of assets were also viewed as some of the least acceptable by respondents in (Schicks, 2013)', as were reducing education (deemed unacceptable by 80% of her sample respondents who made this sacrifice) and reducing food consumption (deemed unacceptable by 73 % of respondents who made this sacrifice). The last three distress above were not used by (Schicks, 2013).

The responses to some of the questions on distress events are subjective, in particular distress event (1)-(4) above. For the *eatless* event, different people may interpret 'less than usual' in a different way, possibly depending on the volatility over time of their disposable income. Moreover, money is fungible and the 'inability to pay' for an expense might also reflect partly a changing opportunity cost of the expense itself and of time. Borrowing-fuelled business expansion may for instance increase the opportunity cost of education and be a stronger driver of the 'schoolless' event than a lack of funds for some households. The extent to which this is the case depend on how forward-looking and risk-averse individuals are, i.e. to what extent uncertain payoffs from education far into the future are discounted. Arguably, the more severe the distress event and the longer lasting its potential adverse effects, the more a household will try to prevent it at all cost and thus the more it captures genuine distress. In this light, it is worth noting that the *schoolless* event of school dropout is mostly a unidirectional event: of all follow-up respondents, 37.8% dropped their child out of school, while only 3.66% answered affirmatively to the 'opposite' question: 'During the last 6 months, have you been able to put a child, who was out of school for at least one term, back into school?'.

A measure of financial distress, the financial distress index, was constructed as the sum of positive answers to the above questions 1 to 6. For the medical dimension of financial distress (*healthless*), there is some misclassification, as only 285 out of 855 respondents got sick or injured the last month; for those households with no member falling ill or getting injured, this indicator takes on the value of 0. Hence, for some households the distress index is artificially too low, but results are not strongly affected by this choice[10]. To give distress events that are more serious for a respondent more weight, we also construct a continuously distributed financial distress index from the binary indicators using polychoric PCA (Kolenikov & Angeles, 2009) - this is the main outcome variable.

[10]To anticipate the mean effect estimates in the Results Section, we re-estimate the ATT estimates with a financial distress index with polychoric PCA-weights as outcome that excludes the *healthless* indicator. The results are: (1) OLS: coef. 0.199*** (st. error 0.074); (2) OLS-EB: 0.212** (0.108) (3) AIPW: 0.143* (0.078) (4) Panel FE: 0.471*** (0.136); (5) Panel FE-EB: 0.475** (0.188), where * p<0.1, ** p<0.05, *** p<0.001.

3.3. Control variables

In line with the literature, we first of all included standard demographics, such as the sex, age and education level of the respondent (i.e. the borrower for formal borrowers, and the head of the household or his wife in the case of other households). 2

Household's initial (i.e., pre-borrowing) socio-economic status is an important potential confounder, as it would influence both borrowing as well as financial struggling. Household income often contains a lot of ill-behaved measurement error, with potential biased estimates as a result (Azzarri et al., 2010; Millimet, 2011). Furthermore, post-treatment bias would be a concern, as one of the likely mechanisms through which borrowing potentially affects financial distress risk is through changes in income, and we do not have baseline income data for the subsample of borrowers. Household asset counts are likely to be more stable over time then income and subject to less (ill-behaved) measurement error. It also allows us to increase statistical power by being able to use more observations (some households were not able to come up with reasonably accurate income numbers). The wealth index was constructed based on asset counts on 25 assets including non-productive assets and housing characteristics (flooring, roofing, walls, and number of rooms). Using polychoric PCA, the wealth index was constructed as the linear combination of those asset indicators that maximize the proportion of explained variance in this first component. This proxy for socio-economic status is not perfect either as it is potentially endogenous: through the *assetless* indicator in the distress index, assets feature in both the regressand as well as in a regressor. However, omitting the wealth index as a control variable in regressions of distress on borrowing would leave out a potentially important confounder. Excluding the wealth index from the regressions hardly changes our results[11].

[11] Across the 5 estimation methods used in the mean effect estimates in this paper, the coefficient on microcredit uptake always retains its statistical significance, and the change in percentage points terms is at most 1.0%

From the perspective of financial distress risk, not just the current mean of household income is important, but also its volatility: for two households with identical characteristics including average income, the household whose income is more volatile has a higher likelihood of facing a liquidity shortage. We therefore asked respondents to what extent they considered their household income stable (answer options: very stable, stable, unstable, very unstable). Moreover, an indicator variable (referred to as 'Shock took place') is included in the regressions taking on the value of 1 if at least one of the following 4 events occured: (1) a member of the extended family fell sick and the respondent had to pay for larger medical or hopital expenses during the last month, (2) a household member lost his or her job during the last month, (3) a member of the extended family got married for which the respondent had to pay a substantial amount of money during the last month, (4) a large and special household expenditure was made (related to weddings and funerals of members of the household) during the last 30 days.

Financial literacy has been identified in the literature as a correlate of overindebtedness (measured as amount of credit card debt, arrears on consumer credit and self-reported excessive financial burdens of debt) in the UK (Gathergood, 2012) and in the US (Lusardi & Tufano, 2009). However, financial literacy was unrelated with Schicks (2013)' measure of overindebtedness in her sample from urban Ghana. More generally, randomized trials of financial literacy programs have been conducted in various countries to evaluate their impacts. Cole et al. (2009) offered randomly selected unbanked households in India and Indonesia financial literacy education, and found modest increases in the likelihood of opening a bank account for uneducated and financially illiterate households. The likelihood of opening a savings account was more affected by offering a small monetary incentive than by the financial literacy training. The discussion in the Theoretical Model Section suggested financial literacy skill levels may be a source of effect heterogeneity of microcredit uptake impacts. Our proxy for financial literacy levels was constructed as the sum of correct answers to a set of basic numeracy questions, and a set of five questions involving percentages and interest, slightly adapted from Bandiera et al. (2010). The questions whose answers make up the numeracy score and financial literacy score variables are listed in Appendix A.

Another key financial variable is a dummy variable taking on the value of 1 if at least one household member is a member of a Rotating Credit and Saving Association (ROSCA). Note that we do not control for loan characteristics or other (informal) loans the household may have outstanding at the time of the interview in order to prevent the introduction of post-treatment bias. The take-up of other (informal) loans may be partly the effect of microcredit uptake. For loan characteristics, we do not have the loan characteristics of the first loan cycle for long-term borrowers in the baseline survey.

