Shocks and Technology Adoption: Evidence from Electronic Payment Systems *

Nicolas Crouzet, Apoorv Gupta and Filippo Mezzanotti†

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

Abstract

We study the diffusion of technologies subject to positive adoption externalities. Using a dynamic technology choice model, we argue that, in the presence of externalities, large, temporary shocks can lead to persistent waves of adoption. In line with this mechanism, we show that the Indian demonetization of 2016 — a large, temporary reduction in the availability of currency — caused a persistent move to electronic payment networks. However, we also show that the response exhibited substantial state-dependence: persistent adoption responses occurred in areas where adoption externalities prior to the shock were likely to be stronger, consistent with the model's predictions. Thus, while temporary interventions can have persistent *average* adoption effects, they also have the potential to exacerbate initial differences. We also provide evidence that the adoption wave did not fully offset the effects of the cash crunch on consumption.

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

*We thank Pat Akey, Anthony DeFusco, Sabrina Howell, Ravi Jagannathan, Dean Karlan, Josh Lerner, Konstantin Milbradt, Avri Ravid, George Siddharth, and participants at the 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 crucial component of the link from innovation to growth is the diffusion of new technologies (Hall and Khan, 2003). 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 (Munshi, 2004; Conley and Udry, 2010; Gupta et al., 2019), to lack of financial development (Comin and Nanda, 2019).

For technologies characterized by positive externalities — that is, when the benefits of adoption increase as the use of the technology becomes more widespread —, coordination problems can be a key obstacle to diffusion. Individual firms may expect either high or low adoption rates by other firms. Consistent with these expectations, they may either adopt or reject the technology, giving rise to multiple possible equilibria.[1] In these situations, understanding how firms may coordinate on technology adoption, or fail to, is an important question, for both research and policy.

In this paper, we study the extent to which large economic shocks (or policy interventions) can help resolve these coordination problems and accelerate the pace of technology diffusion. We analyze a simple dynamic technology adoption model where, because of positive externalities, firms' adoption decisions are complements. Using a novel empirical setting, we show that – consistent with the model – a *temporary* aggregate shock can lead to a *persistent* wave of technology adoption. However, we also show that responses to the shock exhibit strong state-dependence. Persistent adoption waves only occur when pre-shock externalities are large; otherwise, adoption shifts are only as persistent as the shock itself. Thus, our results show that while large, temporary shocks can lead to persistent changes in the *average* pace of technology diffusion, they can also tend to exacerbate *differences* in adoption in the long run.

Our analysis focuses on a particular technology: electronic payment systems. Electronic payment systems are an example of a technology exhibiting network externalities (Katz and Shapiro, 1994; Rysman, 2007). The returns to opting into a payment system for a particular firm depends on the size of the network, that is, the number of customers using the system; in turn, the number of customers depends on how many firms accept the payment system in their operations. This makes the decisions to opt in complements across firms, thus generating the type of coordination problems in adoption described above.

The setting for our analysis is the Indian demonetization of 2016. On November 8th, 2016, the Indian

[1] Technologies characterized by externalities are very common, in particular in the new economy, where many products are network based or structured as multi-sided markets. In their classic work, Katz and Shapiro (1985) discusses several sources of externalities. In particular, they highlight how externalities can arise both directly – in situations where the number of users affect the quality of the product – or indirectly – where the number of users affect the value of other add-on products (e.g. hardware/software) or the type of postpurchase services (e.g. car). Furthermore, for very new products, externalities can also arise from learning about the quality of the products and understand its costs and benefits (Suri, 2011).

Government announced that it would void the two largest denominations of currency in circulation and replace them with new bills. At the time of the announcement, the voided bills accounted for 86.4% of total cash in circulation. The public was not given advance warning, and the bills were voided effective immediately. A two months deadline was announced for exchanging the old bills for new currency. In order to do so, old bills had to be deposited in the banking sector. However, withdrawal limits, in combination with frictions in the creation and distribution of the new bills, meant that immediate cash withdrawal was constrained. As a result, cash in circulation fell and bank deposits spiked. Cash transactions became potentially harder to conclude, but more funds were available for use in electronic payments. Importantly, this shock was very large in size but temporary. In first approximation, things improved dramatically during the month of January and the situation normalized starting in February.[2]

We first show that the demonetization led to a large increase in the use of electronic payment systems. 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 then transferred to merchants' bank accounts via the app.[3] In aggregate, the overall activity in the platform roughly doubled in size several times during the two months following the announcement. Additionally, we show that adoption was *persistent*, though the shock was not. In aggregate, there was no significant mean-reversion in adoption or transaction volumes once cash withdrawals constraints were lifted. Aggregate adoption effects are also visible in the use of debit card payment terminals, but appear much weaker for credit card payments and mostly driven by the intensive margin. Overall, the aggregate data thus suggests that the shock led to a wave of adoption of electronic payments.

