

Grohmann, Antonia; Herbold, Steffen; Lenel, Friederike

Working Paper

Repayment under flexible loan contracts: Evidence from Tanzania

DIW Discussion Papers, No. 1884

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Grohmann, Antonia; Herbold, Steffen; Lenel, Friederike (2020) : Repayment under flexible loan contracts: Evidence from Tanzania, DIW Discussion Papers, No. 1884, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/222933>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

1884

Discussion
Papers

Repayment under Flexible Loan Contracts: Evidence from Tanzania

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

IMPRESSUM

© DIW Berlin, 2020

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
<http://www.diw.de>

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
<http://www.diw.de/discussionpapers>

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
<http://ideas.repec.org/s/diw/diwwpp.html>
<http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html>

Repayment under Flexible Loan Contracts:

Evidence from Tanzania*

Antonia Grohmann[†] Steffen Herbold[‡] Friederike Lenel[§]

July 14, 2020

Abstract

We study repayment and delinquency in an innovative loan contract that offers borrowers a wide range of flexibility. Using a large administrative dataset, we perform unsupervised pattern analysis to study how borrowers repay within the framework of this loan. We identify eight clusters that can be grouped into three distinct repayment types. We show that borrowers with fluctuating incomes and limited consumption smoothing resources use the loan's flexibility more and that farmers in particular adjust their repayment to cash flow. Finally, we show that high use of the loan's flexibility is associated with repayment difficulties, yet typically does not result in eventual default; whereas borrowers who face repayment difficulties that are likely driven by economic shocks face a high probability of default.

Keywords: flexible loans, loan contracts, repayment, microfinance

JEL Classification: D14 (Personal Finance), G21 (Banks, Microfinance Institutions), C55 (Large Datasets)

*We thank for helpful comments, Jenny Aker, Kristina Czura, Stefan Klonner, Lukas Menkhoff, Lisa Spantig, and seminar participants as Aarhus, GIGA Hamburg and Göttingen, as well as participants at Africa and Europe Forward together conference 2018, DIAL Development Conference Paris 2018, International Conference on Globalization and Development 2018, SABE/IAREP conference London 2018. Financial support by the German Research Foundation (DFG, grant RTG 1723 and CRC 190) is gratefully acknowledged.

[†]Aarhus University, Fuglesangs Alle 4, 8210 Aarhus, Denmark, agrohmann@econ.au.dk and DIW Berlin

[‡]Institute of Computer Science, University of Göttingen, Germany, Goldschmidtstr. 7, 37077, Göttingen, herbold@cs.uni-goettingen.de

[§]RTG Globalization and Development, University of Göttingen, Platz der Göttinger Sieben 5, 37073 Göttingen, Germany, friederike.lenel@uni-goettingen.de

1 Introduction

A well-developed financial system can spur economic growth and reduce inequality (Giné and Townsend 2004; Levine 2005; Beck et al. 2007; Beck and Demirgüç-Kunt 2008). In particular for low-income people, access to financial services and the possibility to finance larger investments can provide means to escape poverty and can thereby improve individual welfare (Cull et al. 2009; Banerjee et al. 2019). Yet, while it is shown that small businesses and farmers in developing countries often have returns to capital that are higher than the market interest rate (De Mel et al. 2008), they are typically unable to capture these high marginal returns due to their limited access to financial services (Banerjee and Duflo 2014). Traditionally, many people living in developing countries have been under-served by the financial system. Lenders are less willing to lend to poor borrowers, due to asymmetric information, moral hazard, lack of collateral, and the often weak legal system that makes contract enforcement difficult. The poor are believed to lack the necessary financial literacy and self-control to borrow larger amounts – as a consequence, monitoring costs are very high compared to loan sizes. Furthermore, low-income people are often self-employed, their income fluctuates considerably throughout the year and they are more prone to shocks, which makes lending to them risky.

Microfinance institutions (MFIs) deal with these constraints by offering small loans under rigid repayment schedules. Loans need to be repaid through fixed, typically weekly installments with the first installment due briefly after loan disbursement. Compared to loans in developed countries no physical collateral is required; instead, lending by and large takes place in a group with joint liability and so physical or financial capital is replaced by social capital. Recently, these features have come under criticism. It is argued that the fixed repayment schemes are not compatible with the highly fluctuating income of many borrowers, group lending exposes borrowers to substantial social pressure, and the limited loan size inhibits the often needed large investments. As a consequence, microfinance loan structure can lead to overindebtedness and encourage sub-optimal investment decisions (Pearlman 2010; Shoji 2010; Field et al. 2012; Field et al. 2013; Banerjee et al. 2015). In fact, these features might partly explain the limited effects of microfinance at reducing poverty (Angelucci et al. 2015; Augsburg et al. 2015; Banerjee et al. 2015) and often limited take up (Crépon et al. 2015).

In this paper, we examine an innovative loan product that enables people that are traditionally under-served by the financial system to finance a large investment. The financial product addresses many of the criticisms of microfinance: the loan size is relatively large, the repayment schedule is very flexible and instead of a social collateral, the purchased asset serves as collateral. We investigate, how people use the repayment flexibility, whether this is related to the borrowers' financial circumstances, who becomes delinquent under such a flexible repayment schedule and what drives eventual default.

The financial product that we study is a loan tied to a solar panel home system, which can be used for both business and consumption purposes. The lender operates in different countries in East-Africa, where the electrification rate in rural areas is around 10%, and there is thus high demand for alternative electrification sources (World Bank 2017). The loan size is large compared to the typical loans offered by MFIs: between 600 US-\$ and 1,300 US-\$. Repayment is flexible. Borrowers have three years to repay the loan, but can decide when and how much they pay. Each payment also charges the system

— similar to a pay-as-you-go device – and the system switches off automatically if payments are not sufficient. There is thus immediate punishment of non-payment. Every year the system is allowed to be switched off for an accumulated period of 30.5 days (‘grace-period’). When the grace-period is exceeded and the borrower is not able to get back on track the system is deinstalled. The loan contract thereby offers a number of different flexibility features that allow borrowers to align their payments to their cash flows; at the same time the solar system and the electricity provided act as a form of collateral. Our analysis is based on a large administrative data set that provides information on daily repayments over 5.5 years for around 40,000 borrowers in Tanzania. We link this data to information on the borrowers’ socioeconomic background that is collected as part of the loan eligibility interview as well as to location specific information.

We conduct an unsupervised cluster analysis of the repayment behavior in the first year to identify the various ways of how borrowers repay within the framework of such a flexible loan contract and classify borrowers accordingly. We find that borrowers make use of all the elements of the flexible repayment structure. The cluster analysis reveals eight typical repayment patterns. These eight clusters can be further bundled into three broad repayment groups. Within each group the clusters display similar behaviors with some being more extreme versions of the others. A considerable number of borrowers repay in a very steady manner, similar to what would be required in many standard loan contracts. The majority of borrowers, however, repays very frequently: they pay small amounts, often more than once a week and some pay daily. The last group consists of borrowers that make very large payments very rarely but typically do not wait until their balance is consumed.

We next relate the repayment patterns to borrowers’ socioeconomic background to study whether borrowers with higher income fluctuations and fewer resources to smooth consumption are more likely to use the flexibility the contract offers. Indeed, we find that borrowers with steady incomes, such as wage-employed, and borrowers with more resources that allow for consumption smoothing are more likely to repay steadily, while borrowers that tend to have varying cash flows and less resources are more likely to repay very frequently small amounts. We then investigate in more detail to which extent borrowers adjust their payments according to the variations in their cash flows by analyzing farmers’ repayment over the crop cycle using local weather data. Indeed farmers make use of the loan’s flexibility according to their financial circumstances. They pay larger amounts during the harvest period when crop yields materialize and can thereby provide for periods of lower income.

Finally, we analyze who becomes delinquent in the context of such a flexible loan, i.e. who exceeds the allowed grace-period, and whether delinquency is linked to the use of flexibility. We find that borrowers with higher income fluctuations and less resources have a higher likelihood to face repayment difficulties. Furthermore, repayment behavior in the first year can clearly predict delinquency in later years, with borrowers that repaid steadily in the first year being least likely to become delinquent in later years. Investigating repayment behavior before delinquency allows us to classify different types of delinquency; in particular, delinquency that is likely a result of external circumstances (such as income shocks), delinquency because the repayment may not be feasible for the borrowers, and delinquency due to insufficient track-keeping of the payments made and the grace-period used, which is a consequence of the flexible repayment scheme. The last group makes up the largest share among borrowers who

become delinquent. Yet, these borrowers typically get back on track, while borrowers that suddenly stop repaying, probably due to a shock, are the most likely to default on their loan completely.

This study contributes to the literature on innovative financial products to finance large investments, in particular in the area of collateral requirements, loan sizes and repayment flexibility. Social collateral in terms of joint liability loans is still very common, but there is little evidence that it actually improves repayment rates (Giné and Karlan 2014) and there seems to be a general shift towards single-liability loans in microfinance (De Quidt et al. 2018). The only study we are aware of on an alternative form of collateral is Jack et al. (2016), who analyze uptake of loans for rainwater harvesting tanks. They show that being able to collateralize a loan with the asset that is being purchased instead of having the loan secured by guarantors greatly increases loan take up, which in turn improves access to water and individual welfare. The evidence on larger loan sizes is mixed. While Gine and Mansuri (“Money or Management? A Field Experiment on Constraints to Entrepreneurship in Rural Pakistan”) find no impact on business performance from being assigned a larger loan at random, Berge et al. (2015) report positive effects of providing larger long term loans to microfinance clients on business performance and living conditions, yet only when combined with business training. Existing studies on repayment flexibility analyze the impact of making one contract feature more flexible (e.g. reducing installment frequency or allowing to delay an installment). The evidence is mixed: while overall higher flexibility seems to increase business profits, some studies also find an increase in default as borrowers invest in riskier projects (Field and Pande 2008; Field et al. 2013; Czura 2015; Barboni and Agarwal 2018; Battaglia et al. 2018).

Different to the papers discussed above, we examine how borrowers repay in a non-experimental setting following their payment behavior over a long time period using administrative data.¹ The loan schedule we analyze allows borrowers to align repayment completely to their individual cash flows. To the best of our knowledge there is no prior evidence on such a repayment scheme. The rich administrative data allows us to study all nuances of the use of flexibility. The payment patterns and dynamics reveal interesting insights into borrowers’ circumstances and preferences, which data from repayment schemes with fixed installments do not allow for (either a borrower repays at the point of time when the installment is due or not). We thus provide first evidence on a financial product that allows borrowers due to its unique features – in terms of collateral, repayment flexibility and sanctioning mechanisms – to make a welfare-improving investment that they would not have been able to make with the conventional loan products on the market. The product can serve as an innovative example to addresses several of the short-comings of typical microfinance loans.

Clear lessons on how under-served borrowers can finance large investments can be drawn from this study. First, we can show that different people use the flexibility in different ways, when given the choice. At the same time, repayment is adjusted to income streams. This clearly shows that there is demand for different forms of flexibility and that one loan contract does not serve all borrowers well. More research is needed, but one can imagine a scenario where different borrowers select into loan

¹The only study that we are aware of that follows a similar approach are Laureti et al. (2017) who analyze bank account transactions of microfinance clients in Bangladesh, in particular their savings and loan repayment behavior, and link it to shocks such as festivities and political protests. They show that poor borrowers do adjust their saving and borrowing according to their financial needs.

contracts based on their demand for flexibility. When looking at the delinquency analysis, patterns found in other studies still apply. Poorer people and people with unstable incomes are still more likely to become delinquent and economic shocks seem to be the most important reason for default. On the other side, the extensive use of the contract’s flexibility features is not associated with a higher likelihood of eventual default.

The remainder of the paper is structured as follows. In Section 2, we describe the context of the study and the data we use. In Section 3, we identify typical repayment patterns and classify borrowers accordingly; we contrast these patterns with borrowers’ socioeconomic characteristics and study the circumstances of delinquency in such a flexible repayment scheme. Section 4 concludes.

2 Context and Data

2.1 Setting

In this study, we analyze data from a company that sells solar panel home systems on credit. The company operates in several countries in East Africa. We focus on Tanzania, where the company started its operations in 2011 and had about 100,000 customers by the end of 2018 of which 95% bought the system on credit (see Figure 7 in Appendix A.1).² The solar panel home system comes with a number of additional appliances (TV, lamps, radio, mobile phone charger etc.) and is typically sold through the company’s own outlets that are located in towns throughout the country. The target customer group are low-income households living in rural areas, where the rate of electricity is still very low and the demand for alternative electricity sources is high.³ Borrowers can use the system both for private consumption and to generate income. The company offers different system types to match the varying electricity needs and payment abilities of the different customer groups. The systems typically cost between 600 US-\$ and 1,300 US-\$, a substantial financial investment compared to the average income in Tanzania.⁴ The types differ in the panels’ power (between 80 Watts and 200 Watts) and the additional appliances that a customer receives with the system.⁵

Interested borrowers go through a detailed phone assessment interview that tests whether they are eligible for the loan. If borrowers are approved, they pay a downpayment of 5% of the total system costs, receive the system and have then three years to repay the remaining amount. There is an incentive to repay faster: borrowers receive a 10% discount if they repay within 2 years and a 20 % discount if they repay within one year. Furthermore, if borrowers pay the system upfront they have to pay only 75% of the total system price. These incentives make up the implicit interest rate of the loan, which is relatively cheap compared to many loans offered by microfinance institutions in Tanzania,

²As there were many operational changes and trials when the company was first set up, we exclude the first two years in the following analysis.