3.4. Econometric approach and identification strategy

To identify the impact of the take-up of microcredit, one would need to randomly deny credit to some applicants and not to others. The problem with this approach is that very few lenders would want to turn down creditworthy borrowers. Another problem is that applicant-borrowers that are rejected may look for substitutes elsewhere. It is therefore more feasible to evaluate the impact of changes in credit *access*. Karlan & Zinman (2009) for instance, provide loans to a random subset of marginally rejected clients. This solves the first problem of lenders having to turn down applicants deemed creditworthy. By considering impacts of credit *access*, the possibility of applicants taking up substitutes is addressed. On the one hand, the pool of applicants initially deemed uncreditworthy by the lender is an interesting group, as positive impacts found would suggest lending criteria may be too stringent. A limitation of this approach however is that the impact of microcredit on applicants deemed creditworthy is not evaluated. The approach we take is to compare applicants for microcredit to those who already received it. A key underlying assumption is that the client recruitment and self-selection mechansisms is constant over time. We will evaluate this assumption in the Robustness checks subsection. Potential advantages of this approach compared to randomizing credit access to marginally uncreditworthy applicants, is (i) impacts are evaluated on the treated (i.e., the Average Treatment Effect on the Treated (ATT)), including inframarginal borrowers, and (ii) statistical power challenges due to rejected applicants taking up substitutes, are alleviated.

To examine the relationship between financial distress with formal borrowing and other covariates, then, we first use the baseline data to estimate regression models of the form

$$y_i = \beta \times microcredit_i + X_{it}\gamma + \epsilon_i \tag{10}$$

where y_i is the continuously distributed financial distress for household i, $formal_i$ takes on 1 if the household contains a formal borrower and 0 otherwise, and X_i is a vector of controls. The key assumption underlying this approach is that selection mechanism into microcredit uptake is time-invariant over the time span considered. To relax this assumption, a second type of analysis is conducted on the longitudinal data:

$$y_{it} = \beta \times microcredit_{it} + X_{it}\gamma + \alpha_i + \epsilon_{it} \tag{11}$$

where α_i are household fixed effects. The fixed effects filter out any time-invariant differences between treated (those having microcredit at baseline) and controls (those applying for credit at baseline). Such characteristics that could plausibly be considered time-invariant (over the course of a few years) include the applicant's enterpreneurial ability, risk preferences, and the personality of loan officers. The identifying assumption for the fixed effects estimation is that selection into microcredit in terms of individual-specific trends in financial distress is time-invariant. More on this in the Robustness checks section.

4. Results

4.1. Descriptive statistics

Table 2 reports summary statistics. The average number of distress events experienced in our sample during the six months preceding the interview is 1.87.

When a cutoff of $\geq$3 distress events is used, 31.7% of respondents face severe financial distress. As for the components making up the financial distress index, having to cut back on food consumption quantity and/or quality and running out of money are the most commonly experienced distress events, followed by having an open bill at a shop or elsewhere. With 61% of the sample having at least once been unable to pay for medicine when falling ill or getting injured prior to the interview, the level of financial distress in this sample appears high.

4.2. The mean impact of microcredit uptake on financial distress

Results on the average treatment effects on the treated on financial distress from both the cross-sectional and the panel household fixed effects analyses, are reported in Table 3. The cross-sectional analysis on the baseline data has the advantage of a sample with a larger number of households/cross-sectional observations, whereas the panel data analysis has the advantage of filtering out any time-invariant unobservables.

Additional control variables included (not reported) are a count variable taking on the value of 1 if the baseline survey took place in September 2013, the value of 2 if the baseline survey took place in October 2013, and so on. This variable, as well as this variable squared, are included to control for seasonality effects[12]. The results from both types of analyses coincide qualitatively, but the effect size estimate for food consumption from the FE estimation is larger (in absolute value) than the corresponding OLS regression partial effect estimate. In addition to OLS and linear panel regressions, we also run these estimations after entropy balancing (Hainmueller, 2011). Entropy balancing relies on a maximum entropy reweighting scheme that calibrates unit weights so that the reweighed treatment and control group satisfy balance conditions for (in our case,) the first, second and third sample moments of the covariates as well as all their pairwise

[12]Another approach, adding a dummy control variable for each calendar month in which the baseline survey took place, produced very similar results. Because the latter approach resulted in a lower goodness-of-fit, we opt for the approach described in the main text instead.

Variable(s)	Observations	Mean (s.d.): full sample	Mean: applicants	Mean: borrowers	Diff.	T-test (p-value)
Outcome variables						
Disc. distress index	855	1.891 (1.402)	1.726	1.922	0.197	0.1521
Cont. distress index	855	-0.048 (0.971)	-0.165	-0.026	0.138	0.1456
Wealth index	852	6.396 (4.630)	5.23	6.575	1.344	0.0083***
Eatless	855	0.453 (0.498)	0.476	0.448	-0.027	0.5734
Schoolless	855	0.1343 (0.321)	0.065	0.147	0.083	0.0130**
Assetless	855	0.100 (0.300)	0.048	0.109	0.061	0.0375**
Runoutofmoney	855	0.442 (0.497)	41.255	82.205	40.95	0.1553
Healthless	855	663.21 (472.009)	749.347	646.979	-102.368	0.0266**
Any open bill	855	0.572 (0.495)	0.468	0.591	0.123	0.0108**
Covariates						
Female	855	0.769 (0.422)	0.726	0.777	0.051	0.2192
Age	854	37.944 (10.589)	33.653	38.753	5.1	0.0000***
Household size	855	5.588 (2.731)	4.871	5.723	0.852	0.0014***
Muslim	855	0.217 (0.413)	0.218	0.217	0	0.9918
Compl. primary educ.	855	0.353 (0478)	0.331	0.357	0.026	0.5717
O-level educ.	855	0.276 (0.447)	0.194	0.292	0.098	0.0248**
A-level educ.	855	0.042 (0.201)	0.056	0.04	-0.017	0.3901
Diploma/univ. degree	855	0.069 (0.254)	0.056	0.071	0.015	0.5469
Numeracy score	855	2.628 (1.234)	2.75	2.605	-0.145	0.2297
Fin. lit. score	855	3.476 (1.485)	3.887	3.398	-0.489	0.0007***
Saving group	855	0.742 (0.438)	0.621	0.764	0.143	0.0008***
Shock took place	855	0.601 (0.490)	0.54	0.612	0.072	0.1327
Income reported: stable	855	0.382 (0.487)	0.339	0.395	0.056	0.237
Unstable	855	0.518 (0.500)	0.581	0.506	-0.075	0.1278
Very unstable	855	0.065 (0.247)	0.065	0.065	0.001	0.9725

