Banking the Unbanked: What do 280 Million New Bank Accounts Reveal about Financial Access?

Sumit Agarwal
Shashwat Alok
Pulak Ghosh
Soumya Kanti Ghosh
Tomasz Piskorski
Amit Seru*

This Version: September 2022

Abstract

We study the Indian government's largest financial inclusion program that led to over 280 million new bank account openings. Using administrative account-level data, we shed light on the demand and supply of banking services to the poor. Banking usage is initially low but increases over time. Account usage is greater in regions with a high pre-program theft rate indicating that safekeeping is an important function of banking for the unbanked. Individuals also use banking access to manage their liquidity, increasing saving balance during periods of positive income shock and decreasing the balance during periods of high liquidity need. Banks also extend credit to newly banked customers that moves with borrowers' liquidity needs. Concurrently, household survey evidence reveals a decline in individuals' borrowing from informal sources. There is also suggestive evidence that consumption volatility decreases in regions more exposed to the program. These results are consistent with a latent demand among unbanked for banking services that allow for safekeeping and liquidity management through savings and access to formal credit. There is an increase in delinquency rates on credit extended to newly banked consumers, suggesting that unbanked borrowers are of lower credit quality. Quantitatively small deposit inflow from newly banked consumers (0.8% of pre-program deposits) coupled with their lower credit quality likely dissuaded banks from "banking the unbanked" in the absence of government intervention.

JEL codes: C93; D14; G21; O16; O12

Keywords: Banking, Financial Inclusion, Financial Literacy, Big Data, Financial Access, Savings, Spending, Debit Card, Consumer Finance, Household Finance

*First Version: December 2015. We would like to thank Susan Cherry, Sam Liu and Vasudha Nukala for excellent research assistance. We thank Shilpa Agarwal, Abijit Banerjee, Thorsten Beck, Souphala Chomsisengphet, Ross Levine, Andy Rose, Bernie Yeung, Wenlan Qian and seminar participants at BIS, NUS, Reserve Bank of India, ISB Summer Research Conference, Helsinki Finance Summit, IIM-NYU Stern India Research Conference and World Bank for helpful comments and suggestions. We gratefully acknowledge the support from the NSE-NYU initiative for Indian markets, and EY institute of emerging marketing studies for providing the data and the necessary financial assistance for this project. Corresponding Author - Agarwal: National University of Singapore, ushakri@yahoo.com; Alok: Indian School of Business, shashwat_alok@isb.edu; Ghosh: Indian Institute of Management, Bangalore, pulak.ghosh@iimb.ac.in; Ghosh: State Bank of India, kantisoumya@gmail.com; Piskorski, Columbia Graduate School of Business and NBER, tp2252@columbia.edu; Seru: Stanford Graduate School of Business, Hoover Institution, SIEPR and NBER, aseru@stanford.edu.

I Introduction

There is a big debate about the role of financial markets and products in shaping consumer welfare and real economic activity (Mian and Sufi (2015)).[1] In developed economies, such as the U.S., there is an increasing discussion that the financial sector may have become inefficiently large and products offered to households may have become excessively complex.[2] In contrast, in many developing countries, there has been a significant push to increase the usage of financial products – to "complete" the market (Demirgüç-Kunt, Honohan, and Beck (2008)). While there are several studies that evaluate the real effects of access to finance for firms, lack of data has meant there is limited evidence on how access to formal financial products impacts households (e.g., see Dupas, Karlan, Robinson, and Ubfal (2018), Breza, Kanz, and Klapper (2020), Breza and Kinnan (2021)). This paper contributes to this emerging literature by using micro and regional data to evaluate household usage of banking services and lending patterns around the largest financial inclusion program in the world.

Our paper studies the Pradhan Mantri Jan Dhan Yojna ("JDY") launched in India on August 28, 2014. JDY is the world's largest financial inclusion program, with the aim of providing access to banking services for all unbanked households in India. It provided convenient access to saving accounts through a debit card and mobile banking to more than 280 million previously unbanked individuals.[3] Access to banking services can directly benefit lower-income households at the micro-level through safekeeping, savings, spending, and reduction in transaction costs. First, putting money in a bank lowers the risk that the money will be stolen. Second, access to a savings account may allow individuals to better manage their savings and liquidity needs and thereby smooth consumption (Jack and Suri (2014)). Specifically, access to savings technology may allow individuals to build a precautionary buffer that can be used in times of need. For instance, access to a bank account allows consumers to earn interest on their savings and provides incentives to save more. Further, savings in the bank account could help circumvent behavioral biases that would otherwise have caused them to spend this money (Thaler and Benartzi (2004), Ashraf, Karlan, and Yin (2006)). Finally, allowing access to a bank account reduces transaction costs of transferring money to family for subsistence and saving needs. These benefits to households notwithstanding, banks may not supply this service to such households – in the absence of a financial inclusion initiative like JDY – for profitability reasons or due to some other frictions.

In this study, we first document the initial uptake (extensive margin) under the JDY program. Subsequently, exploiting cross-sectional, regional, and time-series heterogeneity, we highlight the

[1] More than 60 countries have adopted financial inclusion as one of the key reform agendas. Financial inclusion is a key aspect of several of the United Nations' sustainable development goals (2014). This thrust is driven by the fact that approximately two billion adults around the world do not have financial access. Of those who have access, approximately 40% actually use it. In India alone, there were approximately 450 million unbanked adults as of 2013 (http://rbi.org.in/Scripts/BS_SpeechesView.aspx?Id=827 [accessed on January 8th, 2016].

[2] See for example Greenwood and Scharfstein (2013) and Philippon (2015).

[3] Easier access is important in developing countries where formal access to bank branches may be costly due to larger distances and lack of proper infrastructure. Similar to phone banking services for the unbanked, debit cards provide for easier access through unmanned ATMs and kiosks.

importance of different functions of savings bank accounts on the demand for banking services by the unbanked.[4]

Second, using both account-level microdata and aggregate data available at the regional level provided by the Reserve Bank of India, we study whether expanding access to traditional savings bank account also results in increased access to credit. At the aggregate level, our analysis compares relative changes in lending patterns in regions with greater exposure to JDY to those with lower exposure around program implementation. Access to banking services can impact lending in several ways. One, such a program could allow new capital to come into the formal banking system in the form of new deposits, relaxing the capital constraints. If such deposits were quantitatively large, it would allow banks to increase lending to their clients. Alternatively, cost of acquiring customers as well as potentially lower credit quality might have kept unbanked outside the formal banking sector before the program. JDY could have lowered the cost of acquiring such customers thereby allowing credit to be extended to some of them. Notably, quality of lending (proxied by default rates on loans), as well as magnitude of new deposits would inform us on the pre-program lack of supply of banking services to the target population by profit-seeking banks.

Our micro-level analysis relies on proprietary micro-level data that includes a random sample of approximately 1.5 million accounts opened under JDY from August 2014 to May 2015 by one of the largest banks in India. This allows us to capture the usage of banking services during the first ten months of the program. This bank is one of the largest Indian banks based on deposit and lending base. In addition, we obtain data from the same bank on around 50 thousand regular accounts of individuals broadly similar to JDY households opened during the same period ("non-JDY" accounts). This dataset provides us with precise account-level information on monthly account balance, withdrawal, deposit, inward, and outward remittance transactions, along with demographic information on the consumers. We also supplement this data with regional and aggregate statistics provided by the central bank, which is available to us over a longer time period.

We begin by documenting substantial outreach of the program (i.e., the extensive margin). In particular, the program led to a massive increase in the number of households having access to formal banking services. The number of accounts steadily increased at a rate of 14% new accounts per month since the start of the program. As of March 29, 2017, we find 282 million new accounts and 220 million debit cards issued under JDY in the aggregate data. Moreover, 75% of the accounts maintain some positive balance. These facts are consistent with those obtained by using the micro-level data from our bank and extrapolating the estimates to the national level over the longer horizon. We also find that the average monthly balance maintained in JDY accounts is INR 504 (USD 7) or about 62% of the rural poverty line in India.[5]

Along the intensive margin, we find that approximately 81% of the new consumers do not

[4]We also compare the usage patterns of banking services of households who got access to banking under JDY with low-income households who access banking services outside the program to examine how usage of JDY bank accounts evolves over time.

[5]INR 504 translates into USD 7 at the 2016 nominal exchange rate of INR 68 per USD. On a PPP basis, this translates into USD 23.

deposit any money after account opening in the first six months since the account opening. About 12% of individuals perform one deposit transaction while only 7% perform two or more deposit transactions. The statistics are qualitatively similar for cash withdrawals, with approximately 87% of the sample not withdrawing any cash after opening the account, about 5% withdrawing cash only once and 8% withdrawing cash two or more times. In terms of the types of transactions done by households banked under JDY, inward and outward remittances are the most common transactions. Account usage increases over time as individuals gain familiarity with banking, suggesting that some of the effects of financial inclusion may manifest over the long-term.[6]

We next analyze regional and time-series variation to understand the determinants of account usage. To understand whether safekeeping function is indeed valued by the unbanked, we exploit regional variation in crime rate. To this end, we split our sample of the JDY accounts into those in districts with above (high-crime districts) and below (low-crime districts) median pre-program crime per capita, and examine account usage over time.[7] We find that both level and monthly growth in account balance is significantly higher in regions with a higher ex-ante crime rate. The likelihood of performing at least one monthly transaction is approximately 3.5% higher for customers in regions with high ex-ante crime rate, a large effect relative to the average usage rate of 17% in our data.

Next, we examine if access to savings accounts affects how individuals manage their liquidity needs. Specifically, we seek to understand if individuals build a saving buffer during good times and withdraw money during periods of increased liquidity need. To generate exogenous variation in income shock, we exploit variation in the timing of rainfall. As is well known in our setting, positive rainfall shock has been shown to be associated with an increase in nominal wages (Kaur (2019); Adhvaryu, Chari, and Sharma (2013)). We classify a district as experiencing a positive rainfall shock if the rainfall intensity in the district is higher than one standard deviation above the district's historical mean. Consistent with our conjecture, we find that the account balance of individuals increases on average if they are located in districts that experience a positive rainfall shock.

To generate exogenous variation in the need for liquidity, we exploit time-variation in Hindu marriage months (April to May and October to December). As is well known in our setting, individuals are likely to incur higher expenses during Hindu marriage months due to the cost of traveling and gifts associated with marriage ceremonies (Krishna (2004)). Consistent with individuals using JDY accounts to manage their liquidity needs, we find that account balance goes up two months prior to Hindu marriage months, and falls in the marriage months. The decrease in savings balance persists up until two months later.

Further, we assess if banks also extend credit to newly banked customers that moves with their liquidity needs, potentially smoothening the impact of negative liquidity shocks. We obtain data on

[6]We also compare the account activity of JDY households to similar households that were banked without direct government intervention (non-JDY households) and find that the usage patterns under the program gradually converge over time to those of similar individuals who were banked outside the program. Our estimates suggest this convergence occurs on average within seventeen to thirty months since an account opens.

[7]Crime includes theft, robbery, and burglary as defined in the Indian Penal Code.

loans granted to JDY account holders and find evidence that supports the findings on account usage – loans to newly banked go up during the marriage season and fall in times of positive liquidity shock. Overall, this evidence suggests that access to bank accounts can potentially enable ex-ante unbanked individuals to better manage their savings and liquidity needs. Moreover, both savings and transactions go up over time for individuals that are banked under the program. This evidence is consistent with learning by individuals that results in an increase in usage over time as they gain familiarity with banking services.

Next, we exploit spatial (regional) variation in the implementation of this program to investigate how access to consumer savings accounts is related to broader economic outcomes such as household borrowing and the nature and purpose of borrowing. To do this, we construct four ex-ante measures of JDY program exposure: (i) number of adults per unit bank branch in a region – this captures the extent of bank branch penetration, (ii) fraction of bank branches owned by state-owned banks in a region – since privately owned banks are more likely to open branches in higher-income areas with greater financial inclusion, (iii) fraction of unbanked households in a region – this captures the extant level of banking access in each area and (iv) a comprehensive financial inclusion index obtained from CRISIL that uses three parameters as inputs to the index: bank branch penetration, deposit penetration, and credit penetration.[8] A higher value of all four measures indicates a lower level of financial inclusion. We compare changes in the regional outcome variables in regions with greater ex-ante program exposure relative to regions with lower program exposure around program implementation.

We begin our regional analysis by verifying that our ex-ante measures of regional JDY exposure in a region before the program indeed correlate with the subsequent intensity of treatment from the program. We observe that there is a strong positive association between our ex-ante exposure measures with both the number of JDY accounts opened and the total amount deposited in these accounts. Next, we examine whether JDY is associated with an increase in bank lending. Using aggregate data provided by the Reserve bank of India, we observe an increase in aggregate lending in areas with greater ex-ante JDY exposure. These results are consistent with our micro-level analysis on credit access around periods of positive and negative liquidity shock. However, to the extent that one may be concerned regarding the plausibility of these estimates, we verify these effects are present in our proprietary bank data and find an increase in the amount of loans granted in regions with greater JDY exposure relative to those with lower exposure. We find that the total amount of new deposits brought under JDY is small relative to overall deposits in the banks before the program. In particular, the INR 630 billion deposited in JDY accounts is a mere 0.8% of the pre-JDY deposits in the banking sector. Thus, it is unlikely that the additional lending in more exposed regions solely reflects a relaxation of bank financial constraints due to new deposit inflow. Rather, our findings suggest that JDY, by taking away the margin of cost of customer acquisition, may have allowed banks to meet the unmet demand for credit for some households that did not have prior access to formal banking products.

[8] CRISIL is a global analytical company providing Ratings, Research and Risk & Policy Advisory services.

To provide further evidence on the increase in credit access, we use data from a time-series panel of a household survey conducted by the Centre for Monitoring Indian Economy (CMIE) and examine the association between ex-ante exposure to JDY with the sources and reasons for borrowing, household savings, and consumption. Focusing on the low-income households defined as those with below median monthly income in our sample, we find that consistent with the findings based on bank data, there is an increase in the fraction of households borrowing from banks in regions with greater ex-ante JDY exposure relative to regions with lower exposure. We also observe a contemporaneous decrease in the fraction of households borrowing from non-bank sources such as informal moneylenders, chit funds, friends, and family. To the extent that many of these informal creditors may engage in high-cost loans, this evidence is consistent households substituting to less costly bank credit whenever possible. We also find some suggestive evidence that some of the increased borrowing goes towards funding medical expenditure needs, which might be productivity enhancing if such purchases were curtailed or deferred when only high-cost funds were available.

We also analyze if such increased lending led to increased credit risk for the bank. To examine this conjecture, we use proprietary lending data obtained from the bank and compare the difference between the average monthly default rate on newly originated loans during the program period (September 2014 to December 2015) and the average default rate on loans originated just prior to the program period (January 2014 to July 2014).[9] The default rate is 0.2 to 0.4 percentage points higher for loans originated during the program period, suggesting that some of the ex-ante unbanked borrowers are less creditworthy than previously banked borrowers. Moreover, the increase in default rate is economically significant given that the average delinquency rate in our sample is 2.1 percent. Overall, lower creditworthiness of the newly banked and limited new capital brought in by these individuals – compared to potentially relatively high costs of customer acquisition – might have prompted banks to rationally not serve these consumers before JDY program.

We conclude by examining whether the ability to manage their liquidity needs through access to a savings account allowed households to smooth consumption (Jack and Suri (2014)). Relying on consumer survey data we find a decrease in the monthly volatility of consumption expenditure in regions more exposed to JDY. This is consistent with consumers using bank accounts to manage their liquidity needs.

This paper makes several contributions to the existing literature on financial inclusion. First, uunlike most of the prior literature, which relies on field experiments and financial inclusion interventions with limited breadth and scope, we study the largest financial inclusion program in the world. Extant literature has used survey instruments to measure access, usage of financial services, and other household outcomes. Prior literature highlights that survey instruments, particularly when asking questions about finance, could be biased (Johnson, Parker, and Souleles (2006)). In contrast, we rely on administrative data, which allows us to measure targeted households' use of banking services directly. Further, by comparing the new deposit capital brought into the financial

[9]Default rate is defined as the proportion of loans originated in a given month that become 60-day delinquent or 90-day delinquent within a year from loan origination.

system with ex-ante deposit levels, along with the default rates of new loans granted under the JDY program, we are able to provide suggestive evidence of the reason for the lack of supply of banking services to the poor.

Our work is also related to the broad theoretical and empirical literature on financial inclusion. Theoretical work in this literature highlights that access to financial services can help low-income individuals move out of poverty (Aghion and Bolton (1997), Banerjee and Newman (1993), Callen, De Mel, McIntosh, and Woodruff (2019)). The focus of several empirical papers that have examined this issue is on understanding the broader impact of increased access to banks on aggregate income and labor market outcomes (Burgess and Pande (2005), Bruhn and Love (2014))). However, micro-level evidence on the usage of banking services by poor individuals is scanter (e.g., see Cole, Sampson, and Zia (2011), Dupas and Robinson (2013), Prina (2015), Dupas et al. (2018), Célerier and Matray (2019), Bachas, Gertler, Higgins, and Seira (2021)).[10] Our evidence documenting high take-up but low initial subsequent usage of bank accounts are broadly consistent with the findings reported in RCT studies conducted by Dupas and Robinson (2013) and Dupas et al. (2018).[11]

While we do not benefit from a well-designed RCT, our empirical setting allows us to make a number of contributions to the literature. First, we find additional evidence that suggests that poor households learn as they become more familiar with banking services over time even with access to regular bank accounts. This indicates that the real impact of financial inclusion programs could manifest over the long term as more and more individuals gradually start using these services. Second, our analysis allows us to shed light on the determinants of account usage across regions and time. For instance, we document that both levels and the rate of growth in account usage is higher for accounts in areas with a greater incidence of thefts and robberies. Consistent with the survey evidence in Breza et al. (2020) regarding the effect of bank accounts on the ability to deal with unanticipated shocks, we show that individuals build a savings buffer during periods of positive income shock and draw on the savings during periods of high liquidity need using administrative data from banks.

Third, RCTs are by design limited in scope and not suitable to examine broader regional effects

[10] Cole et al. (2011) examine the reasons for low demand for financial services among the poor using an RCT study and find that compared to encouraging financial literacy, modest subsidies to bank account holders may work better to incentivize bank account usage. However, they find both low take-up and subsequent low usage even among those offered financial incentives to open bank accounts. In contrast, using an RCT design, Prina (2015) documents both high take-up and high usage among bank accounts offered to female household heads in Nepal. In a field experiment conducted in Bangladesh, Breza et al. (2020), study the impact of access to payroll accounts on largely unbanked factory workers. In their setting, receiving wages in a bank account engenders incentives to engage with the financial technology, thereby resulting in learning. Consistent with "learning-by-doing" those provided access to payroll bank accounts use their bank accounts more as compared the control group which received access to regular bank accounts. Using surveys, they find that payroll account holders report a greater ability to cope up with unanticipated shocks. These effects are similar among individuals who got access to a regular bank account.

