Cash is King:

The Role of Financial Infrastructure in Digital Adoption*

Bhavya Aggarwal† Nirupama Kulkarni‡ S. K. Ritadhi§

January 6, 2020

Please do not circulate or cite without prior permission

Abstract

This paper examines whether temporarily increasing the costs of cash transactions can induce households to permanently switch to formal financial transactions through digital payments, using the demonetization episode in India as a laboratory. The demonetization episode discontinued 86 percent of the cash in circulation overnight, causing a temporary increase in the cost of cash transactions. Using a differences-in-differences design, this paper shows that areas more exposed to cash shortage witnessed a 22 percent increase in digital payments and effects persisted over the long term. However, areas with high baseline informality witnessed an increase in digital transactions *only* when the requisite financial infrastructure in the form of point-of-sales terminals was already in place. However, there were no spillover effects on household participation in financial assets relative to gold and real estate investments. Our results suggest that financial infrastructure is an important determinant of whether interventions such as demonetization can successfully induce households to switch to formal financial transactions over the long run.

JEL Classification: E5, O23

Keywords: Formalization, demonetization, development and finance, cash economy

*All opinions expressed are those of the authors and don't necessarily reflect those of the Reserve Bank of India or CAFRAL. We would like to thank Viral Acharya, Abhay Aneja, Pranab Bardhan, Arnab Basu, Pallavi Chavan, Bikramaditya Datta, Prachi Deuskar, Gaurav Khanna, Aprajit Mahajan, Dilip Mukherjee, Priya Mukherjee, Abhinav Narayanan, Satish Rath, R. Ravikumar, Raghuram Rajan, Anand Srinivasan, Carly Trachtman, Christopher Udry and Sagar Wadhwa for several useful comments and suggestions. Kavya Ravindranath provided excellent research assistance.

†University of California - Berkeley Email:bhavyaaggarwal2012@gmail.com

‡Centre for Advanced Financial Research and Learning, Email:nirupama.kulkarni@gmail.com

§Reserve Bank of India. Email:ritadhi@gmail.com

1 Introduction

A distinguishing feature of many developing economies is a large informal sector providing employment to the majority of the labour force. However, while the informal sector is large in terms of employment, it is typically characterized by low productivity and wages, and has a limited share in national income (La Porta and Schleifer (2014)). Moreover, both enterprises and workers in the informal sector have limited access to formal financial institutions for the purposes of credit and savings, increasing their vulnerability to shocks, while excluding them from the government's tax base. This has led researchers to examining various interventions to induce informal enterprises (and workers) to join the formal economy. These have typically aimed at reducing the transaction costs of formalization, providing information regarding the benefits from formalizing, and increasing access to formal financial institutions.[1] The economic impact of most such interventions on firms and households' decision to enter the formal economy have however been modest (La Porta and Schleifer (2014)).

An alternative approach recently explored has been to increase the costs associated with informality (De Andrade et al. (2014); La Porta and Schleifer (2014)). This paper examines the efficacy of this approach by exploiting a unique large-scale natural experiment to empirically identify whether an increase in the cost of cash transactions can induce households to participate in the formal economy through digital payments. We study the "demonetization episode" — the shock decision by the Indian government in November 2016 to discontinue, with immediate effect, two of the largest currency denominations accounting for 86 percent of the cash in circulation.

The policy mandated citizens to deposit their discontinued currency into bank accounts and was accompanied by severe restrictions on currency withdrawals in the subsequent 3 months. This was compounded by the constraints faced by the central bank's printing presses to replenish the discontinued currency with new currency, all of which resulted in the policy causing an adverse reduction in cash supply across the economy which lasted till April 2018, and was most severe until June 2017 (Chodorow-Reich et al. (2018)).

The present paper uses a difference-in-difference framework with novel regulatory data on digital payments to study whether this increase in the cost of cash transactions induced households to participate in the formal economy through a) digital payments – which was unaffected by the policy –

[1] Transaction costs to enter the formal economy are typically of two types: registration costs for firms and the cost of opening bank accounts for households. As workers in the informal sector typically have lower skills in the form of educational qualifications, additional transaction costs involves the paperwork associated with firm registration/opening of bank accounts.

as a substitute for cash, and b) investments in formal financial instruments. Of key interest is whether households with a relatively high likelihood of operating in the informal sector switched to the formal economy through either of the consumption or investment channels mentioned above.

As discussed by Chodorow-Reich et al. (2018), the primarily role of cash is to facilitate economic transactions. This is particularly true in the informal sector which both anecdotally, and empirically, have been documented to be highly dependent on cash as a medium of exchange for wage payments, payments for services, and consumption. With 80 percent of the Indian workforce operating in the informal sector, demonetization (henceforth referred to as "treatment") serves as an ideal natural experiment to study whether an increase in the cost of cash transactions due to an unanticipated reduction in cash supply can force households operating in the informal economy (and mostly reliant on cash) to participate in the formal economy through digital payments as a substitute for cash. In this regard, this paper views non-cash digital payments to be in the realm of the formal economy while cash transactions are considered to be a part of the informal economy. Moreover, the scale of the intervention along with the exceptional secrecy surrounding its announcement meant that while every household was affected by the shock, almost no household anticipated the same.

For causal identification, we exploit the currency management framework of the central bank to generate cross-sectional variation in treatment intensity. Specifically, the central bank – the Reserve Bank of India (RBI) – uses a "hub and spoke" model for currency distribution. Under this framework, the central bank's printing presses distribute currency to 4,000 currency chests, which in turn allocate currency across bank branches and ATM terminals.[2] We exploit the variation in the least Euclidean distance between regions and currency chests to construct the cross-sectional variation in treatment intensity. Our fundamental assumption is that transportation costs and logistical challenges in cash disbursement are an increasing function of distance from currency chests. Consequently, currency replenishment would have taken longer in regions located farther from currency chests, prolonging the treatment's duration and intensity in such areas. We empirically verify this assumption by documenting that the treatment resulted in a 12 percent decline in cash withdrawals from ATM terminals in regions located farther from currency chests in the near term (within 12 months of the treatment), but this negative effect dissipates over the long-term (between 12 and 18 months of the treatment) as the cash supply recovered to its pre-treatment levels. This supports our hypothesis that the cash supply shock was more severe in regions located farther from currency chests.

[2]While the printing presses are operated by the central bank, the currency chests are managed by individual bank branches under a broad agreement with the central bank.

We use this cross-sectional variation in treatment intensity in a difference-in-difference framework to identify the effect of the cash shortage on digital payments. We measure digital payments using proprietary monthly zip code level data on transactions undertaken using debit and credit cards issued by a major national vendor – RuPay. We show that zip codes located farther from currency chests and thereby, having a higher exposure to the treatment, experienced a 16 (22) percent increase in the number (volume) of transactions undertaken through point-of-sales (POS) terminals using credit/debit cards. This positive effect persisted over the long run even as the treatment's negative impact on cash withdrawals steadily declined. This ascertains that a *temporary* increase in the cost of cash transactions can induce households to participate in the formal economy over the long run.

In support of our empirical design, we document that changes in digital payments and ATM cash withdrawals exhibited comparable trends in the pre-treatment period across zip codes situated farther from currency chests, relative to those located nearer to currency chests. The coefficients are robust to the inclusion of state-time fixed effects and district specific time-trends, as well as a permutation based placebo test where we randomly assign zip codes to high and low distance from currency chests.[3] As currency chests are likely to be located in areas with high economic activity,[4] the estimated treatment effects on digital payments if anything are a lower bound as areas with potentially lower economic activity ex-ante, respond most to the treatment.

To identify whether the treatment effects were driven through the adoption of digital payments by households transacting hitherto in the informal sector, we test for differential treatment effects across regions with a high degree of informality in the pre-treatment period: namely zip codes located in districts with a high share of rural households and informal workers. We choose the former as our primary measure of informality as 90 percent of the rural non-farm labour force is employed in the informal sector. The results however show that the treatment effects were muted in regions characterized by high informality – the positive impact of demonetization on digital payments instead were concentrated in zip codes located in districts with a relatively lower share of rural households and a relatively higher share of formal sector workers.

We explain the limited effect of the treatment in regions with high informality by exploring the channels through which the treatment might operate. We rule out that the limited effect of the treatment in rural areas was driven by income hiding motives, local informal networks temporarily reduc-

[3] Districts are the third level in India's administrative structure, below federal and state. It is comparable to the U.S. county. Zip codes are situated within districts.

[4] This is also reflected in our data which shows that that zip codes located farther from currency chests have significantly lower population and financial infrastructure.

ing a household's dependence on cash, or a disproportionate reduction in rural consumption. Instead, we find support for the contention that the limited availability of financial infrastructure in rural areas (as observed from the strong negative correlation between the share of rural households and POS terminals (bank branches) per capita in a district/zip code) diminished the treatment's impact in such areas.

Using a standardized financial infrastructure index based on the pre-treatment availability of POS terminals and bank branches at both the district and the zip code levels, we first show that the treatment effects are enhanced in areas with high financial infrastructure. Subsequently, we test for differential treatment effects across rural zip codes, conditional on the zip code's pre-treatment levels of financial infrastructure. The results confirm our hypothesis that the lack of financial infrastructure forms a major hindrance to the adoption of digital payments in response to the shock: zip codes in the top quartile of financial infrastructure no longer exhibit any differential treatment effects across rural zip codes and the net treatment effect is also positive and statistically significant. On the contrary, for zip codes falling outside the top quartile of financial infrastructure, the treatment continues to have a significantly lower effect on digital payments in rural zip codes. These results underline the critical role played by financial infrastructure in facilitating the adoption of digital payments in response to the cash supply shock.

As the intervention directed citizens to deposit the discontinued currency in banks, we also examine whether the treatment caused any spillovers on households' participation in financial instruments. We expect this to occur if citizens elected to invest any excess liquidity hitherto held as cash, or learnt more about financial instruments in the course of bank visits. For household investment choices, we use data from the Consumer Pyramids – a large nationally representative household panel which contains self-reported information on household participation in formal financial instruments. The results however do not support the contention that the treatment generated spillovers on household participation in financial instruments – in fact, we find that post treatment, households' participation in financial instruments declined, in areas wuth high treatment intensity, particularly over the long-term.

Unlike the case with digital payments where the treatment had a limited effect only in rural areas, we find that the treatment reduced households' participation in financial instruments across both urban and rural households. Consistent with the results on digital payments, the treatment does induce positive spillovers on household participation in financial instruments in districts with high financial infrastructure (measured using bank branches per capita): in such districts, the treatment caused a small increase in urban households' participation in financial instruments over the near term.

Collectively, our results highlight the necessity of financial infrastructure for the treatment to induce informal households to participate in the formal economy, through both the consumption and investment channels. Borrowing De Andrade et al. (2014) terminology, our paper documents that applying the "stick" to raise the costs of informality affects formalization only after the necessary "carrots" in the form of financial infrastructure have already been put in place.

Our paper adds to the large literature studying informality and development. Existing papers such as De Andrade et al. (2014), De Mel et al. (2013) and Jaramillo (2009) show that increasing the ease of formalization had little impact on firms' switching to the formal sector. On the contrary, De Andrade et al. (2014) showed that increasing the costs of remaining informal had modest positive effects on firms' entry into the formal sector. Our paper adds to this literature by empirically identifying whether a similar strategy can also increase households' participation in the formal economy. To this effect, we exploit an unique natural experiment in India which increased the cost of cash transactions through an extensive disruption to cash supply. Importantly, we show that increasing the costs of informality raises household participation in the formal economy, conditional on the availability of financial infrastructure. Our paper in this regard reconciles the two existing approaches to induce formalization: namely, lowering the costs of formalization through improved access to finance, while increasing the costs associated with informality. Moreover, whereas much of the literature studying informality and development rely on field experiments, we use a natural experiment applicable to a population in excess of 300 million households, reducing concerns regarding external validity.[5]

By underlining the role of financial infrastructure, our paper contributes to the literature studying the effect of financial infrastructure and development. While Burgess and Pande (2005) and more recently, Prina (2015) document the benefits from the opening of rural bank branches and low cost bank accounts, Dupas et al. (2018) show through field experiments across 3 countries that access to bank accounts has limited benefits in the form of take-up and usage by households.[6] By showing that the treatment effects were concentrated in areas with a high density of POS terminals and bank branches, we document the vital role played by financial infrastructure in facilitating households'

[5]For instance, Prina (2015) uses a field experiment in Nepal and shows that low cost savings accounts lead to higher savings by households. This is in contrast to much of the literature which finds muted effects of such accounts. In a discussion of her results, Prina (2015) attributes the positive effects to her urban sample.

[6]Burgess and Pande (2005) show that the mandated opening of rural bank branches in India accelerated the process of poverty alleviation.

participation in digital payments.

Finally, our paper adds to a growing literature studying the economic effects of "demonetization" which generated a widespread liquidity shock in the world's fifth largest economy due to the sudden discontinuation of a large volume of currency. Chodorow-Reich et al. (2018) examine this event to test the Keynesian hypothesis regarding the role of cash and show that the shock results in a decline in both local GDP and employment, while inducing households to adopt digital payments. Crouzet et al. (2018) also document similar findings in their study of whether demonetization facilitated technology adoption through network effects. On the contrary, the current paper exploits the shock as a treatment intervention increasing the costs of informal transactions and causally identifies its impact on household participation in the formal economy. While our baseline results are consistent with the findings of the two above-mentioned studies, we make two additional contributions: first, we document the widespread heterogeneity in the treatment's impact across rural areas with high informality, and second, we showcase the critical role played by financial infrastructure for such a shock to induce households to switch to the formal economy through digital payments (financial instruments). The remainder of the paper is organized as follows: Section 2 provides a broad overview regarding the informal sector in India, the demonetization episode which serves as the treatment intervention, and the central bank's currency management. Section 3 discusses the two major data sources in the paper and our mapping of zip codes to the nearest currency chest. Section 4 describes the paper's empirical strategy and Section 5 documents the key findings of the paper. We subject our results to a set of robustness checks in Section 6 and Section 7 offers some conclusions.

2 Informal Sector, Cash Supply Shock and Currency Circulation in India: Background

We begin with a brief overview of the informal sector in India and describe the policy intervention of interest. Next, we outline the channels through which the intervention could have affected households' participation in the formal economy. We end with a discussion on currency circulation in India and how we use the location of currency chests as a source of variation in treatment intensity.

2.1 The Informal Sector in India

Akin to other developing economies, India's informal sector is massive in terms of employment. According to the National Sample Survey (NSS) in 2011-12, 80 percent of working age adults were employed in the informal sector.[7] Out of these, 71 percent were employed in enterprises with under 6 employees. Despite the overwhelming employment share, the informal sector only contributed 45 percent towards the national income (Kotwal et al. (2011)).[8] While average annual economic growth exceeded 7 percent between 1999 and 2011, the share of employment in the informal sector declined only moderately from 88 to 82 percent,[9] contrary to the expectations voiced in La Porta and Schleifer (2014) that the process of development and creative destruction would result in an automatic decline in informality.

The NSS also informs us that the informal sector is concentrated within rural areas: in 2011-12, 90 percent of rural non-farm workers were employed in the informal sector, contrary to 75 percent of urban workers.[10] The informal sector was also characterized by a high share (55 percent) of self-employed workers. Workers employed in the informal sector also tended to have lower skills with only 51 percent of informal sector workers having completed secondary education, relative to 72 percent of formal sector workers.

Districts with a high share of informal sector workers also had lower levels of financial infrastructure and per capita consumption. Here, we consider districts to have a "low" share of informal workers if the share of informal workers in the district falls in the bottom quartile.[11] The remaining districts are classified as "high" share. In 2011, districts with a "high" share of informal workers had 76 bank branches per million persons while districts with a "low" share of informal workers had 102 bank branches per million population. The average monthly household per capita expenditure in districts with a high share of informal sector workers was also 30 percent lower than that in districts with a low share of informal sector workers.[12]

Collectively, these stylized facts suggest that an overwhelming share of the non-farm workforce is employed in the informal sector, with the majority being either self-employed or working in very small enterprises. There has been limited growth in formal sector employment, even during a period of robust aggregate economic growth. Workers in the informal sector are also less skilled in terms of

[7]The NSS conducted representative household surveys on employment and unemployment across India. Working age adults refer to adults aged between 15 and 65 and we restrict the sample to employed workers in the non-farm sector. We define workers working in the informal sector as those who work in enterprises with less than 10 (20) workers when using (not using) electricity.

[8]This too excludes the farm sector - upon including the farm sector, it rises to 58 percent.

[9]In fact, the share of workers employed in very small non-farm enterprises employing under 6 workers grew in this period from 47 to 55 percent.

[10]We exclude rural farm workers as the vast majority of farm workers are classified as informal workers.

[11]The share of informal workers in the non-farm workforce in districts in the bottom quartile is 82 percent or less.

[12]In rupee value, the average monthly per capita expenditures in districts with a high share of informal workers was Rs. 1,385.

educational qualifications and regions with a relatively high share of informal workers have significantly lower financial infrastructure (bank branches per capita).

2.2 Currency in the Indian Economy and Demonetization

Along with a large informal sector, the Indian economy is also characterized by a high level of currency in circulation. In comparison to other major economies, India had one of the highest cash to GDP ratios in 2015, surpassed only by Japan and Hong Kong (Rogoff, 2016). As seen from Figure A1 (Appendix 7), currency-in-circulation as a share of GDP remained stable at 12 percent between 2011 and 2015, with 80 percent of the currency being accounted for by two of the highest denominations – namely Rs. 500 or Rs. 1,000.[13]

On the evening of November 8, 2016, the Prime Minister, Mr. Narendra Modi, made the shock announcement through a televised address that the two largest currency denominations – namely Rs. 500 and Rs. 1,000 – would cease to remain official tender effective midnight, November 9, 2016.[14] The remaining currency denominations and coins were left unaffected and the government undertook to replace the discontinued denominations with a fresh set of Rs. 500 bills and introduced the Rs. 2,000 denomination.

Citizens were provided with a 50 day window, expiring on December 31, 2016, to deposit the discontinued currency in their possession into bank or post office saving accounts. In lieu of anticipated shortages to cash supply, a ceiling of Rs. 20,000 was placed on weekly cash withdrawals from bank accounts for the upcoming weeks.[15] Citizens were eligible to exchange a limited amount of old currency denominations with the new ones in any bank branch upon producing a proof of identity. The initial daily ceiling for such transactions was Rs. 4,000.[16] Restrictions were also imposed on withdrawals from ATM terminals in this period – ATM terminals were closed on November 9 and 10, and subsequently, withdrawals were capped at Rs. 2,000, raised later to Rs. 4,000.[17] No restrictions however were placed on electronic bank transfers, cheques, and digital payments undertaken through e-wallets

[13]The remaining denominations were Rs. 100, Rs. 50, Rs. 20, Rs. 10, Rs. 5, Rs. 2 and Rs. 1, in addition to Rs. 5, Rs. 2 and Rs. 1 coins. In terms of relative magnitude, the nationally representative consumption surveys conducted by the National Sample Survey Organization estimated the average per capita monthly consumption of an urban household to equal Rs. 2,000 - equivalent to 2 (4) Rs 1,000 (Rs. 500) currency bills.

[14]Legally, the government was empowered to undertake this move through Section 26(2) of the RBI Act which allows the government the right to determine the denominations of currency in circulation. Formally, as Parliament was not in session during this period, the government first issued an ordinance to execute the decision to withdraw these currencies, and subsequently, ratified the ordinance through an act of Parliament.

[15]The Prime Minister promised that the ceiling would be revised upwards over time as the supply of the new currency bills increased.

[16]The initial ceiling was supposed to be in place between November 8 and November 24, 2016

[17]Unlike the ceiling on bank transactions, no specific timeline was provided for cash withdrawals from ATMs.

or debit/credit cards.[18]

The government initially justified the disruption as a harsh but necessary step in the fight against corruption and undisclosed wealth, citing that much of corruption is conducted through cash transactions, which accrue to the "shadow" economy.[19] Over time, additional rationales were provided by the government for the policy and in November 2018, the Finance Minister claimed that the primary objective of demonetization was to push households to participate in the formal economy and incorporate excess cash into the banking channel (The Indian Express, November 2018).

We highlight three key elements of "demonetization" to justify it as an exogenous intervention which increased the cost of informal transactions. First, both the policy and its timing came as a complete surprise to citizens as only a select group of advisors associated with the Prime Minister had prior knowledge about the event. This lack of information makes demonetization a truly exogenous event as households neither anticipated the shock, nor could have planned their response in advance. Demonetization in this regard is comparable to an exogenous treatment intervention.

Second, while the discontinuation of 86 percent of the currency in circulation, combined with printing press constraints and restrictions placed on currency withdrawals from banks and ATMs significantly increased the transaction costs associated with cash transactions, absolutely no restrictions were imposed on transactions conducted through credit and debit cards. This permits us to explore whether this exogenous treatment intervention was successful in inducing households to adopt credit/debit cards to conduct their necessary transactions, in lieu of cash. Thereby, our paper considers cash transactions to be in the domain of the informal economy, while digital payments conducted using credit and debit cards are deemed to be within the realm of the formal economy. We also examine whether the treatment had a persistent effect on digital payments conducted over the medium and long-term using credit/debit cards, or whether households returned to cash once the cash supply was restored to its pre-treatment levels.

Finally, the ubiquitous nature of currency in the Indian economy implied that the treatment affected almost all segments of the population and across all geographical regions. The widespread nature of the treatment however precludes the existence of a pre-defined control group which can serve as a valid counterfactual.[20] In this regard, we exploit the framework of the central bank's cur-

[18] E-wallets are akin to prepaid cards.

[19] For instance, during the speech, the Prime Minister gave the example that a substantial fraction of real estate purchases are often undertaken through cash. A second reason offered was that a part of the undeclared cash was used to fund terrorism and purchase weapons.

[20] For instance, Agarwal and Qian (2014) identify consumption responses to an unanticipated income shock induced by

rency supply in the economy to generate cross-sectional variation in regional treatment intensity.

2.3 RBI's Supply of Currency and Regional Variations in Intensity of Shock

As the central bank, the RBI is responsible for currency management and maintaining currency supply across the economy. For currency circulation, the central bank follows a "hub-and-spoke" model. The RBI operates 4 printing presses to print currency which is then sent to 19 issue offices, from where it is distributed to 4,034 currency chests (CC).[21] The CCs serve as repositories of currency and are maintained by commercial banks. From the CCs, the currency is distributed to individual bank branches, which then supply cash to the ATMs. In all, the 4,000 CCs allocate currency to over 135,000 bank branches and 200,000 ATMs. Figure 3b shows that the location of CCs is uniformly spread across India.

As we detect household participation in the formal economy through zip code level transactions made from POS terminals, the ideal experiment would have been to randomly assign zip codes to treatment and control status. In the absence of such random assignment, the second best strategy would have been to exploit any cross-sectional variation in zip codes' pre-treatment cash dependency using a strategy similar to Autor et al. (2013).[22] Unfortunately, neither of these two variables are observable to researchers. While the central bank maintains data on CC cash flows, it is aggregated at the level of district and does not disclose the cash flowing to individual zip codes.

In the absence of zip code-level variation in pre-treatment cash dependency, we exploit instead the variation in the distance between CCs and zip codes. The core assumption is that transportation costs and logistical challenges are an increasing function of distance, implying that transmitting currency to zip codes located farther from a CC would involve higher costs and entail greater time. In a period of significant disruption and currency shortage, it is plausible that the logistical challenges in replenishing currency to areas located farther from CCs would be further accentuated, necessitating a longer time duration. This would heighten the severity of the cash supply shock in zip codes located farther from CCs, increasing the treatment's intensity in these zip codes.

Resultantly, this paper constructs its cross-sectional variation in treatment intensity based on the Euclidean distance between a zip code and its nearest CC. Importantly, as CC locations are fixed,

the growth dividend paid by the Singapore government in 2012. In their case, the substantial proportion of foreigners residing in Singapore – and ineligible for the growth dividend payout – provide a valid counterfactual group to study the effect of the income shock on eligible recipients.

[21] These are strategically located, with one each in eastern, southern, central and western India.

[22] An ideal candidate would have been the amount of cash flowing to all bank branches in the zip code from currency chests in a calendar year, scaled by the economic activity in the zip code for that year.

the set of zip code-CC distances is by construction orthogonal to the treatment. Additionally, there were also no new CC openings in the post-treatment period (or immediately prior to the treatment), ruling out concerns about strategic manipulations in the nearest zip code-CC distance. We describe the computation of this distance in the subsequent section.

2.4 Cash Supply Shock and Household Participation in the Formal Economy: Potential Channels

Before describing our data and empirical strategy, we briefly discuss the two major channels through which the treatment could have positively affected households' participation in the formal economy. The first is what we term the *substitution effect*. As the policy discontinued 86 percent of the currency in circulation and operational constraints on printing presses limited immediate replenishment of the discontinued currency, it severely reduced the overall availability of cash in the economy. This was compounded with the restrictions placed on cash withdrawals from both bank branches and ATM terminals.[23] As the informal economy is heavily reliant on cash transactions, the sharp decline in cash availability could have resulted in households switching to digital payments to conduct their consumption as a substitute for cash transactions. We would expect this effect (at least in the near-term) to be most pronounced for households which already possessed the necessary infrastructure (in the form of debit/credit cards) and had access to firms and retail stores equipped with digital payment systems such as POS terminals.

The second channel we consider is the *information* channel. As the policy required households to deposit their existing set of Rs. 500 and Rs. 1,000 denominations by physically visiting bank branches, it is plausible that interactions with bank officials would have provided information to households regarding the benefits arising from digital transactions. Additionally, it is also possible that households could have learned about financial instruments being offered at the bank during these visits.[24] Thus, the information obtained through bank visits pertaining to digital transactions and financial instruments could also have positively impacted households' participation in the formal economy.

Alternatively though, the following factors which could have limited the treatment's impact on household participation in the formal economy, particularly for households operating in the informal sector. First, if informal transactions were a medium for tax evasion, it is possible that households would select to opt out of digital payments and financial instruments in an effort to hide their wealth.

[23] For instance, Chodorow-Reich et al. (2018) estimated that the cash supply shortage continued till at least June 2017.

[24] While households had the option of exchanging their old currency bills with the new ones, they could only do so for a very limited amount of currency, as mentioned in Section 2.2.

This is particularly plausible given the strong anti-corruption pitch accompanying the policy intervention. Second, the disruption in cash supply could have resulted in a liquidity shock for households operating in the cash-dependent informal economy, causing an aggregate decline in their consumption, and thereby, lower digital transactions. Relatedly, Chodorow-Reich et al. (2018) established that demonetization resulted in a 3 percent decline in GDP in the first 9 months after the policy. If this was concentrated in the cash-dependent informal economy, it could have negatively impacted the income of workers employed in the informal sector, imposing constraints on their consumption and investments. Third, households could have taken recourse to informal credit networks to temporarily reduce their dependence on cash to undertake necessary transactions. Finally, the lack of financial infrastructure and financial inclusion can limit the impact of the shock on households' participation in the formal economy. The availability of POS terminals along with credit/debit cards are a necessary condition for conducting digital payments. Similarly, access to bank branches is essential for households to undertake investments in financial instruments (as well as obtaining credit/debit cards). As discussed in Section 2.1 and depicted in Figure A2 (Appendix 7), rural areas and regions with a high share of informal workers had significantly lower financial infrastructure, both in terms of bank branches, and POS terminals, and this could have limited the treatment's impact. Thus, ex-ante, the causal impact of the shock on households' participation in the formal economy is ambiguous and this paper empirically identifies this effect.