³In 2016, less than 10% of the population living in rural areas in Tanzania had access to electricity (World Bank 2017)

⁴The annual GNI per capita in current terms in Tanzania for the years 2014 to 2017 was 970\$ <https://data.worldbank.org/indicator/NY.GNP.PCAP.CD?view=chart>

⁵The solar home systems come with a four year warranty and close customer support. In case of technical problems or questions on the system or the repayment modalities, customers can call a toll-free number for support. If needed, a technician is sent to fix the problem.

where annual interest rates typically vary between 25% and 60%.⁶

Borrowers repay the loan using mobile money, typically through MPesa. They either have their own MPesa account or transfer through an MPesa agent.⁷ Each loan payment also charges the system – similar to a pay-as-you-go device. The system is equipped with a sim-card. When a payment is made the system is charged by the number of days that the payment translates to. While the company suggests monthly instalments, borrowers can deviate from this payment scheme; they are in principle free in deciding the amount and timing of their payments. For example, if the total loan amount was 720 US-\$ to be repaid within 3 years, then a payment of 20 US-\$ would charge the system for 1 month but a daily payment of 66 US-cents would be similarly sufficient. A borrower can make a new payment before the old payment is consumed and can also make larger payments and thereby repay the loan faster. When the payment is consumed and no new payment is made, the system switches off automatically. As an illustration, Figure 1 depicts the repayment behavior of a random sample of borrowers who received the system end of April 2017. Each line depicts one borrower’s repayment behavior over the first nine months. The x-axis indicates the date of payment, while the y-axis depicts the size of a payment in terms of number of days a system was charged. It already becomes apparent that borrowers repay in very different patterns.

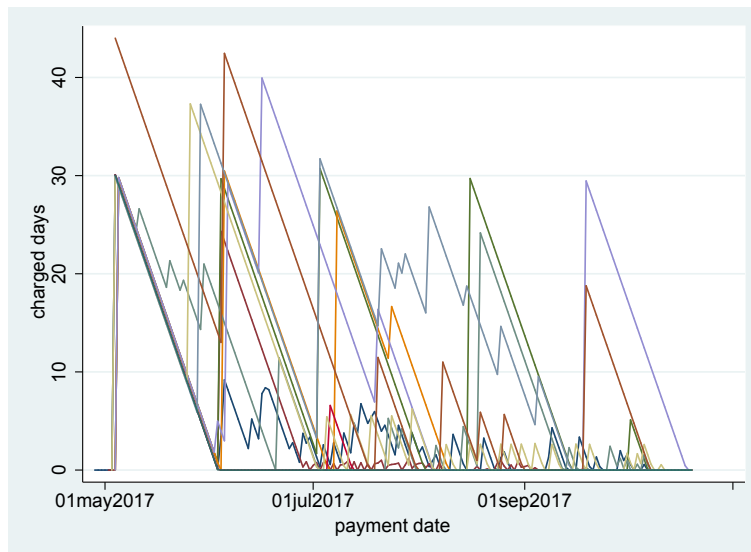


Figure 1: Example of Repayment Behavior

A payment thus provides access to electricity and at the same time repays the loan that was taken out for the system. The only restriction of the repayment plan is the number of days a system is allowed to be switched off (i.e. the number of unpaid days). Borrowers have an accumulated period of 30.5 days of ‘grace period’ per year. If this period is exceeded, the borrower is considered to be delinquent and is attended to by a loan field officer (LFO). The LFO reminds the borrower to repay,

⁶See <https://www.mftransparency.org/microfinance-pricing/tanzania>. The rate of inflation was between 5.5% and 3% in the time period we are observing (<https://data.worldbank.org/indicator/FP.CPI.TOTL.ZG?locations=TZ>).

⁷As of May 2018, over 200,000 MPesa agents were spread around the country (Biztech Africa Industry News, <https://www.biztechafrika.com/article/vodacom-tanzania-marks-10-years-m-pesa>).

first through messages and phone calls, then through on-site visits.⁸ If a borrower is not able or not willing to repay, the system is de-installed and the borrower is not able to buy another system from the company.

2.2 Socio-demographic background

As part of the loan-eligibility interview a large set of variables on borrowers' socio-economic characteristics is collected which include information on occupation, on wealth and resources that can be relied on in case of emergency as well as on prior experience with financial services:⁹ The questions asked during the assessment interview have changed over the years, therefore not all variables are available for all customers. Furthermore, we have information on borrowers' location through GPS coordinates that are recorded when the solar system is installed.

In order to compare the borrower group to the general Tanzanian population, Table 1 presents the variables of interest for all customers that bought one system between July 2013 and November 2018.¹⁰ 17% of borrowers are women and the average age is around 39 years. Borrowers live in relatively large households with 4.5 people in a household on average. Most of the people in our sample are farmers (46%), or are self-employed in other ways (33%); only 22% are wage employed. Many self-employed people run more than one business. During the loan eligibility interview, potential borrowers are asked about the intended purpose of the solar home system: 8% of borrowers indicate that they intended to use the system (also) for business purposes.¹¹ The majority of the borrowers live more than 10km away from the nearest town which we classify as living in rural areas. Compared to the Tanzanian population, borrowers are slightly older and have larger household sizes; they are furthermore more likely to run their own business or to be wage employed.¹²

In terms of wealth and emergency resources, which can be used to smooth consumption, 45% of borrowers hold land, whereas 21% have land that they trade or rent out. 27% employ a domestic worker, 18% hold animals that can be sold in case of emergency and only 7% of borrowers live in a house with a roof that is made up of tiles. Borrowers are asked about their monthly income and expenditures to calculate monthly disposable income. As this data is very messy, we generate an indicator on whether a borrowers' disposable income is above the median in the sample. We use these six indicators to generate a simple wealth index, following the literature on multidimensional

⁸Borrowers have to pay for the number of days they exceeded grace period. Until Sept 2017, the unpaid period was apportioned on the outstanding instalments; since then each borrower has to pay first for the unpaid period before the system is switched on.

⁹This data is not unproblematic. Prospective borrowers have an incentive to provide inaccurate information to questions they expect to affect their eligibility. While there are verification and consistency checks throughout the interview, some of the recorded information might still be biased.

¹⁰Table 11 in Appendix A.1 reports the information for the restricted sample that we use for the main part of our analysis.

¹¹Note that this is likely an underestimation of the true proportion of borrowers who earn money with the system - many started to earn money with the system later in their repayment cycle. Surveys with a subset of the borrowers indicate that one out of four borrowers earn money with the system at some point.

¹²According to the 2014-2015 wave of the Living Standards Measurement Survey, 29% of households are headed by a women. The average age of the population is 22, while the average age of a household head is 44. The average number of household members is 3.7. Half of the household heads report farming as their main activity, where as about 17% run their own business and 13.5% are wage employed (<https://microdata.worldbank.org/index.php/catalog/lsms>).

Table 1: Borrower Characteristics

Variable	Mean	SD	Min	Max	Median	Obs.
female	0.17	0.38	0.00	1.00	0.00	96070
hh members	4.45	1.94	1.00	30.00	4.00	95832
age	39.20	11.07	15.00	89.00	37.00	50435
farmer	0.46	0.50	0.00	1.00	0.00	76718
wage-employed	0.22	0.41	0.00	1.00	0.00	76718
self-employed	0.33	0.47	0.00	1.00	0.00	76718
number of businesses	1.22	0.98	0.00	5.00	1.00	54222
rural (10km)	0.56	0.50	0.00	1.00	1.00	67313
use of formal finance	0.20	0.40	0.00	1.00	0.00	54222
use of informal finance	0.28	0.45	0.00	1.00	0.00	54222
system loanamount (in 100 US Dollar)	10.34	3.71	2.41	25.43	9.26	88744
business purpose	0.08	0.26	0.00	1.00	0.00	52800
trading or renting land	0.21	0.41	0.00	1.00	0.00	54222
monthly disposable income - above median	0.47	0.50	0.00	1.00	0.00	52785
owns land	0.45	0.50	0.00	1.00	0.00	54222
employs domestic workers	0.27	0.44	0.00	1.00	0.00	54222
owns animals than can be sold in emergency	0.18	0.38	0.00	1.00	0.00	54222
roof made out of tiles	0.07	0.25	0.00	1.00	0.00	96266
wealth index	1.61	0.99	0.00	4.00	2.00	51767

Notes: Borrowers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists.

poverty (Alkire and Foster 2011). In terms of experience with formal financing, only 20% of the sample currently hold or have previously held a formal loan. 28% are member of an informal financial institution such as savings and credit cooperative (SACCO). In later regressions we control for the size of the loan, which is about 1000 US-\$ on average. There are larger differences between our sample and the general Tanzanian population in regards to use of finances. According to the 2017 wave of the FINDEX data only 5.3% of the population borrowed from a formal financial institution, while 9.5% borrowed from a savings club.¹³

2.3 Data

Our main analyses is based on the administrative repayment data of the company. We primarily study borrowers' daily repayments to assess whether they use the flexibility that the loan contract allows. In particular, we are interested in the timing, frequency and amount of payments as well as timing, frequency and duration of system shutdowns. For each day in the life of a borrower, we know the amount of payments a borrower made and how much charge-time such payment translates to. We use this information to calculate by how much a system was still charged when a payment was made as well as whether and for how long a system was switched off in case no payment was made.

We include all borrowers in our data analysis, who bought the system on credit and for whom we have at least one year of repayment data. We exclude borrowers who started before July 2013, who bought more than one system (as in these case repayment cannot be attributed to a specific system) or whose loan was rescheduled. This leaves us with a total of 38,407 borrowers and 32,955,679 borrower-day observations for a time period from July 2013 to November 2018 .

¹³See https://globalfindex.worldbank.org/#about_focus

Table 2: Repayment behavior of median borrower by repayment year

YEAR	num payments	avg. paiddays	sd. paiddays	min paiddays	max paiddays	avg. balance	sd. balance	no. of shutdowns	shutdown days	no. of obs
1	17.00	20.81	7.26	3.75	31.99	1.40	2.42	8.00	8.06	38405
2	25.00	8.32	2.41	1.93	30.07	0.57	0.90	11.00	13.05	37016
3	19.00	7.52	2.38	1.88	26.91	0.55	0.84	9.00	9.98	22692
Avg.	20.00	11.92	4.24	2.50	30.50	0.80	1.30	9.00	10.10	

Notes: Borrowers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists. Borrowers for which less than one year of repayment data exists or who bought more than one system excluded.

Table 2 shows important features that describe the median repayment behavior of the borrowers over the three years. It becomes apparent that the repayment behavior is not as it is required in most conventional loan contracts. 50% of the borrowers pay 20 or more times per year and on average for 12 days.¹⁴ The average balance on the account at the point when the borrower makes the next payment is less than one day. The median borrower has his system shut down for 8 to 11 times per year and accumulates around 10 days of grace period per year; the average shut down thus lasts for just about one day. From Table 2 we can further infer that there is considerable heterogeneity in the way that people repay their loans. Both the standard deviations of the paid days and the standard deviation of the balance are very large.¹⁵

3 Analysis

The repayment scheme allows for a variety of ways to repay the loan: borrowers can repay any amount at intervals that they choose, meaning that they can repay, for example, very small amounts every other day or large amounts few times a year. To understand how borrowers use the loan’s flexibility, in the following, we first identify typical repayment patterns and classify borrowers accordingly (Section 3.1). We then investigate which customer characteristics are linked to specific repayment patterns, focusing in particular on income fluctuations and access to consumption smoothing resources, and study whether borrowers that face highly varying cash flows throughout the years make use of flexibility to match their repayment to their cash flows (Section 3.2). Finally, we analyze who becomes delinquent under such a flexible repayment schedule and what drives eventual default (Section 3.3).

3.1 Specifying Repayment Types

Our repayment data constitutes a time series per customer, i.e. daily data with information about the remaining balance and payments made. We aim to understand how and in which combination borrowers use the different elements of flexibility provided to them. As already apparent from the summary statistics depicted in Table 2, there is large variation in how borrowers use the flexibility that the loan provides. In order to allow for the complexity of different repayment patterns without any prior on potential payment types we use an unsupervised clustering approach without ex-ante

¹⁴The higher average paiddays in the first year are a result of the first installment that the borrower makes which has to cover at least 1 month.

¹⁵A more detailed description of the repayment feature distributions can be inferred from the kernel density plots depicted in Appendix B.3.

classification labels to classify borrowers according to their repayment behavior in the first year. We focus on the first year for two main reasons. First, we can thereby fix the length of time to be the same for each borrower.¹⁶ Second, focusing on repayment patterns in the first year only allows us to relate it to repayment behavior in later years.