[11] The bank accounts offered in these studies offered no interest rate on deposit and levied a withdrawal fee. This makes it difficult to disentangle the impact of monetary transaction costs from the lack of demand for banking services. In our setting, the interest rate offered on savings is significant (4%-6% depending on the bank), and comparable to the rates offered to any other ordinary savings account holder. Moreover, the first four withdrawals in any month are free. This allows us to abstract from financial transaction costs as an explanation for low usage of these accounts. Another paper, released subsequent to our study – Chopra, Prabhala, and Tantri (2017)– also examines account usage of PMJDY accounts and provides external validation of our micro-level findings.

of financial inclusion. In contrast, our setting allows us to shed some light on such effects by examining the evolution of economic outcomes in regions more and less exposed to the program. We provide evidence of an increase in credit access in areas more exposed to the program using micro-level loan accounts and data on regional credit. Finally, to the best of our knowledge, ours is the first paper to focus on a large-scale financial access policy program with hundreds of millions of targeted households (see also Higgins (2020) for evidence on financial inclusion program in Mexico). Our study sheds light on the reasons underlying the lack of supply of banking services to the poor. Specifically, quantitatively small deposit inflow from newly banked consumers (0.8% of pre-program deposits) coupled with their lower credit quality likely dissuaded banks from "banking the unbanked" in the absence of government intervention.

Our work is related to the large literature highlighting the positive link between financial development and economic growth (e.g., King and Levine (1993), Jayaratne and Strahan (1996) Rajan and Zingales (1998), and Black and Strahan (2002)). However, much of this literature focuses on the broader country-level financial development and the impact of access to finance for firms on economic growth. In contrast, the literature evaluating the role of increased access to consumer-level financial products on both micro-level individual outcomes and the broader aggregate economy is small. We further the work in this area by studying the largest experiment in expanding access to banking services for low-income individuals.

Finally, our work is also broadly related to empirical studies evaluating the micro and regional effects of large-scale programs aimed at mortgage and consumer credit markets (e.g., Mayer, Morrison, Piskorski, and Gupta (2014), Agarwal, Chomsisengphet, Mahoney, and Stroebel (2015), Agarwal, Amromin, Chomsisengphet, Landvoigt, Piskorski, Seru, and Yao (2017); Agarwal, Amromin, Ben-David, Chomsisengphet, Piskorski, and Seru (2017), Ganong and Noel (2020), DeFusco, Johnson, and Mondragon (2020)) as well to studies evaluating polices aimed at stimulating household consumption (e.g., Johnson et al. (2006), Mian and Sufi (2012), Baker and Yannelis (2017), Granja, Makridis, Yannelis, and Zwick (2022), Baker, Farrokhnia, Meyer, Pagel, and Yannelis (2021)).

II Institutional Context, Data, and Summary Statistics

II.A Prime Minister Jan Dhan Yojna's Program Background and Objectives

The incumbent Prime Minister of India, Mr. Narendra Modi, launched the Pradhan Mantri Jan-Dhan Yojna (JDY from now), a national financial inclusion mission program, on 28 August 2014. The primary stated objective of this program is to ensure access to basic financial services (e.g., savings and deposit accounts, remittance) in an affordable manner to the low-income strata of India. JDY's ultimate aim is to ensure bank account access for each household in India.

The main features of JDY that distinguish this scheme from earlier financial inclusion programs are: (i) universal access to banking facilities along with financial literacy programs to improve the understanding of financial products for effective use; (ii) provision of basic bank accounts such as zero-balance accounts with "RuPay" debit cards and an overdraft facility of INR 5,000 (USD

73) after six months of satisfactory transaction record; (iii) provision of insurance facilities such as accidental insurance cover of INR 1 lakh (0.1 million) to all account holders and life insurance cover of INR 30,000 (USD 440) to those who have opened the account by 26 January 2015;[12] (iv) provision of mobile banking to conduct simple transactions such as transferring funds and checking balance and (v) access to micro insurance and pension schemes in the second phase of the program. Moreover, while prior programs focused on village level metrics of banking access, JDY was explicit in its objective to provide access to each household.

To ensure program success, the government followed a multi-pronged approach. First, districts that are regional administrative units were further sub-divided into 1,53,488 sub-service areas (SSA) comprising of 1000-1500 villages. To ensure accountability at the bank level, each state-owned bank was allotted different SSAs based on their existing presence in the region. The bank in charge of an SSA is mandated to provide a fixed-point banking outlet through either a brick and mortar branch or a business correspondent (BC) who act as a representative of the banks within 5 km of every village. Towards this end, the government would set up an additional 50,000 business correspondents, 7,000 brick and mortar branches, and 20,000 ATMs across all state-owned banks. One key feature of the program is the greater thrust on the use of technology to reduce the cost of banking access. Among other technologies, first, it includes micro-ATMs provided to each BC that enables them to function as a mini bank branch and offer basic banking services such as deposit, withdrawal, fund transfer, and balance inquiry. A second feature is a provision of no-frills basic mobile banking for each account holder through SMS that allows for fund transfer and balance inquiry. In addition, BCs were trained to create awareness regarding the benefits of JDY accounts such as overdraft facility and insurance provisions. The availability of BCs at their doorstep and mobile banking potentially reduces a significant barrier to bank account opening for the poor, the cost and time taken to travel to a bank (Demirguc-Kunt, Klapper, Ansar, and Jagati (2017)). Further, apart from the use of doorstep banking through BCs, banks also set up camps in villages to be able to open a larger number of accounts in the early days of the program. To ensure financial viability, various subsidies were made available to banks.

PMJDY also significantly reduced another pre-existing barrier to account opening, the requirement of documents, and the existing complex application procedure through the use of a biometric authentication system and a one-page simplified application form. Even in the absence of any identification, an account can be opened on a temporary basis for a period of one year that could be made permanent on providing the documents later (Demirguc-Kunt et al. (2017)).

Finally, an extensive monitoring process was set up to ensure program success with implementation committees at the district level, state level, and the central level. Each bank was required to set up an MIS portal to monitor the functioning of BCs. This is linked to a portal maintained by the

[12]The accidental insurance will initially be for five years while the life insurance covers a person until they turn 60. The eligibility criteria stipulate that the insured needs to have a valid RuPay debit card. Only one person in every household can avail this insurance. A claim under the Personal Accidental Insurance under PMJDY is payable if the Rupay Cardholder has performed a minimum of one successful financial or non-financial card transaction within 90 days prior to the date of accident.

department of financial services (DFS). A banking committee is required to monitor the progress from the DFS portal at a weekly level. This further monitored by the district level committee fortnightly, state-level committee monthly, and eventually the finance minister's team quarterly.

JDY has spurred a heated debate amongst policy circles and academics regarding the long-term implications of the program for household welfare and real economic activity. A similar initiative called the "no-frills" account scheme launched by the Reserve Bank of India in 2005 did not see much activity among those who opened the accounts. Some commentators have suggested that JDY might also follow a similar trajectory.[13]

II.B Data Sources

Our main data is based on a proprietary dataset obtained from the largest public bank in India. This bank is one of the largest Indian banks based on deposit and lending base. In the pre-program period, this bank held about 22% of the deposit base in the country. We obtain data on two types of accounts. First, our JDY group comprises a random sample of 1,518,865 accounts opened under the JDY launched by the government of India. We have nine months of information on these accounts – i.e., between September 2014 and May 2015. We also subsequently obtained data on loans granted to these accounts. The loan data provided by the banks include the loan amount and the date of the loan grant and includes loans applied for and granted to JDY account holders in our sample over the period August 2014 to March 2017.

In addition, we obtain data on a random sample of 50,089 non-JDY accounts opened during the same sample period. These accounts are selected to be demographically close to the JDY sample, with the difference being that these individuals have access to banking services outside JDY. The Non-JDY accounts were selected in two steps. First, the sample of accounts opened outside the program was restricted to those with the opening balance within two standard deviations of the average opening balance for JDY accounts. From this set, a random sample was drawn. For all individuals in our sample, we have monthly information on the average monthly balance, cash deposit transactions, cash withdrawal transactions, remittances, and access to debit cards, among other things. The data also contains a rich set of demographics about each individual, including age, gender, marital status, mobile phone ownership, and the district of residence.[14]

Our micro-level data is aggregated at the account-month level. For instance, we compute the Cash Deposit Amount (Cash Withdrawal Amount) by summing all deposit (withdrawal) transactions by an individual in a month. Likewise, we aggregate all monthly inward and outward remittance transactions for each account. The average monthly balance is the average of the daily account balance in a month.

We supplement this dataset with district-level data on GDP from Indicus Analytics, the population from the latest Census of India (2011), and aggregate district-level lending data from the

[13] http://www.thehindubusinessline.com/opinion/a-pointless-number-chase/article7735093.ece [accessed on Jan 15, 2016]

[14] Districts are territorial administrative units in India that are similar to counties in the USA.

Reserve Bank of India (RBI).

Finally, we use data from a time-series panel of a household survey conducted by the Centre for Monitoring Indian Economy (CMIE). We use this data to examine the association between ex-ante exposure to JDY with the sources and reasons for borrowing, household savings, and consumer spending patterns.

Admittedly, our sample period varies across tests based on data availability. In particular, the bank's proprietary data on JDY accounts is only available until May 2015. The data on loans granted to these account holders covers the loans granted until March 2017. To maintain consistency across tests, in all our analyses, permitting data availability, we track the outcome variables until March 2017. Appendix A provides details of all the variables used in our study, the data source, and the time span for which the data is available.

II.C Aggregate Summary Statistics

Using the data provided by the central bank that coordinates with the ministry of finance, we are able to get an aggregate overview of the program. Table 1, Panel A, shows key aggregate statistics for the Indian economy on the outset of the program. Table 1, Panel B shows that as of 29th March 2017, 282 million accounts were opened under JDY.[15] In addition, 220 million debit cards were issued, with around INR 630 billion (USD 9 billion) deposited in these JDY accounts. Thus, an additional capital approximating USD 9 billion became part of the Indian financial system as a result of the JDY program as of March 2017. However, the amount deposited in JDY accounts is a mere 0.8% of the pre-JDY deposits in the banking sector. The number of accounts with positive balance has progressively increased – from around 23.0% in September 2014 to about 75% as of 29th March 2017.[16] These statistics closely track the positive balance trends in a random sample of accounts that form our micro-level data.[17]

The banks participating in the JDY program are divided into three kinds: public sector – i.e., owned by the government and with a national presence, rural regional –i.e., owned by the government and with local presence created with the mandate to primarily service the rural areas, and major private banks. A vast fraction of accounts opened under JDY were in public sector banks. Specifically, around 80% of the accounts opened as of March 2017 were in public sector

[15] We get very similar aggregate statistics when we take the information on account openings in our micro-level data and scale them to the national level. In particular, the new account openings in our micro-level data in the first ten months are 1.5 million (approximately 4% random sample of the total JDY accounts opened by our bank). Scaling this by the ratio of aggregate deposits (INR 80,000 billion) to deposits in our bank (INR 17,000 billion) in the period before JDY implementation (as of the end of March 2014), we find that the number of new accounts opened in the first ten months is 176 million. This is comparable to 158 million new JDY accounts opened based on the data we obtain from the central bank for the period August 2014 to May 2015.

[16] On the inauguration day of JDY, 1.5 crores (i.e., 15 million) bank accounts were opened. Guinness World Records recognized this achievement and provided a certificate that says the most bank accounts opened in one week as part of a financial inclusion campaign is 18,096,130 and was achieved by banks in India from 23 to 29th August 2014.

[17] In particular, in our micro-level data, about 43% (46%) of the accounts had positive balance as of April 2015 (May 2015) (the tenth month). In the data obtained from the central bank, this statistic is 44% and 46% as of April 2015 and May 2015 respectively.

banks, with the largest players being State Bank of India (SBI) (31%), Punjab National Bank (PNB) (8%) and Bank of Baroda (7%).

The majority of the unbanked in India are in rural regions, with estimates from the central bank suggesting that only 40% were banked in such regions. As a result, the government expected that most of the new accounts opened under JDY would be in rural regions. Consistent with this view, of the total number of accounts opened, 61% were from rural regions, and 39% were from urban regions. The substantial fraction of accounts opened in urban areas is not surprising, given that around 40% of low-income individuals residing in urban areas did not have bank accounts prior to JDY.

II.D Account Level Summary Statistics

Panels C and D of Table 1 reports the summary statistics for key variables used in our account level analysis. Panel C shows monthly amounts for financial transactions across the two samples used in our analysis. The average monthly balance is INR 504 (approximately USD 7) for the JDY sample and INR 2420 (USD 35) for the non-JDY sample. The low balance for the JDY sample is not surprising, given that these accounts cater to the individuals below the poverty line or just above the line. The poverty line in India is INR 816 (USD 12) per month for rural areas and INR 1000 (USD 15) per month for urban areas. In percentage terms, INR 504 is 62% of the monthly poverty line. Thus, the average monthly balance maintained in these accounts is economically consequential, given their monthly income levels. The average monthly balance maintained by individuals in the non-JDY sample is about six times those in the JDY sample. This relative difference is sensible since individuals in the non-JDY samples consist of individuals with higher income levels as compared to those in the JDY sample. Consequently, the average monthly balance and transaction amounts are higher for such individuals.

In Panel D, we report dummy variables that identify individual-months for each of the five kinds of transactions. Approximately 36% of the individuals in the JDY sample and 92% individuals in non-JDY sample operate accounts with a positive balance. Focusing on the last row of Panel B that relates to overall usage, we see that about 17% of individuals in the JDY sample, and 59% in non-JDY use the accounts monthly for at least one of the four purposes: deposit, withdrawal, inward or outward remittance. We note that these statistics are not directly comparable as they do not account for potential differences across these groups, such as the average age of the accounts. Consequently, in our empirical analysis, we will compare the usage patterns of banking services over time among individuals in each of these groups, controlling for a number of observable characteristics, and individual fixed effects.

II.D.1 Program Reach (Extensive Margin)

We begin by providing some aggregate statistics on the program take-up. The program started with about 15 million accounts opened up the first day itself. Since then, the number of accounts opened has increased at a significant pace. Panel A of Figure 1 presents time series data on the

number of JDY accounts opened. Starting with approximately 54 million accounts at the end of September 2014, the total number of accounts opened has been growing at a monthly rate of 14% and has reached approximately 282 million accounts as of March 29, 2017. The largest fraction of these accounts have been opened by the public sector (state-owned banks), followed by regional rural banks and finally privately-owned banks. Similarly, total deposits (Figure 1, panel (b)) from INR 42 billion to INR 630 billion. The number of debit cards issued has gone up from about 19 million as of September 2014 to 220 million as of March 2017, representing a monthly growth of about 34% (Figure 1, panel (c)). State-owned banks have opened a significant fraction of the new accounts opened as well as debit cards issued.

We also find that the fraction of accounts with a positive balance (Figure 1, panel (d)) has been growing over time – from around 23.0% in September 2014 to about 75% as of 29th March 2017. In our sample – covering the first ten months of the program – about 36% of accounts maintain a positive balance. This fraction is higher (44%) for accounts that are more than six months old. This is comparable to the aggregate average of 44% of accounts with a positive balance as of May 2015. Since then, the percentage of JDY accounts with a positive balance nationally has gone up to 77%.

From panel (d) of Figure 1, we also learn that there is cross-sectional variation in the number of positive-balance accounts across the type of banks. The fraction of users with a positive balance seems to be highest for rural banks, followed by state-owned banks. The higher fraction of positive balance accounts in rural banks suggests that there may have been latent demand for banking services in rural areas, but banks may not have found it profitable to provide services to the unbanked population. In contrast, the fraction of positive balance accounts is the lowest for the private sector banks. The unbanked living in urban areas may have easier access to banks, but given their low income and saving may not have found it optimal to open a bank account. To the extent that private banks cater to urban areas, they may have opened more JDY accounts with zero balance.

II.D.2 Usage of Banking Products under the Program (Intensive Margin)

While a large number of accounts have been opened and an economically significant number of consumers maintain some savings in these accounts, initial usage of these accounts remains quite low. Figure B1 reported in Online Appendix B presents the summary for the frequency of four kinds of banking transaction performed by consumers: Cash Deposits (Panel (a)), Cash Withdrawals (Panel (b)), Inward Remittances (Panel (c)), and Outward Remittances (Panel (d)) during first six months since an account opening. Panel A suggests that around 81% of the consumers in our sample do not deposit any money after account opening. About 12% of individuals perform one deposit transaction, and about 7% perform two or more deposit transactions. The statistics are qualitatively similar for cash withdrawals, with approximately 87% of the sample not withdrawing cash, about 5% withdrawing cash only once, and about 8% withdrawing cash two or more times.

Focusing on panels (c) and (d) of Figure B1 reported in Appendix B, we learn that remittance

seems to be the most common transaction performed by the individuals in our sample. This suggests that remittances are important for low-income individuals in India. This is not surprising given that many workers in India migrate to other states away from their families for employment (Banerjee and Duflo (2007), Morten and Oliveira (2016)). Thus, the increase in ease and reduced transaction costs of remittances through the JDY bank account may be an important benefit of the program. In percentage terms, approximately 34% (see panel (c)) of individuals receive money in their account via inward remittance. When we look at the distribution of the number of such transactions, we find that 17% of individuals receive inward remittance only once during the first six months since an account opening, while about 17% receive remittance two or more times. Similarly, about 21% of account holders send remittance (see panel (d)) at least once. With regards to the number of transactions, about 15% of individuals send remittances only once, while about 8% send remittances two or more times. However, the percentage of heavy users, that is, those performing these transactions ten or more times, is extremely low at less than 1%.

Next, we explore heterogeneity in the usage of these accounts. In Figure B2 reported in Online Appendix B, we present the frequency of usage by males and females in our sample. Similar to our baseline summary Figure B1, we find that overall usage is low for both males and females. However, the frequency of cash withdrawal and remittance transactions is relatively higher for males as compared to females. In Figure B3 reported in Online Appendix B, we split our sample into married and non-married account holders. Here, we find that the frequency of banking transactions is significantly higher for married individuals. For instance, we find that the proportion of married consumers performing at least one deposit transaction is substantially higher at 36% (Figure B3, panel (a)) compared to the sample average of 19%. In comparison, only 16% of the unmarried individuals performed one or more deposit transactions. In terms of the frequency of usage, the proportion of married individuals with just one deposit transaction is 23%. In contrast, the frequency of those with two or more transactions is approximately 10%. The proportion of individuals performing at least one cash withdrawal transaction is also higher for the married consumers at 9% (Figure B3, panel (b)) compared to 3.5% for the unmarried sample.

The difference in relative terms for remittances is not as stark as that for deposit transactions. In essence, we do not find a significant difference between married and unmarried individuals in terms of inward remittances (Figure B3, panel (c)). However, again the fraction of married individuals performing at least one outward remittance transaction is significantly higher at 37% compared to 18% for the unmarried sample (Figure B3, panel (d)). In Figure B4 reported in Online Appendix B, we report these statistics after splitting the sample into four quartiles based on the age of the account holder. Here we find that usage remains similar across individuals of different age groups. Overall, we learn from these figures that the usage of the JDY bank accounts is initially quite infrequent. Marital status appears to be the most significant characteristic among the ones we consider with regards to predicting the likelihood of relatively high usage of these accounts.

III Multivariate Account Level Analysis

III.A Empirical Methodology

In our micro-level analysis, we are interested in assessing behavior of individuals – such as usage patterns and cross-sectional heterogeneity in usage. We focus on the period that spans 10 months after the commencement of the program and examine the evolution of usage over time. Formally, we use the following regression specification:

$$Y_{it} = \beta_0 + \beta_1 \times Age_{it} + \theta_i + \text{Account Opening Month}_t + \epsilon_{it} \quad (1)$$

Y_{it}, is a bank account related outcome variable for individual i at time t (year-month). The coefficient of interest is β_1, which captures the monthly change in outcome variables for the JDY accounts. We also include account opening month fixed effects to control for potential seasonality and customer fixed effects (θ_i) to control for time-invariant differences across individuals.