3 Data and Summary Statistics

This section discusses the primary datasets used in the paper and describes the cross-sectional variation in treatment intensity. We also describe some descriptive trends to motivate the empirical results.

3.1 Digital Payments

The data on digital transactions is provided by the National Payments Corporation of India (NPCI). This is proprietary data collected by the NPCI, comprising of transactions conducted using RuPay credit/debit cards across all point-of-sales (POS) terminals issued by RuPay.[25] The data, aggregated to the zip code level, is available at a monthly frequency between January 2016 and April 2018. This provides us with a 28-month zip code-level panel with 10 months of pre, and 18 months of post-treatment observations. In addition to digital payments, the NPCI also provides us with monthly zip

[25]RuPay is a domestic network of credit and debit cards, handled by the NPCI, and supervised by the RBI. In terms of representation, in April 2018, our data included 35 percent of all POS terminals in the country and accounted for 5 percent of the value of transactions undertaken using credit and debit cards.

code-level data on ATM withdrawals from all operating ATMs.

To compute each zip code's exposure to the treatment, we first compute the zip code-CC distance using the latitude and longitude of the centroid of each zip code from the All India Pincode Directory maintained by the Government of India. We combine this with the RBI's list of physical addresses of all CCs and express each zip code's distance to the nearest CC as:

$$DistCC_i = \min[Dist(z_i, z_c)]; \forall c \in \mathbf{C} \tag{1}$$

In (1) *Dist* is the Euclidean distance between zip code i and currency chest (CC) c, based on the geo-coordinates of zip code i (z_i) and CC (z_c). This distance is calculated for all operating CCs (**C**) and we obtain the minimum value from this vector of distances to define $DistCC_i$ as the distance to the nearest CC from zip code i. From (1), the median zip code to nearest CC distance – *DistCC* – is 18 kilometres, while the mean distance is 19 kilometres. A little under a sixth of the zip codes (2,734 zip codes) have a CC located within the zip code, implying $DistCC = 0$. Section 2.3 hypothesized that cash replenishment to zip codes located further from CCs would take longer, amplifying the intensity of the treatment in such areas. To this effect, we classify zip codes with a relatively high exposure to the treatment using the indicator $HighDistCC_i$ which equals 1 if $DistCC_i > \pi_{DistCC}$; π_{DistCC} being the median value of *DistCC* across all zip codes. Section 6.1 verifies that our results are robust to alternate distance thresholds to classify zip codes as *HighDistCC*.

3.2 Household Consumption and Financial Instruments

We use the household-level Consumer Pyramids (CP) database, maintained by the Centre for Monitoring the Indian Economy (CMIE), to pinpoint the specific channels through which the treatment affects digital payments, along with identifying whether the treatment had spillovers on households' participation in financial instruments. The CP is a household-level panel collected through extensive household surveys. The database was initiated in 2014 and covers 27 states and 514 districts.[26] Each household is assigned to a survey wave and interviewed thrice every year.[27] The surveys themselves are conducted continuously through the year and the data identifies the precise month within each wave in which the household is surveyed. Each survey wave covers approximately 135,000 households with 34,000 households being surveyed each month. While there is limited attrition in the data,

[26]Our paper limits the sample to the twenty major states in India, excluding the states of Goa, Jammu and Kashmir, Meghalaya, Tripura, and the union territories of Chandigarh and Puducherry.

[27]The survey waves are identified as January-April; June-August and September-December respectively.

we restrict ourselves to the 119,482 households for whom we are able to construct a balanced panel between September 2014 and August 2018, providing us with 12 observations – 6 pre and 6 post-treatment – for every household. One drawback of the CP data however is that it oversamples urban households, with 60 percent of the sample residing in urban areas. We address this by weighting our results using the sample weights provided by the CP.[28]

The CP data covers household investments, borrowing, consumption and income, in addition to household demographics. The consumption module enquires of households their monthly expenditures in the four months preceding the survey across 57 consumption items, including food, clothing, human capital and durables. Similarly, the investment module enquires about investments made by households along the extensive margin in the 4 months preceding the survey across 8 financial instruments requiring participation in formal financial markets, in addition to gold and real estate.[29] For the consumption module, we average the household's monthly consumption across the 4 months to obtain an average consumption value corresponding to each survey wave. We normalize this by the household size to obtain $MPCE_{hdt}$ – denoting the average monthly per capita expenditure for household h, residing in district d and interviewed in month-year t. To determine households' participation in financial instruments, we use a binary measure – $AnyFin_{hdt}$ which equals 1 if the household invested in either of the 8 financial instruments in the period $[t-4, t]$. Across the 6 survey waves in the pre-treatment period (between September 2014 and August 2016), the average MPCE for rural households was Rs. 1,987 and the corresponding MPCE for urban households was Rs. 2,862.[30] During this period, 21 percent of households within a survey wave made some investment in financial instruments and the corresponding figures for gold and real estate is 5 percent.

3.3 Other Data Sources

In addition to the NPCI and CP data, we also use data from the National Sample Survey (NSS) and the Basic Statistical Returns (BSR) to obtain covariates of interest for identifying treatment heterogeneity.[31] As the paper seeks to identify whether an increase in the costs of informality can induce informal

[28]The weights reflect the inverse of the sampling probability for each household.

[29]These include both risk-free and risky instruments. The complete list is: a) fixed deposits; b) post office savings; c) national savings certificates; d) Kisan Vikas Patra; e) provident fund; f) mutual funds; g) shares and h) life insurance.

[30]For the sake of comparison, the average monthly per capita expenditures reported by the NSS in 2011 was Rs. 1,310 for rural and Rs. 2,708 for urban households.

[31]The National Sample Surveys are nationally representative household surveys conducted by the National Sample Survey Organization (NSSO), headed by the Ministry of Statistics and Programme Implementation. We use Schedule 10 of the 68th round of the NSS survey, covering household demographic and employment characteristics as measured in 2011-12. The Basic Statistical Returns is an annual dataset of the Reserve Bank of India which collects account-level information on deposits and credits across all branches of scheduled commercial banks in India.

households to adopt digital payments, we need to identify the treatment's impact in regions with high informality. To this effect, we construct district and zip code level measures of informality. At the level of the district, our principal measure of informality is the share of rural households, which we supplement with the share of informal sector workers and the share of self-employed workers.[32] We consider the share of rural households as our prime measure of informality based on the discussion in Section 2.1 where we documented that the majority of rural non-farm workers operate in the informal sector. This also makes our measure of informality consistent across the multiple data sources as the CP data identifies rural households and zip codes can also be classified as rural or urban (see discussion below). There is however no corresponding variable in the CP database capturing households operating in the informal sector; neither is there any measure of zip code level informality.

For each indicator of informality, we categorize districts as "low" or "high", based on their ranking, relative to that of the median district in 2011. Thus, a district is classified as "high rural" (and thereby high informal) if the share of rural households in the district exceeds the median share of rural households across all districts in 2011-12. Finally, in the absence of zip code covariates, we also use the NSS to obtain other district-level covariates of interest such as district population, monthly per capita consumption, and the share of workers with secondary or higher education.

Due to the paucity of administrative data at the zip code level, we use the BSR, published annually by the RBI, to infer whether a zip code is urban or rural and also impute the zip code's population. To achieve this, we use the data on bank branch locations published by the RBI to first map bank branches to zip codes.[33] To identify whether a zip code is rural, we use the RBI's classification of branches by population centres and consider a zip code to be "rural" if every bank branch operating in the zip code is either "rural" or "semi-urban".[34] Based on this classification, 51 percent of the zip codes in our sample are rural.[35] To impute the zip code's population, we assign to each zip code a fraction of the

[32]We consider a worker to be employed in the formal sector if the worker works either in an enterprise employing over 20 workers, or in an enterprise employing over 10 workers and using electricity. The classification of self-employed worker is based on the responses in the NSS employment-unemployment survey about the employment type of each worker. Based on the household responses and sampling weights assigned to household by the NSS, we compute the share of rural households, informal sector workers and self-employed workers in each district.

[33]The RBI for each year provides a complete list of bank branches in operation as of 31st March that year. The list contains a unique code for each branch and the physical address of the branch. We extract the pincode for each branch from the physical address and use the unique branch code to map the branch to the BSR data. This provides us with a mapping of branches to zip codes. Through this exercise, we successfully match 85 percent of the banks in our sample to a zip code. 18 percent of the zip codes in our sample lack a bank branch and we drop them from this exercise.

[34]The RBI classifies branches as "semi-urban" if they operate in an area with a population between 10,000 and 100,000 and "rural" if they operate in an area with a population less than 10,000. The remain two categories are urban and metropolitan for which the respective population cutoffs are 100,000-1 million and in excess of 1 million.

[35]A stricter classification would consider zip codes to be rural if every branch in that zip code is "rural". Under this classification 17 percent of the zip codes would be rural. We consider this to be an extreme classification of rural areas, in

district's population proportionate to the district's share of deposit accounts attributable to that zip code, based on the data ending in March 2016.[36]

We combine the data from the NPCI and BSR to construct district and zip code-level measures of financial infrastructure. Section 2.4 acknowledged that participation in the formal economy through digital payments require the presence of both POS terminals and credit/debit cards, which are issued by banks. In this regard, we create a financial infrastructure index at the district (zip code) level by adding the standardized indices of bank branches and POS terminals per capita at the district (zip code) level. For bank branches (POS terminals), we use the total number of bank branches (POS terminals) in operation in the district (zip code) in March (January) 2016. Both measures are normalized by the district (zip code) population in millions with the zip code population being imputed in the manner described above.

3.4 Descriptive Trends

Figure 4 plots the monthly trends in ATM and POS transactions based on the NPCI data. We aggregate monthly transactions across all zip codes and normalize it by the number (volume) of transactions conducted in January 2016. This provides us with an index of digital payments relative to January 2016. The left hand figure depicts the level of ATM transactions which exhibits a stable trend in the number (volume) of transactions between January and October 2016 with very limited growth. This is followed by a steep drop in both the number and volume of transactions in November 2016 – the volume of ATM transactions shrunk by almost 50 percent in December 2016, relative to January 2016 – indicating the severity of the shortage in cash supply due to the sudden withdrawal of high value currency denominations. While the number (volume) of ATM transactions started recovering since January 2017, they returned to the pre-treatment levels only in August 2017.[37] In contrast, the number (volume) of transactions conducted through POS terminals exhibited a flat trend till October 2016 and rose ten-fold through December 2016, relative to January 2016. Although the number (volume) of digital transactions dropped subsequently between January and April 2017, it still remained significantly higher (by a magnitude of 5 or more) than the pre-treatment levels till the last period in our sample (April 2018), suggesting that the number (volume) of digital payments shifted to a new (and higher) equilibirum in the post-treatment period.

light of the fact that over 60 percent of Indian households reside in rural areas.

[36] As we calculate bank branches and deposit accounts in March 2016, this includes all the deposit accounts opened under the Pradhan Mantri Jan Dhan Yojana (PMJDY) scheme, which was a major push towards financial inclusion.

[37] As a point of comparison, the aggregate level of cash supply in the economy recovered to the pre-shock levels only in April 2018 - 16 months after the shock.

Next, we disaggregate the monthly trends by zip codes' distance from CCs. Figure 5A shows the aggregate trends in monthly transactions made from ATM terminals across zip codes located far from CCs (*HighDistCC = 1*), versus zip codes located near CCs (*HighDistCC = 0*). Akin to Figure 4, we normalize the monthly transactions from POS (ATM) terminals by the transactions conducted in January 2016. ATM withdrawals in both sets of zip codes fall sharply in the aftermath of the treatment intervention, reflecting the decline in cash supply. While the monthly trends in ATM withdrawals are identical for both groups of zip codes, the level of ATM withdrawals (scaled) are consistently lower in zip codes located farther from CCs than those near CCs in the months immediately succeeding the shock. This is contrary to the monthly trends in Figure 5B which document a sharp increase in transactions undertaken through POS terminals in the post-treatment period in zip codes located farther from CCs, particularly in terms of transaction volumes.[38] Critically, there is no evidence of pre-trends in POS transactions in the 10 months prior to the treatment intervention. Both in terms of transaction counts and volumes, transactions made at POS (ATM) terminals (relative to January 2016) vary identically in the pre-treatment period across zip codes located relatively far from CCs — vis-a-vis zip codes located relatively close to CCs.

To provide some preliminary intuition about whether the treatment induced informal households to participate in the formal economy, we split our sample and separately plot the monthly trends in digital payments for zip codes located in districts with high informality, measured using the share of rural households (Figure A5). The outcome of interest is transactions undertaken through POS terminals and monthly transactions are normalized by the January 2016 transaction levels. As stated previously, districts are classifed as "high rural" if the share of rural households in the district exceeds the median share of rural households across all districts in 2011-12. We see that the sharp increase in POS transactions seen in Figure 5B is only observed in zip codes located in districts with a relatively low share of rural households. For zip codes located in districts with a high share of rural households, while there is an increase in transactions from POS terminals, there is no *differential* effect for zip codes located farther from CCs. The trends are very similar if we measure informality by the district's share of workers employed in the informal sector (Figure A10, Appendix 7).

Figure 7 plots the non-parametric relationship between digital payments and zip codes' relative distance from CCs by comparing the change in average digital payments (in logs) across the pre and post-treatment periods. We plot the means of these average differences in 1 kilometre bins of distance

[38]The transaction counts in zip codes located farther from CCs also increase sharply in the post-treatment period but only 7 months after policy implementation.

from CCs. As the differences are computed within zip codes, Figure 7 essentially presents the unconditional first difference of the treatment's impact on digital payments. To ascertain whether zip codes located in districts with high informality responded most to the treatment, we disaggregate our sample by districts' informality (share of rural households). In Figure 7, the outcome of interest is transaction volumes, while in Figure A6 the outcome of interest is the number of transactions. We see a positive relationship between the change in digital payments post-treatment – both in terms of transaction counts and volumes – and zip codes' distance from CCs, but *only* for zip codes located in districts with low informality. The positive relationship is strongest in zip codes located in districts where the share of rural households is in the bottom quartile. The relationship is particularly weak in zip codes with the highest levels of informality – zip codes located in districts in the top quartile of rural households – and modest for zip codes located in districts between the 25th and the 75th percentile. The figures are almost identical if we measure informality using the share of informal sector workers in the district (Figure A11, Appendix 7).

Collectively, the descriptive trends suggest that while the treatment had induced a persistent long-run impact on household adoption of digital payments, the treatment effects were concentrated in regions with relatively low informality and the treatment had a modest effect in areas with high informality. We now describe the empirical strategy to test the descriptive trends through a rigourous identification of the treatment effects.

4 Empirical Strategy

This section presents the empirical strategy to identify the causal impact of an increase in the cost of cash transactions on households' participation in the formal economy.

4.1 Cash Supply Shock and Digital Payments

The treatment's impact on digital payments is identified using the reduced form specification:

$$ln(Y_{idst}) = \alpha_i + \delta_{st} + \theta_d t + \beta HighDistCC_i * Post_t + \phi f(ATMPOS_{idst}) + \epsilon_{idst} \tag{2}$$

In (2), the unit of observation is zip code i, located in district d and state s,[39] and observed in month-year t. The primary outcomes of interest are logged number (volume) of transactions conducted through POS terminals. *Post* is a dummy equaling 1 for all month-years in the post-treatment

[39] States and districts form the second and third tier of administration in India. Districts are comparable to the U.S. county.

period (post November 2016). Our preferred specification splits the post period into "near-term" and "long-term" – the former captures the one year period succeeding the shock - namely November 2016 to November 2017; the latter covers the remaining 6 months between November 2017 and April 2018. This permits us to identify the persistence of the shock on digital payments. *HighDistCC* is a dummy equaling 1 if the zip code is located relatively far from CCs, signifying a higher exposure to the treatment. The classification is based on the Euclidean distance between a zip code and the nearest CC, described in Section 3.1.

Specification (2) includes zip code and state-month-year fixed effects – α and δ – along with district-specific linear time-trends – $\theta_d t$. These allow us to flexibly control for time-invariant zip code characteristics, month-year specific factors secularly affecting all zip codes within the state, and factors varying linearly over time within districts (but uniformly across zip codes). The state-month-year fixed effects control for any state-level policies which are not targeted to any particular region. f is a fourth-order polynomial in the number of POS and ATM terminals operating within the zip code. This controls for any targeted roll out of financial infrastructure in the post-treatment period.[40] The threat to causal identification from omitted variables arises from time-varying factors within zip codes such as public policies targeted to select zip codes within districts.[41] The identifying assumption for a causal interpretation of β is that changes in digital payments would have been identical across zip codes located relatively near and far from CCs in the absence of the treatment.

In (2), β identifies the average percentage change in transaction levels conducted from POS terminals in the post-treatment period in zip codes located farther from CCs, relative to zip codes located near CCs. The identification strategy is subject to two concerns regarding the actual location of CCs and the potential mismatch between zip codes and the nearest CC. First, while demonetization was wholly unanticipated, the physical location of CCs is most likely to be strategic. In particular, we would expect CCs to be located in zip codes with relatively high economic activity. This is confirmed in Table A1 (Appendix 7) where we see that zip codes located near CCs have a significantly higher population, and a higher likelihood of being located in a tier-I city, indicating that they would have higher levels of economic development.[42] We contend however that this biases us against rejecting our null hypothesis: if zip codes located farther from CCs have lower economic activity and the penetra-

[40]While there is limited growth in the number of ATM terminals in this period, there was a sharp increase in the months of December 2016 and January 2017 in the number of POS terminals.

[41]A potential candidate is the rural employment guarantee scheme – MNREGA. If the government disproportionately targeted MNREGA to zip codes located farther from currency chests in this period, we would be unable to control for that in the empirical specifications.

[42]Tier-I cities are cities with a population exceeding 100,000.

tion of digital technology is positively correlated with economic development, we would also expect a lower level of transactions from POS terminals in such places.[43] This implies that the β estimated in (2) would possibly be estimating the lower bound of the treatment's impact on digital payments.

The second concern relates to measurement error arising from a mismatch between zip codes and the nearest CC. Thus, it is plausible that a zip code is nearest in terms of Euclidean distance to a certain CC but the bank branches operating in that zip code receive currency from a separate CC, located at a farther distance.[44] In such situations, it's possible that we are erroneously classifying a zip code as "low distance" (*HighDistCC* = 0) when in reality it is "high distance" (*HighDistCC* = 1). This classical measurement error should attenuate the β in (2) towards 0, conditional on the mismatching being orthogonal to other factors affecting POS transactions in the post-treatment period. Nonetheless, we empirically verify that our findings are not driven by any spurious correlation through a placebo test in Section 6 where we randomly assign zip codes to "high" and "low" distances from CCs.

In addition to identifying the average treatment effect on digital payments, we also estimate a month-by-month impact of the treatment using a distributed-lag framework.

$$ln(Y_{isdt}) = \alpha_i + \delta_{st} + \theta_d t + \sum_{j=-11}^{18} \beta_j HighDistCC_i * Shock_{t+j} + \phi f(ATMPOS_{idt}) + \epsilon_{isdt} \tag{3}$$

In (3), β_j estimates the percentage change in digital payments in zip codes located farther from CCs, for each month j, relative to the month prior to the treatment (October 2016), which forms the reference period. This specification also tests for pre-trends in the outcomes: if our identification strategy is valid, we would expect $\beta_j = 0; \forall j \in \{-11, .., -2\}$.

4.2 Differential Impact of Cash Supply Shock on Digital Payments Across District Characteristics

We identify whether the treatment induced informal households to adopt digital payments by testing for differential treatment effects across regions which had a relatively high degree of informality based on pre-treatment characteristics. As described in Section 3.3, we being by focusing on 3 district-level

[43]This would be particularly true if the availability of the necessary financial infrastructure such as POS terminals — essential for conducting digital transactions — is positively correlated with local economic development.

[44]This is possible as ATM terminals receive their supply of currency from bank branches and not directly from CCs. The mismatch is more likely particularly in the presence of natural obstacles and other topographical factors can give rise to such a situation. Thus, if the closest CC to a zip code is located on the other side of a hill through which no motorable road passes, the zip code might be receiving currency from an alternate CC, which is possibly located at a greater Euclidean distance, but involving lower travel time. Other factors giving rise to such mismatches is if bank branches in the zip code opt to receive currency from currency chests maintained by their own banks.

pre-treatment indicators of informality sourced from the NSS — namely the share of rural households, share of workers employed in the informal sector, and the share of self-employed workers. To test the channels which generate the differential response to the shock, we also identify heterogeneous treatment effects across regions with high informality, conditional on pre-treatment levels of corruption and financial infrastructure.

For each characteristic of interest, we group districts as either "high" or "low", based on the district's ranking relative to the sample median. Thus, a district is classified as "high rural" if the district's share of rural households in 2011-12 exceeds the median share of rural households in 2011-12 across all districts. Based on this classification, we use a triple-interaction to test for heterogeneous treatment effects across pre-treatment regional characteristics:

$$\begin{aligned} ln(Y_{idst}) = \alpha_i + \delta_{st} + \theta_d t + \beta_1 HighDistCC_i * Post_t + \beta_2 Post_t * DistChar_d^k \\ + \beta_3 HighDistCC_i * Post_t * DistChar_d^k + \phi f(POSATM_{idt})\epsilon_{idt} \end{aligned} \quad (4)$$

In (4), *DistChar* is a dummy denoting the district characteristic of interest and equals 1 if the district is classified as "high', as described above. The coefficient of interest is β_3, comparing the differential treatment effect across zip codes located in districts with a high value of characteristic k, conditional on the zip code being located far from CCs. β_1 estimates the direct effect of the treatment — in zip codes located far from CCs but in districts with a low-value of characteristic k; β_2 estimates the treatment effect in zip codes located close to CCs, but in districts with a relatively high value of characterisic k.[45] The remaining interaction terms are absorbed by the zip code and state-month-year fixed effects. The sum of β_1, β_2 and β_3, capture the net treatment effect in zip codes far from CCs and located in districts with a high value of characteristic k.

4.3 Cash Supply Shock and Household Outcomes

We adopt an approach similar to Sections 4.1 and 4.2 to identify the treatment's impact on household level outcomes. We use the household-level specifications to first infer the channels through which the treatment is operating and subsequently, to identify whether it generated spillovers in households' participation in financial instruments. One major distinction between the two approaches arise due to

[45] β_2 in this regard can be considered to capture the average post-treatment trend in digital payments for zip codes located in districts with a high value of characteristic k.

differences in the administrative level of the data. While the data on digital transactions is at the level of zip code, we can only map households to districts, necessitating a redefining of the cross-sectional variation in treatment intensity at the district level.

To obtain district-level variation in treatment intensity, we first define for district d, *DistrDistCC* as:

$$DistrDistCC_d = \pi_{DistCC_i}; \forall i \in d \tag{5}$$

In (5), π denotes the median and *DistrDistCC* represents the median value of *DistCC* - distance between the zip code and the nearest CC - across all zip codes located within district d. *DistrDistCC* therefore reflects the distance between the representative zip code located in the district and the nearest CC. We choose the median as opposed to the mean to avoid outlier values from influencing our distance measure in either direction. Based on this formulation, the median value of *DistrDistCC* across all districts is 20.8 kilometres.[46] The critical assumption for this to be an unbiased measure of district-level variation in treatment intensity is that the district's population is not concentrated in zip codes which are in either tail of the within-district *DistCC* distribution.

We exploit this district-level variation in treatment intensity to identify the impact of the treatment on household participation in financial instruments:

$$Y_{hdst} = \psi_h + \delta_{st} + \theta_d t + \beta HighDistrDistCC_d * Post_t + \phi X_{hdt} + \epsilon_{idst} \tag{6}$$

In (6), the unit of observation is household h, residing in district d and interviewed in month-year t. *HighDistrDistCC* is a dummy equaling 1 if $DistrDistCC_d > \pi_{DistrDistCC}$ where $\pi_{DistrDistCC}$ denotes the median value of *DistrDistCC* across all districts in the sample. ψ and δ denote household and state-month-year fixed effects while θ is a district-specific linear time trend.[47] X is a vector of household level covariates which can possibly affect the outcome of interest.[48] β measures the impact of the treatment on household outcomes in districts where the median zip code is located farther from CCs, relative to districts where the median zip code is located close to CCs. The identifying assumption for a causal interpretation of β is that household outcomes across districts where the median zip code

[46] The mean value of *DistrDistCC* is 22.2 kilometres.

[47] The month-year classification is based on the month of the survey.

[48] The covariate set includes the following household-level variables: number of female members; average household age; number of children in the household; number of elderly members; household size; average years of education in the household; a dummy equaling 1 if the household has any secondary educated member or member having completed higher education; dummies for whether any member of the household is hospitalized or on regular medication.

is relatively far from CCs would have varied identically in the absence of the treatment, relative to districts where the median zip code is located relatively close to CCs.

The primary household outcome of interest is the binary variable *AnyFin* equaling 1 if the household made any investment along the extensive margin in any financial instrument between months t and $t-4$. Apart from these, we also test the impact of the shock on a dummy equaling 1 if the household has made any investment in gold or real estate in the past 4 months. This allows us to observe whether the treatment induced households to switch from investing in physical assets to financial assets in the post-treatment period. To identify whether households operating in the informal sector respond to the treatment, we test for differential treatment effects across household characteristics in the spirit of specification (4). As the CP data has no information on whether household members are employed in the informal sector, we test for differential treatment effects across rural households to identify whether the treatment led to differential responses across households which have a high likelihood of operating in the informal sector.

5 Results

This section documents the key findings of the paper. We first establish that the policy intervention resulted in a sharp decline in cash withdrawals from ATMs in regions located farther from CCs. Next, we show the impact of the treatment on digital payments, measured through transactions undertaken at POS terminals, and test for differential treatment effects across regions with high informality. We then identify the channels which could have possibly dampened the treatment's impact on informal households' participation in digital payments. The final set of results examine whether the treatment generated spillovers in households' participation in financial instruments.

5.1 Cash Supply Shock and Adoption of Digital Payments

5.1.1 Treatment Effects on Cash Withdrawals

We first establish that zip codes with higher exposure to the treatment experienced a reduction in cash withdrawals from ATM terminals. This follows Chodorow-Reich et al. (2018), who contend that this testifies to the negative impact of the treatment on cash as banks held surplus cash in this period due to deposits of the discontinued currency. We start with the distributed lag specification in (3) which compares the impact of the treatment for zip codes located far from CCs, versus those near CCs, individually for each month. The coefficients are plotted in Figures A7A and A9A (Appendix 7). The month preceding the treatment - October 2016 - is taken as the base period. Figures A9A (Appendix 7) and A7A establishes a steep decline in both the number, and the volume of ATM transactions undertaken in zip codes located far from CCs in the immediate aftermath of the treatment. For instance, in the month of the treatment, ATM transactions, (both the number and volume) were 40 percent lower in zip codes relatively far from CCs and remain at least 10 percent lower, in the first 9 months following the treatment. Critically, we identify no differential trends in cash withdrawals in the 9 months prior to the treatment intervention across zip codes located near and far from CCs.

