Accepted: 22 September 2025

DOI: 10.1111/ecca.70013

ORIGINAL ARTICLE

Economica LSE

# Repayment under flexible loan contracts: evidence based on high-frequency data

**Antonia Grohmann**[1] | **Friederike Lenel**[2] | **Steffen Herbold**[3]

[1]Aarhus University

[2]University of Göttingen and Potsdam Institute for Climate Impact Research

[3]University of Passau

**Correspondence**
Friederike Lenel, Potsdam Institute for Climate Impact Research (PIK), P.O. Box 60 12 03 14412 Potsdam, Germany
Email: Friederike.Lenel@pik-potsdam.de; friederike.lenel@uni-goettingen.de

**Abstract**

We study repayment in the context of an alternative financial product that enables the purchase of a large asset—a solar panel home system—while offering complete repayment flexibility. Using a large administrative dataset on daily repayment of 38,400 borrowers in Tanzania over 5.5 years, we perform unsupervised pattern analysis to classify repayment behaviour. We show that borrowers with fluctuating incomes make more use of the loan's flexibility, and adjust their repayment to cash flow. In further analysis, we show that in particular, farmers use the flexibility provided by the loan contract to adjust to seasonal fluctuations and to deal with unexpected shocks. Our results indicate that low-income households can finance large assets through innovative financial approaches that allow aligning payments to financial circumstances.

**KEYWORDS**

large assets, flexible repayment schemes, loan contracts, repayment patterns, high-frequency data

**JEL CLASSIFICATION**

D14; G20; C55

## 1 | INTRODUCTION

The possession of large physical assets can have important impacts on the lives of the poor (Banerjee *et al.* 2015a) and can serve as a means to escape poverty traps (Balboni *et al.* 2021). This has been shown not only for productive assets such as livestock or farm equipment (Bandiera *et al.* 2017; Tadesse and Zewdie 2019), but also for a range of non-productive assets such as televisions, smartphones, home appliances or solar panel home systems that can improve quality of life (Aker *et al.* 2012; Stojanovski *et al.* 2017; Kafle *et al.* 2018). Yet while there is considerable evidence on the positive welfare effects of asset ownership, less is known about how the poor can finance such assets (Kumaraswamy *et al.* 2020).

For many decades, microfinance institutions (MFIs) have been one of the few financial actors providing credit to the poor. In a setting where asymmetric information, lack of collateral, and an often a weak legal system that makes contract enforcement difficult, MFIs typically offer small loans under very rigid repayment schedules (Armendáriz de Aghion and Morduch 2005). More recently, these features have come under criticism. Loans are deemed to be too small to have a meaningful impact, and the rigid repayment schemes are not compatible with the highly fluctuating income of many borrowers. It is argued that as a consequence, the microfinance loan structure can lead to overindebtedness and encourage suboptimal investment decisions (Pearlman 2010; Shoji 2010; Field *et al.* 2012, 2013; Banerjee *et al.* 2015b). In response, more recently a number of alternative financial products have been developed that help the poor to access large assets and allow for more repayment flexibility. Yet evidence on the use of such products is still scarce.

In this paper, we study such an innovative financial product that allows poor borrowers to purchase a large asset while offering a very flexible repayment scheme. We use high-frequency observational data to investigate how people repay when they have payment flexibility, and whether the use of flexibility is related to the borrowers' financial circumstances. We study how borrowers in rural areas adjust their payments in response to seasonal fluctuations, and analyse the impact of local vegetation shocks on repayment behaviour.

The financial product that we study is a loan for 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 electricity sources (World Bank 2017). The loan size is large: between US\$600 and US\$1300, depending on the system that the borrower acquires. 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 (PAYG) 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 (the 'grace period'). When the grace period is exceeded and the borrower is not able to get back on track, the system is de-installed. The loan contract thereby offers different flexibility features that allow borrowers to align their payments to their cash flows; at the same time, the solar panel home system and the electricity provided act as a form of collateral. The financial product has a number of innovative elements: the loan amount is large, the repayment schedule is very flexible, and the purchased asset serves as collateral—thus the arrangement resembles a hire purchase contract.

Our analysis is based on a large administrative dataset that provides information on daily repayments for around 38,400 borrowers in Tanzania over 5.5 years. 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 first conduct an unsupervised cluster analysis of the repayment behaviour in the first year to identify the various ways in which 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, which can be further bundled into three broad repayment groups: steady, frequent and irregular repayers. We then study whether borrowers with higher income fluctuations and fewer resources to smooth consumption are more likely to use the flexibility that the contract provides. 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. On the other hand, borrowers who tend to have varying cash flows and fewer resources are more likely to repay small amounts very frequently. Finally, we investigate in more detail to which extent borrowers adjust their payments according to the variations in their cash flows. Analysing farmers' repayments over the local crop cycle, we find that farmers make use of the loan's flexibility according to their financial circumstances. In particular, they pay more frequently and lower amounts during the growing season, and larger amounts when crop yields materialize. We further analyse the adjustment of repayment by examining how borrowers in rural areas respond to local vegetation shocks experienced during the growing season. We find that borrowers, and especially farmers, adjust their repayment when experiencing harvest loss: they pay smaller amounts, more frequently. Off-days increase, yet to a lesser extent than the reduction in the total payment amount, indicating that the flexibility of the repayment scheme allows borrowers to buffer some of the negative effects of income loss.

Our study is one of the first empirical analyses of repayment behaviour in the context of a flexible loan product. We show that when given the choice, borrowers make use of the flexibility provided in different ways, depending on their general financial circumstances, and adjust their repayment to temporal cash flow variations. Although we can make little judgement about welfare implications or the optimality of borrowers' choices, this study shows that there clearly is demand for many different forms of flexibility in repayment of large assets, and that rigid loan contracts, or even loan contracts that offer only one form of flexibility, do not serve all borrowers equally well.

This paper contributes to two strands of the literature. First, by studying a product with a highly flexible repayment structure, we add to the discussion of flexibility in lending. Low-income people are often self-employed, their income fluctuates considerably throughout the year, and they are more prone to shocks (Banerjee 2001). This makes lending to them very risky for traditional banks. MFIs usually offer small loans under rigid repayment schedules: loans need to be repaid through fixed, typically weekly instalments, with the first instalment due soon after loan disbursement (Armendáriz de Aghion and Morduch 2005). A number of studies analyse the impact of MFIs offering slightly more flexibility in loan repayment. These studies are mostly based on randomized control trials (RCTs), varying one element of flexibility at a time. Early research suggests that decreasing repayment frequency from weekly to monthly does not necessarily lead to higher delinquency rates (Field and Pande 2008) and can help to alleviate borrowers' financial stress (Field *et al.* 2012). The evidence on the effects of offering payment deferrals is mixed. Brune *et al.* (2022) find that allowing borrowers to defer their monthly repayments altered their investment behaviour but also increased default rates. In the context of group lending, Czura *et al.* (2020) show that the option to defer payments can reduce social pressure and increase default. On the other hand, Battaglia *et al.* (2024) find that giving borrowers the option to postpone repayment improved business outcomes and reduced default. Supporting this, Barboni and Agarwal (2023) show that when borrowers can choose between a cheaper rigid contract and a costlier but flexible contract that allows them to reshuffle their repayment throughout the year in case of need, borrowers tend to self-select to the contract that suits them best, resulting in improved business outcomes and no increase in default. Along the same lines, Fiorin *et al.* (2023) show that offering payment deferrals to delinquent borrowers can have a positive effect on repayment.

Moreover, our study speaks to the emerging literature on innovative financial products in low-income settings that seek to better align loan terms with borrowers' individual circumstances. Berge *et al.* (2015) find that providing larger long-term loans in combination with business training to microenterprises can improve business outcomes and living conditions. Similarly, Bandiera *et al.* (2022) show that access to larger loans supports rural borrowers in diversifying into non-agricultural activities. Bari *et al.* (2024) find sustained improvements in business outcomes and household consumption when existing microfinance clients finance productive assets via hire purchase contracts, where asset ownership gradually increases with repayment.[1] Jack *et al.* (2023) demonstrate that collateralizing loans with the purchased asset, namely rainwater harvesting tanks, instead of having the loan secured by guarantors greatly increases loan take up, which in turn improves access to water and individual welfare. Gertler *et al.* (2024) analyse a school fee loan that is collateralized through a solar panel home system, which is switched off if the school fee loan is not repaid. They find that such 'digital collateral' improves repayment compared to an unsecured loan.[2]

In contrast to previous studies, this paper focuses on how borrowers *use* the flexibility that their loan product provides. While there has been a rise in alternative financial products that allow low-income households to acquire larger assets and adjust their repayment to their financial circumstances, to the best of our knowledge, this is the first empirical study that examines the use of such a loan product. The only other study that we are aware of that studies explicitly the usage of a flexible financial product is by Laureti *et al.* (2017), who study a microfinance product, namely a savings-and-loan account that allows clients to flexibly deposit and withdraw money. Using high-frequency observational data, our study can complement the existing experimental literature in a number of ways. First, being based primarily on administrative data, our analyses are not burdened with the reporting biases faced by many other studies on loans and repayment. Second, we are able to observe borrowers' repayment behaviour in a real-life context and over a longer period than is possible in most RCTs. Since we have a time series that contains daily data from borrowers, we can examine in detail how borrowers respond to season events and shocks. Third, while previous studies examine the specific effect of one element of flexibility, the loan schedule that we analyse allows borrowers to align repayment completely to their individual cash flows, and the rich administrative data allow us to study all nuances of the use of such flexibility. We are therefore able to answer questions that could not be answered using an RCT.

Although the rich observational data on which we base our study have a number of clear advantages, the data also come with different limitations. First, we study existing borrowers that have selected into a flexible loan contract. Thus we cannot say anything about the impact of such a flexible product, for example as compared to a traditional microfinance product with rigid repayment terms. Second, since they are self-selected, the group of borrowers is not representative of a specific population in Tanzania. In fact, as we show, the borrowers in our sample are likely slightly richer compared to the average rural population. Nonetheless, the sample probably represents the increasing share of households that are willing and able to invest well into solar powered electricity (World Bank 2022). Finally, besides the observational data on loan repayment, we have only a limited number of variables that can be used to describe the borrowers' profiles, hence our analyses, in particular those linking repayment patterns to borrowers' socioeconomic background, might suffer from omitted variable biases. We address these concerns throughout the paper.

The remainder of the paper is structured as follows. In Section 2, we describe the context of the study and the data that we use. In Section 3, we identify typical repayment patterns, classify borrowers accordingly, and link these patterns to borrowers' socioeconomic characteristics, while in Section 4, we study the implications of cash flow variations more specifically, linking repayment to the crop cycle and vegetation shocks. We discuss alternative explanations for our results in Section 5. Section 6 concludes.