Next, we examine whether and how customer demographics and regional characteristics explain differences in usage across JDY account holders. Formally, we use the following regression specification:

$$Y_{it} = \beta_0 + \beta_1 \times Char_{it} + \beta_2 \times Age_{it} + \beta_3 Char_{it} \times Age_{it} + \theta_i + \text{Account Opening Month}_t + \epsilon_{it} \quad (2)$$

Y_{it}, is a bank account related outcome variable for individual i, at time t (year-month). $Char_{it}$ refers to a region-specific characteristic such as crime rate or customer group (JDY vs. non-JDY). β_1 captures the baseline time-invariant difference between cross-sectional groups. As in equation 1, age is the number of months since account opening. Thus, β_2 captures the differences in account usage over time. The coefficient of interest is β_3, which captures the monthly change in outcome variables for the individuals in one group relative to the other groups.

III.B Average Monthly Balance and Overall Account Usage

We begin our formal analysis by analyzing the average monthly balance maintained and financial transactions performed by individuals with JDY accounts. More specifically, we focus on understanding the dynamics of usage of banking services over time in a regression framework.

We estimate the specification outlined in equation (1). Table 2, panel A reports our results based on our analysis of monthly account balance maintained and overall account usage by individuals. The dependent variable in these tests is the average monthly balance in column (1), the natural log of one plus the account balance in column (2), positive balance dummy in column (3), and positive usage (*Anyusage* from now) dummy in column (4).

The results in column (1) show that the average monthly balance maintained by individuals increases with time since the account opening. The coefficient on Age of Account suggests that the monthly balance maintained by non-JDY sample increases by INR 94 with each month since

opening. This captures the increase in savings over time and is both economically and statistically significant. In column (2), we use the natural log of the monthly account balance as the dependent variable to examine the monthly growth in the account balance. The coefficient estimate of 0.29 translates into a 34% month-on-month increase in the account balance.

Even amongst the low-income households, there can be substantial variation in income, and consequently, in their savings and account balances. So a small fraction of accounts may drive the growth in the monthly balance reported above. Thus, it may be more meaningful to analyze whether or not these individuals maintain some positive balance in their accounts. Consequently, in column (3), we seek to understand whether there is an increase in the proportion of accounts with some positive balance. The dependent variable in these tests is a positive balance dummy, which takes the value one for account-months with a positive balance and zero otherwise. Column (3) shows that the proportion of accounts maintaining a positive balance grows at a rate of 4.5% per month. This represents an approximately 20% monthly increase.[18] This suggests that while many of the JDY individuals may not maintain any balance in their accounts initially, their likelihood of maintaining some savings in these accounts increases as time passes.

In column (4), we investigate the evolution of the banking transactions performed by JDY individuals. The dependent variable (Anyusage dummy) takes the value one for account-months in which some banking transaction was performed and zero otherwise. Consistent with earlier results, we find that the proportion of JDY individuals performing at least one banking transaction in a month increases over time.

Overall, these results indicate a latent demand for banking among the unbanked. Further, the use of banking services gradually increases with time since the account opening. This could be the result of some learning on the part of JDY individuals, likely due to them becoming increasingly familiar with banking services over time. These results suggest that the real effects of JDY may fully manifest over the long term as more individuals gradually start using the account.

III.C Types of Banking Transactions

We next study the type of banking transactions performed on the individual's bank accounts. We begin by analyzing deposit and withdrawal transactions and report these results in panel B of Table 2. We use three variables to capture the deposit and withdrawal transactions: (i) cash deposit amount (cash withdrawal amount) captures the total cash deposit (withdrawal) by an individual in his account in a month, (ii) number of deposit transactions (number of withdrawal transactions) and (iii) a dummy variable, Cash Deposit Dummy (Cash Withdrawal Dummy), that identifies whether an individual performed at least one deposit (withdrawal) transaction in a month. Columns (1) to (3) present the results for deposit transactions.

The coefficient estimate of INR 3.41 in column (1) translates into an approximately 3% increase in the deposit amounts relative to one-month-old JDY accounts in our sample. Focusing on column

[18] About 22% of the one-month-old JDY account-months maintain a positive account balance during our sample period. Thus 4.5% monthly increase represents a 20% relative increase.

(2), we find that number of deposit transactions performed in a month decreases over time. At the same time, column (3) shows that the monthly increase in the probability of performing at least one deposit transaction is 0.4%. This represents an 8% increase in the likelihood of deposits relative to the average likelihood of a deposit transaction (5%) by the JDY individuals during the first month since the account opening. Overall, we find that the JDY account holders make fewer deposit transactions of larger amounts as the account gets old.

We report the results for withdrawal transactions in columns 4–6. Consistent with our results on deposit transactions, we find that both the amount of withdrawal and the likelihood of withdrawal are higher for older accounts. Finally, in Panel C of Table 3, we analyze inward and outward remittances performed by individuals. Consistent with the evidence for deposit and withdrawal transactions, we observe an economically significant dynamic increase in these kinds of transactions.

III.C.1 Who Uses the Accounts?

We next analyze cross-sectional heterogeneity in account usage using equation 2. Table 3 reports these results. Specifically, we examine the average difference in usage across individuals based on gender, age, marital status, and mobile ownership. We are unable to control for customer fixed effects in these tests as it would absorb the time-invariant demographic characteristics.

Focusing on columns 1–8, we find that the average monthly account balance is higher for males, older individuals, married individuals, and those who own mobiles. In columns 9–12 (13–16), we focus on the likelihood of maintaining a positive balance and performing any transaction. We find that marital status and mobile ownership are the strongest predictors of account usage. Specifically, the likelihood of maintaining a positive balance is 13% higher for married individuals and 18.4% higher for individuals who own mobiles. This makes sense as marital status and mobile ownership is likely to be correlated with income. Consistent with this idea, we also find that the likelihood of performing some banking transactions is 8.5% (12.1%) higher for married individuals (those who own mobiles).[19,20]

[19] In unreported tests, we examine cash deposit and withdrawal transactions and find qualitatively similar results. In subsequent tests, we only report our analysis of Average monthly balance, Positive balance dummy, and *Anyusage* dummy in the interest of brevity. The analysis of cash deposit, withdrawal, and remittance transactions, although unreported, are available from authors on request.

[20] In Table B1 reported in Appendix B, we control for customer fixed effects and examine the cross-sectional heterogeneity in the dynamics of usage over time. We find that, on average, the increase in monthly balance is higher for males. However, the rate of monthly growth in the account balance is lower for males relative to females suggesting that the balance maintained by females should converge over time to that of male account holders. Along similar lines, the fraction of accounts maintaining a positive balance also grows at a higher rate for females. In contrast, age, marital status, and mobile ownership are positively correlated with the monthly increase in the account balance and the fraction of accounts that maintain some savings in their account.

III.D Determinants of Demand for Banking

III.D.1 Safekeeping Role of Bank Accounts

One benefit of having access to a savings bank account is that it provides security from crimes such as theft and robbery. To examine if safekeeping is an important determinant of usage among the unbanked, we exploit variation in regional crime rates as a proxy for the demand for safekeeping. One would expect that individuals in regions with higher theft (including robbery) rates may be inclined to keep their money in bank accounts.

To examine this thesis, in Table 4, we split our sample of the JDY accounts into those in districts with above (high-crime districts) and below (low-crime districts) median pre-program crime per capita and examine account usage over time. Crime includes all reported instances of theft, robbery, and burglary during the year 2013. We control for State FEs, account opening month FEs and customer characteristics in these tests. We are unable to control for customer fixed effects as it would absorb the pre-program high per capita crime dummy. Also, we were unable to match 106 districts in the JDY sample with the crime dataset. Hence, the sample size is relatively smaller in Table 4 as compared to Tables 2 and 3.

Consistent with our conjecture, we find that the fraction of accounts maintaining a positive balance is 3.4 percentage points higher for account holders in high-crime districts. In a similar vein, the fraction of accounts performing a financial transaction is also 3.5 percentage points higher for account holders in high-crime districts. Importantly, focusing on the coefficient on the interaction term (*Account Age X High Theft Dummy*), we find that the rate of a monthly increase in the account balance and likelihood of maintaining a positive balance is higher for accounts in areas with higher crime rates.[21] In column 4, we find that the likelihood of performing a banking transaction in a month decreases over time for users in crime-prone areas. However, we find that this estimate is not robust to controlling for customer fixed effects in Table B2 of the Appendix B.

Overall, these results are consistent with the importance of the safekeeping role of savings bank accounts for the low-income population.

III.D.2 Liquidity management

This section seeks to examine if access to savings bank accounts allows ex-ante unbanked individuals to manage their liquidity needs. Specifically, we examine if individuals build a savings buffer during periods of positive income shock and draw down their deposits during times of increased liquidity need. To generate exogenous variation in the timing of income shock, we follow prior literature that highlights that positive rainfall shocks in India are associated with an increase in nominal wages (Kaur (2019); Adhvaryu et al. (2013)). Good rainfall is associated with an increase in agricultural yield, which results in an increase in labor demand and positively affects the local economy, leading to higher wages. We classify a district as experiencing a positive rainfall shock if the rainfall

[21] In columns 1–3 of Table B2 reported in Appendix B, we find that these results remain robust to controlling for customer fixed effects.

intensity in the district is more than one standard deviation above the district's historical mean. Focusing on panel A of Table 5, we find that the account balance of individuals increases on average if they are located in districts that experience a positive rainfall shock.[22] This increase persists for one month after the shock and reverts subsequently. Interestingly, we find that the overall likelihood of performing financial transactions falls during positive rainfall shock months. This is likely because individuals may prefer not to visit a bank during the rainfall season. This effect also persists for up to one month later and diminishes two months later.

In panels B and C of Table 5, we examine JDY accounts usage around periods of increased liquidity needs. To generate exogenous variation in the need for liquidity, we exploit time-variation in Hindu marriage months. Hindu marriage months are based on the Hindu calendar. They are April to May and October to December. Prior literature highlights that marriages are associated with significant expenditure for the poor in India and that such expenses affect household savings and credit decisions (Kumar (2008); Anukriti, Kwon, and Prakash (2022)). While marriages especially entail heavy expenditure on feasts, gifts, and dowry for the bride's family, marriages may affect other households too due to social norms regarding gifts for the groom and bride.[23]

Consistent with the individual using JDY accounts to manage their liquidity needs, in panel B of Table 5, we find that account balance goes up two months prior to Hindu marriage months and starts decreasing in the marriage months. In panel C, we focus on the likelihood of maintaining a positive balance and performing a financial transaction and observe a similar pattern.

Overall, these results suggest that individuals use access to a savings account to manage their savings and liquidity needs. We expect that this would also allow them to smooth their consumption.

III.D.3 Benchmarking Banking Usage with Non-JDY Account Holders

In this section, we compare the usage patterns of JDY account holders with a group of non-JDY account holders. Non-JDY customers self-select to open a bank account and are banked outside the program. The purpose of having a comparison group is two folds. First, we seek to understand how the usage of accounts opened under JDY, who were exogenously treated, differs from the non-JDY group who self-select to open an account outside the program. Second, we want to understand whether or not there is convergence in usage across these two sets of groups over time. To the extent that accounts opened under JDY represent a banking supply shock rather than a demand for banking services by the low-income households, the null hypothesis for these tests is low and infrequent usage of these accounts relative to the non-JDY individuals. Formally, we use the following regression specification:

[22] We were unable to match 87 districts in the JDY sample with the rainfall dataset. Hence, the sample size is relatively smaller in Table 4 as compared to Tables 2 and 3.

[23] In A study of 35 Indian Villages inquiring into the causes of poverty in India, Krishna (2004) shows that households decline into poverty due to high-interest debt associated with health expenses and marriage ceremonies.

$$Y_{it} = \beta_0 + \beta_1 \times JDY_{it} + \beta_2 \times Age_{it} + \beta_3 JDY_{it} \times Age_{it} + \text{Account Opening Month}_t + \theta_{district} + \epsilon_{it} \quad (3)$$

Where the dependent variable, Y_{ijt}, is a bank account related outcome variable for individual i, at time t (year-month). JDY is a dummy variable that takes a value of one if the account is opened under the JDY program and zero for accounts in the non-JDY sample. β_1 captures the baseline time-invariant difference between JDY and non-JDY individuals. Age is the number of months since the account opening. Thus, β_2 captures the differences in account usage over time. The coefficient of interest is β_3, which captures the monthly change in outcome variables for the JDY accounts relative to those in the non-JDY group. We control for customer characteristics which includes the account holder's age, sex, marital status, and mobile ownership. We also include account opening month fixed effects to control for potential seasonality. In untabulated analysis, we repeat these tests after controlling for customer fixed effects and find our estimates to be robust.

Accounts opened for very low-income (Below Poverty Line, BPL) individuals after the commencement of the program would, by definition, be a part of the JDY sample. As we have noted, our comparison group consists of non-JDY accounts opened for low-income individuals who are close to but above the BPL. Recall from Table 1 (Panels C and D) that the average account balance and usage statistics for these non-JDY individuals are higher than for individuals in the JDY sample. To the extent these differences are time-invariant, this should be captured by the coefficient β_1 and would not confound our primary coefficient of interest β_3.

Table 6 reports the results of these tests. Column (1) shows that the average monthly balance maintained by individuals increases with time since account opening. The coefficient on Age of Account in column 2 suggests that the monthly balance maintained by the non-JDY sample increases by 1.6% with each month since opening. This captures the increase in savings over time. However, the magnitude of the increase in the average monthly balance with the age of the account is significantly greater for the JDY sample. The coefficient on the interaction term (β_3) in column (1) shows that relative to the non-JDY sample, the average monthly balance maintained by the JDY sample increases by INR 79 every month subsequent to account opening. This effect is both economically and statistically significant. In percentage terms, INR 79 represents a 16% monthly increase in account balance relative to the average balance maintained in these accounts. Further, focusing on column 2, the monthly growth rate in balance is 28% higher for the JDY sample.

In column (3), we seek to understand whether there is a relative increase in the proportion of JDY accounts with some positive balance. The dependent variable in these tests is a positive balance dummy, which takes the value one for account-months with a positive balance and zero otherwise. Column (3) shows that, relative to the non-JDY accounts, the magnitude of the monthly increase in the proportion of accounts maintaining a positive balance is 3.5% higher for the JDY accounts. Relative to the average proportion of one-month-old JDY accounts maintaining a positive balance during our sample period, this represents an approximately 16% monthly increase.[24]

[24] About 22% of the one-month-old JDY account-months maintain a positive account balance during our sample

Finally, in column (4), we investigate the evolution of the banking transactions performed by JDY individuals. The dependent variable (anyusage dummy) takes the value one for account-months in which some banking transaction was performed and zero otherwise. Consistent with earlier results, we find that relative to the non-JDY individuals, the proportion of JDY individuals performing at least one banking transaction in a month increases over time. Interestingly, as the coefficient on Age of Account suggests, this fraction is decreasing for the non-JDY sample.[25]

The relative magnitude of the estimates suggests that the fraction of JDY accounts maintaining a positive balance (performing a banking transaction) converges overtime to those of similar individuals who were banked outside the program within 17 (30) months since an account opening. How should we interpret these findings? Prior to JDY, approximately 40% of the households had access to bank accounts. The fraction of households with bank accounts went up dramatically to approximately 100% post JDY. Theoretically, low-income households could have opened bank accounts even prior to JDY but may have chosen not to, given that they did not have much use for such services – possibly because they might not have the right financial literacy background. Alternatively, there may be a demand for such services, but it may not have been profitable for banks to provide services to the lower strata. Our tests on initial usage tell us that it is the latter rather than the former which might be at play. This suggests that one of the common concerns – that the program may meet the same fate as previously unsuccessful financial inclusion programs in India may be unwarranted. Having said that, the gradual increase in usage is consistent with the demand of individuals improving as they gradually learn and benefit from access to bank accounts.

IV Credit Access

IV.A Liquidity Management and Credit Access: Micro-level Analysis

Next, we examine if access to JDY account also results in credit access. There can be at least a couple of reasons why banks might increase lending following the introduction of JDY. First, the new capital in the formal banking system by means of JDY deposits could relax the bank capital constraints. This would allow banks to increase lending to their clients. Second, information asymmetries between new customers and lenders (or other costs and frictions in acquiring new customers) may imply that a program like JDY allows banks to meet the demand for credit for some households that previously operated outside of the formal banking sector. Towards this end, we obtain proprietary data regarding the loans granted to the JDY account holders in our sample. About 13,751 loans were granted between August-2014 and March-2017.

We begin by examining cross-sectional heterogeneity in who gets a loan in Table B4 reported in Online Appendix B. The dependent variable in these tests is a loan dummy that takes the value one for customers who received a loan and zero otherwise. Consistent with our evidence on account

period. Thus 3.5% increase represents a 16% relative monthly increase for JDY accounts.

[25] In Table B3 reported in Appendix B, we repeat these tests after controlling for customer fixed and obtain qualitatively similar results. Further, in unreported tests, we also look at deposit and withdrawal transactions and again find that the monthly increase in usage is again higher for JDY account holders than non-jdy account holders.

usage, we find that, on average, masculine gender, age, marital status, and mobile ownership are positively correlated with loan access.

In Table 7, we examine the likelihood of receiving a loan around positive and negative liquidity shocks. In these tests, we restrict our sample to the set of customers who received a loan. The panel starts in the month an individual's account was opened and ends in the month when the individual receives a loan. The dependent variable in these tests is the Loan Grant Dummy that takes the value zero in months after the JDY bank account was opened and switches to one in the month the loan was granted.

In panel A of Table 7, we examine the likelihood of access to credit following a positive liquidity shock. As in section III.D.2, our measure of positive liquidity shock is based on whether the rainfall intensity in the customer's home district is more than one standard deviation above the district's historical mean. At the margin, a positive liquidity shock should reduce the demand for credit. Consistent with this thesis, we find that there is a decrease in the likelihood of receiving a loan during high rainfall months and one month after. Further, this effect diminishes two months following the positive rainfall shock.

In panel B, we examine whether the individuals are more likely to get a loan prior to the Hindu marriage months. As discussed in Section III.D.2, marriages in India are associated with huge expenses (negative liquidity shock). The underlying reasoning for these tests is to examine if individuals are more likely to receive a loan in the period leading up to the marriage season relative to other time periods. Consistent with the idea that access to JDY accounts allows at least some individuals to meet their credit demand, we find that there is an increase in the likelihood of receiving a loan two months prior and one month prior to the Hindu marriage months.

Summarizing, consistent with the evidence presented in Section III.D.2, we find that access to JDY accounts allows some customers to access credit and manage their liquidity shocks.

IV.B Aggregate Credit Outcomes

Our micro-level credit access analysis is based on the proprietary data provided by our bank for loans granted to JDY account holders. In this section, we seek to examine whether our micro-level findings are indeed representative of an overall increase in access to credit for the unbanked. Accordingly, we explore a number of regional outcome variables such as aggregate bank lending, delinquency rates, source and purpose of borrowing around JDY implementation. The broad goal is to inform on the effect of large-scale financial inclusion programs, such as JDY, on aggregate economic outcomes. The challenge in using JDY as an experiment to infer its effect on the larger economy is that the effect may be confounded by other contemporaneous macroeconomic policy changes or time trends. To alleviate such confounds, we exploit regional heterogeneity in the level of financial inclusion just prior to the program.