3.1.1 Method

There are multiple ways to cluster time series data (for reviews see Liao (2005) and Aghabozorgi et al. (2015)). One can either use the raw time series data, use parameters of time series models, or use features extracted from the data. We decided against using clustering of raw data. The dimension of the data would be very large with each point in time being one dimension. This makes the interpretation of the cluster results very difficult. Moreover similar behavior that is just shifted in time would not be classified as similar.¹⁷ Using the parameters of time series models would resolve the problem of capturing time-shifted behavior correctly, but the interpretation of the clusters which are described in terms of time series parameters would remain difficult. Furthermore, the hyper parameters of the time series models would have a very strong influence on the cluster result.¹⁸ Therefore, we decided to use the third strategy, i.e., to define features that describe the time series data. To fulfill our needs, we require features that describe the payment behavior over time in the first year. One important property that all features must fulfill is that they may not use any information about the future, i.e., information from the second or third year of payments.¹⁹ To capture the several dimensions of repayment, the selected features reflect the payment behavior in terms of how often payments are made, how large the payments are, how often and how long the system is shut down, and how many charge days are remaining, when new payments are made. The nine selected features are depicted in Table 3.

To capture as much of the richness of the data as possible, we decided to use an implementation of a distribution based clustering that determines Gaussian mixture models using the expectation maximization (EM) clustering algorithm (Scrucca et al. 2016).²⁰ Gaussian mixture models describe

¹⁶Borrowers vary greatly in the time they need until they repaid their loan; yet more than 90% take at least a year.

¹⁷For example, consider two customers A and B. A starts with a downpayment that charges the system for 50 days, B starts with a downpayment that charges the system for 30 days. From then both pay regularly always charging their system for 20 days. From the payment behavior perspective, both are extremely similar. However, when we consider the raw values, there is always an absolute difference of 15 charged days between the customers.

¹⁸For example, for ARIMA time series models the hyper parameters define how many steps in the past values and how much of the previously introduced white noise is used to calculate the next time series value, as well as the order of differencing that is used to make the data stationary. In our case, good hyper parameters may vary not even from cluster to cluster, but also from customer to customer. However, to effectively use the parameters of a time series model for clustering, we would require that the hyper parameters are always the same, which would mean that many time series models would be using bad hyper parameters resulting in a bad fit.

¹⁹While this may sound trivial, this led to the exclusion of potentially helpful features like the shutdown durations, as a shutdown may start in the first but only finish in the second year.

²⁰There are a number of different algorithms to derive the cluster groups, e.g., hierarchical clustering (Ward 1963) density based clustering (Ester et al. 1996), centroid-based clustering (MacQueen 1967), and distribution-based clustering (Dempster et al. 1977). Hierarchical clustering and density based clustering have the drawback that data points that are close to each other describe the clusters. Thus, there is no closed form of a cluster description (e.g., a formula) that we can use for subsequent analyses or to describe the borrower groups. While centroid-based clustering provides a closed form of a cluster description in form of a centroid, this centroid only describes the borrower groups using the mean of all its borrowers. No information regarding how features are related or how they are spread within a cluster is part of the cluster description.

Name	Description
num payments	number of distinct payments made
avg. paiddays	mean number of days that the system was recharged with during payments, excluding the initial downpayment.
min. paiddays	minimum number of days that the system was recharged with during the payment, excluding the initial downpayment.
max. paiddays	maximum number of days that the system was recharged with during the payment, excluding the initial downpayment.
sd paiddays	standard deviation of the days that the system was recharged with during payments, excluding the initial downpayment.
avg. balance	mean number of charged days remaining when a borrower recharged the system with another payment.
sd balance	standard deviation of the charged days remaining when a borrower recharged the system with another payment.
shutdowndays	number of days the system was shut down
no of shutdown	number of times a system was shutdown.

All features are based on the borrowers' first repayment year.

Table 3: Features for cluster analysis

data in terms of components, whereas each component is a multivariate Gaussian distribution. Thus, each borrower group we determine is described by the mean value, variance, and co-variance of our features. Using this clustering approach has multiple advantages for our case. First, all borrowers can be assigned to the most likely cluster, i.e., the cluster whose distribution best explains the behavior of the borrower. Additionally, we know for each cluster the probability that a borrower is assigned to this cluster. This way, we can also identify borrowers that are typical for a cluster, i.e., have an extremely high probability that they belong to the cluster. Second, by considering the mean values of the gaussian mixtures of each clusters, we have a clear description of the most important feature of each cluster, which we can use for comparisons of the clusters and identification of the type of borrower that is within the cluster. Third, because EM is distribution based, we can use the likelihood that the gaussian mixture models explain the data to estimate the number of clusters k required to explain our data. Thus, we can see how much better the explanation of the data is with more clusters and stop adding more clusters if the gain is only very small.²¹

3.1.2 Results

We first calculated for each borrower's first year of repayment the nine selected features described in Table 3. We then conducted the cluster analysis stepwise with 1 to 20 clusters. Comparing the Bayesian Information Criteria of the resulting models suggests an optimal number of 8 clusters (see Figure 8 in Appendix B.1). All clusters and their repayment features are depicted in Table 4.²²

²¹Using EM clustering has the drawback that the complete description of the gaussian mixtures models requires the covariance matrices for each cluster. This means we have for each cluster not only the k mean values, but also a $k \times k$ covariance matrix. However, for our case mean values are sufficient as we only need to identify which kind of borrow is within a cluster.

²²The correlation across features is depicted in Figure 9 in Appendix B.4

Table 4: Mean of Features of Year 1 by Cluster Group

	num payments	avg. paiddays	sd. paiddays	min paiddays	max paiddays	avg. balance	sd. balance	no. of shutdowns	shutdown days	no of obs	%
1	11.42	32.79	5.51	28.98	46.89	5.13	5.26	4.23	10.41	3,663	9.54
2	13.96	26.84	8.16	6.91	33.14	3.79	5.28	4.78	6.97	4,171	10.86
3	12.08	30.72	0.62	30.33	32.46	1.48	1.56	4.38	5.16	5,211	13.57
4	41.25	11.06	8.29	2.54	31.63	1.52	2.79	19.00	12.90	7,418	19.31
5	62.52	9.56	9.78	1.85	49.34	4.26	6.15	22.79	25.15	5,233	13.63
6	127.98	3.42	4.49	0.43	30.64	0.53	1.37	60.05	43.56	6,019	15.67
7	11.85	36.68	22.55	13.37	83.29	10.80	13.25	4.70	15.28	5,119	13.33
8	17.99	41.72	28.94	15.11	99.94	35.66	21.68	7.82	51.96	1,573	4.10
all	43.10	20.78	9.32	10.84	45.30	4.96	5.53	18.65	19.36	38,407	100.00

Notes: Borrowers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists. Borrowers for which less than one year of repayment data exists or who bought more than one system excluded.

The distribution of features across clusters already indicates that repayment cannot be described by one specific feature but is a combination of a number of different aspects.²³ Based on the feature distribution we classify the eight clusters into three main repayment types: steady repayer, frequent repayer and irregular repayer.

Steady Repayers (cluster groups 1-3):

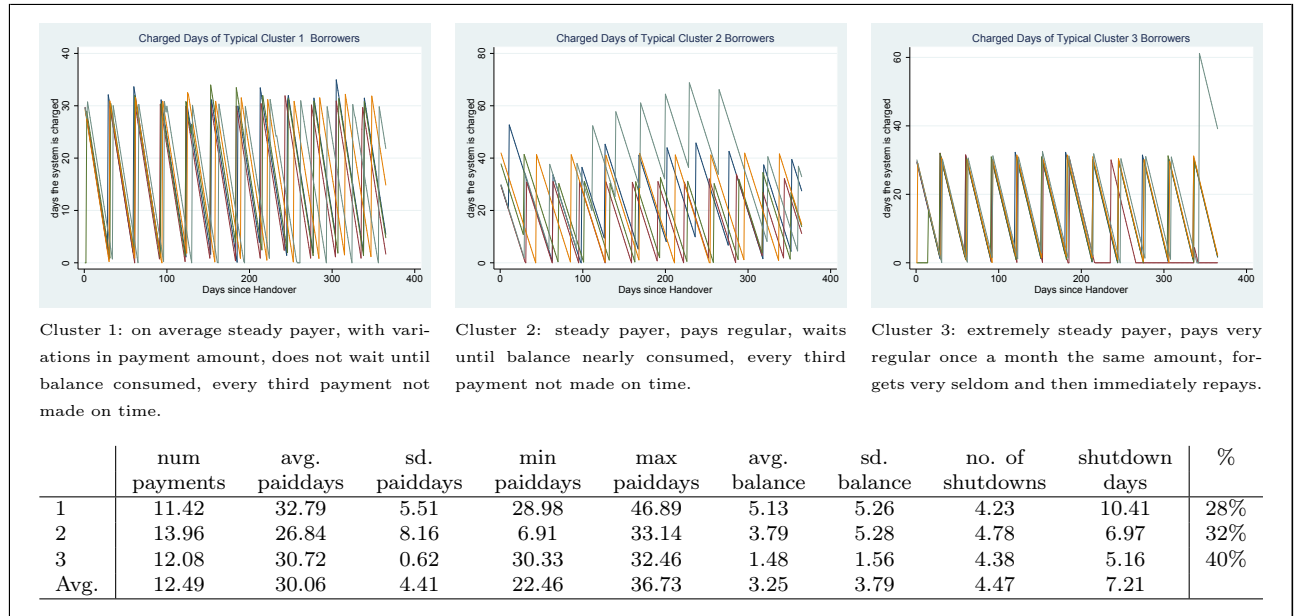


Figure 2: Cluster Groups of Steady Repayer

Borrowers in the first three clusters can be defined as Steady Repayers. They are characterized by monthly payments that are in accordance with the monthly installments suggested by the company and a limited use of grace period. Figure 2 depicts the repayment behavior of five typical borrowers from each of these three groups.²⁴ Cluster 3 contains the most extreme sort of

²³Note that in Table 4 only averages are depicted. Within groups certain features can vary considerably (see Appendix B.4).

²⁴The depicted borrowers in this and the following two figures are randomly drawn from those borrowers that had the 5% highest probability of being assigned to the respective cluster.

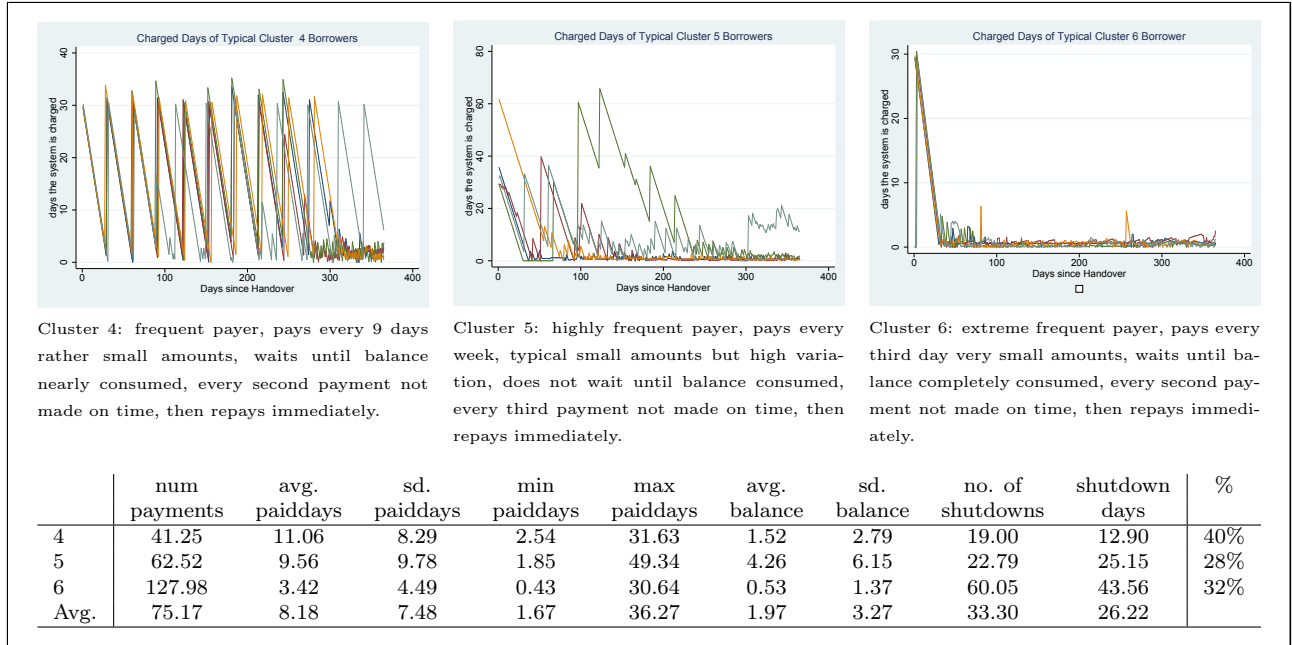


Figure 3: Cluster Groups of Frequent Repayer

repayment behavior within this classification. Borrowers in this cluster repay the same amount at the same time each month. They pay exactly enough to keep the solar system running for one month. Hence, the balance on their account is almost used up when they make their next payment. At the same time, borrowers in this group have very few occasions during which their system is shut down, indicating that they typically have sufficient cash on hand and hardly ever forget to repay. Borrowers in the other two cluster groups still pay very regular but there is some more variation in the timing and amount and longer shutdown durations. One third of all borrowers can be classified as Steady Repayers. They repay in a manner that is in line with repayment behavior required in most standard loan contracts in developed countries. This indicates that it is possible for poor people with low incomes to repay a loan in a steady and regular manner even when the contract offers more options of flexibility.