Table 2: Variable means and standard deviations for the baseline data, and t-tests of the difference in their sample means across applicant and borrower subsamples.

interactions. Hainmueller (2011) show that after such reweighing, the treatment effect estimate based on observational data comes very close to the experimental benchmark of the famous Lalonde data. The procedure has been shown by Zhao & Percival (2016) to be doubly robust: if either the (logit) propensity score model or the outcome regression model is correctly specified, the mean causal effect estimator is consistent. Whereas the estimation approach works very well if either the outcome or propensity score model is specified correctly, entropy balancing is biased when neither is specified correctly, but the inverse probability weighing (IPW) estimator is unbiased in this case (Hirano et al., 2003; Zhao & Percival, 2016). On the other hand, if either model is correctly specified, IPW can be off the mark, and performs worse than entropy balancing. Therefore, we also estimate the ATT nonparametrically by means of augmented inverse probability weighing (AIPW) (Rubin & van der Laan, 2008). In the tails of the propensity score distribution, overlap may be limited, which affects estimates negatively in terms of precision and bias. In the estimates of columns (2),(3) and (5), as advised by Crump et al. (2009), we therefore trim observation lying outside of the interval $[\alpha,\ 1 \quad \alpha]$ to increase precision and robustness of the estimates. Their data-driven method gives $\alpha = 0.08$, dropping 17% of observations.

Reassuringly, all estimates approaches come to qualitatively the same conclusion, which is in line with the prediction of Proposition 1: microcredit uptake increases financial distress on average. Moreover, the linear regression based estimates after the entropy balancing-preprocessing, columns (2) (baseline data) and (5) (panel date with household fixed effects), are relatively similar in magnitude.

Table 3: Average treatment effects on the treated on the financial distress index.

	(1) OLS	(2) Entropy-bal. OLS	(3) AIPW	(4) Linear panel	(5) Entropy-bal. Linear panel
Microcredit	0.202***	0.2774**	0.2064***	0.4355***	0.3451*
	(0.01)	(0.11)	(0.08)	(0.14)	(0.18)
Female	-0.0682	0.1318		-0.4084	-0.8721***
	(0.13)	(0.09)		(0.35)	(0.13)
Head of household	0.0239	0.1093		-0.0690	0.0412
	(0.08)	(0.09)		(0.12)	(0.17)
Age in years	0.0016	0.0217**		-0.0170	-0.0081
	(0.01)	(0.01)		(0.01)	(0.02)
Household size	0.0544***	0.0640**		0.0234	0.0128
	(0.02)	(0.03)		(0.03)	(0.03)
Muslim	-0.0170	-0.0746		-0.6211***	-0.3053
	(0.09)	(0.09)		(0.20)	(0.34)
Primary education completed	-0.0833 (0.10)	-0.0385 (0.11)		-0.0542 (0.15)	-0.0738 (0.18)
Secondary education	-0.2552**	0.0739		-0.4124**	-0.3392*
	(0.12)	(0.12)		(0.17)	(0.19)
Tertiary education	-0.3182*	-0.1467		-0.5217	-0.3260
	(0.17)	(0.21)		(0.38)	(0.60)
Financial literacy score	-0.0275 (0.02)	-0.0148 (0.02)		0.0005 (0.03)	0.0019 (0.03)
Wealth index	-0.0153	-0.0111		-0.0148	-0.0210
	(0.02)	(0.02)		(0.03)	(0.03)
Member of a saving group	-0.2050* (0.12)	-0.2140* (0.11)		-0.3049* (0.16)	-0.0525 (0.17)
Household faced a shock	0.3870*** (0.09)	0.4417*** (0.08)		0.2416** (0.11)	0.2789** (0.13)
Income reported stable	0.2420* (0.13)	-0.0713 (0.26)		-0.2131 (0.32)	-0.2498 (0.34)
Income reported unstable	1.0493*** (0.14)	0.6521** (0.26)		0.7504** (0.34)	0.6980* (0.36)
Income reported very unstable	1.4022*** (0.22)	0.7284** (0.30)		1.3371*** (0.38)	1.1268** (0.45)
Household FE				✓	✓
N	850	589	589	1169	802
Wald	0.37	0.46		0.40	0.45
BIC	2683.80	.	.	917.41	562.31

* p<0.1, ** p<0.05, *** p<0.001

(1) Balance was not achieved on the financial literacy score variable by the entropy balancing algorithm, so it is included as a control variable post-balancing.

Standard errors in parentheses; robust standard errors in columns (1),(4),(5).

The results are in line with Proposition 1 postulated in Section 2: microcredit uptake on average increases financial distress. The effects are somewhat stronger after entropy balancing and in the household fixed effects estimates, suggesting that the effects estimated, if anything, are lower bounds on the true effect of microcredit uptake.

5. Robustness of the mean estimates

The inclusion of loan applicants as control group for microcredit borrowers addresses an important source of potential selection bias: the non-random self-selection by individuals into microcredit. Still, given the observational, quasi-experimental nature of the data, various concerns may linger regarding the strength of identification and remaining sources of bias in the estimates of interest. We will discuss and address these one by one.

5.1. Concerns about reverse causation and post-treatment bias

What comes first: borrowing or distress? To see that our design identifies the counterfactual of interest, it helps to visualize the timing of events. The clearest counterfactuals are obtained when restricting the sample of borrowers to those who received their first loan at least a year ago, as the possibility of reverse causation can then be excluded, see Figure 3. The difference-in-difference analysis is such that in survey wave 1, the treated group consists of the 'borrowers', while the applicants constitute the control group. By the time wave 2 arrives, both groups are in a treated state. Some of the applicants may have applied (partly) because they were in distress, but this holds too for those who are already borrower in round 1: some of them may have applied for a loan because they were in distress at that time. Unobserved confounding may still be an issue of course, one to which we will return in 5.3 below.

