Protecting the Blind Side: Can loan characteristics cushion the impact of an exogenous shock on Micro Credit repayments?

Ashish Desai, Rudra Sensarma and Ashok Thomas

Indian Institute of Management Kozhikode, IIMK Campus, Kozhikode - 673 570, Kerala, India.

Abstract

Studies on microfinance credit repayment behaviour have investigated various reasons for the limited ability to repay ranging from borrower profile to loan product types. However, sparse attention has been given to the repayment ability of borrowers in the face of covariate shocks. This study combines big data (Google Mobility Data and Night Light Data) and government administrative data (Covid-19 infection rate) to proxy for the pandemic caused economic shock. We analyse the interplay of covariate shocks with product characteristics to analyse repayment behaviour of micro-borrowers. Using an original monthly panel data of 2345 women entrepreneurs in a rural region of India, our results from hybrid panel regressions show that loan product characteristics mitigate the impact of an exogenous shock on repayment behaviour. Our findings can serve as valuable inputs for building risk-based pricing models for poor customers and may help in formulation credit policies and credit assessment tools for the micro borrower segment.

Key Words

Financial Literacy, Loan Repayment, Hybrid Regression, Women Empowerment, Covariate Shock.

1. Introduction

Access to credit is a challenge for women entrepreneurial activities in the developing world (Brixova & Kiyotaki, 2016). While microfinance institutions have been instrumental in enhancing the resources of women micro entrepreneurs by providing them access to credit through product innovation (Cull and Murdoch 2017, Armendáriz & Morduch, 2000), the borrowers' limited repayment ability remains a challenge. Several studies have tried to identify the reasons for loan defaults by focusing on the characteristics of the borrowers such as age, gender, education, and occupation (Bhatt & Tang, 2002; Nawai & Shariff, 2012), while others have focused on loan types and sizes (Gebeyehu, *et al*., 2013; Ojiako & Ogbukwa, 2012). However, the impact of a covariate shock on the repayment behaviour of micro borrowers has been scarcely studied.

We exploit a large exogenous shock in the form of the Covid-19 pandemic to study the moderating role of loan product characteristics on the effect of economic shocks on loan repayment. This is the first study to combine big data (Google Mobility Data and Night Light Data) and government administrative data (Covid-19 infection rate) to proxy for the pandemic caused economic shock and study the role of loan product types in moderating the effect of a covariate shock on repayments. We employ an original monthly panel data of 2345 women entrepreneurs in a rural region of India that provides us a unique setting of micro borrowers who faced the same exogenous shock but vary by their socio-economic characteristics and loan types.

In the late 1990s and mid 2000s there were some scepticisms expressed on the ability of Microfinance Institutions (MFIs) to balance between outreach and sustainability (Hermes & Hudon, 2018). However, empirical evidence has shown that MFIs can balance both outreach and profitability (Cull *et al.*, 2007, Conning & Morduch, 2011). Healthy Return on Asset (ROA), along with burgeoning consumer demand resulted in an unprecedented portfolio growth (CAGR of 25%)[1] of Indian MFIs over the past two decades. However, the growth of the sector was impeded by the Covid-19 pandemic. Pandemic not only had mortality and health impact, but also serious economic impacts due to government-imposed lockdowns. Lockdowns, which were institutionalised to save lives, came at the cost of livelihood (Murdoch, 2020).

Collapse of livelihood during the Covid-19 pandemic raised another important issue in the form of sustenance and performance of MFIs in face of exogenous covariate shocks. Often MFIs' credit scoring is based on the ability of an individual borrower to repay, and it relies on group cohesiveness/ social pressures for supporting willingness to repay. The credit scoring process of MFIs, to a certain degree, emulate the response of individuals to idiosyncratic risk (Kanwar, 2005; Lieli & White, 2010) and relies on social capital for loan repayment. Conventional credit scoring processes of MFIs fall short during an exogenous covariate shock such as Covid- 19 owing to the business income of a large percentage of borrowers getting impacted by the correlated event of natural hazards. Increased Non-Performing Assets (NPAs) due to covariate shock pose an existential threat to MFIs who have relatively small portfolios and are concentrated in terms of their geographical coverage (Murdoch, 2020).

When there are exogenous shocks (floods, drought, earthquake etc) which impact the MFI industry, institutions need to find mechanisms to mitigate the consequences (Czura, 2015; Fischer, 2013; Barboni, 2017). In the literature, there have been considerable efforts to detect the reasons for loan defaults and studies have focused on the attributes of borrowers such age, gender, education, and occupation (Bhatt & Tang, 2002; Nawai & Shariff, 2012, Rosenberg, 2010). Another set of studies have focused on type of loans and their characteristics such as their interest rate, group vs individual (Gebeyehu, *et.al*, 2013; Ojiako & Ogbukwa, 2012;

[1] Sa- Dhan's Bharat Microfinance report 2020, available at: http://www.sa-dhan.net/bharat-microfinance-report/

Tedeschi, 2008). There is also some discussion in the literature that focuses on the mitigating effects of micro credits in smoothening consumption in the presence of household shocks (Islam & Maitra, 2009; Isto *et al*.,17) or idiosyncratic shocks. Sparse attention has been given to the fact that while using the loan to fund their livelihoods, borrowers may lose income generating power of the loan due to sustained covariate shocks.

During the Covid-19 pandemic that acted as a covariate shock, the number and amount of new loan disbursements by MFIs dropped and Portfolio at Risk (PAR) increased. To understand the intensity of the covariate shock, we include data from the Google Community Mobility Report (GM) (Google; 2020)[2] which reflects the effects of human activities due to the exogenous shock caused by the pandemic and Night Light Data (NTL) derived from Google earth. Satellite imagery is increasingly being used to estimate and, in some cases, measure many characteristics of the surface of the earth that are associated with human impacts and or activity (de Sherbinin *et al*., 2002). Several studies have reported high correlation of NTL data with economic activity (Donaldson & Storeygard, 2016). In this paper, we use big data from Google Mobility survey, along with profile and transactional data of all customers from a rural branch of a women's cooperative bank. We also complement the findings by using NTL data derived from Google Earth Engine and Covid-19 infection rate as an alternative proxy for the covariate shock.

The paper has some unique contribution to the literature. Firstly, the paper uses an original panel data set of rural women entrepreneurs which limits the problems of survey bias. Secondly, the paper makes use of big data, such as NTL data and Google Mobility data, to assess impact of covariate shock. Thirdly, the paper compares results across four surrogates for

[2] The data are currently available from the URL: https://www.google.com/covid19/mobility/. Google states that the reports will be available publicly "for a limited time". Google has introduced some noise into the data to help prevent the identification of individuals. They claim that this does not reduce the usefulness of the data in research, but it will naturally introduce some measurement error

exogenous shock – i.e., NTL, Google Mobility – Residential Mobility, Google Mobility – NTL Mobility and Covid-19 infection rate. Fourthly, the paper uses hybrid regression model which overcomes the limitations of fixed effects and random effects regression models. In essence, we use an amalgamation of big data (Google Mobility Data and NTL Data to capture lockdown effects during Covid-19), government administrative data (infection rate of Covid-19), a customer-level monthly panel data containing borrower profile as well as bank transactional data to analyse the impact of exogenous covariate economic shocks on loan repayment behaviour of women micro entrepreneurs.

Applying hybrid panel regression to the dataset of 2345 borrowers enables us to investigate if the product type which includes business attributes of small women entrepreneurs amplify or moderate the impact of an exogenous shock on borrower loan repayment behaviour. First, we find that an exogenous shock in Covid-19 pandemic reduces loan repayments. Repayment behaviour is better in the case of younger borrowers, borrowers with higher financial literacy and higher usage of digital transactions. Next, we find that the type of loan product (whether it is an individual loan, group loan or an overdraft) plays an important role in repayment in face of exogenous shocks. For instance, with rising extent of the exogenous shock, repayment performance for individual loans is inferior as compared to joint liability group (JLG) loans, while that of Cash Credit (overdraft) is not different, which attests to the importance of product design. During the covariate shock, location of shop in a marketplace is associated with a better propensity to repay loans. Similar results are obtained from a battery of robustness checks using different measures of the exogenous shock i.e., mobility data, NTL data and Covid-19 infection rate.

Our results provide useful insights on the financial resilience and vulnerability of micro credit customers who are poor women entrepreneurs. Incorporation of these insights in credit evaluation and product design can enable MFI's[3] to build healthier portfolios. Our findings can serve as a valuable input in building risk-based pricing model for Base of Pyramid (BOP) customers and can help in formulation credit policy and credit assessment tool. The findings can help MFIs to innovate and structure innovative non joint liability group (JLG) products. Additionally, regulators and policy makers can also use these findings to explore interventions that may be required to build resilience in the borrowers. Our study points at additional avenues for research using big data as a lead indicator for economic shock and its implications.

The paper has the following structure. Section 2 discusses the background literature and Section 3 presents the research questions. Next, Section 4 describes the Data and Descriptive Statistics. Section 5 presents the Estimation Methodology and Section 6 contains the Results and Discussion. Section 7 provides robustness check by using NTL Data, Work Mobility and Covid-19 Infections as Proxy for Economic shock and the paper concludes with implication and policy recommendation in Section 8.

2. Background literature

2.1 Women Rural Entrepreneurs and Access to Credit

Half of the world's population and one-third of the global manufacturing workforce are women (De Groot, 2001). There have been significant contributions by women to economic growth over the past century (Ghani *et al.*, 2012). Economic participation not only leads to women empowerment, but also benefits family, community, and overall state of economy. A strong correlation exists between the empowerment of women and the development of the economy (Duflo, 2012). Empowering half of the potential workforce has yielded significant

[3] MFI in the paper represents all financial institutions; irrespective of regulatory license lending to base of pyramid borrowers.

economic benefits which go beyond promoting gender equality (Duflo, 2005; World Bank, 2012).

A key difference between developed and emerging economies is that the former has a higher share of non-agricultural sector than the former, where the government or formal institutions create jobs. On the other hand, in emerging markets, including India it is left to the populace to create jobs for itself using its own ability (Sethuraman, 1998; Rustagi, 2011). India has about 63 million MSMEs which contribute to nearly 6% of GDP from manufacturing activities and 24% of GDP from services.[4] Out of this, about 8 million MSMEs are run by women of which 98% are in the micro category. The existence of a widespread gender gap among male and female entrepreneurs is well-known (Global Entrepreneurship Monitor, 2018). With Sustainable development goals (SDG 5) focusing to achieve gender equality and empower women, focus has been on exploring women owned MSME (Sarmah *et al.*, 2022)

An overwhelming majority of women entrepreneurs remain in the informal sector, are unregistered, lack comprehensive formal accounting, income, and business transaction documentation. Formal financial institutions are ill equipped to assess creditworthiness to MSMEs (Biswas, 2014). Information asymmetry, lack of credit history and inadequate collateral make it difficult for all entrepreneurs, especially women entrepreneurs, to access loans in developing countries (Panda, 2012; Sandhu *et al*., 2012; Thampy, 2010).

2.2 Micro Finance in India

While Micro finance Institutions (MFIs) are regulated entities classified as Non-Banking Finance Company (NBFC) and supervised by Reserve Bank of India, in this paper MFI lending refers to MFI- NBFC as well as Cooperative banks or Small Finance Banks who

[4] MSME at a glance report by the Ministry of Micro, Small & Medium Enterprises (MSME), Governemnt of India, available at: https://msme.gov.in/knowledge/msme-glance-english

are in the business of lending to micro and nano women entrepreneurs. MFI lending prioritises outreach based on requirement of small ticket loans; however, there is empirical evidence that MFIs can achieve outreach and remain profitable (Hermes & Hudon, 2018; Cull *et al.*, 2007).

The MFI industry has built itself on the edifice of Joint Liability Group (JLG) lending by providing loans to predominately women entrepreneurs. These loans are considered operationally intensive and riskier than loans to wealthier clients who have a regular income (Assefa *et al.*, 2013). JLG loans are characterized by group guarantees by coborrowers, reference checks and character checks. Significant efforts are invested in gathering information and monitoring loans (Morduch & Aghion, 2000). The loan has weekly or monthly collection, accompanied by group meetings to keep the group cohesion and camaraderie (Morduch, 1999). Businesses need flexible loans, for trade, which are seldom provided by formal financial institutions. Even JLG loans seem to be too rigid and inflexible for the customers (Karlan & Mullainathan, 2007). When there is an event causing exogenous shock or business face unpredictable cash flows, cash strapped micro credit customers are not able to service the loan.

In India, MFIs primarily provide group lending only; while most individual loans are provided to customers with a history of lending (Morduch, 2000; Aragon *et al.,* 2020, Bansal, 2006; Paal & Wisemann, 2006,). Certain lenders have introduced Overdraft - Cash Credit, as well as individual term loan products and it has shown early evidence of improving business performance of nano entrepreneurs (Aaragon *et al.*, 2020). Credit products for micro entrepreneurs have typically been structured from a supply side point of view, not with a lens of what is needed from the demand side. Under the veil of higher risk and operational cost, structured financial products for micro entrepreneurs are deemed mostly unaffordable (Choudhury *et al*.,2022; Kodongo & Kendi 2013).