## 2 | CONTEXT AND DATA

### 2.1 | Setting

We analyse data from a company that sells solar panel home systems on credit. The company was founded in 2010 as a pro-social business, and operates now in several countries in East Africa. At the time of our study, it employed over 1000 people. The company started as a pure solar energy company, while the customer finance operations were added later. We focus on Tanzania, where the company started operations in 2011 and had about 100,000 customers by the end of 2018, of which 95% bought the system on credit (see Figure A1 in Appendix A). The solar panel home system comes with a number of additional appliances (a television, lamps, a radio and a mobile phone charger—depending on the system type) and is typically sold through the company's own outlets that are located in towns throughout the country. The company offers a number of additional appliances (e.g. a stereo system, fan, phone charging station or hair cutter) that can be purchased with cash or on credit. The solar panel 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. Compared to some of its competitors, the company offers products that can be classed as more high end. The target customer group comprises low to middle income households living in rural areas, where the rate of electrification 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 US$600 and US$1300—a substantial financial investment compared to the average income in Tanzania.[4] The types differ in the panels' power (between 80 W and 200 W) and the appliances that a customer receives with the system.

Prospective borrowers go through a detailed phone assessment interview that evaluates whether they are eligible for the loan. If borrowers are approved, then they make a downpayment of 5% of the total system costs, receive the system, and then have three years to repay the remaining amount. There is an incentive to repay faster: borrowers receive a 10% discount if they repay within two years, and a 20% discount if they repay within one year. Furthermore, if borrowers pay for the system upfront, then 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 MFIs in Tanzania, where annual interest rates typically vary between 25% and 60%.[5]

Borrowers repay the loan using mobile money, typically through MPesa. They either have their own MPesa account, or transfer through an MPesa agent.[6] The cost of transferring money to the company is covered by the company through an agreement with the mobile money operators. There are thus no transfer costs for the borrower. Each loan payment also charges the system—similar to a PAYG 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 are allowed to deviate from this payment scheme; they are in principle free to decide the amount and timing of their payments. For example, if the total loan amount was US$720 to be repaid within three years, then a payment of US$20 would charge the system for one month, but a daily payment of 66 cents would be similarly sufficient. Despite the company suggesting a monthly repayment amount, all repayment decisions have to be taken actively, and there is no default repayment option. Further, within the MPesa system, it is not possible to set up automatic payments.

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 repayments of a random sample of ten borrowers who received the system at the end of April 2017. Each line depicts one borrower's repayment behaviour over the first nine months. The horizontal axis indicates the date of payment, while the vertical axis depicts the size of a payment in terms of the number of days for which a system was charged. It already becomes apparent that borrowers repay in very different patterns.

Each 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 that 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, then the borrower is considered to be delinquent and is attended to by a loan field officer (LFO). The LFO reminds the borrower to repay, first through messages and phone calls, then through on-site visits.[7] If a borrower is not able or not willing to repay, then the system is de-installed, and the borrower is not able to buy another system from the company.

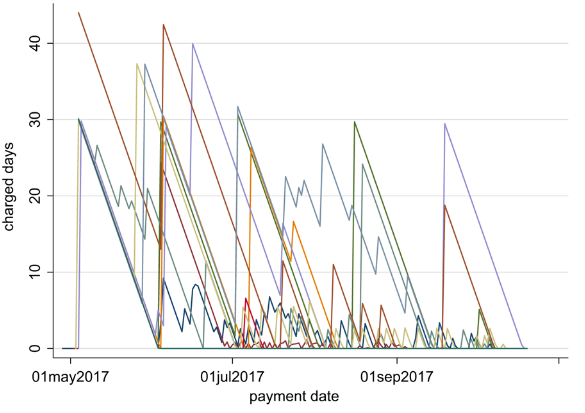



**FIGURE 1** Example of repayment behaviour. *Notes*: Payments in the first six months of ten randomly selected borrowers who received the system on 30 April 2017.

## 2.2 | Characteristics of borrowers

Individuals that consider purchasing a solar panel home system on credit form a selected sample that is not representative of the general population. As part of the loan eligibility interview, information on a large set of variables on borrowers' socioeconomic characteristics is collected, which includes occupation, wealth and resources that can be relied on in case of emergency, as well as prior experience with financial services.[8] The questions asked during the assessment interview have changed over the years, therefore not all variables are available for all borrowers. Furthermore, we have information on borrowers' location through GPS coordinates that are recorded when the solar panel home system is installed, and can therefore include geospatial information.

Table 1 presents variables of interest for all borrowers who bought a system between July 2013 and November 2018—the time period of our analysis. (Table A1 in Appendix A reports the information for the restricted sample that we use for the main part of our analysis.)[9] Some 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 borrowers are farmers (46%), or are self-employed in other ways (33%); 22% are wage-employed. Many self-employed people run more than one business. Borrowers are overall younger, have larger households, and are more likely to be self-employed or wage-employed than the general Tanzanian population. According to the 2014–15 wave of the Living Standards Measurement Survey, 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, whereas about 17% run their own business, and 13.5% are wage-employed.[10] The comparatively large proportion of wage-employed borrowers is partly due the fact that the company ran campaigns that specifically targeted teachers and nurses living in rural areas.

Prospective borrowers are asked about the disposable income that they have available at the end of each month that in principle could be used for the loan repayment. On average, borrowers report having US$29 left at the end of a month. However, as is quite common in such a context, the income information is very messy. For the later analysis, we therefore create a wealth index with wealth-related information that is available for a large proportion of the borrowers, namely land ownership and renting, employment of domestic workers, ownership of livestock, roof material, and whether the disposable income is above the sample's median. It is difficult to compare these measures to wealth indicators of the Tanzanian population. However, given the up-front deposit that the company requires, which varies between US$30 and US$65, it is likely that borrowers are slightly wealthier than the typical rural household in Tanzania.[11]

**TABLE 1** Characteristics of borrowers.

| Variable | Mean | S.D. | Min | Max | Median | Obs. |
|---|---|---|---|---|---|---|
| Female | 0.17 | 0.38 | 0.00 | 1.00 | 0.00 | 96,070 |
| Household members | 4.45 | 1.94 | 1.00 | 30.00 | 4.00 | 95,832 |
| Age | 39.20 | 11.07 | 15.00 | 89.00 | 37.00 | 50,435 |
| Farmer | 0.46 | 0.50 | 0.00 | 1.00 | 0.00 | 76,718 |
| Wage-employed | 0.22 | 0.41 | 0.00 | 1.00 | 0.00 | 76,718 |
| Self-employed | 0.33 | 0.47 | 0.00 | 1.00 | 0.00 | 76,718 |
| Number of businesses | 1.22 | 0.98 | 0.00 | 5.00 | 1.00 | 54,222 |
| Monthly disposable income (US$) | 29.11 | 28.36 | 0.21 | 340.43 | 22.08 | 51,788 |
| Owns land | 0.45 | 0.50 | 0.00 | 1.00 | 0.00 | 54,222 |
| Trading or renting land | 0.21 | 0.41 | 0.00 | 1.00 | 0.00 | 54,222 |
| Employs domestic workers | 0.27 | 0.44 | 0.00 | 1.00 | 0.00 | 54,222 |
| Owns animals that can be sold in emergency | 0.18 | 0.38 | 0.00 | 1.00 | 0.00 | 54,222 |
| Roof made out of tiles | 0.07 | 0.25 | 0.00 | 1.00 | 0.00 | 96,266 |
| Wealth index | 1.62 | 0.99 | 0.00 | 4.00 | 2.00 | 51,133 |
| Member of savings club | 0.28 | 0.45 | 0.00 | 1.00 | 0.00 | 54,222 |
| Credit experience | 0.20 | 0.40 | 0.00 | 1.00 | 0.00 | 54,222 |
| Distance to next town (km) | 16.30 | 17.79 | 0.00 | 168.44 | 10.11 | 67,967 |
| Commercial bank within 10 km | 0.47 | 0.50 | 0.00 | 1.00 | 0.00 | 67,967 |
| Community bank or MFI within 10 km | 0.45 | 0.50 | 0.00 | 1.00 | 0.00 | 67,967 |
| Savings club within 10 km | 0.61 | 0.49 | 0.00 | 1.00 | 1.00 | 67,967 |
| Mobile money agent within 10 km | 0.79 | 0.41 | 0.00 | 1.00 | 1.00 | 67,967 |
| System loan amount (US$100) | 10.34 | 3.71 | 2.41 | 25.43 | 9.26 | 88,744 |
| Business purpose | 0.08 | 0.26 | 0.00 | 1.00 | 0.00 | 52,800 |

*Notes*: Characteristics of all borrowers who bought a system between July 2013 and November 2018 on credit. Financial access information based on Financial Sector Deepening Trust (2014).

The average loan size is about US$1000. In terms of experience with formal financing, 20% of the sample report currently or previously holding a formal loan, and 28% report being member of a savings club (such as SACCO or VICO). Compared to the Tanzanian population, borrowers have more experience in the use of formal or informal financing. According to the 2017 wave of the FINDEX data, 5.3% of the population borrowed from a formal financial institution, while 9.5% borrowed from a savings club.[12] The majority of the borrowers live more than 10 km away from the nearest town. Still, half of the borrowers have a bank or MFI available in a 10 km radius, and 60% live close to a savings club. For the vast majority (79%), there is at least one mobile money agent in a 10 km radius.[13]

During the loan eligibility interview, prospective borrowers are asked about the intended purpose of the solar panel home system: 8% of borrowers indicate that they intend to use the system (also) for business purposes. This is likely an underestimation of the true proportion as many start 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 uses the system for income generation.

## 2.3 | Data

Our main analysis 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

**TABLE 2** Payment behaviour of median borrower by payment year.

| Year | No. of payments | Average paid days | S.D. paid days | Min paid days | Max paid days | Average balance | S.D. 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 | 38,405 |
| 2 | 25.00 | 8.32 | 2.41 | 1.93 | 30.07 | 0.57 | 0.90 | 11.00 | 13.05 | 37,016 |
| 3 | 19.00 | 7.52 | 2.38 | 1.88 | 26.91 | 0.55 | 0.84 | 9.00 | 9.98 | 22,692 |
| Av. | 20.00 | 11.92 | 4.24 | 2.50 | 30.50 | 0.80 | 1.30 | 9.00 | 10.10 | |

*Notes*: Median payment characteristics by payment year for the sample of borrowers used in the main analysis.

allows. In particular, we are interested in the timing, frequency and amount of payments, as well as the timing, frequency and duration of system shutdowns. For each day in the life of a borrower, we know the amount of payments that the borrower made and how much charge-time a 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 or whose loan was rescheduled.[14] 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.