Table 2 estimates the average effect of the treatment on ATM withdrawals across the entire time period using (2). The outcome variable in columns (1)-(2) is measured in transaction counts while columns (3)-(4) measure the outcome as transaction volumes. Columns (1) and (3) estimate a parsimonious specification with only zip code and state-time fixed effects while columns (2) and (4) include a fourth-order polynomial in the number of ATM and POS terminals in the zip code for each month, and a full set of district-specific time-trends. While the inclusion of district-time trends cause some attenuation in the long-term coefficients, the results document a sharp decline in cash withdrawals in zip codes located farther from CCs in the first 12 months of the treatment ("near term"). This reverses over the next 4 months ("long-term") as the cash supply returns to its pre-treatment levels: while there is small increase in the number of transactions from ATM terminals, the effect on the volume of transactions is a precise 0.

The results in Table 2 reaffirms the findings of Figures A7A and A9A (Appendix 7) and establishes what can be considered as the "first stage" of the paper – the treatment resulted in a sharp decline in cash withdrawals from ATMs in the 12 months succeeding the shock in regions located relatively far from CCs. That this occured in a period of rising bank deposits indicates that these regions faced a severe reduction in currency availability (Chodorow-Reich et al. (2018)). We now estimate whether households adjusted to this reduction in cash availability by switching to the use of debit/credit cards to conduct their economic transactions.

5.1.2 Treatment Effects on Digital Payments

We begin by identifying the month-by-month impact of the treatment on digital payments using the distributed lag specification in (3) and plot the monthly coefficients in Figures A7B and A9B (Appendix 7). Consistent with the descriptive trends in Figure 5, Figure Figures A9A (Appendix 7) and A7B records a sharp jump in both the number, and the volume of transactions conducted through POS terminals. In terms of magnitudes, the number (volume) of transactions from POS terminals rises in excess of 20 (30) percent in the month of the treatment in zip codes farther from CCs (and thereby more exposed to the shock). The differential increase in digital payments in zip codes located farther from CCs in response to the treatment also rule out that the decline in ATM withdrawals in these areas were driven by an aggregate demand shock. This bolsters our claim that the decline in ATM withdrawals in zip codes further from CCs was attributable to the shock due to slower replenishment of currency to these regions (Chodorow-Reich et al. (2018)).

Interestingly, while the increase in POS transactions (number and volume) is persistent over the long-term, the 95 percent confidence intervals for the coefficients overlap for almost all months $3 \leq t \leq 18$, ruling out that the treatment generated growth effects in digital payments, while irreversibly increasing the transaction levels. Importantly, akin to the results on ATM withdrawals, we also find little evidence of differential pre-trends in POS transactions. Only 3 out of the 18 pre-treatment coefficients are statistically significant at the 5 percent level and there is no discernible differential trend in the pre-treatment period across the two sets of zip codes. This lends support to our DiD design and suggests that the observed differences between zip codes located further from CCs, versus those located near CCs, reflect time-invariant differences which are absorbed by the zip code fixed effects, and the two sets of zip codes exhibit no differential pre-trends in outcomes during the lead-up to the treatment.

We use (2) to estimate the average treatment effect across the entire post-treatment period. The baseline results are reported in Table 3. The outcome variable in the first two columns is measured as the number of transactions from POS terminals while the latter two columns measure transaction volumes. Columns (1) and (3) use a parsimonious specification with only zip code and state-time fixed effects while columns (2) and (4) include district-specific time trends and a fourth-order polynomial in the number of POS and ATM terminals in the zip code. Standard errors are clustered by zip code.

In our most preferred specification which includes district time trends, the treatment in the near-term results in a 15 (18) percent increase in the number (volume) of transactions undertaken from POS terminals located in zip codes with a higher exposure to the treatment (located farther from CCs). The impact is persistent over the long-term, confirming that there is no abatement in the treatment's impact on digital payments even when the level of cash withdrawals from ATMs across these two sets of zip codes are statistically indistinguishable. This affirms a limited switching out amongst households which adopted digital payments post-treatment, once the cash supply started returning to its pre-treatment levels. In terms of absolute values, the coefficients suggest that relative to the pre-treatment means in zip codes located close to CCs, zip codes located far from CCs witnessed an additional 100 transactions from POS terminals over the long-run, worth Rs. 203,786.[49]

The lack of zip code-level demographic or economic data precludes us from testing the stability of our coefficients to the inclusion of zip code level time-varying covariates, with the exception of ATM and POS terminals. In this regard, we verify the stability of our coefficients to the inclusion of district-level covariates in Table A2 (Appendix 7), where we replace the district-specific time trends with a rich set of district covariates measured in 2011-12, interacted with a post-treatment indicator.[50] The results from this alternate specification is almost identical to the baseline results in Table 3.

Finally, Table A3 (Appendix 7) show our most restrictive specification where we replace the state-month-year fixed effects and district-specific linear time trends with district-month-year fixed effects. This fully controls for all time-varying factors within districts and the results remain unchanged relative to the baseline results. The fact that neither the inclusion of district covariates, nor the district-month-year fixed effects alter our main results assures us that our baseline results are not generated by any omitted time-varying factors at the district-level.

5.1.3 Differential Effects Across Zip Codes in Districts with High Informality

Sections 5.1.1 and 5.1.2 established that the treatment caused a large and persistent increase in digital payments in zip codes which were most affected by the cash supply shock. We now test the key question of the paper: was the increase in digital payments driven by households which had a relatively higher likelihood of operating in the informal economy? We examine this first by testing for differential treatment effects across regions with relatively high informality based on pre-treatment district characteristics. If the treatment effects are amplified in zip codes located in districts with relatively high informality, we can infer that the treatment was effective in inducing household participation in the formal economy (through digital payments) amongst households which had a higher likelihood of operating in the informal sector. We undertake this exercise using (4) and restrict the outcomes of interest to transactions undertaken through POS terminals.

Consistent with the descriptive trends in Figures A5 and 7, the results in Table 4 inform that the

[49] These are calculated as 0.168*591.75 and 0.226*901,706.4, respectively.

[50] In Table A2 (Appendix 7), we replace the district-specific time-trends with the following district-level covariates from the NSS 2011-12 surveys: share of rural population, share of formal sector workers, share of wage workers, share of unemployed workers, share of self-employed workers, share of individuals with secondary or higher education and household monthly per capita consumption expenditure. We also include bank branches per capita in the district in 2015-16.

treatment effect was limited in regions with relatively higher informality. Across columns (1)-(6), we identify a negative coefficient on the triple interaction for zip codes located in districts with a relatively high share of rural households, informal workers and self-employed workers. This implies that the adoption of digital payments in response to the treatment was significantly lower in zip codes located in districts with relatively higher levels of informality. Note that the direct effect of the treatment – β_1 in (4) – is in each instance positive and statistically significant, indicating that the treatment effects were concentrated in zip codes located in districts with relatively more urban households, and lower share of informal sector workers (this is directly shown in Table A5 (Appendix 7).

These results are confirmed in Figures A12 and A13 (Appendix 7) where we plot the triple interaction coefficients using the distributed lag specification. The plots show that while there is no differential impact of the treatment on ATM withdrawals in zip codes located in districts with high informality, the impact of the treatment on digital transactions in such zip codes is significantly lower. Thus, while the treatment resulted in comparable reductions in cash supply in regions with high and low informality, its impact on digital transactions was significantly lower in regions with high informality.

5.2 Mechanisms

The previous two sections established that while the treatment generated a significant and persistent increase in the level of digital payments, the impact was significantly lower in regions where households have a relatively higher likelihood of operating in the informal sector. We now identify the possible channels which might have muted the treatment's impact in regions with relatively high informality. Section 2.4 posited three possible channels which might limit the impact of the treatment - namely income hiding motives, reduction in aggregate consumption and the lack of financial infrastructure. We now evaluate the role of each of these channels in explaining the limited impact of the treatment in regions with relatively higher informality.

5.2.1 Treatment Effects in Regions with High Corruption

The government's rationale for undertaking the policy of demonetization was to detect undisclosed wealth stored as cash. To this effect, the strong anti-corruption pitch surrounding the policy might have deterred individuals from undertaking transactions in the post-treatment period. This would be particularly true if individuals with a high level of undisclosed income/assets chose to avoid detection by minimizing their formal sector transactions (through digital payments or the formal banking system). If the primary source of undisclosed income is corruption and corruption and economic development are negatively correlated, we would expect regions with higher corruption to be less urban and also more prone to income non-disclosure. This can potentially explain the limited impact of the treatment in rural areas. To this effect, we test for differential treatment effects across rural areas, conditional on the pre-treatment corruption levels.

A key challenge in this regard is the lack of reliable data on relative levels of corruption, even at the aggregation of states. We adopt Fisman et al.'s (2014) strategy and use Transparency International's 2006 classification to sort states into high and low corruption bins, based on their ranking relative to the median corruption rank across all states.[51] The results in Table A6 (Appendix 7) rule out that the limited impact of the treatment in zip codes situated in relatively rural districts is limited only to the high corruption states. The triple interaction coefficient remains negative and statistically significant across both high and low corruption states, indicating that the limited effect of the treatment in regions with high informality is not explained by widespread income hiding in rural areas.

5.2.2 Treatment Effects on Household Consumption

The second mechanism considered is the consumption channel which can be affected either through the treatment's impact on aggregate incomes, or as a shock to household liquidity due to the reduction in cash supply. With regard to the income channel, Chodorow-Reich et al. (2018) document that the shock resulted in a 2 percent reduction in local GDP growth. If this is concentrated in the informal sector dependent rural areas, the treatment's impact on digital payments in rural areas can be limited due to an overall reduction in rural households' consumption. We explore if this is driving the limited take up of the treatment in rural areas by testing for differential treatment effects on household consumption in rural regions.

Data on household consumption is obtained from the CP and the dependent variable is the average monthly per capita consumption (MPCE) undertaken by households in the four months prior to the survey month. This provides us with three observations for every household in each calendar year and we restrict our sample to surveys undertaken between September 2014 and August 2018, providing us with 6 pre-treatment and 6 post-treatment observations per household. As we can only observe households' location at the level of the district and not the zip code, we measure households' exposure to the treatment based on the district's exposure to the treatment, described in Section 3.2.

[51] According to this classification, the states of Tamil Nadu, Haryana, Jharkhand, Assam, Rajasthan, Karnataka, Madhya Pradesh and Bihar are deemed to be "high corruption".

The treatment's impact on household consumption is shown in Table **??**. The dependent variable in each instance is measured in logs and each specification includes household and state-survey month fixed effects, along with district-specific time-trends and household specific demographic covariates.[52] Standard errors are clustered by household. Akin to Section 5.1, we disaggregate the post-treatment period into the near and the long-term.

Column (1) of Table **??** shows that households residing in districts with a higher exposure to the treatment witnessed a moderate increase in consumption in the near term (one year post treatment), and a slightly larger reduction over the long-term. The difference between the near and the long-term coefficients is statistically significant at the 1 percent level. However, as seen in column (2), the decline in consumption over the long-term is not concentrated amongst rural households. The triple interaction coefficients are small in magnitude and not statistically significant, ruling out that the limited treatment effects in rural areas are driven by a disproportionate drop in rural consumption.

Column (3) explores whether the treatment had differential effects across households possessing a credit card, which would have enhanced their ability to conduct transactions in the absence of cash. We define a dummy which equals 1 if the household had a credit card in the pre-treatment period and test for differential treatment effects across this subset of households. Expectedly, column (3) estimates that households owing a credit card witnessed an additional 3 percent increase in their household consumption in the near-term, while the decline in household consumption over the long-term was also relatively limited in such households. The results are consistent with the findings of Crouzet et al. (2018), who showed that households with access to financial infrastructure such as credit cards and digital payments were able to cushion themselves from the negative effects of the cash supply shock.

Columns (4) and (5) explore whether the treatment had differential effects on rural households' consumption, conditional on households' pre-treatment levels of liquidity. Here, we split our sample by households whom we expect to be liquidity constrained based on their pre-treatment income and savings. Specifically, we consider households to have low liquidity if both their per capita income and rate of savings in the pre-treatment period (September 2014-August 2016) fall below the median household per capita income and savings rate across the sample. Column (4) tests for differential treatment effects across rural households, conditional on the household having low liquidity in the pre-treatment period. While the triple interaction coefficient in the near-term is small and imprecisely

[52] The household covariates include: a) average age; b) average yrs of education; c) 5 dummies for whether the head of household is occupied in white collar occupation, farming, business, small business or wage labour; d) household size; e) number of females in the household; f) number of children in the household.

estimated, the long-term triple interaction coefficient is negative and significant. Thus, rural households which were likely to be liquidity constrained and had a higher exposure to the treatment faced a sharper decline in their consumption over the long-term. Conversely, for households not classified as liqudity constrained (column (5)), we are unable to detect any differential treatment effect across rural households.

In summary, the estimates from Table **??** rule out that the limited impact of the treatment on digital transactions in rural areas is driven by an aggregate decline in consumption by rural households. While we are unable to fully rule out that rural households facing ex-ante liquidity constraints did not experience a disproportionate reduction in their consumption, it is worth noting that the negative impact of the treatment on digital transactions in zip codes located in districts with a high share of rural households was observed across both the near and the long-term, while the consumption decline occurs solely over the long-term. Thus, while liquidity constraints amongst rural households could have been a possible channel for lower consumption over the long-term (and consequently, lower digital transactions), it still does not explain the limited impact of the treatment on digital transactions over the near-term in regions with a high share of rural households.

5.2.3 Treatment Effects in Regions with Low Financial Infrastructure

The final channel examined to explain the limited impact of the treatment on digital transactions in rural areas is the relative lack of financial infrastructure, as seen in Figure A2 (7) In this regard, we test for differential treatment effects across rural zip codes after partioning our sample by the pre-treatment level of financial infrastructure.

Prior to that, we demonstrate in Table A7 (Appendix 7) that the treatment effects are amplified in zip codes located in districts with relatively high financial infrastructure. The triple interaction coefficients are positive and significant across all three measures of financial infrastructure – the standardized index of financial infrastructure combining POS terminals and bank branches per capita, and also individual measures of POS terminals and bank branches per capita. Importantly, across all measures of financial infrastructure, the direct effect of the treatment, particularly over the near-term is not statistically distinguishable from 0. This reflects that the positive impact of the treatment is driven by zip codes located in districts with high financial infrastructure.

Having established that the treatment effects are amplified in regions with high financial infrastructure, we now test whether the muted treatment effects in zip codes located in districts with high informality can be attributed to the lack of financial infrastructure. As districts in India are sufficiently large and it is possible for a district with a high share of rural households to have a handful of urban pockets which can drive the results, we conduct this exercise entirely at the level of the zip code. This also makes the unit of heterogeneity consistent with the paper's primary unit of analysis - the zip code. As described in Section 3.3, we use the number of deposit accounts and branch types from the BSR to determine whether a zip code is rural and construct the standardized index of zip code financial infrastructure.

Table A8 tests for differential treatment effects across rural zip codes and zip codes with relatively higher (above median) financial infrastructure, using the zip code level measures of heterogeneity (as opposed to the district). The results are consistent with those obtained in Tables 4 and A7 (Appendix 7) – the treatment effect is significantly lower in rural zip codes and significantly higher in zip codes with high (above median) financial infrastructure. Subsequently, we classify zip codes into three buckets of financial infrastructure – a) "low", where the financial infrastructure index falls below the 25th percentile; b) "intermediate", where the financial infrastructure index falls between the 25th and 75th percentile; and c) "high", where the financial infrastructure index lies above the 75th percentile. If the lack of financial infrastructure explains the limited treatment effects in rural zip codes, we would expect no differential treatment effects across rural zip codes, conditional on the zip code having high financial infrastructure.

The results in Table 5 are consistent with our hypothesis. We identify no differential treatment effect across rural zip codes once the sample is restricted to the top quartile of zip code financial infrastructure (columns (5) and (6)). The triple interaction coefficients are small and not statistically significant. The direct effect of the treatment remains positive and statistically significant and the sum of the coefficients too is positive and significant, confirming that the treatment had a comparable positive impact on digital payments across both rural and urban zip codes. Alternatively, in zip codes with either low or intermediate levels of financial infrastructure (columns (1)-(4)), we identify a negative and statistically significant coefficient on the interaction term in almost every instance, suggesting that the treatment's impact on digital transactions was lower for this subset of zip codes. Also, the direct effect of the treatment in columns (1) and (2) is almost 50 percent smaller to those estimated in columns (3)-(6), suggesting that the treatment had a muted impact on digital payments in the absence of financial infrastructure, even in relatively urban zip codes. The results therefore underline that the lack of financial infrastructure in rural areas served as a major hindrance to the adoption of digital payments in response to the shock and limited the treatment's impact on informality. In areas with sufficiently high financial infrastructure, the treatment was indeed successful in generating a positive impact on digital transactions, even in rural areas which would be expected to have high informality.

One potential concern with the empirical findings of this section is that a region's financial infrastructure could be correlated with other demographic and economic factors. Thus, one would expect districts with high financial infrastructure would also have higher per capita expenditures and a more educated workforce. Therefore, it is possible that our measure of financial infrastructure is serving as a proxy for these characteristics instead of capturing financial infrastructure. To assuage such concerns, we first show in Table A10 (Appendix 7) that even after restricting the sample to zip codes located in districts with relatively high per capita expenditures, a high share of educated workers, and a high share of salaried workers, the treatment effects are a) muted in zip codes located in districts with relatively low financial infrastructure and b) amplified in zip codes located in districts with relatively high infrastructure, particularly over the near term. This affirms the critical role played by financial infrastructure in enabling the adoption of digital payments, even in regions with relatively high consumption and a high share of salaried workers, reassuring us of the validity of our measure for financial infrastructure. We also re-estimate the results in Table 5 after explicitly controling for district-level factors, interacted with a post-treatment dummy. The results are shown in Table A9 (Appendix 7) and are qualitatively similar to those obtained in Table 5. This reassures us that our measurement of financial infrastructure is not a proxy for other regional characteristics such as income and education.

Taken in perspective of the prior findings in the literature, our results suggest that applying the "stick" and increasing the cost of cash transactions in the informal economy can induce households to switch to the formal economy, but conditional on the "carrots" in the form of financial infrastructure having been distributed. Our results signify that the availability of financial infrastructure is a necessary pre-condition for a shock comparable to demonetization to affect households' behaviour in adopting digital payments.

5.3 Spillover Effects on Household Participation in Financial Assets

This section identifies whether the treatment by mandating individuals to visit bank branches to deposit their discontinued currency also generated spillovers on household participation in financial instruments. This would occur if households chose to invest any excess liquidity, previously held as cash, in informal instruments. We use specification (4) and the outcome of interest is the binary indicator *AnyFin*, which equals 1 if the household invested in any financial instrument in the past 4 months. The specifications are estimated using a linear probability model. As the CP data only informs us about investments made along the extensive margin, the coefficients should be interpreted as changes in participation likelihoods in response to the treatment.

Column (1) of Table 6 shows that the treatment had no impact on households' likelihood of participating in financial assets over the near term and a small negative impact on households' likelihood of participating in financial assets over the long term. Columns (2) and (3) of Table 6 shows that the limited impact of the treatment on household participation in financial assets is not driven by the substitution of households away from financial assets towards physical assets: households residing in districts with a higher exposure to the treatment also had lower participation likelihoods in physical assets.

Column (4) tests for differential treatment effects across rural households who we expect would have a higher likelihood of operating in the informal sector. The results suggest that while there was no differential impact of the treatment on rural households' likelihood of participating in financial assets over the long-term, the treatment had an additional negative impact on the likelihood of rural households to participate in financial assets over the near term. The direct effect of the treatment in column (4) is positive and statistically significant, suggesting that urban households residing in districts with a higher exposure to the treatment exhibited increased participation in financial assets. This is in line with previous results documenting the limited impact of the treatment on digital payments in rural areas.

Column (5) tests whether the muted effect of the treatment on household participation in financial assets is linked to supply side constraints such as the availability of financial infrastructure. As all the financial instruments considered (barring post-office savings, stocks and mutual finds) are operated through commercial banks, we measure financial infrastructure using per capita bank branches in the district in 2016.[53] The triple interaction coefficient is positive and significant over both the near and the long-term, while the direct effect of the treatment is negative. This indicates that the treatment did induce households to participate in financial instruments, but only in districts with a high density of bank branches.

Finally, columns (6) and (7) test for differential treatment effects across rural households, conditional on the availability of bank branches. We undertake an approach similar to Section 5.2.3 where

[53] This is plausible as out of the 8 financial instruments considered in the paper, the highest levels of participation is in fixed deposits and provident funds, both of which are operated through the banking system.

we split our sample of households into those residing in districts with high (above median) and low bank branches per capita and subsequently, test for differential treatment effects across rural households. The results from column (6) however provide partial support to the hypothesis that the non-participation in financial instruments amongst rural households is driven by the limited availability of bank branches. For households located in districts with a high density of bank branches, the triple interaction coefficient is negative over the near-term and positive over the long-term. The direct effect of the treatment in these districts is positive over the near-term and 0 over the long-term. The coefficient is sizeable in magnitude reflecting a 25 percent increase in household participation in financial instruments for urban households residing in districts with relatively high bank branches per capita.[54] In contrast, column (7) shows that for households residing in districts with a low density of bank branches, the coefficient on the triple interaction is negative over both the near and the long-term and the direct effect of the treatment is also negative. Thus, the treatment induced participation in financial instruments only amongst urban households over the near-term, conditional on the availability of financial infrastructure. For rural households which have a higher likelihood of being employed in the informal sector, the treatment had a negative effect on participation in financial instruments, irrespective of their access to financial infrastructure.

A comparison of the results in columns (6) and (7) of Table 6 yet again underlines the critical role played by financial infrastructure to facilitate households' participation in financial instruments in response to an extreme shock to informality. The treatment induced higher participation in financial instruments only in districts with a high level of financial infrastructure in the form of bank branches. The "stick" again is effective in inducing formalization only when the "carrots" also are in place.

6 Robustness

This section tests the sensitivity of the baseline results to alternate distance thresholds and restrictions to the time period under consideration. We conclude with a placebo test to verify that our results are not driven by any spurious correlation arising due to mismatches between zip codes and the nearest currency chest.

[54] This is calculated as 0.062/0.24. The denominator is based on the mean level of household participation in financial instruments in the pre-treatment period for households living in districts which had a lower exposure to the treatment and a high bank branches per capita.

6.1 Robustness to Alternate Specifications

The paper's identification strategy rests on assigning zip codes to "high" and "low" treatment intensity, based on whether the distance to the nearest CC exceeds (is less than) the median distance to the nearest CC across all zip codes. The median distance between the zip code and the nearest CC in the sample is 11 kilometres and a potential concern arises whether the results are sensitive to this specific distance threshold. To guard against this possibility, we re-estimate (2) across a broad range of distance thresholds between 0 and 24 kilometres. Thus, when the threshold is 0 kilometres, all zip codes which have a CC within the zip code are assigned to "low" distance (*HighDistCC = 0*) and the rest are assigned to "high" distance (*HighDistCC = 1*). We plot the coefficients obtained from this exercise in Figure 10 (Appendix 7) and see that the results would have been almost identical had we opted for any threshold between 6 and 12 kilometres, instead of 11 kilometres. This rules out that the results are being driven by the specific 11 kilometre threshold used to classify zip codes as "high" and "low" distance from CCs.

Along similar lines, we test the sensitivity of our results to the omission of outlier zip codes which are located either very close to, or very far from CCs. We re-run our baseline specifications after restricting the sample to zip codes located in the middle quartile of the *DistCC* distribution – between 4.6 and 17.4 kilometres from the nearest CC. Table A11 (Appendix 7) show that the results remain unaltered to the exclusion of these zip codes.

We undertake a third robustness check by restricting our treatment period between November 2016 to June 2017. This is because the Goods and Services Tax (GST) was introduced in July 2017 and is perceived to have had a negative impact on the economy as the new regulations constricted business activity (Business Standard, 2018). If this is indeed true and the negative effect was correlated with distance from CCs, it could impact our results in either direction. We thereby restrict our time horizon to the pre-GST period and re-run the baseline specifications. The near term effects of the treatment on both ATM cash withdrawals and transactions from POS terminals remain very comparable, as seen from Table A12 (Appendix 7).

Our empirical findings of the treatment's impact is based on a difference-in-difference framework using monthly zip code data on digital payments. Bertrand et al. (2004) however caution researchers about the possibility of serial correlation in the outcome variable in such a longitudinal panel, which lead to biased hypothesis testing through a deflation of the standard errors. Their suggestion is to collapse the high frequency data to average pre and post-treatment outcomes. Likewise, we collapse our monthly zip code data into 3 periods – namely the pre-treatment, post-treatment near term (within 1 year post treatment) and post-treatment long-term (1 year post treatment). For each period, we obtain the within zip code monthly average value of transactions from ATM and POS terminals and re-run our baseline specifications using zip code and district-time period fixed effects. Table A13 (Appendix 7) confirms that our results remain unaltered if we collapse our monthly data into the pre and post-treatment time periods.

6.2 Placebo Test

In Section 4.1, the mismatch of zip codes to the nearest currency chest had been flagged as a key threat to the empirical strategy. We had argued then that any mismatch would be akin to a classical measurement error which would attenuate the treatment effect towards 0. Nonetheless, to verify that our findings on digital payments is not driven by any systematic mismatching of zip codes to currency chests, we conduct a placebo test by randomly assigning zip codes to "high" and "low" distances from CCs, and re-estimating specification (2) using the *HighDist* dummy generated by this random assignment and repeat this process 100 times. If the results are being driven by some spurious correlation between zip code-specific time-varying unobservables which are positively correlated with both digital payments and the distance to nearest CC, we would expect the results from this placebo test to replicate our findings.

We plot the coefficients from 100 placebo tests as an empirical CDF. Reassuringly, none of the coefficients estimated using the placebo test are close in magnitude to the coefficients estimated in Tables 2 and 3. Overall, out of the 100 specifications estimated, 5 (6) out of 100 near-term (long-term) coefficients are statistically significant at the 95 percent level – exactly what we would expect from a random allocation of zip codes to high and low distances from CCs. This exercise rules out that our treatment effects are generated through any systematic mismatching of zip codes to the nearest CC.

7 Conclusion

This paper empirically identifies whether an increase in the cost of informal transactions in a developing economy can induce households to switch to the formal economy. We use the demonetization episode in India as an exogenous treatment intervention which increased the cost of cash transactions due to an extensive shortage in cash supply. To generate cross-sectional variation in treatment intensity, we exploit the central bank's system of currency management and classify zipcodes with high and low exposure to the treatment.