In order to shed light on the role of externalities in the transmission of this shock to adoption decisions, we start by analyzing a dynamic technology adoption model. Our model is based on the framework of Burdzy et al. (2001). Firms face a choice between two technologies under which to operate, one of which is subject to positive externalities — the flow profits from operating under this technology increase with its rate of use by firms overall. Moreover, the relative benefit of adopting the technology with externalities is subject to aggregate shocks.[4] We show that in response to a large, *temporary* shock to the relative value of

[2]In the paper, we show that cash in circulation restarted to grow vigorously with the new year and we highlight how Government removed all limitations by the end of January. Furthermore, the new evidence on search data confirms the timing.

[3]The costs associated with the adoption of this technology for merchants are small; there are no usage fees, and all that is required to join the platform is to have a bank account and a mobile phone, both of which had high ownership rates in India by 2016(Agarwal et al., 2017). Nevertheless, in what follows, we discuss the role of fixed costs, and argue that they are unlikely to give a full account of the transmission of the shock to adoption.

[4]The presence of these common shocks helps eliminate the potential equilibrium multiplicity arising from complementarities in adoption decisions. The model 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

the two technologies, the total number of users increases *persistently*. Moreover, a distinctive implication of externalities is that the number of *new* users joining the platform remains higher than the pre-period, *even* after the shock has receded.[5] This reflects the fact that, with externalities, the initial increase in adoption triggered by the shock increases the relative value of adoption for other firms; this "snowball" effect can thus generate endogenous persistence in the response of adoption.

Aside from the persistent rise in new users, a key implication of the model is that adoption responses exhibit state-dependence. Specifically, the adoption response to a given shock depends positively on the level of adoption prior to the shock. When the initial adoption rate is low, a large shock may temporarily raise adoption. However, once the shock is undone, the adoption rate will tend converge back to its initially low level. By contrast, when the initial adoption rate is sufficiently high, the same transitory shock may lead to full and permanent adoption. In the model, the initial adoption rates fully capture the initial strength of externalities in adoption; the key prediction is thus that the pre-shock strength of externalities determines the "tipping point" beyond which the transitory shock can generate endogenous persistence in adoption.

We then turn to the electronic data in order to carefully document the dynamics of adoption of electronic wallet technologies by firms during the Indian demonetization. As a first step, we estimate the causal impact of the cash contraction on adoption activity at district-level. This panel analysis allows us to overcome some of limitations of the aggregate event-study, in particular when looking outside the immediate reaction to the shock. We show that variation across districts in the importance of chest banks — local branches in charge of the distribution of new bills — can be exploited to identify heterogeneity in the exposure to the shock at local-level. Using this design, we show that districts more exposed to the demonetization experienced a larger and more persistent increase in adoption following the demonetization, consistent with the predictions of the model. Additionally, higher exposure also predicts a bigger increase in the number of new firms joining the platform, and this is true even after the cash crunch has receded and restrictions on cash withdrawals have been lifted. In light of the theoretical model, this latter result can be only rationalized in the context of a model where complementaries in adoption are important. The paper also carefully discusses several tests that support the causal interpretation of these results.

Then, we test the model's predictions regarding state-dependence in the data, and find strong support for them. As a first step, we provide *firm-level* evidence on state-dependence. 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 the firm level.[6] Overall,

a dynamic coordination problem, instead of a one-shot, static model. The latter difference is important, as it allows us to distinguish between short- and long-run effects of the shock. See Burdzy et al. (2001) for a more detailed discussion of the relationship of this framework with the global games literature.

[5] We show, in particular, that this prediction would not obtain if the main barrier to adoption were fixed costs.

[6] The analysis is conducted at pincode level, which is a geographical unit in India similar to a 5-digit zipcode in the US.

we find that adoption by neighboring firms in the same area has a strong, positive effect on the usage of the technology. In line with the prediction of the model, this effect is particularly larger during the shock period, when firms transition to the new technology. Importantly, these results do not simply capture variation across locations and industries, since they hold when we augment the analysis with a wide set of zipcode by time and industry by time fixed-effects.

