Shocks and Technology Adoption: Evidence from Electronic Payment Systems *

Nicolas Crouzet, Apoorv Gupta and Filippo Mezzanotti†

First version: September, 2018
This version: January, 2019

Abstract

We provide new evidence on the diffusion of technologies subject to positive adoption externalities. Using data on Indian electronic payment systems, we show that the 2016 demonetization - which led to a temporary reduction in cash - caused a permanent increase in the adoption of electronic payment systems by retailers. We show that these dynamics are consistent with a technology choice model with positive externalities in adoption. A number of distinct predictions of the model — in particular, history-dependence in adoption and the importance of spillovers between neighboring firms — receive strong support in the data. Furthermore, evidence on the impact of the demonetization on household consumption suggests that this rise in adoption played an important role in limiting the costs of the shock. Our results support the view that, when coordination problems slow down the diffusion of technology, aggregate shocks can act as coordination devices and shape adoption dynamics.

Keywords: Complementarity, Technology Diffusion, Fintech, Demonetization.
JEL Classification: O33, E51, E65

*We thank Pat Akey, Anthony DeFusco and Ravi Jagannathan and participants at Entrepreneurial Finance and Innovation Workshop for helpful comments and discussions. We gratefully acknowledge financial support from the Financial Institutions and Markets Research Center, Kellogg School of Management. We are also grateful to the staff at the wallet company for help with their data. The data is shared solely for the purpose of academic research. No user data has been shared in any form. The wallet company does not have any role in drawing inferences of the study and the views expressed herein are solely of the authors.

†Crouzet: Northwestern University; Gupta: Northwestern University; Mezzanotti: Northwestern University.

1 Introduction

A large literature in economics has focused on exploring how technological innovation affects long-term growth. A crucial component of the link from innovation to growth is the diffusion of new technologies, as opposed to their invention (Hall and Khan, 2003). In practice, the adoption of new technologies by firms is often a slow process, encumbered by many potential barriers (Rosenberg, 1972). The empirical literature offers several examples of firms failing to use efficiency-enhancing technologies (Mansfield, 1961) or processes (Bloom et al., 2013), for reasons ranging from the presence of organizational frictions (Atkin et al., 2017), to slow social learning (Conley and Udry, 2010; Munshi, 2004), to lack of financial development (Comin and Nanda, 2014).

In a context where the decision to adopt a technology generates externalities for other firms, barriers to adoption may arise simply as a result of coordination problems. Individual firms may expect either high or low adoption rates of a new technology by competitors. Consistent with these expectations, they may either adopt or reject the technology, giving rise to multiple possible equilibria. In these situations, understanding how firms may coordinate on technology adoption, or fail to, is an important question, for both research and for policy — all the more so in contexts where adoption has implications for other sectors or for consumers.

In this paper, we study the role that aggregate economic shocks can play in resolving these coordination problems. Using a novel empirical setting, we provide evidence that an aggregate shock can lead to permanent shifts in technology adoption, even when the shock itself is temporary. We show that such permanent shifts in response to transitory shocks can arise in a dynamic technology adoption model where firms' decisions are complements. We then show that other key predictions of the model — in particular, path-dependence in adoption — are strongly supported by the data, suggesting that complementarities played a key role in shaping adoption dynamics in our setting. Finally, we document how shifts in adoption can help mitigate negative effects of the shock itself.

Our analysis is focused on a particular technology: electronic payment systems. Electronic payment systems are a good example of the coordination problems we analyze for two reasons. First, electronic payment systems are networks and therefore their adoption is subject to strong externalities (Katz and Shapiro, 1994). The returns to opting into the payment system for a particular firm depends on the size of the network, that is, the existing number of other firms accepting the payment system in their operations. This makes the decisions to opt into the network complementary across firms, thus generating the type of coordination problems in adoption described above. Second, the specific instance of electronic payment system we study, an electronic wallet, is characterized by sunk adoption costs that are essentially zero (a point we discuss in more detail in the paper). This feature allows us to isolate coordination problems from

sunk costs, another potential barrier to technology adoption (Basker, 2012).

We analyze electronic payment systems in a setting where the returns to their adoption was subject to a large, temporary shock: the Indian demonetization of 2016. In the fall of 2016, the Indian Government announced that it would void the two largest denominations of currency in circulation, replacing them with new bills. Households and firms were not given advance warning of the swap, but instead they had a grace period of almost two months to exchange the cash. The voided bills represented about 86% of total cash in circulation. During this grace period, old currency would need to be deposited in the banking sector, and the amounts deposited could then be withdrawn in new bills. However, several frictions in the creation and distribution of the new bills meant that agents were not immediately able to withdraw cash. As a result, cash in circulation fell and bank deposits spiked, a point we document in section 2. Cash transactions became potentially harder to conclude, but more funds were available for use in electronic payments.

We first show that the Indian demonetization was associated with a wave of adoption of various electronic payment technologies. We focus primarily on data from the largest Indian provider of non-debit card electronic payments. This payment platform operates as a digital wallet. The digital wallet consists of a mobile app that allows consumers to pay at stores using funds deposited in their bank accounts. Payment is also received by retailers using the app. As mentioned above, the costs associated with the adoption of this technology for firms are small or inexistent; all that is required is to have a bank account and a mobile phone, which both had high ownership rates in India by 2016 (Agarwal et al., 2017).

To start, we use both an aggregate event-study and cross-sectional variation to provide evidence that the cash swap caused an increase in the uptake of this technology. In aggregate, the increase of the use of the electronic wallet around the demonetization shock was large, with the number of retailers accepting the wallet, the number of transactions, and the total value of transactions carried through this platform roughly doubling every week during the grace period. For the cross-sectional analysis, we exploit variation in the exposure to demonetization across Indian districts to argue that the effect of demonetization on adoption was causal. Indian districts had varied exposure to the shock because their local banking sectors were not all equally prepared to execute the swap between old and new bills, a point we examine in detail in section 3. Certain districts experienced sharper increases in bank deposits, with bank customers unable to withdraw new currency. Using a treatment measure based on the excessive deposit growth during the demonetization, we show that adoption of electronic payments was significantly larger in these highly affected locations.[1]

Additionally, we show that adoption was *persistent*, though the shock was not. In aggregate, there was no visible mean-reversion in adoption or transaction volumes once the new currency was successfully put back

[1]Specifically, the excessive deposit growth measures the change in deposit relative to the average deposit growth in the previous two years. The paper will discuss in detail this measure and provide empirical evidence in support of this approach.

in circulation. In the cross-section, districts with a higher exposure to the shock also experienced a stronger permanent increase in adoption. We find some evidence that relative differences in adoption across districts narrowed down after the new currency was put back in circulation, but overall adoption in highly affected districts stayed higher relative to other districts. Finally, we show that - with some caveats - adoption effects are also visible in the use of debit card payment terminals, while the effect appears much weaker for credit card payments.

The evidence we present supports the view that the demonetization, though a transitory shock, caused a permanent increase in the adoption of the technology. However, it does not directly speak to the question of whether externalities played a role in shaping these adoption dynamics. In order to answer this question, we start by analyzing an industry model of technology adoption with externalities. Our model is based on the framework of Burdzy et al. (2001). Firms face a choice of two technologies under which to operate; one of which is subject to positive externalities — the flow profits from operating under this technology increase with the rate of use of the technology in the industry as a whole. Moreover, the relative benefit of adopting the technology is subject to aggregate shocks. Aggregate shocks in this model help eliminate the potential equilibrium multiplicity arising from complementarities in the adoption decision. By inducing common shifts in the benefits from adoption, shocks effectively help firms coordinate on adoption decisions.[2]

The central insight of this model is that complementarities generate a form of history-dependence. Specifically, the way in which use of a technology in a particular industry or district will respond to a sequence of aggregate shocks depends on how widely firms perceive the technology to be initially used. In the model, in an industry or district with low initial adoption rates, a large shock may temporarily raise adoption. However, once the shock is undone, the adoption rate will converge back to its initially low level. By contrast, in industries with a higher initial rate of use of the technology, the same transitory shock may lead to full and permanent adoption.

We use the model to argue that the prediction of history-dependence hinges on the existence of complementarities, and thus provides a way to test for their importance in explaining the adoption dynamics we observe in the data. Specifically, using simulated data from the model, we show how a number of district- and firm-level regressions can be used to discriminate between models with and without complementarities. At the district level, these regressions test for the dependence of the adoption path on the initial level of

[2]The model we analyze is closely related to the literature on global games and equilibrium selection (Carlsson and Van Damme, 1993; Morris and Shin, 1998, 2003). This literature has also analyzed the effects aggregate (public) signals in environment where agents' actions are complements (Morris and Shin, 2002). The two key differences of the framework we study is with global games models is that (a) firms have no private information on the returns to adoption; (b) firms solve a dynamic coordination problem, instead of a one-shot, static model. The latter difference leads to the prediction that adoption should be path dependent; this is particularly important in our empirical analysis, as we show that there is strong evidence of path dependence in adoption in the data. See Burdzy et al. (2001) for a more detailed discussion of the relationship of this framework with the global games literature.

adoption. At the firm level, the regressions test whether a firm's individual decisions are a function of the share of existing firms in the district already using the technology.

We then compare these predictions to the data on the Indian demonetization experiment. First, we examine the importance of history dependence and initial conditions at the district-level. We find that districts located close to a major electronic payment hub experienced relative larger increases in adoption during the months of the demonetization. The idea is that being located in proximity to a large hub proxies for the benefits of adoption initially perceived by firms.[3] To identify these effects, we estimate a difference-in-differences model where we compare districts within a state that differ in terms of distance to a major electronic payment hubs. We find that in the months of the demonetization the increase in adoption was much larger for districts close to large hubs. This result is not driven by differential trends and also holds when we control for other district-level characteristics as well as distance to the local state capital. Thus, the data features the same *district-level* history dependence as predicted by the model.

Second, we show that *firm-level* adoption decisions depend on the share of neighboring firms already using the technology. Using a specification derived from the analysis of the model, and similar to Munshi (2004), we estimate how the intensity of adoption by other firms in the same zipcode and business area affects the adoption decision at firm-level.[4] Overall, we find that adoption by neighboring firms in the same area has a strong, positive effect on the usage of the technology. Consistent with the predictions of the model, this effect is particularly large after the shock as firms transition to the new technology. Importantly, these results do not simply capture variation across location and industry, since they hold when we augment the analysis with a wide set of zipcode by time and industry by time fixed-effects.

We conclude the paper with an analysis of how technology adoption may have helped mitigate the real economic effects of the shock. Following the approach discussed before and using a large panel of households in India, we find that the districts experiencing stronger excessive deposit growth — districts for which the demonetization arguably represented a bigger shock — experienced a deeper contraction in household consumption in the months immediately after the shock. The contraction in consumption is mostly temporary and it is concentrated in the subset of the consumption basket that is more likely to be unnecessary and infrequent. We show that these effects are much smaller for households that are more likely to have an easy access to electronic payments technology. The paper discusses several robustness checks to these results.

We believe that our paper contributes to the existing literature in three areas. First, our results contribute to the work studying the process of diffusion of technologies across firms. The literature has provided evidence

[3] We also directly test for dependence on initial adoption rates at the district level, and find significant effects there too. However, we argue that this type of analysis is problematic for both conceptual and econometric issues. This calls for using other proxies for initial benefits to adoption.

[4] In reality, the analysis is conducted at pincode level, which is a geo-unit in India similar to a 5-digit zipcode in the US. Section 5 provides more details on the pincodes in India.

on a number of potential barriers to technology diffusion. Atkin et al. (2017) suggest that the presence of organizational frictions may prevent adoption of cost-reducing technologies. At the same time, several papers highlight the role of social learning in technology adoption, in particular in the agricultural context(e.g. Conley and Udry (2010), Munshi (2004). Lastly, using cross-country data, Comin and Nanda (2014) emphasize the role of financial development for fostering technological diffusion.[5] Our results provide novel evidence on the importance of coordination frictions in technology adoption. We show that complementarities played an important role in adoption following the Indian demonetization shock. More broadly, our results shed light on the role that policy might play in incentivizing technology adoption in the face of coordination problems. They suggest that large, temporary policy shocks — such as transitory taxes on the use of particular technologies — may be sufficient to permanently shift adoption equilibria toward more desirable outcomes, at least when externalities in adoption are present.

Second, our paper contributes to the understanding of the impact of the demonetization on the Indian economy. We believe that our paper makes two important contributions here. First, we show that the cash swap led to a persistent increase in electronic payment technologies, and argue that complementarities are key to understanding these adoption dynamics. In this dimension, this work is related to the paper by Agarwal et al. (2018), which document similar aggregate shift in the use of electronic payments during the demonetization in India. Relative to this paper, our research proposes a specific economic mechanism that can account for adoption dynamics — the complementarity of adoption decisions — and we provide a novel characterization of adoption dynamics highlighting the importance of path-dependence, both at the firm- and district-level. Second, we provide causal evidence that the reduction in the availability of cash led to a reduction in household consumption. These results are consistent with the effects on Indian output estimated by the contemporaneous paper by Chodorow-Reich et al. (2018). While our focus is mostly on the importance of network externalities during this period, our household-level analyses also extend their district-level analysis by showing that the negative effects of the shock may be particularly large for poorer households.

Lastly, our paper also provides novel evidence to the growing literature on fintech. Most of the literature has focused on the impact of financial technologies as a way for firms and households to obtain financing (Buchak et al., 2018; Tang, 2018; Fuster et al., 2018; de Roure et al., 2018; Howell et al., 2018), but much less work has been devoted to their impact on payment systems. However, previous evidence suggests that better payment technologies can play an important role in alleviating poverty and foster economic activity in poor countries (Yermack (2018), Suri and Jack (2016)). Our results highlight in particular their lower adoption costs relative to more traditional electronic payments. We show that aggregate participation in

[5]For a complete review of the literature of technology adoption, see work by Foster and Rosenzweig (2010).

traditional systems, such as credit cards, did not change as much as for the electronic wallet. Furthermore, our insights on how temporary shock can have permanent impact on adoption can be useful for policy-makers and business leaders interested in the expansion of these markets. This discussion is particularly relevant given recent policy debate on the costs of cash and cash alternatives in modern economies (Rogoff, 2017).

The rest of the paper is organized as follows. Section 2 discusses the specifics of the Indian demonetization experiment, and in particular its effects on banking deposits. Section 3 provides evidence on the causal effect of demonetization on the adoption of electronic payments. Section 4 analyzes a dynamic model of technology adoption with complementarities, and highlights how its predictions can be used to test for the presence of complementarities in the data. Section 5 applies these tests to the electronic payment data. Section 6 documents further the real effects of the shock, and Section 7 concludes.

2 Background

This section describes the setting which we study in this paper: the 2016 Indian demonetization, and in particular, its effects on banking sector deposits.

2.1 Main events

On November 8, 2016, Indian Prime Minister Narendra Modi announced a large-scale demonetization. As per the announcement, starting midnight that day, the two large denomination bills of Rs.500 and Rs.1,000 which corresponds to approximately $10 and $20 respectively - would be illegal tenders and not be accepted as means of transaction payment. At the same time, the Government planned to replace the voided notes with new Rs.500 and Rs.2,000 notes. With this unexpected policy, the Government was aiming to identify individuals holding large amounts of black money. In fact, in order to convert the old cash into new cash, individuals had to deposit the notes into their bank accounts by the deadline of December 30, 2016. People still holding these denominations after the deadline would be subject to pecuniary penalties or even prison time.

The demonetization voided around 86.9% of the nominal stock of currency. We can directly quantify the contraction of cash that is available to citizens by looking at the amount of cash in circulation reported by the Indian Government and plotted in Figure 1. To be clear on what this figure captures, the amount of money in circulation does not mechanically drop with the Government decision, but it declines as the voided notes are deposited in the banking sector. From this figure, we find that the amount of cash in circulation dropped in November by almost 50%. Because the deadline to deposit the voided notes is at the end of the year, the amount of cash in circulation declines even more in December, and starts to increase only with the

new year. The shortage of cash affected almost every household as cash was the major transaction medium for both households and businesses. Cash levels kept increasing post-January and reached pre-shock levels at the end of spring. Thus, the policy led to a large cash shortage, suggesting that the transition between old and new notes was not as smooth as the Government may have initially had planned. In principle, people could have brought back to the bank the old notes and immediately received the new ones, with little costs for households and firms. However, several issues made this smooth transition impossible.

First, the amount of cash required to be printed was several magnitudes higher relative to the average daily printing quantity , causing slow-downs in the actual production of new currency. Second, there were several issues in the distribution of the new currency bills, which exacerbated the initial delays. After printing, the new notes were distributed to a subset of financial institutions that acted as currency chest bank. The role of these institutions was to provide the new notes to banks operating in their local markets. Therefore, the geographical distribution of these intermediaries and their relationship with the rest of the banking sector affected the timing and quantity of the new notes available in a district. More broadly, the structure of the local banking market significantly impacted the printing and diffusion of the new currency, as confirmed in our analysis below.

Lastly, the retail banking sector was unprepared to the large increase in business caused by the sudden need of millions of families to deposit and withdraw money. In fact, the deposits of old currency or the withdrawal of new one had to go through the regular branching network, which was overwhelmed by the large increase in business. In this regard, the press has reported several cases of massive branch queues, which made it even more expensive for regular household to access the new notes. For instance, in a survey of 214 households in 28 slums in the city of Mumbai, 88% households reported waiting for more than 1 hour for ATM or bank services, between 11/09/2016 and 11/18/2016. In the same survey, 25% households reported waiting for more than 4 hours. In line with this result, another randomized survey conducted over nine districts in Indian by a mainstream newspaper, Economic Times, showed that the average number of visits to either a bank or ATM increased from an average of 5.8 in month before demonetization to 14.4 in the month after demonetization.[6]

2.2 The demonetization and bank deposits

In the previous section, we discussed how the decision of the Indian Government to void a substantial part of the cash in November of 2016 led to a sizable shortage of cash. The flipside of this shortage of cash has been the increase in bank deposits that the country experienced over the same period. In this section, we

[6]Lastly, many other idiosyncratic events affected the process of moving to the new currency denominations. For instance, the new notes had a different size than the old ones, and therefore ATMs machines across the country had to be refitted before they could dispense the new Rs.500 currency notes.

explain how these two quantities are related and we argue that bank deposits can be used to proxy for the shortage of cash in the economy. This discussion is important for two reasons. First, understanding the relationship between these quantities will be important to explain the rise in electronic payments. Second, as we discuss in detail, we believe that information on deposit growth can be used to identify variation in the intensity of the shock across Indian districts.

The reason why bank deposit surged as cash became scarce is quite intuitive. In order to convert the old notes, households had to deposit them in a bank by December 31. Since the new notes were distributed through the bank branch network, the idea was that individuals could then withdraw cash from the bank as soon as the new bills were available. However, the transition was characterized by delays and therefore the new currency was not available in large quantities until the New Year. These two facts together the need to deposit cash and the inability to withdraw it led to the contraction of cash in the real economy and a contemporaneous surge in deposits.

The aggregate time series of the amount of deposits provides evidence that is very consistent with this narrative. In particular, data shows an exceptional and unprecedented increase in bank deposits during the last quarter of 2016. In figure 2, we plot by quarter the average of yearly growth in deposit across all districts from quarter 1 of 2013 to quarter 1 of 2017. The figure shows that in the year before the shock deposits in the Indian economy had been growing steadily quarter-by- quarter. This growth was relatively small — the average deposit growth across districts was around 12% annualized. However, in the last quarter of 2016, the average growth in deposits reached an unprecedented 30% relative to the previous year. The average level of deposits were still higher — about 20% — in the first quarter of 2017 relative to the previous year, suggesting that the effects on deposits did not dissipate fast. This is consistent with the evidence on the cash in circulation, which was still lower than the pre-demonetization levels in the first quarter of 2017. Only starting with the second quarter of 2017 do we see deposits slowly returning to normal levels.

This analysis already highlights one of the key idea here: the voided cash did not disappear from the economy; it was still available to Indian families in their bank accounts. As soon as a family deposited the cash, they were able to use it electronically, since electronic transactions provide a way to transfer money between bank accounts. The same clearly also was true for firms.

The second point we want to make in this section is to argue that deposit growth could be used to capture some of the impact of the demonetization in the area. The first-best option would be to measure the impact of the demonetization using data on the amount of voided cash in a certain area. However, this information is impossible to obtain at a disaggregated level, since the Government does not track notes across the country. We believe that a measure based on growth in deposit represents a good second-best option to proxy the amount of cash balances that households had to forego because of the new Government

policies. In particular, later we will use an excessive deposit growth, where we normalize deposit growth in the shock period with the growth that the district experiences normally.[7]

Given our discussion of the role of deposits in the transition, it is intuitively clear why deposits may be used to proxy the effect of the shock in a local area. In practice, we think there are three important pieces of evidence that supports this idea. First, the aggregate evidence provided suggests that the aggregate behavior of deposits tracks well the impact of the shock over time. This fact confirms the connection between the two phenomena. Second, the data on the dispersion of deposit growth at district-level suggests that the demonetization not only affected aggregate deposits, but also the distribution of deposits across districts. To look into this issue, figure 3 presents the distribution of deposit growth across Indian districts in the last quarter of 2016 versus 2015. The difference between the two is stark. The 2016 shows much more dispersion, while the previous year distribution is very tight around a small positive growth. On top of the high dispersion, two things are worth noting. First, the growth in deposit was essentially positive in most districts in India. This is a remarkable fact, given that in normal time this figure would be more widespread. For instance, in the last quarter of the previous year more than 15% of districts experienced a decline in deposit. Second, many districts experienced an unprecedented growth in deposits, which is also very rare on the previous year.[8]

Third, we can show that the heterogeneity in deposit growth is correlated with factors that are likely to predict the speed of flow of cash into the economy. In the previous section, we argued that the structure of the local banking industry played an important role in determining the extent to which the demonetization led to a sizable shortage in cash. We have particularly highlighted the role of chest-banks in helping the flow of cash into the system. In fact, more chest-banks will affect the availability of cash in the local market, either directly since chest banks have direct access to cash for their customers or indirectly, through the positive impact on other banks operating in a district.[9] Therefore, we expect that in districts with a larger density of chest banks, the replacement of the old notes was more efficient, therefore causing less of a shortage of cash.

Indeed, we find evidence that is consistent with this. In particular, table 1 reports a simple regression where we explain deposit growth at the district level normalized to have mean zero and standard deviation one - with a score that is equal to one minus the share of chest banks in the area. In column (1), we use

[7] Specifically, we will use a measure that is equal to deposit growth in the last quarter of 2016 minus the average quarterly growth in the previous two years.

[8] The result is essentially the same if we compare 2016 with data from 2014 on deposit growth dispersion.

[9] There is some anecdotal evidence of conflicts in the distribution of cash between different banks. In particular, given the shortage of new currency bills, chest bank may have favored their own operational needs for cash relative to the mandate of helping the distribution of new notes in the economy. Few anecdotes seem to support this hypothesis. For example, in the district of Gurgaon in Northern region of the country, non-chest banks filed a complaint with the Central Banks against the chest-banks for providing preferential treatment to the parent bank branches for the distribution of cash.

the continuous version of the score; we find that districts characterized by fewer chest banks experienced a larger increase in deposit during the last quarter of 2016. The same result is confirmed when we use a dummy variable, equal to one for districts in the bottom quartile in terms of chest bank density (column 3). This result is robust when controlling for the logarithm of total deposit in the district, suggesting that the density of chest banks does not simply capture some heterogeneity in wealth across region. Furthermore, in the Appendix Table B.1, we discuss a set of placebo tests in which we show that the same strong relationship between chest bank and deposit growth cannot be replicated before the shock.

These results are consistent with the idea that deposit growth can provide relevant information on the size of the cash-contraction at local level. As we present results using this measure in the paper, we will also discuss more in detail the shortcomings of this approach and the various set of tests that we developed to try to assuage the main concerns related to it.

3 The effects of demonetization on payment technology adoption

In this section, we will use data on the Indian demonetization to examine the relationship between technology adoption and large economic shocks. Our focus will be on the adoption of technologies that allow firms and consumers to transact without the use of cash. In particular, our data will cover the universe of transaction for the leading Indian company providing wallet payment technologies, which allow individuals and businesses to exchange payment using the phone.

The objective of this section is to provide some stylized facts about how the adoption of these new technologies followed the large cash-contraction caused by the demonetization. These analyses will highlight a couple of results, which are important to understand how economic shocks can affect adoption. First, we show that the demonetization had a significant impact on the adoption of non-cash payment technologies, measured using our data on wallet payment technologies. Second, even if the shock is only temporary, the shift in adoption is permanent. Third, we find that this positive response to the shock was not unique to the wallet payment technologies examined here, but similar results can be found also for traditional electronic payment methods (e.g. credit or debit cards). However, these effects at least at the extensive margins appear to be smaller, suggesting that the initial cost in adopting the new technology is important to explain the patterns in the data.