Using proprietary data on digital payments made using credit/debit cards through POS terminals as our key measure of household participation in the formal economy, we see that the treatment had a positive and significant impact on digital payments in areas with higher exposure to the shock. Importantly, the treatment effects were persistent over the long-run, even as the cash supply returned to its pre-treatment levels. The treatment effects are amplified in regions with high financial infrastructure in the form of pre-treatment levels of bank branches and POS terminals per capita and muted in regions with high informality - namely zipcodes located in districts with a high share of rural households and a high share of informal workers. In the latter regions, the treatment has a positive effect on digital payments only if the district has a relatively high level of POS terminals per capita, and a relatively high share of households with a credit card.

We also test whether the treatment had any spillovers on household participation in financial instruments. We find that while the average treatment effect on household participation in financial instruments is negative, it becomes positive, particularly over the long-run in districts with high financial infrastructure in the form of bank branches per capita, even for rural households. The results in this regard highlights the critical role played by financial infrastructure for such a treatment to induce households residing in areas with high informality to participate in the formal economy over the long run.

References

Burgess, R. and R. Pande (2005). Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment. *American Economic Review 95.3*, 780–795.

Chodorow-Reich, G., G. Gopinath, P. Mishra, and A. Narayanan (2018). Cash and the Economy: Evidence from India's Demonetization. *National Bureau of Economic Research No. 25370*.

De Andrade, G. H., M. Bruhn, and D. Mckenzie (2014). A Helping Hand or the Long Arm of the Law? Experimental Evidence on What Governments Can Do to Formalize Firms. *The World Bank Economic Review 30.1*, 24–54.

De Mel, S., D. Mckenzie, and C. Woodruff (2013). The Demand for, and Consequences of Formalization Among Informal Firms in Sri Lanka. *American Economic Journal: Applied Economics 5.2*, 122–50.

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.

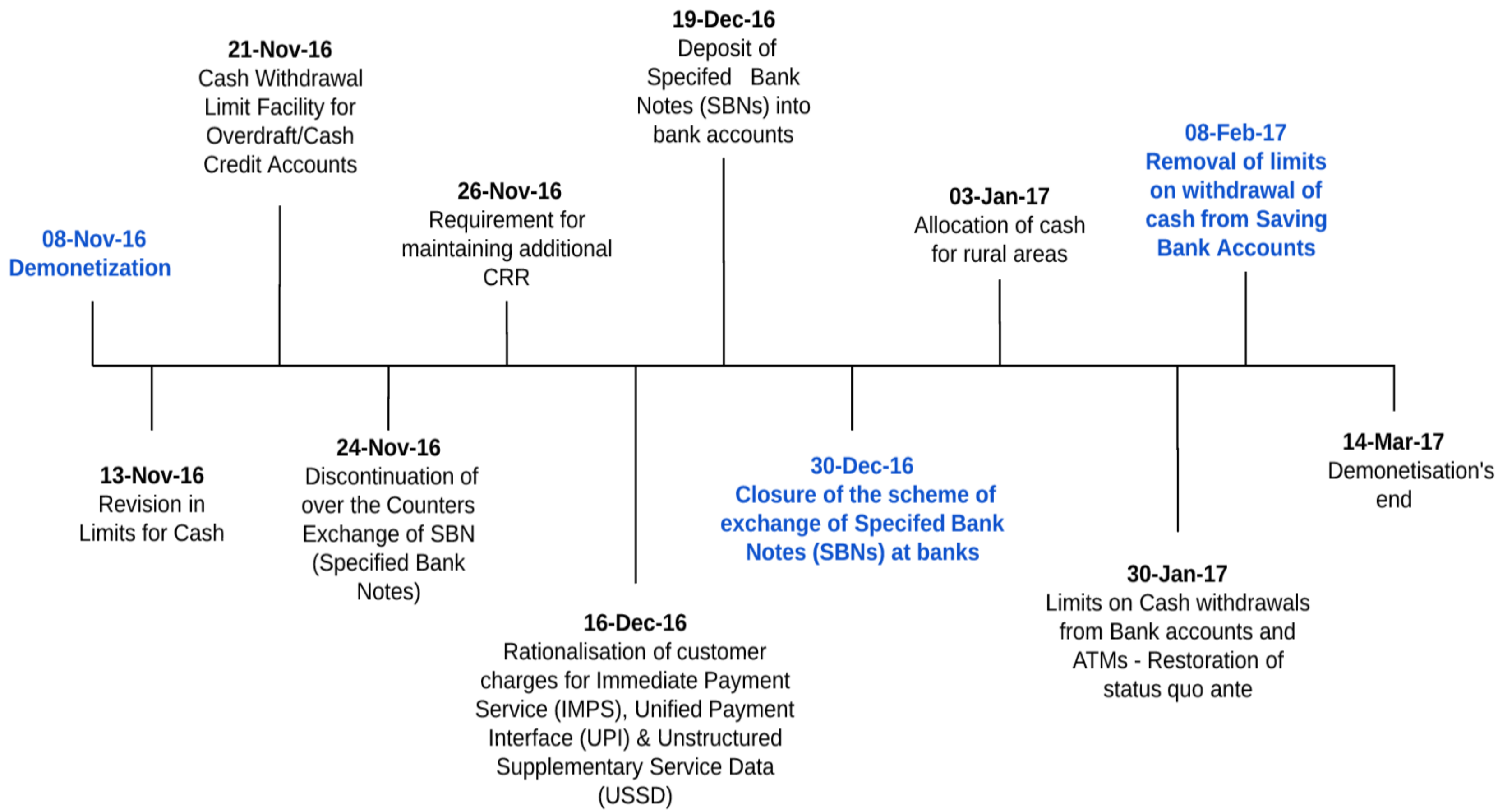
Jaramillo, M. (2009). Iss there Demand for Formality Amongst Informal Firms? Evidence from Microfirms in Downtown Lima. *German Development Institute Discussion Paper No. 0013*.

Kotwal, A., B. Ramaswami, and W. Wadhwa (2011). Economic Liberalization and Indian Economic Growth: What's the evidence? *Journal of Economic Literature 49.4*, 1152–99.

La Porta, R. and A. Schleifer (2014). Informality and Development. *Journal of Economic Perspectives 28.3*, 109–26.

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

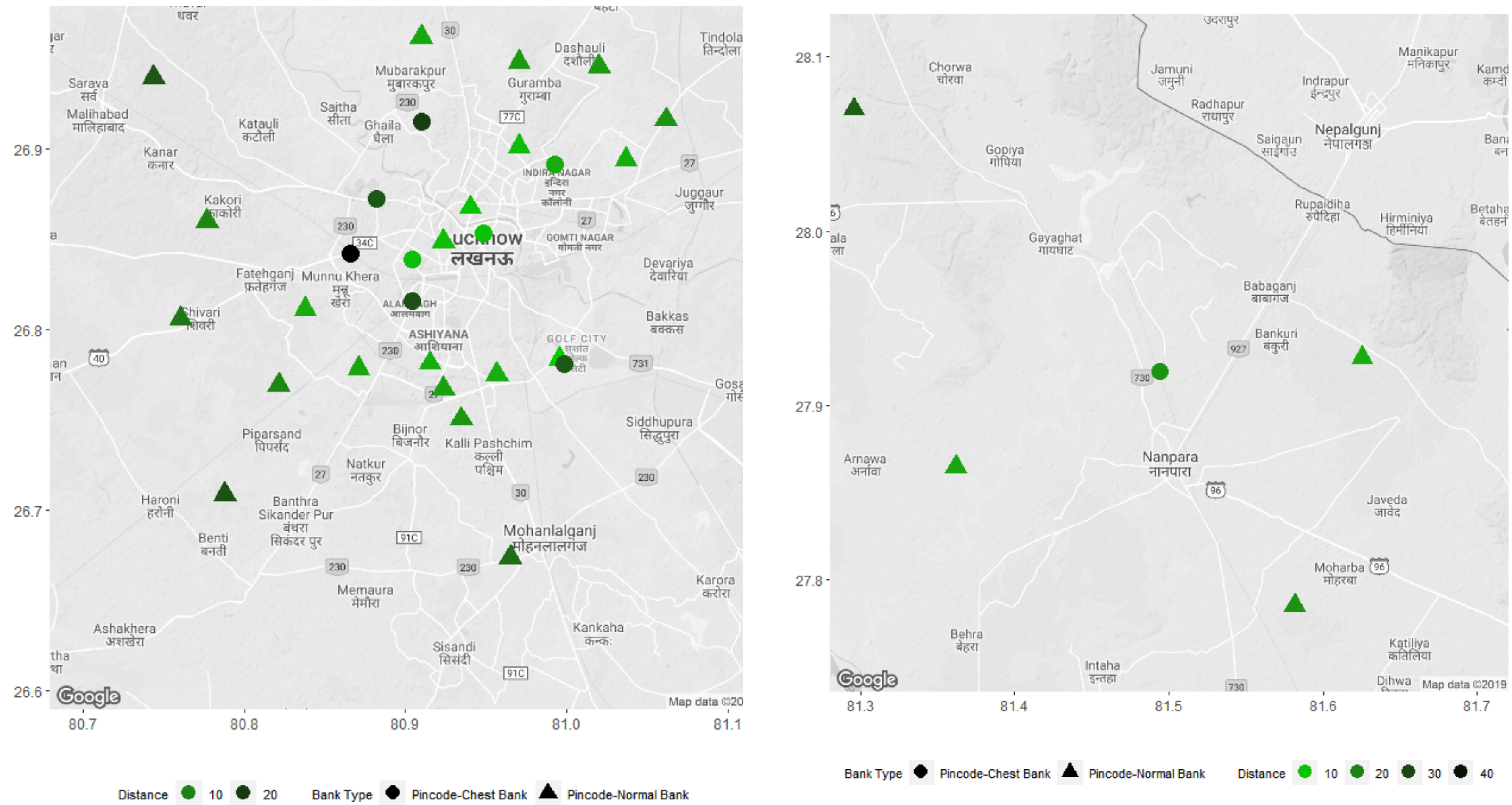
Figure 1: Timeline of events



The above figure shows the timeline of regulatory changes announced post demonetization. It starts from the day of the announcement of demonetization to the last official date. SBNs (specified bank notes) refer to the bank notes which were discontinued due to demonetization.

Figure 2: Distribution of currency chests and banks in representative districts

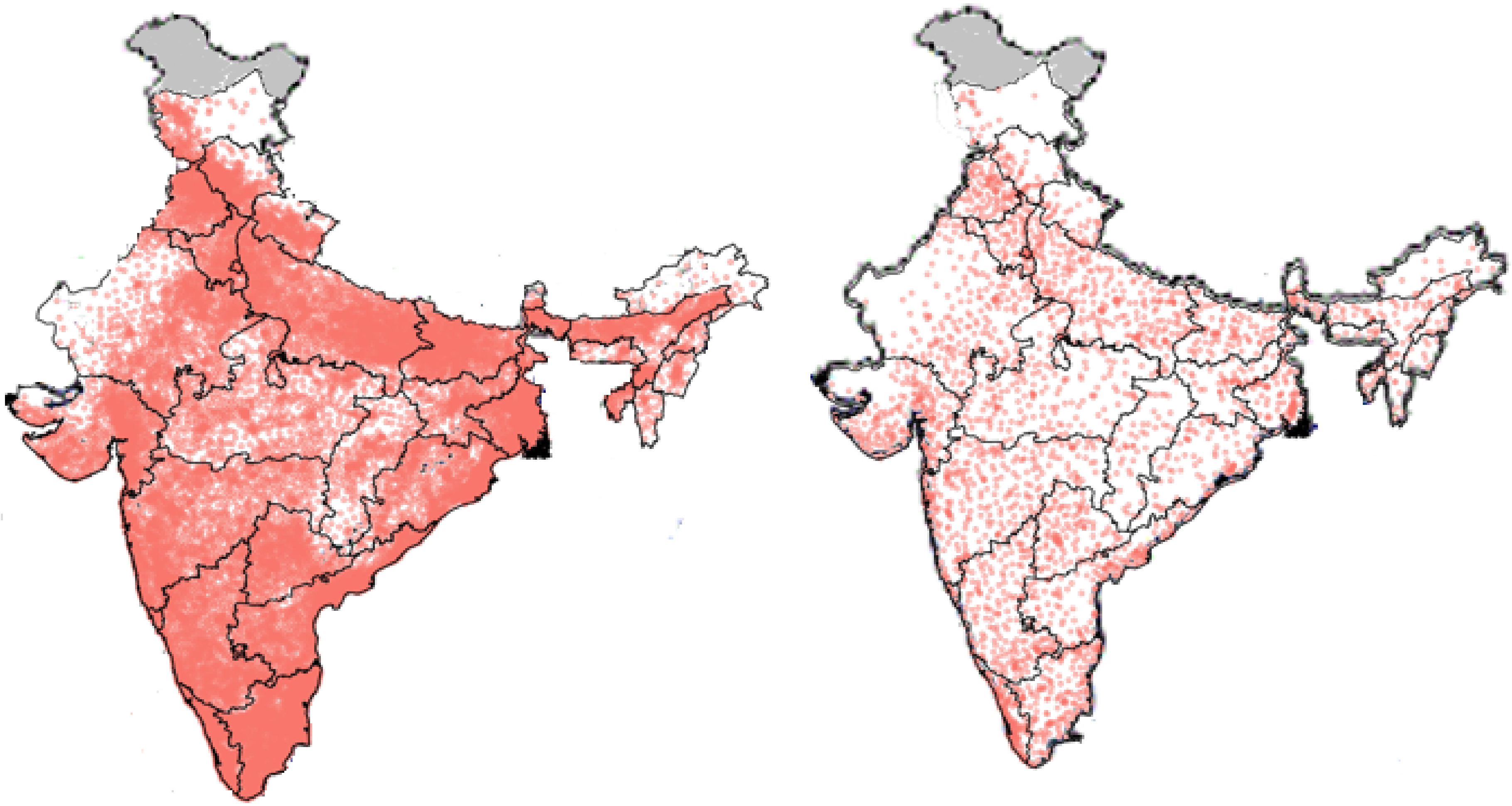
(A) **Lucknow (Urban District)** (B) **Bahraich (Rural District)**



The above figures plot the differences in the distribution of chest banks across two districts. Panel A shows the relatively urban district of Lucknow; panel B shows the relatively rural district of Bahraich. Triangles represent the zip codes of normal banks and circles represent the zip codes of chest banks. Darker shaded triangles represent normal banks that are located far away from a chest bank.

Figure 3: Map of zip codes used in analysis and location of currency chests

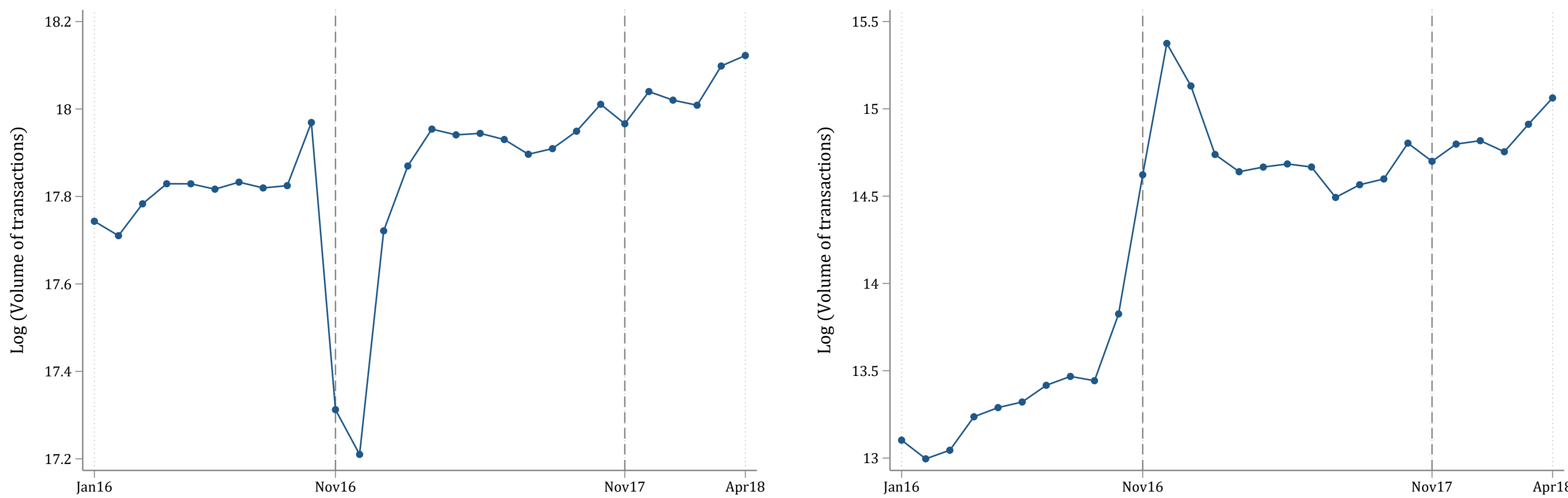
(A) **Zip codes in analysis** (B) **Location of currency chests**



The above figures plot the coverage of zip codes and currency chests across India. Panel (a) plots the location of the zip codes used in our analysis and for which data is available. Panel (b) shows the location of the currency chests. Shaded grey areas are areas for which there is no data available.

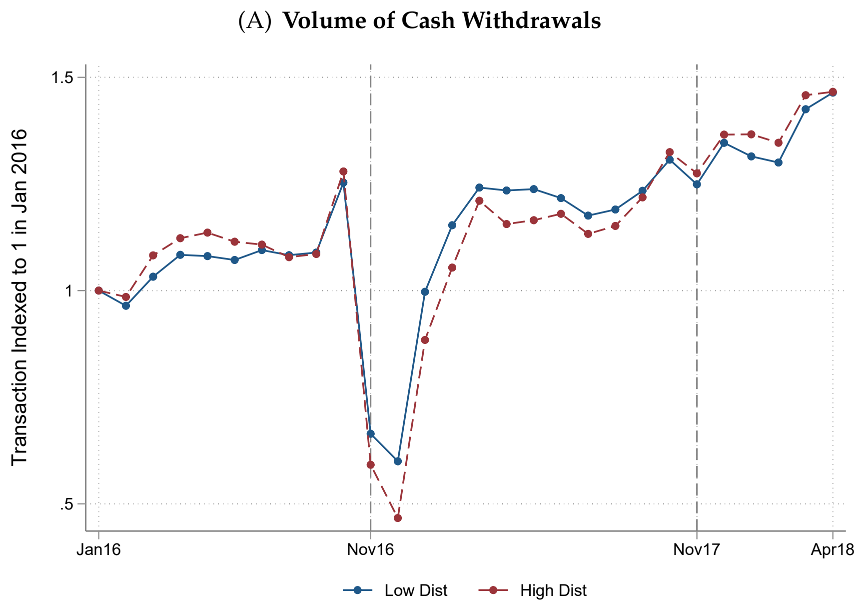
Figure 4: Monthly volume of cash withdrawals and digital transactions

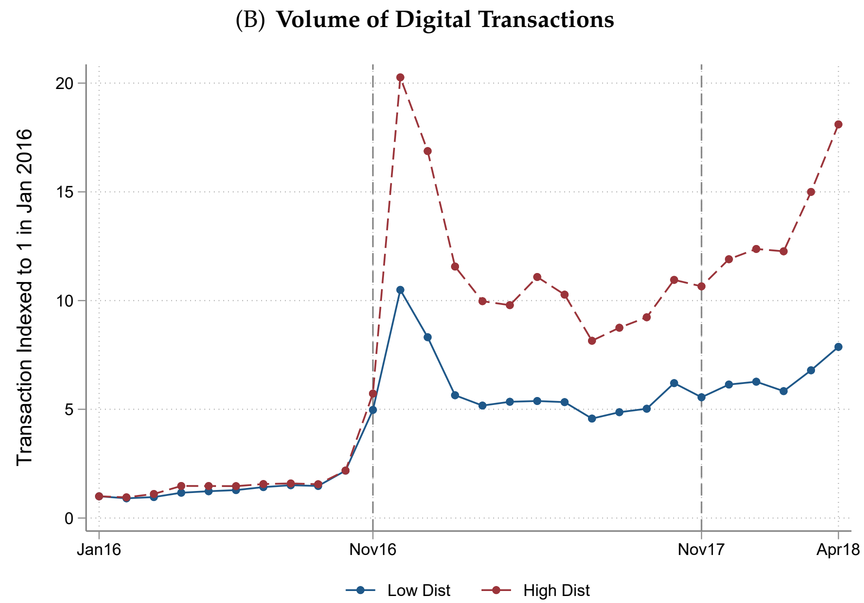
(A) **Volume of Cash Withdrawals** (B) **Volume of Digital Transactions**



The above figures plot the volumes of monthly cash and digital transactions between January 2016 and April 2018. Panel A represents cash withdrawals from ATM terminals; Panel B represents the volume of digital transactions from POS terminals.

Figure 5: Heterogeneity in volume of cash withdrawals and digital transactions by distance to nearest currency chest

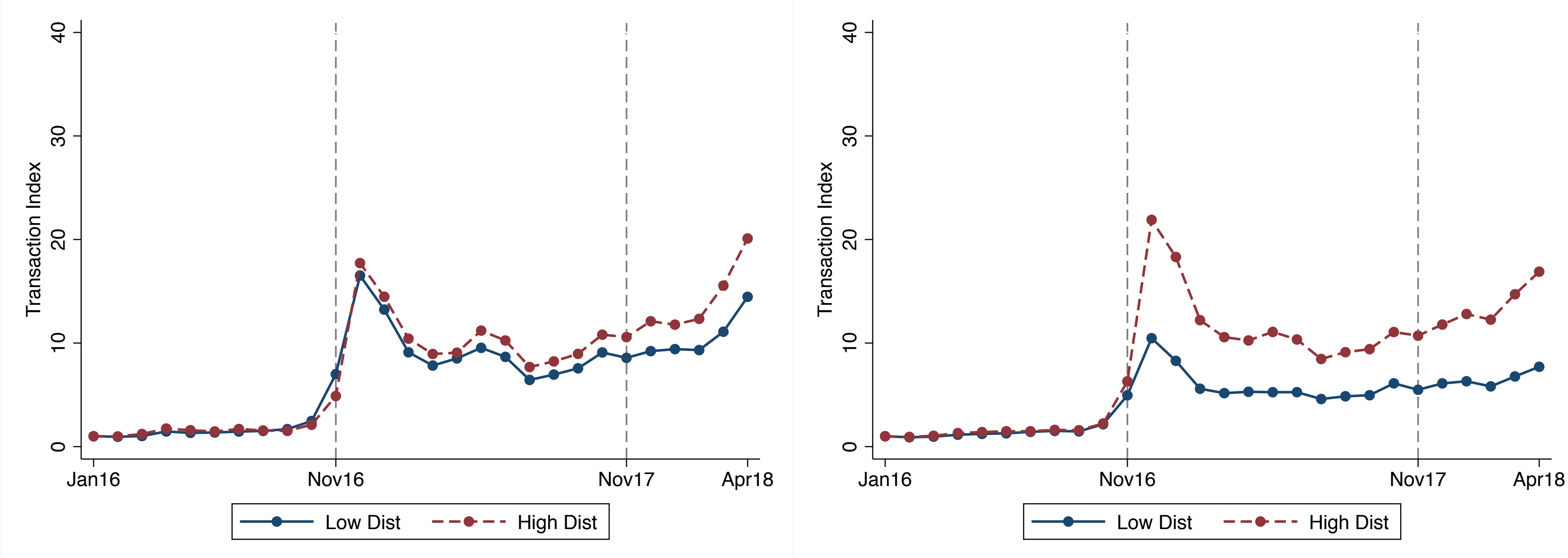




The above figures plot the monthly cash and digital transactions between January 2016 and April 2018, by the zip codes' distance from currency chests (CC). Panel A shows the volume of cash withdrawals from ATM terminals. Panel B shows the volume of digital transactions from POS terminals. Each point on the graph represents the total volume of transactions undertaken in each month, relative to January 2016 *High Dist* refers to pincodes where the pincode's distance to the nearest CC exceeds the median pincode-nearest CC distance in the sample.

Figure 6: Heterogeneity in effect on digital transactions across rural and urban districts: volumes

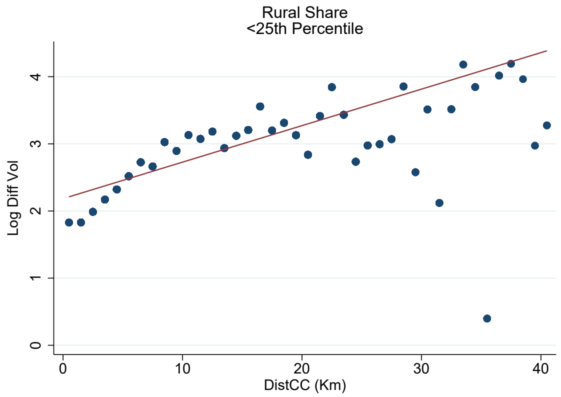
(A) **Volume of Digital Transactions: High Rural** (B) **Volume of Digital Transactions: Low Rural**



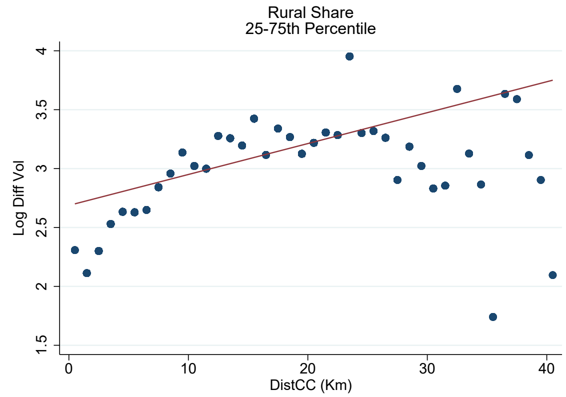
The above figures plot the volume of monthly digital transactions between January 2016 and April 2018, by zip codes' distance from currency chests (CC). Panel A splits the sample by zip codes located in districts with a relatively high share of rural households and Panel B splits the sample by zip codes located in districts with low share of rural households. Each point on the graph represents the total volume of transactions undertaken in each month, relative to January 2016. Districts are categorized into high and low rural based on the median share of rural households across all districts in 2011. *High Dist* refers to zip codes where the zip code's distance to the nearest CC exceeds the median zip code-nearest CC distance in the sample.