Frequent Repayers (cluster groups 4-6):

Borrowers in clusters 4, 5 and 6 can be classified as Frequent Repayers (depicted in Figure 3). Borrowers in this group frequently pay very small amounts. The three groups differ in their frequency, with cluster 6 being the most extreme case. On average, borrowers in this group make a payment about every five days. Payments are very small and only cover the time between payments, hence the balance is mostly consumed when payments are made. At the same time, every other payment is not made on time. These two behaviors combined result in a large number of shut downs, with an average of 26 accumulated shut down days. The shutdown days of the extreme frequent repayers (cluster 6) even exceed the 30.5 days of shut down that a borrower can have without being considered in default. However, the average shut-down duration is very

short: borrowers in this group have their system typically shut down for less than one day until they make the next payment. The group of Frequent Repayers covers the largest number of borrowers. 49% of all borrowers exhibit such a payment pattern. This shows that amongst low-income borrowers, many prefer to repay very frequently as is common for many microfinance loan contracts. However, borrowers also make extensive use of the option of making larger payments at times - a possibility that typical microfinance loan contracts do not provide.

Irregular Repayers (cluster groups 7 & 8):

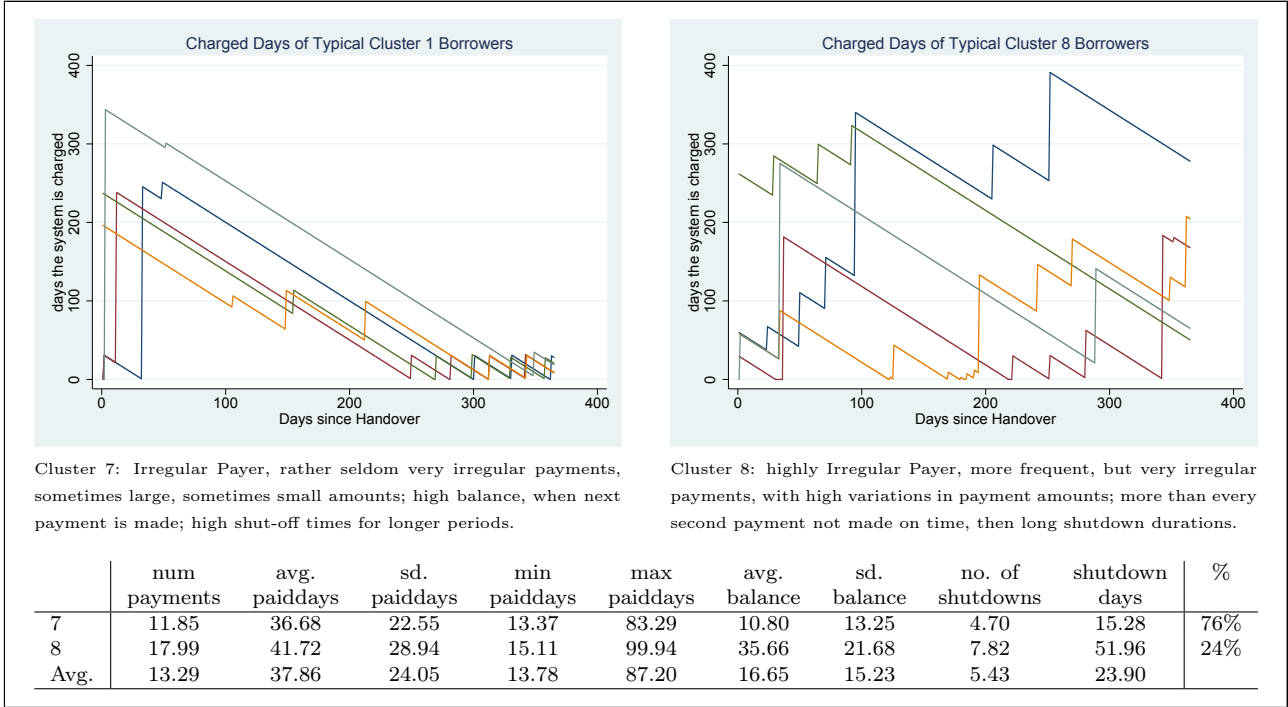


Figure 4: Cluster Groups of Irregular Repayer

The third group, consisting of cluster group 7 and 8, containing Irregular Repayers (depicted in Figure 4). Borrowers in this group pay very rarely and the amount that is paid fluctuates considerably. In the most extreme case, group 8, borrowers repay on average every 20 days. There is a large variation in the amount of payment made varying between 14 and 87 days on average. Borrowers in this cluster have only few shut downs on average, but these tend to be long and build up over time. When borrowers in the irregular group make a payment, they typically do not wait until their balance is consumed (in the extreme case of cluster 8 the average balance is 36 days when the next payment is made). Borrowers in group 7 also pay very irregular but are less extreme in terms of shut downs. Irregular Repayers make up the smallest group. About 17% of borrowers can be classified as irregular payers. It is less clear what drives this behavior.

To summarize, the pattern analysis reveals clear differences in the way that borrowers use the flexibility that the loan contract provides to them. About one fifth does not use the flexibility at all. These people

repay very steadily, as it would be required in most loan contracts in developed countries. The largest proportion of borrowers, however, do use the flexibility and combine the different elements of flexibility in various ways.

3.2 Repayment Patterns and Financial Circumstances

3.2.1 Determinants of Repayment Types

Are certain socioeconomic characteristics related to a specific repayment behavior? In particular, are borrowers with higher income fluctuations and less resources to smooth consumption more likely to use the loan contract's flexibility?

We estimate the following specification:

$$CG_{i,d,my}^J = \alpha^J I_i + \beta^J R_i + \gamma^J X_i + \xi_{my}^J + \delta_d^J + \epsilon_{i,d,my}^J \quad (1)$$

$CG_{i,d,t}^J$ describes the probability that a borrower i living in district d and having received the system at the month-year my belongs to the repayment group J , with $J = \{SP, FP, IP\}$. The probability is derived by the expectation maximization algorithm described in the previous section. We aim to explain a borrower i 's probability displaying a certain repayment pattern CG_i^J by the income fluctuations she is facing, I_i , the resources she has available to smooth consumption, R_i , and individual characteristics, X_i . In addition, we include handover date fixed effects, ξ_{my} , and district fixed effects, δ_d . We estimate the likelihood for each of the three repayment groups separately through a generalized linear model with a logit transformation of the response variable CG^J and a binomial distribution. Standard errors are clustered at district level to account for regional variation.

I includes variables that proxy a borrower's income fluctuations throughout the year: namely, the borrower's main occupation, which we broadly classify in farmer, self-employed and wage-employed, the number of businesses a borrower has and whether the borrower indicated at the loan assessment interview to earn additional money with the system. R includes variables that proxy a borrower access to resources to smooth consumption over time, namely a wealth index, use of formal finance, and use of informal finance. Finally, X includes individual characteristics, namely whether the borrower is female, the number of household members, whether the borrower lives in a rural area (at least 10km from the next town) and the loan the borrower took for the system, which can to some extent also be an indicator of the borrower's wealth, as larger loan amounts were only given to more affluent borrowers.

As some of the variables are only available for a subset of the borrowers, we restrict the analysis first to variables that are available for nearly all borrowers and then expand the analysis incorporating additional informative variables at the expense of losing approximately half of the observations.

Results are depicted in Table 5. Columns (1) and (2) report estimated effects for the likelihood of belonging to the Steady Repayer group, columns (3) and (4) for the likelihood of belonging to the Frequent Repayer group, and columns (5) and (6) for the likelihood of belonging to the Irregular Repayer group. Columns (2), (4) and (6) describe the estimated effects for the restricted sample.²⁵ We

²⁵See Tables 13 to 15 in Appendix C.1 for the results for each of the eight cluster groups

see a clear pattern: borrowers that work in occupations with more irregular income (such as farmers or self-employed), that have several small businesses and those that use the system to earn additional money have a higher likelihood to repay very frequently, while wage employed borrowers repay very steadily on a monthly basis. This indicates that borrowers use the flexibility to align their payments to their individual cashflows. Furthermore, borrowers with access to consumption smoothing resources, such as wealth and formal loans, are more likely to pay steadily on a monthly basis instead of paying small amounts very frequently, while use of informal financial services is rather associated with frequent payments. Irregular repayment patterns can be explained less well. Borrowers repaying irregular seem to be more wealthy and have fewer side businesses, furthermore they seem to neither use formal nor informal financing. It is conceivable that these are borrowers that earn the main bulk of their income only once or twice per year, e.g. farmers with large land plots and no side businesses, and then make their payments.

Table 5: Use of Flexibility and Socioeconomic Characteristics

	Prob SP		Prob FP		Prob IP	
	(1)	(2)	(3)	(4)	(5)	(6)
female	1.177 (0.040)***	1.165 (0.055)***	0.850 (0.027)***	0.887 (0.039)***	1.009 (0.033)	0.995 (0.062)
hh members	1.049 (0.008)***	1.050 (0.013)***	0.928 (0.007)***	0.925 (0.010)***	1.043 (0.008)***	1.075 (0.014)***
rural (10km)	0.930 (0.033)**	0.982 (0.045)	1.038 (0.054)	0.991 (0.054)	1.052 (0.061)	1.049 (0.074)
system loanamount (log)	1.174 (0.063)***	1.224 (0.078)***	0.851 (0.052)***	0.842 (0.052)***	1.005 (0.068)	1.019 (0.071)
farmer	0.774 (0.025)***	0.809 (0.032)***	1.100 (0.034)***	1.110 (0.043)***	1.233 (0.038)***	1.107 (0.066)*
wage-employed	1.569 (0.047)***	1.591 (0.078)***	0.655 (0.021)***	0.693 (0.033)***	0.971 (0.039)	0.938 (0.074)
number of businesses		0.951 (0.024)**		1.095 (0.025)***		0.906 (0.026)***
business purpose		0.755 (0.067)***		1.373 (0.101)***		0.808 (0.089)*
wealth index		1.041 (0.024)*		0.919 (0.018)***		1.110 (0.031)***
use of formal finance		1.278 (0.065)***		0.859 (0.036)***		0.898 (0.061)
use of informal finance		0.866 (0.053)**		1.200 (0.061)***		0.881 (0.055)**
Constant	0.636 (0.119)**	0.173 (0.028)***	0.521 (0.099)***	1.709 (0.285)***	0.271 (0.068)***	0.317 (0.060)***
Observations	35146	18789	35146	18789	35146	18789
Handoverdate FE controls	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: GLM with logit transformation of response variable & binomial distribution. Relative risk estimator, standard errors in parentheses, clustered on district level. Prob SP - Probability of paying in accordance with clusters of steady repayers. Prob FP - Probability of paying in accordance with clusters of frequent repayers. Prob IP - Probability of paying in accordance with clusters of irregular repayers. Borrowers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists. Borrowers for which less than one year of repayment data exists or who bought more than one system excluded. */**/** denote significance levels at 10/5/1 percent respectively.

3.2.2 Cash flow fluctuations and repayment

The analysis above reveals a clear relationship between occupation and repayment patterns. Borrowers that work in occupations with higher cashflow variation tend to use the flexibility to a larger extent. But do borrowers make use of the provided flexibility to match their payments directly to their income flows, i.e. paying more when they have more cash on hand? We do not have good information on borrowers' exact cashflows but for farmers we can assume that their income is particularly high during harvest period. For the following analysis, we thus focus on repayment of farmers over the crop cycle.

In a country such as Tanzania, where seasons and crop cycles vary considerably across regions, there is no clear definition as to when the relevant local growing, harvesting and lean season starts and ends. We define the local growing seasons on ward level using the method developed by Dunning et al. (2016), where the onset and cessation of a rainy season are measured using local daily precipitation data (CHIRPS). We thereby have a very precise measure of the start and end of the local growing season and can thus define time periods as being within or outside the growing season. While we herewith have a good proxy for each ward's start of the harvest season (which sets in after the growing season), we do not know it's length and thus when the lean season sets in. Furthermore, we are agnostic as to when exactly farmers sell their harvest. We deal with this by splitting each of the two time periods, growing season and outside growing season, which are defined on the ward level, in three parts of equal length: start of (outside) growing season, mid of (outside) growing season and end of (outside) growing season.

We then analyze repayment during these so defined periods, assuming that the first two periods of the outside growing season are likely the time periods when crops are harvested and harvest income is realized. If farmers make use of the repayment flexibility, we should expect higher payments in these periods.

Table 6: Average Daily Payment Amount and Season Weeks

	Farmers		Maize Farmers Only		Non-farmers	
	(1)	(2)	(3)	(4)	(5)	(6)
Outside Growing Season	0.122 (0.008)***		0.187 (0.011)***		0.044 (0.007)***	
Outside Growing x Start Season		0.078 (0.014)***		0.095 (0.021)***		0.039 (0.012)***
Outside Growing x Mid Season		0.210 (0.013)***		0.315 (0.019)***		0.089 (0.011)***
Outside Growing x End Season		0.073 (0.013)***		0.137 (0.019)***		0.003 (0.011)
Mid Season		-0.079 (0.013)***		-0.114 (0.020)***		-0.030 (0.011)***
End Season		-0.011 (0.013)		-0.073 (0.020)***		-0.008 (0.011)
Constant	1.037 (0.005)***	1.069 (0.010)***	0.982 (0.008)***	1.047 (0.014)***	1.073 (0.005)***	1.087 (0.008)***
Observations	741849	741849	296295	296295	766772	766772

Notes: Borrower fixed effect regression; standard errors in parentheses. Borrowers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists. Borrowers for which less than one year of repayment data exists or who bought more than one system excluded. */**/** denote significance levels at 10/5/1 percent respectively.