3.1 Data

In the empirical part of this paper, we study the adoption of the new technology by measuring the increase in use of electronic payments in India in 2016. In particular, we measure the adoption of electronic payments

using data from one of the leading digital wallet companies in the economy.[10]

This company allows individual and businesses to undertake transactions between each other using only their mobile phone. In order to use their services, a customer needs to link the bank account to the application. Then, merchants can use a uniquely assigned QR code to accept payments directly from the customers into their mobile wallet which can then be transferred to their bank account. Therefore, the technology is in many respect similar to a credit card or other more traditional electronic payment systems. However, relative to these other electronic payment technologies, adoption costs are much lower, since retailers and consumers can access the electronic wallet almost instantaneously without the need of anything more than a phone and a bank account. In our data, we observe sales amount and number of transactions on a weekly basis for anonymized merchants using the platform between May 2016 to June 2017. For each merchant, we also know the location of the shop at city-level and the detailed industry. For a random sub-sample of shops, the location is provided at much more detailed level of 6-digit pincode. A pincode in India are equivalent to a five-digit zip-code of US and was created by postal service in India.[11]

There are two key advantages of this data relative to other data on electronic payment available for this setting. First, it is relatively high-frequency, since we can aggregate the data at week or monthly level. Second, the transaction is geo-localized, therefore allowing us to aggregate them up at the same level as other data sources used in this study. These two features together allow us to use this data to examine the shift towards electronic payment systems in India at the time of the demonetization. In the rest of the analyses, we also use other sources of data. First, we obtain information related to banking deposits, bank branches and chest bank branches by district from the Reserve Bank of India website. Second, we obtain data on other district characteristics from the 2001 Census of India. This includes information on district-level growth, sex ratio, literacy rate, population, number of villages with electricity/water access/telephone access, and number of households.

3.2 The adoption of electronic payment technologies

3.2.1 Aggregate evidence

As a first step in this analysis, we want to understand whether the contraction in available cash led to an increase in adoption of electronic payment technologies. As previously discussed, in this first part we will focus on wallet electronic payment as a laboratory to examine the shift to electronic payment technologies.

We start by showing that the demonetization led to an overall uptake in electronic payments. In order

[10] During the period of our study, the mobile platform was the biggest firm in the market providing mobile transaction service. Since March 2017, few other public platforms emerged because of the government's initiative of cashless economy.

[11] India has a total of 19,238 pincodes out of which 10,458 pincodes are covered in our dataset.

to explore this dimension, we use the data at weekly level and we examine the weekly growth in the number of transaction and amount changed around the week in which the demonetization was announced and implemented. This approach is equivalent to an event study design, and therefore the underlying assumption is that apart from the Governments decision no other factor affected the incentive to use this technology.

The results are reported in the two panels of figure 4. In the months before the demonetization, the weekly growth in the usage of the wallet technologies had been positive on average but relatively modest. However, in the week following the demonetization the shift towards this payment method was dramatic. In particular, the week after the demonetization the transaction grew by more than 150%, while amounts transacted increased by almost 200%. Furthermore, for the whole month after the shock, weekly growth rates were consistently around 100%. In general, we find these patterns both looking at the number of transactions and the total amount transacted.

One important observation is that this initial positive effect on adoption from demonetization did not dissipate with time, even when cash in circulation reached its pre-crisis level again. In other words, this evidence suggests that a temporary shift in the availability of cash led to a permanent increase in the usage of electronic payment technologies. In particular, the data suggests a slow-down in aggregate growth starting in January - which is when the new cash starts re-entering in the system - and a drop in the overall use of the technology in February.[12] However, after this small adjustment, the average growth rate over the next couple of months remain on average small but positive, confirming that users did not abandon the platform as the original mean of payment – cash – becomes widely available again.[13]

3.2.2 Cross-sectional evidence

Next, we show that the demonetization did not only lead to a generalized increase in the use of electronic payment technologies, but that this increase was particularly strong in areas where disruptions in the swap between old and new notes were large. In order to explore this hypothesis, we develop an identification strategy that compares the uptake in electronic payments across districts that experienced different contractions in cash, which we proxy with the excess increase in deposit during the demonetization period relative to the growth rate during normal times.

The intuition behind this approach is simple. As discussed before, all voided currency had to go through the banking sector by the end of the year to be converted. Therefore, if we remove the average deposit growth that a district experience during normal times, deposits may differentially change across districts

[12] We believe that this decline is related to the introduction of a small fee in February and an increase in competition and entrance by other electronic payment companies.

[13] In some ways, this result is even more surprising given the increase in competition for electronic wallet payments in the Spring 2017. Since our data covers only one company, the overall trend of the technology will then be even more positive.

during the demonetization quarter for two main reasons. First, districts may differ in the nominal amount of cash voided that they were initially endowed with. Second, the local banking sector in a district may be more or less efficient in swapping old and new currency. in the economy. Independent from the specific channel, the growth in deposit in a district provides us with a direct proxy of the amount of money that households deposited in the banking sector and they could not withdraw during the same time.

To be precise, this approach will deliver meaningful results only under two assumptions. First, it must hold that growth in deposit at district level is correlated with the effect of the demonetization. In the previous section, we provided both time-series and cross-sectional evidence to show that the growth in deposit in last quarter 2016 was really an unprecedented event and a likely result of the demonetization.

Second, we also need to assume that excess growth in deposit is uncorrelated with other factors that may have also caused a relative increase adoption in technology. Indeed, districts that received an high-abnormal increase in deposit are on average different than the others. In particular, Table 2 compares districts with high deposit growth to districts with low deposit growth in the last quarter of 2016 on a large set of observable characteristics, sourced from the 2001 Census.[14] Comparing the top quintile to the rest of the distribution, we find that highly affected districts are smaller, have less deposit, and lower literacy rate. However, they do not particularly differ in other poverty metrics, such as the share of temporary housing, access to water, school, or phones. This result suggests that it will be important for the results to control for a wide set of covariates, to reduce the difference on observable across treatment.

While these controls are important, they may not be sufficient to causally interpret our results. Therefore, to bolster our results, we also undertake two other exercises. First, as we present the results, we will show that areas characterized with high deposit growth really change behavior right around the Government announcement, and not before. In other words, these results on deposit growth really do not capture differential trend across districts. Second, in the last part of the paper we use household-level consumption data to develop a set of placebo tests that validate this approach. With the electronic payment data, one of the data constraint is that data is sparse before May 2016. This constraint limits the possibility to run placebo tests for our main specification. However, we have also collected consumption data dating back to 2014. This data will be used to assess the real costs of the shock, but they can also be used to explore the quality of this identification. In particular, the data allows us to estimate the same specification during periods where there were no sharp change in the cash available. Using this data, we will show that during normal times, relative difference in excess deposit growth are not particularly predictive of household behavior in our data. These results are discussed later in Section 6.

[14] A notable difference here is the amount of deposit, which is sourced from the Central Bank data and it refers to the second-quarter of 2016 (the one before the demonetization).

To examine this question, we estimate a standard difference-in-difference model, where the main treatment variable is a measure based on excess-deposit growth in the last quarter of 2016. We construct this measure as the difference between the growth in total deposit in last quarter 2016, scaled by the average quarterly deposit growth over the previous two years. In other words, we estimate:

$$\log(y_{d,t}) = \alpha_t + \alpha_d + \delta\left(\mathbf{1}_{\{d\in\mathcal{T}\}} \times \mathbf{1}_{\{t\geq t_0\}}\right) + \Gamma_t' Y_d + \epsilon_{d,t}, \tag{1}$$

where t is time (a month in this setting), and d indexes the district, t_0 is the time of the shock (November 2016) and $\mathcal{T}$ denotes the treatment group, which we define based on the size of excess deposit growth in district d around the shock, as explained below. The equation is estimated with standard errors clustered at district level, which is the level of the treatment (Bertrand et al., 2004). The sample consists of all districts in India over the period of interest, with the exceptions of largest metropolitan areas.[15] However, results are robust to their inclusion. Lastly, in the main regression we include data between May 2016 and February 2017.

As we mentioned, we augment the baseline specification with a wide set of district level controls, Y_d. These contain share of people in the village with a phone, share of people in temporary houses, the logarithm of the total population, the sex-ratio, and the literacy rate. Furthermore, we control for difference in wealth, measured by the log of total deposit in the quarter before the shock. Each control is interacted with a post to allow for heterogeneous effects before and after the shock. As we mentioned before, all these controls are measured before the shock. In particular, all controls but for the deposit - which refer to deposit in the second quarter 2016 - comes from the Census of India 2001.

Table 3 presents results from this analysis. First, we look at the results splitting the sample across those districts in the top quintile of our treatment measure versus the rest of the sample. Looking at the number of transactions, we find that locations which experienced an exceptional growth in deposits saw their electronic transactions increase by 70% more than locations where growth in deposits was smaller. Similarly, the same highly affected districts experienced an 86% increase in transaction amounts. Using the continuous definition of the treatment, we obtain very similar results, in both economic and statistical significance. If anything, the result with the continuous measure appears to be more precisely estimated.

In the two panels of figure 5, we explore more precisely the timing of the effects using the continuous treatment definition. In particular, these figures plot the coefficients δ_t from the regression model:

$$\log(y_{d,t}) = \alpha_t + \alpha_d + \delta_t\left(\mathbf{1}_{\{d\in\mathcal{T}\}} \times \mathbf{1}_{\{t\geq t_0\}}\right) + \Gamma_t' Y_d + \epsilon_{d,t}. \tag{2}$$

[15] The reason for this decision is to make sure that results were not driving the dynamic of electronic payments in very large metropolitan areas. In particular, we drop the district of New Delhi, Mumbai, Chennai, Bangalore, and Kolkata.

These coefficients are normalized relative to October, which is the month before the shock. Looking at the amount and number of transactions, two results stand out. First, we find the findings from the estimation of the difference-in-differences model (2) are not simply driven by differential trends between high- vs. low-affected regions. In general, the effects estimated in the pre-periods are small in magnitude and non-significant. Second, we find that the shift in adoption is sizable already in the first month of the demonetization. However, the estimated effect is even larger in December and January, suggesting some delays in response.

The figure also confirms that the shock led to a permanent heterogeneity in adoption across districts. In particular, we find that the coefficients are still sizable and positive in the months after January. The results show that this gap between highly and low-affected districts is reduced as cash comes back in the economy, but it remains relatively large. Since the overall uptake of the electronic wallet did not decline post-January, this result is likely driven by less affected regions increasing their relative adoption over time. Anyhow, this result is consistent with what we observe in aggregate, and confirms that the temporary shock led to long-term effect in the adoption of the payment technology.

3.2.3 Fintech vs. other electronic payment systems

Because of data availability, the section has explored how electronic payment methods changed during the demonetization using digital wallet payment as measure of electronic payment adoption. As we discussed earlier, our data provider is the leading player in the digital wallet market in India and therefore our data is likely to represent the largest majority of transactions in this market. However, a large share of electronic payments are undertaken using traditional technologies, like credit or debit cards.

In this section, we use data on debit and credit cards market to shed some lights on this market and relate its dynamics to what we have found using our data. Unfortunately, the only publicly available data we were able to obtain were monthly-level time series aggregate at national level.[16] Therefore, while the set of analyses that we can undertake with this data a limited, we believe that evidence in this area can provide useful evidence to corroborate our main results.

As a first step, in the top panel of figure 6, we look at the use of credit card at debit cards either in terms of change in transaction or amount used at POS over time. The figure clearly shows that credit card experienced essentially no economically significant increase in its usage around the demonetization. A different picture instead holds for debit card, whose usage increased at high-rate during the months following the demonetization. In some way, the time-series for debit card transactions tracks well the aggregate finding

[16] Monthly data aggregated at the bank-level is made available by the Reserve Bank of India at https://rbi.org.in/scripts/atmview.aspx

in wallet payment. In both, we find a large increase in the first three months after, and a return to normal growth rate in the following months, without a significant reversal in usage. Therefore, also in this case, the temporary cash-contraction had effects on the long-term usage of debit card. However, it is important to point out that growth activity in debit card segment is smaller than what we have found for electronic payment. We obtain overall the same results using either the amount that is transacted or the number of transactions.

Next, we look at the extensive margin of the usage of traditional electronic payments method. In particular, we look at the decision to obtain new cards and the introduction of new point-of-sales (POS). These results are reported in the bottom panel of figure 6. In general, consistent with the lack of an increase in usage, credit cards do not experience a significant change in demand. At the same time, while debit card issuance increased significant post-demonetization, this effect does not appear in the data until a couple of months after the shock. This delayed response is consistent with the presence of large switching costs from cash to debit card. Lastly, we also document a sizable increase in the number of POS.

There are two main takeaways from the analysis of these other electronic payment systems. First, the shift into electronic payment systems extended beyond wallet payment technologies, and into to more traditional electronic payment methods. This result suggests that the dynamics we have uncovered with our digital wallet data are likely to represent economic forces that are also important outside this niche market. Furthermore, this is consistent with the work by Agarwal et al. (2018), which focuses exactly on this question and has better data than us on this dimension. Second, these results also seem to highlight the importance of switching costs between technologies in characterizing the response to the shock. In particular, credit cards, which arguably have higher switching costs than debit card or digital wallets, experienced essentially no change in activity around the shock. At the same time, while individuals with a debit cards were able to increasingly use this mean of payments to complete their transactions, our results show that debit cards did not represent a particularly useful method to replace payments for those that did not already have one at the time of the shock.

Summarizing, this section has shown that the demonetization shock had an effect on the adoption of electronic payment systems. The aggregate effect is easy to observe; it was large and persisted beyond the duration of the shock itself. Our cross-sectional evidence, which deposit growth as a metric of shock exposure, suggests that the effect was causal, and confirms that it persisted beyond the period of disruption in the money supply. We next turn to discussing and testing potential economic mechanisms for these findings.

4 Theory

In this section, we analyze an industry model of technology adoption. The model's purpose is to help us understand two keys features of the data: first, the fact a large, common shock led to a wave of adoption; second, the fact that adoption persisted even though the shock was transitory. The model posits an economic mechanism, positive externalities in adoption, which we show has testable implications. The next section will study these testable implications in the data.

4.1 Motivation

The model we study is a variant of the framework of Frankel and Pauzner (2000) and Burdzy et al. (2001). The central feature of this framework is that the value of adoption for any individual firm increases with the number of other firms that are currently using the technology; that is, the technology features positive externalities. These externalities lead to complementarities in adoption decisions; that is, a firm is more likely to adopt if more firms have already adopted. This, in turn, could in principle lead to equilibrium multiplicity: both high and low adoption may be equilibrium outcomes.

However, the framework we study also features aggregate uncertainty about the value of adoption for all firms. This aggregate uncertainty helps select a unique equilibrium. Firms have unique equilibrium strategies, in which adoption (or abandonment) are triggered by large realizations of the shock. Thus, the presence of externalities can create coordinated adoption dynamics in response to aggregate shocks, consistent with the qualitative observations for the setting we study described in the previous section.

Aside from the fact that they generate coordinated adoption dynamics, positive externalities are, plausibly, a key characteristic of electronic payment systems. Electronic payment systems are fundamentally a connected network of firms; as pointed out by, for instance, Katz and Shapiro (1985), positive externalities are a typical feature of network goods. As more firms join the electronic payment network, customers' value of using electronic payments increases; in turn, this further increases firms' incentive to join the network as well. Electronic payment systems could also be subject to positive externalities for reasons other than the network nature of the technology. For instance, it may be that initial adoption by a few firms lowers uncertainty about the quality of the technology, allowing other firms to further adopt the technology as well (Munshi, 2003). Our model does not take a precise stance on the cause of these positive externalities, and simply assumes that they exist; but we think they are likely to be a central economic feature of electronic payments technologies (and in fact, the estimates of the following section provide evidence of their importance.)

Finally, we note that other mechanisms than complementarities in adoption decisions could potentially

explain the observation that a large, transitory shock led to a large, permanent adoption wave. A simple example could be sunk adoption costs. If the electronic payments technology is in general more desirable than using cash, but its adoption is subject to sunk costs, the shock may trigger adoption, and firms may stay locked into the new technology even after the shock is undone. In the setting described above, adoption costs are likely to be very low. The technology we consider has virtually no sunk adoption costs for most retailers. Retailers can start accepting payment by simply downloading a free app and signing up. There are no initial membership fees, and there is no need to purchase special equipment; the payment system can be used virtually immediately. The only potential costs of adoption have to do with requirements to verify bank information for firms that want to accept more than Rs. 20,000 in monthly payments (approximately US$330 in 2016), which can take up to 72 hours. Importantly, after a recent Government policy, bank accounts were relatively widespread in India in 20016 (Agarwal et al., 2017). Thus, generally, we view the sunk cost mechanism as an unlikely explanation for adoption dynamics in this setting. Moreover, we will show that a number of features of adoption dynamics in the data are unlikely to be well-explained by a sunk cost model, in particular the dynamic dependence on past adoption rates.

4.2 Model

Time is discrete: $t = 0, \Delta, 2\Delta, ...$ The industry is a collection of infinitely-lived firms, indexed by $i \in [0, 1]$. They are risk-neutral and discount the future at rate e^{-r}. At different points in time, firm i must choose between operating under two technologies, which we index by $\{0, 1\}$. For firm i in period t, flow profits from operating under technology $x_{i,t} \in \{0, 1\}$ are given by:

$$\mathbf{\Pi}(M_t, X_t; x_{i,t}) = \begin{cases} M_t & \text{if} \quad x_{i,t} = 0 \\ \\ \underline{M} + CX_t & \text{if} \quad x_{i,t} = 1 \end{cases}, \tag{3}$$

where:

$$\underline{M} > 0, \quad C \geq 0, \tag{4}$$

and X_t is the fraction of firms in the industry that are also operating under technology 1, that is,

$$X_t = \int x_{i,t} di. \tag{5}$$

The parameter $\underline{M}$ captures the worst-case (flow) profits from operating the technology with complementarities, that is, the flow profits when no other firms is operating the technology. Technology 1 is the technology exhibiting externalities (and the adoption of which is therefore subject to complementarities.)

The parameter C controls the strength of complementarities; the maximum flow profit from the technology with complementarities is $\underline{M} + C$, and when $C = 0$, adoption decisions are not complementary.

Finally, M_t represents the money stock, or more generally the aggregate state affecting the *relative* value of adopting technology 1. We assume that it follows the discrete-time equivalent of a Brownian motion with drift μ and volatility σ, subject to Poisson jumps:

$$M_t = M_{t-\Delta} + \mu\Delta + \sqrt{\Delta}\sigma\left(\epsilon_t + J_t\right), \quad \epsilon_t \perp J_t, \text{ i.i.d.} \tag{6}$$

Here, $\epsilon_t \sim N(0,1)$, and J_t is assumed to have the three-point distribution:

$$J_t = \begin{cases} 0 & \text{w.p. } e^{-\lambda\Delta} \\ J & \text{w.p. } \frac{1}{2}\left(1 - e^{-\lambda\Delta}\right) \\ -J & \text{w.p. } \frac{1}{2}\left(1 - e^{-\lambda\Delta}\right) \end{cases} \tag{7}$$

where λ governs the arrival rate of jumps and J governs the size of jumps.

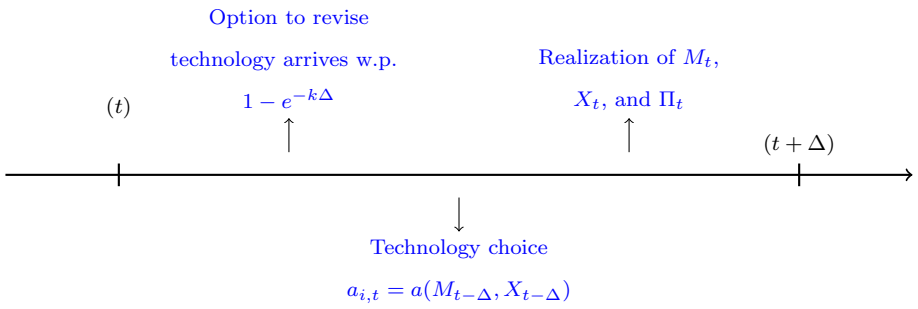


Figure 1: timing of actions and events during a period.

There are frictions to switching between technologies. Specifically, during each increment of time Δ, a mass $1 - e^{-k\Delta} \in [0,1]$ of firms is allowed to re-optimize their technology choice, where $k \geq 0$ is a parameter capturing switching frictions. When $k \to +\infty$, all firms reoptimize each period, while when $k = 0$, firms are permanently locked into their initial technology adoption choice.

The timing of actions within period t is depicted in figure above. Note that firms make their technology choice at the beginning of period t, before either the money stock M_t or the current fraction of adopters X_t are determined. Their information set at the moment of making the technology choice thus only contains the history of publicly observable signals $\mathcal{H}_{t-\Delta} = \{M_{t-n\Delta}, X_{t-n\Delta}\}_{n\geq 1}$ up to and including time $t - \Delta$.

Given that all firms that re-optimize their technology choice in a given period face the same problem going forward, we focus on symmetric equilibria. Moreover, we restrict attention to pure strategies that

only condition on public information. Denoting by $a_{i,t} = a_t \in \{0,1\}$ the technology chosen by firms that reoptimize during time t, the technology under which a particular firm operates then follows:

$$x_{i,t} = \begin{cases} x_{i,t-\Delta} & \text{w.p.} \quad e^{-k\Delta} \\ a_t & \text{w.p.} \quad 1 - e^{-k\Delta} \end{cases}. \tag{8}$$

We can use this to write the dynamics of the fraction of adopters as:

$$\begin{aligned} X_t &= X_{t-\Delta} + \left(1 - e^{-k\Delta}\right) \left\{ \underbrace{(1 - X_{t-\Delta})a_t}_{\text{firms opting into the technology}} - \underbrace{X_{t-\Delta}(1 - a_t)}_{\text{firms opting out of the technology}} \right\} \\ &= \left(1 - e^{-k\Delta}\right) a_t + e^{-k\Delta} X_{t-\Delta}. \end{aligned} \tag{9}$$

Finally, let $V_{a,t}$ be the value of operating under technology $a \in \{0,1\}$ at the beginning of period t. Define the net gain from adoption as:

$$W_t = V_{1,t} - V_{0,t}. \tag{10}$$

The technology choice in period t is then given by:

$$a_t = \mathbf{1}\Big\{W_t \geq 0\Big\}. \tag{11}$$

Given that both the exogenous shock $\{M_t\}_{t\geq 0}$ and the fraction of adopters $\{X_t\}_{t\geq 0}$ are both first-order Markov processes, the decision rules and value functions take the form of mappings from $\mathcal{R} \times [0,1]$ to the relevant space. An equilibrium of the model is thus an adoption rule $\mathbf{a}$ mapping $\mathcal{R} \times [0,1] \to \{0,1\}$ and value functions $\{\mathbf{V}_0, \mathbf{V}_1, \mathbf{W}\}$ mapping $\mathcal{R} \times [0,1] \to \mathcal{R}$, such that, for any realization of the shock $\{M_t\}_{t\geq 0}$ and any given the initial distribution of adopters, $X_{-\Delta}$, the adoption path generated through the law of motion given by (9) is consistent with the technology choice rule given by (11).

4.3 The effects of complementarities in adoption decisions

In order to explain how adoption dynamics can be affected by the presence of complementarities, we start by discussing the equilibrium of the model without complementarities.

4.3.1 The model without complementarities ($C = 0$)

In the absence of complementarities, the technology choice is entirely determined by the level of the aggregate shock, $M_{t-\Delta}$, regardless of the current adopter share $X_{t-\Delta}$. Appendix A shows that the expected relative gain from adopting the technology with complementarities is given by:

$$W_t = \frac{\Delta}{1 - e^{-(r+k)\Delta}} \left(\underline{M} - M_{t-\Delta} - \frac{\Delta}{1 - e^{-(r+k)\Delta}} \mu \right) + C \mathbb{E}_{t-\Delta} \left[\sum_{n \geq 0} e^{-(r+k)\Delta n} X_{t+\Delta n} \Delta \right], \tag{12}$$

where the first term captures the *net* benefits to adopting the technology with complementarities even if there are no other adopters (so that flow benefits to the technology with complementarities are given by $\underline{M}$), whereas the second term captures the net benefits due to complementarities and potential adoption by other firms. When $C = 0$, this second term is equal to zero, and the value of adoption becomes:

$$W_t = \mathbf{W}(M_{t-\Delta}) = \frac{\Delta}{1 - e^{-(r+k)\Delta}} \left(\underline{M} - M_{t-\Delta} - \frac{\Delta}{1 - e^{-(r+k)\Delta}} \mu \right), \tag{13}$$

so that firms adopt technology 1, if and only if:

$$M_{t-\Delta} \leq \mathbf{M} = \underline{M} - \frac{\Delta}{1 - e^{-(r+k)\Delta}} \mu, \tag{14}$$

regardless of the current adopter share $X_{t-\Delta}$. This implies that the model cannot generate history-dependence, in the sense that the dynamics of adoption only depend on the current value of the shock, and not on the past history of shocks, as summarized by the value of the current adopter share, $X_{t-\Delta}$.