5.2. Survivorship bias

One source of sample selection bias in the first survey wave is that, whereas we (tried to) interview(ed) all loan applicants available (from loan officers), for the subsample of formal borrowers we mostly sample those who are still active

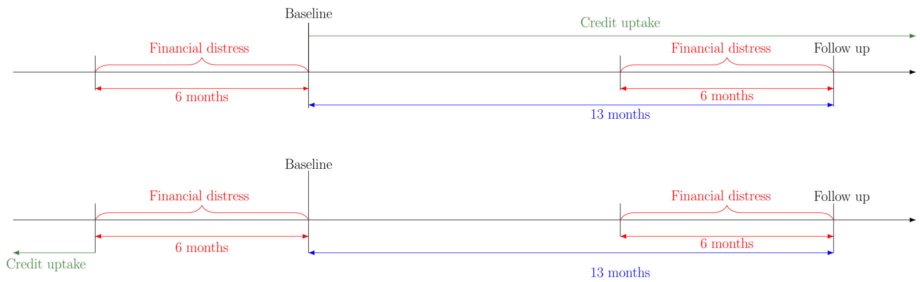

Figure 3: Time line of the determination of credit uptake and financial distress for the treatment group (upper timeline) and control group (bottom timeline) for the estimations where the subsample of borrowers (i.e., the treated) are restricted to those having taken up their credit at least a year prior to the baseline survey.

borrower. We did try to locate ex-clients, but the success rate in finding (and interviewing) them, is lower than for borrowers who are still active client (i.e., who still have a loan outstanding at the time of the baseline survey). The reason is that loan officers change branch, ex-clients change phone numbers and move to a different part of the city or outside of the city, etc. Whenever this process is non-random, and the 'positive' and 'negative' loan drop-outs would not happen to cancel out, then survivorship bias would bias estimates comparing applicants and borrowers. Clients may stop borrowing (earlier than others) because of 'positive' reasons (f.i., they found a better (from their perspective) substitute, are able to save more so don't 'need' a loan anymore) or because of 'negative' reasons: they did not have a good repayment record on their loan, obviating the dynamic incentives for repeat-borrowing, were able to stay within terms but only by making severe 'sacrifices', did not (perceive to) reap non-pecuniary benefits from being member of a borrowing group (e.g., increases in social capital), or even defaulted during a loan cycle. Especially if, on average, individuals stop borrowing (or we failed to interview them) because of distress, this would be unfortunate given our special emphasis on this outcome variable. Since we found that borrowers experience more distress than applicants, such distress-induced reduced likelihood of survivorship would have led to a downward bias in the effect estimate.

	Fin. distress index
$\leq$ 1 years formal borrower	0.294**
	(0.115)
	N=389
$\leq$ 3 years formal borrower	0.244**
	(0.124)
	N=500
$\geq$ 2 years formal borrower	0.244**
	(0.124)
	N=316
$\geq$ 4 years formal borrower	0.223
	(0.154)
	N=205
2-3 years formal borrower	0.282**
	(0.127)
	N=227

Table 4: Short, long and medium-term effect estimates of credit uptake.

A client dropout study conducted by the MDI found that the reasons clients give for departing tend to relate more to the lending methodology than problems with loan repayment. Most drop-outs complain about inadequate loan amounts and terms, mandatory weekly meetings (the dropout study was conducted in 2012, whereas during the time of our survey the MDI only offered biweekly and monthly repayment schemes), and having to pay for group members who default.

Survivorship bias, if any, could be either positive or negative. Those who drop out, may do so for 'positive' reasons (find a better substitute source of capital through an active information seeking process, or, not 'needing' a loan anymore due to re-investing increased business profits obtain through loan-induced investments that were succesful). Alternatively, individuals may not re-apply for a new loan cycle if (with low levels of financial literacy) the loan terms turned out not suitable to their cash flows, or if they are struggling financially (possibly due to a shock). To investigate this, we merge the baseline survey data with MIS data of 31 October 2013 as well as MIS data of 31 March 2014. The dependent variable takes on the value of 1 if the individual is still present in the data on 31 March 2014 (which is the case if and only if he or she is still borrowing), and 0 otherwise. We regress this attrition indicator on the outcomes under study at baseline, and a set of controls, including the time passed since becoming client until the date of interview and the time passed since the time of interview and the 31st of March 2014. The results, presented in Table 9 in Appendix C, are mixed. The overall picture that arises, is that those who are most likely to remain borrower are male, have tertiary education, relatively stable income flows, and have loans with joint rather than individual liability. Experiencing a shock is associated with a higher likelihood of dropout. Conditional on the covariates, the coefficient on the financial distress index is close to zero and statistically not significant, indicating survivorship bias may not be much of a concern.

The longer the (average) time lapsed since taking up the loan for the subsample of formal borrowers, the more severe the survivorship bias problem would plausibly be, as a relatively larger share of initial borrowers from will have stopped borrowing. Therefore, as a robustness check, we gauge the sensitivity of the effect estimates to restrictions on the subsample of borrowers with regards to the number of years that passed since they became client[13]. The estimates, reported in Table 4, show that the coefficient on microcredit remains quite stable for different subsampling definitions, alleviating concerns about survivorship bias. The fact that the coefficient on microcredit for the food intake index regressions is relatively stable when excluding the borrowers who received their loan at most 1 year ago, or at most than 3 years ago, suggests that the estimated effects of microcredit uptake on food consumption reflects more than just a transitory 'sacrifice' of reducing food consumption during business expansion. Analogously, the stability of the estimated coefficients on microcredit in the financial distress regressions when excluding short-term borrowers of at

[13]The date the individual became client in the system is the date they registered their first loan application at the MDI.

most one year, suggests that our financial distress index does not just capture 'illiquidity' (a concept we described in the Theoretical Model Section), but that it is indeed reflective of more systematic financial distress.

5.3. Any remaining bias from unobserved confounding

The quasi-experiment we consider is that two individuals apply for credit at two different points in time, a few years apart, and that their difference in timing of applications is (quasi-)random. A concern however, may be that individuals with more promising investment projects (higher $[r \cdot r_{high}]$ as in our theoretical model) would apply for microcredit 'earlier' after a branch opens. Furthermore, loan officers may first recruit the most promising and least risky clients, and when later on, when local credit markets are relatively more saturated and the least risky borrowers already have a microcredit, are forced to recruit more risky potential clients or potential clients with lower returns to capital. On the other hand, those who apply later may be less risky as some of them are 'recommended' to a client recruiting loan officer by existing clients as is not uncommon at the MDI. If social networks are formed by assortative matching on aspects that are correlated with delinquency risk, then loan officers may choose to take recommendations from long-term, successful client borrowers who recommend them other low-risk potential clients.