2.3 Shocks

Low-income households in developing countries are vulnerable to the effects of covariate shocks as well as idiosyncratic shocks. Covariate shocks are events that occur in the area where the household or borrower is located, such as natural disasters, thereby impacting all borrowers in that area. On the other hand, shocks that are limited to households, such as health emergencies, are termed idiosyncratic shocks. Idiosyncratic risk could be unique to an individual asset while covariate risk impacts a large cohort of people, more often than not in the same locality.

In emerging economies there are limited public safety nets, and hence negative income shocks might have long term consequences if the group or individual at risk does not have the ability to absorb the shocks. Economic shocks impact the way people live, work and operate which gets manifested in their financial behaviour. Previous economic shocks, such as dot-com burst or 2008 crisis, affected a specific group of workers, occupations, or industries, but the Covid-19 pandemic was a widespread exogenous shock, which impacted the whole world simultaneously. Covid-19 induced mobility lockdown resulted in temporal changes in certain occupations and value propositions (Kramer and Kramer, 2020).

The literature highlights various mechanisms that households use to smoothen their consumption in the face of shocks, such as by selling assets and livestock, increasing labour supply, cutting back on non-food expenditures, using savings and credit and most importantly relying on informal of risk share between friends and family (Dercon, 2002; Morduch, 1995). However, during a covariate shock, many of these mechanisms might fail. Using wider networks of family and friends, informal risk sharing networks have been highlighted as a key mechanism through which households smoothen shocks, (Townsend, 1994; Udry, 1994; Chiappori *et al*., 2014; Kazianga & Udry, 2006; Rosenzweig, 1988; Fafchamps & Lund, 2003; Jack & Suri, 2014; Blumenstock *et al*., 2016). During a covariate shock such as the Covid-19 crisis, households have lesser ability to rely on social networks as everyone was affected by the shock (Yang, 2008; Yang & Choi, 2007).

MSMEs tends to be sensitive to economic shocks impacting their ability to service loan repayments (Afrah *et al*., 2021). The Covid-19 crisis brought to light the failure of typical mechanisms used by entrepreneurs for coping with a covariate shock such as by cutting expenditure or selling assets, thereby limiting an entrepreneur's ability to repay debt, resulting in portfolio default (Dercon, 2002; Morduch, 1995; Mahmud & Reily 2021).

2.4 MFI Loan product and credit assessment

MFIs generate their revenue in the form of interest charged on the credit extended to low-income individuals. The loan repayments may be relatively uncertain and necessitates a product structure that incorporates credit evaluation policy to govern the MFI's loan management operations (CBK, 2015, Wara *et al*., 2018). Product structures have written guidelines that set the terms and conditions for customer qualification criteria, procedure for making collections and steps to be taken in case of customer delinquency (Mosin, 2009). It helps MFIs control exposure, default risk and ensure no breach of regulatory requirements (Acharya *et al*., 2013). Lack of clear product level credit policies can lead to default of repayment on microfinance loans. The important components of MFI product structure are eligibility criteria of customers including customer's business and individual profiles; business prerequisites – size, type of business, location of business, average bank balance; relationship with bank – type and number of other banking relationship the entity might have (especially liability). The above parameters are included in a credit scoring formula. Typically, each customer is scored and loan product eligibility and pricing (interest and fee) are linked to the score (Owusu *et al*., 2015; Kimondo, 2013).

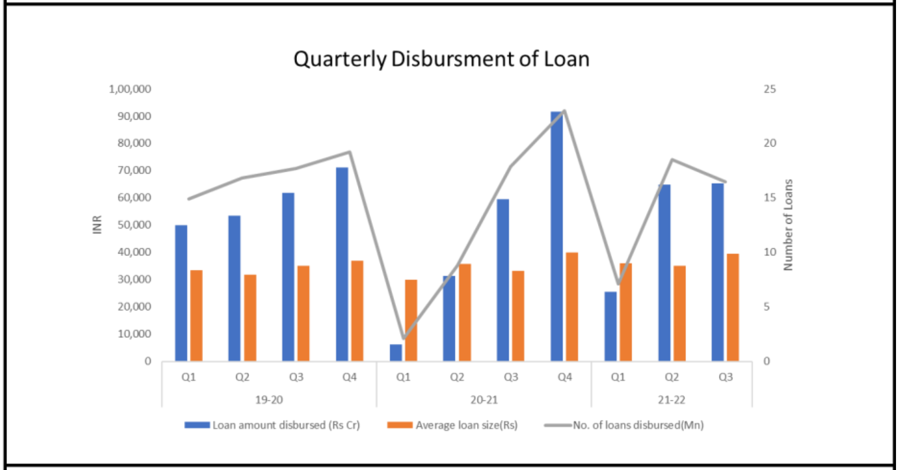
2.5 Performance of MFI Industry During Covid-19 Shock

The Microfinance sector deals with the most vulnerable segments of society at an unprecedented scale[5] and pandemic has been one of the most challenging exogenous shocks for the industry. Management of assets being the key in Microfinance business, net profit (post risk-based provisioning) is affected by repayment of loans. A non-repayment of even 5% of the customer base has a significant impact on the return on assets (ROA).

During Covid -19 pandemic, new disbursements of MFI loans dropped and Portfolio at Risk (PAR) increased (figures 1 and 2). Health of MFI portfolio can be gauged by PAR, which is bucketed into risks at 30, 60 and 90 days. PAR also has impact on liquidity management and Credit rating, which in turn impacts cost and availability of capital including both debt and equity.

Figure 1: Quarterly Disbursement of Loan

Figure 2 : Quarterly Disbursement of Loan

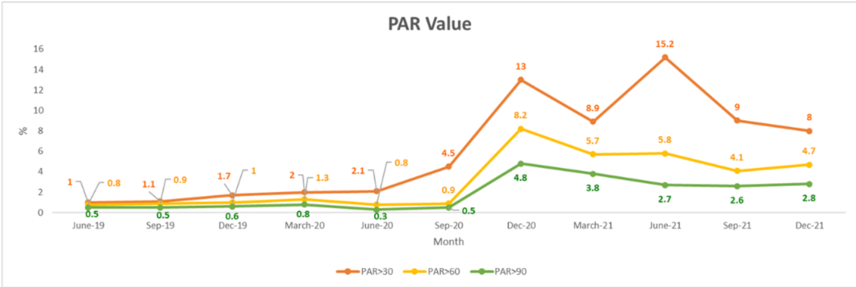


Source : MFIN

[5] 58 million households – Source Sa-dhan report 2021

Figure 2: Quarterly Disbursement of Loan

Figure 3 : PAR Value



Source : MFIN

3. Research Question

Microfinance Credit assessment typically analyses the borrower's ability to repay from their cash flows, and there is a paucity of analysis on the factors behind willingness to repay in the context of customers with limited credit history (Eze & Ibekwe, 2007; Papias & Ganesan 2009, Bhatt & Tang 2002; Nannyonga, 2000). Apart from interest rates. Microfinance product is defined by disbursement structure, frequency, and method of repayment, which in turn plays a role in repayment behaviour (Nannyonga, 2000; Jain & Mansuri, 2003; Sharma & Zeller1997; Derban *et al.*, 2005).

In the literature, there have been considerable efforts to identify the reasons for loan defaults based on borrower characteristics on the demand side and loan types and sizes on the supply side. However, existing studies have ignored the mitigating effects of microcredit to sustain financial behaviour in the face of covariate shocks.

To fill the gap in the existing literature we address the following research questions:

- What is the effect of borrower and loan attributes on repayment behaviour?

- In the face of an exogenous covariate shock, do different loan products exhibit different repayment behaviour?

4. Data and Descriptive Statistics

We include the data from the Google Community Mobility Report popularly known as Google Mobility (GM) (Google, 2020)[6]. which captures the effects on human activities due to the covariate shock of Covid-19 (Gatalo *et al*., 2020). Google collects and stores individuals' commuting information through a GPS linked to Google maps. During the pandemic, GM was released in the public domain by Google since 2020 and is based on tracking data that measures clustering of individuals. GM has been made available online for India at the state, union territory and up to district levels since February 15, 2020.

While other mobility indices use the number of vehicles or the number of trains, GM shows the actual movement and activity of people (Badr *et al*., 2020). Thus, it reflects changes in social and economic behaviour as well as impact of policies and regulations such as lockdowns (Ilin *et al*., 2021; Luther, 2020). GM reports aggregated data at different categories of activities, e.g., workplaces, residential, retail, as compared with a baseline period before the epidemic. For example, data for a particular Monday are compared to corresponding data from the baseline series for a Monday. The values represent the relative changes in percentage compared to the baseline days, not the absolute number of visitors. For instance, a value of -50 in the workplaces data set on a Monday, indicates a 50% drop compared to the Monday in the reference period. Similarly, a positive value indicates an increase in mobility compared to the reference period. GM provides the percentage changes in activity for 6 key activity categories (groceries and pharmacies, parks, transit, retail, and recreation, residential, and workplaces)

[6] The data are currently available from the URL: https://www.google.com/covid19/mobility/ . Google states that the reports will be available publicly "for a limited time". Google has introduced some noise into the data to help prevent the identification of individuals. They claim that this does not reduce the usefulness of the data in research, but it will naturally introduce some measurement error

compared to the baseline days before the advent of COVID-19 (5 weeks, from January 3 to February 6, 2020) (Google, 2020; Chan *et al.*, 2020; Buckee *et al.*, 2020).

The Google Residential Mobility (RM) shows a change in time spent at residential locations, while other categories measure a change in total visitors. To analyse impact of shock, we adopt Google's Residential Mobility (RM) data. Higher the RM value, lower is the economic activity as citizens are confined to their residence due to rising infection rates and lockdowns. We also perform analysis using Workplace Mobility (WM) dimension in GM and Covid-19 infection and per capita covid infection rate (CIR); higher the WM lower is the impact of lockdown and higher the CIR, higher is the impact of lockdown. While CIR, represents health shock, RM and WM represents mobility variations as a proxy for the intensity of the economic shock associated with lockdown. We also complement our analysis with use of NTL data (details in section 7).

To capture the effect on micro borrowers, we collect data from a rural region of India. Our sample comes from Mhaswad branch located in Mann Taluka[7] ;Satara district of the Western state of Maharashtra, where we collaborated with a rural cooperative bank viz. Mann Deshi Mahila Sahkari Bank, as a partner in collection of data. Mann Deshi Bank by charter of their organisation and their Memorandum and Articles of Association can only lend to women. As per the bank's credit policy, the women need to be micro entrepreneurs to avail of a loan. While, all the customers had declared their occupation as micro entrepreneurs, it was also duly verified by the credit approval and disbursement processes of the bank.

Data of all customers of Mann Deshi Bank's Mhaswad branch was chosen for the analysis. This gave us a sample of 2345 women micro entrepreneurs for a period 25 month

[7] Mann Taluka is one of 11 Talukas (a lower administrative unit) within the Satara district. Mann Taluka represents 104 villages and the town of Mhaswad with a total population of 225,634. The literacy rate in Man Taluka is 64% which is much lower than the averages within Satara, 82.87%, and Maharashtra, 82.34%.

June 2019 to June 2021. Mhaswad branch was the oldest branch of Mann Deshi and has been in existence for 25 years. During the Covid-19 pandemic, Satara was one of the highly infected districts. It was one of the 9 districts in Maharashtra that collectively, contributed to around 50% of the entire state of Maharashtra's fresh COVID-19 cases.

We use the customer profile data and transactional data of the 2345 women entrepreneurs over a period of 25 months that covers pre-Covid, Covid and post-lockdown periods (June 2019-June 2021). The average age of the entrepreneurs was 38 years and they were predominantly running nano enterprises – i.e. home based or cottage industry or very small retail stores. All the borrower level variables such as loan repayment, digital transaction, monthly balance, loan value etc. were retrieved from the core banking system (CBS) of the bank, except the customer's financial literacy score. It should be noted that though RBI announced a moratorium permitted to extend the moratorium on payment of all instalments in respect of loan, many MFIs did not stop collection effort and protocol. Mann Deshi bank was one of such banks.

We hand collected financial literacy scores using a questionnaire-based survey conducted at the branch[8] . The list of variables is shown and explained in Table 1.

Table 1: Variable – Source and Description

Variable	Source	Nature	Description
			Dependent Variable
Loan Repayment	CBS	Time Variant	Has the borrower repaid monthly EMI on due date. It is a binary value – 0 for not repayment and 1 for repayment. It should be noted, Loan repayment is through debit order set at disbursement or was actionized through manual collection as in case of Cash Credit product.
			Independent Variable
Covid Infection Rate	Covidindia.org	Time Variant	Number of positive cases of Covid-19 detected in the month in a district divided by population of the district.