Figure 7 illustrates the implications of this lack of history dependence. The top panel of figure 7 reports a particular sample path for the aggregate shock M_t. The diffusion governing M_t is assumed to have no drift ($\mu = 0$); the sample path plotted is chosen to so that two symmetric jumps occur at the beginning of months 3 and 6.[17] The bottom panel of figure 7 shows that the negative jump triggers a brief period of adoption. But once it is undone, the share of adopters reverts progressively back to 0.

The top panel of figure 8 shows this trajectory in the state-space $(M_{t-\Delta}, X_{t-\Delta})$. The economy starts from point A, where $M_{-\Delta} = 0.4$ and the initial adopter share is $X_{-\Delta} = 0$. After staying in the no-adoption region for a few weeks, the downward jump moves the economy to the adoption region, to point B. While the money stock is low, the economy remains in the adoption region and the adopter share grows, out to point C. After the money stock M_t recovers, the economy jumps from point C to point D, in the no adoption region. At that stage, the adopter share declines gradually, to point E at the end of the two years of simulation. In

[17] The time period chosen in these simulations, Δ, is very short (approximately a day, given the choice of discount rate), to conform with the continuous-time limit studied below in the case of complementarities.

this figure, the lack of history dependence can be visualized easily: the line $\mathbf{M}$ separating the adoption and the no adoption region in the top panel of figure 8 is flat. This means that adoption dynamics (the sign of the change in $X_{t-\Delta}$) are determined entirely by the current level of $M_{t-\Delta}$, and are in particular independent of the level of current adoption, $X_{t-\Delta}$.

4.3.2 The model with complementarities ($C > 0$)

In the model with complementarities ($C > 0$), the adoption decision depends on how firms that are currently re-optimizing their technology choice anticipate the fraction of adopters to evolve in the future. This is visible in the formula for the value of adoption given above, in equation 12, which depends positively on:

$$\mathbb{E}_{t-\Delta}\left[\sum_{n\geq 0} e^{-(r+k)\Delta n} X_{t+\Delta n}\Delta\right] \tag{15}$$

when $C > 0$. In principle, this could lead to a situation where multiple equilibria are possible: expectations of high adoption rates may lead re-optimizing firms to chose adoption, thus validating the expectation, and vice-versa. However, in the continuous-time limit, we have the following result.

Lemma 1 (Frankel and Pauzner, 2000)**.** *In the continuous-time limit $\Delta \to 0$, so long as $\sigma > 0$, there is unique equilibrium, characterized by a threshold $\mathbf{M}(X_{t-\Delta})$ such that firms re-optimizing their technology choice adopt the technology with complementarities, if and only if:*

$$M_{t-\Delta} \leq \mathbf{M}(X_{t-\Delta}).$$

Moreover, the frontier $\mathbf{M}(.)$ is upward-sloping in $X_{t-\Delta}$.

Proof. This is a variant of Theorem 1 in Frankel and Pauzner (2000); see the appendix for a derivation specific to our setting. □

This result is obtained through iterated deletion of strictly dominated strategies. In this process, the existence of aggregate shocks ($\sigma > 0$) is crucial. The shocks can move the economy to strict dominance regions, which then allows the progressive elimination of strategy profiles, as explained by Frankel and Pauzner (2000).[18] In simulations of the model, we also obtain the adoption frontier by applying iterated

[18]The intuition for the proof can be roughly described as follows. For any value of the current adopter share, $X_{t-\Delta}$, there always exist sufficiently low values of the stock of money, $M_{t-\Delta}$, such that adoption is optimal even if a firm anticipates that no one else will adopt in the future. Denote the highest such value of $M_{t-\Delta}$ by $\underline{\mathbf{M}}^{(0)}(X_{t-\Delta})$; in any equilibrium, firms will adopt the technology at least when $M_{t-\Delta} \leq \underline{\mathbf{M}}^{(0)}(X_{t-\Delta})$. Next, consider the following strategy: adopt the technology, *if and only if*, $M_{t-\Delta} \leq \underline{\mathbf{M}}^{(0)}(X_{t-\Delta})$. Assume that reoptimizing firms expects others to follow this strategy. Then the optimal strategy of re-optimizing firms will be to chose adoption for some values of $M_{t-\Delta}$ that are strictly higher than (but close to) $\underline{\mathbf{M}}^{(0)}(X_{t-\Delta})$. This is because, so long as there are shocks, if $M_{t-\Delta}$ is close enough to $\underline{\mathbf{M}}^{(0)}(X_{t-\Delta})$, the probability that a

deletion of strictly dominated strategies; we describe the numerical procedure in appendix.

The adoption dynamics implied by this lemma are illustrated in the bottom panel of figure 8. This panel plots the path of the economy under the same mean-reverting shock to the aggregate money stock as in the previous section. The initial adoption dynamics are similar to the model without complementarities: the economy jumps from point A to point B when the negative shock to M_t occurs, and adoption starts. However, when the negative shock is undone, the economy moves to point D, which is now below the adoption frontier, given that the adopter share has risen. Thus, even after the recovery in the money stock, the economy remains in the adoption region. Absent further upward jumps, the economy will converge to full adoption, as illustrated in the bottom right panel of figure 9.

The model with complementarities thus displays a form of history dependence. Adoption dynamics depend not only on the current and expected path of the money stock, but also on the current adopter share, and this creates the possibility for transitory shocks to have permanent effects on the adopter share. The bottom panel figure 9 indicates that this result is due to the fact that the frontier $\mathbf{M}(.)$ is upward-sloping — it is the main difference with case of no complementarities. This is because the discounted present value of the future adopter share (expression 15) is strictly increasing in the initial share $X_{t-\Delta}$. Intuitively, the fact that there are frictions to adoption ($k < +\infty$) means that the adopter share displays some persistence. Thus, when it is relatively high, firms re-optimizing their adoption decision can expect it stay high for at least small periods of time. The higher expected adopter share in the short-run raises the incentive to adopt; thus the aggregate money stock must be higher in order to dissuade firms from adopting.[19]

4.4 Testable predictions

As discussed above, the key distinction between the versions of the model with and without complementarities is the fact that, when there are complementarities, the frontier $\mathbf{M}(X_{t-\Delta})$ is an upward-sloping function of the current fraction of adopters, $X_{t-\Delta}$. This feature of the model with complementarities generates history-dependence in the response of the model to shocks. In what follows, we describe some simple diagnostic regressions which capture this history-dependence, and hence distinguish between data generated from a model without complementarities ($C = 0$) and data generated from a model with complementarities ($C > 0$).

negative shock will push it back below $\underline{\mathbf{M}}^{(0)}(X_{t-\Delta})$ (where adoption occurs with probability one) is non-zero, which makes current adoption attractive. Thus, there is another frontier $\underline{\mathbf{M}}^{(1)}(X_{t-\Delta}) > \underline{\mathbf{M}}^{(0)}(X_{t-\Delta})$ such that adoption is optimal, if and only if, $M_t \leq \underline{\mathbf{M}}^{(1)}(X_{t-\Delta})$. This argument can be iterated forward, and leads to a limiting frontier $\underline{\mathbf{M}}(X_{t-\Delta})$. A similar limiting frontier can be obtained by starting the argument from the region where adoption is never optimal; denote by $\overline{\mathbf{M}}(X_{t-\Delta})$ this other frontier. The two frontiers must coincide. Otherwise there would be pairs $(\underline{M}_{t-\Delta}, X_{t-\Delta})$ and $(\overline{M}_{t-\Delta}, X_{t-\Delta})$ such that $\underline{\mathbf{M}}(X_{t-\Delta}) < \underline{M}_{t-\Delta} < \overline{M}_{t-\Delta} < \overline{\mathbf{M}}(X_{t-\Delta})$ such that adoption would be optimal at $(\overline{M}_{t-\Delta}, X_{t-\Delta})$ but not at $(\underline{M}_{t-\Delta}, X_{t-\Delta})$, which contradicts the fact that adoption is more valuable for lower values of $M_{t-\Delta}$. The resulting frontier of lemma 1 is then given by $\mathbf{M}(.) = \underline{\mathbf{M}}(.) = \overline{\mathbf{M}}(.)$.

[19] The appendix reports an explicit derivation of the adopter share as a function of future adoption decisions and the current adopter share, $X_{t-\Delta}$. This derivation highlights that, in the continuous-time limit, the frontier can only be independent of $X_{t-\Delta}$ if there are no frictions, that is, if $k \to +\infty$.

In order to make our approach more concrete, we test it on simulated data. We interpret the model as representing a single district, indexed by d. We assume that, in each district, the path of the money stock $M_{d,t}$ is given by:

$$M_{d,t} \quad = M_{d,t-\Delta} + \mu\Delta + \sqrt{\Delta}\sigma\left(\epsilon_t + \eta_d J_t\right), \tag{16}$$

where ϵ_t and J_t are distributed as before, $\eta_d \sim U([0,1])$, and:

$$\eta_d \perp \epsilon_t \perp J_t, \text{ i.i.d.} \tag{17}$$

All districts are exposed to the same fundamental fluctuations to the the money supply (ϵ_t, J_t); but their exposure to the aggregate jump risk in the money supply, J_t, is heterogeneous, as captured by η_d. We then simulate data from $N_d = 1000$ districts, each containing $N_f = 100$ firms, for a period of two years.[20] Each simulation corresponds to particular sample paths for $\{\epsilon_t\}_{t\geq 0}$, as well as a particular draws for district-specific initial conditions $\{X_{0,d}\}_{d\in[1,\ldots,N_d]}$ and district exposure to the shock, $\{\eta_d\}_{d\in[1,\ldots,N_d]}$. Across simulations, we assume that the initial condition is drawn from a uniform distribution on the set $[0, 0.05]$, so as to fix adoption to low overall levels prior to the shock. Across simulations, we also fix the sequence of realizations of J_t; specifically, we assume that there is downward shock at the beginning of month 10, followed by an upward jump at the beginning of month 12. This meant to mimic the unexpected and transitory decline in the money stock discussed in the previous section.

4.4.1 District-level predictions: effects of shocks to M_t

As a preliminary step, we verify that, in this model, the difference-in-differences analysis of the previous section indeed captures the causal effect of the shock on adoption. Specifically, we estimate the following regression:

$$X_{d,t} = \alpha + \beta \mathbf{1}_{\{t\geq t_0\}} + \gamma \mathbf{1}_{\{d\in\mathcal{T}\}} + \delta\left(\mathbf{1}_{\{d\in\mathcal{T}\}} \times \mathbf{1}_{\{t\geq t_0\}}\right) + \epsilon_{d,t}, \tag{18}$$

where $X_{d,t}$ is the average of the variable $x_{i,d,t}$ (the technology status of firm i in district d at time t) across firms in district d. In this regression, t_0 is the period starting with the date of the shock (that is, month 10 and afterward in the simulations), and $\mathcal{T}$ is the set of "treated" districts. In the results below, we classify a district as treated if the change in $M_{d,t}$ which it experiences between months 9 and 10 is above the media change in $M_{d,t}$ across all districts between months 9 and 10. Estimates of the coefficient δ are reported in table 4. Both in the model with and without complementarities, the difference-in-differences estimator is positive and significant; furthermore, its magnitude is approximately identical across models. Similarly, we

[20] The underlying model is simulated at the daily frequency, and sampled at either the weekly or monthly frequency.

can estimate dynamic treatment effects on the simulated data, as follows:

$$X_{d,t} = \alpha_t + \delta_t \mathbf{1}_{\{d \in \mathcal{T}\}} + \epsilon_{d,t}. \tag{19}$$

The estimates are reported in figure 9. The dynamic treatment effects estimated on simulated data are positive at all horizons after the shock. They differ qualitatively between the model with and without complementarities; in particular, treatment effects are increasing over time in the model with complementarities, consistent with the intuitions developed above on the persistent effects of transitory shocks when there are complementarities. However, at intermediate horizons, these effects cannot be distinguished statistically from each other. More generally, these estimation results indicate that the mere observation that the shock has positive effects on subsequent adoption does not suffice to distinguish the model with complementarities from the model without complementarities.

4.4.2 District-level predictions: dependence on initial conditions

We next test the more fundamental feature of the model with complementarities, namely the fact that initial conditions help determine the path of subsequent adoption. In order to test this prediction, we estimate a regression analogous to the difference-in-differences estimator of the effect of the shock, but defining the treatment to be the initial share of adopters, $X_{0,d}$. Specifically, we estimate again:

$$X_{d,t} = \alpha + \beta \mathbf{1}_{\{t \geq t_0\}} + \gamma \mathbf{1}_{\left\{X_{0,d} \geq \hat{X}_0\right\}} + \delta \left(\mathbf{1}_{\left\{X_{0,d} \geq \hat{X}_0\right\}} \times \mathbf{1}_{\{t \geq t_0\}} \right) + \epsilon_{d,t}, \tag{20}$$

on the same set of simulations, where, for $\hat{X}_0$, we use the median adoption rate in period 0. The results are reported in table 5. They contrast from the results of table 4 in that the difference-in-differences estimator is now close to zero (though statistically significant) in the case of no complementarities, while it is significantly positive in the case of complementarities. The null that the two coefficients are equal can now be rejected at the 5% (and in fact also at the 1%) level.

In the model with complementarities, dynamics also depend sharply on initial adoption rates, whereas they do not in the model without complementarities. To illustrate this, we estimate the following regression on simulated data:

$$X_{d,t} = \alpha_t + \delta_t \mathbf{1}_{\left\{X_{0,d} \geq \hat{X}_0\right\}} + \epsilon_{d,t}. \tag{21}$$

The estimates of the coefficients δ_t are reported in figure 10. In data simulated from the model without complementarities, the effects are small (and in fact negative, due to mean-reversion) at all horizons. In data simulated from the model with complementarities, by contrast, they are positive and large at all horizons

after the shock. Thus, estimates of a positive dependence of subsequent adoption rates on initial adoption levels in the data is likely to be a sign of the presence of adoption complementarities, in a way that can be quantitatively tested using specifications 20 and 21.

4.4.3 Firm-level predictions

At an even more fundamental level, the key distinction between the model with and without complementarities is that the schedule $\mathbf{M}$ is flat as a function of M_t when there are no complementarities, but upward sloping when there are complementarities; in terms of model notation, the null hypothesis to be tested is:

$$\frac{\partial a_t}{\partial X_{t-\Delta}}(X_{t-\Delta}, M_{t-\Delta}) = 0, \tag{22}$$

where $a_t = \mathbf{a}(M_{t-\Delta}, X_{t-\Delta})$ denotes the adoption decision of firms that are allowed to re-optimize their technology choice in period t.

A difficulty of this prediction is that we do not directly observe the adoption decision of firms in the data; we only observe their current technology choice, $x_{i,d,t}$. (Specifically, we observe whether they have access to the payments technology during a particular week or not; we do not know whether a particular firm receive the possibility to revise their adoption choice.) Within the context of the model, these two objects are related through:

$$x_{i,d,t} = \begin{cases} x_{i,d,t-\Delta} & \text{w.p.} \quad e^{-k\Delta} \\ a_{d,t} & \text{w.p.} \quad 1 - e^{-k\Delta} \end{cases}. \tag{23}$$

This suggests estimating the following firm-level panel regression:

$$x_{i,d,t} = \alpha + \rho x_{i,d,t-\Delta} + \beta M_{d,t-\Delta} + \gamma X_{d,t-\Delta} + \epsilon_{d,t} \tag{24}$$

on simulated data. Table 6 reports coefficient estimates for α, ρ and γ using the same simulated data as above. The coefficients ρ and β are, respectively, significantly positive and significantly negative in both versions of the model. Moreover, the null that they are equal across models cannot be rejected. However, the coefficient γ is positive and large in the case of the model with complementarities, while it is small (and negative, again likely due to mean-reversion) in the model without complementarities, and the null of their equality across models can be rejected at the 5% level. To the extent that movements in $x_{i,d,t}$ are driven by changes in a_t, this result is consistent with the fact that in the model with complementarities, $\frac{\partial a_t}{\partial X_{t-\Delta}}(X_{t-\Delta}, M_{t-\Delta}) = 0$, while in the model with complementarities, $\frac{\partial a_t}{\partial X_{t-\Delta}}(X_{t-\Delta}, M_{t-\Delta}) > 0$. A finding of a positive and significant coefficient on the lagged share of adopters in empirical estimates of equation 24

is thus likely to provide a rejection of the null of no complementarities.

Finally, we can estimate a month-by-month equivalent of the specification above, in order to test whether the dependence on adopter share in the district is stronger during the period of the shock. The regression we estimate is:

$$x_{i,d,t} = \alpha_t + \rho_t x_{i,d,t-\Delta} + \beta_t M_{d,t-\Delta} + \gamma_t X_{d,t-\Delta} + \epsilon_{d,t}. \tag{25}$$

Figure 11 reports the estimates of the coefficients γ_t, normalized to be zero in month 9 (the month preceding the shock), for the model with and without complementarities. The figure shows that the positive estimates for the coefficient γ in the case of complementarities come from the period after the shock. This is because prior to the shock, districts are all in the no-adoption region, as so adoption is simply declining over time; that is, $(M_{d,t}, X_{d,t})$ are in a region where a is flat with respect to both variables. The slope of a with respect to $X_{d,t}$ can only be detected after the shock, when the adoption decision changes for some districts (but not all).

Generally, results from specifications 24 and 25 indicate that the presence of a positive and significant dynamic dependence between firm-level adoption decisions, $x_{i,d,t}$, and district-level adoption shares, $X_{d,t-1}$, *after* controlling for the shock to the money stock, $M_{d,t}$, can help discriminate between the model with and without complementarities.

This section has shown that a model that embeds one of the plausibly key features of electronic payment systems, complementarities in adoption decisions, can generate some of the qualitative patterns documented in the previous section: specifically, a large aggregate shock affecting the relative payoff between technologies can lead to adoption waves; the adoption waves can persist even if the shock is transitory. Moreover, we used to model to highlight specific empirical tests of the presence of complementarities. The following section applies these tests to the data.

5 Understanding the dynamics of adoption

In this section, we come back to the data on electronic payment in order to test some of the key predictions of the model. As discussed in the previous section, the mere increase in adoption could be explained also by a simple model with fixed costs. More evidence should be provided to support the claim that complementarities played a crucial role in adoption during the demonetization.

In order to test for complementarities, we follow the discussion in the theory section and the results obtained with simulated data. In particular, we first will provide evidence that is consistent with the path dependence in the data. Then, we will also show strong evidence of the importance of externality across

firms in the process of adoption.

5.1 History-dependence in district-level data

As highlighted in the previous section, in a context where complementarities are important, we expect to find that the initial level of adoption in a region will positively influence the increase in future adoption. The intuition is that the presence of complementarities increases the marginal benefit of adopting for firms operating in an area with a larger initial share of adopters. As a result, the higher marginal benefit makes these firms more reactive to a shock.

We start by using the specification highlighted in the previous section as a district-level test of the dependence on initial conditions, equation 18. Specifically, we estimate:

$$X_{d,t} = \alpha_t + \alpha_d + \delta \left(I_d \times \mathbf{1}_{\{t \geq t_0\}} \right) + \Gamma_t' Y_d + \epsilon_{d,t}, \tag{26}$$

This specification compares districts with high vs. low adoption rate pre-shock, where the adoption rate pre-shock is measured by I_d. The results are reported in Appendix Table B.2 . We find that indeed, regions with high initial adoption experienced significantly larger adoption ex-post. From the standpoint of the model section 4, this result can be interpreted as a rejection of the null that complementarities did not influence the adoption process (i.e. that $C = 0$).

However, when applied to the data, specification 18 may be problematic. First, the model will always suffer of endogeneity, since we are effectively using some function of the outcome variable in the pre-period to define the treatment variable. Even controlling for other covariates is not solving this issue. Second, the scope of the externality may extent beyond the district. In particular, if externalities in payment adoption are due to a shared customer base, and if this customer base is mobile across districts, then the relevant measure of existing penetration of electronic wallet may not be the penetration in a single district, but the penetration in the region surrounding the district.

Given these concerns, we next test how the increase in adoption differs depending the distance between the district and areas in which the usage of electronic wallets was large *prior* to November. This alternative specification can be thought of as another test of the dependence on initial conditions, with a definition of initial conditions that is both less restrictive, and more plausibly uncorrelated with other subsequent determinants of technology adoption.

Specifically, we run a simple difference-in-difference model analogous to 26 where we compare the usage of wallet technologies around the demonetization period across districts that are differentially close to a major digital wallet districts. In particular, we define a district to be an electronic payment hub if there were more

than 1,000 of active firms pre-demonetization, and then we measure distance to the hub as the minimum of the distance between the district and each hub. Results are essentially identical if we use a threshold of 500 merchants to define the hub districts. It is important to point out that we exclude from this analysis the districts defined as major digit wallet hubs. The main coefficient of interest in the one on the interaction between a post-dummy and the measure of distance at district level. [21]

There are two main concerns with this model. First, the model may be capturing variation across geographical areas in the country. To get around this issue, we saturate the model with a wide set of state by month fixed-effects, which essentially allows us to focus on variation across district but within state in distance to the major hub. The idea is that two districts may belong to relatively homogeneous areas, but one of two because it is positioned in a specific size of a state may be particularly close to a major hub of electronic payment adoption and therefore it may be particularly influenced by its initial level of adoption. Notice that we will show that results also hold without this set of fixed-effects.

The second concern is that even within a state distance from the large electronic payment hubs may simply capture differences between urban and rural areas. To rule out this hypothesis, we also control in every specification from the distance to the capital district of the state. Like in many countries (apart the US), the state capital is generally in the most urban and populous city of a state. Therefore, if our estimator is only capturing the distance from a city, then this covariate should absorb any effect. As usual, this control is interacted with time effect. We also show the results including the same district-level covariates used in previous analyses.

In other words, this specification is a difference-in-difference model of the following form:

$$X_{d,s,t} = \alpha_{st} + \alpha_d + \delta \left(D_d \times \mathbf{1}_{\{t \geq t_0\}} \right) + \gamma \left(\tilde{D}_{d,s} \times \mathbf{1}_{\{t \geq t_0\}} \right) + \Gamma_t' Y_d + \epsilon_{d,t}, \tag{27}$$

where t indicates time, defined at monthly level in this analysis, d indexes the district and s identifies the state of the district. D_d is district's distance to the nearest electronic wallet hub and $\tilde{D}_{d,s}$ is district's distance to the capital district of the state. The equation is estimated with standard errors clustered at district level, which is the level of the treatment (Bertrand et al., 2004). It is important to point out that we exclude from this analysis the districts defined as major digit wallet hubs.[22] Furthermore, we use the data between May 2016 and February 2017. The main coefficient of interest is δ - which provides the difference in level of adoption pre and post demonetization depending on how far the district is from its closest electronic

[21] There are five electronic payment hubs with more than 1,000 active firms pre-demonetization and are spread evenly across the country: Delhi (North), Kolkata (East), Mumbai (West), Bangalore and Rangareddy (South). These five districts are also among top 6 metropolitan cities by population in India, suggesting that the penetration of electronic payment was limited to main urban centers in India. We identify these hubs based on their penetration levels in the month of September 2016.

[22] Notice that this exclusion does not affect our results, if anything results are stronger.

wallet hubs. Later, we will also examine the dynamic of the effect.

These results are reported in Table 7. In the columns (1) and (4), we report the baseline regression where we only control for the distance to the capital district. Across both outcomes - amount and number of transactions - we find that the districts farther away from major hubs experienced lower increase in transaction in the post-demonetization period. The same result holds in the following columns - (2) and (5) - where we augment the specification with state by month fixed-effects. Overall, the addition does not change the qualitative results, as we keep finding that further away districts experienced lower growth. The same happens when, in columns (3) and (6), we add the covariates control interacted post-dummy. If we take the most conservative of the estimates, we find that 100 km increase in distance translates into a 18% lower increase in transactions.

These results are robust in two ways. First, they do not depend on the definition of a major hub as a location with more than 1,000 active firms before the demonetization. For instance, in an unreported Table, we find that the results are essentially identical if we use a threshold of 500 firms. This result is not surprising, since this alternative definition leads to almost the same set of electronic wallet hubs. Second, the results do not depend on the linearity of the treatment. Specifically, we change the specification 27 by replacing the distance to nearest hub with dummy for distance to hub based on three thresholds : 200 km, 300 km and 400 km. Table B.3 of Appendix shows that are robust to alternative definitions of distance to major hubs.