Next, we provide also evidence in line with state-dependence at the *district level*. As a preliminary step, we show that districts with larger initial adoption responded relatively more, in line with the model's predictions. Furthermore, we test whether the rise in adoption during the demonetization period was stronger in areas that were located closer to those districts that already had very high adoption level before the shock (*hubs*). The idea behind this test is that being located close to these electronic payments hubs increases the strength of externalities prior to the shock, since customers are more likely to travel across closely located districts. Building on this intuition, we exploit within-state variation in distance, and we show that the increase in adoption is larger in closer districts. Importantly, this result is not driven by differential trends and hold controlling for a variety of district-level observable characteristics.

The evidence of state-dependence in adoption is not only important from a positive standpoint; it also has implications for understanding whether, and how, large policy interventions have the potential to shape technology choices.[7] Indeed, our analysis of the model shows that the particular form of state-dependence which we highlight disappears when interventions are sufficiently persistent.[8] That is, sufficiently persistent interventions will trigger full adoption regardless of initial conditions. This highlights a potential trade-off when designing such interventions: on the one hand, very persistent interventions may be costly or distortionary; on the other, very temporary interventions will tend to generate state-dependent responses, and thus accentuate initial differences in adoption.

We conclude by highlighting the fact that despite the surge, electronic payment systems did not offset the effects of the cash crunch on consumption. Using novel household data, combined with the chest bank exposure measure described above, we show that total consumption contracted relatively more in more exposed districts. In particular, a standard-deviation increase in the exposure to the shock corresponds to an almost 4% larger decline in consumption. However, we also show that this contraction was completely temporary and larger in the subset of the consumption basket that is more likely to be unncecessary.

The rest of the paper is organized as follows. Section 2 provides some background on the demonetization

Section 4 provides more details on pincodes in India.

[7]From a positive standpoint, in our analysis of the model, we show that the particular form of state-dependence we document in the data, where larger pre-shock externalities imply bigger responses, cannot be generated when fixed costs are the main drivers of adoption decisions.

[8]In particular, the intervention needs to be more persistent than the typical adjustment speed of firms. Section 3 discussed this in more detail.

and documents aggregate adoption effects. Section 3 analyzes our simple model of technology adoption and derives key predictions. Section 4 tests these predictions in the electronic wallet data. Section 5 documents consumption responses to the shock, and section 6 concludes.

1.1 Contribution to the literature

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 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 (2019) emphasize the role of financial development for fostering technological diffusion.[9] 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.[10] 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 also contributes to the growing literature of fintech. Despite the importance of payment technologies in this industry, a large part of the literature in this area has focused on the impact of fintech in funding markets, either for households or firms (Buchak et al., 2018; Tang, 2018; Fuster et al., 2018; de Roure et al., 2018; Howell et al., 2018). In this area, our paper provides a general framework to understand how adoption of fintech technologies can be affected by policy shocks. While the model can easily be applied to technologies outside the fintech sphere, the data application developed in this paper suits particularly well this industry, where new products are usually characterized by low adoption costs and strong externalities in adoption. Furthermore, our paper clearly highlights how fintech can add value to companies and consumers in electronic payments by lowering adoption costs. In particular, we have shown that adoption costs of traditional payment technologies were a real constraint for firms during the demonetization. This evidence is consistent with other works in the area (e.g. Yermack (2018), Suri and Jack (2016)), that have highlighted the important role that better payment technologies can have in improving economic conditions in poor

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

[10]This evidence is also consistent with the contemporaneous paper by Higgins (2019), which explores a policy change in Mexico and studies how a permanent increase in debit card usage affects both the supplier and consumer response in the local market.

countries. This discussion is also particularly relevant given recent debate on the costs of cash and cash alternatives in modern economies (Rogoff, 2017).