We analyze this by estimating a borrower-fixed effect model regressing average daily repayments per week on the local seasons. Results are reported in Table 6. Columns (1) and (2) report results for all farmers. As the cycles of some crops are less dependent on rainy seasons (e.g. coffee, tobacco, bananas) we restrict the sample in columns (3) and (4) to farmers producing a crop whose cycle follows the rainy season closely – namely maize, the main cash crop in Tanzania. As a comparison, we also analyze the impact of seasonality on non-farmers’ repayment (columns (5) and (6)).

Indeed, we find that farmers pay larger amounts outside the growing season (about 12%). Particularly so in the second part, where payment increases by 20%, which is probably when yield income materializes. As expected, the results are even stronger for maize farmers, who pay 30% more in the second part outside the growing season compared to the beginning of the growing period. The effect is considerably smaller for non-farmers, but still their income and thus their repayment seems do be seasonality dependent – this is not unreasonable as many non-farmers run small businesses (such as shops or bars) which benefit from the increase in farmers’ income and their spending capacities during harvest period.

We see similar patterns when looking at the average number of payments per week (see Table 7). Borrowers pay more frequent and smaller amounts when less cash is on hand during the growing season and less frequent outside the growing season, in particular in the second and third part. Again, the effects are more pronounced for maize-farmers and much less pronounced for non-farmers.

Table 7: Number of Payments per Week and Season Weeks

	Farmers		Maize Farmers Only		Non-farmers	
	(1)	(2)	(3)	(4)	(5)	(6)
Outside Growing Season	-0.106 (0.002)***		-0.142 (0.004)***		-0.067 (0.002)***	
Outside Growing x Start Season		0.105 (0.004)***		0.224 (0.008)***		0.043 (0.004)***
Outside Growing x Mid Season		-0.147 (0.004)***		-0.192 (0.007)***		-0.086 (0.004)***
Outside Growing x End Season		-0.248 (0.004)***		-0.411 (0.007)***		-0.143 (0.004)***
Mid Season		0.107 (0.004)***		0.160 (0.007)***		0.052 (0.004)***
End Season		0.172 (0.004)***		0.295 (0.007)***		0.089 (0.004)***
Constant	0.911 (0.002)***	0.813 (0.003)***	1.188 (0.003)***	1.030 (0.005)***	0.884 (0.002)***	0.835 (0.003)***
Observations	741849	741849	296295	296295	766772	766772

Notes: Borrower fixed effect regression; standard errors in parentheses. Borrowers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists. Borrowers for which less than one year of repayment data exists or who bought more than one system excluded. */**/** denote significance levels at 10/5/1 percent respectively.

In summary, we see evidence that borrowers make use of the flexibility that the repayment scheme provides; they pay pre-emptively - i.e. they pay larger amounts when cash is on hand and are thereby able to provide for the months of lower cashflow.

3.3 Repayment difficulties

The loan contract allows borrowers to choose a repayment schedule that fits their circumstances. But can borrowers in such a setting repay successfully, or is a flexible repayment contract a risk, likely to corrupt payment discipline, as proponents of rigid repayment schedule argue (Lahie et al. 2017)? This will be the focus of the following section. We first investigate who faces repayment difficulties in such a loan setting. We then distinguish between different types of delinquency by studying repayment as well as shut-down dynamics just before a borrower becomes delinquent and link these types to the use of flexibility as well as eventual default.

3.3.1 Determinants of delinquency

In most regular microfinance schemes, one can only observe whether a borrower defaults. In our context, we observe whether a borrower already had difficulties before defaulting.²⁶ We can distinguish between the following four states: (1) whether a system is shut-off due to insufficient payments – this type of non-payment is contractually allowed as long as it does not accumulate to more than 30.5 days per year; (2) whether a system is shut-off for more than 30.5 days accumulated over one year, i.e. the borrower exceeded the allowed grace-period; (3) whether a borrower’s system was shut-off for over three consecutive months, what we in the following call ‘quasi-default;’ and finally (4), whether the system was de-installed, i.e. the borrower defaulted. Once a borrower exceeded the allowed grace period, there are several reminders and in-person visits by the loan field officer before the system would be de-installed. This process was not standardized during the time period we analyze and thus the time that passed from when the borrower exceeded the allowed 30.5 days until a system was de-installed is to a certain extent at the discretion of the loan field officer. For the following analysis, we therefore mainly focus on the second state: we define a borrower to be delinquent, if her system has been shut off for more than 30.5 days a year, i.e. the borrower cumulatively missed more than 30.5 days of payment within one year. In the first and second year, 25% of the borrowers become delinquent at least once. Most of them get back on track, only a small proportion defaults eventually.

In Table 8 we report results on the socio-economic determinants of delinquency, estimating the following specification, which mirrors Specification 1:

$$DEL_{i,d,my} = \alpha I_i + \beta R_i + \gamma X_i + \xi_{my} + \delta_d + \epsilon_{i,d,my} \quad (2)$$

where $DEL_{i,d,my}$ indicates whether a borrower i living in district d who has received his system in the month-year my became delinquent at some point during his repayment cycle. I_i , R_i , X_i , ξ_{my} and δ_d are defined as in Specification 1. Standard errors are clustered on district level.

Results are reported in Table 8. Column (1) reports the results for the full sample, while column (2) reports results for the restricted sample. Again, we see that income fluctuations and access to resources play a crucial role. Farmers are more likely than self-employed to face repayment difficulties,

²⁶For all of the following analyses, we exclude borrowers who became delinquent in the first 6 months of their loan cycle to account for borrowers who received the system due to misinformation in the application process (e.g. either borrowers who should not have received a loan in the first place or borrowers who made the purchase decision based on inaccurate information).

wage-employed are less likely. Customers that had planned to use their system for their business are also more likely to become delinquent.²⁷ On the other side, richer households and households with more household members, which can also be interpreted as a form of resource to smooth consumption, are less likely to come into repayment difficulties.

Table 8: Probability of Becoming Delinquent and Socioeconomic Characteristics

	(1)	(2)
female	1.009 (0.031)	1.020 (0.044)
hh members	0.965 (0.008)***	0.960 (0.011)***
rural (10km)	1.082 (0.052)*	1.039 (0.042)
system loanamount (log)	0.972 (0.050)	0.979 (0.077)
farmer	1.062 (0.056)	1.110 (0.046)**
wage-employed	0.788 (0.033)***	0.806 (0.041)***
number of businesses		1.027 (0.026)
business purpose		1.178 (0.083)**
wealth index		0.949 (0.018)***
use of formal finance		0.994 (0.045)
use of informal finance		0.988 (0.042)
Observations	33267	17597
Handoverdate FE controls	Yes	Yes
District FE	Yes	Yes

Notes: Logit estimation. Relative risk estimator, standard errors in parentheses, clustered on district level. Probability of exceeding the 30.5 days of allowed grace period. Borrowers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists. Borrowers for which less than one year of repayment data exists or who bought more than one system excluded. Borrowers who became delinquent in first 6 months excluded. */**/** denote significance levels at 10/5/1 percent respectively.

These findings are reminiscent of our earlier findings on the determinants for repayment patterns. The question thus arises whether repayment behavior, in particular the use of flexibility, can predict delinquency in the long run. To limit mechanical correlations we regress delinquency in the third year on repayment patterns in the first year, more specifically on the probability that a borrower followed a frequent repayment pattern and the probability of following irregular repayment patterns (omitting

²⁷On first sight, this seems surprising: one could have expected that borrowers who generate income with their system have additional means to repay the loan and should thus be less likely to face repayment difficulties. Yet, the type of business for which the solar panel is typically used (e.g. small shops, bars, village cinemas, phone charging) often underlie considerable income fluctuations. Furthermore, if a borrower plans to start a new business with the help of the system (e.g. barber shops) there is a risk that the planned business does not work out.

the probability of steady repayment). Results are reported in Table 9; from Column (4) onwards we restrict the sample to customers who have not become delinquent before the third year to further reduce potential mechanical correlation.

Table 9: Probability of Becoming Delinquent in Year 3 and Repayment Pattern in Year 1

	(1)	(2)	(3)	(4)	(5)	(6)
Prob FP	2.364 (0.082)***	2.388 (0.089)***	2.132 (0.257)***	1.684 (0.109)***	1.687 (0.106)***	1.117 (0.481)
Prob IP	1.396 (0.074)***	1.411 (0.080)***	2.308 (0.372)***	0.831 (0.062)**	0.819 (0.065)**	1.041 (0.454)
female		1.002 (0.050)	0.987 (0.114)		1.136 (0.074)*	1.303 (0.440)
hh members		0.986 (0.009)	1.014 (0.027)		0.981 (0.016)	1.109 (0.090)
rural (10km)		1.078 (0.043)*	1.021 (0.088)		1.061 (0.048)	1.097 (0.316)
system loanamount (log)		1.049 (0.057)	1.177 (0.195)		1.003 (0.095)	1.793 (1.043)
farmer		0.986 (0.056)	1.103 (0.120)		0.962 (0.066)	1.261 (0.504)
wage-employed		1.011 (0.052)	1.090 (0.173)		0.971 (0.061)	1.018 (0.576)
number of businesses			0.945 (0.054)			0.683 (0.171)
business purpose			1.599 (0.200)***			1.704 (0.787)
wealth index			0.895 (0.050)**			0.857 (0.173)
use of formal finance			1.227 (0.143)*			1.456 (0.682)
use of informal finance			0.829 (0.114)			0.490 (0.262)
Observations	22704	20855	5426	13563	12335	2616
Handoverdate FE controls	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Logit estimation. Relative risk estimator, standard errors in parentheses, clustered on district level. Probability of exceeding the 30.5 days of allowed grace period in year 3. Prob FP - Probability of paying in accordance with clusters of frequent repayers. Prob IP - Probability of paying in accordance with clusters of irregular repayers. Prob IP (probability of paying in accordance with clusters of irregular repayers) omitted category. Restricted to borrowers who are in the loan cycle at least for two years. Columns (4), (5) and (6): restricted to borrowers who did not become delinquent in year 1 and 2. Borrowers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists. Borrowers for which less than one year of repayment data exists or who bought more than one system excluded. */**/** denote significance levels at 10/5/1 percent respectively.

Indeed, repayment patterns in the first year clearly predict repayment difficulties in later years. Borrowers who paid very frequently in the first year are more than twice as likely to face payment difficulties in the third year compared to steady payers; also irregular payers are more likely to face payment difficulties. When we exclude borrowers who became delinquent in the first two years, repayment patterns in the first year still have strong predictive power for delinquency in the last year. Yet, the interpretation changes. Borrowers who paid irregular in the first year and did not become delinquent in the first two years have a very low likelihood to become delinquent in the last year. When all socio-economic characteristics are included (column 6), the sample becomes small which makes interpretation difficult. Overall, once conditioning on repayment patterns, socio-economic cha-

racteristics play a less important role in predicting payment difficulties. Surprisingly, access to formal finance options is positively correlated with delinquency. Potentially, these are borrowers that became over-indebted in the course of the repayment cycle.

3.3.2 Types of delinquency

Delinquency can have a number of reasons which are important to distinguish - in particular as different forms of delinquency are likely associated with different borrower characteristics and can have different implications for eventual default. One can differentiate between three major causes for delinquency. Delinquency due to a sudden income shock or an unexpected increase in expenditures ('sudden delinquency'); delinquency because of a general difficulty to repay such as expenditures often exceeding income ('expected delinquency'); and delinquency because of insufficient track-keeping of the payments made and the grace-period used ('creeping delinquency') which is a risk inherent to very flexible repayment schemes. While we cannot observe the actual reason for delinquency, the repayment data, in particular the use of the grace-period before delinquency, can provide an indication for the underlying reason. For the analysis, we focus on the point of time when a borrower became delinquent for the first time. Borrowers who became delinquent after a long period of grace-period while having paid on a very regular basis before-hand are likely to have experienced an income shock. Delinquency of these borrowers is defined as 'sudden delinquency' (9% of all delinquencies). Furthermore, borrowers who have used grace-period in the past several times and for longer periods seem to have general difficulties in making the payments on time - we define this delinquency as 'expected delinquency' (36% of all delinquencies). Finally, borrowers who missed their payment frequently before becoming delinquent but then repaid quickly have likely not kept track of their payments and the number of grace days have slowly accumulated - this is classified as 'creeping delinquency' (55% of all delinquencies). The exact classification strategy is outlined in Appendix D.1..

In Table 10, we report results from a multinomial logit, comparing determinants for sudden, expected and creeping delinquency to no delinquency. Indeed, determinants for delinquency differ once distinguishing between the different types of delinquency. For example, while customer with a large number of businesses are more likely to face creeping delinquency they seem to be less likely to have general difficulties in repaying - i.e. to face expected delinquency. Access to consumption smoothing resources such as wealth has no predictive power for sudden or expected delinquency but is associated with a lower likelihood of creeping delinquency. Access to informal finance (i.e. through saving groups) is linked to a lower likelihood of sudden delinquency. Arguably, these saving groups provide means to cope with income shocks.