Lastly, Figure 12 allows us to explore the dynamic of the effect. We construct this Figure by replicating the same analysis as before but estimating a month-specific effect for every month around the demonetization, normalizing the month of October to zero.[23] This Figure reveals two interesting things. First, the main result does not simply reflect differential trends in adoption between districts that are differentially distant from the hubs. This is an important result, because it provides crucial evidence to rule out that our main result was picking up a secular trend in adoption between major hub cities and their neighboring areas. In fact, we find essentially no relative difference in adoption before October.

Second, the Figure reveals that distance played an important role during the most cumbersome months of the currency swap. In fact, we detect some effect already in November, but then most of the relative difference really arises in December and January. As cash comes back to more normal levels and adoption rates become high across the whole country, this effect of distance seems to attenuate.

Overall, this set of results helps shed light on the mechanism behind the technology adoption wave in the aftermath of the demonetization. We show that adoption was stronger in areas that had initially high

[23] In other words, we just estimate the previous equation but where we interact each covariate with month dummies. Furthermore, we also extent the post-period to June 2017. Standard errors are always clustered at district-level.

adoption rates; moreover, the increase was stronger in areas that were relative geographically close to the major hubs for electronic wallet before the shock. Both of these results are consistent with complementarities being important determinants of adoption dynamics.

5.2 Complementarities at the firm level

The theory has also highlighted another feature of the adoption process in the presence of complementarities. In particular, we have shown that firm-level adoption is affected by the previous decisions of companies that are located close to the firm.

Using our data, we can estimate an empirical counterpart of the equation (24) generated by the model. The idea of this model is that we try to explain the size of adoption of a firm with the extent to which other companies in the same were using electronic payments in the previous period. As we have shown before, the presence of complementarities will generate externalities between firms that share the same market. Overall, the approach used here is similar to Munshi (2004), who used a similar empirical model to examine the role of social learning in agriculture.

One of the limitation of the equation dictated by the model is that it requires the explicit knowledge of the cash shock process. While we could use some proxy of the shock process at district level (for instance, using deposit growth previously discussed), then our results would crucially depend on the assumption that the shock process is correctly specified.

To get around this issue, we leverage on the granularity of our data and measure the set of firms that may have strong complementarities with a firm as the set of firms that are co-located in the same area and that belong to the same business area. The idea is that firms in the same location and industry will likely share the same set of customers, which is an important dimension of externality in this market. Furthermore, these firms are more likely to benefit from each other experience in the adoption. Measuring the reference group across both location and industry will allow us to introduce a full set of industry by time and location by time fixed effects. Therefore, rather than having to specify the dynamic of the shock by district, this approach allows us to difference this part out. As we will discuss later, this specification also allow us to solve other interpretation concerns as well.

To be more precise, our baseline model is:

$$x_{i,p,k,t} = \alpha_i + \alpha_{p,t} + \alpha_{k,t} + \rho x_{i,p,k,t-1} + \gamma X_{p,k,t-1} + \epsilon_{i,p,k,t}. \tag{28}$$

Here $x_{i,p,k,t}$ is a measure of technology choice by firm i in industry k and pincode p at time t (where t is a week), for instance a dummy for whether the firm used the platform or its amount of activity in the

platform.[24] At the same time, $X_{p,k,t-1}$ is a measure of size of adoption by other firms in the same pincode and the same industry during the previous week. To be consistent, we measure $X_{p,k,t-1}$ using the same variable as used as outcome, summing that dimension across all firms in the same pincode and industry, always excluding the firm itself. To ease the interpretation of the coefficients, apart from when the outcome is a dummy, we log-transform all the relevant variables.[25]

The model is estimated using weekly data from our electronic wallet company, with firms identified at the pincode level. The sample represents a balance panel of all the firms that used the product between May 2016 and June 2017 and that has the pincode information.[26] We conservatively estimate our standard errors clustering them by pincode, which allows firm errors to be correlated both across time and across space within the same location.

Before we discuss the limitations and extensions to this approach, we present the baseline estimates in table 8. To start, we examine the effect on the amount of transaction at firm level. In the first column, we report the simple regression where we explain the amount of transaction in the electronic wallet platform at the firm level as a function of the amount of transaction done by other firms and controlling to the firm behavior in the previous week. This specification does not include any fixed effects. We find that a higher level of transaction activity by other firms in the same location and business area strongly predicts an increase in the size of transactions that happen at the firm itself. This effect is not only statistically significant but also quantitatively large. In particular, a one-standard-deviation in the size of adoption by neighboring firms leads to a 40% increase in the amount of business completed in wallet technology by the firm, which corresponds to an 18% of the standard-deviation of the outcome variable. In the next column, we add a merchant fixed-effect. This addition allows us to focus only on within-firm variation in the analysis. The results remain highly significant and the coefficient increases in size.

Overall, we find similar results when look at two other outcomes. First, we examine the same analysis looking at the number of transactions, rather than the amount that is transacted. Across the different specifications, the magnitude and the significance are similar to the one discussed with amount. Second, we also look at the probability to be active in the platform. This outcome allows us to focus on the extensive margin of the decision to use the technology. Also in this case, our results are large and highly significant. This outcome in particular is the closest to the one that we use in the theory section.

[24] We define industries in the 14 broad categories: Food and Groceries(14%), Clothing(10%), Cosmetics(2%), Appliances(8%), Restaurants(12%), Recreation(2%), Bills and Rent(1%), Transportation(13%), Communication(12%), Education(3%), Health(7%), Services(4%), Jewellery(1%) and Others(11%).

[25] In other words, we transform each variable to be equal to the log-plus one of the primitive.

[26] We use pincode to identify firms because we want to use the narrowest definition of location that is available in the data. Unfortunately, the actual pincode is only available for a subset of firms. Later in the section, we also present a robustness using district to define location a measure that is broader geographically but available for every firm and we show results do not change.

In this analysis, the main identification concern is that past-decisions by neighboring firms may proxy for unobservable heterogeneity across companies that may affect adoption for reasons that are unrelated with complementarities. Since we have a firm fixed-effect in the regression, this heterogeneity must be time varying in order to explain our results. In particular, the most important concern is that our measure may capture differences in the relative costs of adoption over time. For this technology, direct costs of adopting the technology are essentially constant across firms, but different firms may differ in indirect adoption costs, which may be related to the business area or the industry.

To deal with this concern, we augment our baseline specification with a detailed set of pincode-by-week and industry-by-week fixed effects. The idea behind this approach is that most of differences in adoption costs should arise either because of differences in the type of business or because of the local market in which the firm operates. Therefore, the addition of these dynamic fixed-effects effectively remove all the variation coming from the heterogeneity across industries and local market. Furthermore, this is the closest specification to dynamic specification 24 proposed in the previous section, with the pincode-by-week fixed effects taking the place of local measure of the shock.

The new results are presented in Table 8. In column (3), we start adding the pincode-by-time fixed-effects. Here, we are effectively comparing two firms within the same pincode, where the heterogeneity in exposure to the technology comes from the fact that the two firms are operating in different business areas. Across the three main outcomes, we continue to find that the effect is positive and highly significant, but the magnitude decreases from the previous estimates. In column (4), we move one-step forward and add a detailed set of industry-by-week fixed-effects. Relative to the previous set-up, in this specification we not only compare across firms within the same pincode, but we also adjust the estimates for changes adoption rates in the same industry. Our estimates remain positive and highly significant. This is true across all the outcomes used.

Until now, our analyses have used the whole panel for firms, using data from both before, during, and after the demonetization. However, one may expect the role of complementarities to be different in the three periods. In particular, complementarities could play a more important role when the economy experiences a large increase in the use of electronic payments. In fact, this is what the theoretical model was predicting. To shed light on this question, we repeat the same analyses as before, but rather than estimating one single parameter for the effect of externality in adoption, we estimate a month-specific parameter for each of our outcomes. We complete these analyses using the most saturated model, containing all the fixed-effects discussed before.

The results are presented in the three panels of figure 13. Each figure plots the main coefficient of interest the coefficient on the adoption on neighboring firms for each month in the data and across the

three outcomes. From these analyses of the dynamic of the effects, we can draw two main conclusions. First, the positive effect of externality is not specific to any period. Indeed, in every month, the estimated coefficient is positive and — apart from a few cases — it is always highly statistically significant. Second, we find that the effect of externality is much higher in the months of the demonetization relative to both before and after the shock. In particular, in the figure we also plot two vertical lines that separate the months of November, December, and January from the others. Across the different outcomes, we always find that the effects were much larger during the months in which the cash-contraction was more acute. This result highlights the importance of externality during the months of demonetization.

Before we conclude this section, we want to highlight a couple of relevant robustness tests. As we have mentioned before, we have run these tests using the sub-sample of firms that were geo-localized in the data with pincode information because we wanted to run our analyses with the tightest definition of neighboring firms. However, in Table B.4 in the Appendix we show that results also would have been qualitatively the same if we had defined location based on district, which is a location that is instead available for the whole sample.

Furthermore, in defining the relevant market in the previous analyses we have used other stores that are in the same location and industry. The idea behind these choices is that firms in the same market will likely share the same clientele and therefore they are a perfect target to measure the role of externality. However, another approach is to simply use the level of adoption at pincode level across every industry as the relevant reference for measuring potential externalities. One limitation of this approach is that clearly we cannot control for time-varying pincode effect in this specification, since this effect would be almost collinear with the main variable. In Table B.5 in the Appendix, we show that the results using this alternative approach are indeed similar.

Overall, these results replicate closely the implications of the model with complementarities in technology adoption, as we have shown in the previous section. Like in the model, the business decision to adopt the technology and its utilization is highly affected by the decision of other firms in the area. Furthermore, we find that this effect is particularly strong during the shock period, which is exactly the period in which this transition dynamic is more important. Overall, these results support the importance of ies in explaining the growth in electronic technologies.

6 Other real effects of the shock

This section examines the effects of demonetization on household consumption. Consistent with the results in Section 3 on the rise in use of electronic payment technologies, we show two set of results. First,

the contraction in consumption is mostly limited to the months characterized by the highest reduction in cash. Moreover, effects are concentrated in the unnecessary and infrequent consumption - categories which households could substitute across time which more flexibility. Second, we show that indeed households that were more likely to have access to electronic payments were essentially unresponsive. Taken together, these finding suggests that households were able to limit the costs of the policy change by shifting a substantial part of the transactions on electronic platforms.

6.1 Empirical setting

In this section, we examine how household consumption responded differentially to the the cash swap by comparing consumption responses across districts based on their exposure to the cash shortage. Specifically, we utilize the same identification strategy used in Section 3 to analyze impact of the shock on the adoption of electronic payment technologies - we define a district's exposure to the shock based on the excess deposit growth during the quarter of demonetization for that district.[27]

To examine the changes in consumption behavior by Indian households, we use data from the Consumer Pyramids database maintained by Center for Monitoring Indian Economy (CMIE). This data set has two crucial advantages relative to the widely used National Sample Survey (NSS), which is a consumption survey conducted by the central government agencies. First, the NSS is not available for the period of interest, as it was ran for the last time in 2011. Second, the NSS is a repeated cross-section of households, while CMIE data is a panel data set, covering a set of about 160,000 households on a monthly frequency. The data set provides a representative sample of Indian households, where households selected to be representative based on 371 "homogenous regions" across India. The survey provides information on the monetary amount of household expenses across 15 major different categories, total household income, and some other background information on the members of the households.[28] In particular, the database provides information on household characteristic such as age and occupation for all members of the households. Lastly, the quality of the data is increased by the fact that CMIE collects the data in person, using specialized workers.[29] In particular, each household is interviewed every four months, and in the interview the person is asked about their behavior month by month in the previous period.[30]

[27] Recall that the excess deposit growth is defined as the growth in deposit in the last quarter of 2016 net of the average quarterly growth in the previous two years.

[28] The expense categories include food, intoxicants, clothing and footwear, cosmetics and toiletries, restaurants, recreation, transport, power and fuel, communication and information services, health, education, bills and rents, appliances, equal monthly installments (EMIs) and others.

[29] In developed countries, this type of consumption survey would likely be conducted by phone or e-mail. This approach would be problematic in developing countries, and lead to wide non-reponse rate which may generate bias or lack of representativeness in the data. To avoid this type of issue, CMIE does not run the survey using phone or e-mail, but instead it employs specialized workers that visit the family and conduct the questionnaire in person. According to CMIE, this approach minimizes the number of non-responding households and allow the sample to cover households with low socio-economic status.

[30] Households are interviewed every four months and report their consumption for each of the four months preceding the

With detailed consumption information for a representative sample of about 95,000 households, we estimate the following household-level counterpart of the district level estimating equation 1:[31]

$$\log\left(y_{h,d,t}\right) = \alpha_t + \alpha_h + \delta_t\left(I_d \times \mathbf{1}_{\{t \geq t_0\}}\right) + \Gamma_t' Y_{h,d} + \epsilon_{h,d,t}, \tag{29}$$

where $y_{h,d,t}$ are consumption measures for household h in district d and time t (t is a month). α_t and α_h are time and household fixed effects, I_d is the treatment variable defined below, which is interacted with dummies for the months post-demonetization, $Y_{h,d}$ are controls, that are either at district or individual level. For controls in the regression, we use the same covariates as in the previous set of analysis with the difference that here we also add household level controls like the age of the head of the household and the log of the household income in September 2016. As usual, standard errors are clustered at district level, which is the level of the treatment.

The main coefficient of interested is the set of parameters δ_t, which estimate the relative change in consumption between the treatment and control group separately for each month after the shock, relative to the pre-period. We use a panel between June 2016 and February 2017, thus providing us with a window of four months before and after the shock. As before we use two measures for the treatment variable I_d. The first measure is a dummy variable that takes the value of one if the district falls in the top decile (household-weighted) of the excess deposit measure.[32] Second, we also present the results using the continuous measure for adjusted deposit growth in the district.

6.2 Main results

Table 9 (Panel (a)) shows results for consumption responses based on exposure to the shock. Column (1) shows that relative to the pre-period, consumption drop was about 7% higher in highly affected districts in November 2016 - the month of demonetization. In December, the reduction in consumption relative to the pre-period was essentially zero, suggesting that household were either unable to or did not substitute inter-temporally for the lost consumption in previous month. However, the difference is still significant and similar in magnitude in January. Post January 2017, as shown later in detail, the differences in consumption responses across the two groups essentially disappear as the cash levels returned to normalcy.

To understand better the effect on consumption dynamics, we repeat the same analysis by dividing ag-

survey month. Thus, about 39,500 households are surveyed every month. The execution is planned such that every week nearly 10,000 households are interviewed.

[31] We only consider households for which the age of head of household is between 18 and 75 years as of September 2016. We then only consider households with non-missing information between June 2016 and March 2017, giving us a balanced panel of about 95,000 households.

[32] We also show that results are essentially unchanged if we use other cutoffs in defining the dummy, for instance top quartile.

gregate consumption into finer categories. First, we explore the responses separately across necessary and unnecessary consumption goods. In particular, we define necessary consumption based as groups that represent the major share of household consumption basket and hence has lower probability for intertemporal substitution. This includes expenditure on food, bills and rent, gasoline and other fuels, and toiletries. Second, we instead look at the differential responses by splitting consumption between frequent and infrequent consumption baskets. This categorization is motivated by the fact that not all consumption is frequented across all months equally. Specifically, the infrequent categorization contains expenditure on appliances, clothing and footwear, intoxicants, recreation, restaurants, various monthly installments and any other unclassified expenditure under "miscellaneous".

To be clear, in our minds these two measures should capture the same underlying phenomenon and we do not have any specific prediction that should hold for one class but not the other. In particular, both categories should capture the set of consumption that can be reduced more easily by households, with more limited repercussion on welfare. In some ways, the two measures can be considered as two different proxy of the same underlying variables. Later, we also show how results are conceptually similar with the primitive variables.

Looking across necessary and unnecessary consumption, we find stark differences in the reaction to the shock across these two quantities. Column (2), shows that the shock had a limited effect on the consumption of necessary goods. In particular, the coefficient of interest is never significant at a conventional level and the magnitude is smaller than for total consumption. At the same time, unnecessary consumption contracted drastically. Column (3) shows that households in highly affected district saw a 16% reduction in consumption of unnecessary goods in November. Similar to total consumption, the flow of consumption is still non-significantly lower in December and it contracts again significantly in January. Column (4)-(5) shows that classification based on frequent and infrequent consumption mimic these results - infrequent expenses for highly affected household contracted by more than 28% in November, with no reversal in December and a similar decline in January.

Panel (b) of the Table 9 shows that the results are qualitatively similar when using a continuous variable for district's exposure to the shock. Consumption responses to district's excess deposit growth are all negative and significant across all the coefficients for the month of November. Moreover, relative magnitudes presents a pattern very similar to the previous analyses. First, the economic magnitude for unnecessary is about double than the one for necessary. Second, the effect for infrequent is almost three times larger than for frequent. Third, most of the contraction is very limited in time to November. In Appendix Table B.5, we show that results are qualitatively similar when using a dummy for the treatment splitting the sample at the top 25%.

These results suggest that cash shortage has significant impact on household consumption, but these effects did not persist over time and were for most part concentrated in infrequent and unnecessary categories.[33] We present two additional tests in support of this argument. First, we show that the results are not driven by the way we categorize consumption in two groups. In Appendix Table B.7, we find that there were none to minimal effects on the consumption of food, but relatively large effects on expenses related to recreation, services, or transportation activities.[34] Second, in Appendix Table B.8 we present the result over a longer horizon and we show that - even for the specification based on extreme size shock (top decile of deposit growth) - the effect is primarily concentrated in the initial months following demonetization and the level of consumption came back to normality thereafter.

6.3 Robustness

The results on consumption and adoption of electronic wallet technologies rely on the assumption that our measure of deposit growth does not capture any other economic force that may be affecting the change in behavior, and at the same time is unrelated to the demonetization shock. The results from Section 3 show that - consistent with this hypothesis - our measure of shock intensity does not seem to predict any differential change in adoption of technology before November, but only after the announcement. In this section, we present additional set of results in support of this hypothesis.

6.3.1 Pre-trends

As a first step, we examine the dynamic of the effect to test for the presence of pre-trend in the data. In the two panels of Figure 14, we replicate the same specification as before with the exception that now we estimate a specific parameter for each month both before and after the shock. All parameters are reported relative to consumption in October 2016, which we normalized to zero. Furthermore, we also extend our window of analysis by including more months, both before and after, to provide a more exhaustive picture of the dynamics.

Consistent with the pre-trend analysis using electronic wallet data, we find that the high excess deposit districts did not experience a decline in consumption in the months before demonetization. The contraction in relative consumption happens primarily in the months after the shock, with the strongest effects in November. Consistent with previous results, the effects are stronger in infrequent and unnecessary consumption.[35]

[33] Importantly, the main results do not depend on the inclusion or not of the controls. Removing controls, we still find In particular, that in highly affected areas consumption declined mostly for unnecessary and infrequent purchases. The main difference is that without controls we find that the decline in terms of total consumption is statistically non-significant, but its magnitude is close to the one in our main specification.

[34] Services expenditure covers transportation, education, health and communication expenses.

[35] The results with the continuous variable are essentially equivalent, with the exception that - in line with results shown before - most of the negative results are concentrated in the month of November.

6.3.2 Placebo shocks

Although the results above rule out any pre-trend, it could still be the case that districts with high growth in deposit post demonetization are regions that are particularly sensitive to business cycle fluctuations. To shed light on this issue, we repeat the above analysis but by centering it in a period with no shock. In particular, we consider the possibility of a "fake" shock happening in every month in the year before demonetization. We then test whether districts with high excess deposit growth in the last quarter of 2016 experienced large negative changes in consumption in response to these "fake" shocks.

The two panels of Figure B.1 report the result. To be clear, the figure plots the coefficient that is equivalent to the interaction between November 2016 and the dummy for highly affected districts in the main analysis (equation 29), where we now replace the November 2016 dummy by the "fake" shock month corresponding to each month on the horizontal axis. Importantly, the whole sample period and regressors are constructed like in the main analysis but just centered around the month in the $x-$axis rather than the demonetization month. Panel (a) reports results for unnecessary consumption. Across all the periods examined, we find no statistical difference in consumption growth for high deposit-growth districts. Furthermore, the coefficient in almost every case positive - the opposite sign compared to our main result. The results for infrequent consumption which are reported in panel (b) also presents the same pattern.

6.3.3 Effects from Diwali

It could also be possible that our results are simply capturing some of the effect of Diwali. Diwali is a major Indian festival primarily celebrated by Hindus. As festivals in the country are based on their religious calendars, which typically follows solar or lunar calendar, Diwali does not occur on a fixed date over the years. In particular, its occurrence generally varies between mid-October to mid-November. As for most religious or civil festivities, the concern here is that this type of event may affect the consumption patterns of families in a way that may affect our results. However, our results are not likely to be affected by this event. First, to the extent that the Diwali effects national consumption patterns, any effect should be absorbed by the month fixed-effects. Second, the evidence based on the pre-trend analysis above suggests that there was no abnormal behavior in the months before the shock, which is what we would expect if results simply capture some anomalous movement in consumption patterns due to Diwali festivities in October.

To bolster this point, we examine how consumption reacted in the previous two Diwali - for which we have data - and find no responses in support of the hypothesis that our results could be explained by consumption pattern due to Diwali. The intuition behind this test is that - if what we are capturing in our analysis is the change in consumption post-Diwali - then one should expect to see similar effects in previous years. In

order to explore this dimension, we replicate our main table using data from the previous two years (2014 and 2015). Specifically, we construct measures in the same way in the year of interest and place the Diwali month as the last month of the pre-period. In this way, we mimic exactly how the analysis is ran in 2016.

The results are in Appendix Table B.9. The table reports the estimates of the post-Diwali coefficient, which is the equivalent of our main November analysis in the previous results. Across the two years, we find reassuring results - during 2014 and 2015 there is no evidence that households in district experiencing an exceptional increase in deposits had any fallout in consumption after Diwali. If anything, the coefficient is actually positive, but its size is usually small and - apart from one case - statistically insignificant. This result, together with those previously discussed, show that the concern that Diwali may contaminate the main finding in the paper does not seem to be supported by empirical evidence.

Overall, these tests provide supportive evidence on the idea that excess deposit growth may represent a valid measure to estimate the casual impact of demonetization.

6.4 Consumption responses and electronic payment

The above results show that the demonetization affected household behavior by reducing consumption. This evidence is consistent with the idea that the temporary reduction in credit and the overall disruption caused in the economy by the currency swap had real effects. However, the estimates also suggest that these effects were very concentrated across spaces, times, and type of consumption. In fact, we find that the effect are mostly driven by consumption that can be defined as unnecessary or infrequent and the reduction is significant only during the main months of the demonetization. Furthermore, the effects tend to be driven by districts in which the shock appeared to be very high.

This overall picture is suggestive that - to some extent - households were able to cope with the reduction in available cash and limit some of the adverse effect of the policy. Given the previous results on the large increase in electronic payments during the same period, one hypothesis is that a large share of the payments and transactions for consumption purposes were simply switched from cash to electronic payment. To shed more light on this dimension, we conduct a study examining the heterogeneity of the effect across households based on their likelihood to be able to access electronic payments. To the extent that electronic payments did play a role in reducing the impact of the shock, we should expect smaller (or even null effects) for households with high ability to switch to the new mean of payments.[36]

We extend our previous analysis by simply estimating our main models across different sub-samples, characterized by different likelihood or propensity to access electronic payment methods broadly defined.

[36]Since the ability to use payment technology depends on both the demand and supply of payment methods (i.e. the retailer needs to be able to accept the payment) it does not have to be the case that the effect is zero for individuals with an easier access to alternative payment technologies.

The results of these tests are reported in Table 10. Here, we report the main coefficient - the interaction between households in highly affected districts and the dummy for month of November 2016 - across the various outcomes of interest and across the different sub-groups. Even if we report only this coefficient, the specification and set-up is identical to the one of the main results.

To sort households based on the likelihood of accessing electronic payments, we divide the sample across three dimensions. First, we examine the difference in the estimates across different income groups, constructed by dividing the sample in four quartiles of household income in August 2016 (three months before the shock). The idea is that high-income individuals should be more likely to use electronic payments or shop in stores where electronic payments are accepted. The findings are striking. As before, splitting consumption based on necessary and unnecessary baskets, we find that the group in the top quartile of income did not experience any significant reduction in consumption in either basket. In general, the estimate is much smaller than the other coefficients and it is highly non-significant. However, the effect becomes significant and economically larger for the poorer quartiles, with strongest effect among lowest income group. The same dynamics holds in column (3) and (4), when we split consumption into frequent and infrequent baskets.