In particular, we construct four ex-ante measures that capture different dimensions of financial inclusion. Our first main measure is a proxy for bank branch penetration. It captures the average number of adults serviced by one bank branch in an area (Adults per Unit Bank Branch). Our

second measure is based on the idea that private banks are less likely to expand in financially excluded lower-income areas. In contrast, given their mandate to promote social welfare, state-owned branches are more likely to open branches in such areas. Hence, we use the percentage of state-owned bank branches as a second proxy for financial inclusion (%State-Owned Branches). Our third measure is the percentage of households without bank accounts (%Without Bank Accounts). It is important to note that the mandate for the first phase of the JDY program was to provide 100% banking access to all households. Finally, we also use a comprehensive district-level measure of financial inclusion, which is annually released by CRISIL, and combines three critical parameters of basic financial services: branch penetration, deposit penetration, and credit penetration into one single metric in the form of an index. It is a relative index that has a scale of 0 to 100, with higher numbers indicating lower levels of financial inclusion. Higher values of all four measures indicate lower levels of financial inclusion.

Because new JDY accounts are more likely to be opened in regions that had lower levels of banking access prior to the program, we can trace out the association between JDY intensity and relative changes in economic outcomes using variation in these ex-ante measures of the program exposure. The idea is to compare economic outcomes in regions that had lower levels of financial/banking access before the program, and therefore also more likely to experience a surge in account openings for the poor under JDY, to regions with higher levels of prior banking access. This approach is broadly similar to that used by Mian and Sufi (2012) and Agarwal et al. (2017) in their studies.

It is of course possible that differential changes in economic outcomes in regions that had lower levels of financial/banking access before the program can reflect some fundamental differences across these regions. This is particularly relevant in our context since areas more exposed to the program are on average much poorer than the less exposed regions.

IV.B.1 Ex-Ante JDY Exposure and Program Intensity

We begin by examining the differences in regional income across areas based on our ex-ante exposure measures. In Table B5, we split our sample into two based on median cuts for our four exposure measures. Each observation represents a district, and there are approximately 621 districts in India. Districts in India are similar to counties in the USA and represent a territorial administrative unit. Consistent with financially excluded areas being those with lower income levels, GDP per capita is lower for districts with higher values of all our four exposure measures. Figure 2 shows regional dispersion of our exposure measures.[26]

We next verify that our ex-ante measures of regional JDY exposure in a state before the program correlate with the subsequent intensity of treatment. The data for these tests spans the period from August 2014 to March 2017. The results from these tests are reported in Table 8. In Panel A,

[26] In principle, our exposure measure based on the percentage of households without bank accounts is perhaps most appropriate for our setting given that the mandate of JDY was to provide banking access to all households. However, we report all our regional estimates using all four measures in the interest of robustness and transparency.

we present the results of a regression in which the dependent variable is Log (# of JDY accounts) opened in each state. We control for Log(GDP) to account for regional differences in output. As we observe, there is a strong positive association between the number of JDY accounts opened and our ex-ante exposure measures. Note that we scaled each exposure measure by its standard deviation. Hence the reported coefficient estimates the change in the dependent variable for one standard deviation change in our exposure measures.

In percentage terms, a one-standard-deviation increase in the number of adults per unit bank branch (about 50% relative increase) is associated with a 91% absolute increase in the number of JDY accounts opened in a district.[27] Similarly, a one standard deviation increase in the fraction of state-owned bank branches, the fraction of households without a bank account, and the financial inclusion index is associated with a 53%, 65%, and 85% increase in the number of JDY accounts opened, respectively.

In Panel B of Table 8, we repeat these tests with the Log (total deposits) in JDY accounts opened in each state as the dependent variable. Again, we observe a positive correlation between the total amounts deposited in JDY accounts in each state and the ex-ante exposures measures. This relationship is both statistically and economically significant. In unreported tests, we repeat this analysis using our proprietary micro-level data. Specifically, we aggregate the number of accounts opened and the total amount deposited in these accounts in each state. These tests also confirm a strong correlation between the ex-post intensity of the program and our four ex-ante exposure measures.

Overall, consistent with our account-level analysis, the above results confirm that the program led to a significant number of account openings. Moreover, we find a stronger intensity of bank deposit inflows in more exposed areas. However, this does not necessarily mean that the program increased the overall number of accounts and deposits. The reason is that the program may have adversely affected the private account activity that would have been undertaken in the absence of the program. To shed light on this issue, we also analyze the evolution of deposits for the non-JDY accounts between 2014-2017 and present the results in Table B6 of the Online Appendix B. As we observe from Table B6, we do not find evidence that the program led to a substitution of private accounts with JDY ones. This suggests that the program did indeed lead to an overall net increase in the number of accounts and the amount of deposits in India.

Overall, we note that the more exposed districts experience a relative increase in deposits after the program starts, relative to less exposed ones. However, the overall increase in deposits associated with the program is economically small relative to the pre-program deposits in the banking sector. The total amount deposited in JDY accounts over the period August 2014 to March 2017 (INR 630 Billion, USD 9 Billion) is approximately 0.8% of the pre-JDY deposits in the banking sector (Approximately INR 80,000 Billion, USD 1,176 Billion). In terms of state-level variation, the ratio of the JDY deposits to total pre-JDY deposits in the banking sector varies from a minimum of

[27] The coefficient estimate of 0.652 in panel A of Table 8 translates into ($e^{0.652}$-1)*100=91% higher number of accounts opened.

about 0.2% to a maximum of 5.5%. We plot the evolution of deposit growth in Panel A of Figure 3 in more and less exposed districts. The districts with the above (below) median value of the pre-program percentage of households without bank accounts are represented by the solid (dashed) line. We observe an overall relative increase in deposit growth in more exposed areas. However, the increase in deposits associated with the program is economically small.

IV.B.2 Bank Lending Growth

In this subsection, we investigate whether there was an aggregate increase in the growth rate of bank lending around the program. Specifically, our micro-level analysis discussed in section IV.A was based on credit granted by one bank to the JDY account holders in our sample. We now do regional analysis to assess if the findings on credit growth are applicable more broadly. Towards this end, we use data obtained from the Reserve Bank of India, which provides information on all banks' aggregate lending at the district level.

Panel B of Figure 3 shows the annual growth rate of bank lending in more and less exposed districts. The districts with the above (below) median value of the pre-program percentage of households without bank accounts are represented by the solid (dashed) line. As we observe, more exposed areas experience a relative increase in bank lending after the program starts, relative to less exposed ones. The increase in lending in regions more exposed to the program is especially high starting in 2015, suggesting that the effect of a large-scale financial inclusion program such as JDY on credit access may manifest with a lag.

To investigate this more formally, we examine if regions with greater exposure to JDY experienced greater bank lending than those with limited exposure. Formally, we estimate the following regression model:

$$Y_i = \beta_0 + \beta_1 \text{Exposure Measure} + \epsilon_i \tag{4}$$

Here i refers to a unique district. Y_i is the difference between the average annual growth in bank lending during the program period (2014-2017) and the annual growth in bank lending in the pre-program period (2011-2014). The coefficient of estimate β_1 is a difference-in-differences estimate that captures the change in the growth rate of bank lending before and after the JDY program in districts with high exposure relative to those with low exposure. We focus on growth rates rather than levels as it requires weaker identification assumptions (Angrist and Krueger (1999), Card and Krueger (2000), Mora and Reggio (2019).

Table 9 shows the estimation results. Panel A reports the results based on total lending by all banks. Consistent with Panel B of Figure 3, we find evidence that districts more exposed to the program experienced a significant relative increase in bank lending. In particular, a one standard deviation increase in the exposure measure is associated with between 1.2 percentage points to 2.6 percentage points annual increase in lending in a district. We note that this substantial increase in bank lending is not implausible, even though the JDY households are very poor on average.

This is because the most affected program areas are also mainly poor regions with much lower levels of pre-program bank lending than less affected areas.[28] Hence the addition of 282 million low-income bank customers could lead to a substantial differential increase in bank lending in low-income regions relative to richer areas without having any meaningful effect on the aggregate level of bank lending in India.

We further validate that our micro-level analysis reported in Table 7 and macro-level analysis reported in Table 9 are consistent with each other. To this end, we obtained aggregate district-level lending data from the bank, which is the data source for our account-level analysis. The data provided by the bank spans the period 2011-2015 and comprises total lending done by the bank at the district level. Further, we also obtained data from the bank on default rates on loans granted Pre- and Post-JDY periods. This allows us to examine whether the credit growth during the JDY period is associated with riskier loans.

In panel B of Table 9, and Table 12, we use our proprietary account level loan data obtained from the bank to examine how aggregate lending at the district level changes around the program. Panel B of Table 9 reports the results based on estimation of equation (1), where the dependent variable is the difference between the average annual growth in total loans extended by our bank after the program started (2014-2015) and the average of the same variable during the pre-program period (2011-2014).

We find a statistically and economically significant increase in credit growth in areas with higher ex-ante exposure to JDY relative to those with lower exposure. Specifically, a one standard deviation increase in the exposure measures is associated with between 1.1 percentage points to 2.9 percentage points annual increase in lending growth. Remarkably, the estimates of the increase in lending based on the proprietary bank data is quite similar to the aggregate increase in lending by all banks reported in panel A of Tables 9.

As noted in section IV.B.1, since the deposits coming in due to JDY were economically small, this meaningful increase in lending is not likely due to additional capital being available to financially constrained banks. Rather, JDY may have allowed banks to meet some households' unmet demand for credit.

V Other Tests

V.A Sources of Borrowing

To provide further credence to our results on the effect of JDY on aggregate lending outcomes, in Tables 10 and 11, we use data from a time-series panel of a household survey conducted by Centre for Monitoring Indian Economy (CMIE) to examine how the fraction of households (at the district level) who borrow changes around the program. The survey started in January 2014 and

[28] For example, the average pre-program bank lending level among bottom 25% regions with the lowest level of financial inclusion are about fourteen times smaller than the average bank lending level among the top 25% of regions with the highest level of financial inclusion.

asks households whether they have borrowed from banks and also other sources. While the survey currently covers 505 districts, it started out with only 432 districts in 2014.

In Panel A of Table 10, we re-estimate the regression equation 4, where the dependent variable is the difference between the average monthly growth in the fraction of people who borrowed from banks after the program started (September 2014 to March 2017) and the average of the same variable during the pre-program period (January 2014 to August 2014). Since JDY was primarily targeted towards the low-income households, in these tests, we restrict our sample to households with below-median income.[29] Consistent with our results discussed in sections IV.B.2, we find that there is a statistically significant increase in the fraction of households that borrow from banks in areas with higher ex-ante exposure to JDY relative to those with lower exposure. Specifically, a one standard deviation increase in the exposure measure is associated with between 0.14 percentage points to 0.26 percentage points increase in growth in the number of households that are able to borrow from banks.[30]

In Panel B of Table 10, we examine whether the reliance of households on non-banking sources of credit changes around the program. These non-bank sources of capital include informal moneylenders, chit funds, friends, family, and other informal sources. To the extent that these informal creditors may engage in high-cost loans, consumers with access to formal bank credit may find it optimal to reduce their borrowings from such sources.[31] Consistent with this idea, we find that one standard deviation increase in the JDY exposure measure is associated with between 0.9 percentage points to 1.7 percentage points decrease in monthly growth in the number of households that borrow from non-banking sources.

Overall, these results confirm that the JDY program is associated with increased credit access for low-income borrowers in ex-ante financially excluded areas. Further, consistent with increasing access to formal credit, we observe a drop in households relying on informal sources for credit.

V.B Purpose of Borrowing

In Table 11, we examine the purpose of household borrowing from banks. Focusing on panel A, we find no evidence of an increase in household borrowing for consumption expenditure. Panel B points to a relative increase in borrowing to fund medical expenditure needs in more program-exposed areas. This suggests that one potential benefit of expanding access to banking services is that it may allow consumers to better cope with uncertain health shocks. These results are consistent with Krishna (2004) who finds that health-related expenditure is one of the key determinants of credit demand for the poor in India.

[29] The results are qualitatively similar if we define low-income households as those that belong to the lowest quartile of the income distribution.

[30] In placebo tests, focusing on high-income households (reported in Table B7 of the Online Appendix B), we do not find any significant impact of the program on the fraction of such households borrowing from banks.

[31] Acknowledging that such lenders often engage in usurious lending practices, various states in India have passed laws to protect consumers. For instance, the Kerela Police initiated legal proceedings against a number of money lenders for charging "a usurious rate of interest from their customers, often up to 20 times more than that set by the Reserve Bank of India (RBI)..." Source: The Hindu [accessed on July 2017]

In panel C of Table 11, we examine whether, consistent with an increase in borrowing from banks for the purpose of medical expenditure, there is also a contemporaneous increase in household medical expenditure. Again, the survey asks detailed questions regarding overall expenses on health. We find that a one standard deviation increase in the JDY exposure measure is associated with between 2.5 percentage points to 4 percentage points growth in health expenses.

To ensure that these results are not driven by spurious correlation, we perform placebo tests for low-income households (see Table B8 in Appendix B) with dependent variables that are unlikely to change as a result of JDY. Specifically, we use the expenditure on overseas education, hobby classes, and private tuition fees. To the extent that these entail high costs, which are unlikely to be covered by the benefits provided by JDY accounts, we should not expect any significant effect of JDY on these outcomes. Consistent with this interpretation, we find no evidence of a relative increase in spending among these categories in more program-exposed areas.

V.C Credit Quality

Since JDY was primarily targeted towards low-income households such an increase in credit could manifest itself as an increase in the riskiness of bank loans. We now explore if this conjecture is borne out in the data.

In particular, we examine defaults on loans approved around the program implementation. The bank that supplied data for the account level analysis also provided the proprietary data for these tests for the period January 2014 to December 2015. The dependent variable in these tests is the difference between the average monthly default rate on newly originated loans during the program period (September 2014 to September 2015), and the average default rate on loans originated just prior to the program period (January 2014 to August 2014). The default rate is defined as the proportion of loans originated in a given month that become 60-day delinquent (panel A) or 90-day delinquent (panel B) within a year from loan origination.

These results are presented in Table 12, panels A and B. We find a relative increase in default rates for loans granted during the post-JDY period in the more program-exposed areas. In percentage terms, the number of new loans granted that become 60-day delinquent (Panel A) increased by between 0.2 to 0.4 percentage points. This is economically significant given that our sample's average 60-day delinquency rate is 2.1 percent. In Panel B, we repeat these tests with 90-day delinquent loans and obtain qualitatively similar results.

We conclude from these tests that banks may have optimally chosen to stay away from the unbanked population due to their low creditworthiness. Despite low creditworthiness, banks may still find it optimal to provide traditional savings bank accounts to the poor if it allows them to attract deposit capital. However, the new capital brought into the banks by means of deposits by JDY account holders is an economically insignificant fraction of the levels of the ex-ante deposits. Thus, profit-seeking banks may not have many incentives to cater to the unbanked population in the absence of regulatory push.

V.D Consumption Smoothing

One potential benefit of access to formal banking services is that the availability of savings accounts may enable consumers to smooth consumption and thus make them more resilient to income shocks (Jack and Suri (2014)). Our micro-level analysis of account usage and credit access discussed in Sections III.D.2 and IV.A hinted towards this possibility. To formally examine this hypothesis, in panels A and B of Table 13, we examine the effect of the JDY program and monthly variation in consumption. We proxy for consumption variation using the standard deviation of monthly expenses on food and cooking fuel. The idea is that if banking access via JDY allowed consumers to smooth consumption, ex-post, we should observe a drop in monthly variability of consumption expenditure. The dependent variable in these tests is the difference between the regional average standard deviation of low-income household consumption expenditure during the post-JDY period (September 2014 to March 2017) and the pre-JDY period (January 2014 to August 2014). Consistent with the above thesis, focusing on both Panels A and B, we observe a significant drop in the standard deviation of monthly household expenditure on food and cooking fuel in more exposed areas relative to less exposed areas.

Given that JDY accounts were associated with both accident and life insurance provisions, these results may be driven by access to insurance and not necessarily access to bank accounts. However, as per finance ministry data, only 1,767 accidental insurance claims and 4,165 life insurance claims were paid as of August 2017.[32] Given that the number of insurance claims is insignificant relative to the number of accounts opened, insurance provision alone is unlikely to fully explain the drop in consumption volatility. Rather, these results are in line with our micro-level analysis discussed in section III.D.2 which document that access to JDY accounts allowed customers to manage their liquidity needs better.

VI Conclusion

In this study, we use the largest financial inclusion program in the world to study the impact of financial access on the unbanked. We find that the program led to over 280 million new bank account openings in India (as of March 2017). These individuals received bank accounts, debit cards, as well as an overdraft facility. About 77% of the new accounts maintain a positive balance, and usage increases over time, with inward and outward remittances being the most common transactions performed by the individuals. Account usage is higher for males, married households, older individuals, and for those who own mobiles. The level and rate of growth in the account balance are higher for accounts in crime-prone regions, suggesting safekeeping is an important determinant of banking usage. We also find evidence of an increase (decrease) in savings around positive (negative) liquidity shock, suggesting that ex-ante unbanked individuals use formal banking access to manage their liquidity needs. While the initial usage may appear quite small, it gradually converges overtime to those of similar individuals who were banked outside the program. This

[32] See https://pmjdy.gov.in/ [accessed on September 2017]

evidence is consistent with learning by individuals that results in an increase in usage over time as they gain familiarity with banking services.

At the aggregate level, we find an increase in lending and default rate on new loans in regions with low ex-ante access to banking services. These results are consistent with banks catering to the new demand for formal banking credit by previously unbanked borrowers. Our results shed light on why banks may have chosen to stay away from catering to the unbanked population. First, the new deposit capital brought into the banking system by JDY account holders is economically insignificant as compared to the ex-ante levels. Second, the increased default rate associated with loans granted during the JDY period suggests that many of the previously unbanked borrowers are less creditworthy compared to previously banked ones.

Our paper has implications for the growing body of work in financial inclusion and policymakers. 40% of the world's population is still unbanked, and governments around the world – such as in Indonesia, Malaysia, the Philippines, Brazil – have been thinking of implementing such inclusion programs. Our results can inform policymakers in these countries by presenting evidence on the usage of banking services by previously unbanked and on the evolution of economic outcomes in the first phase of the large-scale financial inclusion program. In particular, our findings suggest that there could be millions of low-income households that are unbanked despite the fact that they display a fairly comparable demand for banking services as banked households with similar demographics. These benefits to the unbanked notwithstanding, profit-seeking banks may not have many incentives to cater to the unbanked population in the absence of regulatory push. Moreover, our results also suggest that the real impact of financial inclusion programs could manifest over the long term as more and more individuals gradually start using these services.

Finally, we note that our findings do not allow us to assess the overall welfare effects of the program. This is a challenging question since such analysis would require, among other aspects, the assessment of program implementation costs, which are difficult to measure.

References

Adhvaryu, A., A. V. Chari, and S. Sharma (2013). Firing costs and flexibility: evidence from firms' employment responses to shocks in india. *Review of Economics and Statistics 95*(3), 725–740.

Agarwal, S., G. Amromin, I. Ben-David, S. Chomsisengphet, T. Piskorski, and A. Seru (2017). Policy intervention in debt renegotiation: Evidence from the home affordable modification program. *Journal of Political Economy 125*(3), 654–712.

Agarwal, S., G. Amromin, S. Chomsisengphet, T. Landvoigt, T. Piskorski, A. Seru, and V. W. Yao (2017). Mortgage refinancing, consumer spending, and competition: Evidence from the home affordable refinancing program. *Working paper*.

Agarwal, S., S. Chomsisengphet, N. Mahoney, and J. Stroebel (2015). Regulating consumer financial products: Evidence from credit cards. *The Quarterly Journal of Economics 130*(1), 111–164.

Aghion, P. and P. Bolton (1997). A theory of trickle-down growth and development. *The Review of Economic Studies 64*(2), 151–172.