Such non-random differential temporal selection into microcredit would be a form of unobserved confounding, potentially generating omitted variable bias. To the extent that a household's 'innate distress risk', or the risk or ambiguity aversion of a potential borrower is time-invariant, it would be filtered out by the household fixed effects in the panel data analysis. The same holds true if the factors affecting the expected return to investment are time-invariant over the course of a few years (think of educational attainment, entrepreneurial skills, etc.). But for the baseline data analysis, it would still be a concern.

Fortunately, we have MIS data from the MDI to shed light on the direction, if any, of the evolution of the selection mechanism. From October 2010 until January 2014, we consider the branches we sampled from (minus one that opened recently) for the loan products that are part of our research. Whereas the number of borrowing clients has declined slowly but steadily (perhaps due to the opening of new branches), from 7725 in October 2010 to 6803 in January 2014. But over the same period, the delinquency rate (>0 days late on a loan) has increased monotonically over time, from 7.96% in October 2010 to 14.01% in January 2014. Hence delinquency risk has increased, perhaps due to increasing credit market saturation in Kampala and the saturation of lower risk potential clients that can be recruited by loan officers. Indeed, the urban Ugandan credit market has been perceived by MDI CEOs as competitive (McIntosh et al., 2005). This is a reason that over time, more risky borrowers are selected into credit uptake, suggests that the estimates reported here are lower bound on the true effects. This is also in line with the fact that the panel analyses with household fixed effects return coefficients that are larger in magnitude than the cross-sectional ones.

To test for changes over time since a branch starts lending, i.e. learning effects, we add a variable for months passed since the branch of the MDI opened as a covariate to the regressions with the outcomes under study[14]. But this variable's coefficient does not reach statistical significance and is close to zero. Third, we conduct various exogeneity tests and sensitivity analyses with respect to selection on unobservables that we will now report.

Whereas stability of a coefficient estimate is not sufficient to give it a causal interpretation, it is necessary in a sense (Lu & White, 2014). We test the stability of the effect estimates with regards to the covariate set, by conducting the Romer's extension of Leamer's Extreme Bounds Analysis. The results are reassuring: in 98% of the covariates subset permutations, the coefficient on microcredit in the estimations on the baseline data, is of the same sign. See also the density of the coefficient estimates in Figure 4. As a second sensitivity analysis, we apply a method developed by Oster (2014), which analyzes coefficient sta-

[14]The branches under consideration existed for an average of 5.25 years (63.6 months, minimum 22 months, maximum 100 months) at the time of the baseline survey.

bility and movements in R^2 when control variables are added to the regression. The assumption is that selection on unobservables is proportional (governed by a coefficient of porportionality, δ) to selection on observables. A value of $\delta = 1$ would imply equal selection on observables and unobservables. More generally, assuming proportional selection on observables and unobservables, a positive δ implies that the coefficient estimate is biased away from zero by selection on unobservables, whereas a negative δ implies that the coefficient estimate is biased towards zero. After specifying a maximum R^2 that would be obtained if all confounders were included in the regression equation (an $R^2 = 1$ is not realistic given measurement error), the δ can be estimated. For $R^2_{max} = 0.7$ we obtain a negative value of $\hat{\delta} = -0.54$ for the sample of formal and informal borrowers, $\hat{\delta} = -0.43$ for the sample of formal only and informal only and $\hat{\delta} = -0.50$ for the sample of formal borrowers and applicants for formal loans[15]. Assuming proportional selection on observables and unobservables, we can thus conclude the following with regard to the coefficient on formal borrowing in the financial distress regressions: the positive and significant estimate in the sample of formal borrowers and applicants is a lower bound on the true effect. The panel fixed effect regressions can in a way (namely, when proportional selection on observables and unobservables is assumed) also be seen as a sensitivitiy analysis with respect to bias from selection on unobservables for the comparison between applicants and formal borrowers, as the household fixed effects filter out all time-invariant household-level unobservable differences between formal borrowers and applicants.

Finally, not all individuals may apply at random times during their (productive) lives. Individuals and their households may apply for credit when their socioeconomic status or business is in a dip (á la Ashenfelter dip). This would invalidate the assumption of no systematic selection based on individual-specific trends in financial distress that underlies our panel data analysis. Note however, that in the subsample of the MDI covering the branches considered in this study,

[15]Using $R^2_{max} = 0.9$, similar results are obtained (available from the authors upon request).

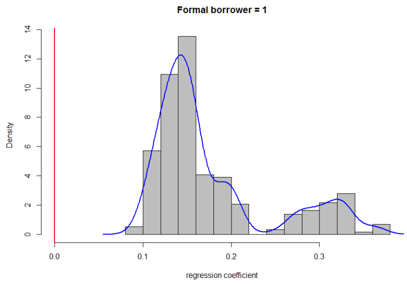

Figure 4: Formal borrowers and applicants.

the median number of days passed since becoming client to receiving a loan is 28 days (inter-quartile range: [13,182]). This suggests that the formal microloans considered in this study are unlikely to be useful in making up for liquidity shortfalls. Second, the financial distress index reflects distress experienced not in the last weeks but over a longer period of the last 6 months, thereby diluting any Ashenfelter dip type selection. If in spite of the aforementioned facts, some of individuals do select into credit due to them experiencing financial distress, then this would bias the estimated coefficients downwards, and the coefficients could be interpreted as lower bounds on the true effect. Note that some applicants may be foreseeing distress while not yet having experienced it (to its full extent), and apply for credit to envisage to use the credit to prevent having to experience the distress event, say a school dropout. However, if a household applies for microcredit to prevent, say, school dropout, then the financial struggling leading up to such an important event would likely show up in the household reported to have run out of money in the months before this decision is taken. However, when excluding the 73 applicants from the analysis who ran out of money at most once, the estimated effect if anything, is stronger (coef. 0.403; st. error 0.160). Similarly, when excluding the 58 applicants who have an open bill, the effects get stronger (coef. .473; st. error .138).

6. Effect heterogeneity

The prediction from the theory in Section 2 (Proposition 1 and 2) was that the microcredit uptake-distress response is larger for households with more volatile income flows, those who do not have access to emergency loans, and those that have low financial literacy skill levels.

We start with the financial literacy score as conditioning variable, which is a count variable running from 0 to 10, but can be treated as a continuous one by adding small white noise, which we do[16]. Abrevaya et al. (2015) proposed a consistent and asymptotically normal semiparametric estimator for the conditional average treatment effect on the treated (CATT) where the conditioning is on a continuously distributed covariate. The left panel of Figure 5 displays the CATT with financial literacy skills as conditioning variable, with different bandwidths. Again we trim the observations with propensity score estimates outside the interval [0.08, 0.92] to increase precision and robustness of the estimates. The right panel of Figure 5 shows the CATT estimates with a bandwidth of $0.55 * \sigma$, with 95% confidence bands. The CATE estimates are in line with Proposition 4: the lower the financial literacy skill levels of the respondent, the stronger the household's financial distress response to microcredit. For financial literacy scores at or above the median of 6, the null of no effect of microcredit uptake cannot be rejected.