Figure 7: Heterogeneity in effect on digital transactions by share of rural households: Volumes

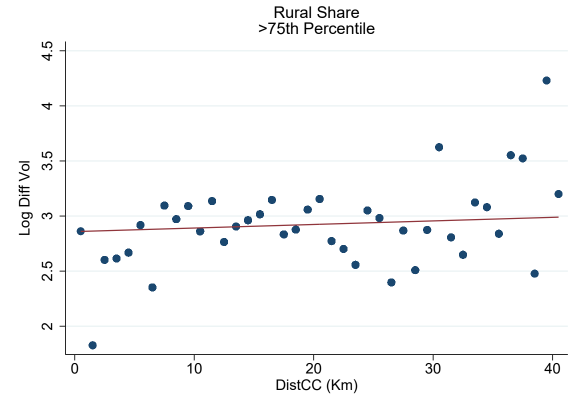
(A) **Rural Share: 25th Pctile**



(B) **Rural Share: 25-75th Pctile**



(C) Rural Share: 75th Pctile

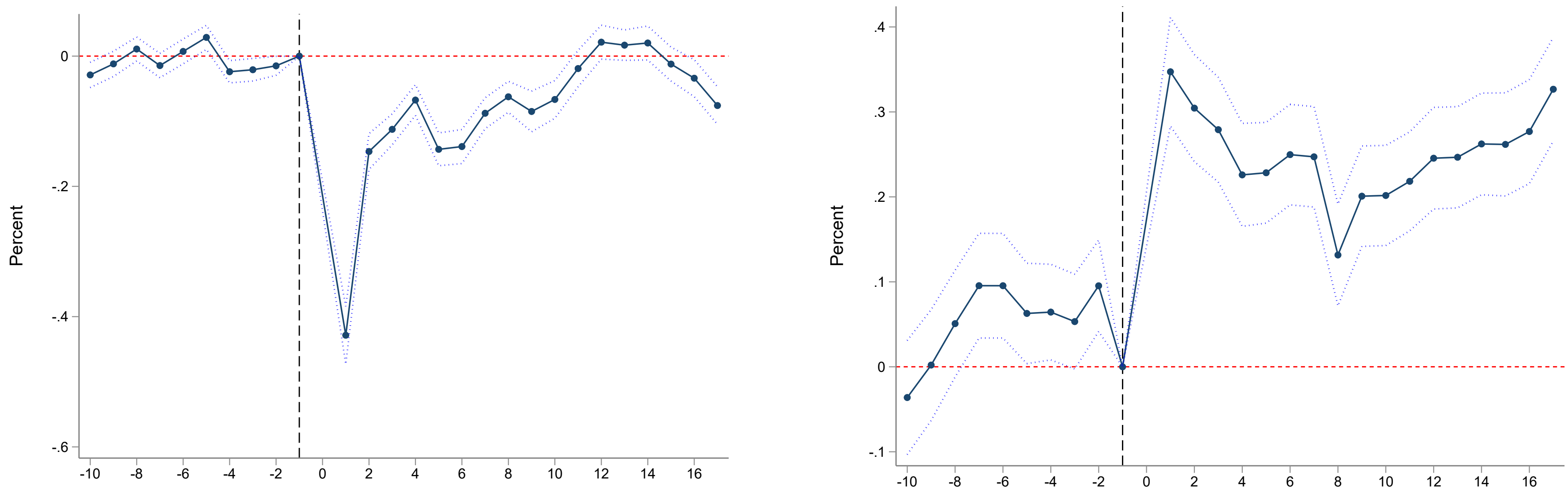


The above figures show binned scatter plots of the difference in the mean transaction levels from POS terminals across the post and pre-treatment periods. The differences are measured in logs. The x-axis represents the distance between zip codes and the nearest currency chest which serves as a proxy for access to cash. A higher distance implies lower access to cash. Panels A–C depicts changes in transaction volumes. Panel A restricts the sample to zip codes located in districts where the share of rural households is below the 25th percentile; panel B restricts the sample to zip codes located in districts where the share of rural households is between the 25th and 75th percentiles; Panel C restricts the sample to zip codes located in districts where the share of rural households exceeds the 75th percentile.

Figure 8: Effects on volumes of cash withdrawals and digital transactions

(A) **Volume of Cash Withdrawals**

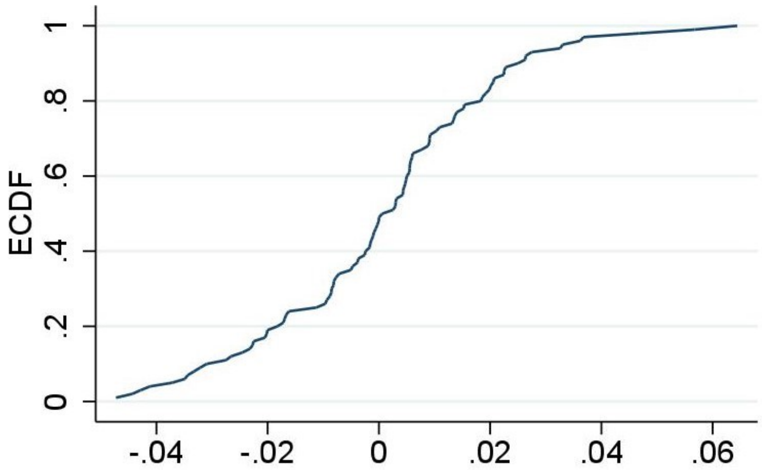
(B) **Volume of Digital Transactions**



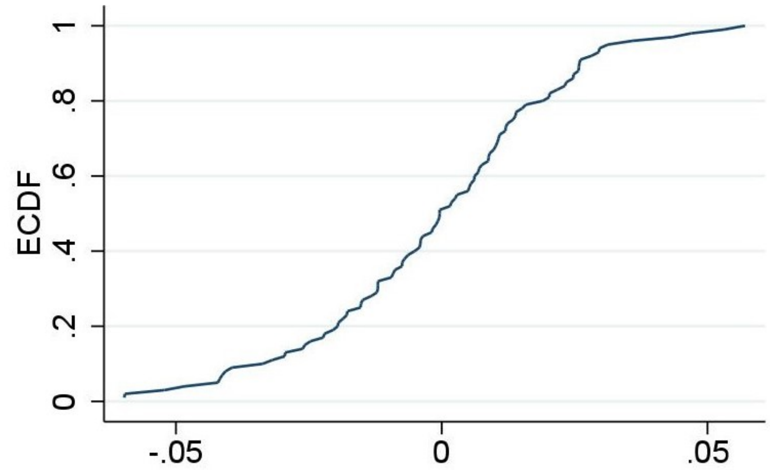
The above figures presents the monthly average treatment effects of the cash supply shock on cash withdrawals and digital transactions. The unit of observation is the zip code. Panel A shows the volume of cash withdrawals from ATM terminals. Panel B shows the volume of digital transactions from POS terminals. The outcome variable is logged in both panels. The light blue dotted line shows the 95 percent confidence interval and the x-axis is the months before and after the treatment intervention. The reference period is October 2016 - the month prior to the shock. All specifications include zip code and state-month-year fixed effects, along with district-specific time trends and a fourth order polynomial in the number of POS terminals in the zip code. Standard errors are clustered by zip code.

Figure 9: Placebo tests: Empirical CDFs from Random Assignment

(A) **Volume of Digital Transactions: Near Term**



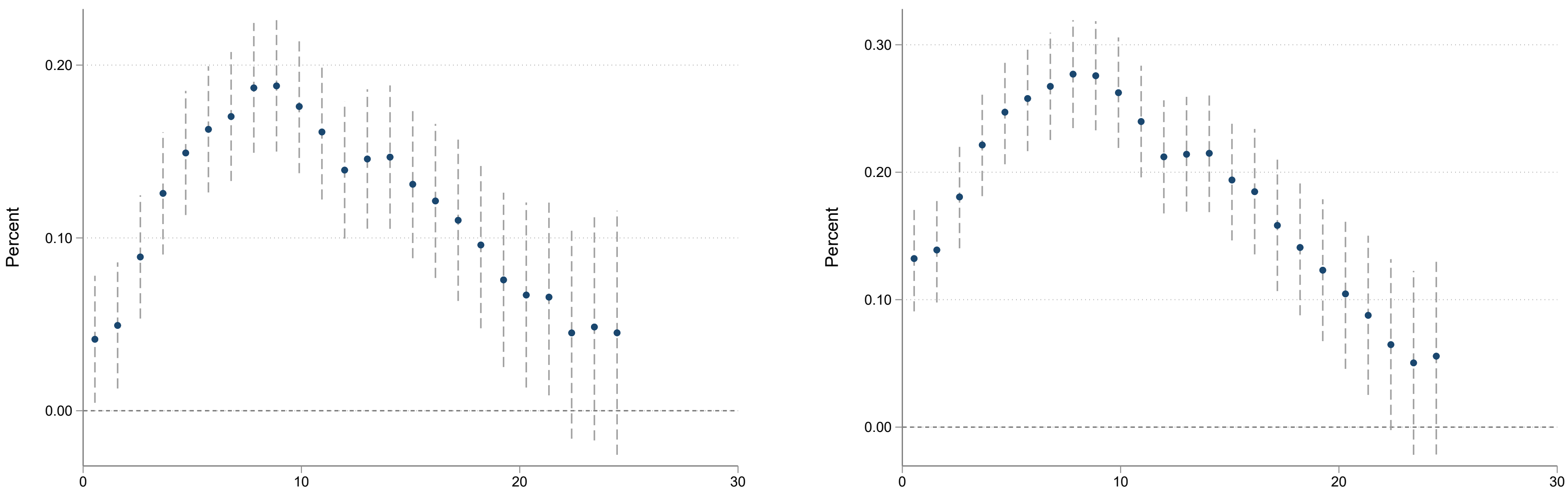
(B) **Volume of Digital Transactions: Long Term**



The above figures present the empirical CDFs from 100 iterations of the baseline results whereby zip codes are randomly assigned to "high" and "low" distance from currency chests. The x-axis shows the coefficient values. The true coefficient over the near-term for transaction covolumes) is 0.154 (0.191); over the long-term, for transaction counts (volumes) it is 0.168 (0.226). The unit of observation in each specification is the zip code. The outcome of interest is logged. Panel A presents the near-term coefficients; Panel B presents the long-term coefficients. The outcome of interest is transaction volumes. All specifications include zip code and state-month-year fixed effects, along with district-specific time trends and a fourth order polynomial in the number of POS terminals in the zip code.

Figure 10: Robustness of effect on digital transactions to alternate distance thresholds

(A) **Near Term** (B) **Long Term**



The above figures represent the robustness of the baseline results to alternate distance thresholds determining the assignment of zip codes to high and low distance from currency chests. The unit of observation is the zip code. The outcome of interest is the volume of digital transactions from POS terminals. Panel A presents the near-term coefficients and Panel B presents the long-term coefficients. The outcome variable is logged. The x-axis represents distance in kilometres. Vertical lines dotted represent 95 percent confidence intervals of the coefficients. All specifications include zip code and state-month-year fixed effects, along with district-specific time trends and a fourth order polynomial in the number of POS terminals in the zip code. Each coefficient corresponds to a 1 kilometre increase in the distance threshold used to determine which zip codes are located farther from currency chests and hence, more exposed to the treatment. The distance threshold in the baseline specifications is 11 kilometres.

Table 1: Descriptive Characteristics Based on Distance From Currency Chests

Panel A

	Low Distance	High Distance	Significance
Share of Zipcodes in Tier I Tier II Cities	0.17	0.05	$***$
Share of Zipcodes with At Least 1 Bank Branch	0.85	0.80	$***$
Number of POS Terminals	54.02	4.65	$***$
Number of ATM Terminals	19.90	4.72	$***$
ATM Transaction Counts (Rs. '000)	24.64	5.02	$***$
ATM Transaction Volumes (Rs. '000,000)	875.00	190.41	$***$
POS Transaction Counts (Rs. '000)	0.02	0.01	$***$
POS Transaction Volumes (Rs. '000,000)	9.61	0.73	$***$
Average ATM Transaction Value (Rs.)	3739.84	3566.35	$***$
Average POS Transaction Value (Rs.)	2026.23	1850.00	

Panel B

	Low Distance	High Distance	Significance
Share of Rural Households	0.63	0.81	$***$
Share of Low Caste Households	0.67	0.68	$***$
Bank Branches Per Million	91.85	65.84	$***$
Avg. Household Age	33.50	32.14	$***$
Total Children	0.97	1.07	$***$
Average Yrs of Education	6.14	4.91	$***$
Head of Household Graduate	0.11	0.07	$***$
Head of Household Not in Labour Force	0.24	0.18	$***$
Head of Household in Agriculture	0.29	0.42	$***$
Head of Household White Collar	0.06	0.05	$***$
Head of Household in Business	0.07	0.05	$***$
Head of Household in Small Business	0.06	0.05	$***$
Head of Household in Industrial Labour	0.28	0.29	$***$
Household Invested in Financial Assets	0.21	0.18	$***$
Household Invested in Gold	0.07	0.06	$***$
Household Invested in Real Estate	0.07	0.06	$***$

Panel A presents descriptive characteristics of zip codes based on distance from nearest currency chest. *Low Distance* zip codes are those for which the nearest currency chest is within 11 kilometres. *High Distance* zip codes are those for which the nearest currency chest exceeds 11 kilometres. Panel B presents descriptive characteristics of districts based on the distance of the median zip code in the district from nearest currency chest. *Low Distance* districts are those for which the median zip code's distance to the nearest currency chest is within 15 kilometres. *High Distance* districts are those for which the median zip code's distance to the nearest currency chest exceeds 15 kilometres.

Table 2: The effect on cash withdrawals

	(1)	(2)	(3)	(4)
	#		Vol.	
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	-.108***	-.107***	-.168***	-.116***
	(.008)	(.008)	(.009)	(.009)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	.045***	.040***	-.079***	-.014
	(.009)	(.008)	(.011)	(.010)
Observations	453764	354919	453764	354919
R^2	.92	.94	.89	.91

This table presents the results estimating the average treatment effect on ATM transactions. The unit of observation is zip code. The dependent variable is transactions from ATM terminals. The dependent variable in columns (1)-(2) is transaction counts; the dependent variable in columns (3)-(4) is transactions volumes. *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *DistCC* is the distance between a zip code and the nearest currency chest (CC). *High Dist CC* equals 1 if the zip code's distance to the nearest CC exceeds the median zip code-nearest CC distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state-month-year fixed effects. Columns (2) and (4) also include district-specific time trends and a fourth order polynomial in the number of ATMs. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table 3: The effect on digital transactions

	(1)	(2)	(3)	(4)
	#		Vol.	
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	.191***	.154***	.396***	.181***
	(.021)	(.017)	(.024)	(.020)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	.287***	.168***	.563***	.226***
	(.025)	(.019)	(.027)	(.023)
Observations	378240	354919	378240	354919
R^2	.89	.93	.85	.89

This table presents the results estimating the average treatment effect on POS transactions. The unit of observation is zip code. The dependent variable is transactions from POS terminals. The dependent variable in columns (1)-(2) is transaction counts; the dependent variable in columns (3)-(4) is transactions volumes. *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *DistCC* is the distance between a zip code and the nearest currency chest (CC). *High Dist CC* equals 1 if the zip code's distance to the nearest CC exceeds the median zip code-nearest CC distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state-month-year fixed effects. Columns (2) and (4) also include district-specific time trends and a fourth order polynomial in the number of POS terminals. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table 4: Heterogeneity of effect across zip codes with high informality

	(1)	(2)	(3)	(4)	(5)	(6)
	Measure of Informality (MoI)					
	Rural		Informal		Self-Employed	
	#	Vol	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	$.244^{***}$	$.287^{***}$	$.203^{***}$	$.239^{***}$	$.194^{***}$	$.230^{***}$
	(.020)	(.025)	(.020)	(.025)	(.021)	(.026)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	$.236^{***}$	$.319^{***}$	$.192^{***}$	$.277^{***}$	$.180^{***}$	$.264^{***}$
	(.022)	(.027)	(.023)	(.028)	(.023)	(.029)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Rural}$	$.084^{***}$	$.168^{***}$				
	(.026)	(.031)				
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Rural}$	$.066^{**}$	$.153^{***}$				
	(.033)	(.041)				
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Rural} * \mathbb{1}_{High\ Dist\ CC}$	$-.238^{***}$	$-.294^{***}$				
	(.033)	(.038)				
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Rural} * \mathbb{1}_{High\ Dist\ CC}$	$-.176^{***}$	$-.248^{***}$				
	(.039)	(.043)				
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Informal}$			$.134^{***}$	$.228^{***}$		
			(.024)	(.029)		
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Informal}$			$.119^{***}$	$.221^{***}$		
			(.031)	(.039)		
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Informal} * \mathbb{1}_{High\ Dist\ CC}$			$-.143^{***}$	$-.183^{***}$		
			(.032)	(.038)		
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Informal} * \mathbb{1}_{High\ Dist\ CC}$			$-.065^{*}$	$-.138^{***}$		
			(.038)	(.043)		
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ SelfEmp}$					$.160^{***}$	$.220^{***}$
					(.024)	(.030)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ SelfEmp}$					$.164^{***}$	$.223^{***}$
					(.032)	(.040)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ SelfEmp} * \mathbb{1}_{High\ Dist\ CC}$					$-.096^{***}$	$-.123^{***}$
					(.031)	(.037)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ SelfEmp} * \mathbb{1}_{High\ Dist\ CC}$					-.028	$-.084^{**}$
					(.036)	(.042)
Observations	352454	352454	352454	352454	352454	352454
R^2	.93	.89	.93	.89	.93	.89

This table presents the results estimating the differential treatment effect on POS transactions across zip codes located in districts with high informality. The unit of observation is zip code. The dependent variable is transactions from POS terminals. The dependent variable in columns (1), (3) and (5) is transaction counts; the dependent variable in columns (2), (4) and (6) is transactions volumes. *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist CC* equals 1 if the zip code's distance to the nearest currecy chest (CC) exceeds the median zip code-nearest CC distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state-month-year fixed effects along with district-specific time trends and a fourth order polynomial in the number of POS terminals. Columns (1) and (2) test for differential treatment effects across zip codes located in districts with a high (above median) share of rural households; columns (3) and (4) test for differential treatment effects across zip codes located in districts with a high (above median) share of informal sector workers; columns (5) and (6) test for differential treatment effects across zip codes located in districts with a high (above median) share of self-employed workers. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table 5: Mechanism: The role of financial infrastructure

	(1)	(2)	(3)	(4)	(5)	(6)
	Zipcode Financial Infrastructure Index					
	$< 25pc$		25-75pc		$> 75pc$	
	#	Vol	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	.112**	.067	.233***	.271***	.225***	.260***
	(.046)	(.050)	(.028)	(.035)	(.041)	(.055)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	.139***	.120**	.223***	.302***	.175***	.253***
	(.053)	(.057)	(.032)	(.039)	(.045)	(.060)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{Rural}$	.068	.138	.229***	.277***	.221***	.266***
	(.077)	(.089)	(.044)	(.053)	(.046)	(.061)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{Rural}$	.139	.252**	.360***	.409***	.200***	.262***
	(.088)	(.102)	(.051)	(.061)	(.053)	(.068)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{Rural} * \mathbb{1}_{High\ Dist\ CC}$	-.163*	-.230**	-.183***	-.219***	-.044	-.035
	(.089)	(.102)	(.055)	(.067)	(.069)	(.089)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{Rural} * \mathbb{1}_{High\ Dist\ CC}$	-.165	-.248**	-.203***	-.238***	.051	.031
	(.102)	(.117)	(.062)	(.076)	(.074)	(.097)
Observations	76450	76450	158561	158561	80759	80759
R^2	.89	.85	.94	.90	.96	.92

This table presents the results estimating the differential treatment effect on POS transactions across rural zip codes, conditional on the zip code's pre-treatment level of financial infrastructure. The unit of observation is zip code. The dependent variable is transactions from POS terminals. The dependent variable in columns (1), (3) and (5) is transaction counts; the dependent variable in columns (2), (4) and (6) is transactions volumes. *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist* equals 1 if the zip code's distance to the nearest currency chest (CC) exceeds the median zip code-nearest currency chest distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state-month-year fixed effects along with district-specific time trends and a fourth order polynomial in the number of POS terminals. Columns (1) and (2) restrict the sample to zip codes in the bottom quartile of zip code financial infrastructure; columns (3) and (4) restrict the sample to zip codes between the 25th and 75th percentile of zip code financial infrastructure; columns (5) and (6) restrict the sample to the top quartile of zip code financial infrastructure. zip code financial infrastructure is based on the pre-treatment financial infrastructure index for each zip code. It is computed as the sum of the standardized pre-treatment indices of POS terminals and bank branches per capita in the zip code. zip codes are considered to be rural if every bank branch in the zip code is either rural or semi-urban. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table 6: Spillover effect on household participation in financial instruments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Household Investment in:						
	Full Sample					High Branch Per Capita	Low Branch Per Capita
	Financial Asset	Gold	Real Estate	Financial Asset	Financial Asset	Financial Asset	
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	-.012**	-.029***	-.029***	.022***	-.040***	.062***	-.018**
	(.005)	(.002)	(.002)	(.005)	(.007)	(.008)	(.008)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	-.029***	-.046***	-.046***	-.016**	-.072***	-.009	-.042***
	(.008)	(.004)	(.004)	(.008)	(.011)	(.013)	(.012)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{Rural} * \mathbb{1}_{High\ Dist\ CC}$				-.045***		-.068***	-.042***
				(.005)		(.007)	(.008)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{Rural} * \mathbb{1}_{High\ Dist\ CC}$				.002		.059***	-.024**
				(.007)		(.010)	(.010)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{Rural}$				.007**		.002	.007
				(.004)		(.005)	(.006)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{Rural}$				-.084***		-.104***	-.055***
				(.005)		(.006)	(.008)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Branch} * \mathbb{1}_{High\ Dist\ CC}$					.041***		
					(.010)		
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Branch} * \mathbb{1}_{High\ Dist\ CC}$					.082***		
					(.016)		
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Branch}$					-.059***		
					(.007)		
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Branch}$					-.069***		
					(.011)		
Observations	1237061	1237061	1237061	1237061	1237061	749739	487322
R^2	.46	.39	.38	.46	.46	.48	.47
Dep Var Mean	.22	.06	.05	.22	.22	.24	.19

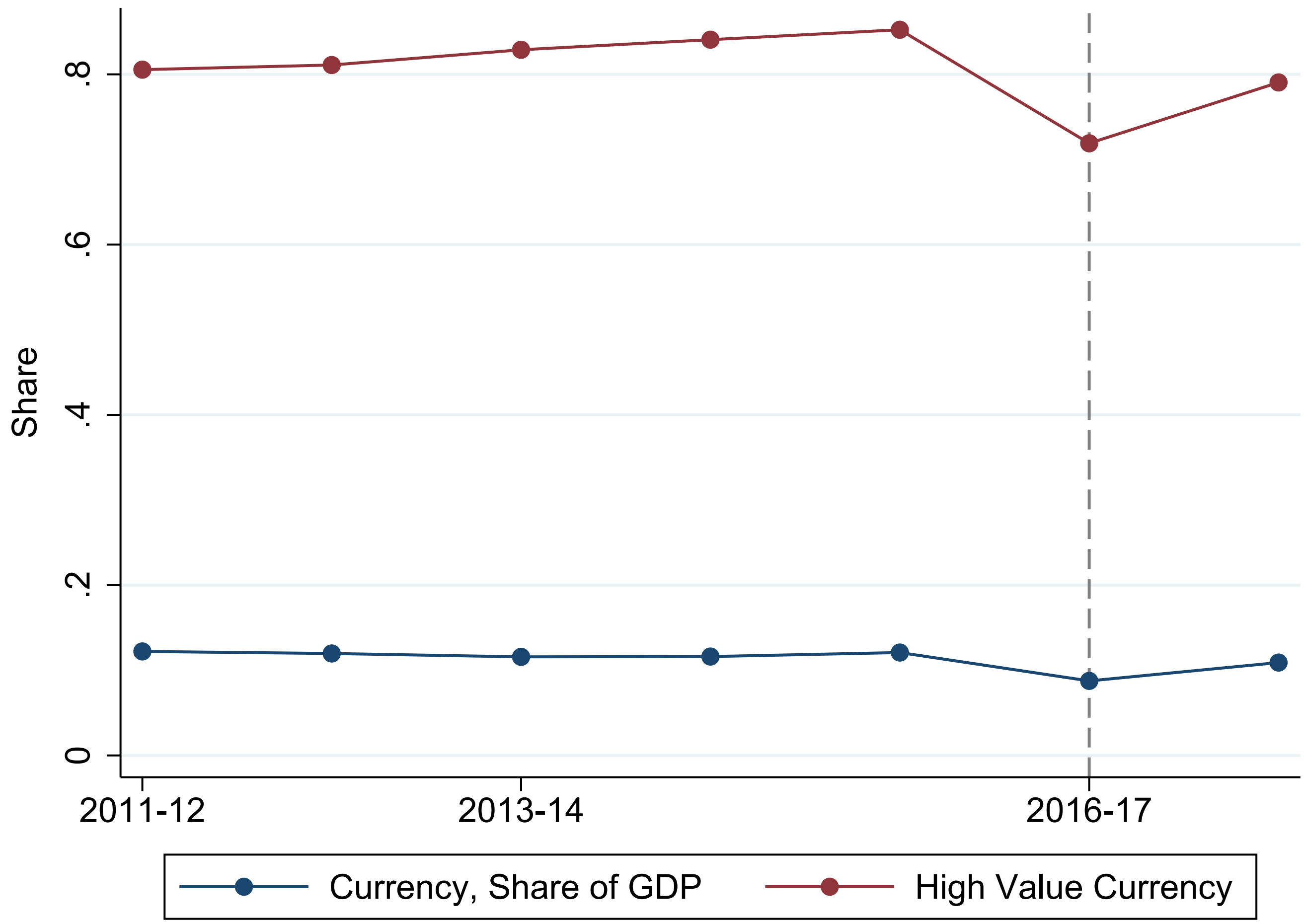
This table presents the results estimating whether the treatment resulted in spillovers on household participation in financial instruments. The unit of observation is household. The dependent variable in column (1) and (4)-(7) is a dummy equaling 1 if the household made any investment in financial instruments in the past 4 months; the dependent variable in column (2) is a dummy equaling 1 if the household invested in gold in the past 4 months; the dependent variable in column (3) is a dummy equaling 1 if the household invested in real estate in the past 4 months. *NearTerm* is a dummy equaling 1 for surveys conducted between November 2016 and January 2018; *LongTerm* is a dummy equaling 1 for surveys conducted after January 2018. *High Dist* equals 1 if the district's median zip code-currency chest (CC) distance exceeds the within-district median zip code-CC distance across all districts. All specifications include household and state-month-year fixed effects along with district-specific time trends and household covariates. Column (4) tests for differential treatment effects across rural households; column (5) tests for differential treatment effects across households residing in disricts with high bank branches per capita; columns (6) and (7) tests for differential treatment effects across rural households after restricting the sample to households residing in districts with high and low bank branches per capita. Standard errors are clustered by household. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Cash is King:
The Role of Financial Infrastructure in Digital Adoption

Online Appendix

Nirupama Kulkarni   S.K.Ritadhi   Bhavya Aggrawal

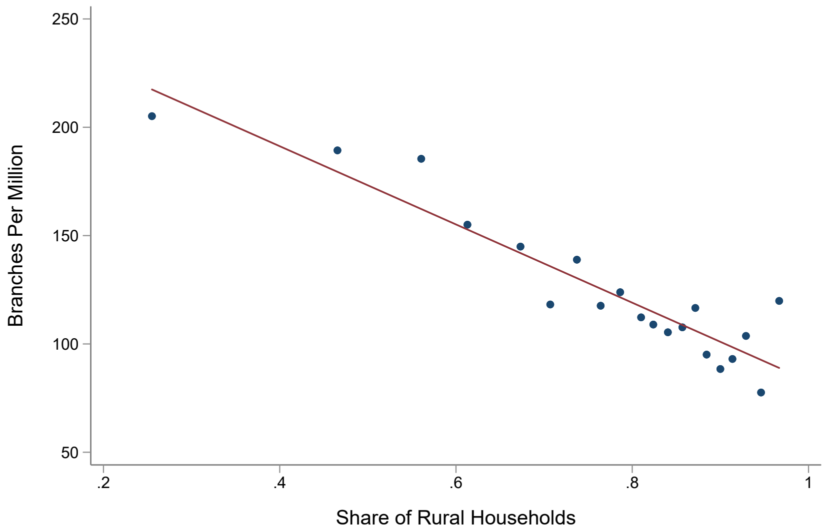
Figure A1: Annual Trends in Currency Circulation



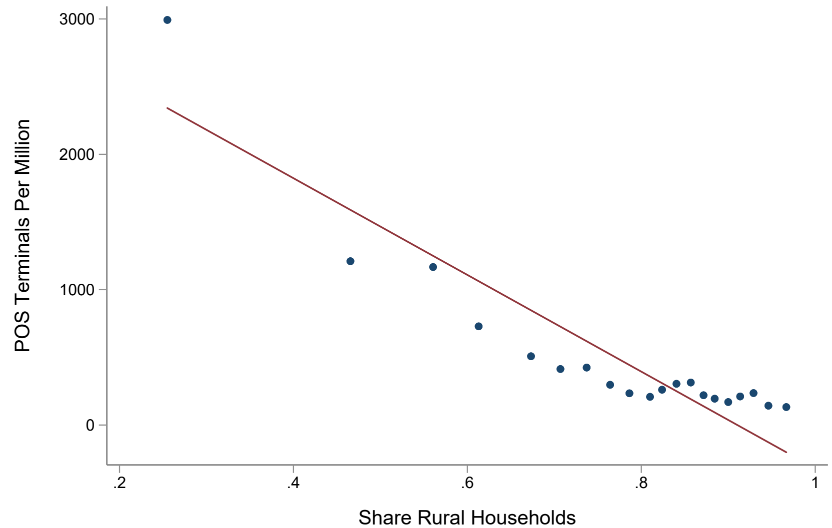
This figure plots the annual trends in currency in circulation as a share of GDP and the share of high value currency as a share of overall currency in circulation. High value currency till November 2016 included denominations of Rs. 500 and Rs. 1,000; since 2016, it includes denominations of Rs. 500 and Rs. 2,000.