Angrist, J. D. and A. B. Krueger (1999). Empirical strategies in labor economics. In *Handbook of labor economics*, Volume 3, pp. 1277–1366. Elsevier.

Anukriti, S., S. Kwon, and N. Prakash (2022). Saving for dowry: Evidence from rural india. *Journal of Development Economics 154*, 102750.

Ashraf, N., D. Karlan, and W. Yin (2006). Tying odysseus to the mast: Evidence from a commitment savings product in the philippines. *The Quarterly Journal of Economics 121*(2), 635–672.

Bachas, P., P. Gertler, S. Higgins, and E. Seira (2021). How debit cards enable the poor to save more. *The Journal of Finance 76*(4), 1913–1957.

Baker, S. R., R. A. Farrokhnia, S. Meyer, M. Pagel, and C. Yannelis (2021). Income, liquidity, and the consumption response to the 2020 economic stimulus payments. *Review of Finance, Forthcoming*.

Baker, S. R. and C. Yannelis (2017). Income changes and consumption: Evidence from the 2013 federal government shutdown. *Review of Economic Dynamics 23*, 99–124.

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

Banerjee, A. V. and A. F. Newman (1993). Occupational choice and the process of development. *Journal of political economy 101*(2), 274–298.

Black, S. E. and P. E. Strahan (2002). Entrepreneurship and bank credit availability. *The Journal of Finance 57*(6), 2807–2833.

Breza, E., M. Kanz, and L. F. Klapper (2020). Learning to navigate a new financial technology: Evidence from payroll accounts. *Working paper*.

Breza, E. and C. Kinnan (2021). Measuring the equilibrium impacts of credit: Evidence from the indian microfinance crisis. *The Quarterly Journal of Economics 136*(3), 1447–1497.

Bruhn, M. and I. Love (2014). The real impact of improved access to finance: Evidence from mexico. *The Journal of Finance 69*(3), 1347–1376.

Burgess, R. and R. Pande (2005). Do rural banks matter? evidence from the indian social banking experiment. *American Economic Review 95*(3), 780–795.

Callen, M., S. De Mel, C. McIntosh, and C. Woodruff (2019). What are the headwaters of formal savings? experimental evidence from sri lanka. *The Review of Economic Studies 86*(6), 2491–2529.

Card, D. and A. B. Krueger (2000). Minimum wages and employment: a case study of the fast-food industry in new jersey and pennsylvania: reply. *American Economic Review 90*(5), 1397–1420.

Célerier, C. and A. Matray (2019). Bank-branch supply, financial inclusion, and wealth accumulation. *The Review of Financial Studies 32*(12), 4767–4809.

Chopra, Y., N. Prabhala, and P. L. Tantri (2017). Bank accounts for the unbanked: Evidence from a big bang experiment. *Working Paper*.

Cole, S., T. Sampson, and B. Zia (2011). Prices or knowledge? what drives demand for financial services in emerging markets? *The journal of finance 66*(6), 1933–1967.

DeFusco, A. A., S. Johnson, and J. Mondragon (2020). Regulating household leverage. *The Review of Economic Studies 87*(2), 914–958.

Demirgüç-Kunt, A., P. Honohan, and T. Beck (2008). *Finance for all?: Policies and Pitfalls in Expanding Access*. World bank.

Demirguc-Kunt, A., L. Klapper, S. Ansar, and A. Jagati (2017). Making it easier to apply for a bank account: a study of the indian market. *World Bank Policy Research Working Paper 8205*.

Dupas, P., D. Karlan, J. Robinson, and D. Ubfal (2018). Banking the unbanked? evidence from three countries. *American Economic Journal: Applied Economics 10*(2), 257–97.

Dupas, P. and J. Robinson (2013). Why don't the poor save more? evidence from health savings experiments. *American Economic Review 103*(4), 1138–71.

Ganong, P. and P. Noel (2020). Liquidity versus wealth in household debt obligations: Evidence from housing policy in the great recession. *American Economic Review 110*(10), 3100–3138.

Granja, J., C. Makridis, C. Yannelis, and E. Zwick (2022). Did the paycheck protection program hit the target? *Journal of Financial Economics 145*(3), 725–761.

Greenwood, R. and D. Scharfstein (2013). The growth of finance. *Journal of Economic Perspectives 27*(2), 3–28.

Higgins, S. (2020). Financial technology adoption. *Working paper*.

Jack, W. and T. Suri (2014). Risk sharing and transactions costs: Evidence from kenya's mobile money revolution. *American Economic Review 104*(1), 183–223.

Jayaratne, J. and P. E. Strahan (1996). The finance-growth nexus: Evidence from bank branch deregulation. *The Quarterly Journal of Economics 111*(3), 639–670.

Johnson, D. S., J. A. Parker, and N. S. Souleles (2006). Household expenditure and the income tax rebates of 2001. *American Economic Review 96*(5), 1589–1610.

Kaur, S. (2019). Nominal wage rigidity in village labor markets. *American Economic Review 109*(10), 3585–3616.

King, R. G. and R. Levine (1993). Finance and growth: Schumpeter might be right. *The quarterly journal of economics 108*(3), 717–737.

Krishna, A. (2004). Escaping poverty and becoming poor: who gains, who loses, and why? *World development 32*(1), 121–136.

Kumar, S. (2008). Does consumption expenditure determine credit demand in rural north india? *Cornell Policy Review*, 1.

Mayer, C., E. Morrison, T. Piskorski, and A. Gupta (2014). Mortgage modification and strategic behavior: evidence from a legal settlement with countrywide. *American Economic Review 104*(9), 2830–57.

Mian, A. and A. Sufi (2012). The effects of fiscal stimulus: Evidence from the 2009 cash for clunkers program. *The Quarterly journal of economics 127*(3), 1107–1142.

Mian, A. and A. Sufi (2015). *House of debt: How they (and you) caused the Great Recession, and how we can prevent it from happening again*. University of Chicago Press.

Mora, R. and I. Reggio (2019). Alternative diff-in-diffs estimators with several pretreatment periods. *Econometric Reviews 38*(5), 465–486.

Morten, M. and J. Oliveira (2016). The effects of roads on trade and migration: Evidence from a planned capital city. *NBER Working Paper 22158*, 43.

Philippon, T. (2015). Has the us finance industry become less efficient? on the theory and measurement of financial intermediation. *American Economic Review 105*(4), 1408–38.

Prina, S. (2015). Banking the poor via savings accounts: Evidence from a field experiment. *Journal of development economics 115*, 16–31.

Rajan, R. G. and L. Zingales (1998). Financial dependence and growth. *The American Economic Review 88*(3), 559–586.

Thaler, R. H. and S. Benartzi (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of political Economy 112*(S1), S164–S187.

TABLE 1: Summary Statistics – Aggregate Statistics

This table presents descriptive statistics. Panels A reports aggregate statistics regarding minimum wage, poverty, aggregate household bank deposits, and India's GDP per capita for the period before the Jan Dhan Yojna (JDY) program. The data on minimum wage, average wage, balance of payments, and per capita GDP is obtained from the planning commission of India as of February 2014. The poverty line estimates are from the Tendulkar Committee report (2005) constituted by the planning commission of India. Data on pre-JDY aggregate household deposits is as of March 2014 and was obtained from the Reserve bank of India. INR to USD conversion is based on the 2016 nominal exchange rate of INR 68 per USD. Panel B reports key aggregates statistics regarding the JDY program as of March 2017 obtained from the website http://pmjdy.gov.in/Archive maintained by the government of India. Panels C and D present descriptive statistics for the proprietary data on JDY and non-JDY bank account samples obtained from one of the largest banks in India. The JDY sample comprises 1,514,307 accounts opened under the JDY program from August 2014 to May 2015. The non-JDY sample comprises 50,023 regular non-JDY accounts opened during the same period. The number of observations, N, corresponds to account months. Note that account openings are staggered across months. All variables are defined in Appendix A.

Panel A: Pre Jan-Dhan Yojana Statistics		
	(1)	(2)
	USD	INR
Minimum wage per day	1.3	89
Average wage per day	3.7	256.52
Poverty Line (Avg monthly per capita expenditure) - Rural	12	816
Poverty Line (Avg monthly per capita expenditure) - Urban	15	1000
Aggregate household deposits in the Indian banking sector in billions	1170	79,558
India GDP per capita	1,431	97,500

Panel B: Jan-Dhan Yojana Progress Statistics	
Number of accounts opened under JDY in millions	282
Number of Debit Cards issues in millions	220
Total deposits in JDY accounts in INR (in billions)	630

TABLE 1: **Summary Statistics – Account-Level Banking Data**

	Panel C: Cash Amounts					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	JDY Accounts			Non-JDY Accounts		
	N	Mean	SD	N	Mean	SD
Average Monthly Balance	9,167,405	504	4,714	398,981	2,420	13,345
Cash Deposit Amount	9,167,405	142	2,121	398,981	1,757	18,164
Cash Withdrawal Amount	9,167,405	161	2,667	398,981	4,427	25,336
Inward Remittance Amount	9,167,405	282	3,955	398,981	4,429	40,063
Outward Remittance Amount	9,167,405	166	3,597	398,981	1,342	32,448

	Panel D: Usage Dummy					
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	JDY Accounts			Non-JDY Accounts		
	N	Mean	SD	N	Mean	SD
Positive Balance Dummy	9,167,405	0.36	0.48	398,981	0.92	0.26
Cash Deposit Dummy	9,167,405	0.04	0.20	398,981	0.22	0.41
Cash Withdrawal Dummy	9,167,405	0.03	0.17	398,981	0.34	0.47
Inward Remittance Dummy	9,167,405	0.11	0.31	398,981	0.25	0.43
Outward Remittance Dummy	9,167,405	0.04	0.21	398,981	0.27	0.44
Overall Usage Dummy	9,167,405	0.17	0.38	398,981	0.59	0.492

TABLE 2: **JDY Account usage and Account Age: Panel A**

This table reports the estimates from a panel regression model examining the relationship between account usage and account age. Panel A includes Average monthly balance in column (1), Log(1 + Average monthly balance) in column (2), Positive balance dummy in column (3), and Anyusage dummy in column (4). Panel B includes Cash deposit amount in column (1), Cash deposit count in column (2), Cash deposit dummy in column (3), Cash withdrawal amount in column (4), Cash withdrawal count in column (5) and Cash withdrawal dummy in column (6). Panel C includes Inward remittance amount in column (1), Inward remittance count in column (2), Inward remittance dummy in column (3), Outward remittance amount in column (4), Outward remittance count in column (5) and Outward remittance dummy in column (6). The regressions control for Customer Fixed effects and Account Opening Month Fixed Effects. The data spans the period September 2014-May 2015. Account-clustered robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Account Balance and Usage				
VARIABLES	Avg monthly balance	Log(1 + Avg monthly balance)	Positive balance Dummy	Anyusage
	(1)	(2)	(3)	(4)
Account Age	94.082***	0.290***	0.045***	0.003***
	(1.015)	(0.001)	(0.000)	(0.000)
Observations	9167405	9142339	9167405	9167405
R-squared	0.598	0.806	0.810	0.437
Customer FE	Y	Y	Y	Y
Account Open Month FE	Y	Y	Y	Y

TABLE 2: **JDY Account usage and Account Age: Panels B and C [Continued]**

Panel B: Cash Deposit and Withdrawal						
VARIABLES	Cash deposit Amount	Cash deposit Count	Cash deposit Dummy	Cash withdrawal Amount	Cash withdrawal Count	Cash withdrawal dummy Dummy
	(1)	(2)	(3)	(4)	(5)	(6)
Account Age	3.412***	-0.004***	0.004***	24.036***	0.010***	0.005***
	(0.456)	(0.000)	(0.000)	(0.534)	(0.000)	(0.000)
Observations	9167405	9167405	9167405	9167405	9167405	9167405
R-squared	0.299	0.378	0.319	0.303	0.455	0.419
Customer FE	Y	Y	Y	Y	Y	Y
Account Open Month FE	Y	Y	Y	Y	Y	Y

Panel C: Inward and Outward remittance						
VARIABLES	Inward remittance Amount	Inward remittance Count	Inward remittance Dummy	Outward remittance Amount	Outward remittance Count	Use Outward remittance Dummy
	(1)	(2)	(3)	(4)	(5)	(6)
Account Age	4.858***	0.003***	0.003***	6.173***	0.002***	0.001***
	(0.774)	(0.000)	(0.000)	(0.706)	(0.000)	(0.000)
Observations	9167405	9167405	9167405	9167405	9167405	9167405
R-squared	0.322	0.445	0.424	0.291	0.459	0.331
Customer FE	Y	Y	Y	Y	Y	Y
Account Open Month FE	Y	Y	Y	Y	Y	Y

TABLE 3: **Relationship between Account usage and Account holder characteristics**

This table reports the estimates from a panel regression model examining the relationship between account usage and account holder characteristics. The dependent variables include Average monthly balance in columns (1), (2), (3) and (4), Log(1 + Average monthly balance) in columns (5), (6), (7) and (8), Positive balance dummy in columns (9), (10), (11) and (12), and Anyusage dummy in columns (13), (14), (15) and (16). The regressions control for District Fixed effects and Account Opening Month Fixed Effects. The data spans the period August 2014-May 2015. Account-clustered robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Account Usage and Balance															
VARIABLES	Avg monthly balance				Log(1 + Avg monthly balance)				Positive balance Dummy				Anyusage			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Account Age	90.379***	58.241***	86.620***	78.306***	0.293***	0.239***	0.291***	0.288***	0.046***	0.041***	0.046***	0.046***	0.001***	-0.003***	0.003***	0.004***
	(1.595)	(1.867)	(0.784)	(0.730)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	30.062***				0.105***				0.010***				0.017***			
	(7.999)				(0.005)				(0.001)				(0.001)			
Age Quartile		-1.114				0.006***				0.002***				0.004***		
		(1.712)				(0.001)				(0.000)				(0.000)		
Married Dummy			189.830***				0.814***				0.131***				0.085***	
			(29.458)				(0.010)				(0.002)				(0.001)	
Mobile Dummy				283.561***				1.179***				0.184***				0.121***
				(20.449)				(0.008)				(0.001)				(0.001)
Observations	9164773	9167405	9167405	9167405	9139724	9142339	9142339	9142339	9164773	9167405	9167405	9167405	9164773	9167405	9167405	9167405
R-squared	0.005	0.005	0.006	0.007	0.075	0.077	0.080	0.089	0.072	0.073	0.075	0.083	0.025	0.027	0.025	0.029
District FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Account Open Month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

TABLE 4: **Safekeeping role of banking: Crime and Account usage**

This table reports the estimates from a panel regression model examining the relationship between account usage and district level crime. High per capita crime dummy takes value 1 for districts with above median value of total instances of theft, robbery, and burglary per capita. We were unable to match 106 districts in the JDY sample with the rainfall dataset. Hence, the sample size is relatively smaller in Table 4 as compared to Tables 2 and 3. The outcome variable is Average monthly balance in column (1), Log(1 + Average monthly balance) in column (2), Positive balance dummy in column (3), and Anyusage dummy in column (4). These regressions control for State Fixed effects and Account Opening Month Fixed Effects. Other controls include gender, age of account holder, marital status and mobile ownership status. The data spans the period August 2014-May 2015. Account-clustered robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

VARIABLES	Avg monthly balance	Log(1 + Avg monthly balance)	Positive balance dummy	Anyusage
	(1)	(2)	(3)	(4)
High Per Capita Crime Dummy	0.472	0.223***	0.034***	0.035***
	(10.188)	(0.007)	(0.001)	(0.001)
Account Age	89.144***	0.287***	0.046***	0.003***
	(1.861)	(0.001)	(0.000)	(0.000)
High Per Capita Crime Dummy × Account Age	8.802***	0.007***	0.001**	-0.002***
	(2.706)	(0.001)	(0.000)	(0.000)
Observations	7233538	7213511	7233538	7233538
R-squared	0.005	0.063	0.057	0.022
State FE	Y	Y	Y	Y
Account Open Month FE	Y	Y	Y	Y
Customer Characteristics	Y	Y	Y	Y

TABLE 5: **Liquidity Management and Account usage: Panel A**

This table investigates if JDY account holders use the bank accounts to manage their liquidity needs. Panel A reports the estimates from a panel regression model examining relationship between account usage and rainfall shocks. Positive rainfall shock is a dummy variable that takes value one when rainfall in a district is more than one standard deviation above the district's historical mean. Similarly, "one month after positive rainfall shock" is a dummy variable takes the value one in the month following the positive rainfall shock. We were unable to match 87 districts in the JDY sample with the rainfall dataset. Hence, the sample size is relatively smaller in Table 5 as compared to Tables 2 and 3. The outcome variables in panel A include Average monthly balance in columns (1)-(3), Log(1 + Average monthly balance) in columns (4)-(6), Positive balance dummy in columns (7)-(9), and Anyusage dummy in columns (10)-(12). Panels B and C report the estimates from a panel regression model examining the relationship between account usage and Hindu Marriage months. Hindu marriage months are based on the Hindu calendar. These are April to May and October to December. Hindu Marriage Months dummy takes value one in these months and zero otherwise. The outcome variables include Average monthly balance in columns (1)-(5) of Panel B, Log(1 + Average monthly balance) in columns (6)-(10) of Panel B, Positive balance dummy in columns (1)-(5) of Panel C, and Anyusage dummy in columns (6)-(10) of panel B. The regressions control for Customer Fixed effects and Account Opening Month Fixed Effects. The data spans the period August 2014-May 2015. Account-clustered robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Rainfall Shock												
VARIABLES	Avg monthly balance			Log(1 + Avg monthly balance)			Positive balance Dummy			Anyusage		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Positive rainfall shock	38.435***			0.102***			0.015***			-0.003***		
	(2.969)			(0.001)			(0.000)			(0.000)		
One month After Positive rainfall shock		10.931***			0.032***			0.002***			-0.009***	
		(3.693)			(0.002)			(0.000)			(0.000)	
Two months After Positive rainfall shock			-29.537***			-0.076***			-0.015***			0.000
			(4.997)			(0.002)			(0.000)			(0.000)
Observations	7627626	7625271	7625271	7605717	7603379	7603379	7627626	7625271	7625271	7627626	7625271	7625271
R-squared	0.585	0.585	0.585	0.806	0.806	0.806	0.810	0.810	0.810	0.437	0.438	0.437
Customer FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Account Open Month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

TABLE 5: **Liquidity Management and Account usage: Panels B and C**

Panel B: Account Balance										
VARIABLES	Avg monthly balance					Log(1 + Avg monthly balance)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Two month Before Hindu Marriage Month	20.927***					0.114***				
	(2.673)					(0.001)				
One month Before Hindu Marriage Month		40.408***					0.180***			
		(2.308)					(0.001)			
Hindu Marriage Months			-12.603***					-0.032***		
			(3.914)					(0.001)		
One month After Hindu Marriage Month				-43.240***					-0.181***	
				(2.417)					(0.001)	
Two month After Hindu Marriage Month					-22.795***					-0.146***
					(3.466)					(0.002)
Observations	9162420	9162420	9162420	9162420	9162420	9137355	9137355	9137355	9137355	9137355
R-squared	0.598	0.598	0.598	0.598	0.598	0.807	0.807	0.806	0.807	0.807
Customer FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Account Open Month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Panel C: Account Usage										
VARIABLES	Positive balance Dummy					Anyusage				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Two month Before Hindu Marriage Month	0.018***					0.001***				
	(0.000)					(0.000)				
One month Before Hindu Marriage Month		0.022***					-0.012***			
		(0.000)					(0.000)			
Hindu Marriage Months			-0.012***					-0.005***		
			(0.000)					(0.000)		
One month After Hindu Marriage Month				-0.026***					0.005***	
				(0.000)					(0.000)	
Two month After Hindu Marriage Month					-0.013***					0.013***
					(0.000)					(0.000)
Observations	9162420	9162420	9162420	9162420	9162420	9162420	9162420	9162420	9162420	9162420
R-squared	0.811	0.810	0.810	0.811	0.810	0.437	0.437	0.437	0.437	0.437
Customer FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Account Open Month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