[8] Refer annexure I

Residential Mobility	COVID-19 Community Mobility Reports (google.com)	Time Variant	Residential Mobility, higher the RM value, lower is the economic activity as citizens are confined to their Residential due to rising infection rate and lockdown.
Workplace Mobility	COVID-19 Community Mobility Reports (google.com)	Time Variant	Workplace Mobility, higher the WM value, higher is the economic activity as citizens are mobile and visit workplace.
Nightlight Data	Google Earth engine	Time Variant	Nightlight data reflets intensity of nightlight in a particular geography, higher the NTL value, higher is the economic activity. We use change in night light for the same month over previous year to measure the impact.
Age	CBS	Time Invariant	Age of the respondent that is recorded in the Core Banking System at the time of loan disbursement
Digital Transaction	CBS	Time Variant	Number of transactions using UPI (Unified Payments Interface – the primary digital financial transactions platform in India), done by the customers in a month. It includes both send and receive transactions.
Average Monthly Balance	CBS	Time Variant	Monthly Average balance in the savings account of Mann Deshi Bank.
Holder of Pygmy deposit	CBS	Time Invariant	Bank offers a product of daily deposit which is collected from depositors of financial services agent. If the respondent has subscribed to this product, then status is 1 else 0.
Financial Literacy Score	Survey	Time Invariant	"Ability to understand, analyse, manage and inform about the financial conditions that affect material well-being of an individual". Financial Literacy Score[9] is questions the respondent can answer correctly, out of the 5 questions10.
Loan Value sanctioned	CBS	Time Invariant	Value of the loan approved (in case of Cash Credit) / disbursed by the bank for the individual borrower. The loan approval value is based on various credit evaluation parameters.
Location of Shop – In Market	CBS	Time Invariant	Is the establishment of the borrower located in main market of the village which is the hub of commercial activity of the village/ town. Typically, Market place is a hub of all economic activity in a rural area, and possibly having an establishment there might imply that economic recovery of is faster.
Product	CBS	Time Invariant	Is the product Cash Credit (Overdraft), JLG (Joint Liability Group lending) or Unnati (Individual Loan)[11]

[9] Scale for financial literacy is adopted from Ćumurović; Hyll, Walter "Financial Literacy and Self-employment", *IWH – Member of the Leibniz Association*, 2016

10 Refer Annexure ii

[11] **Joint Liability Loan (JLG) .** No collateral, unsecured loan . Women form groups of 8-10 and co guarantee each other. It is an individual loan but with group guarantee. The loan is a term loan with a fixed value EMI (equated monthly instalment), with reducing balance interest calculation. There is no penalty on prepayment.

Descriptive summary of the variables is presented in Table 2. It can be observed that standard deviation of loan value sanctioned is large, due to which the values of the variable are converted to log scale.

Table 2: Summary of observed values of variables

Variable	Symbol	Nature	Observations	Mean	Std. Dev.	Min	Max
				Dependent Variable			
Loan Repayment	LR	Dependent Variable	58,625	0.6288	0.4831	0.00	1.0
				Independent Variable			
Covid Infection Rate	CIR	Exogenous Shock	58,625	0.0024	0.0047	0.00	0.020
Residential Mobility	RM	Exogenous Shock	58,625	13.218	11.3156	0.00	36.84
Workplace Mobility	WM	Exogenous Shock	58,625	13.4503	-9.5944	-51.3	3.87
Nightlight Data	NTL	Exogenous Shock	58,625	-0.0126	0.1254	-0.3048	0.20
Age	AGE	Borrower Profile	58,625	38.77	9.56	18.0	73.0
Digital Transaction	UPI	Borrower Profile	58,625	34.37	11.88	6	49
Log_Average Monthly Balance	LAB	Borrower Profile	58,625	0.3796	7.438	-1.0	6.8
Holder of Pygmy deposit	PD	Borrower Profile	58,625	0.504	0.499	2.0	4.0
Financial Literacy Score	FLS	Borrower Profile	58,625	3.181	1.276	2.9	4.6
Log_Loan Value sanctioned	LLV	Borrower Profile	58,625	0.3749	9.210	9.21	12.89
Location of Shop – In Market	LS	Borrower Profile	58,625	0.4289	0.4949	0	1
Product	PDT	Product Type	58,625	2.822	0.5582	1	3

Cash Credit Loan This product is like cash Credit and overdraft. Women vendors across dozens of village weekly markets currently benefit from this cash flow facility to build their working capital. No collateral, unsecured loan. It is an individual loan with no guarantee. This product has flexible repayment schedule , depending on the cashflow on the borrower. The Women can draw down and repay the value multiple times within the approved limit. It is a classical cash Credit/ overdraft product in banking terminology
Unnati Individual Loan - No collateral, unsecured loan . It is individual loan, typically higher ticket size for Women who has displayed appropriate and satisfactory Credit behaviour. There are no group guarantees. The loan is a term loan with a fixed value EMI (equated monthly instalment), with reducing balance interest calculation. There is no penalty on prepayment.

5. Estimation Methodology

The objective of our empirical exercise is to estimate the determinants of loan repayment behaviour. Since our data is a monthly panel of 2345 individuals, we use panel data regression methods for estimation. One of the key advantages of panel data regression is that it can address issues posed by omitted variable bias, thereby controlling for unobservable factors, making it better suited for estimation of linear model as compared to cross sectional Ordinary Least Squares (OLS) regression analysis (Allison, 2009).

A 'between-effect' (BE) regression model uses within-group variations and exploits the differences in the cross-sectional dimension of the data. However, it loses critical information which may be present in the time dimension. Essentially OLS estimation using time averages of the cross-sectional data is a BE model (Wooldridge, 2010). Alternatively, fixed effects (FE) regression produces what is called as 'within-effect' i.e. the response of the dependent variable to changes in the independent variables across the time dimension. The third approach in panel data regression is the random effects (RE) model. RE parameters are weighted averages of 'between-effect' and 'within-effect' estimators. RE model estimates are more efficient compared to FE, provided that the unobserved heterogeneity is not correlated with the explanatory variables and individual effects are randomly distributed. However, a limitation of this model is that we cannot be certain whether its assumptions hold good.

Typically, objective criteria such as the Hausman test (Hausman, 1978) are often used to guide the choice between FE or RE models, the FE model is typically favoured in empirical research due to the assumptions imposed in the RE model. FE model technique provides an unbiased estimate, making them a standard default in panel data modelling (Bell & Jones K, 2015). While FE model can generate consistent estimates (Allison, 2009; Wooldridge, 2010), its limitation is that the effect of time invariant variables cannot be estimated.

The econometric model for our analysis can be specified as:

$Y_{it} = \beta_0 + \beta_1 * X_{it} + \beta_2 * Z_i + \alpha_i + \varepsilon_{it}$……. (1)

Where, Y_{it} is the dependent variable (loan repayment), X_{it} are the time dependent explanatory factors, Z_i are the time invariant explanatory factors, i subscript indexes for cross-sectional units and t is the subscript for time. α_i represents the intercept for cross-sectional units and εit is the random error term. From table 1 we note that only three out of the ten variables are time variant, while the rest are time invariant ruling out the possibility of estimating an FE model. To address this empirical challenge, we employ a recently developed alternative, viz. the hybrid panel regression model (Allison, 2009). Following Mundlak (1978), the 'hybrid' model distinguishes between a level 1 variable that varies between and within clusters (X_{it}), and a level 2 variable that varies only between clusters (Z_i). The level 1 variable is decomposed into a between-effect component ($\overline{x_\iota} = n_i^{-1} \Sigma_{t=1}^{n_i} X_{it}$) and a within-effect component ($X_{it} - \overline{x_\iota}$). Subscript i denotes level 2 (e.g. age) and subscript t denotes level 1 (e.g. Infection Rate). α_i is the level 2 random intercept and ε_{it} is the level 1 error term.

The hybrid panel model can then be rewritten as (Allison, 2009; Shuncks, 2013):

$Y_{it} = \beta_0 + \beta_1 * (X_{it} - \overline{x_\iota}) + \beta_2 * Z_i + \beta_3 * \overline{x_\iota} + \alpha_i + \varepsilon_{it}$ ……. (2)

Where, β_1 is the within-effect estimator, i.e. the fixed-effects estimate and β_3 is the estimator for the between-effect (Mundlak, 1978; Neuhaus & Kalbfleisch, 1998). While it is not necessary to include the cluster mean ($\overline{x_\iota}$) to obtain the within estimate of β_1, its inclusion ensures that the estimates of level 2 parameters are corrected for between-cluster differences in X_{it}. Also, because equation (2) is a random-effects model, we can use it to estimate effects of level 2 variables including time invariant variable Z_i. For the estimate of β_2 to be unbiased, $E(u_i| X_{it}, Z_i) = 0$ and $u_i| X_{it}, Z_i \sim N(0, \sigma_\mu^2)$ still have to hold.

In summary, β_1 produces the same results as FE parameters, while β_3 is the between-effect parameter. The hybrid model also allows us to test whether the within and between effects are similar, i.e., whether the impact of change in one variable in a cross-sectional unit has the same effect as that of the change in that variable to a different cross-sectional unit. This test, which is referred to as an augmented regression test (Jones *et al*., 2007), can be used as an alternative to Hausman specification test (Baltagi, 2008). If between and within effects are the same, i.e. $\beta_1 = \beta_3$, then (2) collapses to (1). This method allows estimating the correct FE or within-effect parameter – which is often the main interest of the analysis – but also allows deciphering the effects of non-time varying factors.

With this background, we first estimate the relationship with RE method and then estimate the hybrid regression model. Our estimable model is as follows:

$$Y_{it} = \alpha + \beta_1 * X_{1i} + \beta_2 * X_{2t} + \beta_3 * X_{it} + \beta_4 (X_{1i*} X_{2t}) + \varepsilon_{it} \ldots\ldots (3)$$

Where, Y_{it} *is LR,* X_{1i} *represents time invariant variables (*Age_i *,* PD_i*,* FLS_i*, ,*LS_i*,* PDT_i*),* X_{2t} *represents time variant variable (*CIR_t*,* WM_t *,* RM_t*),* X_{3it} *represents time and individual variant Level 1 variables (*LLV_{it}*,* LAB_{it}*) and* X_{1i}***X_{2t} *are interaction terms (CIR t ** PDT_i*,* CIR_t*** PDT_i*).*

6. Results and Discussion

We report the results of our main regression as specified in equation (3) with RM as the measure of Covid shock capturing the lockdown effect of people spending more time at home than usual. We use panel regression methods starting with RE estimation as out of the ten variables, seven are time invariant, making FE regression inapplicable.

Using RE method we estimate 3 alternative specifications (refer table 3 for the results). The first model – Model 1 does not have any interactions between independent variables, the second model factors in interaction between infection and marketplace presence (impact of location of shop during Covid) and in the third model, we add additional interaction between infection rate and product (to capture impact of product type during Covid). JLG being one of the most widely accessed and popular products, has been consider as a reference category.

At a portfolio level, the results show that the coefficients of average monthly balance and financial literacy are positive and significant ($p<1\%$), while the coefficient of pygmy deposit is positive and statistically significant at 5%. As expected, the exogenous shock (RM) has a negative and significant ($p<1\%$) effect on loan repayment for all the three models. Loan value proxying the income of the individual, is statistically insignificant; implying not the actual loan sanction plays a significant role in managing loan repayment during exogenous shock. During exogeneous shocks, such as pandemic, when there is enforcement of social distancing and access to ATM is restricted, the number of UPI transaction has a positive and significant effect on loan repayment ($p< 1\%$), highlighting the importance of digital financial literacy as an agency, which helps cope with exogenous shocks. Repayment behaviour for all products is similar as there is no statistical difference in repayment behaviour of JLG, Unnati and Cash Credit (in model 2, repayment behaviour of Unnati is poorer than JLG at 10 % level of significance). However, when exogenous shock variable, interacted with product; repayment of Unnati loan is lower than that of JLG Loan ($p<1\%$) but behaviour of Cash Credit is similar to JLG (refer Models 2 and 3, Table 3) . Location of the shop is positive and statistically significant ($p<1\%$) during an exogenous shock, as post lifting of lockdowns, marketplaces opened before other locations resulting in higher income, positively impacting propensity to repay the loan.