Lastly, our paper also contributes to the understanding of the impact of the demonetization on the Indian economy. On the one hand, our paper show that the large policy shock had some important positive effects by causing a persistent increase in the adoption of new payment technologies. In this dimension, the work closest to us is Agarwal et al. (2018) which – leveraging mostly on time-series variation around the demonetization – documents a general shift into electronic payments. Relative to this work, our paper provides novel cross-sectional evidence on the way the propagation of this technology happened during the demonetization period. This evidence allows us to provide a novel characterization of adoption dynamics highlighting the importance of state-dependence. Furthermore, our paper shows that complementarity in adoption was a key feature of the technology in explaining the aggregate shift. On the other hand, our paper also documents the negative effects of this policy, in the form of a reduction of consumption by households in response to the scarcity of cash in the local markets. These results are consistent with the effects on Indian output estimated by the contemporaneous paper by Chodorow-Reich et al. (2018). While our paper differs greatly in the main focus, we believe that our household-level analysis also significantly extends the understanding of the real effect of the shock by providing evidence on the heterogeneity of the effects across different consumption groups and areas.

2 Background

2.1 The demonetization

On November 8, 2016, at 08:15 pm IST, Indian Prime Minister Narendra Modi announced the demonetization of Rs.500 and Rs.1,000 notes, during an unscheduled live television interview. The announcement was accompanied by a press release from the Reserve Bank of India (RBI), which stipulated that the two notes would cease to be legal tender in all transactions at midnight on the same day. The voided notes were the largest denominations at the time, and together accounted for 86.4% of the total value of currency in circulation. The RBI also specified that the two notes should be deposited with banks before December 30th, 2016. Two new bank notes, of Rs.500 and Rs.2,000, respectively, were to be printed and distributed to the public through the banking system. The policy's stated goal was to identify individuals holding large amounts of "black money," and remove fake bills from circulation.[11]

However, the swap between new and old currency did not occur at once: instead, the public was unable

[11] In its annual report for 2017-2018, the RBI reported that 99.3% of the value of voided notes had been deposited in the banking system during the demonetization.

to withdraw cash at the same rate as it was depositing old notes. As a result, the amount of currency in circulation dropped precipitously during the first two months of the demonetization period. This can be seen in Figure 2, which plots the monthly growth rate of currency in circulation.[12] Overall, cash in circulation declined by almost 50% during November, and continued declining in December.

Partly, this cash crunch reflected limits on cash withdrawals put in place by the RBI in order to manage the transition.[13] But the cash crunch also reflected the difficult logistics of the swap itself. In order to ensure that the policy remain undisclosed prior to its implementation, the RBI had not printed and circulated large amounts of new notes to banks. This caused many banks to be unable to meet public demand for cash, even under the withdrawal limits. The public rushed to obtain the available cash, resulting in long waiting lines at banks.[14]

Importantly, the demonetization did not lead to a reduction in the total money supply, defined as the sum of cash and bank deposits. The total money supply was stable over this period, as reported in Figure 2. In its press release, the RBI highlighted that deposits to bank accounts could be freely used through "various electronic modes of transfer". The public was thus still allowed to transact using any form of non-cash payment, such as cards, cheques, or any other electronic payment method; cash transactions were the only ones to be specifically impaired.[15]

Despite its magnitude, the cash crunch was a temporary phenomenon. Overall, things significantly improved in January and essentially normalized in February. Cash in circulation grew significantly again in January 2017, suggesting that the public was able to withdraw cash from banks (see Figure 2). Furthermore, by January 30th, 2017, the Government lifted most remaining substantial limitations on cash withdrawals.[16] Consistent with the brief disruption period, the general perception of the negative consequences of the demonetization on payment systems significantly improved with the new year (see Figure 3).[17]

[12]The time series for currency in circulation reported in this graph does not mechanically drop with the voiding of the two notes; it only declines as these notes are deposited in the banking sector.

[13]In its initial press release, the RBI indicated that over the counter cash exchanges could not exceed Rs.4,000 per person per day, while withdrawals from accounts were capped to Rs.20,000 per week, and ATM withdrawals were capped to Rs.4,000 per day per card, for the days following the announcement. Additionally, a wide set of exceptions were granted, including for fuel pumps, toll payments, government hospitals, and wedding expenditures. Banerjee et al. (2018) discuss the uncertainty surrounding withdrawal limits and exceptions, and argue that they may have exacerbated the overall confusion during this transition period.

[14]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 (Krishnan, 2017). Another randomized survey conducted over nine districts in Indian by a mainstream newspaper, Economic Times, showed that the number of visits to either a bank or an ATM increased from an average of 5.8 in the month before demonetization to 14.4 in the month after demonetization (`https://economictimes.indiatimes.com/news/politics-and-nation/how-delhi-lost-a-working-day-to-demonetisation/articleshow/56041967.cms`).