For discrete covariates, the sample can simply be split based on the values of the covariate, and treatment effects can be estimated for each subsample, as suggested for instance by (Abrevaya et al., 2015). The treatment effect estimates conditional on household ROSCA membership are in line with Proposition 3: households that do not contain a ROSCA member experience a much stronger microcredit-distress response, and the response becomes statistically insignificant for households that have a(t least one) ROSCA member. The evidence is weaker for income volatility. The effect estimates for the subsample with more

[16] As suggested in the code by Abrevaya et al. (2015), a uniformly distributed random number in the interval (0,1) is added to the variable, and then 0.5 substracted.

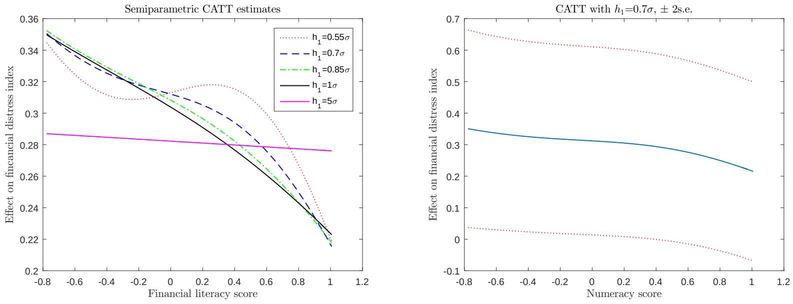

Figure 5: CATE estimates of microcredit uptake on the financial distress index, with the financial literacy score as conditioning variable.

	OLS-EB	AIPW	Panel FE-EB
Not member of ROSCA	0.310** (0.142) [n = 185]	0.402*** (0.121) [n = 184]	0.681** (0.301) [n = 245]
Member of Rosca	0.158 (0.118) [n = 404]	0.098 (0.093) [n = 397]	0.188 (0.208) [n = 557]

Before splitting the sample, observations with a propensity score outside of the $[0.08, 0.92]$-interval were trimmed following the data-driven rule of Crump et al. (2009).

Table 5: CATE estimates of microcredit uptake on the financial distress index, with ROSCA membership as conditioning variable.

volatile household income flows are larger in magnitude than those for the subsample with more stable income, but for some of the estimators, they do not reach statistical significance.

	OLS-EB	AIPW	Panel FE-EB
Income reported stable/very stable	0.149	0.218*	0.155
	(0.130)	(0.121)	(0.275)
	$[n = 228]$	$[n = 230]$	$[n = 309]$
Income reported unstable/unvery stable	0.221	0.220	0.430**
	(0.135)	(0.096)	(0.211)
	$[n = 358]$	$[n = 358]$	$[n = 492]$

Before splitting the sample, observations with a propensity score outside of the $[0.08, 0.92]$-interval were trimmed following the data-driven rule of ?.

Table 6: CATE estimates of microcredit uptake on the financial distress index, with household income volatility as conditioning variable.

7. Conclusion and policy recommendations

To gain understanding of the links between household borrowing behavior and financial distress risk, we developed a simple three-period theoretical model. One, perhaps intuitive, prediction that is derived is that borrowing increases financial distress risk. The predictions regarding the heterogeneity of the credit-distress response is that the response is stronger when the household does not have access to emergency credit, its income is volatile or when financial literacy skills are low.

These predictions were tested against quasi-experimental household panel survey data collected in urban Uganda. Indeed, microcredit uptake is associated with increased risks of, and levels of, financial distress. Given the observational, quasi-experimental nature of the data, these mean impact results are scrutinized and withstand various types of robustness checks. Next, substantial effect heterogeneity is uncovered, which may partly be due to the outcome measures constructed in this study being more effective in capturing relocations in the counterfactual (outcome) distributions as compared to more traditional outcomes such as profits or consumption expenditures. The two important sources of effect heterogeneity identified in this study relate to financial literacy levels and ROSCA membership. The finding that the microcredit-distress response is much weaker or absent when a household has a participant in a ROSCA, suggests a complementary, consumption smoothing role for ROSCAs in formal credit markets. This is in line with the theoretical literature on the insurance function of ROSCAs.

A word of caution regarding extrapolation of our findings: while these findings may be relevant to the spread of general-purpose group loans in other highly penetrated urban credit markets, the effects of credit expansion will be heterogeneous with respect to loan type and can therefore not be extrapolated to other loan types such as individual liability loans or school fee loans. We especially warn against extrapolation of the findings to rural areas, where loan market penetration is much lower, and the growth-inhibiting effect of existing credit market constraints is likely to be more important.

On the policy side, perhaps testing numeracy skills and using it in credit scoring may reduce financial distress in the population. In other, forthcoming work, we analyze the role loan officer incentives play in the recruitment process.

On the methodological front, eliciting respondents' recent experiences of discrete events (which could also include salutary events, such as asset purchases, housing improvements and enrolling a child back into school) may help boost statistical power to detect relocations up and down the counterfactual distribution. It is therefore advised that future randomized trials that aim to evaluate population-level impacts of financial services and financial market interventions include such indicators in the range of outcome measures to more completely and effectively capture their complex and heterogenous impacts.

References

Abrevaya, J., Hsu, Y.-C., & Lieli, R. P. (2015). Estimating conditional average treatment effects. *Journal of Business & Economic Statistics*, *33*, 485–505.

Angelucci, M., Karlan, D., & Zinman, J. (2015). Microcredit impacts: Evidence from a randomized microcredit program placement experiment by compartamos banco. *American Economic Journal: Applied Economics*, *7*, 151–82. doi:10.1257/app.20130537.

Azzarri, C., Carletto, C., Covarrubias, K. et al. (2010). Measure for measure: Systematic patterns of deviation between measures of income and consumption in developing countries. evidence from a new dataset. In *Fifth International Conference on Agricultural Statistics, Kampala, Uganda*.

Baland, J.-M., Guirkinger, C., & Hartwig, R. (2016). Strategic positioning in savings groups: The insurance role of roscas revisited. *mimeo*, .