Table 3: Results from RE Regression (Exogenous Shock – Residential Mobility)

	No Interaction	Only Market Place Interaction	Market Place and Product Interaction
	(1)	(2)	(3)
Variables			
	Exogenous Shock		
Exogenous Shock	-0.0579*** (0.0028)	-0.0812*** (0.0037)	-.0787*** (0.00381)
	Borrower Profile		
Age	-.0080* (0.0045)	-.0078* (.0045)	-.00773* (.0045)
UPI Transaction	1.2634*** (0.0256)	1.2717*** (0.0261)	1.2727*** (0.2615)
Log Average Monthly Balance	0.8187*** (0.1617)	0.8588*** (0.1643)	0.86491*** (0.1647)
Holder of Pygmy deposit	0.4506** (0.2196)	0.4704** (0.2225)	.45780** (0.2235)
Financial Literacy Score	1.0169*** (0.3760)	1.0285*** (0.0381)	1.028*** (0.0382)
Log Loan Value sanctioned	0.2134 (0.1317)	.2185 (0 .1382)	.2219 (0.1386)
Location of Shop – In Market	0.5457*** (0.2114)	-.4183* (0.2316)	-.3962* (0.2333)
	Product Profile		
Individual Loan – Unnati	-0.378* (0.2194)	-.4008* (0.2212)	.04133 (0.2781)
Cash Credit	-0.6420 (0.5234)	-0.6987 (0.5293)	-.7418 (.7442)
	Interaction Term		
Residential Mobility* Location of Shop		0.0612*** (0.0057)	.0604*** (.0058)
Residential Mobility * Individual Loan			-.0285*** (.0105)
Residential Mobility * Cash Credit			.00269 (.03351)

Robust Standard Error in parenthesis ***P Value < 1% , ** P value <5% , * P value < 10%

Notes: The dependent variable is Loan Repayment . Column 1 has explanatory variable of exogenous shock as Google Mobility. Age , Number of UPI transaction , average monthly savings balance , Location of shop – in market place / out of market place , financial literacy score , loan value sanctioned , is the borrower holder of pygmy deposit (recurring deposit scheme) , product *type -Individual Loan (Unnati) , Cash Credit (Overdraft facility) are the explanatory variables. Column 1 has no interaction , column 2 has Interaction with product type and exogenous shock , column 3 has interaction of product type with exogenous shock and location of establishment with exogenous shock. Estimations are conducted using Random Effect Panel Regression. Clustered standard errors in parentheses. *** p< 0.01, **p<0.05, *p<0.1*

We now re-estimate equation (3) using the hybrid regression model (Allison, 2009) as an alternative to both random and fixed-effects models (Neuhaus & McCulloch, 2006; Schunck, 2013). Hybrid model estimates separate effects for the level 1 and level 2 variables. Variables with the W_prefix denote within-cluster effects, variables with the B_prefix denote between-cluster effects, and variables with the R_prefix are those for which the effects are the same as those in RE model.

Table 4 explores whether the results of between and within cohorts are statistically different which would validate use of a hybrid model. The results show that the within-cluster effects are statistically different from the between-cluster effects, as can be seen from the small p-values in the formal tests of the random-effects assumption of orthogonality between the observables and the unobservable. For instance, in the case of model 1 the test for b_ [B UPI Transaction] = b_[W UPI Transaction] has a p-value of 0.0000 and the test for b[B_ Residential Mobility] = b[W_Residential Mobility] has a p-value of 0.0000. Only, the monthly balance does not show statistically significant difference between the within-cluster and between-cluster effects (p value = 0.6345). As two of the three parameters have small p-value (statistically significant at <1%), the evidence tilts towards favouring a hybrid panel regression model over the standard RE model. Similarly, in model 3 we find statistical significance for four out of six variables. Hence it can be concluded that the hybrid model is a valid model for generating insights over and above the RE model[12].

[12] We conducted Hausman Taylor regression method and the results match with results presented in paper

Table 4: Test for difference in the within and between cohorts (For Table 5)

	P-value (Table 4 model 1)	**P-value (table 4 model 2)**	**P-value (table 4 model 3)**
Residential Mobility	0.0000	0.0000	0.0000
Log Average Monthly Balance	0.6345	0.6771	0.6317
UPI Transaction	0.0000	0.0000	0.0000
W_Residential Mobility *Market Place Presence		0.0000	0.0000
W_Residential Mobility * Individual Loan			0.0002
W_Residential Mobility * Cash Credit			0.4524

So, we proceed with hybrid regression to estimate the three models (see results in table 5). The coefficients with prefix W_ give the within-cluster effects. The coefficients with prefix B_ is between-cluster effects. For UPI, the estimated coefficient indicates that a between-individual monthly increase is associated with a small increase in repayment behaviour, while within-cluster increases in UPI transaction is associated with a within-cluster increase in repayment.

At a portfolio level, propensity of repayment of loan decreases as infection rate increases. Higher financial literacy and digital transactions are associated with better loan repayments. While location of shops itself has no impact on loan repayment; however, during a covariate shock, establishment in marketplace showed better repayment behaviour, indicating the importance of location of shop to manage exogenous shocks. Similarly, while products themselves did not have an impact on loan repayment, however compared to JLG, individual (Unnati) loan showed a poorer repayment behaviour, while JLG and Cash Credit product displayed similar behaviour during the period of exogenous shock. The results also displayed decrease in repayment with increase in age. Value of loan sanctioned, and average monthly balance did not have any statistically significant effect on loan repayment.

Table 5: Results from Hybrid Regression (Exogenous Shock – Residential Mobility)

	No Interaction (1)	**Only Market Place Interaction (2)**	**Market Place and Product Interaction (3)**
Variables			
Exogenous Shock			
W_Exogenous Shock Residential Mobility	-0.0037*** (0.001)	-0.0055*** (0.001)	-0.0054*** (0.001)
Borrower Profile			
R_Age	-0.0002** (0.0001)	-0.0002** (0.0001)	-0.0002** (0.0001)
R_Pygmy deposit	0.0075 (0.0068)	0.0075 (0.0068)	0.0075 (0.0068)
R_Financial Literacy Score	0.0125*** (0.0014)	0.0125*** (0.0014)	0.0125*** (0.0014)
R_Loan Value sanctioned	0.0005 (0.0041)	0.0005 (0.0041)	0.0005 (0.0041)
R_Location of Shop – In Market	0.0009 (0.0065)	0.0009 (0.0065)	0.0009 (0.0065)
W_UPI Transaction	0.0271*** (0.0001)	0.0268*** (0.0001)	0.0268*** (0.0001)
B_UPI Transaction	0.045*** (0.0003)	0.045*** (0.0003)	0.045*** (0.0003)
W_Log Average Monthly Balance	0.0146 (0.0109)	0.0138 (0.0108)	0.0139 (0.0108)
B_Log Average Monthly Balance	0.0087 (0.0058)	0.0087 (0.0058)	0.0087 (0.0058)
Product Profile			
R_Individual Loan – Unnati	-0.0075 (0.0064)	-0.0075 (0.0064)	-0.0075 (0.0064)
R_Cash Credit	-0.0088 (0.0153)	-0.0088 (0.0153)	-0.0088 (0.0153)
Interaction Term			
W_ Residential Mobility* Location of Shop		0.0042*** (0.0002)	0.0042*** (0.0002)
W_ Residential Mobility * Individual Loan Unnati			-0.0020*** (0.0003)
W_ Residential Mobility * Cash Credit			0.0004 (0.0010)

Robust Standard Error in parenthesis *P Value < 1% , ** P value <5% , *** P value < 10%

*Notes: The dependent variable is Loan Repayment. Column 1 has explanatory variable of exogenous shock as Google Mobility. Age, Number of UPI transaction, average monthly savings balance , Location of shop – in market place / out of market place , financial literacy score , loan value sanctioned , is the borrower holder of pygmy deposit (recurring deposit scheme) , product type -Individual Loan (Unnati) , Cash Credit (Overdraft facility) are the explanatory variables. Column 1 has no interaction, column 2 has Interaction with product type and exogenous shock, column 3 has interaction of product type with exogenous shock and location of establishment with exogenous shock. Estimations are conducted using Hybrid Panel Regression. Clustered standard errors in parentheses. *** p< 0.01, **p<0.05, *p<0.1*

Further we unpack the determinants of repayment behaviour for various products individually, (see table 6).

Table 6: Results from Hybrid Regression by Product (Exogenous Shock – Residential Mobility)

	Unnati (1)	Cash Credit (2)	JLG (3)
Variables			
W_ Residential Mobility (Exogenous Shock)	-0.0077 *** (0.0004)	-0.0024* (0.0016)	-0.0053*** (0.0001)
R_Age	-0.0001 (0.0005))	0.013 (0.018)	-0.0003*** (0.0001)
R_Pygmy deposit	-	-	0.0208*** (0.0094)
R_Financial Literacy Score	0.0091** (0.0045)	0.0193 (0.0195)	0.0128*** (0.0014)
R_Loan Value sanctioned	0.0007 (0.0123)	0.0129 (0.0211)	-0.0005 (0.0045)
R_Location of Shop – In Market	0.0032 (0.0159)	0.0733 (0.0494)	0.0064 (0.0043)
W_UPI Transaction	0.0272*** (0.0004)	0.0222*** (0.0012)	0.0268*** (0.0001)
B_UPI Transaction	0.0456 *** (0.0009)	0.0446*** (0.0030)	0.0450 *** (0.0003)
W_Log Average Monthly Balance	-0.0139 (0.0367)	-0.2966*** (0.0974)	0.0198** (0.0114)
B_Log Average Monthly Balance	-0.0077 (0.0200)	-0.0489 (0.0701)	0.0048 (0.0064 |)
W_Residential Mobility* Location of Shop	0.0054*** (0.0006)	0.0007 (0.0019)	0.0041*** (0.0002)

Robust Standard Error in parenthesis *P Value < 1% , ** P value <5% , *** P value < 10%

*Notes: The dependent variable is Loan Repayment. explanatory variable of exogenous shock as Google Mobility. Age , Number of UPI transaction , average monthly savings balance , Location of shop – in market place / out of market place , financial literacy score , loan value sanctioned , is the borrower holder of pygmy deposit (recurring deposit scheme) Interaction between residential mobility and location of shop is also factored. Column 1 is for Individual Loan – (Unnati) ; Column 2 represents Overdraft product - Cash Credit ; Column 3 represents Joint Liability loan - JLG loan. Estimations are conducted using Random Effect Panel Regression. Clustered standard errors in parentheses. *** p< 0.01, **p<0.05, *p<0.1*able 9 : Interpretation of Table 8 - Product Analysis using Hybrid Model

For all the products, propensity of repayment of loan increases as the extent of exogenous shock increases. Unnati , JLG followed by Cash Credit in that order was impacted by exogenous shock. We also observe that higher financial literacy (although statistically insignificant for Cash Credit) and number of digital transactions lead to better loan repayments. For JLGwhile higher the financial literacy, higher is the repayment. Previous studies have also highlighted the role of FL in appropriate debt management and loan repayment (Bahovec *et al*., 2021; Kurowski, 2021; Wanjiku & Muturi , 2015); .We also observe lower the age, higher is the repayment. We attribute the reasons to ; firstly younger generation plausibly are more digital literate and can have higher income from accessing digital marketplaces. Secondly , the health advisory during the period of covariate shock; older generation stayed at home Location of shop was statistically insignificant; however when interacted with exogenous shock, the effect was statistically significant for JLG and Unnati i.e. individual loan (statistically insignificant for Cash Credit).

All borrowers of Unnati loans and Cash Credit product are also owners of Pygmy deposit, hence the variable is missing in analysis for those two products, but it is considered for JLG product and it was statistically significant. Between-effect coefficient of balance was insignificant while within effects was significant for JLG loan (insignificant for Unnati i.e., individual loan). However, for Cash Credit, lower the average monthly balance, better was the repayment. This can be attributed to product structure of Cash Credit which has flexible repayment as well as flexible drawdown facility for the borrower. If the borrower repays some part of the dues monthly, the loan is considered regular.

During Covid – 19, the bank stopped fresh drawdowns and initiated physical collection (compared to digital collection or debit order earlier) to avoid NPA. The physical visit was at least biweekly, which ensured the intensity of collection. In order to avoid going to ATM / bank branch, the borrowers would keep with themselves money in cash to fulfil weekly collection demands, reflecting in the negative coefficient for the product variable. Also we infer that during covariate shock, establishment in market place was associated with better repayment behaviour, again indicating the importance of location of shop to manage exogenous shocks.

7. NTL Data, Work Mobility and Covid-19 Infections as Proxy for Economic shock

The societal, commercial, and private value of satellite imagery (and other remotely sensed data) has been increasing at different rates for different applications. Satellite imagery is commonly used to measure physical attributes of the earth's surface typically associated with weather, land cover, temperature, topography, etc. it is increasingly being used to estimate and, in some cases, measure many characteristics of the surface of the earth that are associated with human impacts and or activity (de Sherbinin *et al*., 2002). Studies have reported high correlation of NTL intensity with economic activity (Donaldson & Storeygard, 2016). We carry out an analysis of the determinants of loan repayment behaviour using NTL data as a variable for exogenous shock.

Firstly, night time satellite imagery is profound in obtaining human activity and presence, and a strong correlation exists between night-time lights and on Gross Domestic Product (GDP) measures at the national, state, regional or even more micro granular levels (Sutton *et al*,. 2007; Proville, 2017). NTL data enables estimating economic growth (Henderson *et al*., 2012). In the Indian context specifically, it is well documented that NTL Provides useful approximation to economic activity (Prakash *et al*., 2019; Beyer *et al*., 2020; Gibson *et al*., 2017; Chanda & Kabiraj, 2020). In fact, the economic impact of demonetisation was also assessed by NTL tracking (Chodorow-Reich *et al*., 2020). Secondly, economic variables such as GDP or electricity consumption, are not published at a district level in India. NTL data's availability at high frequency (monthly) and at district level, allows us to use this variable as a proxy for economic impact before, during and after various waves of Covid-19. Thirdly, NTL is an objective measure and is not influenced by biases which can creep in from survey data (Beyer, 2020).