Figure A2: District-Level Urbanization and Financial Infrastructure

(A) **Bank Branches Per Capita**



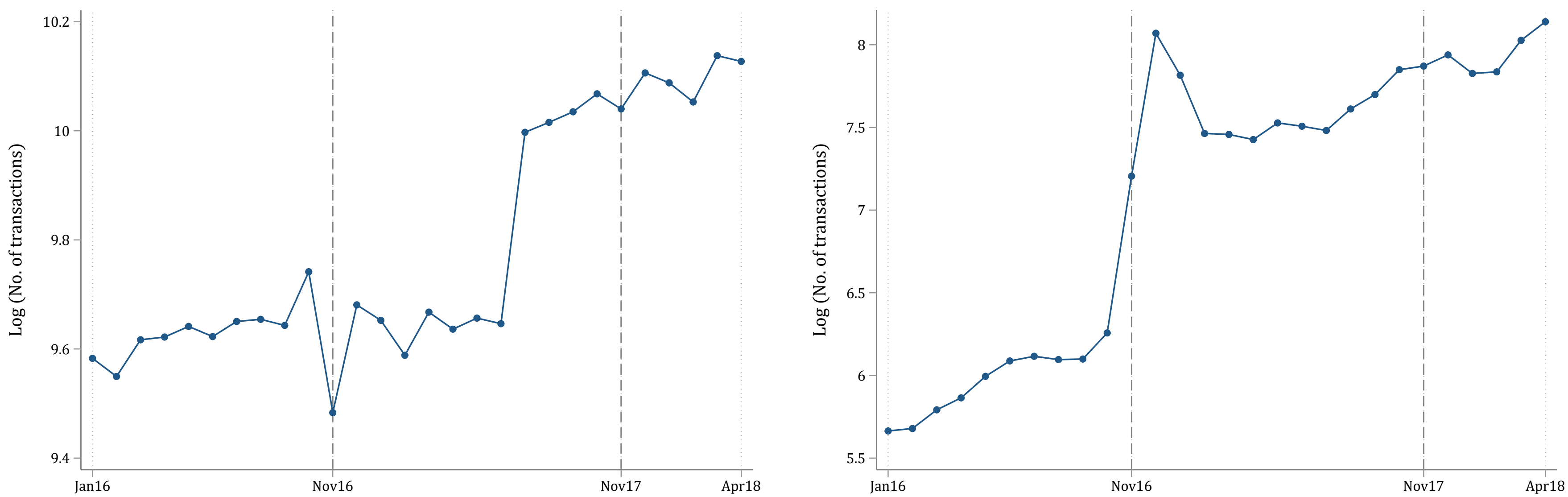
(B) **POS Terminals Per Capita**



The above figure shows the correlation between district financial infrastructure and urbanization using binned scatter plots. The x-axis in both figures is divided into 20 equally bins of the share of rural households in a district. The y-axis measures bank branches per million persons in the left panel, and POS terminals per million persons in the right panel. Each point on the figure represents the unconditional mean of bank branches (POS terminals) within each bin of the share of rural households in the district. The number of bank branches operating in the district is measured in March 2016; the number of POS terminals functional in the district is measured in January 2016.

Figure A3: Monthly Number of Cash Withdrawals and Digital Transactions

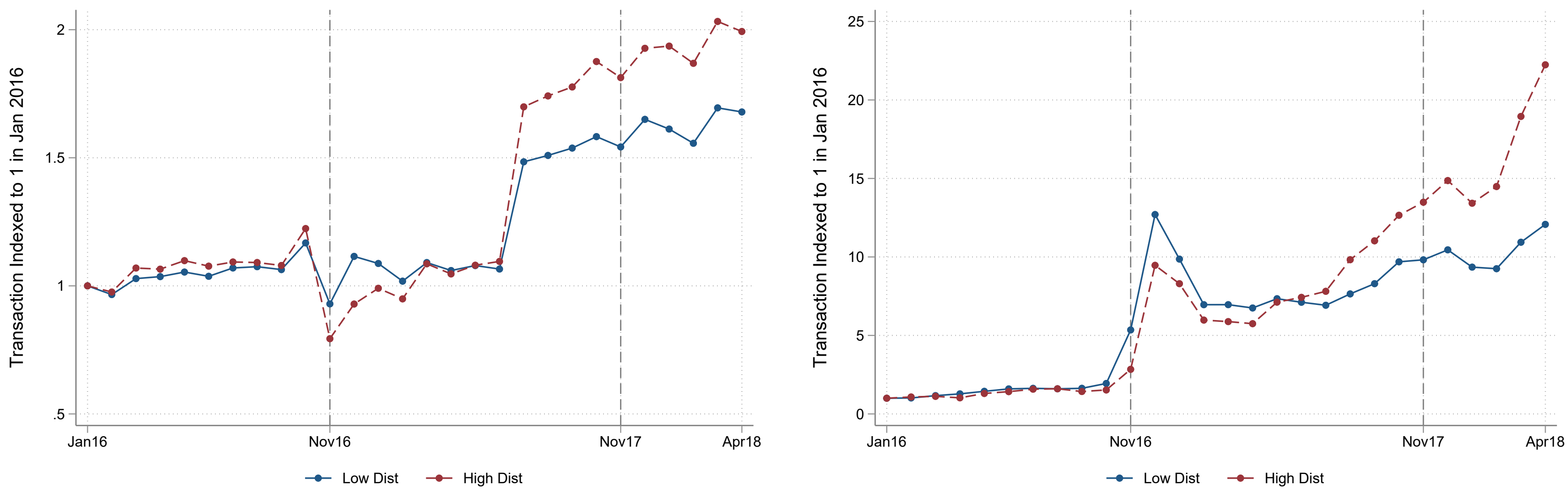
(A) **Number of Cash Withdrawals** (B) **Number of Digital Transactions**



The above figures plot the number of monthly cash and digital transactions between January 2016 and April 2018. Panel A represents the number of cash withdrawals from ATM terminals and Panel B represents the number of digital transactions from PoS terminals.

Figure A4: Monthly Number of Cash Withdrawals and Digital Transactions, by Distance from Currency Chests

(A) **Number of Cash Withdrawals** (B) **Number of Digital Transactions**

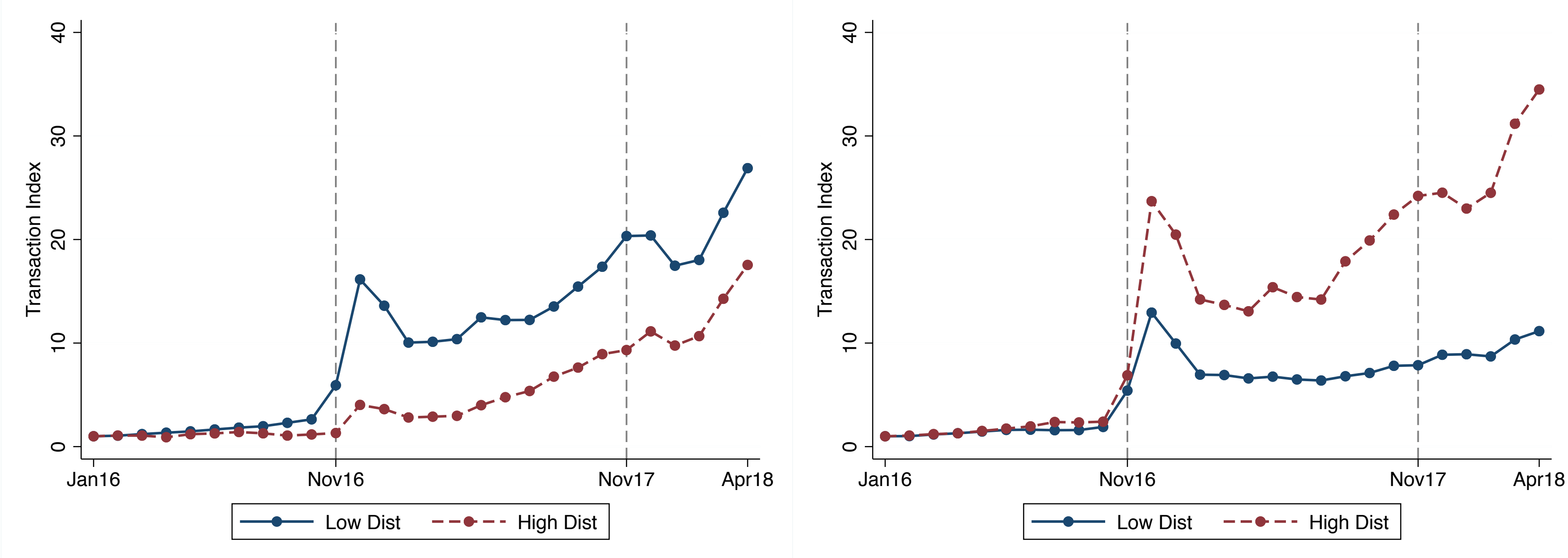


The above figures plot the monthly cash and digital transactions between January 2016 and April 2018, by the pincodes' distance from currency chests (CC). Panel A shows the number of cash withdrawals from ATM terminals. Panel B shows the number of digital transactions from PoS terminals. Each point on the graph represents the total number of transactions undertaken in each month, relative to January 2016 *High Dist* refers to pincodes where the pincode's distance to the nearest CC exceeds the median pincode-nearest CC distance in the sample.

Figure A5: Heterogeneity in effect on digital transactions across rural and urban districts: Numbers

(A) **Number of Digital Transactions: High Rural**

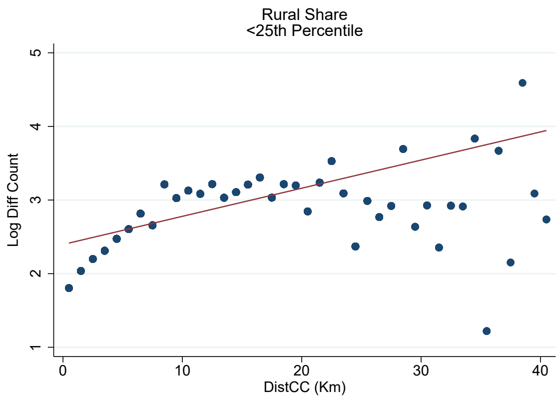
(B) **Number of Digital Transactions: Low Rural**



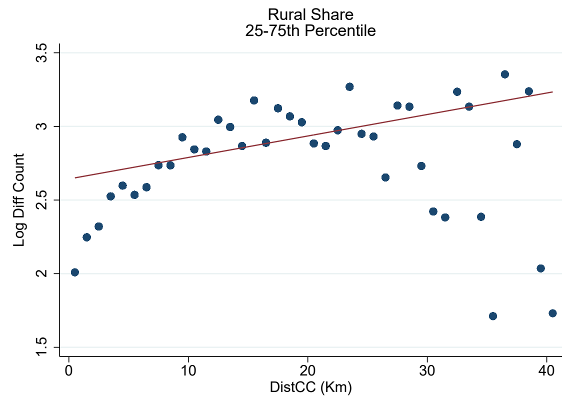
The above figures plot the number of monthly digital transactions between January 2016 and April 2018, by zip codes' distance from currency chests (CC). Panel A splits the sample by zip codes located in districts with a relatively high share of rural households and Panel B splits the sample by zip codes located in districts with low share of rural households. Each point on the graph represents the total number of transactions undertaken in each month, relative to January 2016. Districts are categorized into high and low rural based on the median share of rural households across all districts in 2011. *High Dist* refers to zip codes where the zip code's distance to the nearest CC exceeds the median zip code-nearest CC distance in the sample.

Figure A6: Heterogeneity in effect on digital transactions by share of rural households: Numbers

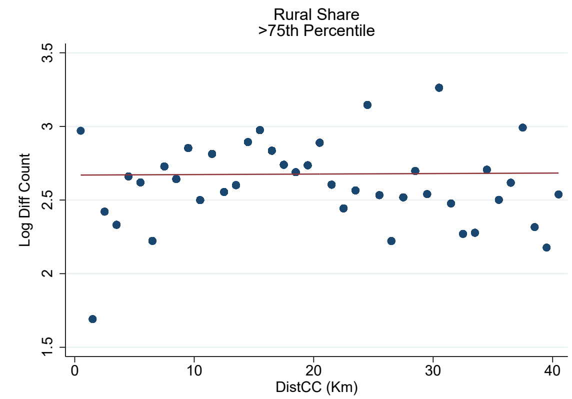
(A) **Rural Share: 25th Pctile**



(B) **Rural Share: 25-75th Pctile**



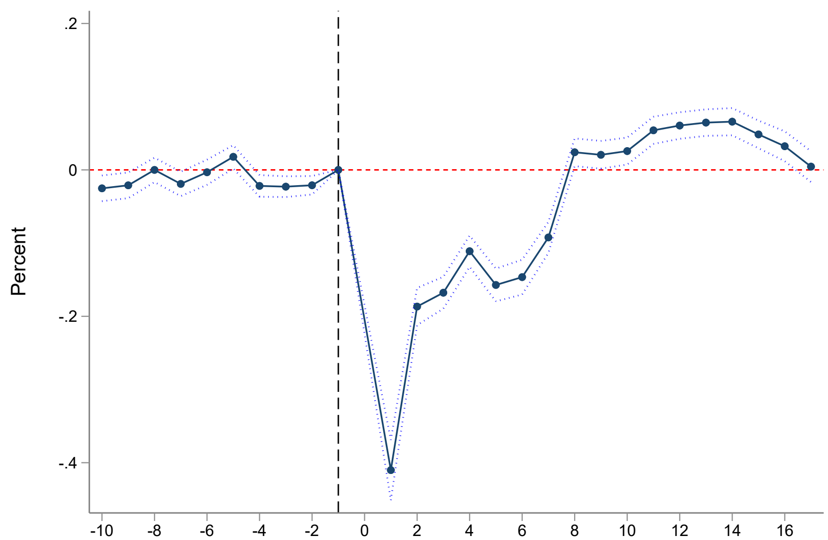
(C) **Rural Share: 75th Pctile**



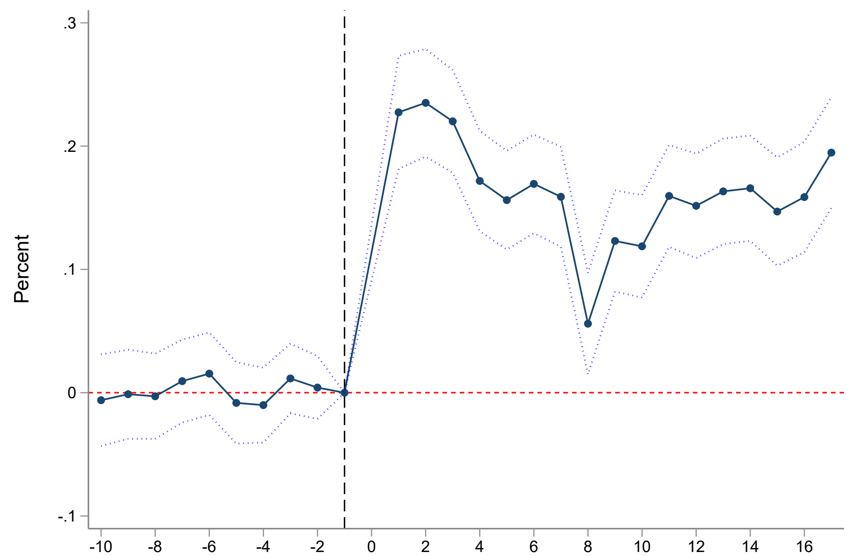
The above figures show binned scatter plots of the difference in the mean transaction levels from POS terminals across the post and pre-treatment periods. The differences are measured in logs. The x-axis represents the distance between zip codes and the nearest currency chest which serves as a proxy for access to cash. A higher distance implies lower access to cash. Panels A–C depicts changes in transaction counts. Panel A restricts the sample to zip codes located in districts where the share of rural households is below the 25th percentile; panel B restricts the sample to zip codes located in districts where the share of rural households is between the 25th and 75th percentiles; Panel C restricts the sample to zip codes located in districts where the share of rural households exceeds the 75th percentile.

Figure A7: Monthly Average Treatment Effects on Transactions from ATM and POS Terminals

(A) **Number of Cash Withdrawals**



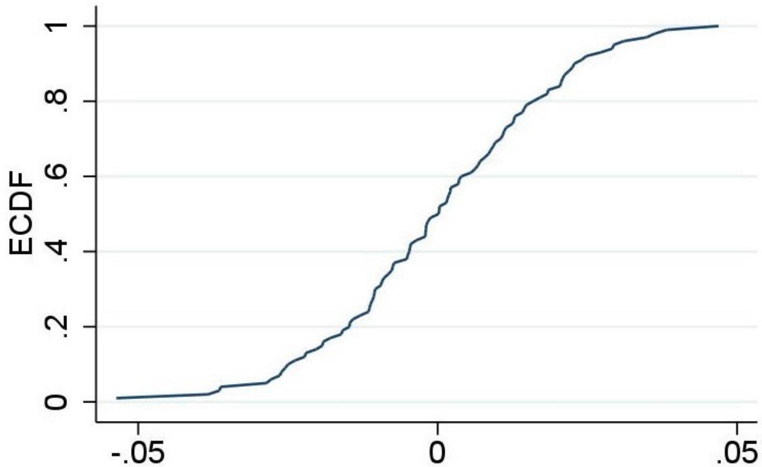
(B) **Number of Digital Transactions**



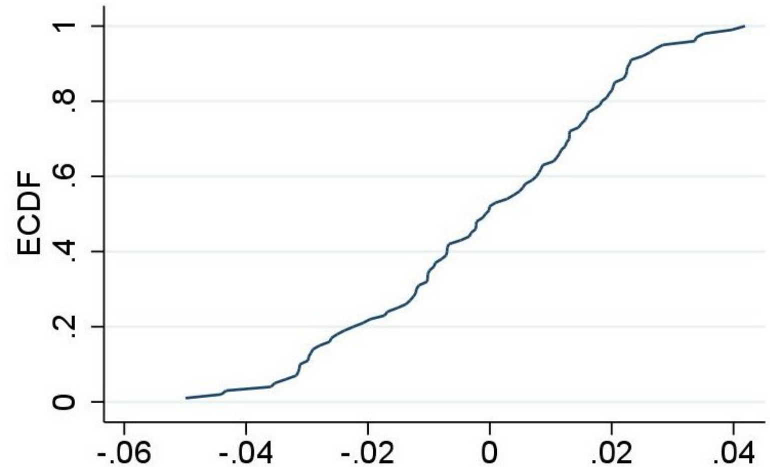
The above figures presents the monthly average treatment effects of the cash supply shock on cash withdrawals and digital transactions. The unit of observation is the zip code. Panel A shows the number of cash withdrawals from ATM terminals. Panel B shows the number of digital transactions from PoS terminals. The outcome variable is logged in both panels. The light blue dotted line shows the 95 percent confidence interval and the x-axis is the months before and after the treatment intervention. The reference period is October 2016 - the month prior to the shock. All specifications include zip code and state-month-year fixed effects, along with district-specific time trends and a fourth order polynomial in the number of POS terminals in the zip code. Standard errors are clustered by zip code.

Figure A8: Empirical CDFs from Random Assignment of Zip Codes to High and Low Distance from Currency Chests

(A) **Number of Digital Transactions: Near Term**



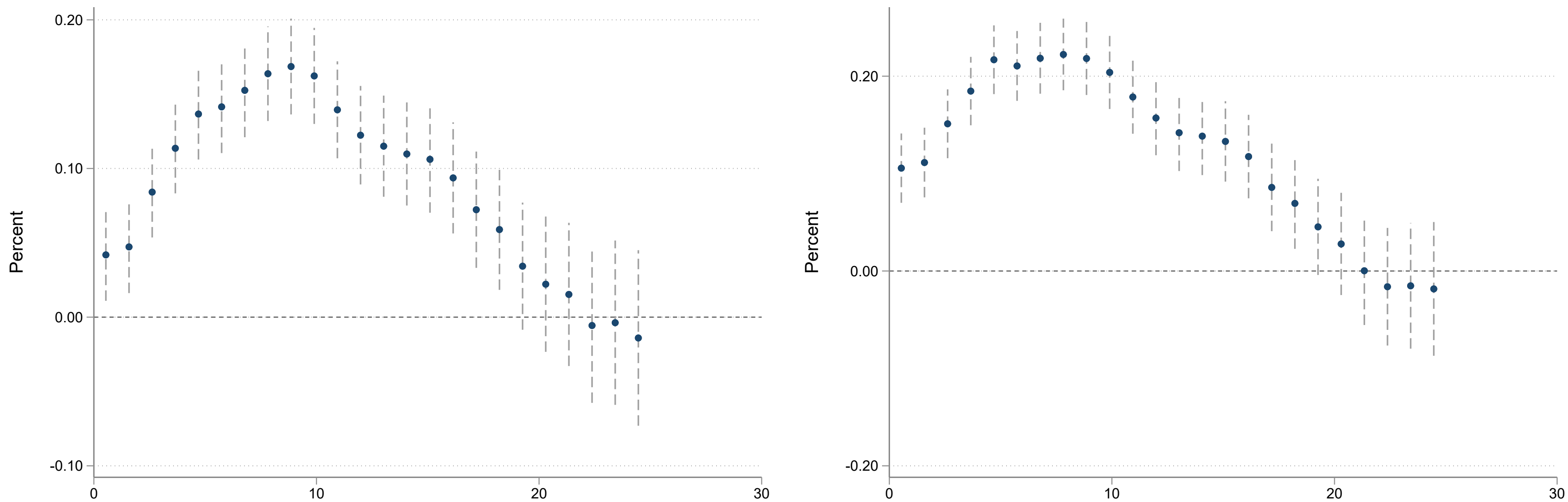
(B) **Number of Digital Transactions: Long Term**



The above figures present the empirical CDFs from 100 iterations of the baseline results whereby zip codes are randomly assigned to "high" and "low" distance from currency chests. The x-axis shows the coefficient values. The true coefficient over the near-term for transaction counts is 0.154 and over the long-term it is 0.168. The unit of observation in each specification is the zip code. The outcome of interest is logged. Panel A presents the near-term coefficients and Panel B presents the long-term coefficients. The outcome of interest is POS transaction counts. All specifications include zip code and state-month-year fixed effects, along with district-specific time trends and a fourth order polynomial in the number of POS terminals in the zip code.

Figure A9: Robustness of Baseline Results to Alternate Distance Thresholds: Number of Digital Transactions

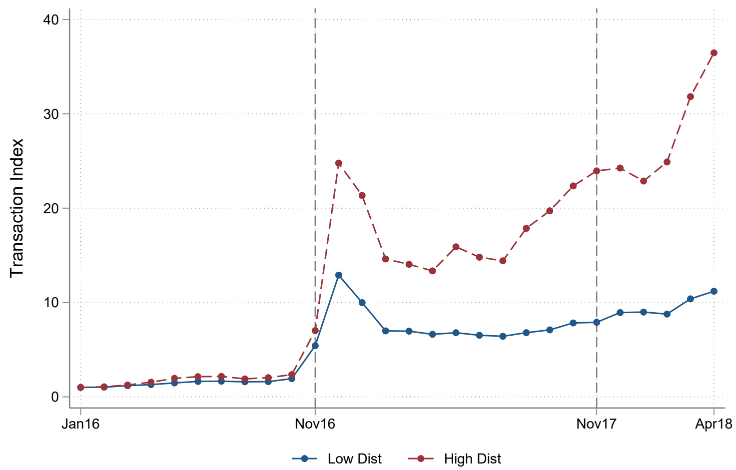
(A) **Near Term** (B) **Long Term**



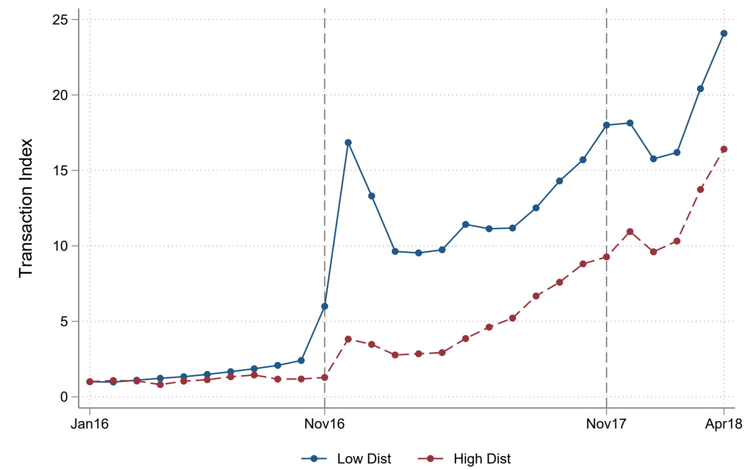
The above figures represent the robustness of the baseline results to alternate distance thresholds determining the assignment of zip codes to high and low distance from currency chests. The unit of observation is the zip code. The outcome of interest is the number of digital transactions from POS terminals. Panel A presents the near-term coefficients and Panel B presents the long-term coefficients. The outcome variable is logged. Vertical lines dotted represent 95 percent confidence intervals of the coefficients. All specifications include zip code and state-month-year fixed effects, along with district-specific time trends and a fourth order polynomial in the number of POS terminals in the zip code. Each coefficient corresponds to a 1 kilometre increase in the distance threshold used to determine which zip codes are located farther from currency chests and hence, more exposed to the treatment. The distance threshold in the baseline specifications is 11 kilometres.