TABLE 6: **JDY Account Level Analysis of Average Balance and Usage Relative to Non-JDY Accounts**

This table reports the coefficient estimates from the following regression model:

$$Y_{it} = \beta_0 + \beta_1 \times JDY_{it} + \beta_2 \times Age_{it} + \beta_3 JDY_{it} \times Age_{it} + \text{Account Opening Month}_t + \theta_{district} + \epsilon_{it}$$

Where i refers to unique bank account, and t refers to year-month. The outcome variable is Average monthly balance in column (1), Log(1 + Average monthly balance) in column (2), Positive balance dummy in column (3), and Anyusage dummy in column (4). Age is the number of months since account opening. JDY is a dummy variable that identifies accounts opened under the JDY program. The excluded category is non-JDY accounts. The regressions control for District Fixed effects and Account Opening Month Fixed Effects. Other controls include gender, age of account holder, marital status and mobile ownership status. The data spans the period August 2014-May 2015. Account-clustered robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

VARIABLES	Avg monthly balance	Log(1 + Avg monthly balance)	Positive balance Dummy	Anyusage
	(1)	(2)	(3)	(4)
JDY	-2024.246***	-4.255***	-0.598***	-0.311***
	(76.831)	(0.015)	(0.002)	(0.002)
Age of Account	8.630	0.016***	0.006***	-0.007***
	(13.699)	(0.002)	(0.000)	(0.000)
Age of Account X JDY	78.976***	0.247***	0.035***	0.010***
	(13.656)	(0.002)	(0.000)	(0.000)
Observations	9512713	9487569	9512713	9512713
R-squared	0.011	0.144	0.126	0.064
District FE	Y	Y	Y	Y
Account Open Month FE	Y	Y	Y	Y
Customer Characteristics	Y	Y	Y	Y

TABLE 7: Liquidity Management and Credit Access

This table reports the estimates from a panel regression model examining the relationship between liquidity needs and credit access. Panels A and B report the results based on rainfall shock and marriage shock respectively. Loan Grant Dummy takes value one in the month the loan was taken. Positive rainfall shock dummy takes value one when rainfall in a district is more than one standard deviation above the district historic mean. Similarly, "one month after positive rainfall shock" is a dummy variable takes the value one in the month following the positive rainfall shock. Hindu marriage months are based on the Hindu calendar. They are April to May and October to December. Hindu Marriage Months dummy takes value one in these months and zero otherwise. The regressions control for Customer Fixed effects and Account Opening Month Fixed Effects. The data spans the period August 2014-March 2017. Account Level-clustered robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

VARIABLES	Loan Grant Dummy		
	(1)	(2)	(3)
Panel A: Rainfall Shock			
Positive rainfall shock	-0.007***		
	(0.001)		
One month After Positive rainfall shock		-0.006***	
		(0.001)	
Two months After Positive rainfall shock			-0.001
			(0.001)
Observations	182960	182960	182960
R-squared	0.067	0.067	0.067
Panel B: Marriage Shock			
Two months Before Hindu Marriage Month	0.007***		
	(0.001)		
One month Before Hindu Marriage Month		0.006***	
		(0.001)	
Hindu Marriage Months Dummy			0.002*
			(0.001)
Observations	182960	182960	182960
R-squared	0.070	0.070	0.069
Customer FE	Y	Y	Y
Account Open Month FE	Y	Y	Y

TABLE 8: Regional Exposure Measures and Ex-Post JDY Intensity

This table reports the OLS estimates where the dependent variable is the Log of Total Number of JDY accounts opened (Panel A) in each state during the period August 2014 to March 2017 and the Log of Total INR Amount Deposited in JDY accounts during the same period (Panel B). The explanatory variable is one of the four following exposure measures: Adults per Unit Bank Branch, % State-Owned Branches, % Households Without Bank Accounts and a comprehensive Financial Inclusion Index. The coefficients reported in this table are scaled such that they measure the change in the dependent variable for one standard deviation increase in the corresponding exposure measure. *, **, and *** indicate significance at 10%, 5%, and 1% levels respectively. Standard errors are reported in parentheses.

Panel A: Log(# of JDY Accounts)				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	0.652***			
	(0.105)			
%age State-Owned Branches		0.428***		
		(0.148)		
%age Households Without Bank Accounts			0.498***	
			(0.129)	
Financial Inclusion Index				0.616***
				(0.109)
Observations	32	32	32	32
R-squared	0.922	0.859	0.880	0.914

Panel B: Log(Total JDY Deposits)				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	0.504***			
	(0.114)			
%age State-Owned Branches		0.305**		
		(0.144)		
%age Households Without Bank Accounts			0.238*	
			(0.139)	
Financial Inclusion Index				0.447***
				(0.120)
Observations	32	32	32	32
R-squared	0.890	0.840	0.833	0.876

TABLE 9: JDY Exposure and Credit Supply

This table reports the OLS coefficient estimates from the regression model where the dependent variable is the difference in the average annual lending growth rate during the JDY period (2014-2017) and the pre-JDY period (2011-2014). The explanatory variable is one of the four following exposure measures: Adults per Unit Bank Branch, % State-Owned Branches, % Households Without Bank Accounts and a comprehensive Financial Inclusion Index. Panel A reports the results based on total lending at the district-level by all banks. The data is obtained form the Reserve Bank of India for the period 2011-2017. Panel B reports the results on total lending at the district-level by the bank which provided us account-level data for our micro-analysis. This data is proprietary and is available for a limited time period, ending in 2015. The coefficients reported in this table are scaled such that they measure the change in the dependent variable for one standard deviation increase in the corresponding exposure measure. *, **, and *** indicate significance at 10%, 5%, and 1% levels respectively. Standard errors are reported in parentheses.

Panel A: JDY Exposure and Credit Supply (All Banks (RBI Data))				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	0.026***			
	(0.003)			
%age State-Owned Branches		0.027***		
		(0.003)		
%age Households Without Bank Accounts			0.012***	
			(0.004)	
Financial Inclusion Index				0.030***
				(0.004)
Observations	516	516	516	514
R-squared	0.101	0.107	0.021	0.094

Panel B: Regional Exposure Measure and Credit Supply (Proprietary Data for one Bank)				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	0.020***			
	(0.005)			
%age State-Owned Branches		0.016***		
		(0.006)		
%age Households Without Bank Accounts			0.011**	
			(0.006)	
Financial Inclusion Index				0.029***
				(0.005)
Observations	516	516	516	514
R-squared	0.023	0.016	0.008	0.052

TABLE 10: JDY Exposure and Sources of Borrowing

This table reports the OLS coefficient estimates from the regression model where the dependent variable is the difference in the monthly growth in the fraction of households borrowing from banks (Panel A) and other sources (Panel B) during the JDY period (September 2014 to March 2017) and the pre-JDY period (January 2014 to August 2014). The explanatory variable is one of the four following exposure measures: Adults per Unit Bank Branch, % State-Owned Branches, % Households Without Bank Accounts and a comprehensive Financial Inclusion Index. The coefficients reported in this table are scaled such that they measure the change in the dependent variable for one standard deviation increase in the corresponding exposure measure. *, **, and *** indicate significance at 10%, 5%, and 1% levels respectively. Standard errors are reported in parentheses.

Panel A : Growth in Fraction of Households Borrowing From Banks				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	0.00227***			
	(0.00087)			
%age State-Owned Branches		0.00136		
		(0.00085)		
%age Households Without Bank Accounts			0.00304***	
			(0.00085)	
Financial Inclusion Index				0.00263***
				(0.00087)
Observations	415	415	415	415
R-squared	0.016	0.006	0.030	0.022

Panel B : Growth in Fraction of Households Borrowing From Other Sources				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	-0.01455***			
	(0.00294)			
%age State-Owned Branches		-0.01536***		
		(0.00286)		
%age Households Without Bank Accounts			-0.00891***	
			(0.00294)	
Financial Inclusion Index				-0.01730***
				(0.00293)
Observations	415	415	415	415
R-squared	0.056	0.065	0.022	0.078

TABLE 11: JDY Exposure and Purpose of Borrowing: Panels A and B

This table reports the OLS coefficient estimates from the regression model where the dependent variable is the difference in the monthly growth in the fraction of households borrowing from banks from households for consumption expenditure (Panel A) and medical expenditure (Panel B) the JDY period (September 2014 to March 2017) and the pre-JDY period (January 2014 to August 2014). The explanatory variable is one of the four following exposure measures: Adults per Unit Bank Branch, % State-Owned Branches, % Households Without Bank Accounts and a comprehensive Financial Inclusion Index. The coefficients reported in this table are scaled such that they measure the change in the dependent variable for one standard deviation increase in the corresponding exposure measure.. *, **, and *** indicate significance at 10%, 5%, and 1% levels respectively. Standard errors are reported in parentheses.

Panel A: Growth in Fraction of Households Borrowing from Bank for Consumption Expenditure				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	-0.00027			
	(0.00023)			
%age State-Owned Branches		-0.00021		
		(0.00022)		
%age Households Without Bank Accounts			-0.00033	
			(0.00022)	
Financial Inclusion Index				-0.00036
				(0.00023)
Observations	415	415	415	415
R-squared	0.003	0.002	0.005	0.006

Panel B: Growth in Fraction of Households Borrowing from Bank for Medical Expenditure				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	0.00015*			
	(0.00007)			
%age State-Owned Branches		-0.00006		
		(0.00007)		
%age Households Without Bank Accounts			0.00017**	
			(0.00007)	
Financial Inclusion Index				-0.00001
				(0.00008)
Observations	415	415	415	415
R-squared	0.009	0.002	0.013	0.000

TABLE 11: JDY Exposure and Purpose of Borrowing: Panel C

Panel C reports the OLS coefficient estimates from the regression model where the dependent variable is the difference in the monthly growth in household health expenditure during the JDY period (September 2014 to March 2017) and the pre-JDY period (January 2014 to August 2014). The explanatory variable is one of the four following exposure measures: Adults per Unit Bank Branch, % State-Owned Branches, % Households Without Bank Accounts and a comprehensive Financial Inclusion Index. The coefficients reported in this table are scaled such that they measure the change in the dependent variable for one standard deviation increase in the corresponding exposure measure. *, **, and *** indicate significance at 10%, 5%, and 1% levels respectively. Standard errors are reported in parentheses.

Panel C: Expense on Health				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	0.040***			
	(0.008)			
%age State-Owned Branches		0.027***		
		(0.008)		
%age Households Without Bank Accounts			0.025***	
			(0.008)	
Financial Inclusion Index				0.035***
				(0.008)
Observations	415	415	415	415
R-squared	0.055	0.026	0.023	0.043

TABLE 12: **JDY Exposure and Bank Loan Delinquency (Proprietary Bank Data)**

This table reports the OLS coefficient estimates from the regression model where the dependent variable is the difference between the average default rate on newly originated loans during the JDY period minus the average monthly default rate (September 2014 to December 2015) on loans originated just prior to the JDY period (January 2014 to August 2014), where default rate is defined as the proportion of loans originated in a given month that become 60-day delinquent (panel A) or 90-day delinquent (panel B) within a year from loan origination. This data is proprietary and is available for a limited time period and for fewer districts. The explanatory variable is one of the four following exposure measures: Adults per Unit Bank Branch, % State-Owned Branches, % Households Without Bank Accounts and a comprehensive Financial Inclusion Index. The coefficients reported in this table are scaled such that they measure the change in the dependent variable for one standard deviation increase in the corresponding exposure measure. *, **, and *** indicate significance at 10%, 5%, and 1% levels respectively. Standard errors are reported in parentheses.

Panel A: Default Rate - 60-day delinquency				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	0.004***			
	(0.001)			
%age State-Owned Branches		0.001		
		(0.001)		
%age Households Without Bank Accounts			0.004***	
			(0.001)	
Financial Inclusion Index				0.002*
				(0.001)
Observations	439	439	439	438
R-squared	0.03	0.003	0.026	0.006

Panel B: Default Rate - 90-day delinquency				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	0.003***			
	(0.001)			
%age State-Owned Branches		0.001		
		(0.001)		
%age Households Without Bank Accounts			0.003***	
			(0.001)	
Financial Inclusion Index				0.002*
				(0.001)
Observations	439	439	439	438
R-squared	0.027	0.003	0.017	0.007

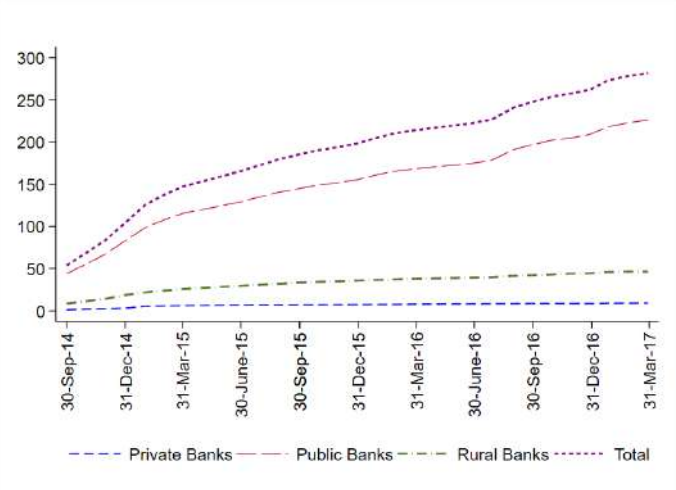
TABLE 13: JDY Exposure and Consumption Smoothing

Panels A and B report the OLS coefficient estimates from the regression model where the dependent variable is the difference in monthly standard deviation of total household food expenditure (cooking fuel expenditure) during the JDY period (September 2014 to March 2017) and the pre-JDY period (January 2014 to August 2014) in panel A (panel B). The explanatory variable is one of the four following exposure measures: Adults per Unit Bank Branch, % State-Owned Branches, % Households Without Bank Accounts and a comprehensive Financial Inclusion Index. The coefficients reported in this table are scaled such that they measure the change in the dependent variable for one standard deviation increase in the corresponding exposure measure. *, **, and *** indicate significance at 10%, 5%, and 1% levels respectively. Standard errors are reported in parentheses.

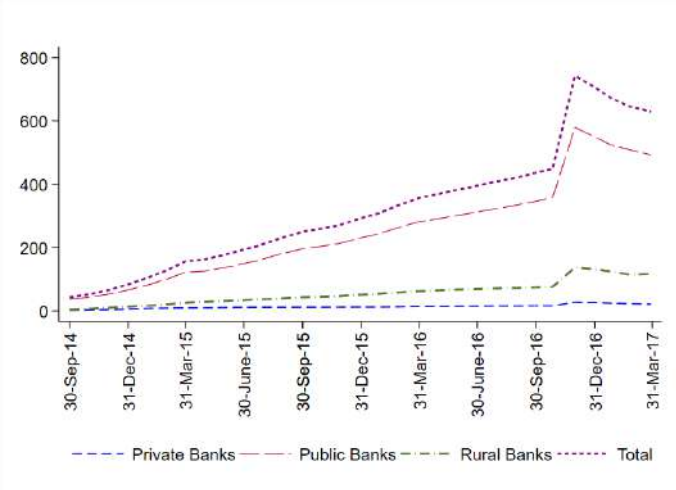
Panel A : Standard Deviation of Food Expenditure				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	-56.44584***			
	(13.97924)			
%age State-Owned Branches		-69.86693***		
		(13.83272)		
%age Households Without Bank Accounts			-22.03077	
			(14.20956)	
Financial Inclusion Index				-90.41337***
				(13.94435)
Observations	415	415	415	415
R-squared	0.038	0.058	0.006	0.092
Panel B : Standard Deviation of Cooking Fuel Expenditure				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	-4.62435*			
	(2.61391)			
%age State-Owned Branches		-4.46825*		
		(2.61457)		
%age Households Without Bank Accounts			-7.85288***	
			(2.59532)	
Financial Inclusion Index				-7.67954***
				(2.66752)
Observations	415	415	415	415
R-squared	0.007	0.007	0.022	0.020

Figure 1: **Time Series of Aggregate Extensive Margin Statistics on JDY**

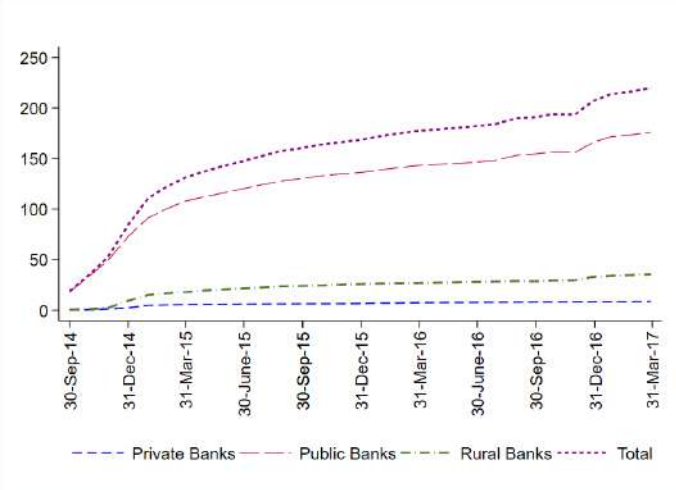
The figure shows time series of aggregate statistics related to the JDY program: Number of Accounts Opened (Panels (a)), Number of Debit Cards issued (Panel (b)), Fraction of accounts with positive balance (panel (c)), and Total amount deposited in JDY accounts (panel (d)). The four lines represent the statistics respectively for JDY accounts opened by private banks, public (government owned) banks, regional rural banks and the total across all banks. The data was obtained from https://pmjdy.gov.in/ website maintained by the government of India for the period September 2014 to March 2017. https://pmjdy.gov.in/ does not report aggregate statistics for the month of August 2014.