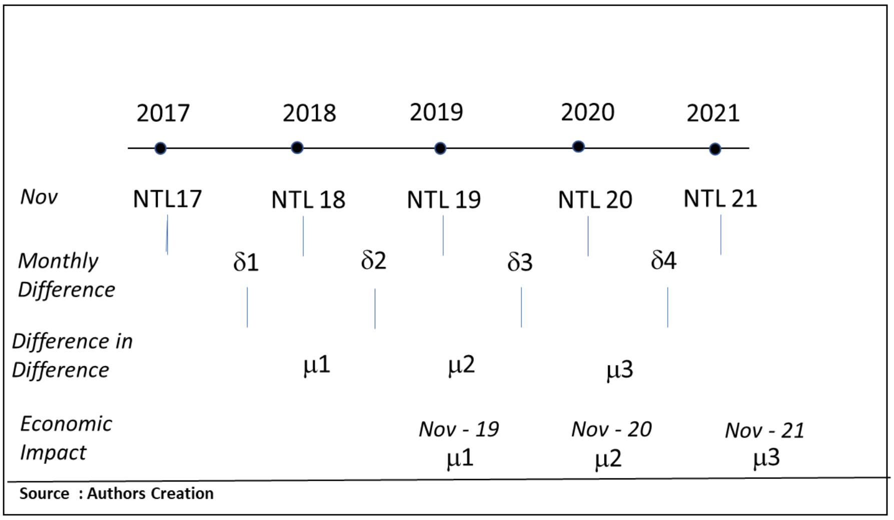
However, Sutton observed using NTL geospatial data, Mean Absolute Deviation of percent error at sub national level GDP estimates is substantially large for China (Sutton *et al* , 2007). The large errors are almost universally observed in states with low GDP, but the aggregation of subnational level estimates to national level GDP estimates turned out to be highly accurate (Sutton et al, 2007). Developed geographies experienced and displayed higher nightlight intensity differential during the differential lockdown period (Beyer, 2020). Despite the crude level of accuracy, approach of using NTL as a robustness check is of immense value to estimate the magnitude of informal economy and impact of an exogenous shock, where data availability is an issue and collection of any data itself is a moral hazard. In summary, NTL observations can be used as a proxy for economic activity, especially over periods or regions where other measures are not available or where the statistical systems are of low quality or when no recent population or economic censuses are available. Similarly, changes in night-time light intensity can be used by economists as an additional measure of income growth when no other measures of income growth are available[13]. The approach and methodology of the same is documented in Annexure II.

We consider the growth of NTL as an indicator for economic activity. To remove the monthly seasonality, we have considered monthly growth as compared to previous year's growth. To measure the impact of Covid-19 as economic shock, we therefore consider second order differences. For example, for November 2019, we calculate difference in NTL intensity between November 2017 and November 2018 referred as δ1 (see figure 4 for the illustration), while the difference between November 2018 and November 2019 is referred as δ2.

[13] worldbank.github.io

Subsequently we calculate difference in δ2 and δ1 (say, μ1), that measures the difference between expected change vs. actual observed change. This value is considered as the measure for an exogenous economic shock.

Figure 4: Calculation Methodology - Covid Effect on NTL



In our regressions we use NTL as an independent variable instead of residential mobility and compare results with those obtained earlier.

In another robustness exercise we use GM's workplace mobility data that depicts mobility of people at places of work, and we also generate results by using per capita Covid-19 infection rates as a measure the covariate shock. A comparative table is reported below, and details are reported in Annexure 2 and 3.

Table 7: Results with Alternative Proxies of Covariate Shock

	Per Capita Infection (1)	**Work Mobility (2)**	**NL (3)**
W_Exogenous Shock	-2.6410*** (0.2128)	0.0357*** (0.0072)	0.1092*** (0.0079)
R_Age	-0.0002** (0.0001)	-0.0002 ** (0.0001)	-0.0002** (0.0001)
R_Pygmy deposit	0.0075 (0.0068)	0.0079 (0.0068)	0.0075 (0.0068)
R_Financial Literacy Score	0.0125*** (0.0014)	0.0125*** (0.0014)	0.0125*** (0.0014)
R_Loan Value sanctioned	0.0005 (0.0041)	0.0005 (0.0041)	0.0005 (0.0041)
R_Location of Shop – In Market	0.0009 (0.0065)	0.0009 (0.0065)	0.0009 (0.0065)
W_UPI Transaction	0.0278* (0.0001)	0.0268*** (0.0001)	0.0278*** (0.0001)
B_UPI Transaction	0.045* (0.0003)	0.045*** (0.0003)	0.045*** (0.0003)
W_Log Average Monthly Balance	0.0134 (0.0110)	0.0120 (0.0108)	0.0134 (0.0110)
B_Log Average Monthly Balance	0.0087 (0.0058)	0.0087 (0.0058)	0.0087 (0.0058)
R_Individual Loan – Unnati	-0.0075 (0.0064)	-0.0075 (0.0064)	-0.0075 (0.0064)
R_Cash Credit	-0.0088 (0.0153)	- 0.0088 (0.0153)	-0.0088 (0.0153)
W_Economic Shock* Market Place	8.1917*** (0.4301)	-0.0015*** (0.0001)	-0.0996*** (0.0116)
W_Economic Shock * Unnati	-7.43300*** (0.7664)	0.0008*** (0.0003)	0.0140*** (0.0207)
W_Economic Shock * Cash Credit	-1.7757 (2.3633)	-0.0008 (0.0008)	-0.0116*** (0.0639)

Robust Standard Error in parenthesis *P Value < 1% , ** P value <5% , *** P value < 10%

*Notes: The dependent variable is Loan Repayment for JLG product. Column 1 has explanatory variable of exogenous shock as Google Mobility. Age , Number of UPI transaction , average monthly savings balance , Location of shop – in market place / out of market place , financial literacy score , loan value sanctioned , is the borrower holder of pygmy deposit (recurring deposit scheme) , product type -Individual Loan (Unnati) , Cash Credit (Overdraft facility) are the explanatory variables. Interaction between, exogenous shock and product type and location of establishment is also factored as dependent variable. Column 1 represents residential mobility as variable for exogenous shock; Column 2 represents NTL data as variable for exogenous shock ; Column 3 represents work mobility as variable for exogenous shock; Column 4 represents covid infection rate as variable for exogenous shock . Estimations are conducted using Hybrid Panel Regression. Clustered standard errors in parentheses. *** p< 0.01, **p<0.05, *p<0.1*

In table 7, we compare different variables representing the exogenous shock. At a portfolio level, consistently for all variables we find that propensity of repayment of loan increases as impact of exogenous shocks decreases. Higher financial literacy, digital transactions and age is associated with higher propensity of loan repayments. While location of shop itself has no impact on loan repayment, however, during a covariate shock (consistent for all variables), establishment in market place showed better repayment behaviour, indicating importance of location of shop to manage exogenous shocks. From a product perspective, Unnati loan and Cash Credit loan showed similar behaviour, however when interacted with exogenous shock, Unnati loan displayed a poorer repayment behaviour and Cash Credit was statistically insignificant except when NTL was used as a variable to depict exogenous shock. Overall the results are consistent and aligned to findings from using RM as a variable for exogenous shock.

When we compare the results for each product using different variables to represent exogenous shocks (see Table 8), for JLG loan the results are consistent in terms of the effect of exogenous shock and digital transactions (both within and between effects). Location of shop is statistically insignificant, except when we use Workplace mobility as a exogenous shock variable; however when interacted with the exogenous shock variable, location of shop does matter in determining the propensity to repay. W_Average monthly balance is statistically significant while B_Average monthly balance is not statistically significant. Similar results are observed in Unnati loan, however age and monthly account balance are not statistically significant.

Table 8: Comparative analysis for JLG Product

	Per Capita Infection (1)	Work Mobility (2)	NTL (3)
Variables			
W_Exogenous Shock	-5.4755 ***	0.0042***	0.1519***
	(0.2971)	(0.0001)	(0.0112)
R_Age	-0.0003***	-0.0003***	-0.0003***
	(0.0001)	(0.0001)	(0.0001)
R_Pygmy deposit	0.0205***	0.0205***	0.0205***
	(0.0094)	(0.0094)	(0.0094)
R_Financial Literacy Score	0.0130***	0.0130***	0.0130***
	(0.0014)	(0.0014)	(0.0014)
R_Loan Value sanctioned	-0.0000	-0.0238	-0.0000
	(0.0000)	(0.0214)	(0.0000)
R_Location of Shop – In Market	-0.0097	-0.0097***	-0.0097
	(0.0087)	(0.0087)	(0.0087)
W_UPI Transaction	0.0277***	0.0267***	0.0277***
	(0.0001)	(0.0001)	(0.0001)
B_UPI Transaction	0.0449***	0.0449***	0.0450 ***
	(0.003)	(0.0003)	(0.0003)
W_Log Average Monthly Balance	0.0196*	0.0183*	0.0200***
	(0.0115)	(0.0113)	(0.0114)
B_Log Average Monthly Balance	0.0043	0.0043	0.0043
	(0.0064|)	(0.0064)	(0.0064|)
W_Covid Infection* Location of Shop	7.9484***	-0.0015***	-0.1673***
	(0.4504)	(0.0002)	(0.0107)

Robust Standard Error in parenthesis *P Value < 1% , ** P value <5% , *** P value < 10%

*Notes: The dependent variable is Loan Repayment for JLG product. Column 1 has explanatory variable of exogenous shock as Google Mobility. Age , Number of UPI transaction , average monthly savings balance , Location of shop – in market place / out of market place , financial literacy score , loan value sanctioned , is the borrower holder of pygmy deposit (recurring deposit scheme) , product type - Individual Loan (Unnati) , Cash Credit (Overdraft facility) are the explanatory variables. Column 1 represents residential mobility as variable for exogenous shock ; Column 2 represents NTL data as variable for exogenous shock ; Column 3 represents work mobility as variable for exogenous shock; Column 4 represents covid infection rate as variable for exogenous shock . Estimations are conducted using Hybrid Panel Regression. Clustered standard errors in parentheses. *** p< 0.01, **p<0.05, *p<0.1*

Table 9 : Comparative Analysis for Unnati loan

	Per Capita Infection (1)	**Work Mobility (2)**	**NTL (3)**
Variables			
W_Exogenous Shock	-14.0655***	0.004***	0.1185***
	(0.9196)	(0.0003)	(0.0342)
R_Age	-0.0001	-0.0001	-0.0001
	(0.0005)	(0.0005)	(0.0005))
R_Pygmy deposit	-	-	-
R_Financial Literacy Score	0.0092**	0.0092**	0.0091**
	(0.0045)	(0.0045)	(0.0045)
R_Loan Value sanctioned	0.0307	0.0307	0.0307
	(0.0729)	(0.0729)	(0.0729)
R_Location of Shop – In Market	0.0034	0.0034	0.0034
	(0.0159)	(0.0159)	(0.0159)
W_UPI Transaction	0.0282***	0.0279***	0.0296***
	(0.0003)	(0.0004)	(0.0003)
B_UPI Transaction	0.0456***	0.0456***	0.0456***
	(0.0009)	(0.0009)	(0.0009)
W_Log Average Monthly Balance	-0.0019	-0.0218	-0.0141
	(0.0374)	(0.0373)	(0.0383)
B_Log Average Monthly Balance	0.0186	0.0186	0.0108
	(0.0200)	(0.0200)	(0.0200)
W_Covid Infection* Location of Shop	12.2345***	-0.0020***	-0.1117***
	(1.5271)	(0.0005)	(0.0548)

Robust Standard Error in parenthesis *P Value < 1% , ** P value <5% , *** P value < 10%

*Notes: The dependent variable is Loan Repayment for Individual Loan -Unnati. Column 1 has explanatory variable of exogenous shock as Google Mobility. Age , Number of UPI transaction , average monthly savings balance , Location of shop – in market place / out of market place , financial literacy score , loan value sanctioned , is the borrower holder of pygmy deposit (recurring deposit scheme) , product type -Individual Loan (Unnati) , Cash Credit (Overdraft facility) are the explanatory variables. Column 1 represents residential mobility as variable for exogenous shock ; Column 2 represents NTL data as variable for exogenous shock ; Column 3 represents work mobility as variable for exogenous shock; Column 4 represents covid infection rate as variable for exogenous shock . Estimations are conducted using Hybrid Panel Regression. Clustered standard errors in parentheses. *** p< 0.01, **p<0.05, *p<0.1*

In Table 10 we find that Cash Credit product displays a unique behaviour. While loan repayment shows dependency on exogenous shock variable and digital transactions, W_average loan balance is significant only for within effects (except for infection rate). Location of shop has a positive impact on the loan repayment, however, when exogenous shock variable interacted with location, we observe in the case of infection rate as a measure of exogenous shock, location in market place has a negative effect on loan repayment. The difference in findings can be explained as follows; firstly, infection rate represents health shock, while other three represent pure economic shocks. Health shock is not necessarily congruent to lockdown or economic shock and lockdown dates and infection rate are not necessarily overlapping. For example, in the delta wave, the per capita infections were on a rise, but there was considerably lower severity in lockdowns as compared to initial waves. Additionally, the collection mechanism for Cash Credit product was physical collection and during the health shock, possible closure of shops resulted in lower collection frequency.