[15]See Chodorow-Reich et al. (2018) for a discussion of the RBI's liabilities, and of key policy and market rates, during the demonetization period.

[16]The limits were progressively relaxed after the announcement; in January 2017, the limits for ATM withdrawals were set to Rs.10,000 per day, and by February 2017, the limits on withdrawals from bank accounts had been raised to to Rs.50,000 per week. By March 2017, all limits on withdrawals from savings accounts had been removed.

[17]Figure 3 reports the monthly plot (09/2016 to 07/2017) of Google Searches for several key words that could be associated with the shock. Data is obtained by Google Trends, and the index is normalized by Google to be 0 to 100, with value of 100 assigned to the day with maximum searches made for that topic. Across all the panels, we find that Google searches connected

The demonetization thus had three key features relevant to our analysis. First, it led to a significant contraction of cash in circulation. Second, it did not change the total money stock, that is, the sum of cash and deposits. As a result, the public could still access and use money electronically once notes had been deposited. Third, it was short-lived: the cash shortage was particularly acute in November and December, but quickly normalized with the new year.

2.2 The adoption of electronic payment technologies

The demonetization was associated with a large uptake in electronic payments, which is visible in measures of both intensive and extensive margin use of various specific payment technologies.

We start by illustrating this using data from one of the leading digital wallet companies in the country. The company allows individual and businesses to undertake transactions with each other using only their mobile phone. In order to use the service, a customer needs to download an application, and link their bank account to the application. Merchants can then use a uniquely assigned QR code to accept payments directly from the customers into a mobile wallet. The contents of the mobile wallet can then be transferred to the merchant's bank account.[18]

Figure 4 reports data for the total number and the total value of transactions executed by merchants using this technology around the week in which the demonetization was announced and implemented.[19] In the months before the demonetization, the weekly growth in the usage of the wallet technology 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, in the week after the demonetization the number of transactions grew by more than 150%, while the value of transactions increased by almost 200%. Furthermore, for the whole month after the shock, weekly growth rates were consistently around 100%.

One important observation is that this initial positive effect of demonetization on adoption did not dissipate with time, even when cash availability constrains were relaxed. In other words, this evidence suggests that a temporary shift in the availability of cash led to a permanent increase in the usage of the platform. In particular, the data suggests a slow-down in aggregate growth starting in January, which is when limits on the circulation of new cash start being relaxed. However, after a small negative adjustment

---

with the demonetization spiked in November, remained high in December, but then significantly dropped in January before returning to the pre-shock levels in February. One exception is the search on "ATM Cash withdrawal limit today" which reaches it maximum on January 31, 2017. This is consistent with the fact that January 31, 2017 was the date when most limits on ATM withdrawals were lifted by the RBI.

[18]There are multiple ways to transact on the digital wallet. First, customers can scan the merchants' unique QR code on the application installed on their smart-phones to complete the transaction. Second, customers can enter the mobile number of the merchant, upon which the merchant receives a unique code which the customer then uses as confirmation to complete the transaction. Third, in the absence of smart-phones or mobile internet availability, customers can call a toll-free number and ask the wallet company to make a transaction using the cellphone number of the merchant. To avail this facility on an ongoing basis, customers needed to be enrolled through a one-time verification process.

[19]We describe and analyze the disaggregated data underlying these graphs in more detail in section 4.

in early February, the average growth rate over the next two months remained on average small but positive, confirming that users did not abandon the platform as cash became widely available again.[20]

Aside from this fintech platform, more traditional electronic payment technologies were also available to the public. To explore traditional electronic payment methods, we collect publicly available data on debit and credit cards activity aggregated at the national level and are reported at the monthly frequency by the RBI.[21] Figure 5 presents these data. In particular, the first panel reports the growth in the number of transactions for both credit and debit cards, across ATM and point of sales (stores). In the second panel we report the growth in the number of cards, again divided across debit and credit cards.

Two findings are important to highlight. First, the permanent increase in electronic payments is not unique to electronic wallet technologies. In particular, the number of transactions at point of sales increased dramatically in both November and December, before going back to a similar pre-shock trend in January. This evidence suggests that the demonetization also led to a permanent increase on the transactions undertaken with debit cards. Second, the short-run increase is completely driven by the intensive margin, unlike for the electronic wallet. In other words, the overall volume in debit card transactions increased only because debit card holders started to use them more frequently. In particular, in the second panel of Figure 5, there is little growth in the number of debit cards during either November and December.