Second, we repeat the same exercise as above based on household's education level. In line with before, we expect that more educated consumers will have a better access to electronic payment technologies. In fact, while income is an important variable to capture the ability to use these tools, education and financial literacy has been shown to play a key role in financial decision making, sometimes over and above income. To explore this dimension, we split the sample in three groups depending whether the head of the households has a college degree, high school degree, or some and no relevant education. Across all four outcomes, we find that the effect for college educated households is much smaller than the others and - apart from infrequent consumption - non-significant. On the flip-side, households with little or no formal educations appears to be highly affected. The effect for high-school household is more mixed, but generally lies between the other two groups.

Third, we show that the main effect cannot be detected for those households in which some member had access to a credit card. Importantly, we measure credit card holding during August 2016 to rule out the concern that credit card holding could be responsive to the shock itself. We consistently find that the main effect on consumption across different basket is always non-significant for households that own a credit card. On the other hand, households without credit card in districts experiencing a large increase in deposits witnessed significant decline in consumption, more so for infrequent and unnecessary purchases. Interesting, the effect is never statistically different across the two groups. One explanation for this result is that the access to a credit card does not necessarily guarantee the ability to use it, since it requires retailer to have the technology to be able to accept credit card. As a result, households experiencing a large shock would be

effected even if they had access to alternative electronic payments.

Overall, the consumption evidence support the idea that electronic payments played a crucial role in moderating the potential negative effect of the demonetization. Consistent with this hypothesis, the shock led to a contraction in consumption but these effects were highly concentrated across space, time, and basket of consumption. Furthermore, households with resources and ability to be able to access electronic payments with relative ease were little affected by the shock.

7 Conclusion

In this paper, we study technology adoption dynamics in the presence externalities in the adoption decision. The technology we study are electronic payment systems; we use the Indian demonetization of 2016 as a laboratory to study how a large, aggregate shock can affect adoption. The shock led to a permanent increase in its adoption by the firms. Our results support the view that, beyond fixed adoption costs, the event served as a coordinating device that allowed firm to overcome coordination frictions in the adoption of the technology.

Using data from the leading electronic wallet payment company, we start the paper by documenting two findings. First, the shock led to an increase in adoption of payment technologies. This positive connection between the shock and the adoption is found both in aggregate and across districts. Second, the adoption was persistent — the number of firms and amount transacted on the payment platform remains high even after the shock dissipates.

Next, we explain these findings through the lens of a dynamic model of technology adoption with positive externalities. The model rationalizes the permanent increase in adoption (even thought the shock was transitory) and highlights the importance of complementarities this persistence.

To further support the role of complimentarities in adoption dynamics, we test two distinct predictions of the model in the data. First, the model predicts some level of history-dependence. In fact, because of complementarities, the level of adoption after the shock depends on the existing level of adoption before the shock. Using district-level data, we find evidence that supports this hypothesis. Second, complimentarities generates spillovers across neighboring firms. When more firms within the same industry and location adopt a technology this period, the marginal adoption benefit also increases, leading to more usage of the technology in the future. Electronic payment usage by firms around the demonetization supports this pattern.

Our focus on electronic payment system technologies allows us to identify complimentarities independently of sunk costs effects. While this ability to separate sunk costs from complementarities is unique to this setup, we believe that the mechanism that we highlight here can be relevant in other contexts where

complementarities in adoption are important. Our results suggest that providing large incentives, for example through innovation credits or subsidies, even for a short period of time, could act as a coordinating device among firms and incentivize them to undertake value-enhancing investments.

References

Agarwal, S., S. Alok, P. Ghosh, S. Ghosh, T. Piskorski, and A. Seru (2017). Banking the unbanked: What do 255 million new bank accounts reveal about financial access? Technical report.

Agarwal, S., D. Basu, P. Ghosh, B. Pareek, and J. Zhang (2018). Demonetization and digitization.

Atkin, D., A. Chaudhry, S. Chaudry, A. K. Khandelwal, and E. Verhoogen (2017). Organizational barriers to technology adoption: Evidence from soccer-ball producers in pakistan. *The Quarterly Journal of Economics 132*(3), 1101–1164.

Basker, E. (2012). Raising the barcode scanner: Technology and productivity in the retail sector. *American Economic Journal: Applied Economics 4*(3), 1–27.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? Technical Report 1.

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts (2013). Does management matter? evidence from india. *The Quarterly Journal of Economics 128*(1), 1–51.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics 130*(3), 453–483.

Burdzy, K., D. M. Frankel, and A. Pauzner (2001). Fast equilibrium selection by rational players living in a changing world. *Econometrica 69*(1), 163–189.

Carlsson, H. and E. Van Damme (1993). Global games and equilibrium selection. *Econometrica: Journal of the Econometric Society*, 989–1018.

Chodorow-Reich, G., G. Gopinath, P. Mishra, and A. Narayanan (2018). Cash and the economy: Evidence from india's demonetization. Technical report, National Bureau of Economic Research.

Comin, D. and R. Nanda (2014). Financial development and technology diffusion.

Conley, T. G. and C. R. Udry (2010). Learning about a new technology: Pineapple in ghana. *American economic review 100*(1), 35–69.

de Roure, C., L. Pelizzon, and A. V. Thakor (2018). P2p lenders versus banks: Cream skimming or bottom fishing?

Foster, A. D. and M. R. Rosenzweig (2010). Microeconomics of technology adoption. *Annu. Rev. Econ. 2*(1), 395–424.

Frankel, D. and A. Pauzner (2000). Resolving indeterminacy in dynamic settings: the role of shocks. *The Quarterly Journal of Economics 115*(1), 285–304.

Fuster, A., M. Plosser, P. Schnabl, and J. Vickery (2018). The role of technology in mortgage lending. Technical report, National Bureau of Economic Research.

Hall, B. H. and B. Khan (2003). Adoption of new technology. Technical report, National bureau of economic research.

Howell, S. T., M. Niessner, and D. Yermack (2018). Initial coin offerings: Financing growth with cryptocurrency token sales. Technical report, National Bureau of Economic Research.

Katz, M. and C. Shapiro (1994). Systems competition and network effects. *Journal of Economic Perspectives 77*(77), 93–115.

Katz, M. L. and C. Shapiro (1985). Network externalities, competition, and compatibility. *The American economic review 75*(3), 424–440.

Mansfield, E. (1961). Technical change and the rate of imitation. *Econometrica: Journal of the Econometric Society*, 741–766.

Morris, S. and H. S. Shin (1998). Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review*, 587–597.

Morris, S. and H. S. Shin (2002). Social value of public information. *american economic review 92*(5), 1521–1534.

Morris, S. and H. S. Shin (2003). Global games: Theory and applications. *Advances in Economics and Econometrics*, 56.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the us labor market. *The Quarterly Journal of Economics 118*(2), 549–599.

Munshi, K. (2004). Social learning in a heterogeneous population: technology diffusion in the indian green revolution. *Journal of development Economics 73*(1), 185–213.

Rogoff, K. S. (2017). *The Curse of Cash: How Large-Denomination Bills Aid Crime and Tax Evasion and Constrain Monetary Policy*. Princeton University Press.

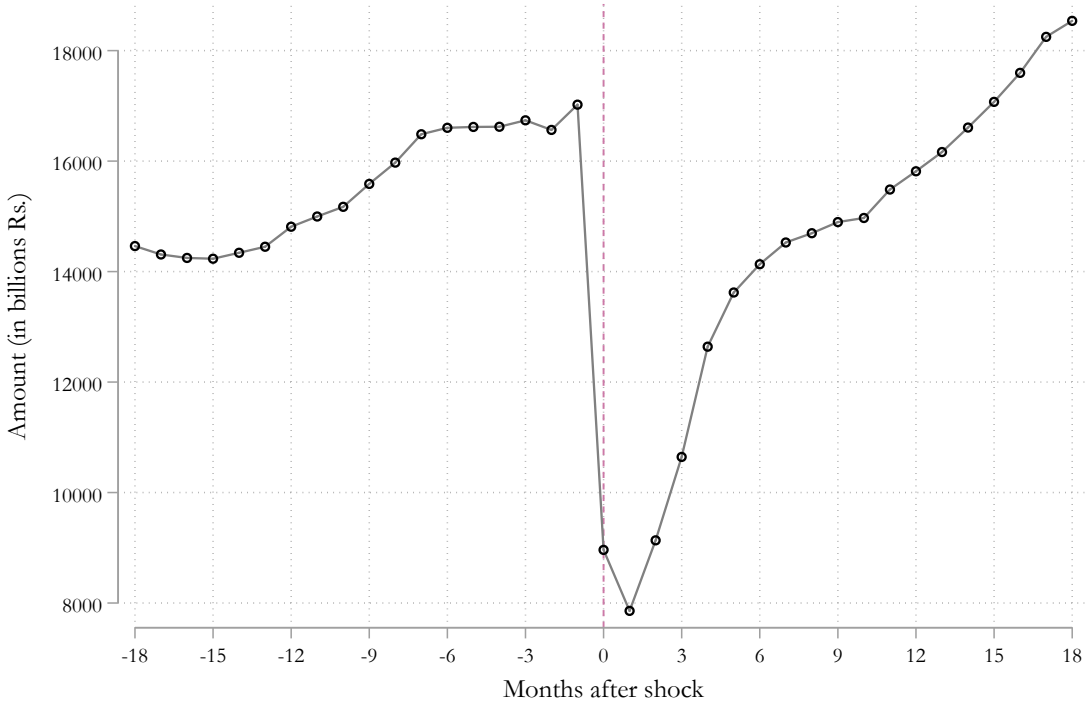
Rosenberg, N. (1972). Factors affecting the diffusion of technology. *Explorations in economic history 10*(1), 3.

Suri, T. and W. Jack (2016). The long-run poverty and gender impacts of mobile money. *Science 354*(6317), 1288–1292.

Tang, H. (2018). Peer-to-peer lenders versus banks: substitutes or complements? *Review of Financial Studies*.

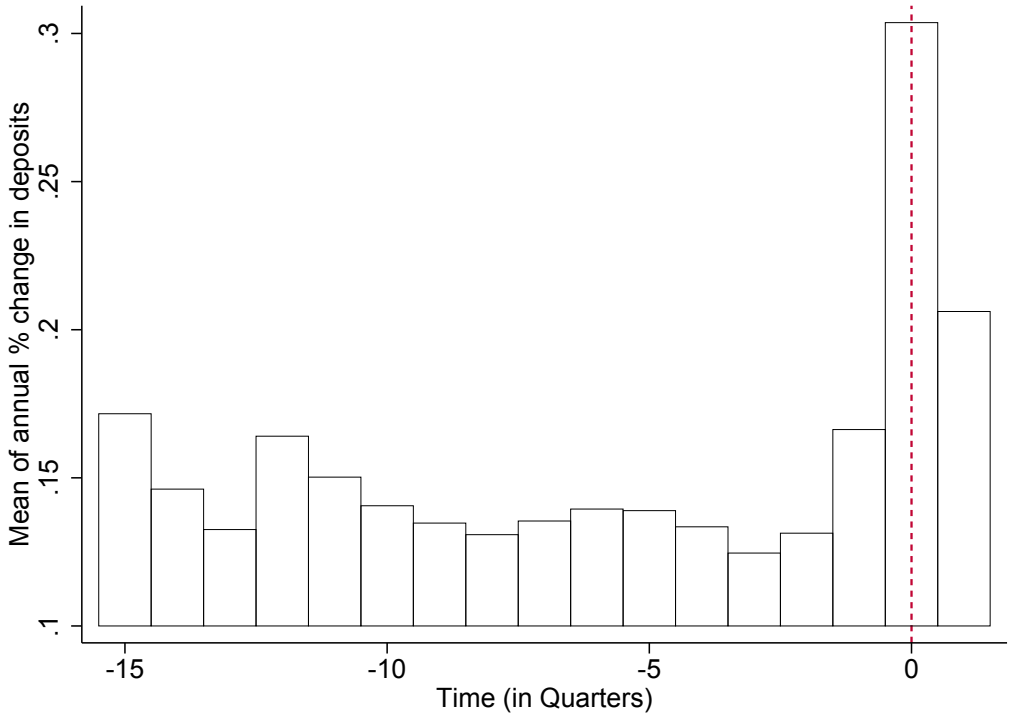
Yermack, D. (2018). Fintech in sub-saharan africa: What has worked well, and what hasn't. Technical report, National Bureau of Economic Research.

Figure 1: Nominal value of currency in circulation



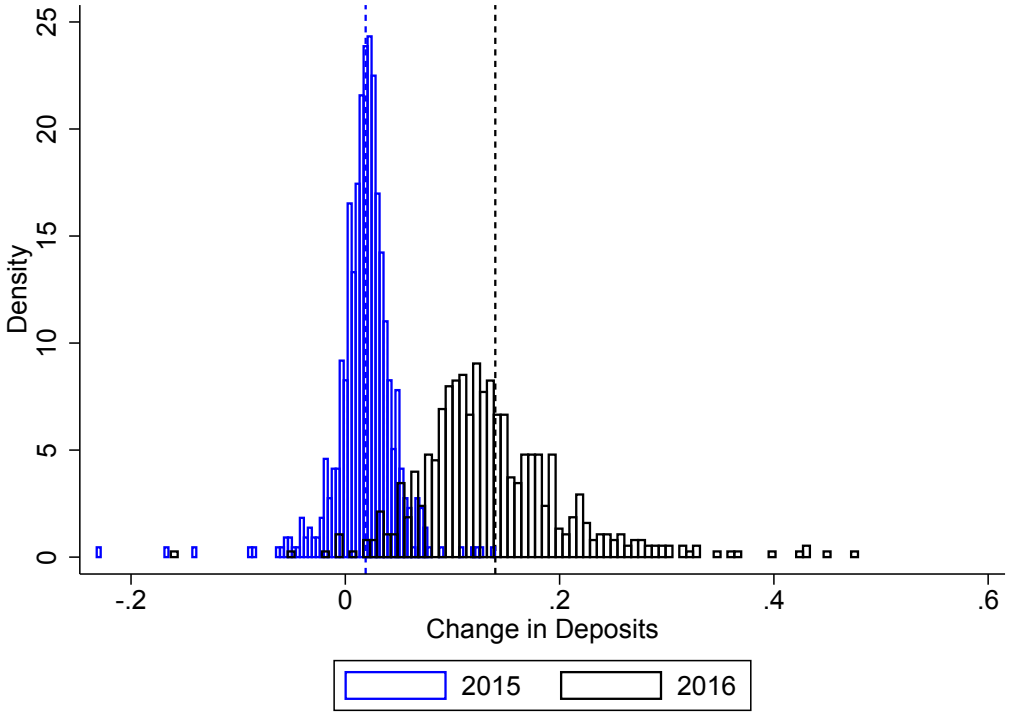
Notes: Nominal value of the stock of currency in circulation in India. Month 0 is the month of November 2016; the figures are end-of-month estimates. Source: Reserve Bank of India.

Figure 2: Average annual growth rate in deposits across districts



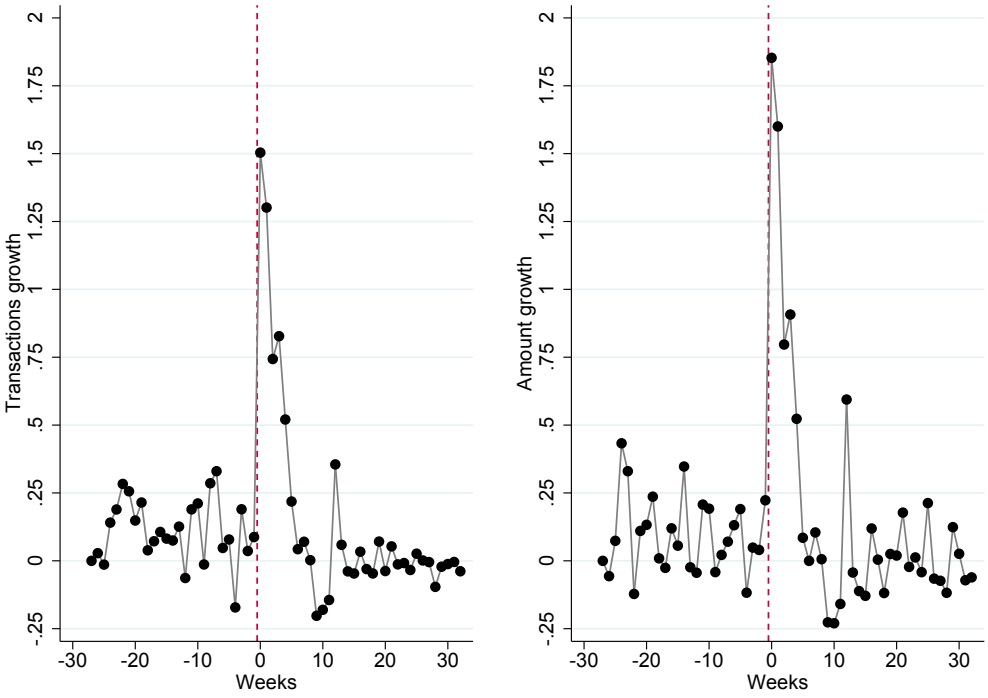
Notes: The graph shows mean of year-on-year growth in banking sector deposits across districts from Quarter 1 of 2013 to Quarter 1 of 2017. Time 0 is the quarter of November 2016. Source: Reserve Bank of India.

Figure 3: Distribution of growth in deposits across districts



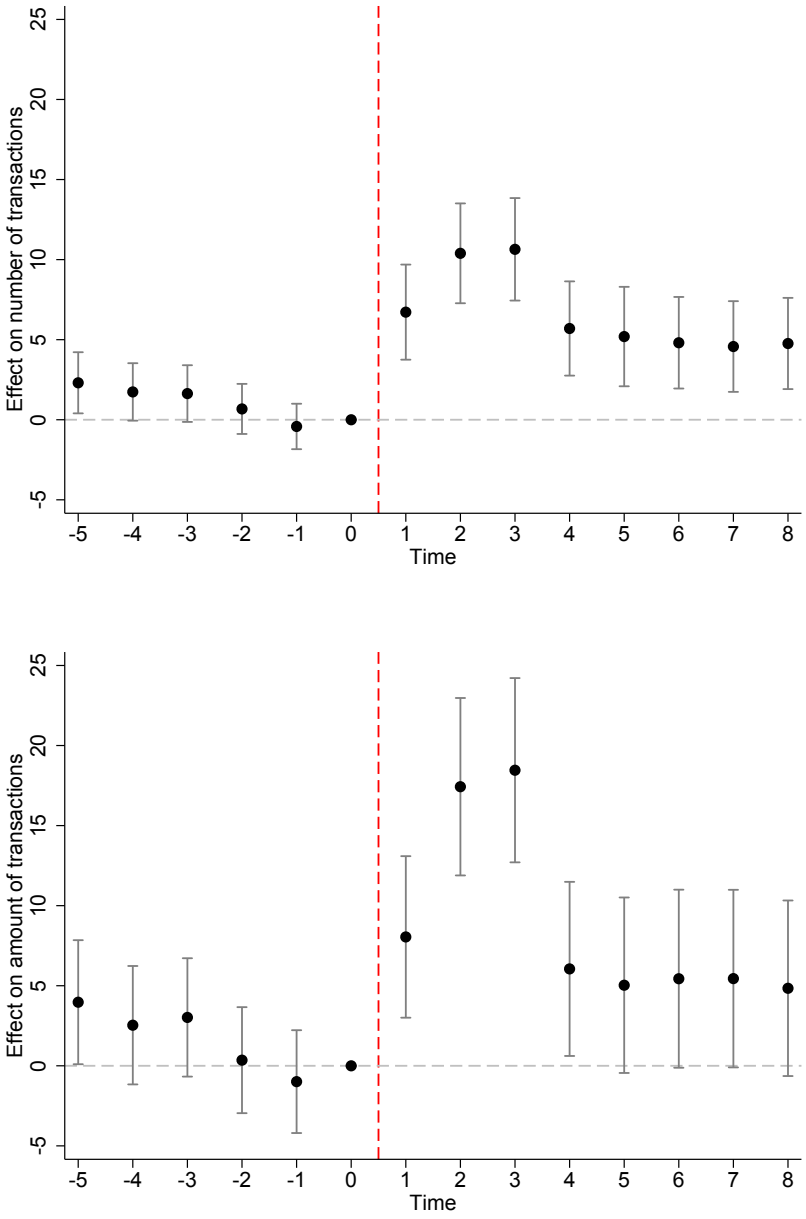
Notes: Distribution across deposits of the growth in total banking sector deposits from October to December during year 2015 (blue) and 2016 (black). The vertical dashed lines represent the corresponding mean deposit growth for these years. Source: Reserve Bank of India.

Figure 4: Amount and Transaction Growth on Mobile Payment Platform



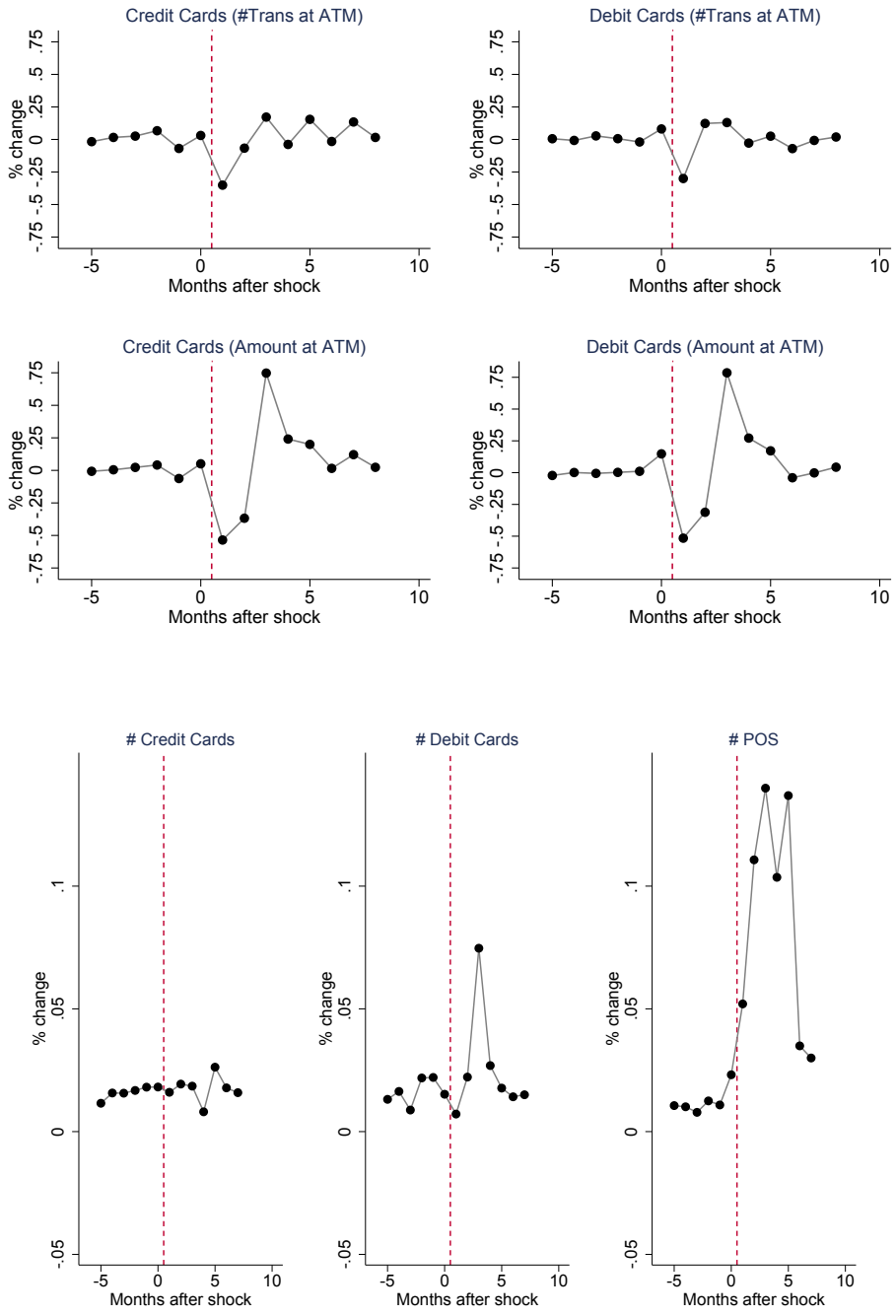
Notes: Week-on-week growth rate in number of transactions (left panel) and amounts (right panel) on the electronic wallet platform. The dashed red line indicates the week of November 8th, 2016. See main text for a discussion of data sources.