Table 2 shows important features that describe the median repayment behaviour of the borrowers over the three years. It becomes apparent that the repayment behaviour is not as required in most conventional loan contracts: 50% of the borrowers pay 20 or more times per year, on average for 12 days.[15] 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 8–11 times per year, and accumulates around 10 days of grace period per year; the average shutdown thus lasts for only about one day. From Table 2, we can further infer that there is considerable heterogeneity in the way in which people repay their loans. Both the standard deviations of the paid days and the standard deviation of the balance are very large.[16]

## 3 | ANALYSIS: REPAYMENT TYPES

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 a 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 (Subsection 3.1), then investigate which customer characteristics are linked to specific repayment patterns, focusing in particular on income fluctuations and access to consumption smoothing resources (Subsection 3.2). We analyse who becomes delinquent under such a flexible repayment schedule, and what drives eventual default, in Online Appendix D.

### 3.1 | Specifying repayment types

Our repayment data constitute a time series per customer, that is, 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 is already apparent from the summary statistics reported in Table 2, there is large variation in how borrowers use the

**TABLE 3** Features used in the cluster analysis (based on a borrower's first repayment year).

| Name | Description |
|---|---|
| No. of payments | Number of distinct payments made |
| Average paid days | Mean number of days that the system was recharged with during payments, excluding the initial downpayment |
| Min paid days | Minimum number of days that the system was recharged with during the payment, excluding the initial downpayment |
| Max paid days | Maximum number of days that the system was recharged with during the payment, excluding the initial downpayment |
| S.D. paid days | Standard deviation of the days that the system was recharged with during payments, excluding the initial downpayment |
| Average balance | Mean number of charged days remaining when a borrower recharged the system with another payment |
| S.D. balance | Standard deviation of the charged days remaining when a borrower recharged the system with another payment |
| Shutdown days | Number of days the system was shut down |
| No. of shutdowns | Number of times a system was shut down |

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* classification labels to classify borrowers according to their repayment behaviour 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.[17] Second, focusing on repayment patterns in the first year only allows us to relate them to repayment behaviour in later years.

### 3.1.1 | Method

There are multiple ways to cluster time series data (for reviews, see Liao 2005; Aghabozorgi *et al.* 2015). One can 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 behaviour that is just shifted in time would not be classified as similar.[18] Using the parameters of time series models would resolve the problem of capturing time-shifted behaviour correctly, but the interpretation of the clusters that are described in terms of time series parameters would remain difficult. Furthermore, the hyperparameters of the time series models would have a very strong influence on the cluster result.[19] Therefore, we decided to use the third strategy, that is, to define features that describe the time series data. To fulfil our needs, we require features that describe the payment behaviour over time in the first year. One important property that all features must fulfil is that they may not use any information about the future, that is, information from the second or third year of payments.[20] To capture the several dimensions of repayment, the selected features reflect the payment behaviour 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 described in Table 3.

To capture as much as possible of the richness of the data, 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).[21] Gaussian mixture

**TABLE 4** Mean of features of year 1 by cluster group.

| | No. of payments | Average paid days | S.D. paid days | Min paid days | Max paid days | Average balance | S.D. 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 | 3663 | 9.54 |
| 2 | 13.96 | 26.84 | 8.16 | 6.91 | 33.14 | 3.79 | 5.28 | 4.78 | 6.97 | 4171 | 10.86 |
| 3 | 12.08 | 30.72 | 0.62 | 30.33 | 32.46 | 1.48 | 1.56 | 4.38 | 5.16 | 5211 | 13.57 |
| 4 | 41.25 | 11.06 | 8.29 | 2.54 | 31.63 | 1.52 | 2.79 | 19.00 | 12.90 | 7418 | 19.31 |
| 5 | 62.52 | 9.56 | 9.78 | 1.85 | 49.34 | 4.26 | 6.15 | 22.79 | 25.15 | 5233 | 13.63 |
| 6 | 127.98 | 3.42 | 4.49 | 0.43 | 30.64 | 0.53 | 1.37 | 60.05 | 43.56 | 6019 | 15.67 |
| 7 | 11.85 | 36.68 | 22.55 | 13.37 | 83.29 | 10.80 | 13.25 | 4.70 | 15.28 | 5119 | 13.33 |
| 8 | 17.99 | 41.72 | 28.94 | 15.11 | 99.94 | 35.66 | 21.68 | 7.82 | 51.96 | 1573 | 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*: Average feature outcome for each cluster group. Cluster groups are sorted by *ex post* labelling.

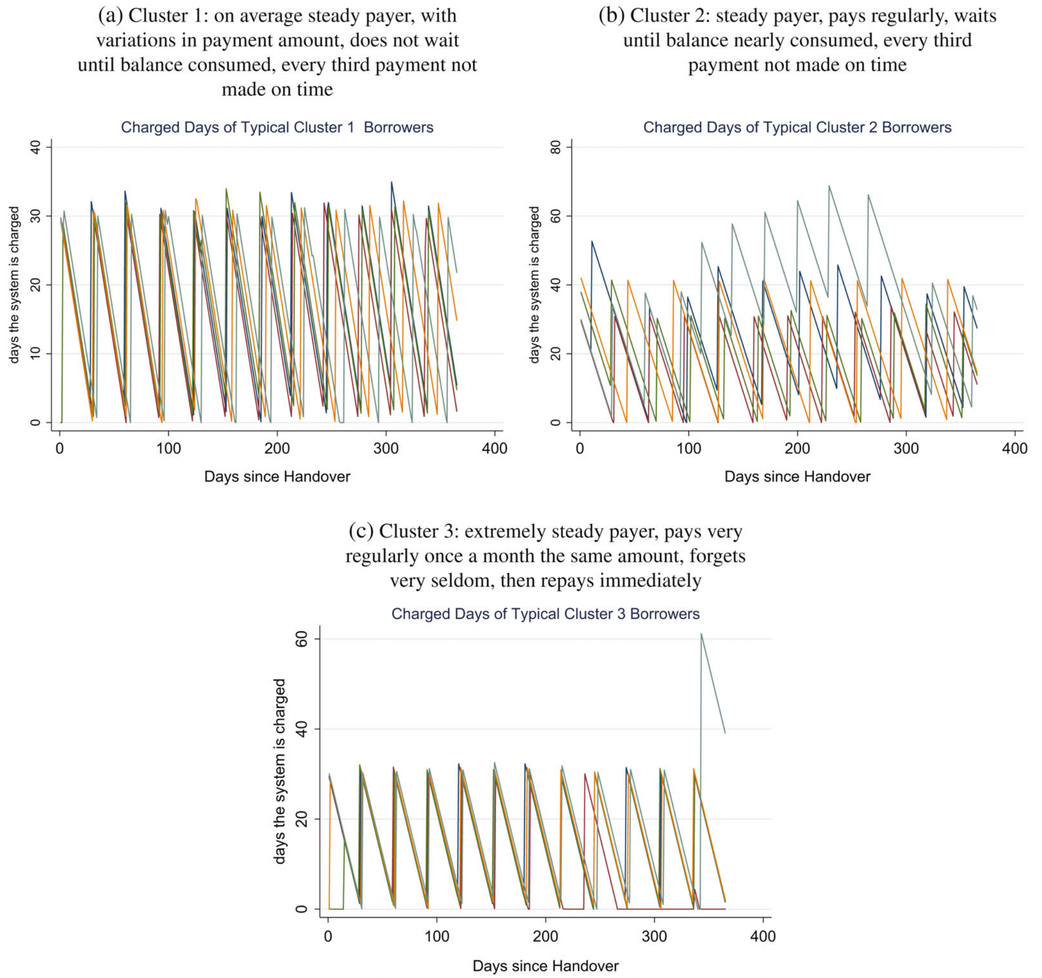
models describe data in terms of components, whereas each component is a multivariate Gaussian distribution. Thus each borrower group that we determine is described by the mean value, variance and covariance of our features. Using this clustering approach has multiple advantages for our case. First, all borrowers can be assigned to the most likely cluster, that is, the cluster whose distribution best explains the behaviour of the borrower. Additionally, we know for each cluster the probability that a borrower is assigned to this cluster. In this way, we can also identify borrowers who are typical for a cluster, that is, have an extremely high probability that they belong to the cluster. Second, by considering the mean values of the Gaussian mixtures of each cluster, 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 who 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.[22]

The clustering is applied to all data, that is, we do not need to distinguish between training data used to fit the clusters, and test data in which we evaluate the clusters. This is common in unsupervised learning, where the goodness of fit, as well as the potential for generalization, is often evaluated through subsequent manual analysis of the clusters, as is the case in our work. Automated evaluation, for example based on a test set, is usually not possible, due to a lack of ground truth data.

### 3.1.2 | Results

We first calculate for each borrower's first year of repayment the nine selected features described in Table 3. We then conduct the cluster analysis stepwise with 1–20 clusters. Comparing the Bayesian information criteria of the resulting models suggests an optimal number of 8 clusters (see Figure 1 in Online Appendix Subsection B.1). All clusters and their repayment features are reported in Table 4, sorted based on our *ex post* labelling.[23]

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.[24] Based on the feature distribution, we classify the eight clusters into three main repayment types: steady repayer, frequent repayer and irregular repayer.




| | No. of payments | Average paid days | S.D. paid days | Min paid days | Max paid days | Average balance | S.D. 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 | 3663 |
| 2 | 13.96 | 26.84 | 8.16 | 6.91 | 33.14 | 3.79 | 5.28 | 4.78 | 6.97 | 4171 |
| 3 | 12.08 | 30.72 | 0.62 | 30.33 | 32.46 | 1.48 | 1.56 | 4.38 | 5.16 | 5211 |
| Av. | 12.49 | 30.06 | 4.41 | 22.46 | 36.73 | 3.25 | 3.79 | 4.47 | 7.21 | |

**FIGURE 2** Cluster groups of steady repayers.

*Steady repayers (cluster groups 1–3)*

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 instalments suggested by the company and a limited use of the grace period. Figure 2 depicts the repayment behaviour of five typical borrowers from each of these three groups.[25]

Cluster 3 contains the most extreme sort of repayment behaviour 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 panel home 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 regularly but there is 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 the repayment scheme of a typical loan contract of a commercial bank, where instalments are due on a monthly basis. This indicates that it is possible for people with low income to repay a loan in a steady and regular manner even when the contract offers more options for 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 cover only 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 behaviours combined result in a large number of shutdowns, and many accumulated shutdown days. The shutdown days of the extreme frequent repayers (cluster 6) even exceed the 30.5 days of shutdown that a borrower can have without being considered in default. However, the average shutdown 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 among low-income borrowers, many prefer to repay very frequently, as is common for many microfinance loan contracts, where repayment is usually required on a weekly basis. 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 and 8)*

The third group consists of cluster groups 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 few shutdowns on average, but these tend to be long and build up over time. Borrowers in the irregular group typically do not wait until their balance is consumed before making a payment (when a payment is made, on average the system could have run for more than another two weeks).