When we relate these three types of delinquency to repayment patterns, established above, in the first year, we see a clear link (see Figure 5). Borrowers that repaid very steadily in the first year have a higher likelihood to become suddenly delinquent than borrowers who repaid very frequently or irregularly. The association is particularly strong between frequent repayers and creeping delinquency. The repayment pattern in the first year of those borrowers that at some point became 'creepingly' delinquent is with over 80% probability associated with the frequent repayment cluster. Thus those borrower that make use of the contract's flexibility by paying very frequently small amounts are more

Table 10: Delinquency Types and Socioeconomic Characteristics

	(1)			(2)		
	Sudden	Expected	Creeping	Sudden	Expected	Creeping
female	1.043 (0.077)	1.089 (0.046)**	0.944 (0.027)**	0.941 (0.151)	1.155 (0.074)**	0.963 (0.036)
hh members	0.950 (0.018)***	1.003 (0.007)	0.939 (0.007)***	0.922 (0.036)**	1.007 (0.012)	0.936 (0.011)***
rural (10km)	1.215 (0.058)***	1.088 (0.032)***	1.067 (0.034)**	1.282 (0.149)**	1.083 (0.049)*	1.018 (0.030)
system loanamount (log)	1.076 (0.143)	0.964 (0.080)	0.973 (0.048)	0.992 (0.318)	0.930 (0.145)	0.984 (0.049)
farmer	1.046 (0.101)	1.112 (0.039)***	1.072 (0.046)	0.971 (0.169)	1.171 (0.063)***	1.114 (0.066)*
wage-employed	0.930 (0.082)	1.030 (0.041)	0.652 (0.032)***	1.104 (0.158)	1.002 (0.072)	0.715 (0.039)***
number of businesses				1.073 (0.060)	0.923 (0.029)**	1.062 (0.024)***
business purpose				2.127 (0.306)***	1.076 (0.053)	1.169 (0.075)**
wealth index				1.004 (0.090)	1.012 (0.025)	0.929 (0.018)***
use of formal finance				1.079 (0.154)	1.015 (0.055)	0.998 (0.054)
use of informal finance				0.723 (0.114)**	0.994 (0.065)	1.008 (0.049)
Observations	34972			18755		
Handoverdate FE controls	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Multinomial logit estimation. Relative risk estimator, standard errors in parentheses, clustered on district level. Probability of exceeding the 30.5 days of allowed grace period in year 3. Prob FP - Probability of paying in accordance with clusters of frequent repayers. Prob IP - Probability of paying in accordance with clusters of irregular repayers. Prob SP (probability of paying in accordance with clusters of steady repayers) omitted category. Restricted to borrowers who are in the loan cycle at least for two years. Columns (4), (5) and (6): restricted to borrowers who did not become delinquent in year 1 and 2. Borrowers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists. Borrowers for which less than one year of repayment data exists or who bought more than one system excluded. */**/** denote significance levels at 10/5/1 percent respectively.

likely to become delinquent due to slowly accumulating unpaid days.

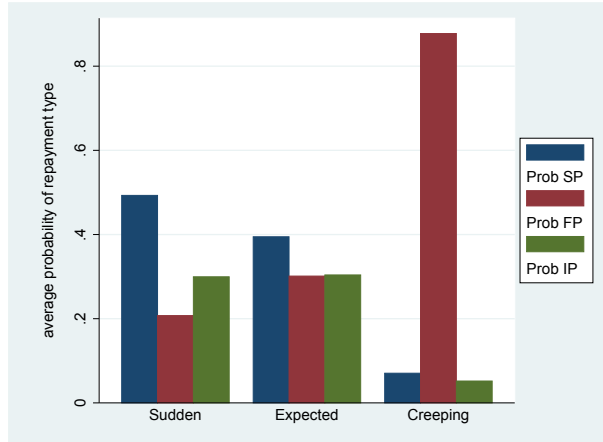


Figure 5: Average repayment type probability by type of delinquency

Yet, not all delinquency results in eventual default. Figure 6 depicts the probability of default by type of delinquency distinguishing between quasi-default (i.e. non-payment for three consecutive months) and actual default (i.e. the system is de-installed). Only 10% of those borrowers that became creepingly delinquent eventually default and less than 25% do not pay for over 3 months, i.e. the majority of these borrowers typically get back on track. This is in stark contrast to borrowers who became suddenly delinquent. Three out of four eventually default or quasi-default.

Thus while the flexibility that the payment scheme provides for can jeopardize on-time payments, it seems not to be the primary reason for eventual default. Nevertheless, creeping delinquency, which is possibly a result of borrowers not keeping track of their payments and thus a result of the flexibility the contract provides, can be costly; both for the company as borrowers need to be reminded and potentially visited as well as for the borrowers who have no access to electricity during the system's shut-off times.

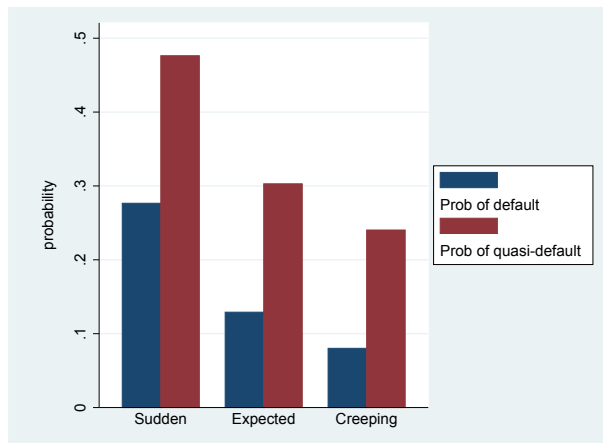


Figure 6: Default probability by type of delinquency

4 Conclusion

Being able to borrow to finance investment is crucial for individual welfare as well as for the economy as a whole. To this day, many poor people and others living in developing countries are underserved by the financial system. Microfinance has provided one alternative for poor people to borrow. Most microfinance operates using a very specific loan contract, where small loans are given out and have to be repaid very frequently, shortly after disbursement. In many microfinance institutions social collateral, in the form of group lending, is used instead of physical collateral. Partly because of this loan structure, microfinance has come under criticism.

In this paper, we examine an alternative way for the financially underserved to borrow. We study a loan contract that is given out to finance the purchase of a solar home panel. Different to most loans, this loan is very flexible and relatively large compared to borrowers' incomes. Furthermore, the solar home system itself acts as collateral. We study repayment and delinquency in the context of such a flexible loan contract using a very rich administrative data set for 40 000 borrowers.

Using an unsupervised pattern analysis we show that there are three repayment types, that are made up of several clusters each. A considerable proportion of borrowers repay very regularly as would be required in most standard loan contracts (steady repayers). The largest proportion of borrowers repay small amounts very frequently (frequent repayers). The last group pays large amounts very seldom (irregular repayers). We are further able to show that borrowers with varying cash flows are more likely to be frequent repayers, while wage employed borrowers are more likely to be steady repayers. Looking at this in more detail, we can show that farmers align their repayment to the local crop cycle and thus their cash flows. In terms of delinquency, we find frequent repayers to be more likely to face payment difficulties. However, we show that delinquency and eventual default are more nuanced. While borrowers that repay frequently are more likely to become delinquent because of short term liquidity problems and forgetfulness, they typically get back on track. On the other side, borrowers who suddenly became delinquent, probably due to an economic shock, are more likely to default completely.

Two main conclusions can be drawn from this study that are important when thinking about an adequate design of loan contracts to serve the poor. First, there is a clear demand for flexible contracts. Borrowers that face varying income flows and borrowers with limited resources to smooth consumption repay in a flexible manner when they are allowed to and they mostly do this in a way that is adapted to their cash-flow and their household's more general situation. Secondly, while use of flexibility is linked to short term delinquency, sudden economic shocks are the most important reason for complete default. This paper hence provides evidence that borrowers may benefit from being offered flexible repayment contracts and that using this flexibility does not automatically lead to deterioration in repayment discipline.

This is a purely observational study. To further understand the effects of innovative contracts on repayment as well as on individual welfare more research is needed. Most experimental studies so far only examine one element of flexibility or collateral. Little is known about contracts that combine different novel elements. This could be an avenue for future research.

References

- Aghabozorgi, Saeed, Ali Seyed Shirخورshidi, and Teh Ying Wah (2015). “Time-series clustering – A decade review”. In: *Information Systems* 53, pp. 16–38.
- Alkire, Sabina and James Foster (2011). “Counting and multidimensional poverty measurement”. In: *Journal of public economics* 95.7-8, pp. 476–487.
- Angelucci, Manuela, Dean Karlan, and Jonathan Zinman (2015). “Microcredit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco”. In: *American Economic Journal: Applied Economics* 7.1, pp. 151–82.
- Augsburg, Britta et al. (2015). “The Impacts of Microcredit: Evidence from Bosnia and Herzegovina”. In: *American Economic Journal: Applied Economics* 7.1, pp. 183–203.
- Banerjee, Abhijit V and Esther Duflo (2014). “Do firms want to borrow more? Testing credit constraints using a directed lending program”. In: *Review of Economic Studies* 81.2, pp. 572–607.
- Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman (2015). “Six Randomized Evaluations of Microcredit: Introduction and Further Steps”. In: *American Economic Journal: Applied Economics* 7.1, pp. 1–21.
- Banerjee, Abhijit et al. (2019). *Can Microfinance Unlock a Poverty Trap for Some Entrepreneurs?* Working Paper 26346. National Bureau of Economic Research.
- Barboni, Giorgia and Parul Agarwal (2018). “Knowing what’s good for you: Can a repayment flexibility option in microfinance contracts improve repayment rates and business outcomes?”
- Battaglia, Marianna, Selim Gulesci, Andreas Madestam, et al. (2018). *Repayment flexibility and risk taking: Experimental evidence from credit contracts*. Centre for Economic Policy Research Discussion Paper Series 13329.
- Beck, Thorsten and Asli Demirgüç-Kunt (2008). “Access to finance: An unfinished agenda”. In: *The world bank economic review* 22.3, pp. 383–396.
- Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine (2007). “Finance, inequality and the poor”. In: *Journal of economic growth* 12.1, pp. 27–49.
- Berge, Lars Ivar Oppedal, Kjetil Bjorvatn, and Bertil Tungodden (2015). “Human and financial capital for microenterprise development: Evidence from a field and lab experiment”. In: *Management Science* 61.4, pp. 707–722.
- Crépon, Bruno et al. (2015). “Estimating the Impact of Microcredit on Those Who Take It Up: Evidence from a Randomized Experiment in Morocco”. In: *American Economic Journal: Applied Economics* 7.1, pp. 123–50.
- Cull, Robert, Asli Demirgüç-Kunt, and Jonathan Morduch (2009). “Microfinance meets the market”. In: *Journal of Economic perspectives* 23.1, pp. 167–92.
- Czura, Kristina (2015). *Do flexible repayment schedules improve the impact of microcredit? Evidence from a randomized evaluation in rural India*. Tech. rep. Munich Discussion Paper.
- De Mel, Suresh, David McKenzie, and Christopher Woodruff (2008). “Returns to capital in microenterprises: evidence from a field experiment”. In: *The Quarterly Journal of Economics* 123.4, pp. 1329–1372.

- De Quidt, Jonathan, Thiemo Fetzer, and Maitreesh Ghatak (2018). “Commercialization and the decline of joint liability microcredit”. In: *Journal of Development Economics* 134, pp. 209–225.
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). “Maximum Likelihood from Incomplete Data via the EM Algorithm”. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 39.1, pp. 1–38.
- Dunning, Caroline M., Emily C. L. Black, and Richard P. Allan (2016). “The onset and cessation of seasonal rainfall over Africa”. In: *Journal of Geophysical Research: Atmospheres* 121.19, pp. 11,405–11,424.
- Ester, Martin et al. (1996). “A Density-based Algorithm for Discovering Clusters a Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise”. In: *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. KDD’96. Portland, Oregon: AAAI Press, pp. 226–231.
- Field, Erica and Rohini Pande (2008). “Frequency and Default in Microfinance: Evidence From India”. In: *Journal of the European Economic Association* 6.2-3, pp. 501–509.
- Field, Erica et al. (2012). “Repayment Flexibility Can Reduce Financial Stress: A Randomized Control Trial with Microfinance Clients in India”. In: *PLoS ONE* 7.9. Ed. by Tiziana Leone, e45679.
- Field, Erica et al. (2013). “Does the Classic Microfinance Model Discourage Entrepreneurship Among the Poor? Experimental Evidence from India”. In: *American Economic Review* 103.6, pp. 2196–2226.
- Giné, Xavier and Dean S Karlan (2014). “Group versus individual liability: Short and long term evidence from Philippine microcredit lending groups”. In: *Journal of development Economics* 107, pp. 65–83.
- Gine, Xavier and Ghazala Mansuri. “Money or Management? A Field Experiment on Constraints to Entrepreneurship in Rural Pakistan”. In: *Economic Development and Cultural Change* forthcoming.
- Giné, Xavier and Robert M. Townsend (2004). “Evaluation of financial liberalization: a general equilibrium model with constrained occupation choice”. In: *Journal of Development Economics* 74.2, pp. 269–307.
- Jack, William et al. (2016). *Borrowing Requirements, Credit Access, and Adverse Selection: Evidence from Kenya*. Working Paper 22686. National Bureau of Economic Research.
- Labie, Marc, Carolina Laureti, and Ariane Szafarz (2017). “Discipline and flexibility: a behavioural perspective on microfinance product design”. In: *Oxford Development Studies* 45.3, pp. 321–337.
- Laureti, Carolina, Alain De Janvry, and Elisabeth Sadoulet (2017). “Flexible Microfinance Products for Financial Management by the Poor: Evidence from SafeSave”. In: *Working Papers CEB* 17.
- Levine, Ross (2005). “Finance and growth: theory and evidence”. In: *Handbook of economic growth* 1, pp. 865–934.
- Liao, T. Warren (2005). “Clustering of time series data – a survey”. In: *Pattern Recognition* 38.11, pp. 1857–1874.
- MacQueen, J. (1967). “Some methods for classification and analysis of multivariate observations”. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. Berkeley, Calif.: University of California Press, pp. 281–297.