Figure 5: District adoption dynamics in electronic payments data based on exposure to shock



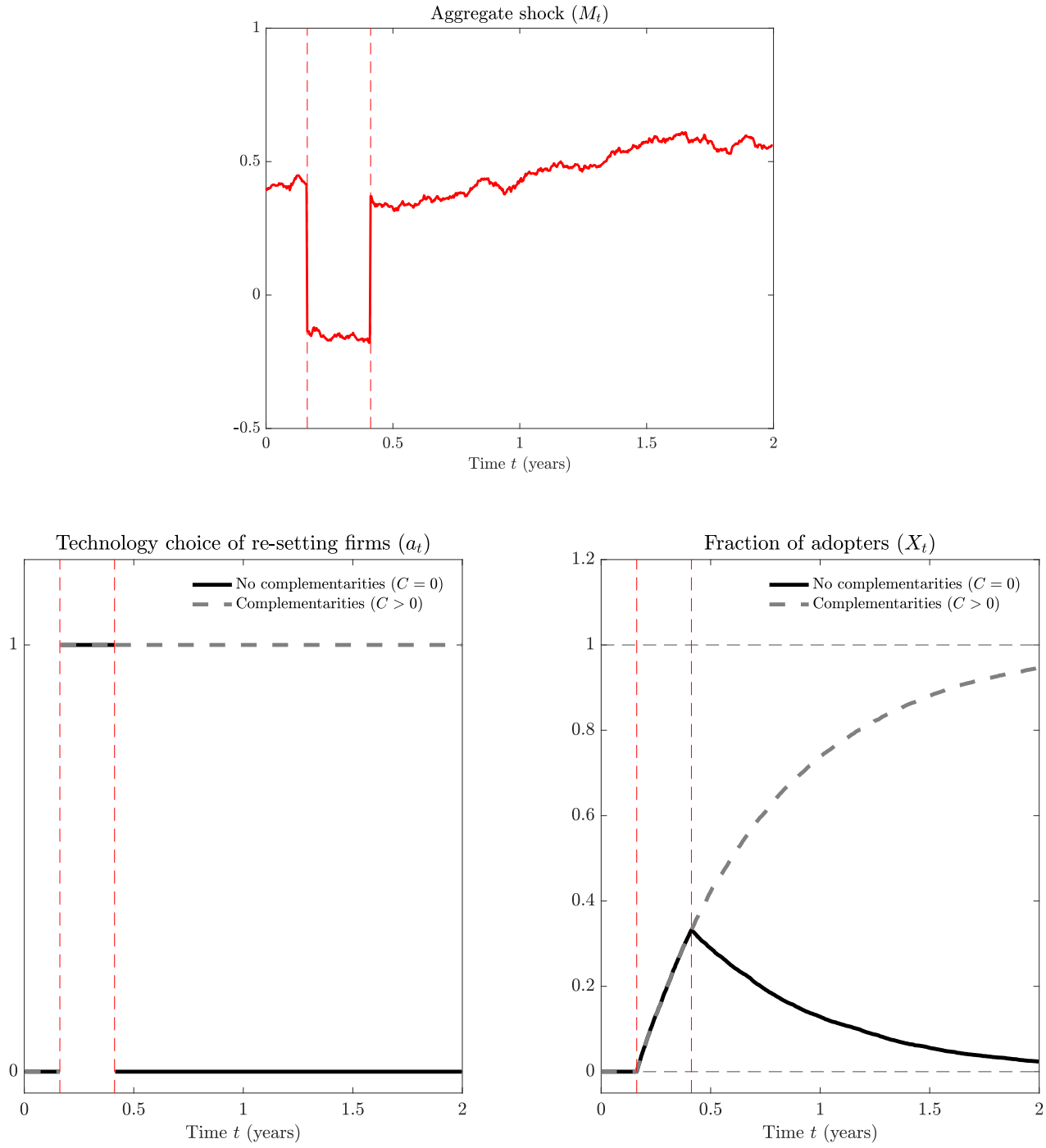
Notes: The figure plot the dynamic treatment effects of the demonetization shock on technology adoption of electronic payment systems. The graphs report the coefficients δ_t from specification 2; the top panel reports effects for the number of transactions, and the bottom panel reports effects for the total amount of transactions. 90% confidence intervals are represented with the vertical lines; standard errors are clustered at the district level.

Figure 6: Changes across alternate electronic payment systems



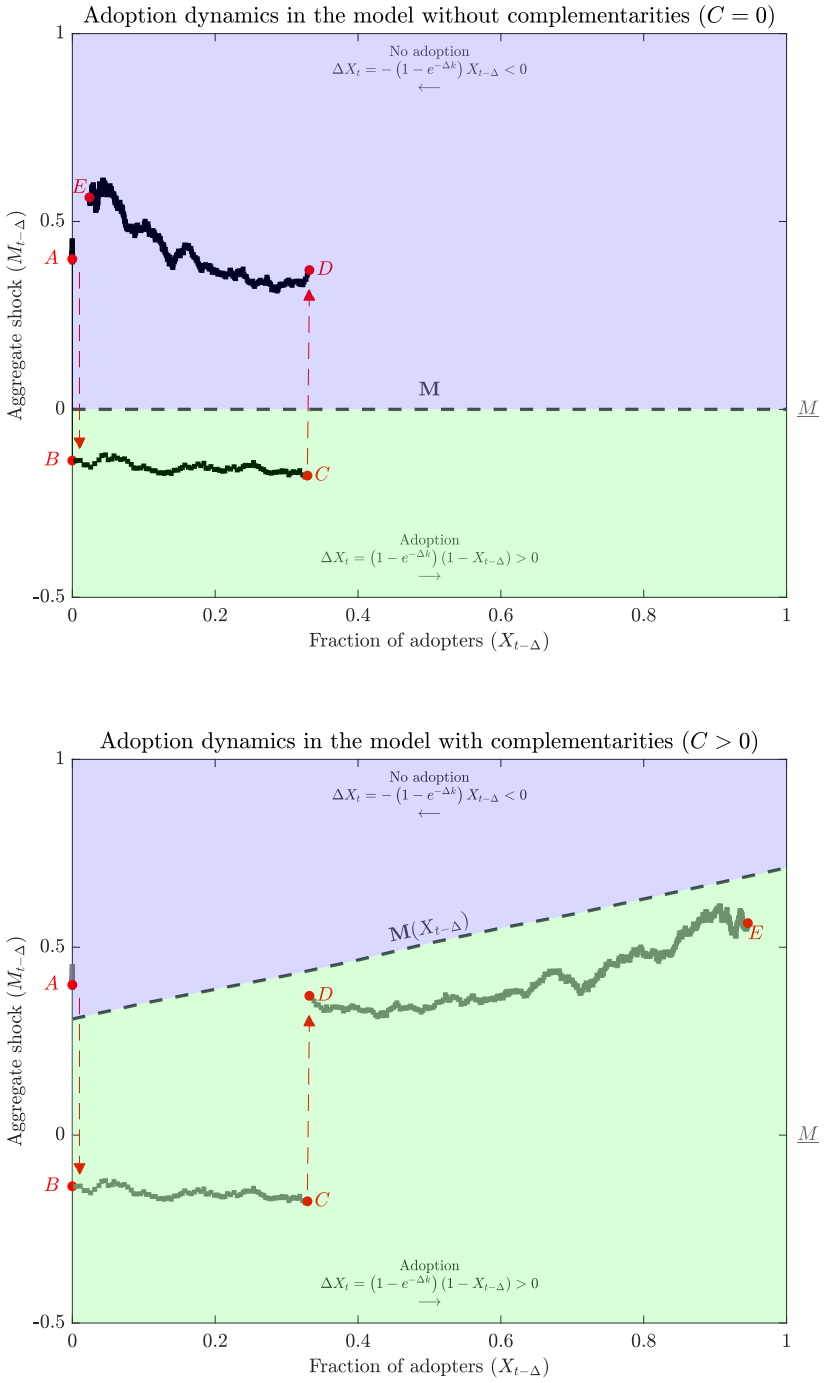
Notes: Change in use of other electronic payment systems for credit cards, debit cards and POS (point of sale machines) around the period of the shock. The top panel reports measures of intensive margin use, and the bottom panel reports measures of adoption. All the data are monthly and at the national level. The dashed red line indicates the month before the shock. Source: Reserve Bank of India.

Figure 7: Technology adoption paths in the model



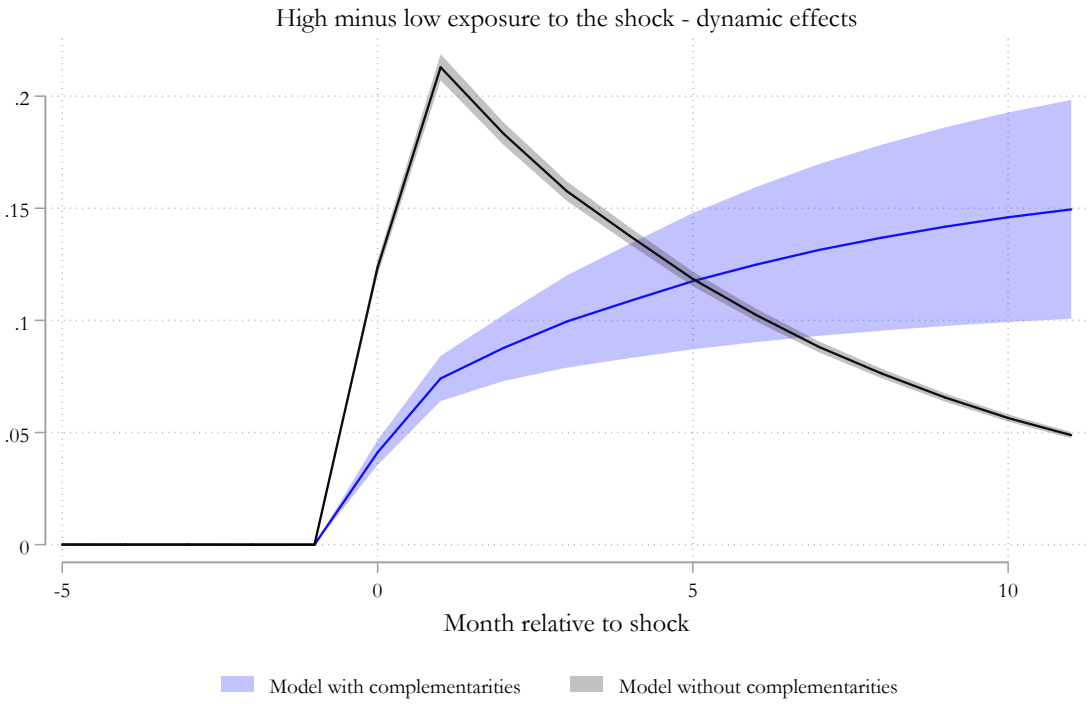
Notes: The figure plots the technology adoption paths with and without complementarities in response to aggregate shock to M_t based on the model. The top panel reports a particular sample path for M_t, chosen so that two symmetric jumps occur at the beginning of months 3 and 6; the red dashed lines indicate the jumps. The bottom panels reports the adoption decision a_t and the fraction of adopters X_t along this sample path, in the model without complementarities ($C = 0$, solid black line) and in the model with complementarities ($C > 0$, dashed grey line). The drift of the aggregate shock is set to $\mu = 0$.

Figure 8: Adoption dynamics in the model



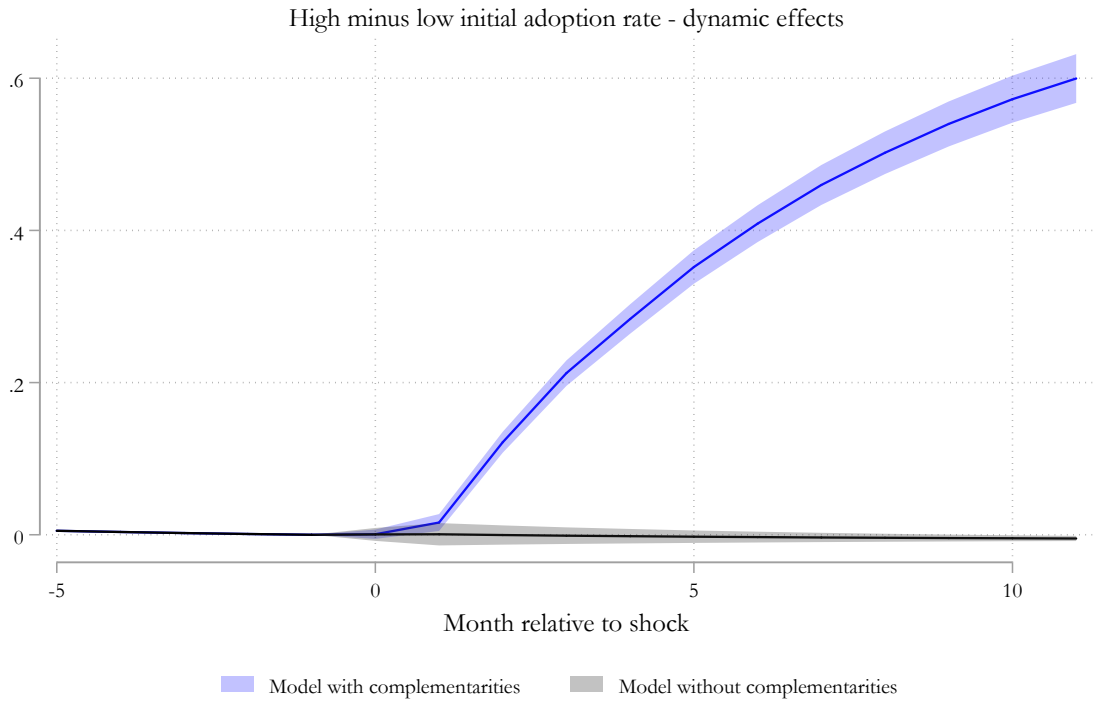
Notes: The figures shows adoption dynamics in the model without complementarities ($C = 0$, top panel) and in the model with complementarities ($C > 0$, bottom panel). In each graph, the dotted lines traces the state of the economy in $(X_{t-\Delta}, M_{t-\Delta})$ space following the sequence of aggregate shocks reported in the top panel of figure 7. The shaded green area is the adoption region, while the shaded blue are is the no-adoption region; the frontier $\mathbf{M}(X_{t-\Delta})$ is marked by the dashed line. The drift of the aggregate shock is set to $\mu = 0$.

Figure 9: District adoption dynamics in simulated data based on exposure to the aggregate shock



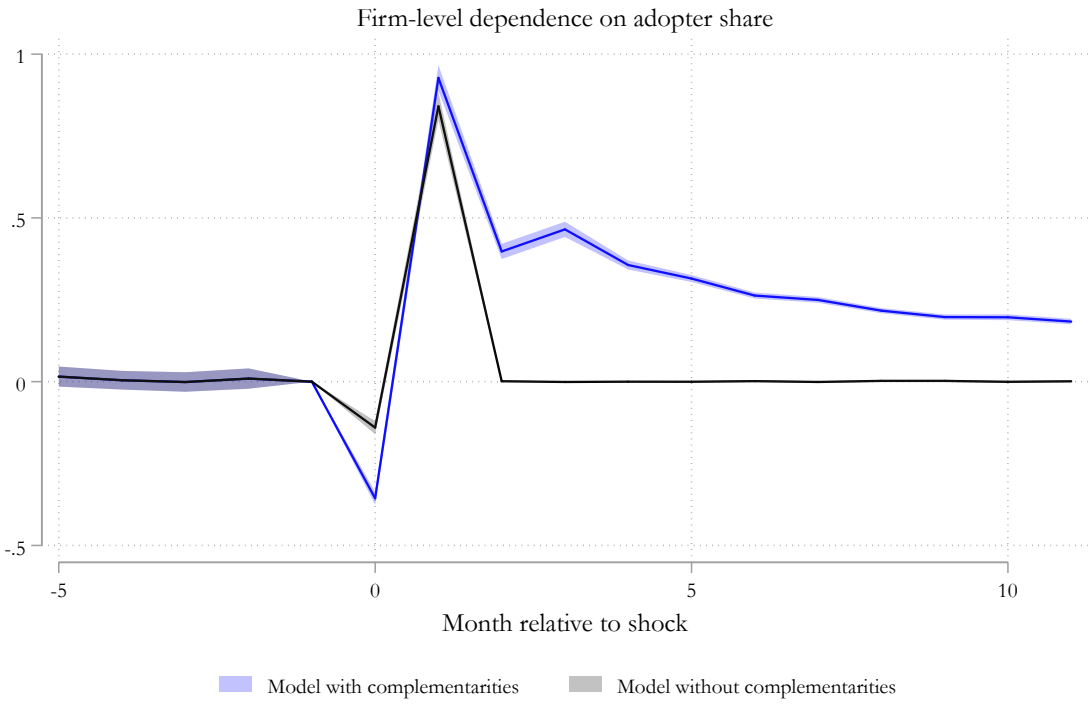
Notes: The figure shows dynamic effects in the difference-in-differences estimator of the effect of the shock $M_{d,t}$, applied to simulated data. The estimated equation is specification 19. The coefficients reported are the coefficients δ_t, with the month-9 coefficient normalized to 0, along with 95% confidence bands.

Figure 10: District adoption dynamics in simulated data based on initial adoption rate



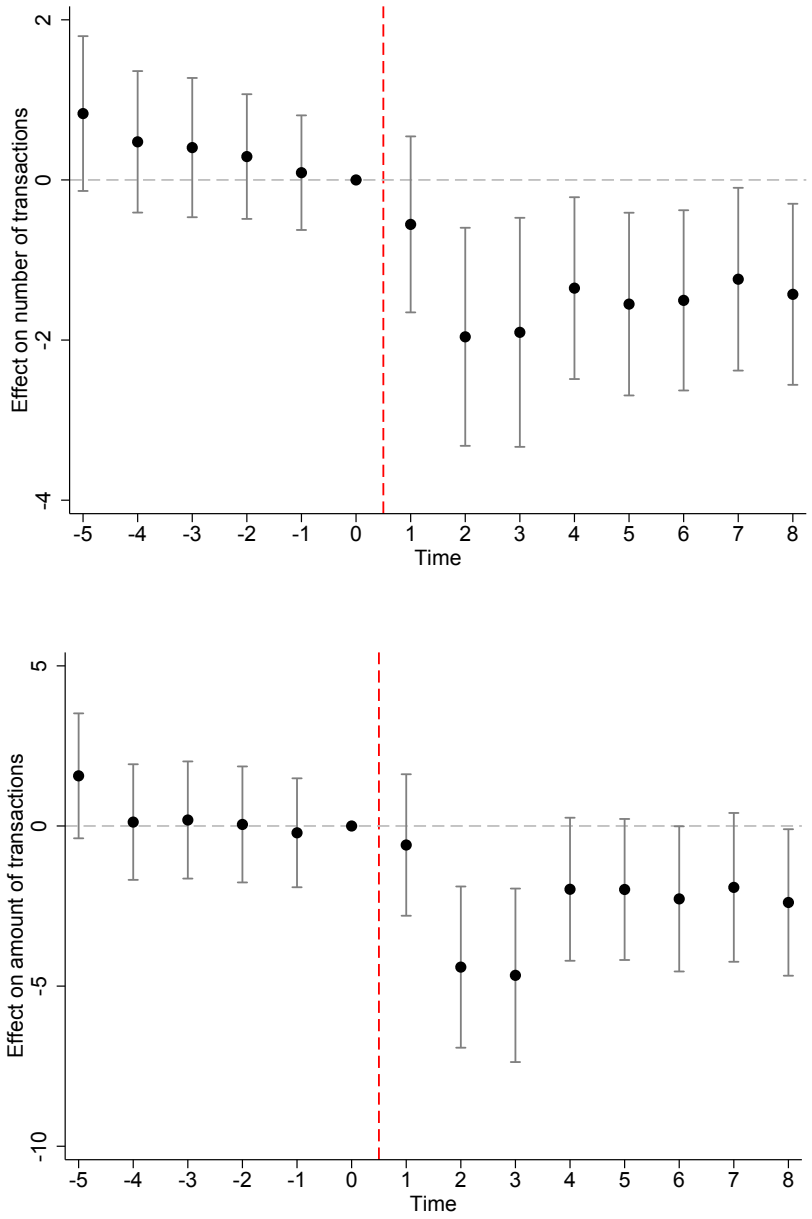
Notes: Dynamic effects in the difference-in-differences estimator of the effect of initial conditions $X_{0,d}$, applied to simulated data. The estimated equation is specification 21. The coefficients reported are the coefficients δ_t, with the month-9 coefficient normalized to 0, along with 95% confidence bands.

Figure 11: Firm adoption dynamics in simulated data based on existing share of adopters



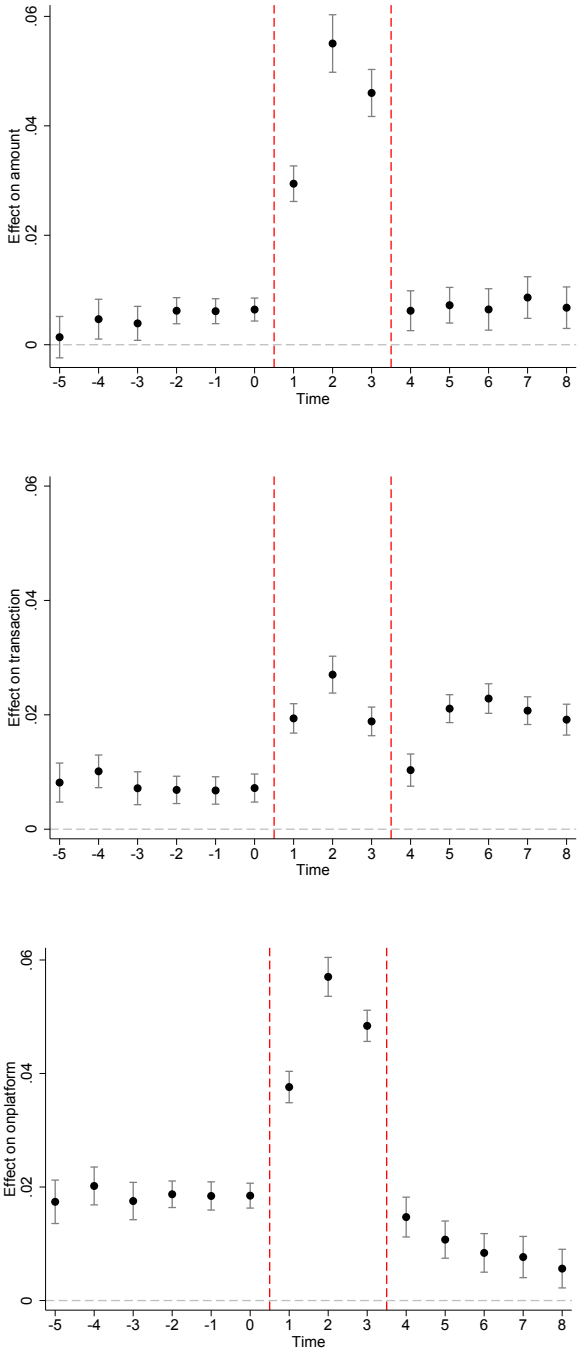
Notes: The figure shows dynamic estimates of the dependence of individual firm's technology choice on lagged adopted share, applied to simulated data. The estimated equation is specification 25. The coefficients reported are the coefficients γ_t, with the month-9 coefficient normalized to 0, along with 95% confidence bands.