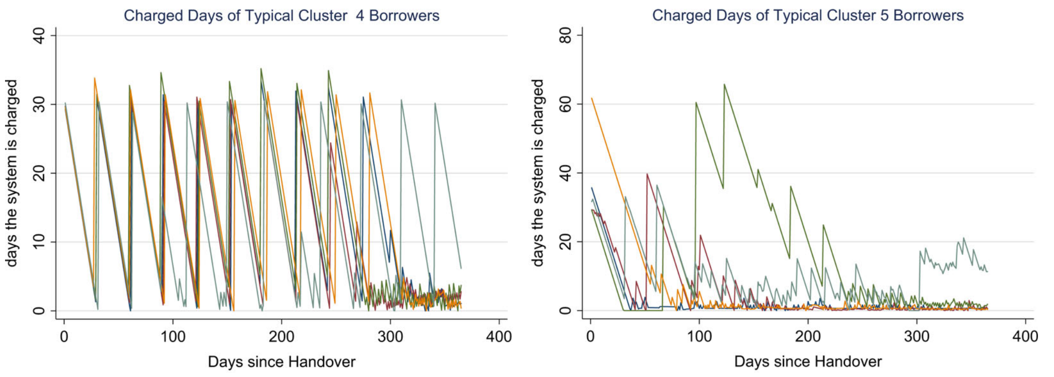
Irregular repayers make up the smallest group. About 17% of borrowers can be classified as irregular payers.

*Summary*

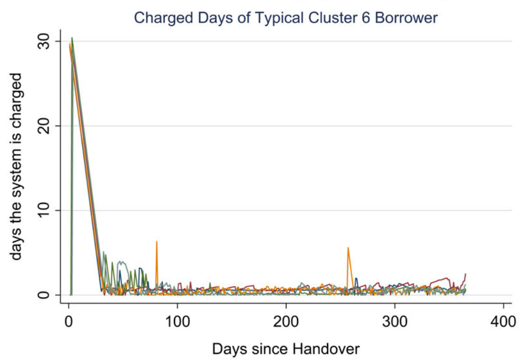
To summarize, the pattern analysis reveals clear differences in the ways that borrowers use the flexibility that the loan contract provides to them. About one-fifth (cluster 3) do not use the flexibility at all. These people repay very steadily, as would be required in most loan contracts offered by commercial banks. The largest proportion of borrowers, however, do use the flexibility, and combine the different elements of flexibility in various ways.

(a) Cluster 4: frequent payer, pays every 9 days rather small amounts, waits until balance nearly consumed, every second payment not made on time, then repays immediately

(b) Cluster 5: highly frequent payer, pays every week, typical small amounts but high variation, does not wait until balance consumed, every third payment not made on time, then repays immediately




(c) Cluster 6: extreme frequent payer, pays every third day very small amounts, waits until balance completely consumed, every second payment not made on time, then repays immediately




| | No. of payments | Average paid days | S.D. paid days | Min paid days | Max paid days | Average balance | S.D. balance | No. of shutdowns | Shutdown days | No. of obs. |
|---|---|---|---|---|---|---|---|---|---|---|
| 4 | 41.25 | 11.06 | 8.29 | 2.54 | 31.63 | 1.52 | 2.79 | 19.00 | 12.90 | 7418 |
| 5 | 62.52 | 9.56 | 9.78 | 1.85 | 49.34 | 4.26 | 6.15 | 22.79 | 25.15 | 5233 |
| 6 | 127.98 | 3.42 | 4.49 | 0.43 | 30.64 | 0.53 | 1.37 | 60.05 | 43.56 | 6019 |
| Av. | 75.17 | 8.18 | 7.48 | 1.67 | 36.27 | 1.97 | 3.27 | 33.30 | 26.22 | |

**FIGURE 3** Cluster groups of frequent repayers.

## 3.2 | Repayment behaviour and socioeconomic background

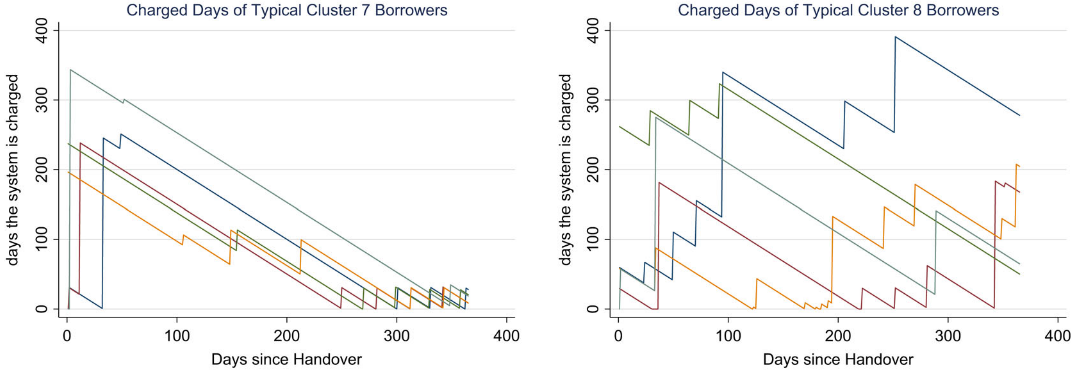
Are certain socioeconomic characteristics related to a specific repayment behaviour? In particular, are borrowers with higher income fluctuations and fewer resources to smooth consumption more likely to use the loan contract's flexibility?

We estimate the joint likelihood with a multinomial logistic regression:

$$\left.\begin{aligned} P(CG^{FP}) &= \exp(\alpha^{FP} I_i + \beta^{FP} R_i + \gamma^{FP} X_i + \eta^{FP}_{lfo} + \zeta^{FP}_{p} + \delta^{FP}_{y})\ k^{-1}, \\ P(CG^{IP}) &= \exp(\alpha^{IP} I_i + \beta^{IP} R_i + \gamma^{IP} X_i + \eta^{IP}_{lfo} + \zeta^{IP}_{p} + \delta^{IP}_{y})\ k^{-1}, \end{aligned}\right\} \tag{1}$$

(a) Cluster 7: irregular payer, rather seldom very irregular payments, sometimes large, sometimes small amounts; high balance, when next payment is made; high shut-off times for longer periods

(b) Cluster 8: highly irregular payer, more frequent, but very irregular payments, with high variations in payment amounts; more than every second payment not made on time, then long shutdown durations




| | No. of payments | Average paid days | S.D. paid days | Min paid days | Max paid days | Average balance | S.D. balance | No. of shutdowns | Shutdown days | No. of obs. |
|---|---|---|---|---|---|---|---|---|---|---|
| 7 | 11.85 | 36.68 | 22.55 | 13.37 | 83.29 | 10.80 | 13.25 | 4.70 | 15.28 | 5119 |
| 8 | 17.99 | 41.72 | 28.94 | 15.11 | 99.94 | 35.66 | 21.68 | 7.82 | 51.96 | 1573 |
| Av. | 13.29 | 37.86 | 24.05 | 13.78 | 87.20 | 16.65 | 15.23 | 5.43 | 23.90 | |

**FIGURE 4** Cluster groups of irregular repayers.

with

$$k = 1 + \exp(\alpha^{FP} I_i + \beta^{FP} R_i + \gamma^{FP} X_i + \eta_{lfo}^{FP} + \zeta_p^{FP} + \delta_y^{FP}) + \exp(\alpha^{IP} I_i + \beta^{IP} R_i + \gamma^{IP} X_i + \eta_{lfo}^{IP} + \zeta_p^{IP} + \delta_y^{IP}),$$

where $CG_{i,p,y}^{J} = 1$ if borrower $i$ living in province $p$, having received the system in year $y$, belongs to cluster group $J$, with $J = \{SP, FP, IP\}$ (steady, frequent, irregular), derived by the EM algorithm described in Subsubsection 3.1.1. The steady repayment group, $CG^{SP}$, is used as reference category. We aim to explain borrower $i$'s probability of displaying a certain repayment pattern $CG_i^J$ by the income fluctuations that she is facing, $I_i$, the resources that she has available to smooth consumption, $R_i$, and individual characteristics, $X_i$. Furthermore, we include fixed effects for the LFO who is in charge of borrowers, $\eta_{lfo}$, as LFOs have considerable leeway in managing their borrowers.[26] In addition, we include province fixed effects, $\zeta_p$, and dummies for the year of handover, $\delta_y$. Standard errors are clustered at the LFO level.

Here, $I$ includes variables that proxy a borrower's income fluctuations throughout the year: namely, the borrower's main occupation (which we broadly classify as farmer, self-employed or wage-employed), the number of businesses that a borrower has, and whether the borrower planned (at the time of the loan assessment interview) to earn additional money with the system (business purpose). Also, $R$ includes variables that proxy a borrower's access to resources to smooth consumption over time, namely her wealth index and credit experience, and whether she is a member of a savings club. We account for financial access by including indicators on whether the borrower lives within a 10 km radius of a community or microfinance bank or savings club. Furthermore, to proxy for market access, we include an indicator on whether the borrower lives within a 10 km radius of the next town. Finally, $X$ includes individual characteristics, namely whether the borrower is female, the number of household members, and the loan amount that

**TABLE 5** Repayment pattern and socioeconomic characteristics—multinomial logit.

| | Frequent repayer (1) | Irregular repayer (2) | Frequent repayer (3) | Irregular repayer (4) |
|---|---|---|---|---|
| Female | 0.827 | 0.928 | 0.849 | 0.900 |
| | (0.035)*** | (0.047) | (0.046)*** | (0.077) |
| Household members | 0.925 | 1.008 | 0.937 | 1.034 |
| | (0.008)*** | (0.008) | (0.012)*** | (0.016)** |
| Farmer | 1.273 | 1.385 | 1.263 | 1.350 |
| | (0.038)*** | (0.053)*** | (0.059)*** | (0.097)*** |
| Wage-employed | 0.613 | 0.760 | 0.630 | 0.734 |
| | (0.022)*** | (0.035)*** | (0.030)*** | (0.071)*** |
| System loan amount (log) | 0.869 | 0.897 | 0.798 | 0.850 |
| | (0.038)*** | (0.050)** | (0.054)*** | (0.076)* |
| Next town within 10 km | 0.958 | 0.932 | 0.966 | 0.929 |
| | (0.044) | (0.058) | (0.066) | (0.061) |
| Community bank or MFI within 10 km | 0.973 | 0.993 | 0.962 | 0.939 |
| | (0.062) | (0.061) | (0.085) | (0.085) |
| Savings club within 10 km | 1.018 | 1.001 | 1.064 | 1.069 |
| | (0.058) | (0.063) | (0.073) | (0.097) |
| No. of businesses | | | 1.073 | 0.967 |
| | | | (0.027)*** | (0.033) |
| Business purpose | | | 1.372 | 1.066 |
| | | | (0.116)*** | (0.142) |
| Wealth index | | | 0.937 | 1.046 |
| | | | (0.022)*** | (0.035) |
| Credit experience | | | 0.778 | 0.751 |
| | | | (0.042)*** | (0.061)*** |
| Member of savings club | | | 1.190 | 0.986 |
| | | | (0.081)** | (0.088) |
| Observations | 35,154 | | 18,680 | |
| LFO fixed effects | Yes | | Yes | |
| Handover year fixed effects | Yes | | Yes | |
| Province fixed effects | Yes | | Yes | |

*Notes*: Multinomial logistic regression. Outcome is the repayment cluster group to which a borrower is assigned based on her payment behaviour in the first year, i.e. steady repayer, frequent repayer or irregular repayer. The omitted category is being assigned to steady repayer. Coefficients are reported as relative risk estimator. Standard errors in parentheses, clustered on LFO level.
*, **, *** denote significance levels at 10%, 5%, 1%, respectively.