Here, it is worth highlighting the differences between electronic wallets and debit or credit cards. Relative to cards, adoption costs for the wallet are much lower, in particular for merchants, since they can access and use the platform almost instantaneously, with nothing more than a phone and a bank account. Furthermore, for small and medium-sized merchants — who make up the bulk of our data — this technology does not imply any direct monetary cost. [22] Higher adoption costs of cards for merchants are consistent with some of the empirical patterns reported in Figure 5. In particular, we find evidence of a delayed increase in the number of debit cards, starting in January 2017 (two months after the demonetization). We also find essentially no change in the use of credit cards, at both the intensive and extensive margin, which may reflect even higher adoption costs.[23]

Overall, the aggregate data on both electronic wallets and debit or credit cards indicate that the demonetization was associated with a large take-up in electronic payment systems. Moreover, the use of these

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

[21]We obtain this data from: `https://rbi.org.in/scripts/atmview.aspx`, which reports monthly data at the bank level on number of debit cards and credit cards outstanding; the number and amount of transactions made using each system; and the source of transactions (at ATM or point of sales).

[22]Merchants using the digital wallet are classified by the provider into three segments: small, medium and large. Small merchants have lower limits on amount they can transact and pay 0% transaction costs. Medium merchants can transfer money to their bank account at midnight every day up to a certain limit. Large merchants can transact any amount but pay a percentage of transfer amount as a fees. Our data only covers small and medium merchants.

[23]In this dimension, our setting is very different from Higgins (2019), which instead studies a technology — debit cards — which still requires, for merchants, a large set-up cost as well as regular fees.

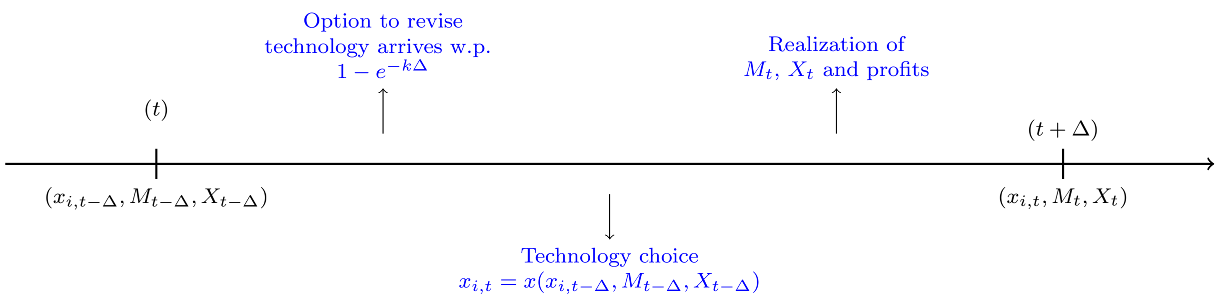


Figure 1: Timing of actions and events during a period.

payment systems by merchants persisted beyond the period of the cash crunch. Consistent with the view that credit and debit cards are subject to larger adoption and usage costs for merchants than electronic wallets, adoption effects for these technologies are more muted than for the electronic wallet, which will be the main focus the rest of our analysis.

3 Theory

In this section, we analyze a simple model of technology adoption. The model nests two mechanisms that could account for the adoption wave documented in the previous section. The first is the existence of externalities across firms in the use of the new technology. The second is the presence of fixed costs associated with switching from the old to the new technology. Additionally, the model is fully dynamic. The model is a variant of Frankel and Pauzner (2000) and Burdzy et al. (2001), in which firms face fixed costs of adjusting the technology under which they operate.

The purpose of the model is twofold. First, it allows us to general testable predictions regarding the characteristics of adoption dynamics in the presence of externalities. Second, it allows us to analyze the role of shock persistence, and in particular how it interacts with externalities to generate state-dependence in adoption.