Figure A10: Heterogeneity in Digital Transactions Across Districts' Share of Informal Workers

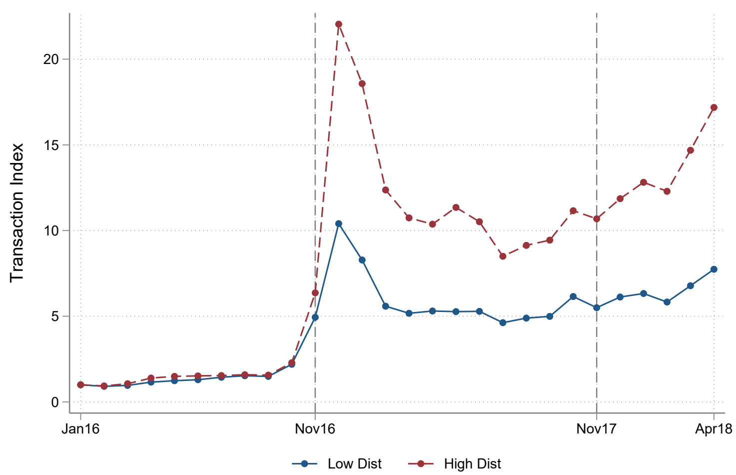
(A) **Number of Digital Transactions: Low Informal Sector**



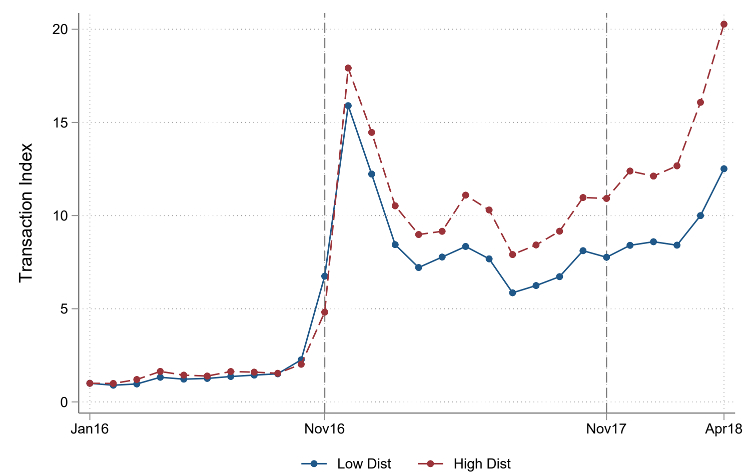
(B) **Number of Digital Transactions: High Informal Sector**



(C) **Volume of Digital Transactions: Low Informal Sector**

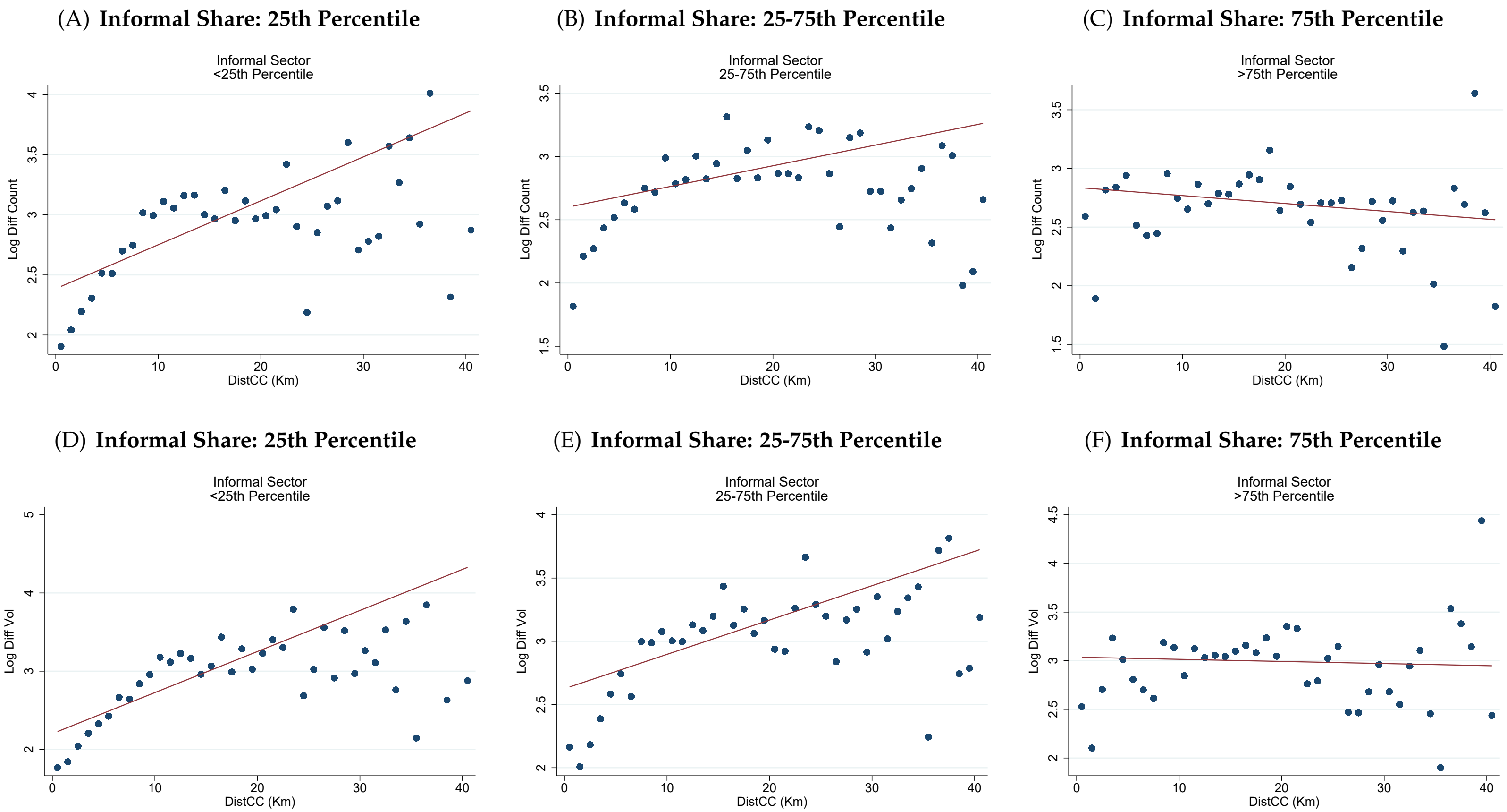


(D) **Volume of Digital Transactions: High Informal Sector**



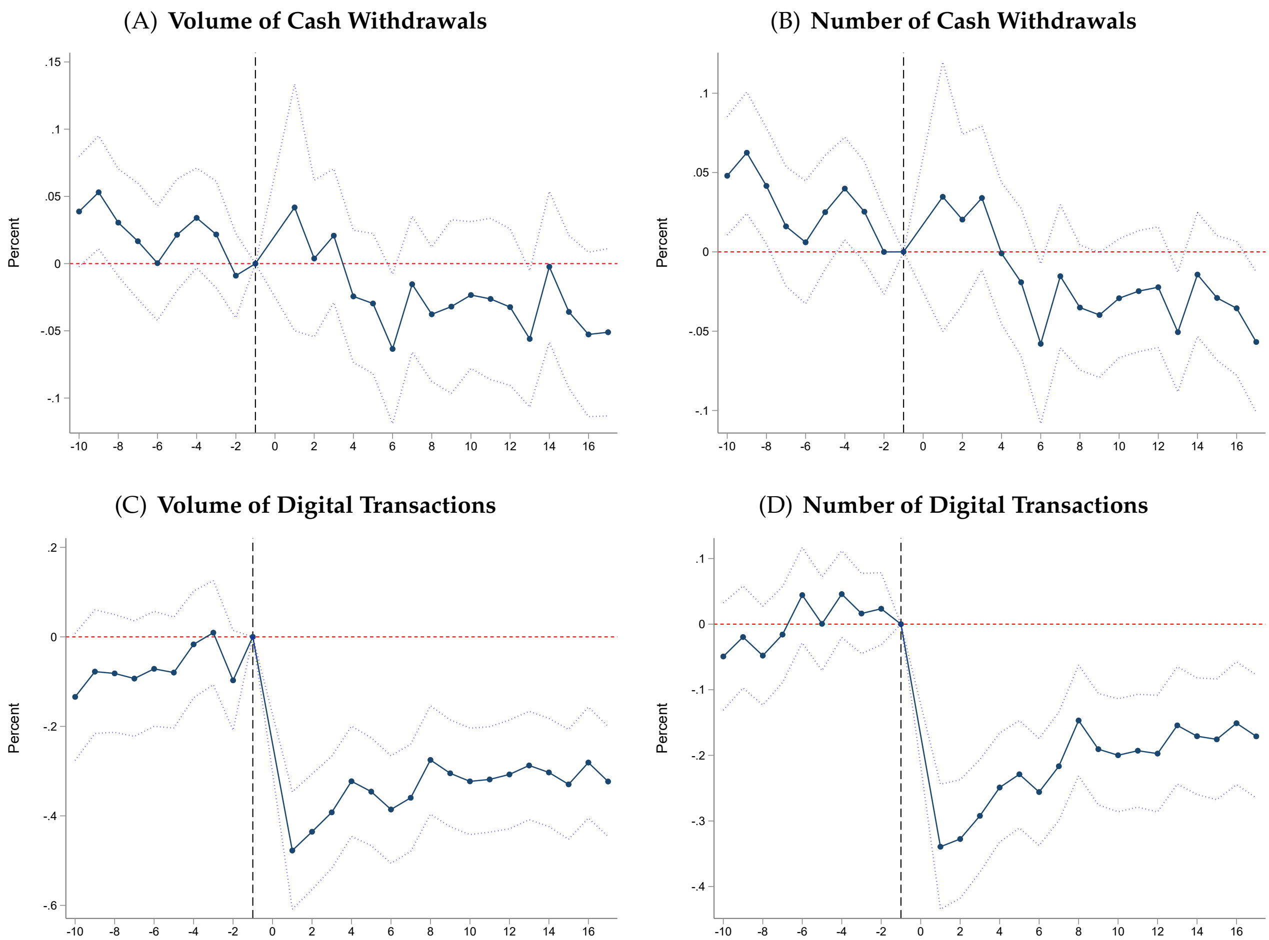
The above figures plot the monthly digital transactions from POS terminals between January 2016 and April 2018, by zip codes' distance from currency chests (CC). Panels A and C split the sample by zip codes located in districts with a relatively high share of informal sector workers; panels B and D split the sample by zip codes located in districts with low share of informal sector workers. Panels A and B present transaction counts; panels C and D present transaction volumes. Each point on the graph represents the total transactions undertaken in each month, relative to January 2016. Districts are categorized into high and low informal sector workers based on the median share of informal sector workers across all districts in 2011. *High Dist* refers to zip codes where the zip code's distance to the nearest CC exceeds the median zip code-nearest CC distance in the sample.

Figure A11: Within Zip Code Changes in Digital Transactions as a Function of Distance from Currency Chests: Heterogeneity Across Share of Informal Sector Workers



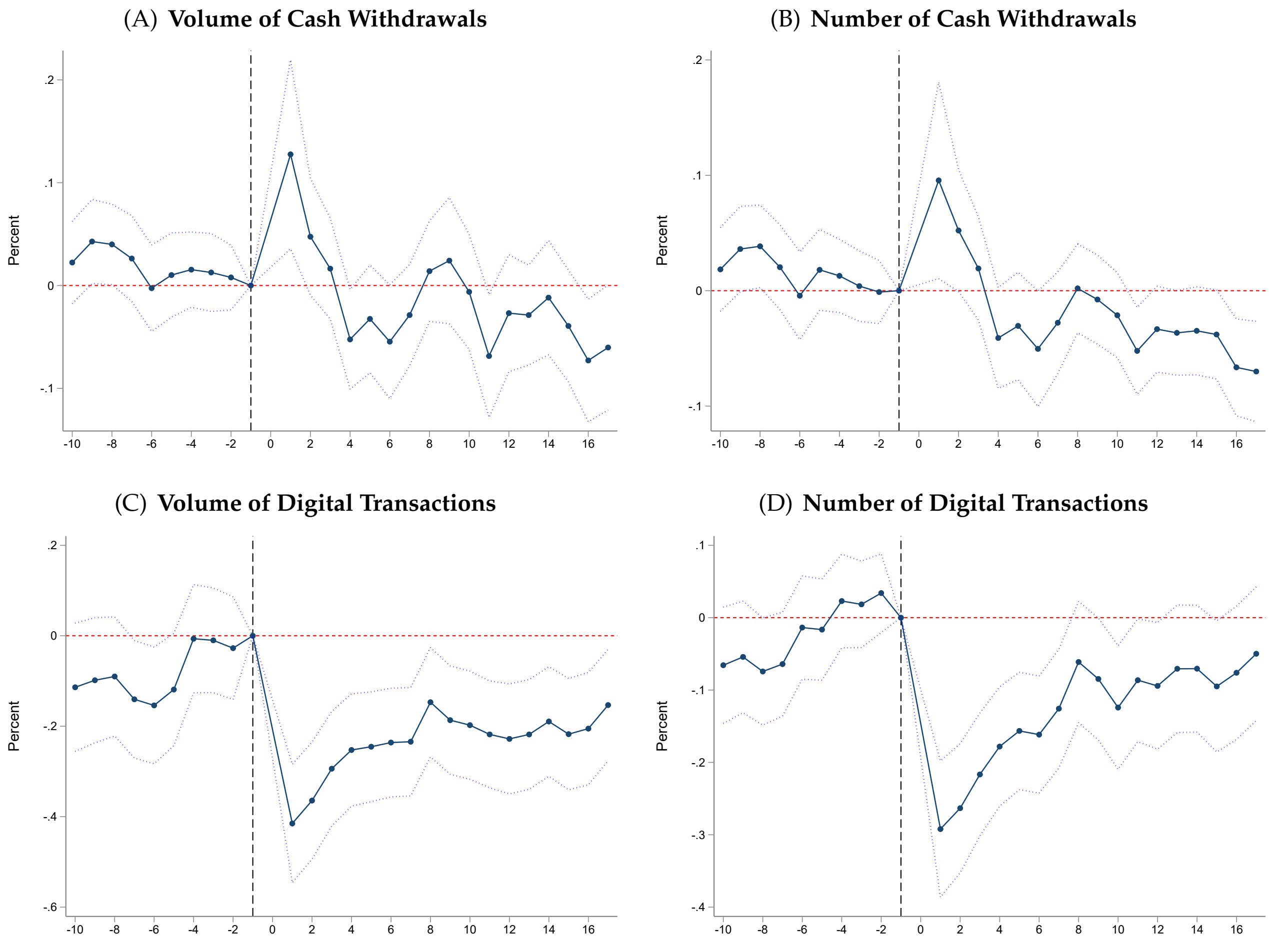
The above figures show binned scatter plots of the difference in the mean transaction levels from POS terminals across the post and pre-treatment periods. The differences are measured in logs. The x-axis represents the distance between zip codes and the nearest currency chest which serves as a proxy for access to cash. A higher distance implies lower access to cash. Panels A–C depicts changes in transaction counts; panels D–F depicts changes in transaction volumes. Panels A and D restrict the sample to zip codes located in districts where the share of informal sector workers is below the 25th percentile; panels B and E restrict the sample to zip codes located in districts where the share of informal sector workers is between the 25th and 75th percentiles; panels C and F restrict the sample to zip codes located in districts where the share of informal sector workers exceeds the 75th percentile.

Figure A12: Differential Treatment Effects Across Zip Codes Located in Districts with a High Share of Rural Households: Monthly Triple Interaction Coefficients



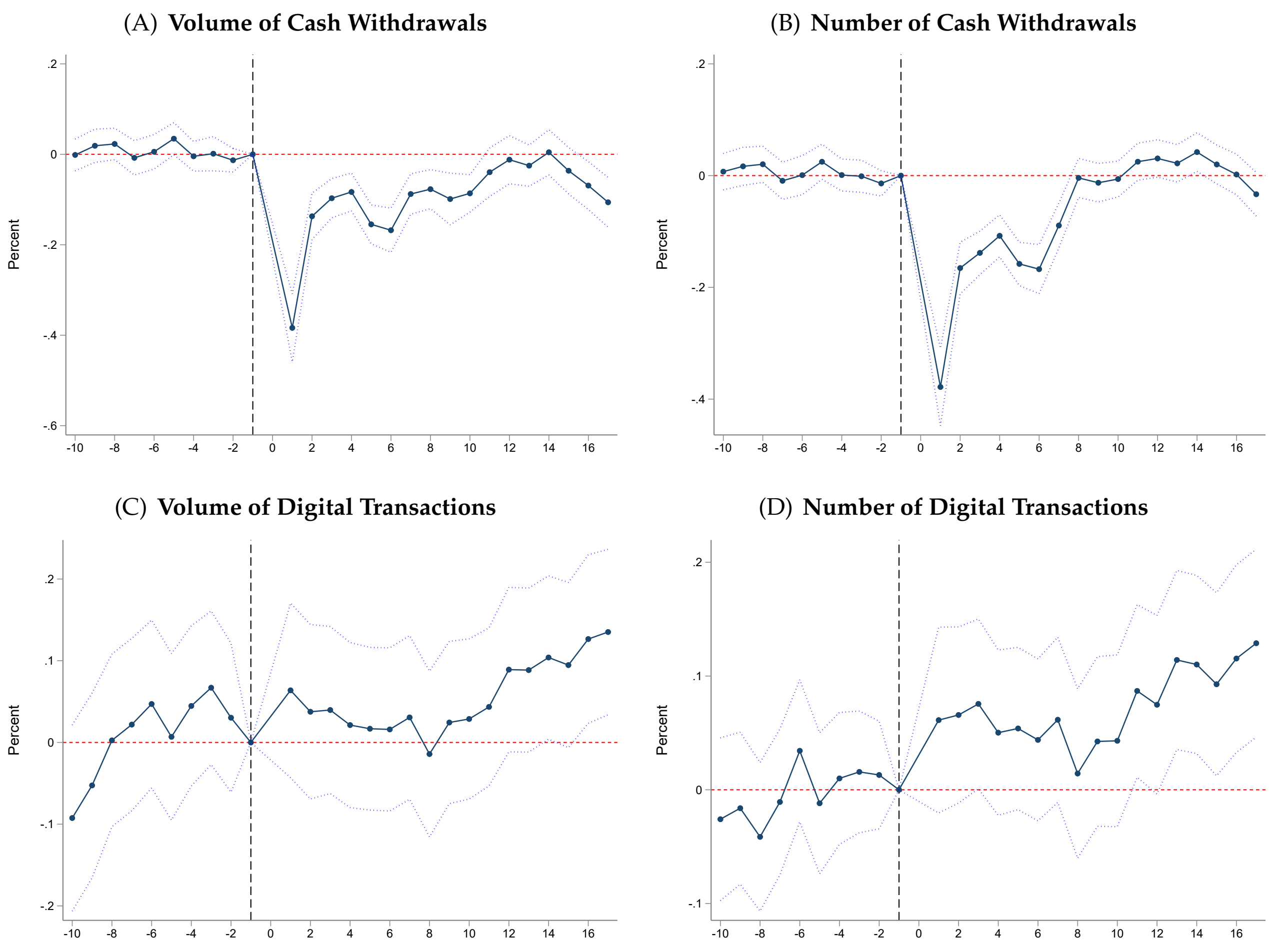
The above figures presents the monthly differential treatment effects across zip codes with a high share of rural households on transactions from ATM and POS terminals. The unit of observation is the zip code. Panels A and B show cash withdrawals from ATM terminals. Panels C and D show digital transactions from POS terminals. Panels A and C show transaction volumes and Panels B and D show transaction counts. The outcome variable in each instance is logged. The light blue dotted lines reflect the 95 percent confidence intervals and the x-axis is the months before and after the treatment intervention. The reference period is October 2016 - the month prior to the shock. All specifications include zip code and state-month-year fixed effects, along with district-specific time trends and a fourth order polynomial in the number of POS terminals in the zip code. Standard errors are clustered by zip code.

Figure A13: Differential Treatment Effects Across Zip Codes Located in Districts with a High Share of Informal Sector Workers: Monthly Triple Interaction Coefficients



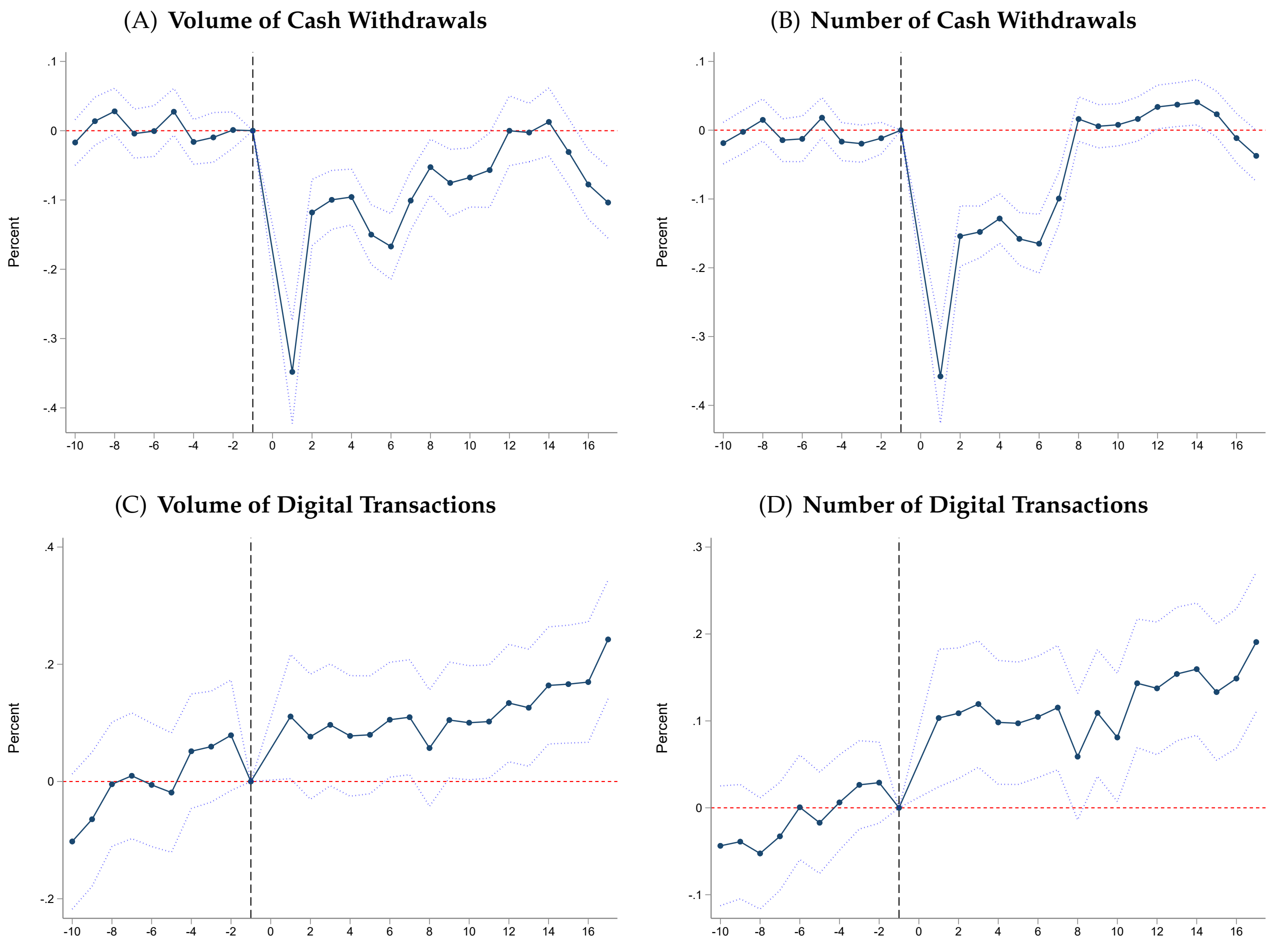
The above figures presents the monthly differential treatment effects across zip codes with a high share of informal sector workers on transactions from ATM and POS terminals. The unit of observation is the zip code. Panels A and B show cash withdrawals from ATM terminals. Panels C and D show digital transactions from POS terminals. Panels A and C show transaction volumes and Panels B and D show transaction counts. The light blue dotted lines reflect the 95 percent confidence intervals and the x-axis is the months before and after the treatment intervention. The reference period is October 2016 - the month prior to the shock. All specifications include zip code and state-month-year fixed effects, along with district-specific time trends and a fourth order polynomial in the number of POS terminals in the zip code. Standard errors are clustered by zip code.

Figure A14: Monthly Average Treatment Effects for Zip Codes Located in Districts with a High Share of Rural Households



The above figures presents the monthly average treatment effects of the cash supply shock on ATM and POS transactions. Panels A and B show cash withdrawals from ATM terminals. Panels C and D show digital transactions from POS terminals. Panels A and C show transaction volumes and Panels B and D show transaction counts. The outcome variable in each instance is logged. The light blue dotted lines reflect the 95 percent confidence intervals and the x-axis is the months before and after the treatment intervention. The reference period is October 2016 - the month prior to the shock. All specifications include zip code and month-year fixed effects, along with district-specific time trends. The sample is restricted to zip codes located in districts with a high share (above median) of rural households. Standard errors are clustered by zip code.

Figure A15: Monthly Average Treatment Effects for Zip Codes Located in Districts with a High Share of Informal Sector workers



The above figures presents the monthly average treatment effects of the cash supply shock on ATM and POS transactions. Panels A and B show cash withdrawals from ATM terminals. Panels C and D show digital transactions from POS terminals. Panels A and C show transaction volumes and Panels B and D show transaction counts. The outcome variable in each instance is logged. The light blue dotted lines reflect the 95 percent confidence intervals and the x-axis is the months before and after the treatment intervention. The reference period is October 2016 - the month prior to the shock. All specifications include zip code and month-year fixed effects, along with district-specific time trends. The sample is restricted to zip codes located in districts with a high share of informal sector workers. Standard errors are clustered by zip code.