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 provided only to more affluent borrowers.

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

Results are reported in Table 5. Coefficients are shown as relative risk estimators in reference to the steady repayment group; that is, a coefficient larger (smaller) than 1 indicates that a borrower is more (less) likely to be classified as a frequent/irregular repayer than as a steady repayer if the variable in question changes. Columns (1) and (2) describe the estimated effects for the full sample, and columns (3) and (4) for the restricted sample.[27] We see a clear pattern: borrowers who work in occupations with more irregular income (such as farmers or self-employed) are more likely to repay very frequently or irregularly, while wage-employed borrowers are more likely to repay steadily on a monthly basis. Furthermore, borrowers who have several small businesses, and those who planned to use the system to earn additional money, have a higher likelihood of repaying very frequently. This indicates that borrowers use the flexibility to align their payments to their individual cash flows. Access to finance or markets cannot explain differences in repayment behaviour. Yet borrowers with access to consumption smoothing resources, such as wealth, and experience with formal loans, are more likely to pay steadily on a monthly basis instead of paying small amounts very frequently, while being a member of a savings club is rather associated with frequent payments. Borrowers paying irregularly also tend to work in occupations with less regular cash flows, but they are more comparable to steadily repaying borrowers in terms of wealth and side businesses. It is conceivable that these are borrowers who earn the main bulk of their income only once or twice per year, such as farmers with large land plots and no side businesses, and then make their payments. Note that due to the limited information available on each borrower (e.g. there is no information on education, limited information on wealth, and no information on other income sources), the analysis might suffer from omitted variable bias. In particular, we have only limited information on wealth, and do not know the borrower's level of education, which can affect both access to formal finance and also loan repayment directly. Nevertheless, we are confident that the link between occupation and repayment patterns persists, as the following analysis also confirms.

# 4 | ANALYSIS: RESPONSE TO VARIATION IN INCOME

## 4.1 | Cash flow fluctuations and repayment

The analysis above reveals a clear relationship between occupation and repayment patterns. Borrowers who work in occupations with higher cash flow 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, that is, paying more when they have more cash on hand? We do not have good information on borrowers' exact cash flows, but for farmers we can assume that their income is particularly high during the 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 of when the relevant local growing, harvesting and lean seasons start and end. We define the local growing seasons at 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. In the following, for ease of reading, we imprecisely call the time period outside the growing season 'harvest season'. Yet we have a good proxy only for each ward's start of the harvest season (which sets in after the growing season); we do not know its 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 what we call harvest season (both defined at the ward level)—into three parts of equal length: start of growing (harvest) season, mid growing (harvest) season, and end of growing (harvest) season.

We then analyse repayment during these so-defined periods, assuming that the first two periods of what we call harvest season are likely the time periods when crops are harvested and harvest

**TABLE 6** Average payment amount and seasonality.

| | Average weekly payment | | | | | |
|---|---|---|---|---|---|---|
| | Farmers | | Maize farmers only | | Non-farmers | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Harvest season | 0.430 | | 0.576 | | 0.141 | |
| | (0.035)*** | | (0.049)*** | | (0.036)*** | |
| Growing season, mid | | −0.440 | | −0.569 | | −0.168 |
| | | (0.060)*** | | (0.086)*** | | (0.062)*** |
| Growing season, end | | −0.374 | | −0.635 | | −0.201 |
| | | (0.060)*** | | (0.086)*** | | (0.062)*** |
| Harvest season, start | | −0.058 | | −0.116 | | −0.059 |
| | | (0.062) | | (0.089) | | (0.065) |
| Harvest season, mid | | 0.772 | | 0.993 | | 0.322 |
| | | (0.058)*** | | (0.083)*** | | (0.060)*** |
| Harvest season, end | | 0.514 | | 0.764 | | 0.132 |
| | | (0.058)*** | | (0.083)*** | | (0.060)** |
| Constant | 5.624 | 5.907 | 5.155 | 5.574 | 5.801 | 5.930 |
| | (0.024)*** | (0.044)*** | (0.035)*** | (0.063)*** | (0.025)*** | (0.045)*** |
| Observations | 533,250 | 533,250 | 224,084 | 224,084 | 447,789 | 447,789 |

*Notes*: Borrower fixed effect regression; standard errors in parentheses. Outcome is the average payment in a week. Onset and cessation of growing seasons are defined using local daily precipitation data. Start, mid and end refer to the first, second and third parts of a season, when it is divided into three parts of equal length. Borrowers living in urban areas are excluded.
*, **, *** denote significance levels at 10%, 5%, 1%, respectively.

income is realized, while the third period is probably reflecting the lean period. If farmers make use of the repayment flexibility, then we should expect higher payments in the harvest periods.

We estimate the following borrower fixed effect model focusing on borrowers in rural areas only:[28]

$$y_{itw} = \alpha + \beta_1 \ Harvestseason_{tw} + \gamma_i + u_{itw}, \tag{2}$$

where $y_{itw}$ is the average payment amount or the number of payments of customer $i$ in week $t$ living in ward $w$. As above, the payment amount is expressed in terms of charge days. $Harvestseason_{tw}$ is a dummy that is 1 if week $t$ is outside of the growing season specified for ward $w$, and 0 otherwise. In a second step, we look at the effect of the seasons in more detail by regressing our outcome variables on dummy variables, indicating whether week $t$ is in the mid or end of the growing season, or the beginning, mid or end of what we call harvest season. Beginning of growing season acts as the base category. In all regressions, $\gamma_i$ denotes the individual borrower fixed effect.

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 analyse the impact of seasonality on non-farmers' repayments (columns (5) and (6)). The graphs in Figure 5 replicate the estimated dynamics.

Indeed, we find that farmers pay larger amounts in the harvest season (about 7%), particularly so in the second part, where payment increases by 13%, which is probably when yield income materializes. As expected, the results are even stronger for maize farmers, who pay 18% more in the second part of the harvest season compared to the beginning of the growing period. The effect

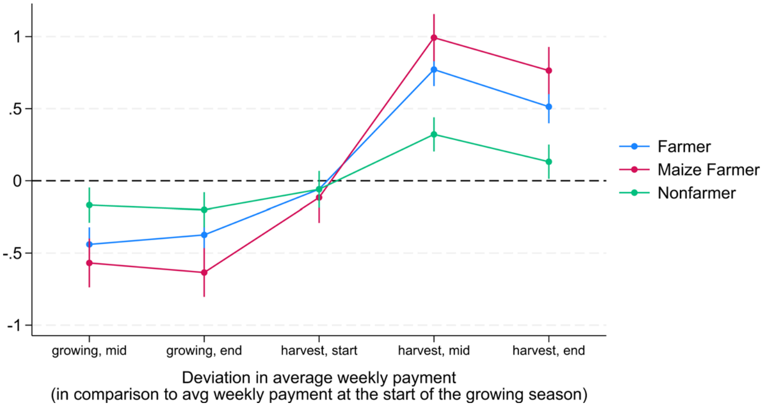



**FIGURE 5** Deviation in weekly payment amount over the season. *Notes*: Coefficients of borrower fixed effect regressions estimating the weekly number of payments over seasons. Onset and cessation of growing seasons are defined using local daily precipitation data. Start, mid and end refer to the first, second and third parts of a season, when it is divided into three parts of equal length.

is considerably smaller for non-farmers, but still their income and thus their repayments seem to be seasonally dependent. This is not unreasonable as many non-farmers run small businesses (such as shops or bars) that benefit from the increase in farmers' income and their spending capacities during harvest period.

In the Online Appendix, we report results comparing weekly payments between farmers and non-farmers directly (see Table 6 in Online Appendix Subsection C.2). The differences in the payment behaviours along the different season periods are highly significant. During the harvest season, farmers pay larger amounts than non-farmers, particularly so in the second part. On the other hand, farmers pay overall lower amounts at the peak of the growing season when probably less resources are available.

We see similar patterns when looking at the average number of payments per week (see Table 7 and Figure 6). Borrowers pay more frequent and smaller amounts when less cash is on hand during the growing season, and less frequently during harvest season, in particular in the second and third parts. The effects are more pronounced for maize farmers (columns (3) and (4)) and less pronounced for non-farmers (columns (5) and (6)). Again, the differences in the number of payments between farmers and non-farmers along the season periods are highly significant (see Table 7 in Online Appendix Subsection C.2).

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

## 4.2 | Shocks to harvest and repayment

So far, we have identified different repayment types and have seen that farmers are more likely to repay flexibly, while workers with fixed incomes are more likely to repay in a steady manner. We further showed that farmers repay less often but larger amounts during harvest season (i.e. once harvest income materializes) than during the growing season, indicating that borrowers adjust their repayment to their income streams.

To provide further evidence, we study the impact of income shocks on loan repayment. More specifically, we analyse whether borrowers who experience vegetation shocks during the growing season adjust their repayment during the following harvest season, when faced with a decline

**TABLE 7** Number of payments per week and seasonality.

| | No. of payments | | | | | |
|---|---|---|---|---|---|---|
| | Farmers | | Maize farmers only | | Non-farmers | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Harvest season | −0.106 | | −0.143 | | −0.068 | |
| | (0.003)*** | | (0.005)*** | | (0.003)*** | |
| Growing season, mid | | 0.105 | | 0.144 | | 0.050 |
| | | (0.005)*** | | (0.008)*** | | (0.005)*** |
| Growing season, end | | 0.175 | | 0.273 | | 0.084 |
| | | (0.005)*** | | (0.008)*** | | (0.005)*** |
| Harvest season, start | | 0.104 | | 0.196 | | 0.038 |
| | | (0.005)*** | | (0.009)*** | | (0.005)*** |
| Harvest season, mid | | −0.148 | | −0.196 | | −0.091 |
| | | (0.005)*** | | (0.008)*** | | (0.005)*** |
| Harvest season, end | | −0.246 | | −0.387 | | −0.137 |
| | | (0.005)*** | | (0.008)*** | | (0.005)*** |
| Constant | 0.905 | 0.808 | 1.163 | 1.018 | 0.895 | 0.848 |
| | (0.002)*** | (0.003)*** | (0.003)*** | (0.006)*** | (0.002)*** | (0.004)*** |
| Observations | 533,250 | 533,250 | 224,084 | 224,084 | 447,789 | 447,789 |

*Notes*: Borrower fixed effect regression; standard errors in parentheses. Outcome is the number of payments in a week. Onset and cessation of growing seasons are defined using local daily precipitation data. Start, mid and end refer to the first, second and third parts of a season, when it is divided into three parts of equal length. Borrowers living in urban areas are excluded.
*, **, *** denote significance levels at 10%, 5%, 1%, respectively.