In summary, repayment behaviour is not just dependent on the exogenous shock but also influenced by product characteristics. Financial institutions can enhance product portfolio quality by defining location of the enterprise, minimum financial literacy, digital skills ability , savings banking account balance in the credit policy of the product. Appropriate product structure and product mix aligned to borrowers' profiles can assist the MFI in managing portfolio risk in event of exogenous shock.

Table 10 : Comparative Analysis for Cash Credit

	Per Capita Infection (1)	Work Mobility (2)	NTL (3)
Variables			
W_Exogenous Shock	10.1714*** (3.7979)	0.0033*** (0.0013)	0.2614*** (0.1025)
R_Age	0.0019 (0.0020)	0.0019 (0.0020)	0.0019 (0.0020)
R_Pygmy deposit	-	-	-
R_Financial Literacy Score	0.0167 (0.0195)	0.0167 (0.0195)	0.0167 (0.0195)
R_Loan Value sanctioned	-0.0343 (0.0633)	-0.0343 (0.0633)	-0.0343 (0.0633)
R_Location of Shop – In Market	0.0660 (0.0494)	0.0660 (0.0494)	0.0660 (0.0494)
W_UPI Transaction	0.0222 *** (0.0011)	0.0221*** (0.0012)	0.0222*** (0.0012)
B_UPI Transaction	0.0454*** (0.0032)	0.0454 *** (0.0032)	0.0454*** (0.0011)
W_Log Average Monthly Balance	-0.3150 (0.0971)	-0.2799*** (0.0970)	-0.2966*** (0.0974)
B_Log Average Monthly Balance	-0.0489 (0.0701)	-0.0457 (0.0691)	-0.0457 (0.0691)
W_Covid Infection* Location of Shop	-10.4031*** (4.4344)	-0.0016 (0.0016)	-0.5157*** (0.1659)

Robust Standard Error in parenthesis *P Value < 1% , ** P value <5% , *** P value < 10%

*Notes: The dependent variable is Loan Repayment for Cash Credit. Column 1 has explanatory variable of exogenous shock as Google Mobility. Age , Number of UPI transaction , average monthly savings balance , Location of shop – in market place / out of market place , financial literacy score , loan value sanctioned , is the borrower holder of pygmy deposit (recurring deposit scheme) , product type - Individual Loan (Unnati) , Cash Credit (Overdraft facility) are the explanatory variables. Column 1 represents residential mobility as variable for exogenous shock ; Column 2 represents NTL data as variable for exogenous shock ; Column 3 represents work mobility as variable for exogenous shock; Column 4 represents covid infection rate as variable for exogenous shock . Estimations are conducted using Hybrid Panel Regression. Clustered standard errors in parentheses. *** p< 0.01, **p<0.05, *p<0.1*

8. Conclusion and Recommendation

MFI has dual objectives viz. social objective of outreach and economic objective of profitability. To align them and balance both optimally, it is imperative that MFIs ensure lower PAR and steady disbursement of Credit. Product on its own has limited implication for repayment behaviour and product characteristics and nuances of a product play important roles in repayment behaviour of micro loans, especially in face of covariate exogenous shocks. The findings can be applied to credit policy of a micro lending institution. For example, a scoring model can be developed with parameters and weights attached to financial literacy score , number of UPI transactions , savings balance. This will enable developing a risk-based pricing for the portfolio. The criteria's need to be embedded in the product credit policy of the institutions, which will enable MFIs to assess propensity of the borrower to repay in the event of covariate exogenous shocks. This can be a proactive strategy for risk management. It can also enable mapping / matching product with a segment / subsegment of customers.

From the analysis, we observe that borrower characteristics such as financial literacy, UPI usage, location of trade establishment play important roles in repayment in face of exogenous shocks. Aligning these criteria in credit evaluation will enable healthier portfolio. MFIs need to invest in financial literacy and digital literacy of borrowers, to enable launching and scaling of complex products such as cash credit product.

From a policy perspective, to build consumer resilience, the regulator can explore mandating MFIs, banks, and non-bank financing companies to invest in financial literacy and digital literacy of customers along with ensuring that all repayments be managed digitally.

Our findings can serve as a valuable input for building risk-based pricing model for Base of Pyramid (BOP) customers and can help in formulation credit policy and credit assessment tool. Fresh disbursement and/ or moratorium offered on repayment can be appropriately structure by MFI. The findings can help MFI's in innovate and structure new innovative non joint liability group (JLG) products. Additionally, the regulator might explore guiding financial institutions to incorporate additional behaviour and profile data fields during Credit assessment.

Use of Big data – NTL and Google Mobility can provide as alternative measurement to exogenous shock. Use of this data is of significant value, when area is smaller, and we are measuring an impact of an event. It is recommended the results from using these techniques are triangulated and checked for robustness by using multiple alternative variables. Use of big data opens up opportunity to evaluate various socio – economic development due to an occurrence of an event, not necessarily exogenous shock. For example, one can evaluate development caused due to building of highways, afforestation, and dam building.

This paper opens avenues to conduct further research on borrower behaviour which possibly have bearing on macroeconomic parameters. It also opens opportunity to explore use of big data to understand impact of an event on socio- economic parameters.

9. References

Acharya, V. V., Almeida, H., & Campello, M. (2013). Aggregate risk and the choice between cash and lines of credit, The Journal of Finance, 68 (5), 2059-2116.

Afrah, N. A., & Mohamed, A. M. (2021). "Economic Impacts of COVID-19 on Microfinance Institutions and Small Businesses: Empirical Survey from Somalia."

Agnihotri, I., Mazumdar, I., & Neetha, N. (2011). "Gender and Migration in India." *National Workshop on Internal Migration and Human Development in India* (Vol. 2).

Alderman, H., & Paxson, C. H. (1994). "Do the poor insure? A synthesis of the literature on risk and consumption in developing countries." *Economics in a changing world*, 48-78.

Allison, P. D. (2009). "Fixed effects regression models." *SAGE publications.*

Armendáriz de Aghion, B., & Morduch, J. (2000). "Microfinance beyond group lending." *Economics of Transition*, *8*(2), 401-420.

Aragón, F. M., Karaivanov, A., & Krishnaswamy, K. (2020). "Credit lines in microCredit: Short-term evidence from a randomized controlled trial in India." *Journal of Development Economics*, *146*, 102497.

Assefa, E., Hermes, N., & Meesters, A. (2013). "Competition and the performance of microfinance institutions." *Applied Financial Economics*, *23*(9), 767-782.

Ayadi, R., & Gadi, S. (2013). "Access by MSMEs to Finance in the Southern and Eastern Mediterranean: What role for Credit guarantee schemes?."

Badr, H. S., Du, H., Marshall, M., Dong, E., Squire, M. M., & Gardner, L. M. (2020). "Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study." *The Lancet Infectious Diseases*, *20*(11), 1247-1254.

Bahovec, V., Barbić, D., & Palić, I. (2015). Testing the effects of financial literacy on debt behavior of financial consumers using multivariate analysis methods. *Croatian Operational Research Review*, 361-371.

Baltagi, B. H., & Baltagi, B. H. (2008). "Econometric analysis of panel data" (Vol. 4). *Chichester: John Wiley & Sons*.

Banerjee, A. V., & Duflo, E. (2007). "The economic lives of the poor." *Journal of economic perspectives*, *21*(1), 141-168.

Bansal, A. K., & Bansal, A. (2012). "Microfinance and poverty reduction in India." *Journal of Management*, *5*(1), 31-35.

Barboni, G. (2017). "Repayment flexibility in microfinance contracts: Theory and experimental evidence on take up and selection." *Journal of Economic Behavior & Organization*, *142*, 425-450.

Bell, A., & Jones, K. (2015). "Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data." *Political Science Research and Methods*, *3*(1), 133-153.

Bell, A., & Jones, K. (2015). "Age, period and cohort processes in longitudinal and life course analysis: A multilevel perspective." *A life course perspective on health trajectories and transitions*, 197-213.

Bhalla, S., & Kaur, R. (2011). "Labour force participation of women in India: some facts, some queries."

Bhatt, N., & Tang, S. Y. (2002). "Determinants of repayment in microCredit: Evidence from programs in the United States." *International Journal of Urban and Regional Research*, *26*(2), 360-376.

Biswas, A. (2014). "Financing constraints for MSME sector." *International Journal of Interdisciplinary and Multidisciplinary Studies*, *1*(5), 60-68.

Blumenstock, J. E., Eagle, N., & Fafchamps, M. (2016). "Airtime transfers and mobile communications: Evidence in the aftermath of natural disasters." *Journal of Development Economics*, *120*, 157-181.

Brixiová, Z., & Kangoye, T. (2016). "Start-Up Capital and Women's Entrepreneurship: Evidence from Swaziland." *Available at SSRN 2861024*.

Brockhaus Sr, R. H. (1980). "Risk taking propensity of entrepreneurs." *Academy of management Journal*, *23*(3), 509-520.

Buckee, C. O., Balsari, S., Chan, J., Crosas, M., Dominici, F., Gasser, U., ... & Schroeder, A. (2020). "Aggregated mobility data could help fight COVID-19." *Science*, *368*(6487), 145-146.

Bull, G., & Ogden, T. (2020). "COVID-19: How does microfinance weather the coming storm." *CGAP blog post*, *25*.

Chan, H. F., Skali, A., & Torgler, B. (2020). "A global dataset of human mobility" (No. 2020-04). *CREMA Working Paper*.

Cao, S., & Leung, D. (2020). "Credit constraints and productivity of SMEs: Evidence from Canada." *Economic Modelling, 88, 163-180.*

Chan, H. F., Skali, A., Savage, D. A., Stadelmann, D., & Torgler, B. (2020). “Risk attitudes and human mobility during the COVID-19 pandemic.” *Scientific reports*, *10*(1), 1-13.

Chen, M., Mao, S., & Liu, Y. (2014). “Big data: A survey.”*Mobile networks and applications*, *19*(2), 171-209.

Chen, C. P., & Zhang, C. Y. (2014). “Data-intensive applications, challenges, techniques and technologies: A survey on Big Data.” *Information sciences*, *275*, 314-347.

Chiappori, P. A., Samphantharak, K., Schulhofer-Wohl, S., & Townsend, R. M. (2014). “Heterogeneity and risk sharing in village economies.” *Quantitative economics*, *5*(1), 1-27.

Choudhury, A., Jones, J., & Opare-Addo, M. (2022). Perceived risk and willingness to provide loan to smallholder farmers in Ghana. *Journal of African Business*, *23*(1), 23-40.

Clark, R., Lusardi, A., & Mitchell, O. S. (2017). “Employee financial literacy and retirement plan behavior: a case study.” *Economic Inquiry*, *55*(1), 248-259.

Conning, J., & Morduch, J. (2011). “Microfinance and social investment.” *Annu. Rev. Financ. Econ.*, *3*(1), 407-434.

Cull, R., Demirgu¨ ,C Kunt, A., & Morduch, J. (2007). “Financial performance and outreach: A global analysis of leading microbanks.” *The Economic Journal*, *117*(517), F107-F133.

Cull, R. and Murdoch, J.(2018).“ Microfinance and Economic Development.” *Handbook of Finance and Development. Edward Elgar Publishing*

Czura, K. (2015). “Do flexible repayment schedules improve the impact of microCredit? Evidence from a randomized evaluation in rural India” (No. 2015-20). *Munich Discussion Paper*.

Danes, S. M., & Hira, T. K.(1987) Money Management Knowledge of College. *Journal of Student Financial Aid*, *17*(1), 1.

Derban, W. K., Binner, J. M., & Mullineux, A. (2005). “Loan repayment performance in community development finance institutions in the UK.” *Small business economics*, *25*(4), 319-332.

De Groot, T. (2001). Challenges Faced by Women in Industrial Development Section III in Women Entrepreneurs in Africa: Experience from Selected Countries. *Vienna: UNIDO*.

Dercon, S. (2002). Income risk, coping strategies, and safety nets. *The World Bank Research Observer*, *17*(2), 141-166.

Diagne, A., Zeller, M., & Sharma, M. P. (2000). “Empirical measurements of households’ access to Credit and Credit constraints in developing countries: Methodological issues and evidence.” (No. 583-2016-39550).

Duflo, E. (2005). “Gender equality in development.” *BREAD Policy Paper*, *11*(4).

Duflo, E. (2012). “Women empowerment and economic development.” *Journal of Economic Literature*, *50*(4), 1051-79.

Elborgh-Woytek, M. K., Newiak, M. M., Kochhar, M. K., Fabrizio, M. S., Kpodar, M. K., Wingender, M. P., ... & Schwartz, M. G. (2013). “Women, work, and the economy: Macroeconomic gains from gender equity.” *International Monetary Fund*.

Erdem, F. (2001). A cultural approach toward risk taking propensity and tolerance for ambiguity of entrepreneurs. *Akdeniz IIBF Dergisi*, *2*, 43-61.

Eze, C. C., & Ibekwe, U. C. (2007). “Determinants of loan repayment under indigenous financial system in Southeast, Nigeria.”

Fafchamps, M., & Lund, S. (2003). “Risk-sharing networks in rural Philippines.” *Journal of development Economics*, *71*(2), 261-287.