Table A1: Descriptive Characteristics of Districts by Distance From Currency Chests

	Panel A: All Zip Codes		
	Low Distance	**High Distance**	**Significance**
Distance from Currency Chest (Km)	4.22	18.94	$***$
Imputed Population	92917	45431	$***$
Distance from Urban Centre (Km)	30.75	49.25	$***$
Fraction Rural	0.65	0.96	$***$
Branches Per Capita (Per Million)	270.12	292.89	
ATM Terminals Per Capita (Per Million)	4.56	0.26	$***$
POS Terminals Per Capita (Per Million)	1.02	0.42	$***$
No. of ATM Transactions	26563.47	5253.89	$***$
No. of POS Transactions	444.89	90.18	$***$
ATM Transaction Volume (Rs. '000)	91720.27	19527	$***$
POS Transaction Volume (Rs. '000)	814.42	54.24	$***$
Average ATM Transaction Value (Rs.)	3655.36	3860.48	$***$
Average POS Transaction Value (Rs.)	2015.24	1842.72	$***$
	Panel B: Exclude Outliers		
	Low Distance	**High Distance**	**Significance**
Distance from Currency Chest (Km)	7.87	13.98	$***$
Imputed Population	34368.76	39579.53	$***$
Distance from Urban Centre (Km)	31.29	43.62	$***$
Fraction Rural	0.80	0.95	$***$
Branches Per Capita (Per Million)	222.99	191.41	
ATM Terminals Per Capita (Per Million)	0.67	0.35	$***$
POS Terminals Per Capita (Per Million)	0.47	0.34	$***$
No. of ATM Transactions	10958.83	5373.93	$***$
No. of POS Transactions	178.37	130.55	$***$
ATM Transaction Volume (Rs. '000)	37165.48	19783.83	$***$
POS Transaction Volume (Rs. '000)	259.28	62.42	$***$
Average ATM Transaction Value (Rs.)	3716.24	3840.84	$***$
Average POS Transaction Value (Rs.)	1995.33	1847.22	$***$

This table presents descriptive characteristics of districts based on the distance of the median zip code in the district from nearest currency chest. *Low DistCC* districts are those for which the median zip code's distance to the nearest currency chest is within 11 kilometres. *High DistCC* districts are those for which the median zip code's distance to the nearest currency chest exceeds 11 kilometres.

Table A2: Average Treatment Effects on Cash Withdrawals/Digital Transactions - Flexible District Covariates

	(1)	(2)	(3)	(4)
	Cash Withdrawals		Digital Transactions	
	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	-.098***	-.121***	.161***	.178***
	(.008)	(.009)	(.017)	(.020)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	-.060***	-.064***	.119***	.150***
	(.014)	(.017)	(.030)	(.032)
Observations	345964	345964	345964	345964
R^2	.94	.90	.92	.88

This table presents the results estimating the average treatment effect on ATM and POS transactions. The unit of observation is zip code. The dependent variable in columns (1) and (2) is transactions from ATM terminals; columns (3) and (4), transactions from POS terminals. The dependent variable is measured in transaction counts in columns (1) and (3) and transaction volumes in columns (2) and (4). *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist* equals 1 if the zip code's distance to the nearest CC exceeds the median zip code-nearest currency chest distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state-month-year fixed effects along with a fourth order polynomial in the number of POS terminals, a fourth order polynomial in distance from CC (interacted with a post-treatment indicator) and the following district-level covariates: share of rural households; average per capita consumption; share of formal sector workers; share of workers with secondary or higher education; household size; average age of workers; share of SC/STs in the district. All the district covariates are based on the 2011-12 NSS Employment Unemployment Household Survey and interacted with a post-treatment indicator. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table A3: Average Treatment Effects on Cash Withdrawals/Digital Transactions - Robustness to District-Month-Year Fixed Effects

	(1)	(2)	(3)	(4)
	Cash Withdrawals		Digital Transactions	
	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	-.076***	-.093***	.124***	.142***
	(.008)	(.009)	(.017)	(.020)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	.056***	-.004	.158***	.214***
	(.008)	(.010)	(.020)	(.023)
Observations	354425	354425	354425	354425
R^2	.94	.91	.94	.90

This table presents the results estimating the average treatment effect on ATM and POS transactions. The unit of observation is zip code. The dependent variable in columns (1) and (2) is transactions from ATM terminals; columns (3) and (4), transactions from POS terminals. The dependent variable is measured in transaction counts in columns (1) and (3) and transaction volumes in columns (2) and (4). *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist* equals 1 if the zip code's distance to the nearest CC exceeds the median zip code-nearest currency chest distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and district-month-year fixed effects along with a fourth order polynomial in the number of POS terminals. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table A4: Differential Treatment Effects on Cash Withdrawals Across zip codes Located in Districts with High Informality

	Measure of Informality (MoI)					
	Rural		Informal		Self-Employed	
	#	Vol	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	-.083***	-.096***	-.093***	-.106***	-.105***	-.118***
	(.010)	(.011)	(.010)	(.011)	(.010)	(.012)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	.066***	.009	.064***	.008	.035***	-.022*
	(.010)	(.012)	(.010)	(.012)	(.011)	(.013)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Rural}$	-.119***	-.063***				
	(.016)	(.017)				
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Rural}$	-.138***	-.058**				
	(.020)	(.024)				
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Rural} * \mathbb{1}_{High\ Dist\ CC}$	-.035**	-.035*				
	(.016)	(.018)				
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Rural} * \mathbb{1}_{High\ Dist\ CC}$	-.056***	-.055***				
	(.017)	(.020)				
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Informal}$			-.106***	-.059***		
			(.015)	(.016)		
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Informal}$			-.105***	-.049**		
			(.019)	(.023)		
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Informal} * \mathbb{1}_{High\ Dist\ CC}$			-.013	-.013		
			(.016)	(.018)		
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Informal} * \mathbb{1}_{High\ Dist\ CC}$			-.053***	-.051**		
			(.016)	(.020)		
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ SelfEmp}$					-.121***	-.107***
					(.016)	(.017)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ SelfEmp}$					-.131***	-.116***
					(.021)	(.023)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ SelfEmp} * \mathbb{1}_{High\ Dist\ CC}$					.007	.013
					(.015)	(.017)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ SelfEmp} * \mathbb{1}_{High\ Dist\ CC}$					.013	.019
					(.016)	(.019)
Observations	352454	352454	352454	352454	352454	352454
R^2	.94	.91	.94	.91	.94	.91

This table presents the results estimating the differential treatment effect on cash withdrawals from ATM terminals across zip codes located in districts with high informality. The unit of observation is zip code. The dependent variable is transactions from ATM terminals. The dependent variable in columns (1), (3) and (5) is transaction counts; the dependent variable in columns (2), (4) and (6) is transactions volumes. *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist CC* equals 1 if the zip code's distance to the nearest currency chest (CC) exceeds the median zip code-nearest CC distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state-month-year fixed effects along with district-specific time trends and a fourth order polynomial in the number of POS terminals. Columns (1) and (2) test for differential treatment effects across zip codes located in districts with a high (above median) share of rural households; columns (3) and (4) test for differential treatment effects across zip codes located in districts with a high (above median) share of informal sector workers; columns (5) and (6) test for differential treatment effects across zip codes located in districts with a high (above median) share of self-employed workers. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table A5: Differential Treatment Effects Across Zip Codes Located in Districts with High Share of Salaried Workers, Educated Workers and High Household Consumption

	(1)	(2)	(3)	(4)	(5)	(6)
	Measure of Formality (MoF)					
	Salaried Worker		Educated Workers		Household Expenditures	
	#	Vol	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	.037	.030	.064**	.095***	.044	.026
	(.028)	(.032)	(.027)	(.032)	(.029)	(.033)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	.049	.077**	.101***	.153***	.070**	.082**
	(.032)	(.036)	(.032)	(.036)	(.034)	(.037)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Salaried\ Worker}$	-.138***	-.240***				
	(.026)	(.031)				
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Salaried\ Worker}$	-.192***	-.286***				
	(.033)	(.042)				
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Salaried\ Worker} * \mathbb{1}_{High\ Dist\ CC}$	.179***	.224***				
	(.033)	(.039)				
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Salaried\ Worker} * \mathbb{1}_{High\ Dist\ CC}$	.187***	.233***				
	(.038)	(.043)				
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Educated\ Worker}$			-.152***	-.233***		
			(.027)	(.032)		
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Educated\ Worker}$			-.144***	-.241***		
			(.036)	(.044)		
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Educated\ Worker} * \mathbb{1}_{High\ Dist\ CC}$			.142***	.125***		
			(.033)	(.038)		
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Educated\ Worker} * \mathbb{1}_{High\ Dist\ CC}$			.115***	.121***		
			(.038)	(.043)		
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Expenditure}$					-.117***	-.204***
					(.030)	(.036)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Expenditure}$					-.135***	-.226***
					(.038)	(.048)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Expenditure} * \mathbb{1}_{High\ Dist\ CC}$					.165***	.229***
					(.033)	(.039)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Expenditure} * \mathbb{1}_{High\ Dist\ CC}$					.152***	.223***
					(.039)	(.044)
Observations	352454	352454	352454	352454	352454	352454
R^2	.93	.89	.93	.89	.93	.89

This table presents the results estimating the differential treatment effect on POS transactions across zip codes located in districts with high share of salaried workers, educated workers and high per capita consumption. The unit of observation is zip code. The dependent variable is transactions from POS terminals. The dependent variable in columns (1), (3) and (5) is transaction counts; the dependent variable in columns (2), (4) and (6) is transactions volumes. *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist* equals 1 if the zip code's distance to the nearest currency chest (CC) exceeds the median zip code-nearest CC distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state-month-year fixed effects along with district-specific time trends and a fourth order polynomial in the number of POS terminals. Columns (1) and (2) test for differential treatment effects across zip codes located in districts with a high share of salaried workers; columns (3) and (4) test for differential treatment effects across zip codes located in districts with a high share of workers with secondary or higher education; columns (5) and (6) test for differential treatment effects across zip codes located in districts with high per capita household expenditures. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table A6: Differential Treatment Effects Across Zip Codes Located in Districts with High Informality, Conditional on Pre-Treatment Levels of Corruption

	(1)	(2)	(3)	(4)
	High Corrupt		Low Corrupt	
	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	.222***	.268***	.258***	.298***
	(.030)	(.039)	(.026)	(.031)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	.213***	.301***	.250***	.332***
	(.032)	(.044)	(.031)	(.035)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Rural}$	.103**	.162***	.073**	.170***
	(.044)	(.054)	(.032)	(.037)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Rural}$	.080	.145**	.056	.158***
	(.056)	(.072)	(.042)	(.051)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Rural} * \mathbb{1}_{High\ Dist\ CC}$	-.220***	-.275***	-.251***	-.307***
	(.050)	(.061)	(.044)	(.049)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Rural} * \mathbb{1}_{High\ Dist\ CC}$	-.169***	-.208***	-.180***	-.279***
	(.056)	(.067)	(.053)	(.056)
Observations	137121	137121	215333	215333
R^2	.93	.89	.93	.89

This table presents the results estimating the differential treatment effect on POS transactions across zip codes located in districts with high informality, conditional on pre-treatment levels of corruption. The unit of observation is zip code. The dependent variable is transactions from POS terminals. The dependent variable in columns (1) and (3) is transaction counts; the dependent variable in columns (2) and (4) is transactions volumes. *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist* equals 1 if the zip code's distance to the nearest currency chest (CC) exceeds the median zip code-nearest CC distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state-month-year fixed effects along with district-specific time trends and a fourth order polynomial in the number of POS terminals. District informality is proxied by the share of rural households in the district. Districts with a high (above median) share of rural households are considered to have high informality. Columns (1) and (2) restrict the sample to states with a high-level of corruption; columns (3) and (4) restrict the sample to states with a low-level of corruption. State corruption levels are based on the index prepared by Transparency International in 2006. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table A7: Differential Treatment Effects Across Zip Codes Located in Districts with High Financial Infrastructure

	(1)	(2)	(3)	(4)	(5)	(6)
	District Measures of Financial Infrastructure					
	Financial Infrastructure Index		POS Terminals Per Capita		Bank Branches Per Capita	
	#	Vol	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	-.006	-.010	-.024	-.048	.034	.017
	(.030)	(.034)	(.032)	(.036)	(.029)	(.032)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	.048	.068*	.018	.037	.072**	.084**
	(.036)	(.038)	(.038)	(.041)	(.035)	(.037)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Fin\ Infra}$	-.146***	-.260***				
	(.031)	(.037)				
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Fin\ Infra}$	-.149***	-.249***				
	(.041)	(.049)				
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Fin\ Infra} * \mathbb{1}_{High\ Dist\ CC}$	.236***	.271***				
	(.035)	(.040)				
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Fin\ Infra} * \mathbb{1}_{High\ Dist\ CC}$	.179***	.235***				
	(.041)	(.045)				
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ POS}$			-.139***	-.275***		
			(.031)	(.036)		
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ POS}$			-.164***	-.273***		
			(.040)	(.049)		
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ POS} * \mathbb{1}_{High\ Dist\ CC}$			.250***	.310***		
			(.036)	(.041)		
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ POS} * \mathbb{1}_{High\ Dist\ CC}$			.211***	.264***		
			(.042)	(.047)		
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Branch}$					-.076**	-.215***
					(.030)	(.035)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Branch}$					-.096**	-.213***
					(.040)	(.047)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Branch} * \mathbb{1}_{High\ Dist\ CC}$					.192***	.249***
					(.034)	(.038)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Branch} * \mathbb{1}_{High\ Dist\ CC}$					.153***	.226***
					(.040)	(.044)
Observations	352454	352454	352454	352454	352454	352454
R^2	.93	.89	.93	.89	.93	.89

This table presents the results estimating the differential treatment effect on POS transactions across zip codes located in districts with high financial infrastructure. The unit of observation is zip code. The dependent variable is transactions from POS terminals. The dependent variable in columns (1), (3) and (5) is transaction counts; the dependent variable in columns (2), (4) and (6) is transaction volumes. *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist* equals 1 if the zip code's distance to the nearest currency chest (CC) exceeds the median zip code-nearest CC distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state-month-year fixed effects along with district-specific time trends and a fourth order polynomial in the number of POS terminals. Columns (1) and (2) test for differential treatment effects across zip codes located in districts with high (above median) pre-treatment financial infrastructure (*Fin Infra*); columns (3) and (4) test for differential treatment effects across zip codes located in districts with high (above median) pre-treatment POS terminals per capita (*POS*); columns (5) and (6) test for differential treatment effects across zip codes located in districts with high (above median) pre-treatment bank branches per capita (*Branch*). District financial infrastructure is the sum of standardized indices of district POS terminals and bank branches per capita in the pre-treatment period. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table A8: Differential Treatment Effects Across Rural Zip Codes and Zip Code Financial Infrastructure

	(1)	(2)	(3)	(4)
	Zipcode Measure of Heterogeneity			
	Rural		High Financial Infrastructure	
	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	.190***	.228***	.069***	.089***
	(.021)	(.025)	(.024)	(.027)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	.172***	.236***	.112***	.166***
	(.024)	(.028)	(.028)	(.031)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{Rural}$	.197***	.244***		
	(.030)	(.036)		
$\mathbb{1}_{LongTerm} * \mathbb{1}_{Rural}$	.281***	.335***		
	(.034)	(.041)		
$\mathbb{1}_{NearTerm} * \mathbb{1}_{Rural} * \mathbb{1}_{High\ Dist\ CC}$	-.168***	-.210***		
	(.038)	(.046)		
$\mathbb{1}_{NearTerm} * \mathbb{1}_{Rural} * \mathbb{1}_{High\ Dist\ CC}$	-.146***	-.191***		
	(.044)	(.052)		
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Fin\ Infra}$			-.107***	-.141***
			(.020)	(.023)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Fin\ Infra}$			-.059**	-.067**
			(.025)	(.028)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Fin\ Infra} * \mathbb{1}_{High\ Dist\ CC}$			.170***	.181***
			(.032)	(.037)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Fin\ Infra} * \mathbb{1}_{High\ Dist\ CC}$			.127***	.133***
			(.036)	(.042)
Observations	317107	317107	317107	317107
R^2	.93	.90	.93	.90

This table presents the results estimating the differential treatment effect on POS transactions across rural zip codes and zip codes with high financial infrastructure. The unit of observation is the zip code. The dependent variable is transactions from POS terminals. The dependent variable in columns (1) and (3) is transaction counts; the dependent variable in columns (2) and (4) is transactions volumes. *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist* is a dummy equaling 1 if the zip code's distance to the nearest currency chest (CC) exceeds the median zip code-nearest CC distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state-month-year fixed effects along with district-specific time trends and a fourth order polynomial in the number of POS terminals. Columns (1) and (2) test for differential treatment effects across rural zip codes; columns (3) and (4) test for differential treatment effects across zip codes with high financial infrastructure. zip code financial infrastructure is based on the sum of the standardized indices of pre-treatment POS terminals and bank branches per capita in the zip code. zip codes are considered to be rural if every bank branch in the zip code is classified as rural or semi-urban. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table A9: Differential Treatment Effects Across Rural Zip Codes, Conditional on Pre-Treatment Zip Code Financial Infrastructure and District Covariates

	(1)	(2)	(3)	(4)	(5)	(6)
	Zipcode Financial Infrastructure Index					
	$< 25pc$		25-75pc		$> 75pc$	
	#	Vol	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	.112**	.067	.227***	.263***	.209***	.236***
	(.046)	(.050)	(.029)	(.035)	(.042)	(.056)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	.041	-.005	.139***	.212***	-.006	.073
	(.089)	(.087)	(.046)	(.052)	(.063)	(.079)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{Rural}$	.072	.142	.226***	.273***	.206***	.241***
	(.077)	(.089)	(.044)	(.053)	(.047)	(.061)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{Rural}$	.056	.170	.315***	.376***	.138***	.213***
	(.093)	(.106)	(.051)	(.062)	(.054)	(.069)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{Rural} * \mathbb{1}_{High\ Dist\ CC}$	-.162*	-.231**	-.180***	-.216***	-.032	-.013
	(.089)	(.102)	(.055)	(.067)	(.069)	(.089)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{Rural} * \mathbb{1}_{High\ Dist\ CC}$	-.084	-.169	-.162***	-.208***	.105	.069
	(.107)	(.120)	(.063)	(.076)	(.075)	(.097)
Observations	76394	76394	156070	156070	75119	75119
R^2	.89	.85	.94	.90	.96	.91

This table presents the results estimating the differential treatment effect on POS transactions across rural zip codes, conditional on the zip code's pre-treatment level of financial infrastructure. The unit of observation is zip code. The dependent variable is transactions from POS terminals. The dependent variable in columns (1), (3) and (5) is transaction counts; the dependent variable in columns (2), (4) and (6) is transactions volumes. *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist* equals 1 if the zip code's distance to the nearest currency chest (CC) exceeds the median zip code-nearest currency chest distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state-month-year fixed effects along with district-specific time trends, a fourth order polynomial in the number of POS terminals, and district-specific covariates interacted with a post-treatment indicator. Columns (1) and (2) restrict the sample to zip codes in the bottom quartile of zip code financial infrastructure; columns (3) and (4) restrict the sample to zip codes between the 25th and 75th percentile of zip code financial infrastructure; columns (5) and (6) restrict the sample to the top quartile of zip code financial infrastructure. zip code financial infrastructure is based on the pre-treatment financial infrastructure index for each zip code. It is computed as the sum of the standardized pre-treatment indices of POS terminals and bank branches per capita in the zip code. zip codes are considered to be rural if every bank branch in the zip code is either rural or semi-urban. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table A10: Differential Treatment Effects Across District Financial Infrastructure: Zip Codes Located in Districts with Low Informality

	(1)	(2)	(3)	(4)	(5)	(6)
	Measure of Formality					
	High Salaried Workers		High Educated Workers		High Household Expenditures	
	#	Vol	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	.022	.080	-.005	-.037	.003	.116*
	(.050)	(.054)	(.052)	(.062)	(.062)	(.065)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Dist\ CC}$	.202***	.260***	.018	.059	.214***	.296***
	(.060)	(.062)	(.060)	(.069)	(.073)	(.076)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Fin\ Infra}$	-.178***	-.201***	-.128***	-.215***	-.294***	-.330***
	(.043)	(.049)	(.046)	(.057)	(.050)	(.056)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Fin\ Infra}$	-.060	-.101	-.150**	-.210***	-.191***	-.263***
	(.057)	(.067)	(.060)	(.078)	(.065)	(.078)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Fin\ Infra} * \mathbb{1}_{High\ Dist\ CC}$	.214***	.193***	.233***	.300***	.224***	.152**
	(.053)	(.058)	(.055)	(.066)	(.064)	(.069)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\ Fin\ Infra} * \mathbb{1}_{High\ Dist\ CC}$	.032	.054	.221***	.262***	.004	.011
	(.064)	(.067)	(.063)	(.073)	(.076)	(.080)
Observations	228116	228116	216508	216508	233913	233913
R^2	.94	.91	.95	.91	.95	.90

This table presents the results estimating the differential treatment effect on POS transactions across zip codes located in districts with high financial infrastructure, conditional on the districts having low informality. The unit of observation is the zip code. The dependent variable is transactions from POS terminals. The dependent variable in columns (1), (3) and (5) is transaction counts; the dependent variable in columns (2), (4) and (6) is transaction volumes. *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist* is a dummy equaling 1 if the zip code's distance to the nearest currency chest (CC) exceeds the median zip code-nearest currency chest distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state-month-year fixed effects along with district-specific time trends and a fourth order polynomial in the number of POS terminals. Columns (1) and (2) restrict the sample to zip codes located in districts with a high share of salaried workers; columns (3) and (4) to zip codes in districts with a high share of educated workers; columns (5) and (6) to districts with high per capita household expenditures. District financial infrastructure is the sum of the standardized indices of POS terminals and bank branches per capita in the district in the pre-treatment period. Standard errors are clustered by zip code. * p <0.10; ** p <0.05; *** p <0.01

Table A11: Average Treatment Effects on Cash Withdrawals/Digital Transactions - Robustness to Dropping Outlier Distances

	(1)	(2)	(3)	(4)
	Cash Withdrawals		Digital Transactions	
	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	-.037***	-.049***	.136***	.155***
	(.012)	(.014)	(.024)	(.029)
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$	.028**	.014	.108***	.164***
	(.012)	(.015)	(.028)	(.033)
Observations	164771	164771	164771	164771
R^2	.91	.86	.91	.85

This table presents the results estimating the average treatment effect on ATM and POS transactions. The unit of observation is zip code. The dependent variable in columns (1) and (2) is transactions from ATM terminals; columns (3) and (4), transactions from POS terminals. The dependent variable is measured in transaction counts in columns (1) and (3) and transaction volumes in columns (2) and (4). *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist* equals 1 if the zip code's distance to the nearest CC exceeds the median zip code-nearest currency chest distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state month-year fixed effects along with a fourth order polynomial in the number of POS terminals and district-specific time-trends. The sample is restricted to zip codes falling in the middle 2 quartiles of the zip code-CC distance distribution. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table A12: Average Treatment Effects on Cash Withdrawals/Digital Transactions - Pre-GST Effects

	(1)	(2)	(3)	(4)
	Cash Withdrawals		Digital Transactions	
	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\ Dist\ CC}$-.114***	-.114***	.153***	.194***	
	(.009)	(.010)	(.017)	(.021)
Observations	208842	203089	203089	203089
R^2	.94	.93	.94	.89

This table presents the results estimating the average treatment effect on ATM and POS transactions. The unit of observation is zip code. The dependent variable in columns (1) and (2) is transactions from ATM terminals; columns (3) and (4), transactions from POS terminals. The dependent variable is measured in transaction counts in columns (1) and (3) and transaction volumes in columns (2) and (4). *NearTerm* is a dummy equaling 1 for all months between November 2016 and June 2017. *High Dist* equals 1 if the zip code's distance to the nearest CC exceeds the median zip code-nearest currency chest distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and state-month-year fixed effects along with a fourth order polynomial in the number of POS terminals and district-specific time trends. The sample is restricted to the pre-GST period, prior to July, 2017. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table A13: Average Treatment Effects on Cash Withdrawals/Digital Transactions - Collapsed Time Periods

	(1)	(2)	(3)	(4)
	Cash Withdrawals		Digital Transactions	
	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\,Dist\,CC}$	-.018**	-.056***	.099***	.195***
	(.008)	(.010)	(.025)	(.030)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\,Dist\,CC}$	.059***	-.015	.121***	.233***
	(.009)	(.011)	(.029)	(.033)
Observations	42481	42481	42481	42481
R^2	.97	.95	.90	.88

This table presents the results estimating the average treatment effect on ATM and POS transactions after collapsing the data to 3 time periods – pre-treatment, near term post-treatment and long-term post-treatment. The unit of observation is zip code. The dependent variable in columns (1) and (2) is average monthly transactions from ATM terminals; columns (3) and (4), average monthly transactions from POS terminals. The dependent variable is measured in transaction counts in columns (1) and (3) and transaction volumes in columns (2) and (4). *NearTerm* is a dummy equaling 1 for the period between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist* equals 1 if the zip code's distance to the nearest CC exceeds the median zip code-nearest currency chest distance in the sample. The outcome variable in each specification is logged. All specifications include zip code and district-time period fixed effects along with a fourth order polynomial in the number of POS terminals. Standard errors are clustered by zip code. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table A14: Average Treatment Effects on District-Level Cash Withdrawals and Digital Transactions

	(1)	(2)	(3)	(4)
	Cash Withdrawals		Digital Transactions	
	#	Vol	#	Vol
$\mathbb{1}_{NearTerm} * \mathbb{1}_{High\,Dist\,CC}$	-.084***	-.062***	.075	.178***
	(.021)	(.022)	(.057)	(.055)
$\mathbb{1}_{LongTerm} * \mathbb{1}_{High\,Dist\,CC}$	-.090***	-.060*	.038	.176***
	(.027)	(.032)	(.073)	(.063)
Observations	13916	13916	13916	13916
R^2	1.00	.99	.99	.99

This table presents the results estimating the average treatment effect on ATM and POS transactions at the district-level. The unit of observation is the district. The dependent variable in columns (1) and (2) is transactions from ATM terminals; columns (3) and (4), transactions from POS terminals. The dependent variable is measured in transaction counts in columns (1) and (3) and transaction volumes in columns (2) and (4). *NearTerm* is a dummy equaling 1 for all months between November 2016 and November 2017; *LongTerm* is a dummy equaling 1 for all months after November 2017. *High Dist* is a dummy equaling 1 if the median zip code-nearest currency chest (CC) distance within the district exceeds the median, within district, zip code-nearest CC distance across all districts in the sample. The outcome variable in each specification is logged. All specifications include district and state-month-year fixed effects along with a fourth order polynomial in the number of POS terminals in the district and district time-trends. Standard errors are clustered by district. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$