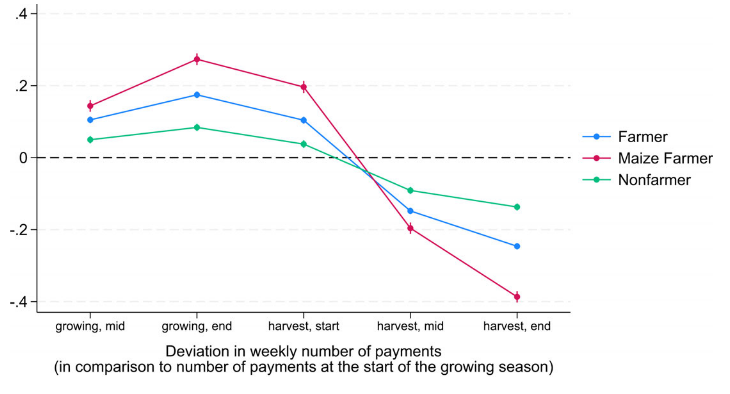



**FIGURE 6** Deviation in weekly number of payments over the season. *Notes*: Coefficients of borrower fixed effect regressions estimating the weekly number of payments over seasons. Onset and cessation of growing seasons are defined using local daily precipitation data. Start, mid and end refer to the first, second and third parts of a season, when it is divided into three parts of equal length.

in yields. We focus on vegetation shocks as most of the borrowers who we study live in rural areas, typically with occupations linked to the agricultural sector. Agricultural shocks, in terms of droughts, pests or diseases, are very common in Tanzania, and are expected to increase over the next decades (Tomalka *et al.* 2020; Rahut *et al.* 2021).

We measure vegetation shocks using the Normalized Difference Vegetation Index (NDVI), which is calculated based on remote sensing data, measuring the presence of chlorophyll pigment and therefore the health of a plant (Tucker *et al.* 2005). Data come from the US Center for Satellite Applications and Research of the National Oceanic and Atmospheric Administration, and have spatial resolution $4 \times 4\,\text{km}^2$ (NOAA STAR 2018). We define local vegetation shocks by computing deviations of the ward-specific seasonal NDVI on a weekly level during the growing season from its long-term average, measured from 1995 to 2014.[29] Local seasons are defined based on rainfall data following the same definition as described above.

More specifically, we estimate

$$y_{isw} = \alpha + \beta_1\ NDVIShock_{(s-1)w} + \beta_2\ LengthofOwnership_{isw} + \gamma_i + u_{isw}, \tag{3}$$

where $y_{isw}$ measures different aspects of repayment behaviour of borrower $i$ living in ward $w$ during the harvest seasons $s$, namely the number of payments, the total payment amount, the average payment amount when a payment is made, and the average balance when a payment is made, as well as the number of days for which the system is shut down. As above, payment amounts and the balance are expressed in terms of charge days. *NDVIShock* measures the accumulated negative percentage deviation of the NDVI in the preceding growing season $(s-1)$ from the NDVI in the usual growing seasons at the level of the borrower's ward $w$. For ease of interpretation, the shock measure is standardized to have mean 0 and a standard deviation 1. We include borrower fixed effects, and control for the number of weeks that borrower $i$ had his system during the harvest season. While the lengths of the seasons are fixed on a ward level and thus captured in the borrower fixed effects, we might observe shorter seasons if borrowers have started or finished repaying during that specific season. Standard errors are clustered at the ward level, that is, the level of our treatment (the vegetation shock). We estimate equation (3) separately for farmers, maize farmers and non-farmers, focusing on rural areas only.

Results are shown in Table 8. Panel A reports the results for borrowers who are farmers, panel B reports results for maize farmers only, and panel C reports results for non-farmers.[30] In the even-numbered columns, borrower fixed effects are included. We see clear implications of vegetation shocks for repayment behaviour. A one standard deviation increase in our local shock measure increases the number of payments of the affected farmers in the following harvest season by 30% compared to its mean; the total payment amount reduces by 5%, the average payment amount reduces by 16%, the average balance when a payment is made reduces by 15%, and the number of shutdown days increases by 22%. Farmers experiencing harvest loss thus respond by paying smaller amounts more often. Interestingly, the number of shutdown days increases by less than the total payment amount decreases, suggesting that the loan's flexibility—namely, the possibility to make larger payments when cash is available to provide for more uncertain times—allows farmers to buffer some of the income loss.

Effects are considerably more pronounced for farmers who produce maize (see panel B of Table 8) for which our shock measure is arguably more precisely defined. Effects can also be observed for non-farmers living in rural areas (panel C), though the estimated effect sizes are smaller and less precisely estimated. This is not surprising as in many areas of Tanzania, the local economy depends strongly on farming. In other words, if local farmers in the area have less cash available due to an agricultural shock, then this also affects other people living in the area.

## 5 | DISCUSSION

Above, we show that borrowers who have more stable incomes and who have resources to smooth consumption repay in a more regular manner, that is, they are more likely to be steady repayers. Furthermore, in particular for borrowers who earn their main income from agriculture, we can show that they make use of the flexibility to match payments to their cash flows, which can allow them to remain solvent in times of income shocks. However, the data in this study are purely

**TABLE 8** NDVI shock and repayment.

| | No. of payments | | Total payment amount | | Average payment | | Average balance | | No. of shutdown days | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| *Panel A: Farmers* | | | | | | | | | | |
| NDVI shock | 1.959 | 3.583 | −1.118 | −4.819 | −1.543 | −3.688 | −0.430 | −0.820 | 0.564 | 1.132 |
| | (0.328)*** | (0.330)*** | (0.336)*** | (0.750)*** | (0.346)*** | (0.427)*** | (0.191)** | (0.250)*** | (0.100)*** | (0.196)*** |
| Customer fixed effects | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ |
| Mean | 12.17 | | 102.85 | | 22.66 | | 5.34 | | 5.02 | |
| Observations | 17,020 | | 17,020 | | 16,592 | | 15,948 | | 17,020 | |
| *Panel B: Maize farmers* | | | | | | | | | | |
| NDVI shock | 1.824 | 4.457 | −3.083 | −10.598 | −2.534 | −6.702 | −0.482 | −1.889 | 0.541 | 1.496 |
| | (0.370)*** | (0.459)*** | (0.521)*** | (0.995)*** | (0.366)*** | (0.545)*** | (0.257)* | (0.359)*** | (0.161)*** | (0.259)*** |
| Customer fixed effects | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ |
| Mean | 14.73 | | 98.25 | | 20.12 | | 4.35 | | 5.65 | |
| Observations | 7644 | | 7644 | | 7479 | | 7235 | | 7644 | |
| *Panel C: Non-farmers* | | | | | | | | | | |
| NDVI shock | 1.679 | 2.159 | −0.170 | −3.534 | −1.061 | −2.007 | −0.260 | −0.775 | 0.315 | 0.762 |
| | (0.325)*** | (0.424)*** | (0.251) | (0.729)*** | (0.283)*** | (0.417)*** | (0.205) | (0.265)*** | (0.140)** | (0.212)*** |
| Customer fixed effects | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ |
| Mean | 12.12 | | 98.95 | | 21.91 | | 5.05 | | 4.74 | |
| Observations | 14,240 | | 14,240 | | 13,824 | | 13,309 | | 14,240 | |

*Notes*: This table shows the estimated effects of a crop shock in the growing season on repayment behaviour in the following harvest season, separated by customers' types of occupation. Uneven-numbered columns: OLS estimates. Even-numbered columns: fixed effect regression using customer fixed effects. Control for length of ownership in current harvest season. Standard errors (in parentheses) are clustered at the ward level. NDVI shock is the accumulated negative percentage deviation of the NDVI in the preceding growing season from the NDVI in the usual growing seasons measured on the level of a borrower's ward.
*, **, *** denote significance levels at 10%, 5%, 1%, respectively.

observational, and we deal with a selected sample. Importantly, people might have entered into this repayment contract with different experiences, expectations and preferences regarding repayment, which influence the way in which they repay, and repayment behaviour may be driven by factors that we are unable to capture in this study. It is thus possible that some of the patterns that we find have explanations other than the ones discussed above. In the following, we discuss some potential alternative explanations for our findings.

*Transaction costs:* Potentially, differences in repayment patterns are due to differences in transaction costs rather than due to income fluctuations or access to resources. Some borrowers may have to travel far to reach a mobile money agent, or they may have poor understanding of mobile money and so have to rely on the help of someone else to make payments. In such cases, borrowers may find it easier to collect payments and transfer larger sums at a time. In theory, this could result in repayment patterns that look like steady or irregular repayment patterns. In practice, we believe that this is unlikely. For the loan, borrowers are required to have a mobile money account, and they are aware that this is the only mode by which to make payments. In the eligibility interview, 89.6% of the borrowers report using mobile money often. Within our data, we do not find any indication that familiarity with mobile money or distance to the agent influence repayment patterns. Even if transaction costs played a role, they should rather work against the relationship that we find between financial circumstances and repayment, as the richer borrowers in our sample are more familiar with mobile money and thus face lower transaction costs; yet we find that they make less frequent payments.[31]

*Habit formation through formal loans:* We find that credit experience can predict steady repayment, and assume that loans act as a consumption smoothing resource. But potentially, borrowers with credit experience are more likely to repay steadily, as they adopt the repayment pattern that they are used to when repaying other loans. Again, we believe that although this is not impossible, it is unlikely. In the loan eligibility interview, applicants were asked which banks they had borrowed from in the past. When analysing the repayment schemes of the reported banks, we find that besides monthly payments, many also offer weekly and bi-weekly payment schemes. To gather further evidence on this, we look at the actual dates on which steady repayers transfer money. If they simply transferred money on the day that they transferred instalments for their other loans, then we should see a spike of large payments made on the first and the last few days of the months. However, we find that payments are equally likely on all days of a month. Therefore we do not believe that habit formation can explain the link between access to formal loans and steady repayment.