Figure 12: District adoption dynamics in electronic payments data based on initial adoption rate



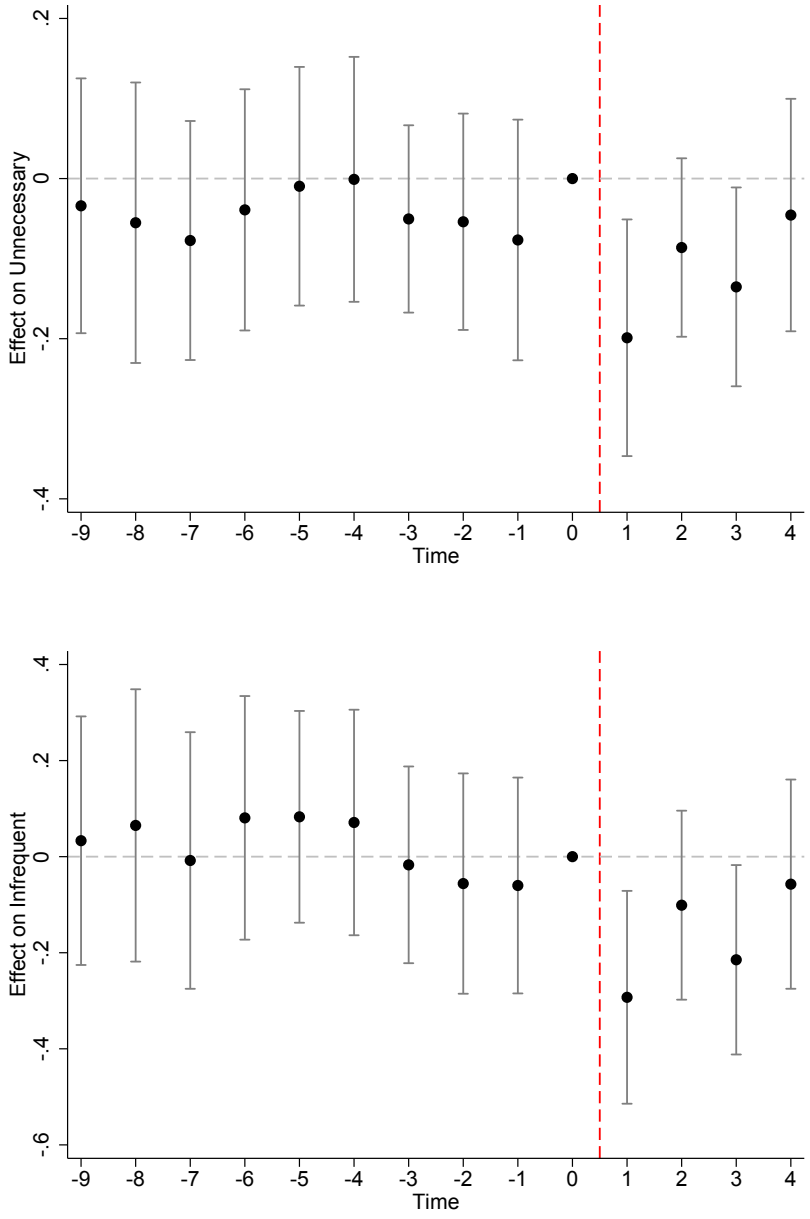
Notes: The figure plot the dynamic effects of adoption across districts based on district's initial adoption rates. The specification we estimate is the dynamic version of equation 26. The top panel reports effects for the number of transactions, and the bottom panel reports effects for the total amount of transactions. The treatment is the nearest distance of a district to a electronic wallet hub district. 90% confidence intervals are represented with the vertical lines; standard errors are clustered at the district level.

Figure 13: Firm adoption dynamics in electronic payments data based on existing share of adopters



Notes: The figure plots month-by-month estimates of the dependence of firm-level adoption rates on the share of other adopters in the industry/pincode. The specification we estimate is a version of equation 28 in which each coefficient is interacted with a weekly dummy; we reported the monthly estimates of the coefficient γ. The top panel reports effects when x is the total value of transactions, the middle panel reports effects when x is the total number of transactions, and the bottom panel reports effects when x is a dummy for whether the firm used the platform over the past week. 95% confidence intervals are represented with the vertical lines; standard errors are clustered at the pincode level.

Figure 14: Consumption responses based on exposure to the shock



Notes: The figure plots month-by-month estimates of consumption responses depending on exposure to the shock. The specification we estimate is a version of equation 29 in which each coefficient is based on the interaction of the treatment variable with a monthly dummy. We report the monthly estimates of the coefficient δ. Treatment is a dummy that takes the value of one if the district falls in the top decile (household-weighted) of the excess deposit measure. The top panel reports the effects when y is the unnecessary consumption basket, and the bottom panel reports when y is the infrequent consumption basket (as described in Section 5). 95% confidence intervals are represented with the vertical lines; standard errors are clustered at the district code level. Source: CMIE Consumption Data.

Table 1: Share of Chest Banks and Deposit Growth

	Deposit Growth			
	(1)	(2)	(3)	(4)
1−share of chest banks	0.223^{***}	0.147^{**}		
	(0.065)	(0.064)		
Low share of chest banks			0.141^{***}	0.124^{***}
			(0.031)	(0.031)
District-level controls	✓	✓	✓	✓
Control for log(Total deposits pre shock)		✓		✓
Observations	575	575	575	575
R-squared	0.216	0.264	0.223	0.279

Notes: The table report results from regression of district-level deposit growth (for 2016 quarter 4) on the density of chest banks in the district. The dependent variable is normalized to have mean zero and standard deviation 1. The independent variables are either one minus the share of chest banks in the district (where share of chest banks is defined as share of deposits in 2015 Q4 in a district held by all branches of banks that operate a currency chest branch in the district), or an indicator function for a low number of chest banks in the districts. Standard error in parentheses; $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 2: High deposit growth (0/1) and district characterisitics (Balance Test)

	(1)	(2)	(3)
Dependent Variable:	**mean**		**difference**
	Treatment	Control	(Treatment-Control)
log(population)	14.000	14.173	-0.172*
	(0.090)	(0.048)	(0.102)
$\log(\text{Deposits})_{pre}$	9.959	10.435	-0.475***
	(0.109)	(0.053)	(0.121)
% villages with phone	0.492	0.486	0.007
	(0.023)	(0.013)	(0.027)
% temporary homes	0.208	0.198	0.010
	(0.018)	(0.010)	(0.020)
% villages with water	0.994	0.994	-0.000
	(0.001)	(0.001)	(0.001)
% villages with school	0.194	0.215	-0.021
	(0.016)	(0.011)	(0.019)
Sex-ratio	923.620	954.402	-30.782***
	(4.615)	(2.874)	(5.427)
Literacy rate	68.857	73.208	-4.351***
	(0.796)	(0.483)	(0.930)
Growth	0.223	0.180	0.043***
	(0.014)	(0.006)	(0.016)

Notes: The table reports the mean of district-characteristics in the treatment and control cells (Column 1 & 2) and their correlation with the treatment variable (Column 3). The treatment variables takes the value of 1 if the deposit growth in the district in the last quarter of 2016 is in the top 20%. Columns (3) report the cofficient of the univariate OLS regression of each variable on the treatment dummy and hence is the difference in mean value of variable between treatment and control groups. Robust standard errors reported in parentheses. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 3: Deposit growth and adoption of electronic payment system

	log(# transactions)				log(amount)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbf{1}_{\{t \geq t_0\}} \times \mathbf{1}_{\{\text{High dep. growth}_{d,t}\}}$	0.928^{***}	0.715^{***}			1.301^{***}	0.864^{**}		
	(0.233)	(0.237)			(0.416)	(0.419)		
$\mathbf{1}_{\{t \geq t_0\}} \times (\text{Deposit growth})_{d,t}$			9.646^{***}	8.355^{***}			17.99^{***}	13.71^{***}
			(1.787)	(1.798)			(3.096)	(3.112)
$\mathbf{1}_{\{t \geq t_0\}} \times (\text{Deposit level})_{d,t}$	1.330^{***}	1.422^{***}	1.427^{***}	1.637^{***}	1.867^{***}	1.347^{***}	2.048^{***}	1.683^{***}
	(0.084)	(0.196)	(0.076)	(0.206)	(0.149)	(0.324)	(0.134)	(0.346)
Month F.E.	✓	✓	✓	✓	✓	✓	✓	✓
District F.E.	✓	✓	✓	✓	✓	✓	✓	✓
District Controls		✓		✓		✓		✓
Observations	5,860	5,770	5,860	5,770	5,860	5,770	5,860	5,770
R-squared	0.489	0.516	0.496	0.520	0.478	0.511	0.485	0.513
Number of districts	586	577	586	577	586	577	586	577

Notes: Difference-in-differences estimates of the effect of the shock on the use of electronic payment system. The estimated specification is equation (1). In Columns (1)-(4), the dependent variable is the log of the total number of transactions carried out using electronic payment system in district d during month t; in Columns (5)-(8), the dependent variable is the log of the total amount (in Rs.) of transactions carried out using electronic payment system in district d during month t. See the main text for a list of controls included in the different specifications. Errors Clustered by District in parentheses. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 4: District adoption rates based on exposure to the aggregate shock in the simulated data

	No complementarities ($C = 0$)	**Complementarities** ($C > 0$)
$\mathbf{1}_{\{d \in \mathcal{T}\}} \times \mathbf{1}_{\{t \geq t_0\}}$	0.116	0.118
	(0.025,0.208)	(0.109,0.126)
Observations per simulation	22,000	22,000
Average R-sq.	0.389	0.688

Notes: This table reports difference-in-differences estimator of the effect of the shock $M_{d,t}$, applied to simulated data. The estimated specification is equation 18. The simulated data is aggregated at the district level and sampled monthly; see text for details. The 95% confidence interval is reported in parentheses.

Table 5: District adoption rates based on initial adoption rate in the simulated data

	No complementarities ($C = 0$)	**Complementarities** ($C > 0$)
$\mathbf{1}_{\left\{X_{0,d} \in \hat{X}_0\right\}} \times \mathbf{1}_{\{t \geq t_0\}}$	−0.009	0.331
	(-0.017,-0.002)	(0.311,0.352)
Observations per simulation	22,000	22,000
Average R-sq.	0.211	0.582

Notes: This table reports difference-in-differences estimator of the effect of initial adoption rates $X_{0,d}$, applied to simulated data. The estimated specification is equation 20. The simulated data is aggregated at the district level and sampled monthly; see text for details. The 95% confidence interval is reported in parentheses.

Table 6: Firms adoption rates in the simulated data

	No complementarities ($C = 0$)	**Complementarities** ($C > 0$)
ρ	0.862	0.863
	(0.861,0.864)	(0.862,0.864)
β	-0.175	-0.177
	(-0.180,-0.170)	(-0.183,-0.171)
γ	-0.016	0.198
	(-0.020,-0.011)	(0.196,0.200)
Observations per simulation	2,100,000	2,100,000
Average R-sq.	0.754	0.837

Notes: This table reports the estimates of the panel data regression model 24 on simualated firm-level data. The coefficient ρ is the autocorrelation to firm's technology choice, $x_{i,d,t}$, while the coefficient β captures the dependence on the stock of money, $M_{d,t-\Delta}$, and the coefficient γ captures the dependence on the existing share of adopters, $X_{d,t-\Delta}$. The simulated data is aggregated at the district level and sampled monthly; see text for details. The 95% confidence interval is reported in parentheses.

Table 7: District adoption rates based on initial adoption in electronic payments data

	log(# transactions)			log(amount)		
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}_{\{t \geq t_0\}} \times$ (Distance to hub)$_d$	-5.736***	-3.666***	-1.791**	-9.201***	-5.111***	-3.194**
	(0.518)	(0.737)	(0.712)	(0.794)	(1.196)	(1.246)
Month & District F.E.	✓	✓	✓	✓	✓	✓
Distance Capital × Month F.E.	✓	✓	✓	✓	✓	✓
State × Month F.E.		✓	✓		✓	✓
District Controls × Month F.E.			✓			✓
Observations	5,800	5,800	5,750	5,800	5,800	5,750
R-sq.	0.585	0.615	0.641	0.520	0.495	0.591
Number of districts	579	579	575	579	579	575

Notes: Difference in differences estimate of the effect of initial conditions, using distance to the nearest hub as a proxy for the initial share of adopters. The specification estimated is equation 27. The dependent variable is either the log of the total number of transactions, or the log of the total nominal value of transactions (plus one in either case). District-level controls are the same as those used in section 3. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 8: Firm adoption based on existing adoption rate in electronic payments data

	(1)	(2)	(3)	(4)
	$x_{i,k,p,t} = \log(\text{amount})_{i,k,p,t}$			
$x_{i,k,p,t-1}$	0.528^{***}	0.437^{***}	0.369^{***}	0.358^{***}
	(0.005)	(0.004)	(0.004)	(0.004)
$X_{k,p,t-1}$	0.090^{***}	0.155^{***}	0.032^{***}	0.015^{***}
	(0.003)	(0.001)	(0.001)	(0.001)
R^2	0.365	0.404	0.455	0.460
	$x_{i,k,p,t} = \log(\#\text{ transactions})_{i,k,p,t}$			
$x_{i,k,p,t-1}$	0.707^{***}	0.617^{***}	0.593^{***}	0.577^{***}
	(0.005)	(0.005)	(0.005)	(0.005)
$X_{k,p,t-1}$	0.032^{***}	0.062^{***}	0.041^{***}	0.017^{***}
	(0.002)	(0.002)	(0.001)	(0.001)
R^2	0.549	0.574	0.601	0.606
	$x_{i,k,p,t} = \mathbf{1}\{\text{On platform}\}_{i,k,p,t}$			
$x_{i,k,p,t-1}$	0.509^{***}	0.404^{***}	0.334^{***}	0.323^{***}
	(0.005)	(0.004)	(0.003)	(0.003)
$X_{k,p,t-1}$	0.046^{***}	0.097^{***}	0.038^{***}	0.022^{***}
	(0.004)	(0.003)	(0.002)	(0.001)
R^2	0.341	0.387	0.443	0.448
Firm F.E.		✓	✓	✓
Industry × Week F.E.			✓	✓
Pincode × Week F.E.				✓
Observations	11,750,558	11,750,558	11,541,757	11,541,757

Notes: The table reports estimates of the dependence of firm-level adoption rates on the share of other adopters in the industry/pincode. The specification we estimate is a version of equation 28 in which each coefficient is interacted with a weekly dummy; we reported estimates of the coefficient γ. The top panel reports effects when x is the total value of transactions, the middle panel reports effects when x is the total number of transactions, and the bottom panel reports effects when x is a dummy for whether the firm used the platform over the past week. Standard errors are clustered at the pincode level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 9: Consumption responses based on exposure to the shock

Panel A: Discrete Measure (0/1)	Total	Necessary	Unnecessary	Frequent	Infrequent
	(1)	(2)	(3)	(4)	(5)
$\mathbf{1}_{\{t=t_0\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0755***	-0.0447	-0.156***	-0.0520	-0.282***
	(0.0257)	(0.0349)	(0.0493)	(0.0352)	(0.0804)
$\mathbf{1}_{\{t=t_0+1\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.00332	-0.0126	-0.0645	-0.0236	-0.105
	(0.0261)	(0.0329)	(0.0525)	(0.0342)	(0.0809)
$\mathbf{1}_{\{t=t_0+2\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0799***	-0.0360	-0.0959**	-0.0366	-0.196**
	(0.0276)	(0.0379)	(0.0478)	(0.0378)	(0.0773)
$\mathbf{1}_{\{t=t_0+3\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0187	-0.0239	0.000509	-0.0203	-0.0354
	(0.0289)	(0.0377)	(0.0618)	(0.0388)	(0.103)
Month F.E.	✓	✓	✓	✓	✓
Household F.E.	✓	✓	✓	✓	✓
Controls × Month F.E.	✓	✓	✓	✓	✓
Observations	852,570	855,288	855,288	855,288	855,288
R-squared	0.018	0.044	0.031	0.043	0.023
# Households	95,032	95,032	95,032	95,032	95,032
Panel B: Continuous Measure	Total	Necessary	Unnecessary	Frequent	Infrequent
	(1)	(2)	(3)	(4)	(5)
$\mathbf{1}_{\{t=t_0\}} \times (\text{Deposit growth})_d$	-0.554***	-0.893**	-1.691***	-0.933**	-2.590***
	(0.186)	(0.452)	(0.527)	(0.466)	(0.813)
$\mathbf{1}_{\{t=t_0+1\}} \times (\text{Deposit growth})_d$	-0.276	0.195	-0.347	0.151	-0.472
	(0.208)	(0.736)	(0.747)	(0.755)	(0.881)
$\mathbf{1}_{\{t=t_0+2\}} \times (\text{Deposit growth})_d$	-0.318	-0.0320	0.00957	-0.00669	-0.0276
	(0.224)	(0.412)	(0.534)	(0.429)	(0.788)
$\mathbf{1}_{\{t=t_0+3\}} \times (\text{Deposit growth})_d$	-0.0507	-0.696	-0.478	-0.667	-0.228
	(0.231)	(0.455)	(0.576)	(0.478)	(0.926)
Month F.E.	✓	✓	✓	✓	✓
Household F.E.	✓	✓	✓	✓	✓
Controls × Month F.E.	✓	✓	✓	✓	✓
Observations	852,570	855,288	855,288	855,288	855,288
R-squared	0.016	0.041	0.029	0.040	0.022
# Households	95,032	95,032	95,032	95,032	95,032

Notes: The table shows difference-in-differences estimate for consumption responses for each month from the demonetization shock (t_0 is November 2016) relative the pre-period (four months). The specification estimated is equation 29. Treatment (I_d) is a dummy that takes the value of one if the district falls in the top decile (household-weighted) of the excess deposit measure (Panel A) and is a continuous measure of district's deposit growth in 2016 final quarter (bottom panel). The dependent variable $\log(y_{h,d,t})$ is either log of total consumption (Column (1)); log of consumption of necessary goods (Column (2)); log of consumption of unnecessary goods (Column (3)); log of consumption of frequent basket (Column (4)); log of consumption of infrequent basket (Column (5)) as defined in Section 6. District-level controls are the same as those used in section 3. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 10: Consumption responses and likelihood to access alternate payment systems

	Necessary	Unnecessary	Frequent	Infrequent
	(1)	(2)	(3)	(4)
$\mathbf{1}_{\{t=t_0\}} \times \mathbf{1}_{\{\text{Top Quartile Income}_h\}}$	0.0318	-0.0598	0.0291	-0.160
	(0.0757)	(0.0843)	(0.0776)	(0.0980)
$\mathbf{1}_{\{t=t_0\}} \times \mathbf{1}_{\{\text{Second Quartile Income}_h\}}$	-0.0483	-0.174***	-0.0546*	-0.319***
	(0.0323)	(0.0505)	(0.0321)	(0.0884)
$\mathbf{1}_{\{t=t_0\}} \times \mathbf{1}_{\{\text{Third Quartile Income}_h\}}$	-0.0610**	-0.178***	-0.0696**	-0.313***
	(0.0304)	(0.0612)	(0.0314)	(0.0958)
$\mathbf{1}_{\{t=t_0\}} \times \mathbf{1}_{\{\text{Bottom Quartile Income}_h\}}$	-0.0729***	-0.184***	-0.0830***	-0.315***
	(0.0249)	(0.0423)	(0.0239)	(0.0900)
$\mathbf{1}_{\{t=t_0\}} \times \mathbf{1}_{\{\text{College Educated}_h\}}$	-0.0221	-0.103	-0.0271	-0.193*
	(0.0541)	(0.0691)	(0.0551)	(0.104)
$\mathbf{1}_{\{t=t_0\}} \times \mathbf{1}_{\{\text{High School Educated}_h\}}$	-0.0203	-0.123**	-0.0266	-0.259***
	(0.0425)	(0.0549)	(0.0427)	(0.0903)
$\mathbf{1}_{\{t=t_0\}} \times \mathbf{1}_{\{\text{None of some education}_h\}}$	-0.0641**	-0.187***	-0.0723**	-0.317***
	(0.0289)	(0.0474)	(0.0292)	(0.0800)
$\mathbf{1}_{\{t=t_0\}} \times \mathbf{1}_{\{\text{Have Credit Card}_h\}}$	-0.0439	-0.0852	-0.0382	-0.176
	(0.0936)	(0.150)	(0.0989)	(0.239)
$\mathbf{1}_{\{t=t_0\}} \times \mathbf{1}_{\{\text{Have no credit card}_h\}}$	-0.0465	-0.160***	-0.0543	-0.287***
	(0.0338)	(0.0485)	(0.0340)	(0.0795)

Notes: The table shows the coefficients on the month of shock (November 2016) for different sub-samples based on difference-in-differences estimating equation 29. The sub-samples are a proxy of likelihood of household to be able to access the electronic payment systems. Treatment (I_d) is the interaction between district's exposure to shock (discrete measure (0/1) as in Panel (a) of Table 9) and different measures of consumer heterogeneity. The top panel is based on quartiles of household income in August 2016. The middle panel is based on education level of head of household in August 2016. The bottom panel is based on a dummy that takes the value of 1 if any member of household owned a credit card in August 2016. District-level controls are the same as those used in section 3. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

A Theory

A.1 The relative value of adoption

Let $t \geq 0$ and define:

$$V_{R,t} = \max(V_{0,t}, V_{1,t}), \tag{30}$$

then the values of a firm with technology $a \in \{0,1\}$ can be written as:

$$\begin{aligned} V_{0,t} &= \mathbb{E}_{t-\Delta}\Big[M_t\Delta + e^{-r\Delta}\left\{\left(1-e^{-k\Delta}\right)V_{R,t+\Delta} + e^{-k\Delta}V_{NA,t+\Delta}\right\}\Big] \\ V_{1,t} &= \mathbb{E}_{t-\Delta}\left[\Big(\underline{M} + CX_t\Big)\Delta + e^{-r\Delta}\left\{\left(1-e^{-k\Delta}\right)V_{R,t+\Delta} + e^{-k\Delta}V_{NA,t+\Delta}\right\}\right] \end{aligned} \tag{31}$$

where the information set for $\mathbb{E}_{t-\Delta}[.]$ is $\{M_{t-\Delta}, X_{t-\Delta}\}$. From this, the value function governing net gains from adoption, W_t, follows:

$$\begin{aligned} W_t &= \mathbb{E}_{t-\Delta}\left[\Big(\left(\underline{M} + CX_t\right) - M_t\Big)\Delta + e^{-(r+k)\Delta}W_{t+\Delta}\right] \\ &= \mathbb{E}_{t-\Delta}\left[\sum_{n\geq 0} e^{-(r+k)\Delta n}\Big(\left(\underline{M} + CX_{t+\Delta n}\right) - M_{t+\Delta n}\Big)\Delta\right]. \end{aligned} \tag{32}$$

The value of adoption is:

$$W_t = \mathbb{E}_{t-\Delta}\left[\sum_{n\geq 0} e^{-(r+k)\Delta n}\Big(\left(\underline{M} + CX_{t+\Delta n}\right) - M_{t+\Delta n}\Big)\Delta\right]. \tag{33}$$

Define:

$$H(M_{t-\Delta}) = \frac{1-e^{-(r+k)\Delta}}{\Delta}\mathbb{E}_{t-\Delta}\left[\sum_{n\geq 0} e^{-(r+k)\Delta n}M_{t+\Delta n}\Delta\right]. \tag{34}$$

We can rewrite the value of adoption as:

$$W_t = \frac{\Delta}{1-e^{-(r+k)\Delta}}\left(\underline{M} - H(M_{t-\Delta})\right) + \Delta C\mathbb{E}_{t-\Delta}\left[\sum_{n\geq 0} e^{-(r+k)\Delta n}X_{t+\Delta n}\right]. \tag{35}$$

Given that the aggregate shock follows the process given by equation (16),

$$\mathbb{E}_{t-\Delta}\left[M_{t+\Delta n}\right] = M_{t-\Delta} + \mu\Delta(n+1), \tag{36}$$

and so:

$$H(M_{t-\Delta}) = M_{t-\Delta} + \frac{\Delta}{1-e^{-(r+k)\Delta}}\mu. \tag{37}$$

Therefore,

$$W_t = \frac{\Delta}{1-e^{-(r+k)\Delta}}\left(\underline{M} - M_{t-\Delta} - \frac{\Delta}{1-e^{-(r+k)\Delta}}\mu\right) + \Delta C\mathbb{E}_{t-\Delta}\left[\sum_{n\geq 0} e^{-(r+k)\Delta n}X_{t+\Delta n}\right]. \tag{38}$$

A.2 Dynamics of the adopter share

The dynamics of the adopter share are:

$$\begin{aligned} X_{t+\Delta n} &= \left(1-e^{-k\Delta}\right) a_{t+\Delta n} + e^{-k\Delta} X_{t+\Delta(n-1)} \\ &= \left(1-e^{-k\Delta}\right) a_{t+\Delta n} + e^{-k\Delta}\left(1-e^{-k\Delta}\right) a_{t+\Delta(n-1)} + e^{-2k\Delta} X_{t+\Delta(n-2)} \\ X_{t+\Delta n} &= \left(1-e^{-k\Delta}\right) \sum_{p=0}^{n} e^{-k\Delta(n-p)} a_{t+\Delta p} + e^{-k\Delta(n+1)} X_{t-\Delta} \end{aligned}$$

Thus we have:

$$\begin{aligned} \mathbb{E}_{t-\Delta}\left[\sum_{n=0}^{+\infty} e^{-(r+k)\Delta n} X_{t+\Delta n}\right] &= \mathbb{E}_{t-\Delta}\left[\sum_{n=0}^{+\infty} e^{-(r+k)\Delta n}\left\{\left(1-e^{-k\Delta}\right)\sum_{p=0}^{n} e^{-k\Delta(n-p)} a_{t+\Delta p} + e^{-k\Delta(n+1)} X_{t-\Delta}\right\}\right] \\ &= \left(1-e^{-k\Delta}\right)\mathbb{E}_{t-\Delta}\left[\sum_{n=0}^{+\infty} e^{-(r+k)\Delta n}\left\{\sum_{p=0}^{n} e^{-k\Delta(n-p)} a_{t+\Delta p}\right\}\right] + \frac{e^{-k\Delta}}{1-e^{-(r+2k)\Delta}} X_{t-\Delta}. \end{aligned}$$