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

Gebeyehu, Z., Beshire, H., & Haji, J. (2013). “Determinants of loan repayment performance of smallholder farmers: The case of Kalu district, South Wollo zone, Amhara National Regional State, Ethiopia.” *International Journal of Economics, Business, and Finance*, *1*(11), 431-446.

Ghani, S. E., Kerr, W., & O’Connell, S. (2012). “What explains big gender disparities in India? local industrial structures and female entrepreneurship.” *Local Industrial Structures and Female Entrepreneurship (October 1, 2012). World Bank Policy Research Working Paper*, (6228).

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

Global Entrepreneurship Monitor (2018). Global report 2017/18.https://www.gemconsortium.org/report/gem-2017-2018-global-report

Goyal, Preeti, and Vanita Yadav. "To be or not to be a Women entrepreneur in a developing country." *Psychosociological Issues in Human Resource Management* 2.2 (2014): 68-78.

Greenspan, A. (2002). “Financial literacy: A tool for economic progress.”*The Futurist*, *36*(4), 37.

Hausman, J. A. (1978). “Specification tests in econometrics.” *Econometrica: Journal of the econometric society*, 1251-1271.

Hermes, N., & Hudon, M. (2018). "Determinants of the performance of microfinance institutions: A systematic review." *Journal of economic surveys, 32*(5), 1483-1513.
Heitzmann, K., Canagarajah, R. S., & Siegel, P. B. (2002). "Guidelines for assessing the sources of risk and vulnerability." *WB Social Protection Discussion Paper, 218*.
Hirway, I. (2012). "Missing labour force: An explanation." *Economic and Political Weekly*, 67-72.
Ilin, C., Annan-Phan, S., Tai, X. H., Mehra, S., Hsiang, S., & Blumenstock, J. E. (2021). "Public mobility data enables COVID-19 forecasting and management at local and global scales." *Scientific reports, 11*(1), 1-11.
Islam, A., & Maitra, P. (2012). Health shocks and consumption smoothing in rural households: Does microCredit have a role to play?. *Journal of development economics*, 97(2), 232-243.
Isoto, R. E., Sam, A. G., & Kraybill, D. S. (2017). Uninsured health shocks and agricultural productivity among rural households: the mitigating role of Micro-Credit. *The Journal of Development Studies*, 53(12), 2050-2066.
Hung, A., Parker, A. M., & Yoong, J. (2009). "Defining and measuring financial literacy."
Jack, W., & Suri, T. (2014). "Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution."*American Economic Review*, 104(1), 183-223.
Jain, S., & Mansuri, G. (2003). "A little at a time: the use of regularly scheduled repayments in microfinance programs." *Journal of development Economics, 72*(1), 253-279.
Jones, A., N. Rice, T. B. d'Uva, and S. Balia. 2007. *Applied Health Economics*. Abingdon, UK: Routledg
Kanwar, S. (2012). "Intellectual property protection and technology licensing: the case of developing countries." *The Journal of Law and Economics*, *55*(3), 539-564.
Kapoor, M., Kelejian, H. H., & Prucha, I. R. (2007). "Panel data models with spatially correlated error components." *Journal of econometrics*, *140*(1), 97-130.
Karlan, D., & Mullainathan, S. (2007). "Is microfinance too rigid?" *Financial Access Initiative Concept Note*.
Kimondo, N. N. (2013). *The effect of credit policy on the financial performance of deposit taking microfinance institutions in Kenya* (Doctoral dissertation, University of Nairobi).
Klasen, S., & Pieters, J. (2012). "Push or pull? Drivers of female labor force participation during India's economic boom." *Drivers of Female Labor Force Participation During India's Economic Boom*.
Kodongo, O., & Kendi, L. G. (2013). Individual lending versus group lending: An evaluation with Kenya's microfinance data. *Review of Development Finance*, *3*(2), 99-108.
Kramer, A., & Kramer, K. Z. (2020). "The potential impact of the Covid-19 pandemic on occupational status, work from home, and occupational mobility". *Journal of Vocational Behavior*, *119*, 103442.
Kumar, A., & Francisco, M. (2005). "Enterprise size, financing patterns, and Credit constraints in Brazil: analysis of data from the investment climate assessment survey" (Vol. 6). *World Bank Publications*.
Kunreuther, H. (1976). "Limited knowledge and insurance protection." *Public policy*, 24(2), 227-261.
Kunreuther, H., & Slovic, P. (1978). "Economics, psychology, and protective behavior." *The American Economic Review*, 64-69.
Kurowski, Ł. (2021). Household's overindebtedness during the COVID-19 crisis: The role of debt and financial literacy. *Risks*, *9*(4), 62.
Lerner, J., & Tufano, P. (2011). "The consequences of financial innovation: a counterfactual research agenda." In *The Rate and Direction of Inventive Activity Revisited* (pp. 523-575). *University of Chicago Press*.
Lieli, R. P., & White, H. (2010). "The construction of empirical Credit scoring rules based on maximization principles." *Journal of Econometrics*, *157*(1), 110-119.
Littunen, H. (2000). "Entrepreneurship and the characteristics of the entrepreneurial personality." *International journal of entrepreneurial behavior & research*.
Lukács, E. (2005). "The economic role of SMEs in world economy, especially in Europe." *European integration studies*, *4*(1), 3-12.
Lusardi, A. (2008). "Financial literacy: an essential tool for informed consumer choice? " (No. w14084). *National Bureau of Economic Research*.
Luther, W. J. (2020). "Behavioral and Policy Responses to COVID-19: Evidence from Google Mobility Data on State-Level Stay-at-Home Orders." *AIER Sound Money Project Working Paper,* (2020-06).
Mabula, J. B., & Han, D. P. (2018). "Financial literacy of SME managers' on access to finance and performance: The mediating role of financial service utilization." *International Journal of Advanced Computer Science and Applications*, *9*(9).
Mahmud, M., & Riley, E. (2021). "Household response to an extreme shock: Evidence on the immediate impact of the Covid-19 lockdown on economic outcomes and well-being in rural Uganda." *World Development*, *140*, 105318.
Morduch, J. (1995). "Income smoothing and consumption smoothing." *Journal of economic perspectives*, *9*(3), 103-114.

Morduch, J. (1999). “The microfinance promise.” *Journal of economic literature*, *37*(4), 1569-1614.

Morduch, J. (2000). “The microfinance schism.” *World Development*, *28*(4), 617-629.

Mosin, S. (2009). Sample Credit Policies and Procedures Manual. *Microfinance institutions*.

Murdoch, J. (2000). “Networks—a new paradigm of rural development?”. *Journal of rural studies*, *16*(4), 407-419.

Nannyonga, H. L. (2000). “Determinants of repayment behavior in the Centenary Rural Development Bank in Uganda.” *The Ohio State University*.

Nawai, N., & Shariff, M. N. M. (2012). “Factors affecting repayment performance in microfinance programs in Malaysia.” *Procedia-Social and Behavioral Sciences*, *62*, 806-811.

Neuhaus, J. M., & Kalbfleisch, J. D. (1998). “Between-and within-cluster covariate effects in the analysis of clustered data.” *Biometrics*, 638-645.

Neuhaus, J. M., & McCulloch, C. E. (2006). “Separating between-and within-cluster covariate effects by using conditional and partitioning methods.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *68*(5), 859-872.

Ojiako, I. A., Idowu, A. O., & Ogbukwa, C. (2014). “Determinants of loan repayment behaviour of smallholder cooperative farmers in Yewa North Local Government area of Ogun State, Nigeria: An application of Tobit model.” *Journal of Economics and Sustainable Development*, *5*(16), 144-153.

Owusu, A., Oppong, A., Agyeiwaa, D., & Abruquah, L. (2015). Reducing loan default rate among Microfinance institutions (MFIs) in Ghana through innovative product design, delivery and efficient loan management. *International journal of economics, commerce and management*, *3*(3).

Gatalo, O., Tseng, K., Hamilton, A., Lin, G., & Klein, E. (2021). Associations between phone mobility data and COVID-19 cases. *The Lancet Infectious Diseases*, *21*(5), e111.

Paal, B., & Wiseman, T. (2011). “Group insurance and lending with endogenous social collateral.” *Journal of Development Economics*, *94*(1), 30-40.

Panda, S. (2018). “Constraints faced by women entrepreneurs in developing countries: review and ranking.” *Gender in Management: An International Journal*.

Pandey, V., Gupta, A., & Gupta, S. (2019). “Labor and welfare impacts of a large-scale livelihoods program: Quasi-experimental evidence from India.” *World Bank Policy Research Working Paper*, (8883).

Papias, M. M., & Ganesan, P. (2009). “Repayment behaviour in Credit and savings cooperative societies: Empirical and theoretical evidence from rural Rwanda.” *International Journal of Social Economics*.

Parasuraman, A. (2000). “Technology Readiness Index (TRI) a multiple-item scale to measure readiness to embrace new technologies.” *Journal of service research*, *2*(4), 307-320.

Rhyne, E., & Otero, M. (2006). “Microfinance through the next decade: Visioning the who, what, where, when and how.”(pp. 5-45). *Boston, MA: Accion International*.

Rosenzweig, M. R. (1988). “Risk, implicit contracts and the family in rural areas of low income countries. *The Economic Journal*, *98*(393), 1148-1170.

Rustagi, P. (2013). “Changing patterns of labour force participation and employment of women in India.” *The Indian Journal of Labour Economics*, 56 (2), 215-241.

Sarmah, A., Saikia, B., & Tripathi, D. (2022). Does Entrepreneur Gender Matter for Entrepreneurial Motivation: Answers from Micro Small and Medium Enterprises (MSMEs) of Assam. *Journal of International Women's Studies*, *23*(5), 20-40.

Sandhu, N., Hussain, J., & Matlay, H. (2012). “Barriers to finance experienced by female owner /managers of marginal farms in India.” *Journal of Small Business and Enterprise Development*.

Schmidt, R. H., & Zeitinger, C. P. (1996). “Prospects, problems and potential of Credit-granting NGOs.” *Journal of International Development*, *8*(2), 241-258.

Schumpeter, J. A., & Nichol, A. J. (1934). “Robinson's economics of imperfect competition.” *Journal of political economy*, *42*(2), 249-259.

Schunck, R. (2013). “Within and between estimates in random-effects models: Advantages and drawbacks of correlated random effects and hybrid models.”*The Stata Journal*, *13*(1), 65-76.

Sen, A. (1981). “Ingredients of famine analysis: availability and entitlements.” *The quarterly journal of economics*, *96*(3), 433-464.

Sethuraman, S. V. (1998). “Gender, informality, and poverty: A global review.” *Gender Bias in Female Informal Employment and Incomes in Developing Countries*.

Sharma, M., & Zeller, M. (1997). “Repayment performance in group-based Credit programs in Bangladesh: An empirical analysis.” *World development*, *25*(10), 1731-1742.

Sutton, P. C., Elvidge, C. D., & Ghosh, T. (2007). Estimation of gross domestic product at sub-national scales using night time satellite imagery. *International Journal of Ecological Economics & Statistics*, *8*(S07), 5-21.

Thampy, A. (2010), "Financing of SME firms in India: interview with Ranjana Kumar, former CMD, Indian bank; vigilance commissioner, Central vigilance commission," *IIMB Management Review*, 22(3), 93-101

Townsend, R. M. (1994). "Risk and insurance in village India." *Econometrica: journal of the Econometric Society*, 539-591.

Udry, C. (1994). Risk and insurance in a rural Credit market: An empirical investigation in northern Nigeria. *The Review of Economic Studies*, *61*(3), 495-526.

Wanjiku, K. J., & Muturi, W. (2015). Effect of financial literacy on loan repayment a case of ecumenical churches loan fund, Kenya. *International Journal of Social Science and Information*, *3*(2), 1657-1675.

Wara, F., Baimwera, B., & Mutegi, D. (2018). Effect of credit policy on growth of microfinance Institutions. *International Journal of Business Management and Finance*, *2*(1).

Wise, S. (2013). The impact of financial literacy on new venture survival. *International Journal of Business and Management*, *8*(23), 30.

Wooldridge, J. M. (2010). "Econometric analysis of cross section and panel data." *MIT press*.

World Bank. (2011). "World bank development report 2012: Gender equality and development". *The World Bank*.

Yang, D., & Choi, H. (2007). Are remittances insurance? Evidence from rainfall shocks in the Philippines. *The World Bank Economic Review*, *21*(2), 219-248.

Yang, D. (2008). Coping with disaster: The impact of hurricanes on international financial flows, 1970-2002. *The BE Journal of Economic Analysis & Policy*, *8*(1).