*Change in usage versus adjustment to cash flows:* Studying repayment across seasons or in the aftermath of vegetation shocks, we find evidence that borrowers adjust their payments to their cash flows, making smaller and more frequent payments when cash is low. An alternative explanation for the observed relationship could be that borrowers change how they use their system across seasons or after experiencing an income shock, which might affect borrowers' valuation of the system and in turn their repayment. For example, they might use the system less after experiencing harvest loss as they are less at home when looking for alternative income sources, which in turn makes the system less relevant for them, potentially affecting repayment. However, this seems unlikely to be the main underlying mechanism, as it would explain only a reduction in total and average payment amounts, but not the increase in payment frequency. Indeed, regressing average electricity consumption on the NDVI shock, we find no impact of the vegetation shock experienced during growing season on average electricity consumption in the following harvest season.

*Summary:* In summary, while we cannot completely rule out other mechanisms, we believe that they are unlikely to drive our findings. The relations that we find when linking different uses of flexibility with borrowers' financial circumstances are reasonable and in line with economic theory.

# 6 | CONCLUSION

In this paper, we study a financial product that enables the purchase of an important asset with a repayment scheme that is adapted to the financial circumstances of borrowers with little income and varying cash flows. We show that borrowers make use of the repayment flexibility in a number of different ways, depending on their financial situation. In particular, borrowers with fluctuating cash flows and those with limited resources to smooth consumption, repay in a more flexible manner. Furthermore, we find evidence that repayment is adjusted to cash flows, and payments are made preemptively; in particular, farmers pay larger amounts when cash is on hand (i.e. during harvest), and are thereby able to provide for months of lower cash flow. We also show that borrowers respond to shocks by repaying smaller amounts more frequently.

Our study indicates that there is scope for financial products targeted to the poor that do not follow the traditional microfinance schemes with small loan sizes and fixed instalments. When provided with flexible repayment schemes, borrowers make use of it and align their payments to their financial circumstances, which might prevent overindebtedness.

The loan product provides people with low-income access to an asset that they would not have been able to acquire otherwise, and which has important implications for their lives. The solar panel home system can improve welfare in a number of different ways. Lights provide safety and enable household members to study and work after sunset; television and radio offer access to information; and electricity in general provides means to diversify income, which is vital in particular for farmers in times of increasing climatic uncertainties (Stemmler *et al.* 2024). Furthermore, beyond the positive effects associated with the use of solar panels and the connected appliances, possession alone can increase the perceived and actual social status of a household, and thereby improve mental wellbeing (Banerjee *et al.* 2015a).

One can imagine other assets to be combined with a similar financial product, such as solar-powered irrigation pumps, refrigerators or certain types of farm equipment. The flexible repayment scheme relies on the fact that non-payment can be directly sanctioned, that is, the system is switched off and can therefore be used as digital collateral. This technology already exists for a number of other products, such as mobile phones and some motor vehicles (Gertler *et al.* 2024). When such technology exists, it can allow for a flexible repayment scheme. Without such immediate sanctions in place, flexible repayment schemes might be less feasible. Furthermore, while repayment via mobile money allows for the inclusion of borrowers living remotely, these still need to be attended regularly. Understanding the complex system is crucial. Borrowers need to understand the particularities of the repayment scheme, and need to be reminded when payments are overdue, which comes with costs. Exploring under which circumstances such flexible repayment works in other settings is an interesting avenue for future research.

Importantly, this study focuses solely on the use of flexibility. It does not comment on welfare implications of different forms of repayment. We further cannot assess if borrowers use flexibility in an optimal way for their circumstances. To understand the consequence of being able to repay flexibly, more research is needed.

## ACKNOWLEDGMENTS

For helpful comments, we thank Jenny Aker, Kristina Czura, Krisztina Kis-Katos, Stephan Klasen, Stefan Klonner, Lukas Menkhoff, Lisa Spantig, and seminar participants at Aarhus University, GIGA, PIK and University of Göttingen, as well as participants at Danish Finance Institute Retreat 2024, GDE Conference 2023, PEGNet 2019, DIAL Development Conference Paris 2019, Africa and Europe Forward Together Conference 2019, International Conference on Globalization and Development 2018, and SABE/IAREP Conference London 2018. Financial support by the German Research Foundation (DFG, grant RTG 1723 and CRC 190) is gratefully acknowledged. Open Access funding enabled and organized by Projekt DEAL.

## ENDNOTES

[1] These types of contract are increasingly gaining in popularity. For example, motorcycle taxi drivers in Kampala, Uganda, can acquire their motorcycle through a down-payment and monthly instalments. Ownership of the motorcycle is then transferred after the last instalment is paid (Kumaraswamy *et al.* 2020).

[2] PAYG systems—an increasingly common approach to expanding access to electricity (see Barry and Creti 2020)—offer another way for low-income households to benefit from otherwise unaffordable assets or services. Under this model, users pay only when they consume the service (e.g. electricity from a solar panel), and only when they have the means to do so. However, unlike hire purchase arrangements, PAYG payments go towards usage rather than ownership of the asset.

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

[4] The annual GNI per capita in current terms in Tanzania for the years 2014–17 was $970 (see https://data.worldbank.org/indicator/NY.GNP.PCAP.CD?view=chart, accessed 10 August 2025).

[5] See https://www.mftransparency.org/microfinance-pricing/tanzania (accessed 10 August 2025). The rate of inflation was between 3% and 5.5% in the time period that we are observing (see https://data.worldbank.org/indicator/FP.CPI.TOTL.ZG?locations=TZ, accessed 10 August 2025).

[6] As of 2018, over 450,000 mobile money agents were spread across the country (Bank of Tanzania 2022).

[7] Borrowers have to pay for the number of days they exceeded the grace period. Until September 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 back on.

[8] These data are not unproblematic. Prospective borrowers have an incentive to provide inaccurate information to questions that 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. In the following analysis, we focus on variables that are considered to be reliable.

[9] As there were many operational changes and trials when the company was first set up, we exclude the first two years in all our analyses.

[10] See https://microdata.worldbank.org/index.php/catalog/lsms (accessed 10 August 2025).

[11] In a detailed phone survey that we conducted with a small subset of borrowers in 2018, the majority of borrowers reported a monthly household income of less than US$74, and thus would be considered as poor. Still, such information needs to be taken with a pinch of salt as it is difficult to retrieve accurate income information in such a setting.

[12] See https://globalfindex.worldbank.org/#about_focus (accessed 10 August 2025).

[13] The information on financial access is based on a geospatial mapping of all cash outlets in Tanzania initiated by Financial Sector Deepening Trust (2014). The latest update was in 2014; as the majority of borrowers have taken up a loan after 2014, financial access is likely underestimated.

[14] We have to exclude borrowers who bought more than one system as in these cases, repayment cannot be attributed to a specific system. Furthermore, for borrowers whose loan was rescheduled (approximately 3% of the borrowers), repayments were shifted manually and the recorded repayment data are thus not reliable.

[15] The higher average paid days in the first year are a result of the first instalment that the borrower makes, which has to cover at least one month.

[16] A more detailed description of the repayment feature distributions can be inferred from the kernel density plots depicted in Online Appendix Subsection B.3.

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

[18] For example, consider two customers A and B. Customer A starts with a downpayment that charges the system for 50 days, and 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 behaviour 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.

[19] For example, for ARIMA time series models, the hyperparameters define how many steps in the past values, and how much of the previously introduced white noise, are 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 hyperparameters may vary not only 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 hyperparameters are always the same, which would mean that many time series models would be using bad hyperparameters, resulting in a bad fit.

[20] While 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 year but finish only in the second year.

[21] There are a number of different algorithms to derive the cluster groups, for example, 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 the form of a centroid, this centroid

describes the borrower groups only 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.

[22] Using EM clustering has the drawback that the complete description of the Gaussian mixture models requires the covariance matrices for each cluster. This means that 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 need to identify only which kind of borrower is within a cluster.

[23] The correlation across features is depicted in Figure 4 of Online Appendix Subsection B.5.

[24] Note that in Table 4, only averages are reported. Within groups, certain features can vary considerably (see Online Appendix Subsection B.4).

[25] The depicted borrowers in Figures 2–4 are randomly drawn from those borrowers that had the 5% highest probability of being assigned to the respective cluster.

[26] Borrowers are assigned to an LFO based on the branch where the system was acquired. Over the time period that we analyse, there are in total 54 LFOs that are each responsible for an average of 700 borrowers.

[27] We also estimate the likelihood for each of the repayment groups separately, using cluster probabilities. Results are reported in Online Appendix Subsection C.1.

[28] We classify wards as rural based on the classification of the National Bureau of Statistics of Tanzania.

[29] For more details on this measure, see Stemmler *et al.* (2024), who use the same data to study labour responses to agricultural shocks in Tanzania.

[30] Note that across the outcomes of interest, the numbers of observations differ. In particular, we have fewer observations for average payment, and again fewer for average balance. This is driven by: (i) cases where during a harvest season a customer made no payment, in which case the total payment amount is zero, but the average payment amount and the average balance are set to missing; and (ii) cases when a customer started his loan cycle during a harvest season and made no further payments except his first payment during that season. In these cases, the balance, and thus the average balance, are set to missing as no prior payments had been made.

[31] The correlation between mobile money use and our asset index is positive and strongly significant at 0.08.

## REFERENCES

Aghabozorgi, S., Shirkhorshidi, A. S. and Wah, T. Y. (2015). Time-series clustering—a decade review. *Information Systems*, **53**, 16–38.

Aker, J. C., Ksoll, C. and Lybbert, T. J. (2012). Can mobile phones improve learning? Evidence from a field experiment in Niger. *American Economic Journal: Applied Economics*, **4**(4), 94–120.

Armendáriz de Aghion, B. and Morduch, J. (2005). *The Economics of Microfinance*. Cambridge, MA: MIT Press.

Balboni, C., Bandiera, O., Burgess, R., Ghatak, M. and Heil, A. (2021). Why do people stay poor? *Quarterly Journal of Economics*, **137**(2), 785–844.

Bandiera, O., Burgess, R., Das, N., Gulesci, S., Rasul, I. and Sulaiman, M. (2017). Labor markets and poverty in village economies. *Quarterly Journal of Economics*, **132**(2), 811–70.

———, ———, Deserranno, E., Morel, R., Rasul, I., Sulaiman, M. and Thiemel, J. (2022). Microfinance and diversification. *Economica*, **89**(S1), S239–S275.

Banerjee, A. (2001). Contracting constraints, credit markets and economic development. SSRN Scholarly Paper; available online at https://ssrn.com/abstract=306990 (accessed 10 August 2025).

———, Duflo, E., Goldberg, N., Karlan, D., Osei, R., Pariente, W., Shapiro, J., Thuysbaert, B. and Udry, C. (2015a). A multifaceted program causes lasting progress for the very poor: evidence from six countries. *Science*, **348**(6236), 1260799.