Moreover,

$$\begin{aligned} \mathbb{E}_{t-\Delta}\left[\sum_{n=0}^{+\infty} e^{-(r+k)\Delta n}\left\{\sum_{p=0}^{n} e^{-k\Delta(n-p)} a_{t+\Delta p}\right\}\right] &= \mathbb{E}_{t-\Delta}\left[\sum_{n=0}^{+\infty} e^{-(r+2k)\Delta n}\left\{\sum_{p=0}^{n} e^{k\Delta p} a_{t+\Delta p}\right\}\right] \\ &= \mathbb{E}_{t-\Delta}\left[\sum_{p=0}^{+\infty} e^{k\Delta p} a_{t+\Delta p}\left\{\sum_{n=p}^{+\infty} e^{-(r+2k)\Delta n}\right\}\right] \\ &= \frac{1}{1-e^{-(r+2k)\Delta}}\mathbb{E}_{t-\Delta}\left[\sum_{p=0}^{+\infty} e^{-(r+k)\Delta p} a_{t+\Delta p}\right]. \end{aligned}$$

So:

$$\mathbb{E}_{t-\Delta}\left[\sum_{n=0}^{+\infty} e^{-(r+k)\Delta n} X_{t+\Delta n}\right] = \left\{\frac{1-e^{-k\Delta}}{1-e^{-(r+2k)\Delta}}\mathbb{E}_{t-\Delta}\left[\sum_{n=0}^{+\infty} e^{-(r+k)\Delta n} a_{t+\Delta n}\right] + \frac{e^{-k\Delta}}{1-e^{-(r+2k)\Delta}} X_{t-\Delta}\right\}, \tag{39}$$

and the value of adoption is given by:

$$\begin{aligned} W_t &= \frac{\Delta}{1-e^{-(r+k)\Delta}}\left(\underline{M} - H(M_{t-\Delta})\right) \\ &\quad + C\left\{\frac{1-e^{-k\Delta}}{1-e^{-(r+2k)\Delta}}\sum_{n=0}^{+\infty}\Delta e^{-(r+k)\Delta n}\mathbb{E}_{t-\Delta}\left[a_{t+\Delta n}\right] + \frac{\Delta e^{-k\Delta}}{1-e^{-(r+2k)\Delta}} X_{t-\Delta}\right\}. \end{aligned} \tag{40}$$

The continuous-time ($\Delta \to 0$) limit of this expression is:

$$W_t = \frac{1}{r+k}\left(\underline{M} - H^{(c)}(M_t)\right) + C\left\{\frac{k}{r+2k}\int_{s\geq t} e^{-(r+k)(s-t)}\mathbb{E}_t\left[a_s\right] ds + \frac{1}{r+2k} X_t\right\}, \tag{41}$$

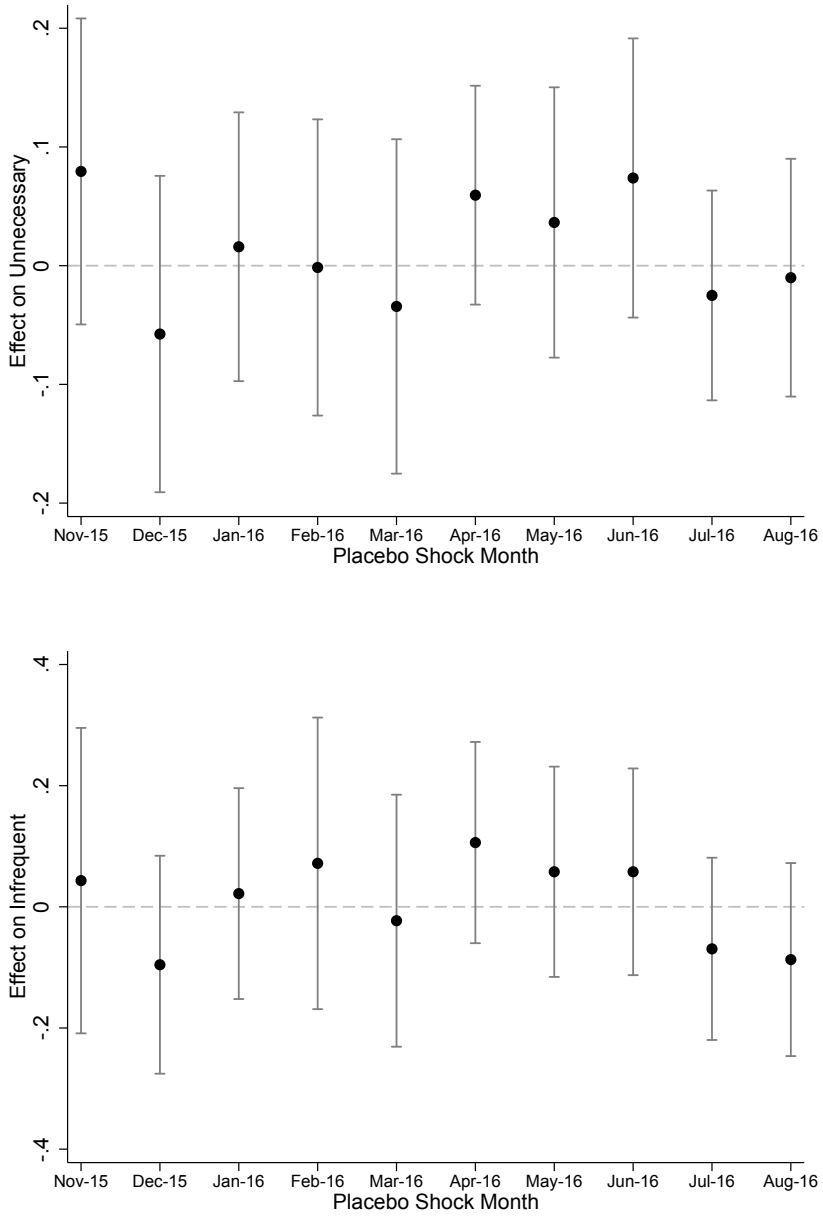
where:

$$H^{(c)}(M_t) = (r+k)\int_{s\geq t} e^{-(r+k)(s-t)}\mathbb{E}_t\left[M_s\right] ds. \tag{42}$$

In particular, since $a_s \geq 0$, the value of adoption can only be independent of X_t in the limit where $k \to +\infty$.

B Figures and Tables

Figure B.1: Consumption responses based on placebo shocks



Notes: The figure plots month-by-month estimates of consumption responses depending on exposure to the shock where we assume the occurrence of a "fake" shock in each month corresponding to each entry on x-axis. The specification we estimate is a version of equation 29 in which each coefficient is based on the interaction of the treatment variable with a monthly dummy. We report the coefficient δ for the month after shock. Treatment is a dummy that takes the value of one if the district falls in the top decile (household-weighted) of the excess deposit measure in the last quarter of 2016. The top panel reports the effects when y is the unnecessary consumption basket, and the bottom panel reports when y is the infrequent consumption basket (as described in Section 5). 95% confidence intervals are represented with the vertical lines; standard errors are clustered at the district code level. Source: CMIE Consumption Data.

Table B.1: Share of Chest Banks and Deposit Growth (pre-shock quarters)

	Deposit Growth						
	2015 Q1	2015 Q2	2015 Q3	2015 Q4	2016 Q1	2016 Q2	2016 Q3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$(1 - \text{share chest banks})_d$	-0.0519	0.0138	0.0640	0.0389	-0.0213	0.00112	-0.0642**
	(0.0443)	(0.0287)	(0.0422)	(0.0244)	(0.0346)	(0.0563)	(0.0293)
District-level Controls	✓	✓	✓	✓	✓	✓	✓
$Log(\text{Total Deposit Pre})_d$	✓	✓	✓	✓	✓	✓	✓
Observations	575	575	575	575	575	575	575
R-squared	0.072	0.086	0.025	0.075	0.113	0.032	0.096

	Deposit Growth						
	2015 Q1	2015 Q2	2015 Q3	2015 Q4	2016 Q1	2016 Q2	2016 Q3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$(\text{Low share of chest banks})_d$	0.0127	-0.0153	-0.0153	0.0122	-0.00234	-0.0297	-0.0246*
	(0.0216)	(0.0162)	(0.0162)	(0.0103)	(0.0159)	(0.0346)	(0.0135)
District-level Controls	✓	✓	✓	✓	✓	✓	✓
$Log(\text{Total Deposit Pre})$	✓	✓	✓	✓	✓	✓	✓
Observations	575	575	575	575	575	575	575
R-squared	0.070	0.087	0.022	0.071	0.113	0.035	0.090

Notes: Regression of district-level deposit growth for all seven quarter before the shock (2016 Q4) on the density of chest banks in the district. The dependent variable is normalized to have mean zero and standard deviation 1. The independent variables are either one minus the share of chest banks in the district (where the share is computed relative to the total number of banks operating in the district) in top panel, or an indicator function for a low number of chest banks in the districts in the bottom panel. Standard error in parentheses; $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table B.2: District adoption rates based on initial adoption in electronic payments data: OLS

	log(# transactions)		log(amount)	
	(1)	(2)	(3)	(4)
$\mathbf{1}_{\{t \geq t_0\}} \times \mathbf{1}_{\{\text{Any adopter}_d\}}$	2.803^{***}		4.864^{***}	
	(0.246)		(0.346)	
$\mathbf{1}_{\{t \geq t_0\}} \times \log(\text{Amount of transactions}_d)$		0.281^{***}		0.230^{***}
		(0.028)		(0.052)
Month F.E.	✓	✓	✓	✓
District F.E.	✓	✓	✓	✓
District Controls		✓		✓
Observations	5,780	5,780	5,780	5,780
R-squared	0.609	0.603	0.578	0.570
Number of districts	578	578	578	578

Notes: The table shows adoption dependence on initial conditions at the district level. The specification estimated is equation 26. In the first line, I_d is a dummy if a district had a positive adoption level before the demonetization. In the second line, I_d the total amount of transactions before the demonetization. The dependent variable is either the log of the total number of transactions, or the log of the total nominal value of transactions (plus one in either case). District-level controls are the same as those used in section 3. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table B.3: District adoption rates based on initial adoption: Alternative specification

	log(# transactions)			log(amount)		
	$\delta = 200$	$\delta = 300$	$\delta = 400$	$\delta = 200$	$\delta = 300$	$\delta = 400$
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}_{\{t \geq t_0\}} \times \mathbf{1}_{\{\text{Distance To Hub} > \delta \text{ km.}\}}$	-0.683**	-0.524**	-0.589***	-0.985**	-0.963**	-1.198***
	(0.302)	(0.224)	(0.209)	(0.495)	(0.393)	(0.385)
Month & District F.E.	✓	✓	✓	✓	✓	✓
Distance Capital × Month F.E	✓	✓	✓	✓	✓	✓
State × Month F.E.	✓	✓	✓	✓	✓	✓
District Controls × Month F.E	✓	✓	✓	✓	✓	✓
Observations	5,750	5,750	5,750	5,750	5,750	5,750
R-squared	0.641	0.641	0.642	0.615	0.615	0.591
Number of districts	575	575	575	575	575	575

Notes: The table shows difference-in-differences estimate of the effect of initial conditions, using distance to the nearest hub as a proxy for the initial share of adopters. The specification estimated is equation 27, replacing D_d with a dummy for distance to hub based on threshold δ ($\mathbf{1}_{\{\text{Distance To Hub} > \delta \text{ km.}\}}$), The dependent variable is either the log of the total number of transactions, or the log of the total nominal value of transactions (plus one in either case). District-level controls are the same as those used in section 3. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table B.4: Firm adoption based on existing adoption rate (district-level)

	(1)	(2)	(3)	(4)
	$x_{i,k,d,t} = \log(\text{amount})_{i,k,d,t}$			
$x_{i,k,d,t-1}$	0.572***	0.474***	0.420***	0.410***
	(0.0100)	(0.0108)	(0.0108)	(0.0105)
$X_{k,d,t-1}$	0.0696***	0.117***	0.0295***	0.00606***
	(0.00257)	(0.00662)	(0.00439)	(0.00134)
R^2	0.398	0.437	0.459	0.463
	$x_{i,k,d,t} = \log(\#\text{ transactions})_{i,k,d,t}$			
$x_{i,k,d,t-1}$	0.776***	0.709***	0.635***	0.624***
	(0.0101)	(0.00933)	(0.0149)	(0.0148)
$X_{k,d,t-1}$	0.0237***	0.0600**	0.116***	0.0212***
	(0.00821)	(0.0301)	(0.00693)	(0.00205)
R^2	0.598	0.615	0.635	0.637
	$x_{i,k,d,t} = \mathbf{1}\{\text{On platform}\}_{i,k,d,t}$			
$x_{i,k,d,t-1}$	0.528***	0.408***	0.378***	0.370***
	(0.00828)	(0.00931)	(0.00857)	(0.00849)
$X_{k,d,t-1}$	0.0158***	0.0314***	0.0198***	0.00489***
	(0.00131)	(0.00180)	(0.00202)	(0.000938)
R^2	0.369	0.419	0.433	0.437
Firm F.E.		✓	✓	✓
Industry × Week F.E.			✓	✓
District × Week F.E.				✓
Observations	58,022,429	58,022,429	58,021,662	58,021,662

Notes: The table reports estimates of the dependence of firm-level adoption rates on the share of other adopters in the industry/district. The specification we estimate is a version of equation 28 at *district-level* in which each coefficient is interacted with a weekly dummy; we reported estimates of the coefficient γ. The top panel reports effects when x is the total value of transactions, the middle panel reports effects when x is the total number of transactions, and the bottom panel reports effects when x is a dummy for whether the firm used the platform over the past week. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table B.5: Firm adoption based on existing adoption rate (allowing for spillovers across industries)

	(1)	(2)	(3)	(4)
	$x_{i,p,d,t} = \log(\text{amount})_{i,p,d,t}$			
$x_{i,p,d,t-1}$	0.533^{***}	0.444^{***}	0.375^{***}	0.358^{***}
	(0.006)	(0.005)	(0.004)	(0.003)
$X_{p,d,t-1}$	0.076^{***}	0.135^{***}	0.023^{***}	0.016^{***}
	(0.002)	(0.002)	(0.002)	(0.001)
R^2	0.364	0.402	0.432	0.441
	$x_{i,p,d,t} = \log(\#\text{ transactions})_{i,p,d,t}$			
$x_{i,p,d,t-1}$	0.711^{***}	0.621^{***}	0.586^{***}	0.579^{***}
	(0.005)	(0.005)	(0.005)	(0.005)
$X_{p,d,t-1}$	0.022^{***}	0.043^{***}	0.021^{***}	0.013^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
R^2	0.548	0.573	0.585	0.590
	$x_{i,p,d,t} = \mathbf{1}\{\text{On platform}\}_{i,p,d,t}$			
$x_{i,p,d,t-1}$	0.496^{***}	0.381^{***}	0.334^{***}	0.323^{***}
	(0.007)	(0.003)	(0.003)	(0.003)
$X_{p,d,t-1}$	0.035^{***}	0.071^{***}	0.027^{***}	0.015^{***}
	(0.002)	(0.001)	(0.001)	(0.001)
R^2	0.347	0.398	0.420	0.428
Firm F.E.		✓	✓	✓
Industry × Week F.E.			✓	✓
District × Week F.E.				✓
Observations	11,750,558	11,750,558	11,750,558	11,749,732

Notes: The table reports estimates of the dynamic specification for adoption based on : $x_{i,p,d,t} = \alpha_i + \alpha_{dt} + \rho x_{i,p,d,t-1} + \gamma X_{p,d,t-1} + \epsilon_{i,p,d,t}$ allowing for spillovers across industries within the same pincode. we reported estimates of the coefficient γ. The top panel reports effects when x is the total value of transactions, the middle panel reports effects when x is the total number of transactions, and the bottom panel reports effects when x is a dummy for whether the firm used the platform over the past week. Standard errors are clustered at the pincode level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table B.6: Consumption responses based on alternative cutoff for exposure to the shock

	Total	Necessary	Unnecessary	Frequent	Infrequent
	(1)	(2)	(3)	(4)	(5)
$\mathbf{1}_{\{t=t_0\}} \times \mathbf{1}_{\{\text{Top Quartile dep. growth}_d\}}$	-0.0454**	-0.0441	-0.0989**	-0.0434	-0.162*
	(0.0199)	(0.0275)	(0.0427)	(0.0282)	(0.0935)
$\mathbf{1}_{\{t=t_0+1\}} \times \mathbf{1}_{\{\text{Top Quartile dep. growth}_d\}}$	-0.0239	-0.0308	-0.0849**	-0.0341	-0.118
	(0.0209)	(0.0303)	(0.0429)	(0.0307)	(0.102)
$\mathbf{1}_{\{t=t_0+2\}} \times \mathbf{1}_{\{\text{Top Quartile dep. growth}_d\}}$	-0.00214	0.0119	0.0206	0.0165	-0.0186
	(0.0215)	(0.0290)	(0.0414)	(0.0298)	(0.0972)
$\mathbf{1}_{\{t=t_0+3\}} \times \mathbf{1}_{\{\text{Top Quartile dep. growth}_d\}}$	0.00332	-0.0218	0.00956	-0.0116	-0.0255
	(0.0246)	(0.0314)	(0.0456)	(0.0325)	(0.109)
Month F.E.	✓	✓	✓	✓	✓
Household F.E.	✓	✓	✓	✓	✓
Controls × Month F.E.	✓	✓	✓	✓	✓
Observations	1,284,702	1,289,304	1,289,304	1,289,304	1,289,304
R-squared	0.022	0.036	0.027	0.035	0.023
# Households	95,032	95,032	95,032	95,032	95,032

Notes: The table shows difference-in-differences estimate for consumption responses for each month from the demonetization shock (t_0 is November 2016) relative the pre-period (four months). The specification estimated is equation 29. Treatment (I_d) is a dummy that takes the value of one if the district falls in the top quartile (household-weighted) of the excess deposit measure. The dependent variable $\log(y_{h,d,t})$ is either log of total consumption (Column (1)); log of consumption of necessary goods (Column (2)); log of consumption of unnecessary goods (Column (3)); log of consumption of frequent basket (Column (4)); log of consumption of infrequent basket (Column (5)) as defined in Section 6. District-level controls are the same as those used in section 3. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table B.7: Consumption responses to the shock based on finer categories

	Food	Recreation	Services	Transportation
	(1)	(2)	(3)	(4)
$\mathbf{1}_{\{t=t_0\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0284	-0.338***	-0.0712*	-0.239**
	(0.0322)	(0.130)	(0.0363)	(0.100)
$\mathbf{1}_{\{t=t_0+1\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0358	-0.373**	-0.0666	-0.310**
	(0.0341)	(0.149)	(0.0422)	(0.121)
$\mathbf{1}_{\{t=t_0+2\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0272	-0.362**	-0.0397	-0.191
	(0.0359)	(0.141)	(0.0459)	(0.128)
$\mathbf{1}_{\{t=t_0+3\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0478	-0.235	0.00346	-0.122
	(0.0394)	(0.156)	(0.0441)	(0.134)
Month F.E.	✓	✓	✓	✓
Household F.E.	✓	✓	✓	✓
Controls × Month F.E.	✓	✓	✓	✓
Observations	855,288	855,288	855,288	855,288
R-squared	0.043	0.021	0.023	0.017
# Households	95,032	95,032	95,032	95,032

Notes: The table shows difference-in-differences estimate for consumption responses for each month from the demonetization shock (t_0 is November 2016) relative the pre-period (four months). The specification estimated is equation 29. Treatment (I_d) is a dummy that takes the value of one if the district falls in the top decile (household-weighted) of the excess deposit measure. The dependent variable $\log(y_{h,d,t})$ is either log of total consumption (Column (1)); log of consumption of food (Column (2)); log of consumption of service goods (Column (3)); log of consumption of transportation goods (Column (4)) as described in Section 6. District-level controls are the same as those used in section 3. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table B.8: Consumption responses to the shock (long-run effects)

	Total	Necessary	Unnecessary	Frequent	Infrequent
	(1)	(2)	(3)	(4)	(5)
$\mathbf{1}_{\{t=t_0\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0726***	-0.0513	-0.159***	-0.0578	-0.309***
	(0.0258)	(0.0346)	(0.0497)	(0.0351)	(0.0822)
$\mathbf{1}_{\{t=t_0+1\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.00187	-0.01000	-0.0581	-0.0179	-0.125
	(0.0266)	(0.0326)	(0.0524)	(0.0335)	(0.0855)
$\mathbf{1}_{\{t=t_0+2\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0763***	-0.0316	-0.107**	-0.0306	-0.247***
	(0.0270)	(0.0378)	(0.0428)	(0.0377)	(0.0798)
$\mathbf{1}_{\{t=t_0+3\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0245	-0.0345	-0.0180	-0.0309	-0.0766
	(0.0283)	(0.0386)	(0.0612)	(0.0393)	(0.104)
$\mathbf{1}_{\{t=t_0+4\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0638**	-0.0730**	-0.115**	-0.0745**	-0.239**
	(0.0265)	(0.0300)	(0.0527)	(0.0294)	(0.102)
$\mathbf{1}_{\{t=t_0+5\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0266	-0.0474	0.0001	-0.0390	-0.0815
	(0.0394)	(0.0364)	(0.0711)	(0.0367)	(0.127)
$\mathbf{1}_{\{t=t_0+6\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0423	-0.0531	-0.0348	-0.0470	-0.164
	(0.0415)	(0.0502)	(0.0767)	(0.0501)	(0.127)
$\mathbf{1}_{\{t=t_0+7\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0328	-0.0228	-0.0443	-0.0185	-0.207
	(0.0475)	(0.0481)	(0.0818)	(0.0482)	(0.134)
$\mathbf{1}_{\{t=t_0+8\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0157	-0.0175	-0.0236	-0.0180	-0.179
	(0.0473)	(0.0533)	(0.0804)	(0.0527)	(0.136)
Month F.E.	✓	✓	✓	✓	✓
Household F.E.	✓	✓	✓	✓	✓
Controls × Month F.E.	✓	✓	✓	✓	✓
Observations	1,797,784	1,803,320	1,803,320	1,803,320	1,803,320
R-squared	0.048	0.043	0.045	0.043	0.042
# Households	95,032	95,032	95,032	95,032	95,032

Notes: The table shows difference-in-differences estimate for consumption responses for each month from the demonetization shock (t_0 is November 2016) relative the pre-period by extending the time-frame of the analysis to include 8 months of pre and post period (instead of 4 months). The specification estimated is equation 29. Treatment (I_d) is a dummy that takes the value of one if the district falls in the top decile (household-weighted) of the excess deposit measure. The dependent variable $\log(y_{h,d,t})$ is either log of total consumption (Column (1)); log of consumption of necessary goods (Column (2)); log of consumption of unnecessary goods (Column (3)); log of consumption of frequent basket (Column (4)); log of consumption of infrequent basket (Column (5)) as defined in Section 6. District-level controls are the same as those used in section 3. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table B.9: Effect of Diwali on consumption (last two years) based on exposure to shock

	Total	Unnecessary	Infrequent
	(1)	(2)	(3)
Diwali 2014:			
$\mathbf{1}_{\{t=\text{Month post Diwali}\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	-0.0001	-0.0114	0.124
	(0.0204)	(0.0566)	(0.132)
$\mathbf{1}_{\{t=\text{Month post Diwali}\}} \times (\text{Deposit growth})_d$	-0.247	-0.964	-0.346
	(0.174)	(0.586)	(0.993)
Month F.E.	✓	✓	✓
Household F.E.	✓	✓	✓
Controls X Month F.E.	✓	✓	✓
Diwali 2015:			
$\mathbf{1}_{\{t=\text{Month post Diwali}\}} \times \mathbf{1}_{\{\text{High dep. growth}_d\}}$	0.0182	-0.0423	0.107
	(0.0233)	(0.0834)	(0.132)
$\mathbf{1}_{\{t=\text{Month post Diwali}\}} \times (\text{Deposit growth})_d$	-0.249	-0.356	4.325**
	(0.411)	(1.106)	(2.153)
Month F.E.	✓	✓	✓
Household F.E.	✓	✓	✓
Controls X Month F.E.	✓	✓	✓

Notes: The table test for effect of festival of Diwali on difference-in-differences estimate for consumption responses for each month from the Diwali month relative the four month before the shock. The specification estimated is equation 29, but by replacing the time of shock as one month before Diwali in 2014 and 2015. Treatment (I_d) is a dummy that takes the value of one if the district falls in the top decile (household-weighted) of the excess deposit measure or is a continuous measure of district's deposit growth in 2016 final quarter. Both the measures are based on 2016 Q4 (the quarter of actual demonetization shock). The dependent variable $\log(y_{h,d,t})$ is either log of total consumption (Column (1)); log of consumption of unnecessary goods (Column (2)); log of consumption of infrequent basket (Column (4)) as defined in Section 6. District-level controls are the same as those used in section 3. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.