- Pearlman, Sarah (2010). “Flexibility matters: do more rigid loan contracts reduce demand for microfinance?”
- Scrucca, Luca et al. (2016). “mclust 5: clustering, classification and density estimation using Gaussian finite mixture models”. In: *The R Journal* 8.1, pp. 205–233.
- Shoji, Masahiro (2010). “Does contingent repayment in microfinance help the poor during natural disasters?” In: *Journal of Development Studies* 46.2, pp. 191–210.
- Ward, Joe H. (1963). “Hierarchical Grouping to Optimize an Objective Function”. In: *Journal of the American Statistical Association* 58.301, pp. 236–244.
- World Bank (2017). *Sustainable Energy for All — Progress toward Sustainable Energy*. Tech. rep. International Energy Agency (IEA) and the World Bank.

A Borrower Characteristics

A.1 Borrower Characteristics

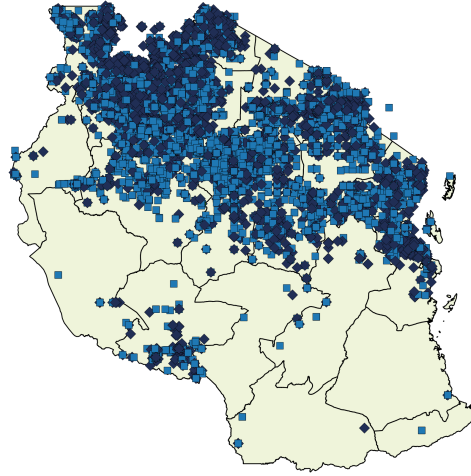


Figure 7: Location of Borrowers across Tanzania

Table 11: Sample Borrower Characteristics

Variable	Mean	SD	Min	Max	Median	Obs.
female	0.19	0.39	0.00	1.00	0.00	37571
hh members	4.48	1.94	1.00	30.00	4.00	37438
age	39.65	10.94	15.00	89.00	38.00	18511
farmer	0.46	0.50	0.00	1.00	0.00	36202
wage-employed	0.24	0.42	0.00	1.00	0.00	36202
self-employed	0.30	0.46	0.00	1.00	0.00	36202
number of businesses	1.27	0.95	0.00	5.00	1.00	19705
rural (10km)	0.58	0.49	0.00	1.00	1.00	37322
use of formal finance	0.18	0.39	0.00	1.00	0.00	19705
use of informal finance	0.15	0.36	0.00	1.00	0.00	19705
system loanamount (in 100 US Dollar)	9.54	3.13	2.45	25.11	8.44	38407
business purpose	0.07	0.25	0.00	1.00	0.00	19525
trading or renting land	0.20	0.40	0.00	1.00	0.00	19705
monthly disposable income - above median	0.57	0.50	0.00	1.00	1.00	19524
owns land	0.59	0.49	0.00	1.00	1.00	19705
employs domestic workers	0.27	0.45	0.00	1.00	0.00	19705
owns animals than can be sold in emergency	0.18	0.39	0.00	1.00	0.00	19705
roof made out of tiles	0.04	0.20	0.00	1.00	0.00	37607
wealth index	1.82	0.97	0.00	4.00	2.00	19469

Notes: Borrowers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists. Borrowers for which less than one year of repayment data exists or who bought more than one system excluded.

B Cluster Analysis

B.1 Selection of Number of Clusters

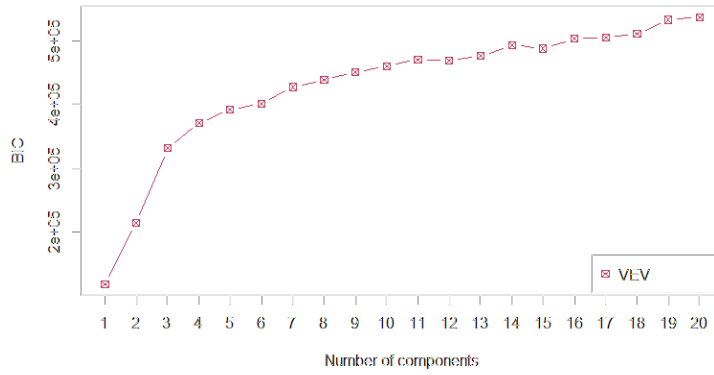


Figure 8: Number of Clusters and associated Bayesian Information Criteria

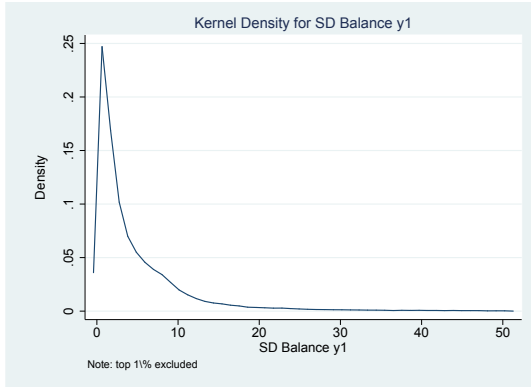
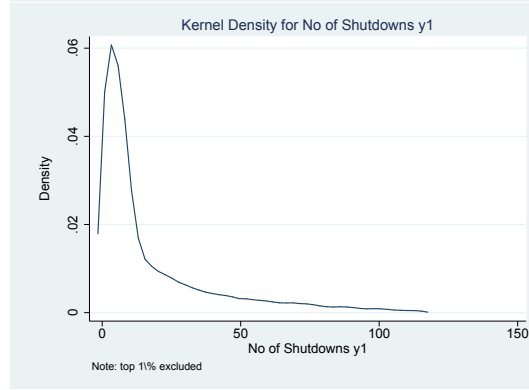
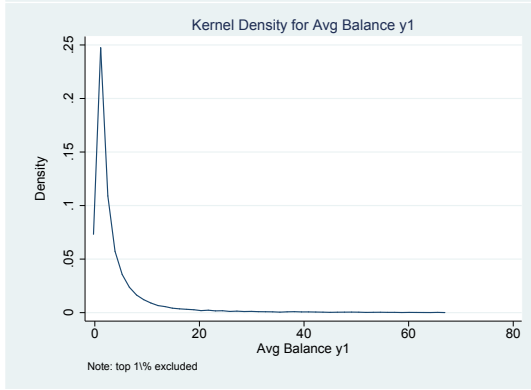
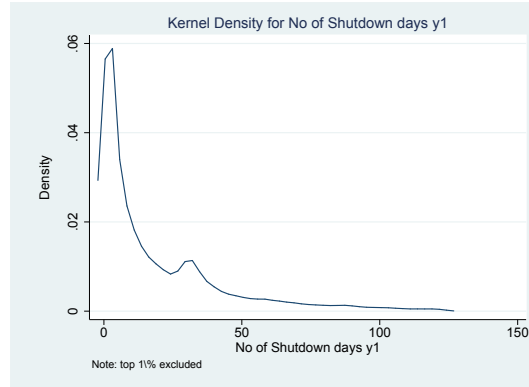
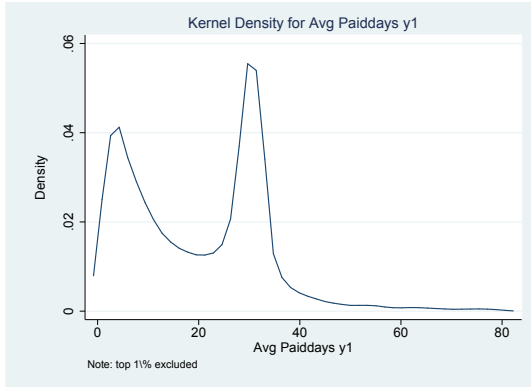
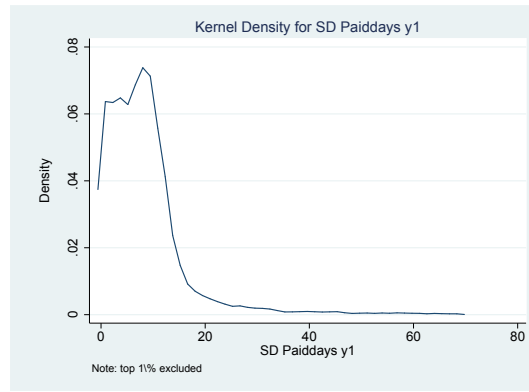
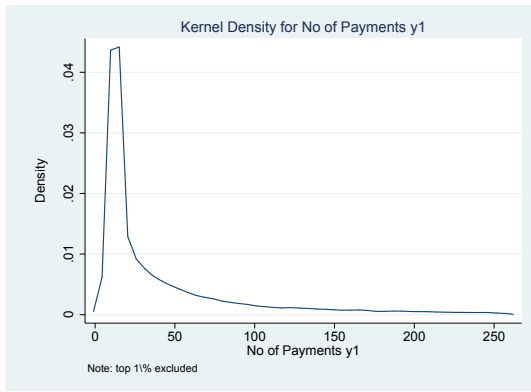
B.2 Precision of Clusters

Table 12: Cluster Precision

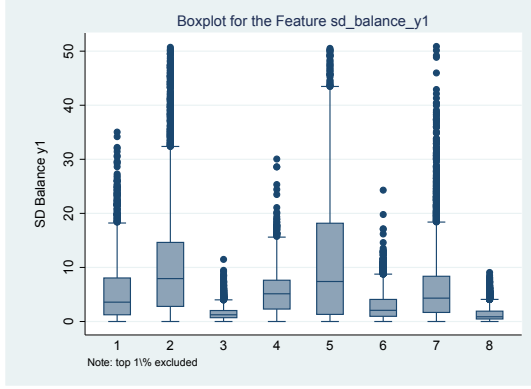
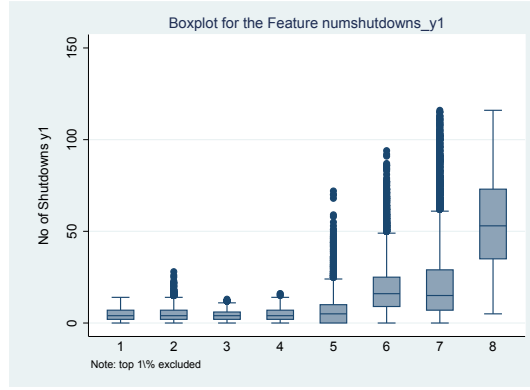
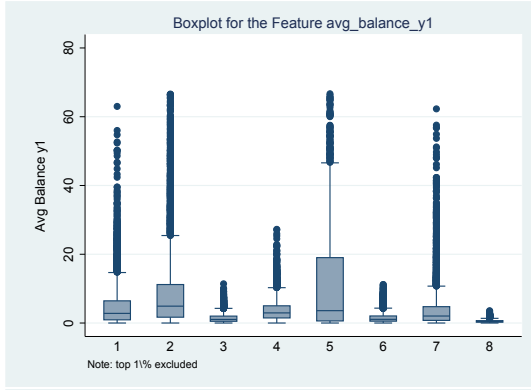
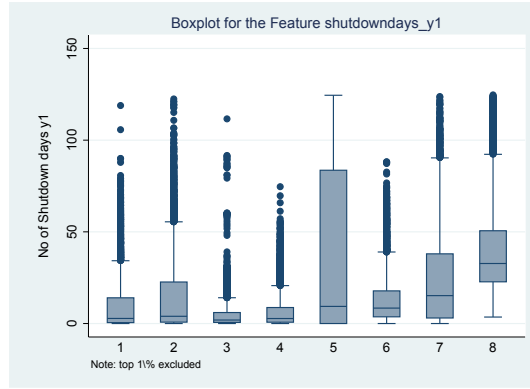
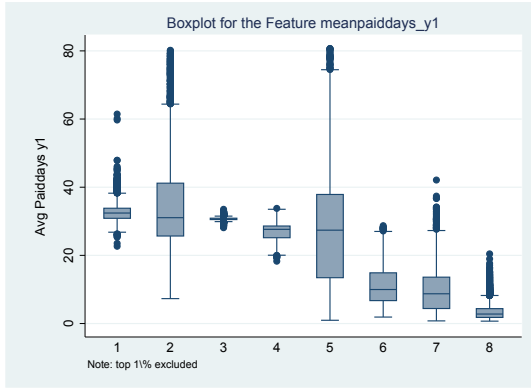
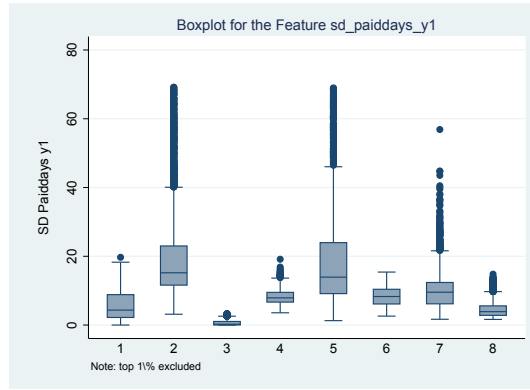
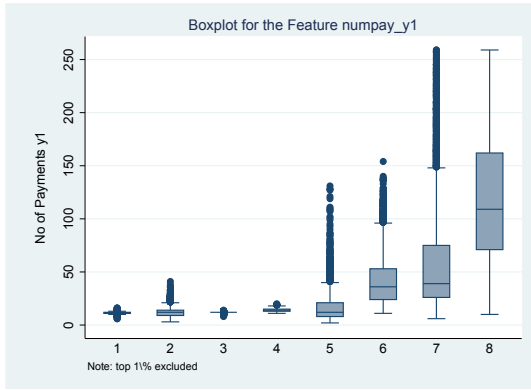
	mean	p5	p10	p25	p50	p75	p90	p95	min	max	count
Prob Cluster 1	0.94	0.64	0.80	0.95	0.99	1.00	1.00	1.00	0.36	1.00	4171
Prob Cluster 2	0.95	0.65	0.77	0.96	1.00	1.00	1.00	1.00	0.35	1.00	3663
Prob Cluster 3	0.97	0.80	0.92	0.99	1.00	1.00	1.00	1.00	0.50	1.00	5211
Prob Cluster 4	0.92	0.60	0.71	0.91	0.99	1.00	1.00	1.00	0.33	1.00	7418
Prob Cluster 5	0.91	0.57	0.67	0.88	0.99	1.00	1.00	1.00	0.34	1.00	5233
Prob Cluster 6	0.95	0.68	0.80	0.97	1.00	1.00	1.00	1.00	0.35	1.00	6019
Prob Cluster 7	0.92	0.59	0.70	0.91	0.99	1.00	1.00	1.00	0.35	1.00	5119
Prob Cluster 8	0.91	0.56	0.65	0.86	0.99	1.00	1.00	1.00	0.35	1.00	1573
Prob SP	0.97	0.83	0.94	0.99	1.00	1.00	1.00	1.00	0.35	1.00	13045
Prob IP	0.95	0.66	0.79	0.97	1.00	1.00	1.00	1.00	0.38	1.00	6692
Prob FP	0.98	0.84	0.96	1.00	1.00	1.00	1.00	1.00	0.38	1.00	18670