———, Karlan, D. and Zinman, J. (2015b). Six randomized evaluations of microcredit: introduction and further steps. *American Economic Journal: Applied Economics*, **7**(1), 1–21.

Bank of Tanzania (2022). National Payment Systems Annual Report; available online at https://www.scribd.com/document/721112153/Annual-Reports-BANK-OF-TANZANIA-NATIONAL-PAYMENT-SYSTEMS-ANNUAL-REPORT-2022 (accessed 20 September 2025).

Barboni, G. and Agarwal, P. (2023). How do flexible microfinance contracts improve repayment rates and business outcomes? Experimental evidence from India. SSRN Scholarly Paper; available online at https://ssrn.com/abstract=4358795 (accessed 10 August 2025).

Bari, F., Malik, K., Meki, M. and Quinn, S. (2024). Asset-based microfinance for microenterprises: evidence from Pakistan. *American Economic Review*, **114**(2), 534–74.

Barry, M. S. and Creti, A. (2020). Pay-as-you-go contracts for electricity access: bridging the 'last mile' gap? A case study in Benin. *Energy Economics*, **90**, 104843.

Battaglia, M., Gulesci, S. and Madestam, A. (2024). Repayment flexibility and risk taking: experimental evidence from credit contracts. *Review of Economic Studies*, **91**(5), 2635–75.

Berge, L. I. O., Bjorvatn, K. and Tungodden, B. (2015). Human and financial capital for microenterprise development: evidence from a field and lab experiment. *Management Science*, **61**(4), 707–22.

Brune, L., Giné, X. and Karlan, D. (2022). Give me a pass: flexible credit for entrepreneurs in Colombia. NBER Working Paper no. 30634.

Czura, K., John, A. and Spantig, L. (2020). Flexible contract, flexible morale? Microcredit design and repayment discipline. CESifo Working Paper no. 8322.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, **39**(1), 1–38.

Dunning, C. M., Black, E. C. L. and Allan, R. P. (2016). The onset and cessation of seasonal rainfall over Africa. *Journal of Geophysical Research: Atmospheres*, **121**(19), 11405–24.

Ester, M., Kriegel, H.-P., Sander, J. and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In E. Simoudis (ed.), *KDD-96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. Washington, DC: AAAI Press, pp. 226–31.

Field, E. and Pande, R. (2008). Frequency and default in microfinance: evidence from India. *Journal of the European Economic Association*, **6**(2–3), 501–9.

———, ———, Papp, J. and Park, Y. J. (2012). Repayment flexibility can reduce financial stress: a randomized control trial with microfinance clients in India. *PLoS ONE*, **7**(9), e45679.

———, ———, ——— and Rigol, N. (2013). Does the classic microfinance model discourage entrepreneurship among the poor? Experimental evidence from India. *American Economic Review*, **103**(6), 2196–226.

Financial Sector Deepening Trust (FDST) (2014). GIS map of financial access; available online at https://www.fsdt.or.tz/resource-knowledge-hub/gis-map-of-financial-access (accessed 20 September 2025).

Fiorin, S., Hall, J. and Kanz, M. (2023). How do borrowers respond to a debt moratorium? Experimental evidence from consumer loans in India. IGIER Working Paper no. 691.

Gertler, P., Green, B. and Wolfram, C. (2024). Digital collateral. *Quarterly Journal of Economics*, **139**(3), 1713–66.

Jack, W., Kremer, M., de Laat, J. and Suri, T. (2023). Credit access, selection, and incentives in a market for asset-collateralized loans: evidence from Kenya. *Review of Economic Studies*, **90**(6), 3153–85.

Kafle, K., Jolliffe, D. and Winter-Nelson, A. (2018). Do different types of assets have differential effects on child education? Evidence from Tanzania. *World Development*, **109**, 14–28.

Kumaraswamy, S. K., Mattern, M. and Hernandez, E. (2020). Assets matter to poor people: but what do we know about financing assets? CGAP Working Paper.

Laureti, C., De Janvry, A. and Sadoulet, E. (2017). Flexible microfinance products for financial management by the poor: evidence from SafeSave. ULB Working Paper no. 17-036.

Liao, T. W. (2005). Clustering of time series data—a survey. *Pattern Recognition*, **38**(11), 1857–74.

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In L. M. Le Cam and J. Neyman (eds), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1: Statistics*. Berkeley, CA: University of California Press, pp. 281–97.

NOAA STAR (2018). Global vegetation health products. *Center for Satellite Applications and Research*; available online at https://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/index.php (accessed 20 September 2025).

Pearlman, S. (2010). *Flexibility matters: do more rigid loan contracts reduce demand for microfinance? Unpublished manuscript*, CAF Development Bank of Latin America.

Rahut, D. B., Aryal, J. P. and Marenya, P. (2021). Understanding climate-risk coping strategies among farm households: evidence from five countries in Eastern and Southern Africa. *Science of the Total Environment*, **769**, 145236.

Scrucca, L., Fop, M., Murphy, T. B. and Raftery, A. E. (2016). mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal*, **8**(1), 205–33.

Shoji, M. (2010). Does contingent repayment in microfinance help the poor during natural disasters? *Journal of Development Studies*, **46**(2), 191–210.

Stemmler, H., Lenel, F., Kis-Katos, K. and Weisser, C. (2024). *Solar technology as a shock-coping device: evidence from rural Tanzania*. Mimeo.

Stojanovski, O., Thurber, M. and Wolak, F. (2017). Rural energy access through solar home systems: use patterns and opportunities for improvement. *Energy for Sustainable Development*, **37**, 33–50.

Tadesse, G. and Zewdie, T. (2019). Grants vs. credits for improving the livelihoods of ultra-poor: evidence from Ethiopia. *World Development*, **113**, 320–9.

Tomalka, J., Lange, S., Röhrig, F. and Gornott, C. (2020). *Climate Risk Profile: Tanzania*. Climate Risk Profiles for Sub-Saharan Africa Series. Bonn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

Tucker, C. J., Pinzon, J. E., Brown, M. E., Slayback, D. A., Pak, E. W., Mahoney, R., Vermote, E. F. and El Saleous, N. (2005). An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *International Journal of Remote Sensing*, **26**(20), 4485–98.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, **58**(301), 236–44.

World Bank (2017). *Global Tracking Framework 2017—Progress toward Sustainable Energy*. Washington, DC: International Energy Agency and World Bank.

——— (2022). *Tracking SDG 7: The Energy Progress Report 2022*. Washington, DC: World Bank.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Grohmann, A., Lenel, F. and Herbold, S. (2026). Repayment under flexible loan contracts: evidence based on high-frequency data. *Economica*, **93**(369), 160–187. https://doi.org/10.1111/ecca.70013

## APPENDIX A. BORROWER CHARACTERISTICS

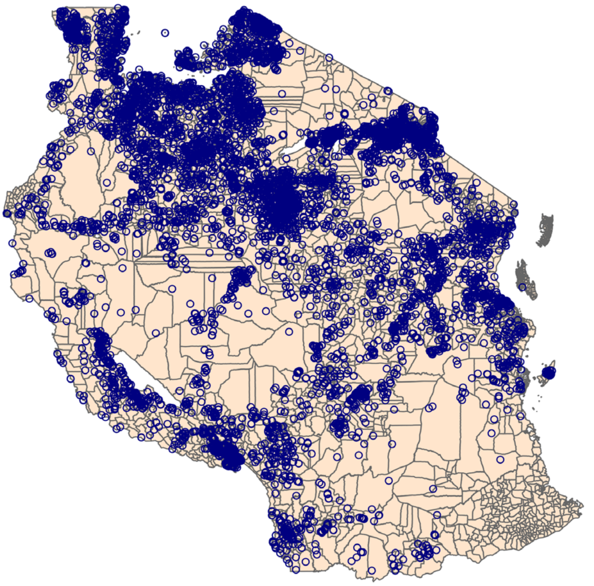


**FIGURE A1** Locations of borrowers across Tanzania. *Notes*: Each circle represents one customer.

**TABLE A1** Sample borrower characteristics.

| Variable | Mean | S.D. | Min | Max | Median | Obs. |
|---|---|---|---|---|---|---|
| Female | 0.19 | 0.39 | 0.00 | 1.00 | 0.00 | 37,571 |
| Household members | 4.48 | 1.94 | 1.00 | 30.00 | 4.00 | 37,438 |
| Age | 39.65 | 10.94 | 15.00 | 89.00 | 38.00 | 18,511 |
| Farmer | 0.46 | 0.50 | 0.00 | 1.00 | 0.00 | 36,202 |
| Wage-employed | 0.24 | 0.42 | 0.00 | 1.00 | 0.00 | 36,202 |
| Self-employed | 0.30 | 0.46 | 0.00 | 1.00 | 0.00 | 36,202 |
| No. of businesses | 1.27 | 0.95 | 0.00 | 5.00 | 1.00 | 19,705 |
| Monthly disposable income (US$) | 33.56 | 25.63 | 0.64 | 340.43 | 27.23 | 19,410 |
| Owns land | 0.59 | 0.49 | 0.00 | 1.00 | 1.00 | 19,705 |
| Trading or renting land | 0.20 | 0.40 | 0.00 | 1.00 | 0.00 | 19,705 |
| Employs domestic workers | 0.27 | 0.45 | 0.00 | 1.00 | 0.00 | 19,705 |
| Owns animals that can be sold in emergency | 0.18 | 0.39 | 0.00 | 1.00 | 0.00 | 19,705 |
| Roof made out of tiles | 0.04 | 0.20 | 0.00 | 1.00 | 0.00 | 37,607 |
| Wealth index | 1.83 | 0.97 | 0.00 | 4.00 | 2.00 | 19,356 |
| Member of savings club | 0.15 | 0.36 | 0.00 | 1.00 | 0.00 | 19,705 |
| Credit experience | 0.18 | 0.39 | 0.00 | 1.00 | 0.00 | 19,705 |
| Distance to next town (km) | 17.00 | 18.39 | 0.00 | 161.00 | 10.38 | 37,674 |
| Commercial bank within 10 km | 0.47 | 0.50 | 0.00 | 1.00 | 0.00 | 37,674 |
| Community bank or MFI within 10 km | 0.42 | 0.49 | 0.00 | 1.00 | 0.00 | 37,674 |
| Savings club within 10 km | 0.59 | 0.49 | 0.00 | 1.00 | 1.00 | 37,674 |
| Mobile money agent within 10 km | 0.78 | 0.41 | 0.00 | 1.00 | 1.00 | 37,674 |
| System loan amount (US$100) | 9.54 | 3.13 | 2.45 | 25.11 | 8.44 | 38,407 |
| Business purpose | 0.07 | 0.25 | 0.00 | 1.00 | 0.00 | 19,525 |

*Notes*: Characteristics of the sample of borrowers who are included in the repayment analysis. Financial access information based on Financial Sector Deepening Trust (2014).