3.1 Model

Economic environment Time is discrete: $t = 0, \Delta, 2\Delta, ...$ There is a collection of infinitely-lived firms, indexed by $i \in [0, 1]$, who are risk-neutral and discount the future at rate e^{-r}; we call this group of firms a "district", by analogy with our empirical setting in the following sections. At different points in time, firm i must choose between operating under one of two technologies, $x_{i,t} \in \{c, e\}$, where c stands for "cash" and e stands for "electronic money," or "e-money," for short. Flow profits of the firm depend on the technology it uses; we assume that they are given by:

$$\Pi(x_{i,t}, M_t, X_t) = \begin{cases} M_t & \text{if } x_{i,t} = c, \\ \\ M^e + CX_t & \text{if } x_{i,t} = e, \end{cases}$$

where:

$$M^e > 0, \quad C \geq 0, \text{ and } X_t \equiv \int_{i \in [0,1]} \mathbf{1}\{x_{i,t} = e\}\, di. \tag{1}$$

This assumption says that flow profits to technology e are increasing in the number of other firms in the district also using e. The parameter $\frac{\partial \Pi}{\partial X_t} = C > 0$ controls the strength of this effect. The assumption that $C > 0$, that is, the presence of increasing external returns to e, is what will generate complementarities in the decision to adopt e.

On the other hand, flow profits to technology c are exogenous, and subject to shocks. These shocks are common to all firms in the district. For simplicity, we refer to M_t as "cash", though it may be thought of as capturing, more broadly, cash-based demand in the district. We assume that cash follows and AR(1) process:

$$M_t \quad = (1 - e^{-\theta\Delta})M^c + e^{-\theta\Delta}M_{t-\Delta} + \sqrt{\Delta}\sigma\epsilon_t, \quad \epsilon_t \sim N(0,1), \text{ i.i.d.} \tag{2}$$

where M^c is the long-run mean of M_t, σ is standard deviation of innovations to M_t, and the parameter θ captures the speed of mean-reversion of the shock.

There are two frictions that might prevent switching between technologies. First, during each increment of time Δ, a mass $1 - e^{-k\Delta} \in [0,1]$ of firms is able to revise their technology choice. This "technology adjustment" shock is purely idiosyncractic, and in particular, it arrives independently the common shock. When $k \to +\infty$, firms can continuously adjust their technology choices, while when $k = 0$, they are permanently locked into their initial choice. We will assume $0 < k < +\infty$, that is, sluggish adjustment.

Second, there are fixed (pecuniary) costs of adopting the technology e. Specifically, a firm must pay a fixed cost κ if it decides to revise its technology from cash to electronic payments. There is no cost of switching from electronic payments to cash, and no cost of staying with the same technology.

The timing of actions within period t is depicted in Figure **??**. 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 is thus only $\{x_{i,t-\Delta}, M_{t-\Delta}, X_{t-\Delta}\}$.

Technology choice Let $V(x_{i,t}, M_{t-\Delta}, X_{t-\Delta})$ be the value of a firm *after* any potential technology revisions, and define:

$$B(M_{t-\Delta}, X_{t-\Delta}) = V(e, M_{t-\Delta}, X_{t-\Delta}) - V(c, M_{t-\Delta}, X_{t-\Delta}).$$

This is relative value of having technology e in place. Appendix A shows that it follows:

$$B(M_{t-\Delta}, X_{t-\Delta}) = \mathbb{E}_{t-\Delta}\left[(\Pi_t^e - \Pi_t^c)\,\Delta + e^{-(r+k)\Delta}B(M_t, X_t) + e^{-r\Delta}(1-e^{-k\Delta})g(B(M_t, X_t))\right] \tag{3}$$

where $\Pi_t^e = \Pi(e, M_t, X_t)$, $\Pi_t^c = \Pi(c, M_t, X_t)$, and $g(B) = \max(0, \min(B, \kappa))$. When there are no fixed costs of switching, $\kappa = 0$, we have $g(B) = 0$. In this case, $B(.,.)$ is simply the expected present value of $\Pi_t^e - \Pi_t^c$, the difference in cash flows from switching from cash to electronic money. With fixed costs, $g(B) \geq 0$; in that case, $g(B)$ captures the relative value of already having technology e, for a firm which receives the technology adjustment shock.

The technology adoption rule for adjusting firms is given by:

$$x(x_{i,t-\Delta}, B_{t-\Delta}) = \begin{cases} c & \text{if } B_{t-\Delta} \leq 0 \\ x_{i,t-\Delta} & \text{if } B_{t-\Delta} \in [0, \kappa] \\ e & \text{if } B_{t-\Delta} > \kappa \end{cases}, \tag{4}$$

where $B_{t-\Delta} = B(M_{t-\Delta}, X_{t-\Delta})$. In particular, firms remain locked in their prior technology choice in the inaction region $B_{t-\Delta} \in [0, \kappa]$. Define $a_{c\to e,t} = \mathbf{1}\{x(c, B_{t-\Delta}) = e\}$ and $a_{e\to e,t} = \mathbf{1}\{x(e, B_{t-\Delta}) = e\}$. Since the arrival of the option to revise is independent of the current technology choice, the change in the number of firms using technology e, $\Delta X_t \equiv X_t - X_{t-\Delta}$, is given by:

$$\Delta X_t = \left(1 - e^{-k\Delta}\right)(1 - X_{t-\Delta})a_{c\to e,t} - \left(1 - e^{-k\Delta}\right)X_{t-\Delta}a_{e\to e,t}. \tag{5}$$