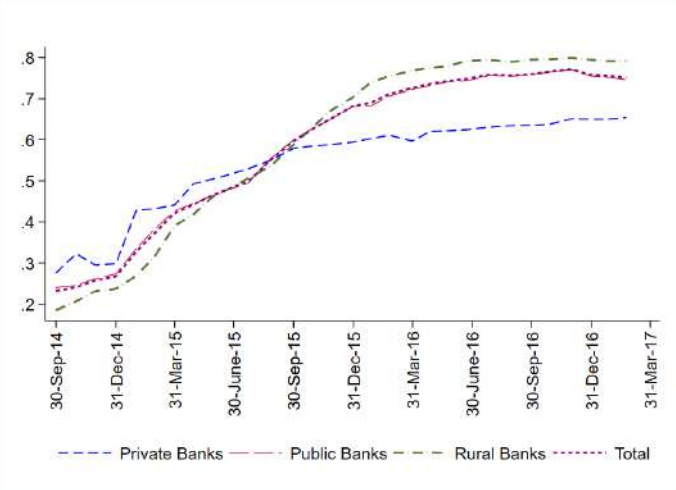
(a) Number Accounts Opened (Million)



(b) Total Deposits in JDY Accounts (INR BIllion)



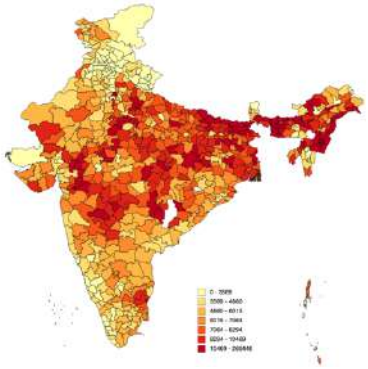
(c) Number Debit Cards Issued (Million)



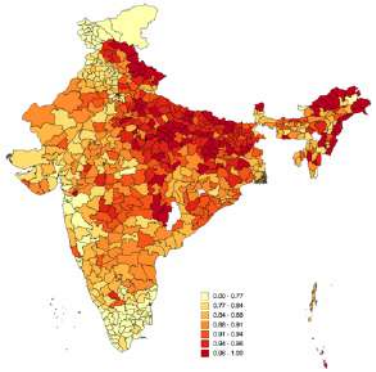
(d) Fraction of Accounts with Positive Balance

Figure 2: **Regional Distribution of Ex-Ante Program Exposure Measures**

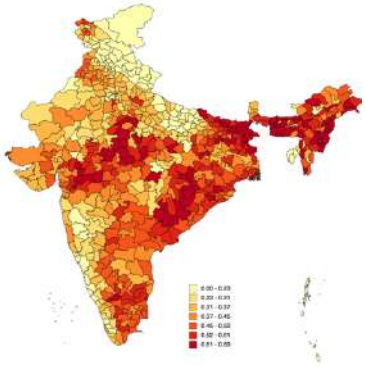
This figure shows the regional distribution of our four program exposure measures in India: Adults per Unit Bank Branch (Panel A, Source: Reserve Bank of India, Period: 2013), % State-Owned Branches (Panel B, Source: Reserve Bank of India, Period: 2013), Fraction of Households Without Bank Accounts (Panel C, Source: Indian Economic Census, Period: 2011) and Financial Inclusion Index (Panel D, Source: CRISIL, Period: 2013). Higher values of all measures indicate lower levels of financial inclusion.



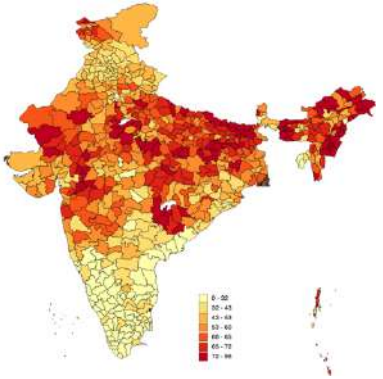
(a) Number of Adults per Unit Bank Branch



(b) % State-Owned Branches

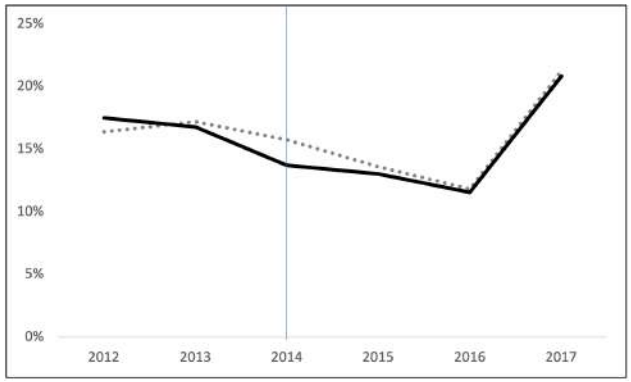


(c) Fraction of Households Without Bank Accounts

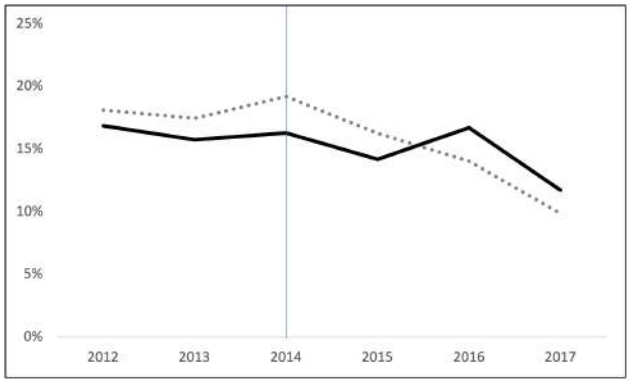


(d) Financial Inclusion Index

Figure 3: **JDY Exposure and Deposit and Bank Lending Growth Rate**
The figure shows the annual percentage growth rates in bank deposits (panel (a)) and bank lending (panel (b)) in the more and less JDY exposed districts (based on the % of Households without Bank Accounts). The districts with the above (below) median value of the pre-program percentage of households without bank accounts are defined as more (less) exposed. The data on aggregate lending and deposits was obtained form the Reserve Bank of India for the period 2011-2017. The more exposed group is represented by the solid line, and the less exposed group is represented by the dashed line.



(a) Deposit Growth Rate



(b) Lending Growth Rate

Appendix A

The Appendix describes some of the key variables used in our regional analysis along with the data source and the time span for which we obtained the data.

Variable Name	Data Source	Time Period
Account Level Data		
Average Monthly Balance	Proprietary Data Source	August 2014 to May 2015
Positive Balance Dummy	Proprietary Data Source	August 2014 to May 2015
Any Usage	Proprietary Data Source	August 2014 to May 2015
Cash Deposit Amount	Proprietary Data Source	August 2014 to May 2015
Cash Deposit Count	Proprietary Data Source	August 2014 to May 2015
Use Deposit Count	Proprietary Data Source	August 2014 to May 2015
Cash Withdrawal Amount	Proprietary Data Source	August 2014 to May 2015
Cash Withdrawal Count	Proprietary Data Source	August 2014 to May 2015
Use Withdrawal Count	Proprietary Data Source	August 2014 to May 2015
Inward Remittance Amount	Proprietary Data Source	August 2014 to May 2015
Inward Remittance Count	Proprietary Data Source	August 2014 to May 2015
Use Inward Remittance Count	Proprietary Data Source	August 2014 to May 2015
Outward Remittance Amount	Proprietary Data Source	August 2014 to May 2015
Outward Remittance Count	Proprietary Data Source	August 2014 to May 2015
Use Outward Remittance Count	Proprietary Data Source	August 2014 to May 2015
Account Age (months)	Proprietary Data Source	August 2014 to May 2015

Male Dummy	Proprietary Data Source	August 2014 to May 2015
Age	Proprietary Data Source	August 2014 to May 2015
Married Dummy	Proprietary Data Source	August 2014 to May 2015
Mobile Dummy	Proprietary Data Source	August 2014 to May 2015
JDY Dummy	Proprietary Data Source	August 2014 to May 2015
Loan Grant Dummy	Proprietary Data Source	January 2014 to March 2017
District Level Variables		
District Per Capita GDP	Indicus Analytics	2013
Crime Per Capita	District-wise crimes: theft, robbery and dacoity under various sections of Indian Penal Code (IPC) crimes during 2013 (data.gov)	2013
Positive rainfall shock dummy	Data from the University of Delaware website and matched to the centroids of the 2001 administrative district boundaries.	1950-2017
Ex-ante Program Exposure (Financial Exclusion) Measures		
Adults Per unit Bank Branch	Reserve Bank of India (RBI)	2013
% State-Owned Branches	Reserve Bank of India (RBI)	2013
% Households Without Bank Accounts	Census 2011 - obtained from IndiaStat	2011
Financial Inclusion Index	CRISIL	2013
Ex-Post JDY Intensity and Outcomes		
Log of Total Number of JDY accounts opened	PMJDY website	August 2014 to March 2017
Log of Total INR Amount Deposited in JDY accounts	PMJDY website	August 2014 to March 2017
Log of Total Deposits in Non-JDY accounts	RBI	August 2014 to March 2017
Change in Annual Growth Rate of Non-JDY Deposits	RBI	August 2014 to March 2017

Difference in the average annual lending growth rate during the JDY period and the pre-JDY period.	RBI and Proprietary Data Source	JDY period: 2011-2014 Post JDY period: 2014-2017
Difference between the average default rate (60-day and 90-day) on newly originated loans during the JDY period minus the average monthly default rate on loans originated just prior to the JDY period.	Proprietary Data Source	JDY period: September 2014 to December 2015 Pre JDY period: January 2014 to August 2014
Sources of Borrowing		
Difference in monthly growth in Fraction of Households Borrowing from Banks during JDY period and the pre-JDY period.	CMIE Consumer Pyramids	JDY period: September 2014 to March 2017 Pre JDY period: January 2014 to August 2014
Difference in monthly growth in Fraction of Households Borrowing from Other Sources during JDY period and the pre-JDY period.	CMIE Consumer Pyramids	JDY period: September 2014 to March 2017 Pre JDY period: January 2014 to August 2014
Purpose of Borrowing		
Difference in monthly growth in Fraction of Households Borrowing from Bank for Consumption Expenditure during JDY period and the pre-JDY period.	CMIE Consumer Pyramids	JDY period: September 2014 to March 2017 Pre JDY period: January 2014 to August 2014
Difference in the monthly growth in household health related expenditure during the JDY period and the pre-JDY period.	CMIE Consumer Pyramids	JDY period: September 2014 to March 2017 Pre JDY period: January 2014 to August 2014
Consumption Smoothing		
Difference in monthly standard deviation of household food expenditure during the JDY period and the pre-JDY period.	CMIE Consumer Pyramids	JDY period: September 2014 to March 2017 Pre JDY period: January 2014 to August 2014

Difference in monthly standard deviation of household cooking fuel expenditure during the JDY period and the pre-JDY period.	CMIE Consumer Pyramids	JDY period: September 2014 to March 2017 Pre JDY period: January 2014 to August 2014
Placebo Household Expenditure Variables		
Difference in the monthly growth in household expenditure on Overseas Education during the JDY period and the pre-JDY period.	CMIE Consumer Pyramids	JDY period: September 2014 to March 2017 Pre JDY period: January 2014 to August 2014
Difference in the monthly growth in household expenditure on Hobby Classes during the JDY period and the pre-JDY period.	CMIE Consumer Pyramids	JDY period: September 2014 to March 2017 Pre JDY period: January 2014 to August 2014
Difference in the monthly growth in household expenditure on Private Tuitions during the JDY period and the pre-JDY period.	CMIE Consumer Pyramids	JDY period: September 2014 to March 2017 Pre JDY period: January 2014 to August 2014

Definitions of Ex-ante Program Exposure (Financial Exclusion) Measures

- *% Adults Per Unit Bank Branch*: the ratio of total number of adults to the total number of bank branches in each state (district).
- *% State-Owned Bank Branches*: the ratio of total number of bank branches that are owned by state-owned banks to the total number of bank branches in each state (district).
- *% Households Without Bank Accounts*: the ratio of total number of households without any bank account to the total number of households in each state (district).
- *Financial Inclusion Index*: a comprehensive district level measure of financial inclusion released by CRISIL, which combines three critical parameters of basic financial services: branch penetration, deposit penetration, and credit penetration into one single metric in the form of an index. It is a relative index that has a scale of 0 to 100 with higher numbers indicating lower levels of financial inclusion. For state level analysis, we use the average value of the index across all districts.

Appendix B

The Appendix reports results of additional tests that are briefly described in the text. Additional details are available from the authors upon request.

TABLE B1: **Account holder characteristics and Account Usage over time**

This table reports the estimates from a panel regression model examining the relationship between account usage and account holder characteristics. The dependent variables include Average monthly balance in columns (1), (2), (3) and (4), Log(1 + Average monthly balance) in columns (5), (6), (7) and (8), Positive balance dummy in columns (9), (10), (11) and (12), and Anyusage dummy in columns (13), (14), (15) and (16). The regressions control for Customer Fixed effects and Account Opening Month Fixed Effects. The data spans the period August 2014-May 2015. Account-clustered robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

VARIABLES	Avg monthly balance				Log(1 + Avg monthly balance)				Positive balance Dummy				Anyusage			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Account Age	90.462***	65.024***	84.414***	75.540***	0.291***	0.203***	0.283***	0.281***	0.046***	0.035***	0.045***	0.045***	0.001***	-0.007***	0.003***	0.004***
	(1.633)	(1.573)	(0.735)	(0.639)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Account Age X Male Dummy	7.121***				-0.003***				-0.001***				0.004***			
	(2.041)				(0.001)				(0.000)				(0.000)			
Account Age X Age Quantile		5.729***				0.017***				0.002***				0.002***		
		(0.399)				(0.000)				(0.000)				(0.000)		
Account Age X Married Dummy			80.677***				0.060***				0.005***				-0.002***	
			(6.563)				(0.002)				(0.000)				(0.000)	
Account Age X Mobile Dummy				88.344***				0.043***				0.002***				-0.006***
				(4.240)				(0.001)				(0.000)				(0.000)
Observations	9159792	9162420	9162420	9162420	9134744	9137355	9137355	9137355	9159792	9162420	9162420	9162420	9159792	9162420	9162420	9162420
R-squared	0.598	0.598	0.598	0.598	0.806	0.807	0.806	0.806	0.810	0.811	0.810	0.810	0.437	0.437	0.437	0.437
Customer FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Account Open Month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

TABLE B2: **Safekeeping role of banking: Crime and Account usage (Customer FEs)**

This table reports the estimates from a panel regression model examining the relationship between account usage and district level crime. High per capita crime dummy takes value 1 for districts with above median value of total instances of theft, robbery, and burglary per capita. We were unable to match 106 districts in the JDY sample with the rainfall dataset. Hence, the sample size is relatively smaller in Table 4 as compared to Tables 2 and 3. The outcome variable is Average monthly balance in column (1), Log(1 + Average monthly balance) in column (2), Positive balance dummy in column (3), and Anyusage dummy in column (4). These regressions control for Customer Fixed effects and Account Opening Month Fixed Effects. The data spans the period August 2014-May 2015. Account-clustered robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

VARIABLES	Avg monthly balance	Log(1 + Avg monthly balance)	Positive balance dummy	Anyusage
	(1)	(2)	(3)	(4)
Account Age	89.373***	0.284***	0.045***	0.003**
	(1.892)	(0.001)	(0.000)	(0.001)
High Per Capita Crime Dummy × Account Age	8.339***	0.015***	0.002***	-0.001
	(2.380)	(0.001)	(0.000)	(0.002)
Observations	7228763	7208736	7228763	7228763
R-squared	0.590	0.804	0.808	0.435
Customer FE	Y	Y	Y	Y
Account Open Month FE	Y	Y	Y	Y

TABLE B3: **JDY Account Level Analysis of Average Balance and Usage Relative to Non-JDY Accounts (Customer FEs)**

This table reports the coefficient estimates from the following regression model:

$$Y_{it} = \beta_0 + \beta_1 \times JDY_{it} + \beta_2 \times Age_{it} + \beta_3 JDY_{it} \times Age_{it} + \text{Account Opening Month}_t + \theta_{district} + \epsilon_{it}$$

Where i refers to unique bank account, and t refers to year-month. The outcome variable is Average monthly balance in column (1), Log(1 + Average monthly balance) in column (2), Positive balance dummy in column (3), and Anyusage dummy in column (4). Age is the number of months since account opening. JDY is a dummy variable that identifies accounts opened under the JDY program. The excluded category is non-JDY accounts. The regressions control for Customer Fixed effects and Account Opening Month Fixed Effects. Other controls include gender, age of account holder, marital status and mobile ownership status. The data spans the period August 2014-May 2015. Account-clustered robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

VARIABLES	Avg monthly balance	Log(1 + Avg monthly balance)	Positive balance Dummy	Anyusage
	(1)	(2)	(3)	(4)
Age of Account	20.382	0.008***	0.005***	-0.004***
	(17.752)	(0.002)	(0.000)	(0.000)
Age of Account X JDY	66.673***	0.258***	0.037***	0.007***
	(17.776)	(0.002)	(0.000)	(0.000)
Observations	9507732	9482589	9507732	9507732
R-squared	0.533	0.817	0.819	0.453
Customer FE	Y	Y	Y	Y
Account Open Month FE	Y	Y	Y	Y

TABLE B4: Relationship between Loans and Account holder characteristics

This table reports the estimates from a panel regression model examining the relationship between account holder characteristics and credit access. The outcome variable is a dummy variable that takes the value one for account holders who received a loan. The regressions control for District Fixed effects and Account Opening Month Fixed Effects. The data spans the period August 2014-March 2017 and consists of loans taken by JDY account holders. Account Level-clustered robust standard errors are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

VARIABLES	Loan Dummy
	(1)
Male	0.006***
	(0.000)
Age Quartile	0.001***
	(0.000)
Married status	0.009***
	(0.001)
Mobile status	0.009***
	(0.001)
Observations	1518865
R-squared	0.021
District FE	Y
Account Open Month FE	Y

TABLE B5: **Regional Summary Statistics**

This table presents summary statistics for the districts in our sample. Panels A and B report the statistics based on the exposure measures, Adults per Unit Bank Branch which is the ratio of total adult population to the total number of bank branches in a district and %age State-Owned Bank Branches in a district which is the ratio of total number of bank branches that are owned by state-owned banks to the total number of bank branches in each district. Higher values for both these measures imply greater ex-ante JDY exposure. We split our sample into two based on these measures. Columns (1) and (2) (columns (3) and (4)) report statistics for the districts with above (below) median value of Adults per Unit Bank Branch (Panel A) and % State-Owned Bank Branches (Panel B). Panels C and D report the statistics based on the exposure measures, % Households Without Bank Accounts, which is the ratio of total number of households without any bank account, to the total number of households in a district and Financial Inclusion Index, which is a comprehensive measure of financial inclusion released by CRISIL based on three critical parameters of basic financial services: branch penetration, deposit penetration, and credit penetration into one single metric in the form of an index. It is a relative index that has a scale of 0 to 100. Higher values for all four measures imply greater (lower) ex-ante JDY exposure (financial inclusion). We split our sample into two based on these measures. Columns (1) and (2) (columns (3) and (4)) report statistics for the districts with above (below) median value of % Households Without Bank Accounts (Panel C) and Financial Inclusion Index (Panel D).