Bandiera, O., Goldstein, M., Rasul, I., Burgess, R., Gulesci, S., & Sulaiman, M. (2010). Intentions to participate in adolescent training programs: evidence from uganda. *Journal of the European Economic Association*, *7*, 548–560. doi:10.1111/j.1542-4774.2010.tb00525.x.

Banerjee, A. V., Duflo, E., Glennerster, R., & Kinnan, C. (2015a). The miracle of microfinance: evidence from a randomized evaluation. *American Economic Journal: Applied Economics*, *7*, 22–53. doi:10.1257/app.20130533.

Banerjee, A. V., Karlan, D., & Zinman, J. (2015b). Six randomized evaluations of microcredit: Introduction and further steps. *American Economic Journal: Applied Economics*, *7*, 1–21. doi:10.1257/app.20140287.

Barrett, C. B., & Carter, M. R. (2013). The economics of poverty traps and persistent poverty: empirical and policy implications. *The Journal of Development Studies*, *49*, 976–990.

Calomiris, C. W., & Rajaraman, I. (1998). The role of roscas: lumpy durables or event insurance? *Journal of development economics*, *56*, 207–216.

Carter, M. R., & Barrett, C. B. (2006). The economics of poverty traps and persistent poverty: An asset-based approach. *The Journal of Development Studies*, *42*, 178–199.

Cole, S. A., Sampson, T. A., & Zia, B. H. (2009). *Financial literacy, financial decisions, and the demand for financial services: evidence from India and Indonesia*.

Crépon, B., Devoto, F., Duflo, E., & Parienté, W. (2015). Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in morocco. *American Economic Journal: Applied Economics*, *7*, 123–150.

Crump, R. K., Hotz, V. J., Imbens, G. W., & Mitnik, O. A. (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, (p. asn055).

CSFI (2012). *Microfinance banana skins, Staying Relevant (2012): Survey of Microfinance Risk*.

CSFI (2014). *Microfinance banana skins, Facing Reality (2014): Survey of Microfinance Risk*.

De Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to capital in microenterprises: evidence from a field experiment. *The Quarterly Journal of Economics*, (pp. 1329–1372).

Fafchamps, M. (2013). Contraintes de crédit, collatéral et prêts aux pauvres. *Revue d'économie du développement*, *21*, 79–100.

Fang, H., Ke, R., & Zhou, L.-A. (2015). *Rosca Meets Formal Credit Market*. Technical Report National Bureau of Economic Research.

Gathergood, J. (2012). Self-control, financial literacy and consumer over-indebtedness. *Journal of Economic Psychology*, *33*, 590–602. doi:10.1016/j.joep.2011.11.006.

Hainmueller, J. (2011). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, (p. mpr025).

Hirano, K., Imbens, G. W., & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, *71*, 1161–1189.

Karlan, D., & Zinman, J. (2009). Expanding credit access: Using randomized supply decisions to estimate the impacts. *Review of Financial studies*, (p. hhp092).

Klonner, S. (2003). Rotating savings and credit associations when participants are risk averse. *International Economic Review*, *44*, 979–1005.

Kolenikov, S., & Angeles, G. (2009). Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer? *Review of Income and Wealth*, *55*, 128–165.

Lu, X., & White, H. (2014). Robustness checks and robustness tests in applied economics. *Journal of Econometrics*, *178*, 194–206.

Lusardi, A., & Tufano, P. (2009). *Debt literacy, financial experiences, and overindebtedness*. Technical Report w14808.

Mantilla, C. et al. (2014). *Financial literacy in Colombia seen through the lens of the interest rate bias*. Technical Report Institute for Advanced Study in Toulouse (IAST).

McIntosh, C., Janvry, A., & Sadoulet, E. (2005). How rising competition among microfinance institutions affects incumbent lenders. *The Economic Journal*, *115*, 987–1004. doi:10.1111/j.1468-0297.2005.01028.x.

Meager, R. (2015). Understanding the impact of microcredit expansions: A bayesian hierarchical analysis of 7 randomised experiments. *Available at SSRN 2620834*, .

Millimet, D. L. (2011). The elephant in the corner: a cautionary tale about measurement error in treatment effects models. *Advances in Econometrics: Missing-Data Methods A*, *27*, 1–39. doi:10.1108/S0731-9053(2011)000027A004.

Oster, E. (2014). Unobservable selection and coefficient stability: Theory and evidence, .

Roodman, D. (2012). *Due diligence: An impertinent inquiry into microfinance*. CGD Books.

Rubin, D. B., & van der Laan, M. J. (2008). Empirical efficiency maximization: Improved locally efficient covariate adjustment in randomized experiments and survival analysis. *The International Journal of Biostatistics*, *4*.

Schicks, J. (2013). The sacrifices of micro-borrowers in ghana–a customer-protection perspective on measuring over-indebtedness. *The Journal of Development Studies*, *49*, 1238–1255.

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

Stango, V., & Zinman, J. (2009). Exponential growth bias and household finance. *The Journal of Finance*, *64*, 2807–2849.

Zhao, Q., & Percival, D. (2016). *Entropy balancing is doubly robust*. Technical Report Working Paper.

Appendix A: The remaining two income brackets in the base case of the theoretical model.

Case 2) Poor: $y_2 + \alpha \cdot i \cdot I = 2 \cdot c_{min}$

Table 7 lists the consumption shortfall under the various states of the world for this income bracket.

State of the world	Cons. shortfall if L>0 (formal loan)	Cons. shortfall if L=0 (no formal loan)	Difference	Probability
y_{low}, r_{high}	$\alpha \cdot i \cdot I$ - ($r_{high} - i) \cdot I$	$\alpha \cdot i \cdot I$	$(i - r_{high}) \cdot I$	$(1-p)q$
y_{low}, r_{low}	$(1+\alpha) \cdot i \cdot I$	$\alpha \cdot i \cdot I$	$i \cdot I$	$(1-p)(1-q)$

Table 7: States of the world and their consequences for expected difference in consumption shortfall w.r.t. $2 \cdot c_{min}$ between households with and without a formal loan.

States of the world (y_{high}, r_{low}) and (y_{high}, r_{high}) are not considered here, as lifetime consumption is above subsistence level under those contingencies. The expected difference in the consumption shortfall due to microcredit equals

$$q(1-p)(r_{high-i}I + (1-p)(1-q)iI \tag{12}$$

Hence, all other initial conditions constant, the take-up of microcredit increases financial distress if

$$(1-q)i > q(i - r_{high}) \tag{13}$$

For a given expected returns to investment $q \cdot r_{high}$, an increase in riskiness of the investment project (a lower q), leads to more financial distress. Neither income level nor income volatility affect the microcredit-expected distress link in this case. Consider now a household, the income of which is even closer to subsistence level, so that financial distress would occur or be possible in all but the most favorable (y_{high},r_{high}) states of the world.