Notes: Probability of belonging to a specific cluster for customers assigned to this cluster. Prob SP - Probability of paying in accordance with clusters of steady repayers. Prob FP - Probability of paying in accordance with clusters of frequent repayers. Prob IP - Probability of paying in accordance with clusters of irregular repayers. Borrowers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists. Borrowers for which less than one year of repayment data exists or who bought more than one system excluded.

B.3 Density of Features within Clusters



B.4 Variations of Features within Clusters



B.5 Correlation between Features

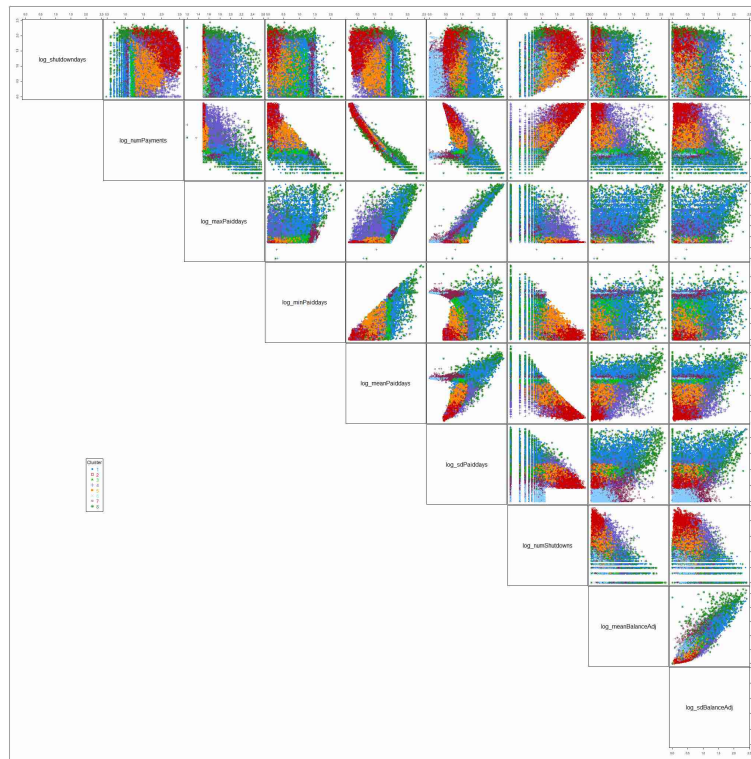


Figure 9: Correlation between Repayment Features

C Repayment Types

C.1 Determinants of Repayment Types by Cluster

Table 13: Probability of belonging to Steady Payment Clusters and Socioeconomic Characteristics

	Cluster 1		Cluster 2		Cluster 3	
	(1)	(2)	(3)	(4)	(5)	(6)
female	0.941 (0.036)	1.010 (0.068)	1.224 (0.081)***	1.259 (0.110)***	1.198 (0.061)***	1.158 (0.084)**
hh members	1.032 (0.009)***	1.029 (0.017)*	1.050 (0.010)***	1.063 (0.017)***	1.023 (0.009)***	1.036 (0.015)**
rural (10km)	0.944 (0.041)	1.017 (0.068)	0.905 (0.044)**	0.885 (0.063)*	0.989 (0.049)	1.022 (0.072)
system loanamount (log)	1.181 (0.070)***	1.298 (0.113)***	0.982 (0.092)	1.244 (0.164)*	1.167 (0.065)***	1.058 (0.103)
farmer	0.845 (0.037)***	0.946 (0.060)	0.874 (0.050)**	0.810 (0.070)**	0.788 (0.034)***	0.773 (0.046)***
wage-employed	1.398 (0.060)***	1.633 (0.139)***	1.239 (0.055)***	1.079 (0.101)	1.413 (0.058)***	1.524 (0.094)***
number of businesses		1.016 (0.043)		0.906 (0.040)**		0.946 (0.031)*
business purpose		0.906 (0.116)		0.663 (0.112)**		0.768 (0.070)***
wealth index		1.025 (0.041)		1.061 (0.046)		1.025 (0.033)
use of formal finance		1.352 (0.096)***		1.085 (0.094)		1.180 (0.094)**
use of informal finance		0.900 (0.083)		0.911 (0.088)		0.879 (0.066)*
Constant	0.118 (0.025)***	0.031 (0.007)***	0.136 (0.040)***	0.065 (0.018)***	0.173 (0.040)***	0.066 (0.015)***
Observations	35146	18789	35146	18789	35146	18789
Handoverdate FE controls	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: GLM with logit transformation of response variable & binomial distribution. Relative risk estimator, standard errors in parentheses, clustered on district level. Customers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists. Customers for which less than one year of repayment data exists or who bought more than one system excluded. */**/** denote significance levels at 10/5/1 percent respectively.

Table 14: Probability of belonging to Frequent Payment Clusters and Socioeconomic Characteristics

	Cluster 4		Cluster 5		Cluster 6	
	(1)	(2)	(3)	(4)	(5)	(6)
female	0.934 (0.033)*	0.961 (0.038)	0.832 (0.035)***	0.900 (0.045)**	0.951 (0.043)	0.959 (0.047)
hh members	0.984 (0.008)**	1.002 (0.012)	0.960 (0.009)***	0.969 (0.012)**	0.918 (0.009)***	0.923 (0.011)***
rural (10km)	1.027 (0.038)	1.024 (0.039)	1.057 (0.047)	1.013 (0.055)	0.987 (0.050)	0.956 (0.048)
system loanamount (log)	0.838 (0.039)***	0.839 (0.042)***	0.856 (0.047)***	0.721 (0.048)***	1.072 (0.066)	1.165 (0.080)**
farmer	1.030 (0.031)	1.040 (0.038)	1.184 (0.041)***	1.191 (0.049)***	0.961 (0.036)	0.972 (0.036)
wage-employed	1.004 (0.034)	1.052 (0.046)	0.745 (0.028)***	0.782 (0.048)***	0.578 (0.026)***	0.682 (0.038)***
number of businesses		1.033 (0.022)		0.981 (0.028)		1.095 (0.026)***
business purpose		1.087 (0.079)		0.984 (0.079)		1.322 (0.095)***
wealth index		0.979 (0.023)		1.030 (0.022)		0.905 (0.018)***
use of formal finance		0.988 (0.049)		0.869 (0.051)**		0.907 (0.037)**
use of informal finance		1.109 (0.055)**		1.097 (0.065)		1.073 (0.047)
Constant	0.182 (0.040)***	0.230 (0.033)***	0.144 (0.027)***	0.816 (0.153)	0.023 (0.012)***	0.102 (0.019)***
Observations	35146	18789	35146	18789	35146	18789
Handoverdate FE controls	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: GLM with logit transformation of response variable & binomial distribution. Relative risk estimator, standard errors in parentheses, clustered on district level. Customers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists. Customers for which less than one year of repayment data exists or who bought more than one system excluded. */**/** denote significance levels at 10/5/1 percent respectively.

Table 15: Probability of belonging to Irregular Payment Clusters and Socioeconomic Characteristics

	Cluster 7		Cluster 8	
	(1)	(2)	(3)	(4)
female	1.023 (0.041)	0.999 (0.069)	0.970 (0.071)	0.989 (0.104)
hh members	1.050 (0.009)***	1.082 (0.016)***	1.012 (0.013)	1.040 (0.019)**
rural (10km)	1.086 (0.070)	1.072 (0.090)	0.952 (0.062)	0.987 (0.090)
system loanamount (log)	0.966 (0.067)	1.123 (0.110)	1.116 (0.103)	0.824 (0.108)
farmer	1.169 (0.042)***	1.036 (0.073)	1.322 (0.072)***	1.249 (0.113)**
wage-employed	0.938 (0.043)	0.875 (0.082)	1.063 (0.073)	1.097 (0.106)
number of businesses		0.931 (0.032)**		0.870 (0.039)***
business purpose		0.743 (0.093)**		1.008 (0.154)
wealth index		1.098 (0.035)***		1.109 (0.047)**
use of formal finance		0.930 (0.072)		0.854 (0.082)*
use of informal finance		0.891 (0.060)*		0.887 (0.089)
Constant	0.239 (0.050)***	0.164 (0.038)***	0.024 (0.010)***	0.120 (0.042)***
Observations	35146	18789	35146	18789
Handoverdate FE controls	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes

Notes: GLM with logit transformation of response variable & binomial distribution. Relative risk estimator, standard errors in parentheses, clustered on district level. Customers who bought a system between 07-2013 and 11-2018, for which information from loan eligibility interview exists. Customers for which less than one year of repayment data exists or who bought more than one system excluded. */**/** denote significance levels at 10/5/1 percent respectively.

D Delinquency

D.1 Specifying Delinquency Types

Table 16 provides summary statistics for the number of consecutive days of non-payment when the borrower became delinquent as well as for the number of non-payment incidences over usage days and the average duration of non-payment before delinquency.²⁸ For the majority of borrowers the non-payment period before delinquency is rather short (just about 1 day), i.e. these borrowers have made considerable use of the grace period before. However, for more than 10% of all borrowers the shut-down period before delinquency is at least 15 days long; these borrowers had used none or very little grace period in this particular year before they became delinquent.

Table 16: Delinquency Analysis: non-payment indicators before grace-period was exceeded for the first time

	mean	sd	p25	p50	p75	p90
Duration of non-payment period	5.00	8.31	0.66	1.01	4.77	17.53
Non-payment incidence (over usage days)	30.93	45.75	6.66	15.22	41.77	68.85
Avg duration of non-payment	1.96	2.47	0.52	1.06	2.51	4.69

All borrowers who became delinquent at least once.

Borrowers who became delinquent in first 6 months excluded.

The majority of borrowers that become delinquent had a large number of late payments in the past. On average, a borrower who became delinquent used to not pay on time at least once a month, the majority did not pay on time every 15 days. However, borrowers typically paid again very fast. The average duration of non-payment is just about 2 days. A clear signal that the sanctioning mechanism of shutting-down the system in case of non-payment worked well.

We use this information to classify delinquent borrowers in three main groups: First, borrowers who tended to pay on time in the past before they became delinquent – i.e. they made little use of grace-period in the past. More specifically, delinquency happened after a long period of non-payment of at least 20 days. 9% of the delinquent borrowers fall into this category of ‘sudden delinquency.’ Second, borrowers whose non-payment period before delinquency was less than 20 days (i.e. they already had made considerable use of the grace period in the past) and who had made use of the grace period quite often, at least every 20 days. These borrowers repay on average quite fast (within less than one day). 56% of all delinquencies fall into this category, which we coin ‘creeping delinquency.’ Finally, borrowers who also had made use of grace period in the past (i.e. their non-payment period before delinquency was less than 20 days), but less often; more specifically, less than every 20 days. Yet, when they did not pay, they did not repay quickly: on average, their system was shut-down due to non-payment for 4 days. We call this type of delinquency ‘expected delinquency’, 35% of all delinquencies.

²⁸For all of the following analysis, we exclude borrowers who became delinquent in the first 6 months of their loan cycle.