Panel A: Exposure Measure - Adults per Unit Bank Branch				
VARIABLES	(1)	(2)	(3)	(4)
	Above Median		Below Median	
	Mean	SD	Mean	SD
Total Region GDP Per Capita in INR Per Capita	50,051	25,849	195,000	1,530,000
All Industry GDP Per Capita in INR Per Capita	12,511	11,207	40,171	138,000
Agriculture GDP Per Capita in INR Per Capita	7,717	8,769	9,316	9,046
Agriculture and Related Per Capita in INR Per Capita	13,187	10,479	15,815	13,869
Manufacturing GDP Per Capita in INR Per Capita	3,584	5,014	19,823	72,157
Banking GDP Per Capita in INR Per Capita	4,387	2,470	62,272	753,000
Adults Per unit Bank Branch	9,461	2,976	4,390	1,335

Panel B: Exposure Measure - % State-Owned Bank Branches				
VARIABLES	Above Median		Below Median	
	Mean	SD	Mean	SD
Total Region GDP in INR Per Capita	55,414	33,480	190,000	1,530,000
All Industry GDP in INR Per Capita	14,704	16,674	37,867	137,000
Agriculture GDP in INR Per Capita	7,870	9,187	9,157	8,642
Agriculture and Related in INR Per Capita	13,752	11,152	15,236	13,408
Manufacturing GDP in INR Per Capita	3,999	6,668	19,351	72,003
Banking GDP in INR Per Capita	4,975	3,687	61,490	751,000
% State-Owned Branches	94	3	79	13

TABLE B5: **Regional Summary Statistics [continued]**

Panel C: Exposure Measure - % Households Without Bank Accounts				
VARIABLES	(1)	(2)	(3)	(4)
	Above Median		Below Median	
	Mean	SD	Mean	SD
Total Region GDP in INR Per Capita	60,282	29,472	185,000	1,530,000
All Industry GDP in INR Per Capita	16,015	12,469	36,544	138,000
Agriculture GDP in INR Per Capita	8,971	9,558	8,045	8,250
Agriculture and Related in INR Per Capita	14,754	11,645	14,224	13,015
Manufacturing GDP in INR Per Capita	5,528	7,169	17,808	72,251
Banking GDP in INR Per Capita	6,379	6,679	60,073	751,000
% Households Without Bank Accounts	56	10	28	9

Panel D: Exposure Measure - CRISIL Financial Inclusion Index				
VARIABLES	Above Median		Below Median	
	Mean	SD	Mean	SD
Total Region GDP in INR Per Capita	54,556	32,558	189,000	1,520,000
All Industry GDP in INR Per Capita	14,541	16,987	37,806	137,000
Agriculture GDP in INR Per Capita	7,776	9,115	9,238	8,709
Agriculture and Related in INR Per Capita	13,803	11,605	15,171	13007
Manufacturing GDP in INR Per Capita	4,013	7,040	19,190	71,641
Banking GDP in INR Per Capita	4,754	4,332	61,166	748,000
Financial Inclusion Index	68	8	38	14

TABLE B6: JDY Exposure and Total Deposits on Non-JDY Accounts

This table reports the OLS estimates where the dependent variable is the Log of Total Deposits in Non-JDY accounts in each state during the period August 2014 to March 2017. The explanatory variable is one of the four following exposure measures: Adults per Unit Bank Branch, % State-Owned Branches, % Households Without Bank Accounts and a comprehensive Financial Inclusion Index. The coefficients reported in this table are scaled such that they measure the change in the dependent variable for one standard deviation increase in the corresponding exposure measure. *, **, and *** indicate significance at 10%, 5%, and 1% levels respectively. Standard errors are reported in parentheses.

Log(Total Non-JDY Deposits)				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	-0.095			
	(0.066)			
%age State-Owned Branches		-0.028		
		(0.068)		
%age Households Without Bank Accounts			-0.095	
			(0.066)	
Financial Inclusion Index				-0.082
				(0.068)
Observations	32	32	32	32
R-squared	0.957	0.954	0.957	0.956

TABLE B7: JDY Exposure and Sources of Borrowing (Placebo Test with High Income Group)

This table reports the OLS coefficient estimates from the regression model where the dependent variable is the difference in the monthly growth in the fraction of high income households borrowing from banks during the JDY period (September 2014 to March 2017) and the pre-JDY period (January 2014 to August 2014) used as a Placebo. The explanatory variable is one of the four following exposure measures: Adults per Unit Bank Branch, % State-Owned Branches, % Households Without Bank Accounts and a comprehensive Financial Inclusion Index. The coefficients reported in this table are scaled such that they measure the change in the dependent variable for one standard deviation increase in the corresponding exposure measure. *, **, and *** indicate significance at 10%, 5%, and 1% levels respectively. Standard errors are reported in parentheses.

Growth in Fraction of Households Borrowing From Banks				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	-0.00344			
	(0.00223)			
%age State-Owned Branches		0.00309		
		(0.00217)		
%age Households Without Bank Accounts			-0.00228	
			(0.00219)	
Financial Inclusion Index				-0.00015
				(0.00225)
Observations	404	404	404	404
R-squared	0.006	0.005	0.003	0.000

TABLE B8: JDY Exposure and Household Expenditure (Placebo Tests)

This table reports the OLS coefficient estimates from the regression model where the dependent variable is the difference in the monthly growth in the household expenditure on overseas education (panel A), hobby classes (panel B), and private tuitions (panel C) during the JDY period (September 2014 to March 2017) and the pre-JDY period (January 2014 to August 2014). The explanatory variable is one of the four following exposure measures: Adults per Unit Bank Branch, % State-Owned Branches, % Households Without Bank Accounts and a comprehensive Financial Inclusion Index. The coefficients reported in this table are scaled such that they measure the change in the dependent variable for one standard deviation increase in the corresponding exposure measure. *, **, and *** indicate significance at 10%, 5%, and 1% levels respectively. Standard errors are reported in parentheses.

Panel A: Expenses on Overseas Education				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	-0.001			
	(0.002)			
%age State-Owned Branches		0.001		
		(0.002)		
%age Households Without Bank Accounts			-0.002	
			(0.002)	
Financial Inclusion Index				0.001
				(0.002)
Observations	415	415	415	415
R-squared	0.000	0.000	0.003	0.000

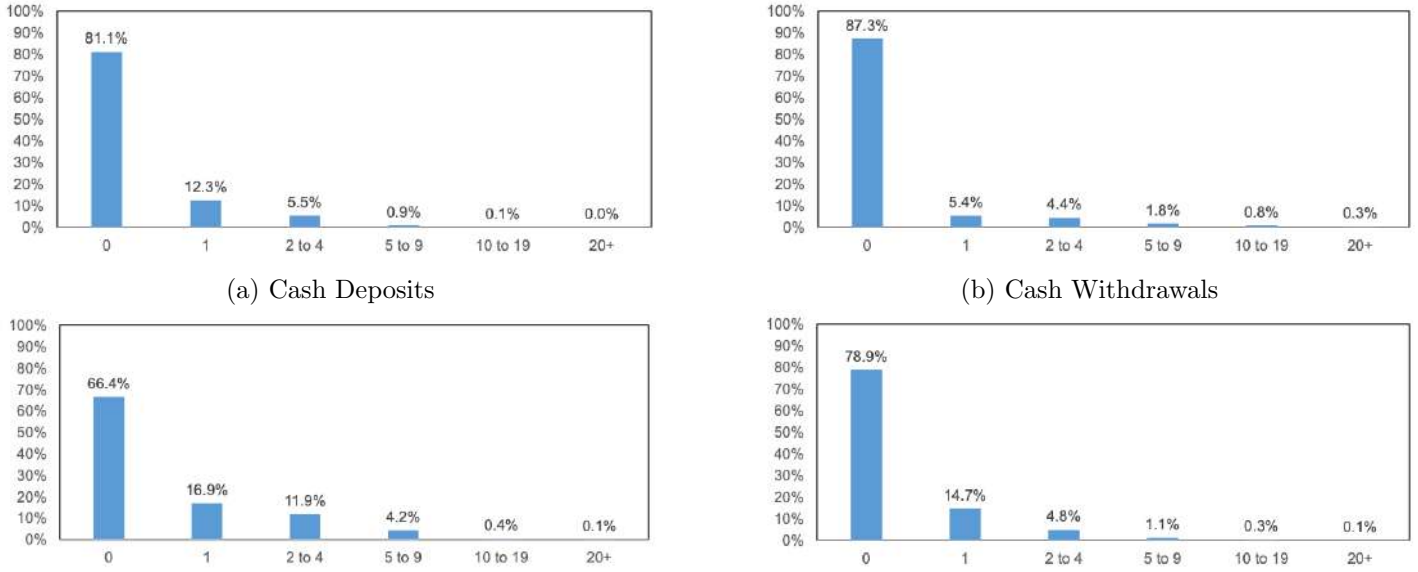
Panel B: Expenses on Hobby Classes				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	-0.001			
	(0.003)			
%age State-Owned Branches		-0.000		
		(0.003)		
%age Households Without Bank Accounts			-0.004	
			(0.003)	
Financial Inclusion Index				0.000
				(0.003)
Observations	415	415	415	415
R-squared	0.001	0.000	0.006	0.000

TABLE B8: [**Continued**]

Panel C: Expenses on Private Tuitions				
VARIABLES	(1)	(2)	(3)	(4)
Adults Per unit Bank Branch	0.018			
	(0.011)			
%age State-Owned Branches		0.007		
		(0.011)		
%age Households Without Bank Accounts			0.010	
			(0.011)	
Financial Inclusion Index				0.016
				(0.011)
Observations	415	415	415	415
R-squared	0.006	0.001	0.002	0.005

Figure B1: **Summary of Banking Transactions Performed**

The figure shows the distribution of the frequency of one of four banking transactions: Cash Deposits (panel (a)), Cash Withdrawals (panel (b)), Inward Remittances (panel (c)), and Outward Remittances (panel (d)) during the first six months since an account opening.



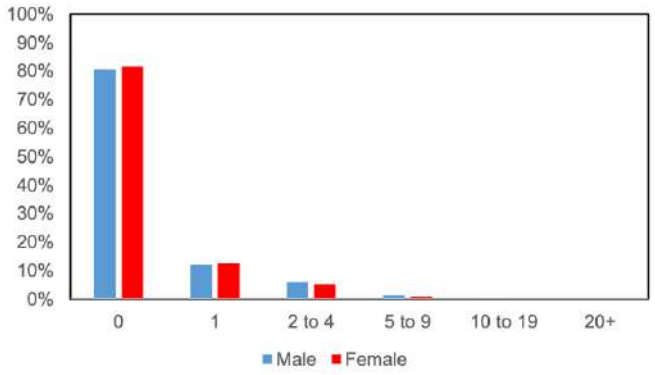
(a) Cash Deposits

(b) Cash Withdrawals

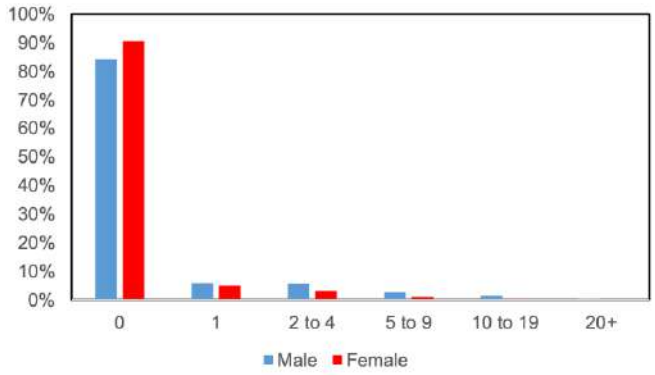
(c) Inward Remittances

(d) Outward Remittances

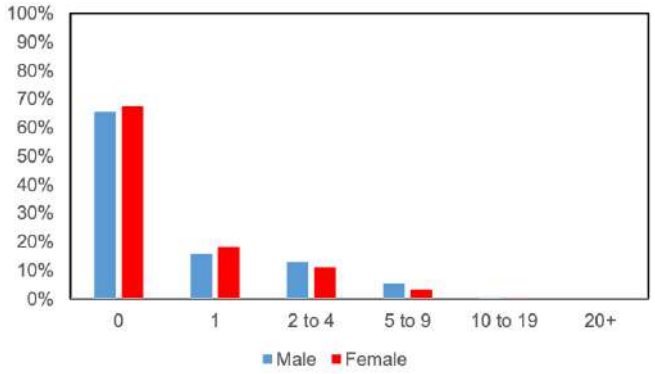
Figure B2: **Summary of Banking Transactions Performed (by Gender)**
The figure shows the distribution of the frequency of one of four banking transactions: Cash Deposits (panel (a)), Cash Withdrawals (panel (b)), Inward Remittances (panel (c)), and Outward Remittances (panel (d)) during the first six months since an account opening across accounts split by the sex (male, female) of the account holder.



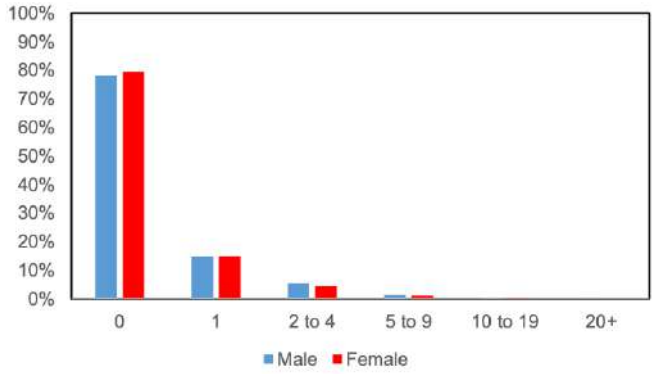
(a) Cash Deposits



(b) Cash Withdrawals

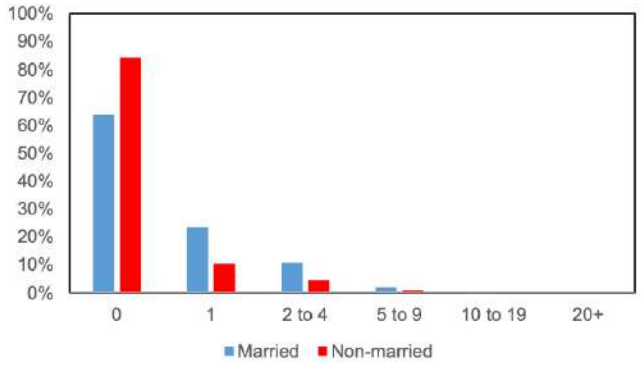


(c) Inward Remittances

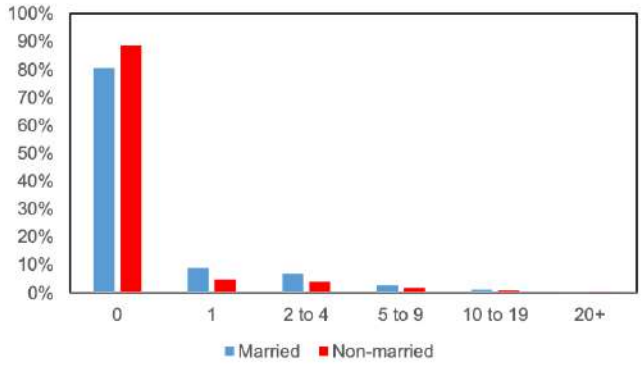


(d) Outward Remittances

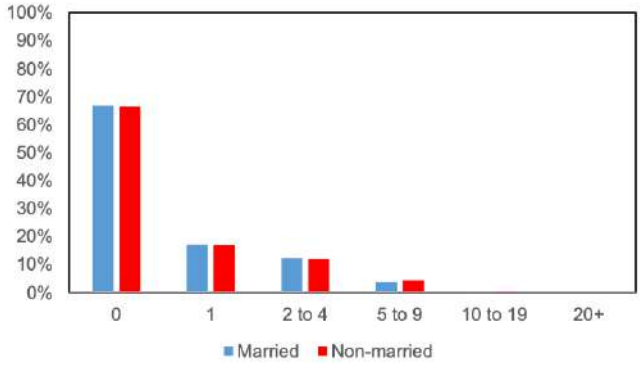
Figure B3: **Summary of Banking Transactions Performed (by Marital Status)**
The figure shows the distribution of the frequency of one of four banking transactions: Cash Deposits (panel (a)), Cash Withdrawals (panel (b)), Inward Remittances (panel (c)), and Outward Remittances (panel (d)) during the first six months since an account opening across accounts split by the martial status (married, unmarried) of the account holder.



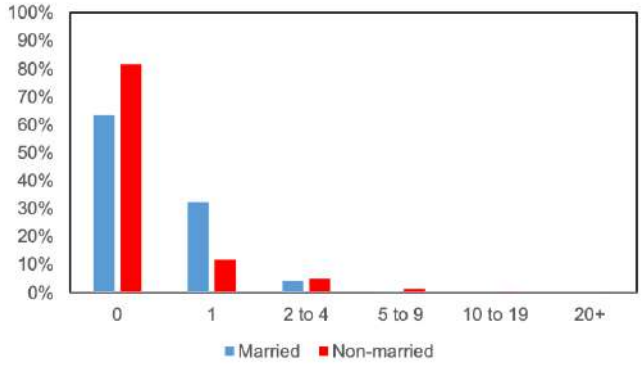
(a) Cash Deposits



(b) Cash Withdrawals

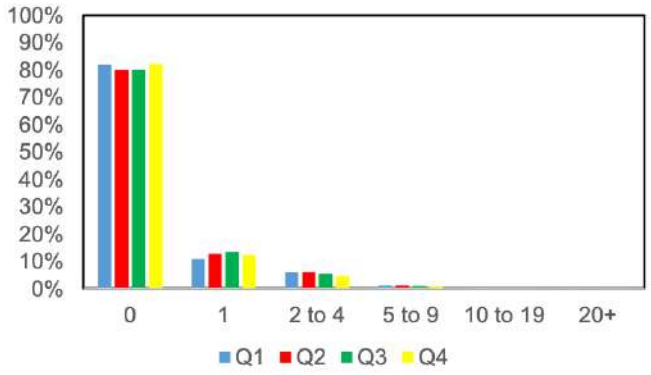


(c) Inward Remittances

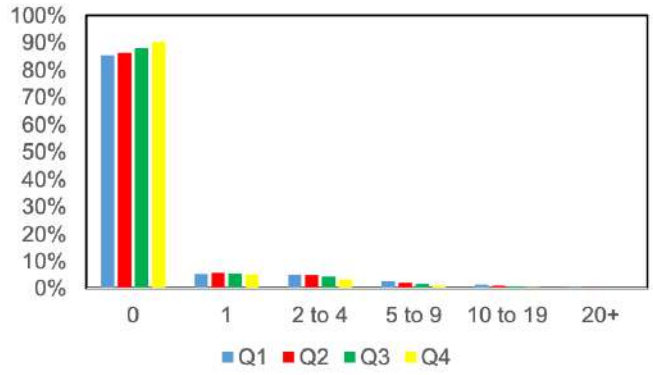


(d) Outward Remittances

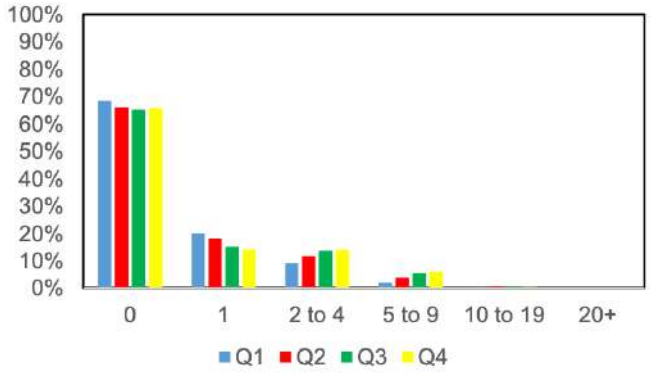
Figure B4: **Summary of Banking Transactions Performed (by Age Quartiles)**
The figure shows the distribution of the frequency of one of four banking transactions: Cash Deposits (panel (a)), Cash Withdrawals (panel (b)), Inward Remittances (panel (c)), and Outward Remittances (panel (d)) during the first six months since an account opening across accounts split by the age quartile of the account holder.



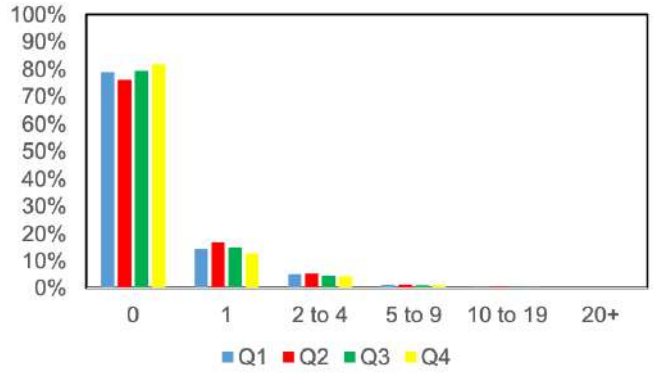
(a) Cash Deposits



(b) Cash Withdrawals



(c) Inward Remittances



(d) Outward Remittances