Case 3) Very poor: $y_{high} + y_2 - \alpha \cdot i \cdot I = 2 \cdot c_{min}$

Table 8 tabulates the states of the world and the corresponding consumption shortfall of the household (ignoring the (y_{high}, r_{high}), scenario. Under (y_{high}, r_{high}) there is no consumption shortfall, and relocations on the upper end of the distribution of consumption are not or to a lesser extent captured by our financial distress index (as compared to relocations in the lower end of the distribution).

State of the world	Cons. shortfall if L>0 (formal loan)	Cons. shortfall if L=0 (no formal loan)	Difference	Probability
y_{high}, r_{low}	$(1-\alpha)\cdot i \cdot I$	$-\alpha \cdot i \cdot I$	$i \cdot I$	$p \cdot (1-q)$
y_{low}, r_{high}	$y_{high} - \alpha \cdot i \cdot I - (r_{high} - i) \cdot I$	$y_{high} - \alpha \cdot i \cdot I$	$-(r_{high} - i) \cdot I$	$q \cdot (1-p)$
y_{low}, r_{low}	$y_{high} + (1-\alpha) \cdot i \cdot I$	$y_{high} - \alpha \cdot i \cdot I$	$i \cdot I$	$(1-p)(1-q)$

Table 8: States of the world and their consequences for expected difference in consumption shortfall w.r.t. $2 \cdot c_{min}$ between households with and without a formal loan.

The difference between the mean consumption shortfall of households with and those without a formal loan is

$$(1-\alpha \cdot p)(1-q) \cdot i \cdot \cdot I - q(1-p)(r_{high} - i) \cdot I \tag{14}$$

If this expression is positive, then formal borrowing increases the expected consumption shortfall:

$$(1-\alpha \cdot p)(1-q) \cdot i \cdot > q(1-p)(r_{high} - i) \tag{15}$$

Inequality (15) shows that the expected increase in consumption shortfall due to microcredit increases in the riskiness of the investment project $(1-q)$, in income volatility $(1-p)$, and in the level of income (α).

Appendix B: The numeracy and financial literacy scores, the food frequency score, and the asset index.

A1: The Financial literacy scores

The numeracy score was constructed as the sum of correct answers to the following four questions:

- What is 25+17?
- What is 49-23?
- What is 12*4?

- What is 56:7?

The following five questions were posed to elicit financial literacy levels, slightly adapted from Bandiera et al. (2010):

- What is 20% out of 3000 UgSh?
- If you could save UGX5,000 per month, how many months would you need to save to get UGX30,000?
- If you needed UGX180,000, how much would you need to save per month (in UgSh) to have the money within one year (12 months)?
- Assume that you saw a radio of the same model on sale in two different shops. The initial retail price was UGX 20,000. One shop offers a discount of UGX 1,500, while the other one offers a 10% discount. Which one is a better bargain?
  - Discount of 10% on 20,000
  - Discount of UgSh 1,500
  - They are equally good
  - Don't know
  - Suppose you have deposited UGX 100,000 in the bank for an interest of UGX 10,000 per year. If you withdraw all the money after 3 years, how much will you get?

The financial literacy score was constructed as the sum of correct answers to the above five answers plus 1 if the number of commercial banks and microcredit institutions known to the respondent ('Please mention as many names of banks in Uganda as possible') was higher than the average number for all respondents, which was 6.33 financial institutions.

A2: The food consumption index

The food frequency questionnaire is listed below.

[ENUMERATOR: For each item, record the number of days the food was eaten PER WEEK]

For each food group, these are the answer options:

- never
- once
- twice
- 3 - 5 times
- almost every day
- not last week, but during last month

The food groups are:

(1) Maize/Millet/Sorghum/Ugali/Pocho, (2) Rice, (3) Wheat flour/bread/chappati/rolex, (4) Cassava (sweet potato), (5) Irish potato, (6) Matooke, (7) Pork, (8) Chicken/duck, (9) Beef, (10) Mutton/lamb/goat, (11) Fish, (12) Dairy products, (13) Beans or peas or lentils, (14) Vegetables (any), (15) Insects (any), (16) Sugar, (17) Salt, (18) Coffee, (18) Beer or other alcoholic beverages, (19) Tea, (20) Soda, (21) Tobacco/Cigarettes.

For the analysis with the food index in this paper, we excluded (18) Beer or other alcoholic beverages and (21) Tobacco/Cigarettes.

A2: The asset index

Counts of the following assets were used to construct the asset index: (1) rooms, (2) chairs, (3) tables, (4) beds, (5) sofas, (6) mirrors, (7) watches, (8) kerosene stoves, (9) gas stoves, (10) televisions, (11) radios, (12) mobile phones, (13) generators, (14) solar panels, (15) light bulbs, (16) bicycles, (17) motorcycles, (18) cars, (19) refridgerators, (20) chicken.

Appendix C: Analysis of attrition (baseline data + MIS data)

Table 9: Marginal effect estimates from a probit regression model predicting dropout as of 31/03/2016 (robust standard errors). The dependent variable takes on the value of 1 if the borrower is still present in the MIS data (i.e., has not dropped out of borrowing from the MDI), and 0 otherwise.

	(1)
Fin. distress score	-0.0072
	(0.01)
Nr. of days passed since becoming client till interview	-0.0002***
	(0.00)
Nr. of days passed since interview till 31/03/2016	0.0002***
	(0.00)
Female borrower	-0.0367***
	(0.01)
Household size	-0.0046
	(0.00)
Completed primary education	-0.0407***
	(0.01)
Compl. secondary educ.	-0.0667***
	(0.02)
Compl. tertiary educ.	-0.0497**
	(0.02)
Financial literacy score	0.0086***
	(0.00)
Income reported (very) unstable	0.0346**
	(0.01)
Shock took place	0.0029
	(0.01)
Village Group loan	0.1238***
	(0.03)
Small Group loan	0.1550***
	(0.03)
N	759
$Pseudo-R^2$	0.4950
Wald	485.46

Robust standard errors in parentheses
* p<0.1, ** p<0.05, *** p<0.01