10. Annexure

Annexure I: Financial Literacy Questionnaire

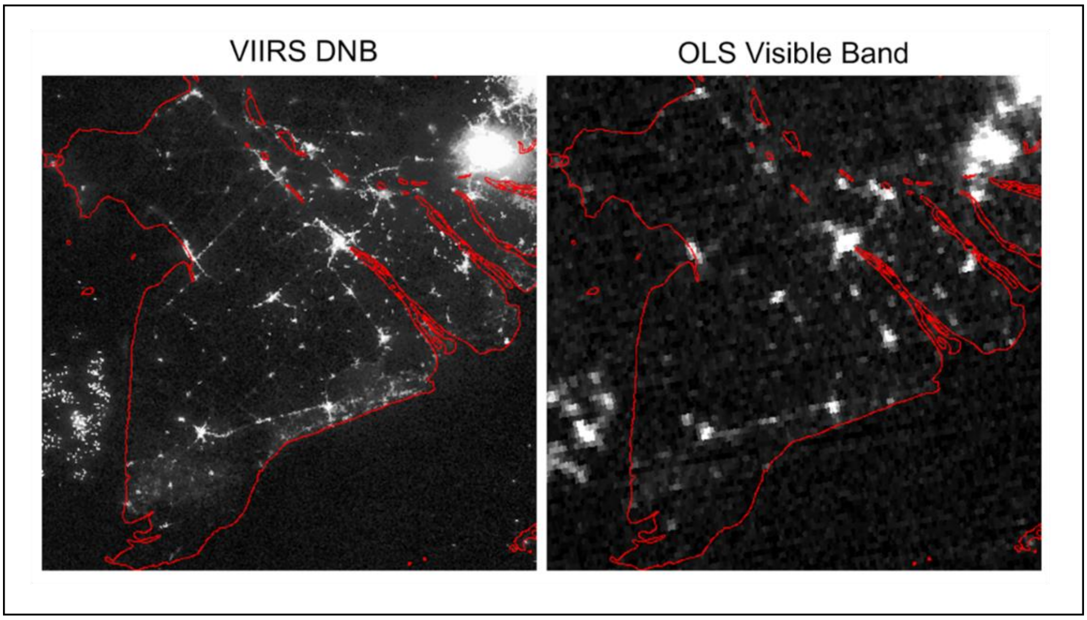
A.	*"You had INR 100 in a savings account and the interest rate was 5% per year. After 5 years, how much would you have in the account if you left the money to grow?"*
B.	*"If you have sales of INR 100, and purchases is of INR 75 , what is your profit? "*
C.	*If 1 person makes 5 chapatis in 1 hour, how many chapatis can 5 people make in 1 hour?*
D.	*If the price of Vada is INR 20 and price of pav is INR 5, what is the price of vadapav?*
E.	*If you have a shop which you don't use, will you rent it out?*

Annexure II: Building the Geospatial Dataset

Selection of Appropriate Image Data set:

For the computation of NTL we use, Google Earth Engine data. GEE is a cloud-based platform that provides geospatial analysis to process a variety of datasets for noncommercial use. GEE provides feeds from NASA – NOAA (National Oceanic and Atmospheric Administration). GEE provides access DMSP-OLS and VIIRS DNB, we have used VIIRS DNB. VIIRS has inherent advantages such as higher resolution, avoiding urban saturation, higher low light urban detection limit, higher quantization, and most important continuity, as DMSP- OLS is end of life.

Figure 6 : Geospatial Images

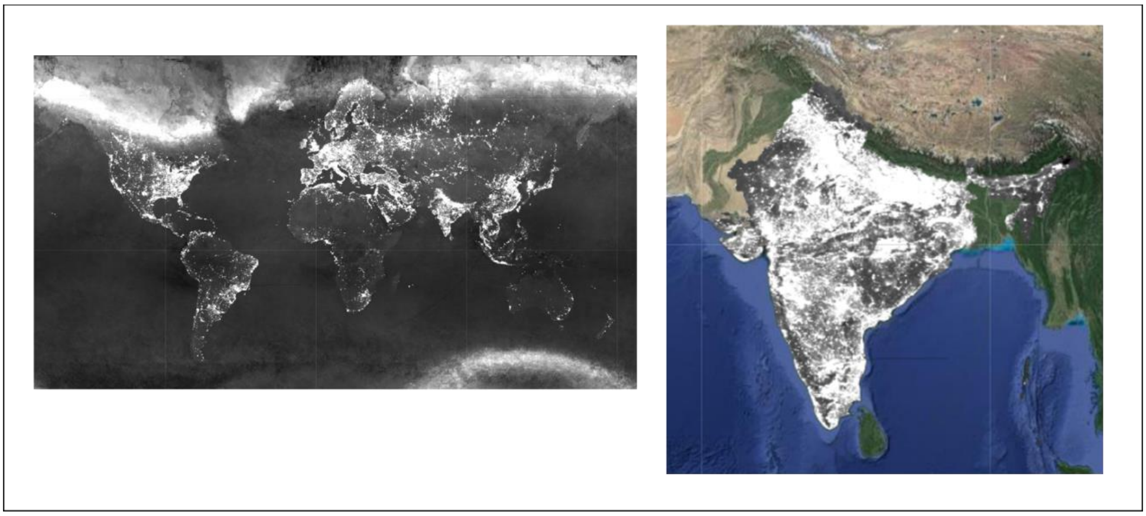


Location: Mekong Delta region Vietnam Date : November 23, 2014. Finer spatial detail and increased sensitivity of the DNB image viz-a-viz to the OLS, Source Worldbank.github.io)

For our analysis we have used VIIRS-DNB sensors (Visible Infrared Imaging Radiometer Suite (VIIRS) – Day and Night Band sensors). VIIRS operates in the visible to near-infrared portion of the spectrum: 400-900 nanometers (nm). This spectral range is ideal for exploring NTL as many human-made light sources provide spectral responses in that range , For example, metal halide lamps, popular on architectural lights, have an increased response at around 550 and 600 nm, and light-emitting diode (LED) lights, which are popular in street lighting, have a spectral signal within this range. VIIRS-DNB sensors provides a resolution of 750 m x 750 m.

Though resolution is at 742 sq m, and for some of the applications such as urban morphology or detection activity in very small human settlements, it is too large as for OLS sensor, it was still designed to be a weather satellite; until a dedicated NTL data satellite is launched, and data is available in public domain, improved sensor in the DNB is used to analyses NTL

Figure 7: Geospatial Images



Source : Google Earth Engine

Source: Google Earth Engine, Author's Analysis

GEE provides, multi-petabyte data catalog alongside a high-performance, intrinsically parallel computation service; which can be accessed by API. It can be accessed and controlled through an Internet-accessible application programming interface (API) making the processing of raster tile easier.

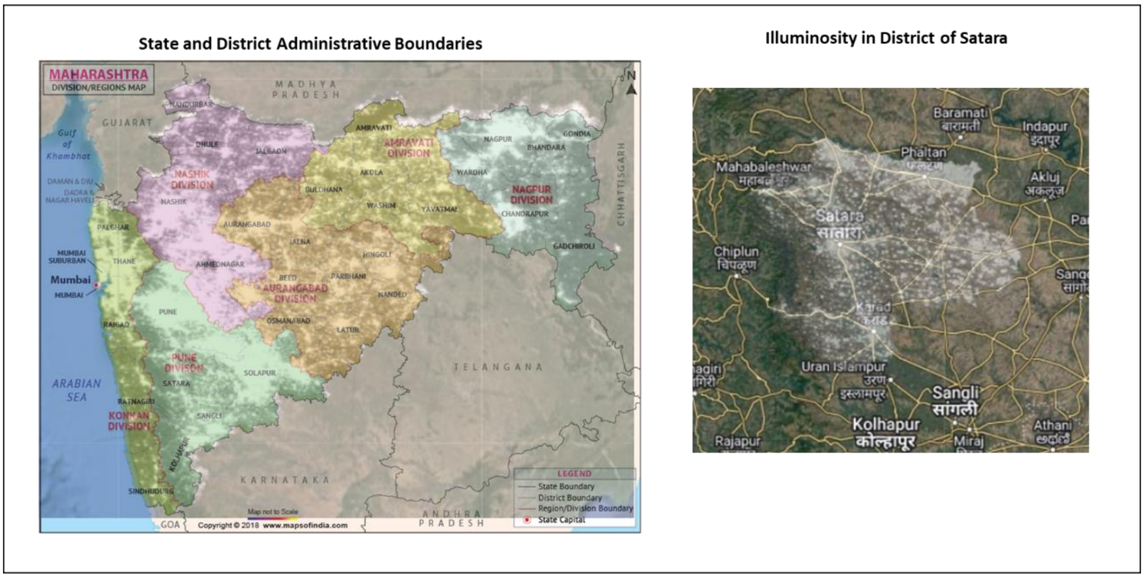
Overlaying of Administrative Boundaries

However, GEE satellite image , there are no administrative boundaries, and it does not have a numeric processable value. To address these challenges, we overlay administrative boundaries. Using http://districts.nic.in/; we first identify list of districts in India. For our study we are interested in Satara District of Maharashtra State. We use shape files from Global Administrative Unit Layers (GAUL) published within GEE. GAUL complies and publishes, information on administrative units for all the countries in the world, providing a contribution to the standardization of the spatial dataset representing administrative units. The GAUL set includes three levels of administrative boundaries and units:

- Level 0 (ADM0): International or country boundaries.
- Level 1 (ADM1): First level administrative boundaries i.e State Boundaries
- Level 2 – Second level -District Boundaries

Our study being relevant to district of Satara, we have used District Boundaries of Satara.

Figure 8: Geospatial Images

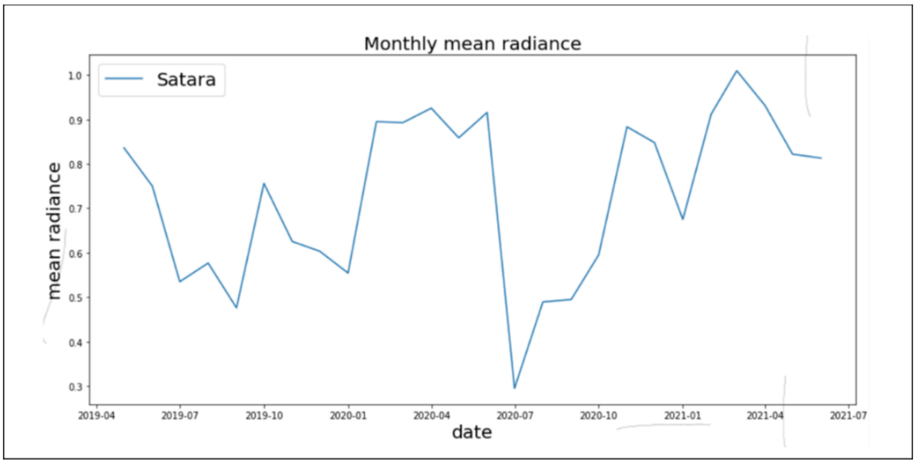


Source: Google Earth Engine , Author's Analysis

Transformation of image to Machine Readable Format

GEE publishes, function to get the mean value of radiance, within a selected data. We utilize this function to get the mean values of radiance and dates in a region. This function is called Reducer. Reducers allows aggregation of GEE data over time, space, bands, arrays and other data structures and produce a single output (e.g., minimum, maximum, mean, median, standard deviation, etc). We have used Monthly Mean over the Level 2 administrative boundaries to calculate NTL luminosity.

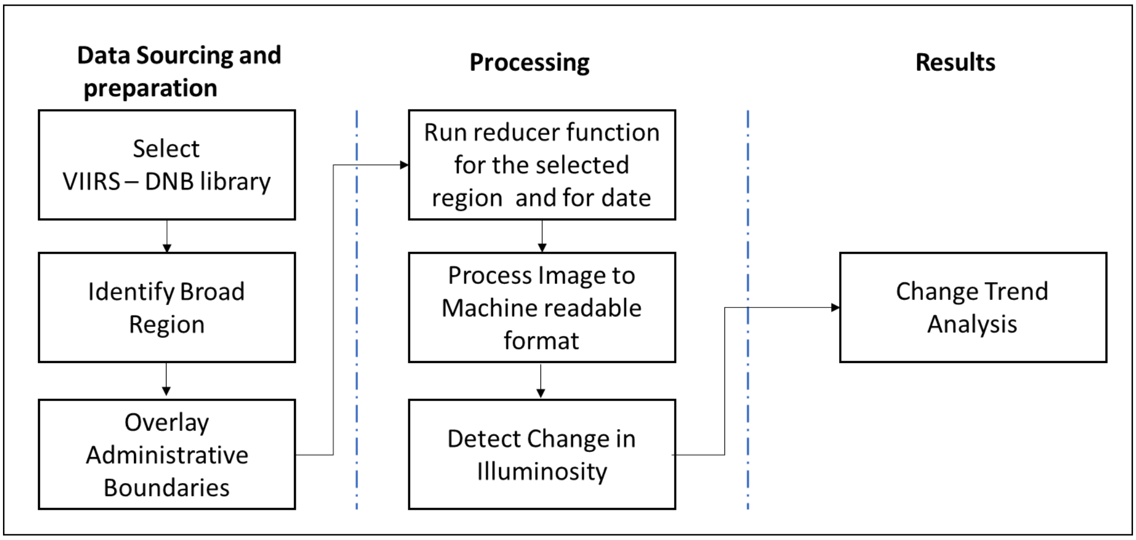
Figure 9: Geospatial Images



Source : Derived value from Google Earth Engine , using Reducer Function, Author's analysis

The overall process can be summarized in the process depicted in Figure 10

Figure 10: Geospatial Images



Source: Author's creation