The first term captures changes due to firms adopting e, while the second term captures changes due to firms abandoning e in favor of c.

Equilibrium An equilibrium of the model is a technology choice rule, x, mapping $\{c, e\} \times \mathbb{R} \to \{c, e\}$, and a function for the gross adoption benefit, B, mapping $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$, such that the technology choice rule and the gross adoption benefit solve the system of equations (3)-(4) when X_t follows the law of motion given by (5), and cash follows the law of motion in (2).

Discussion of key assumptions There are three key assumptions in this model. First, the technology e features positive external returns with respect to adoption by other firms in the industry. We do not provide a precise microfoundation for these external returns, but instead focus on their implications for adoption. Nevertheless, this assumption could capture, for instance, external returns arising in a two-sided market, where a high level of adoption among firms incentivizes customers to adopt the platform, and conversely, a high participation by customers on the platform raises the benefits of adoption for firms. The particular form dependence of profits on X_t assumed in this paper for profits can be thought as a reduced-form way of capturing the external returns that would likely arise in equilibrium in that more complicated model. Alternatively, external returns could arise from spillovers across firms in learning how to use the technology.

The second key assumption is that there are fixed costs in the adoption of the technology. While, for reasons we discussed above, we view pecuniary adoption costs of the specific technology we study in this paper as likely to be low, introducing them in the model allows us to broadly discuss the distinguishing empirical features of complementarities in technology choice relative to fixed costs, which may be more relevant in other settings.

The third key assumption is that firms are unable to continuously adjust their technology choice, but instead must wait, on average, $1/k$ periods before being able to re-optimize their choice (subject to the fixed cost). From a theoretical standpoint, the motivation for this assumption is that some sluggishness in adjustment is necessary in order to break the potential for complementarities to generate multiple equilibria, as emphasized by Frankel and Pauzner (2000). More intuitively, this assumption also captures that firms may have heterogeneous (unobservable) abilities to adjust to market conditions as they change.

3.2 The effects of a cash crunch

We now discuss the effects of a sudden, unanticipated, and large decline in M_t to the technology choices of firms in this model. Throughout, we consider a reduction in M_t of size S at date 0:

$$M_0 = (1 - e^{-\theta\Delta})M^c + e^{-\theta\Delta}M_{-\Delta} - S. \tag{6}$$

Our discussion focuses on the implications of the model for three particular moments: the long-run response of the number of users of e; the long-run response of the number of firms newly switching to e ("switchers," in what follows); and the relationship between initial conditions and long-run responses. Table 1 summarizes the predictions of different versions of the model for these moments. The rest of the section discusses these predictions in more detail.

	Complementarities ($C > 0$ and $\kappa = 0$)	Fixed costs ($C = 0$ and $\kappa > 0$)
P1: Long-run increase in number of users	✓	✓
P2: Long-run increase in likelihood of switching	✓	✗
P3: Adoption depends positively on initial conditions	✓	✗

Table 1: Empirical predictions across versions of the model.

3.2.1 Complementarities ($C > 0$ and $\kappa = 0$)

With externalities ($C > 0$), technology choices depend on firms' expectations about how the number of users of e will evolve in the future. In principle, this could lead to equilibrium multiplicity, with self-fulfilling expectations. However, with common shocks ($\sigma > 0$), Frankel and Pauzner (2000) establish the following result.

Lemma 1 (Frankel and Pauzner, 2000)**.** *So long as $\sigma > 0$, if Δ is sufficiently small, there is a unique equilibrium, characterized by a frontier $\mathbf{\Phi}(.)$ such that firms adopt e, if and only if $M_{t-\Delta} \leq \mathbf{\Phi}(X_{t-\Delta})$. Moreover, the frontier $\mathbf{\Phi}(.)$ is upward-sloping.*

A key feature of the equilibrium is the fact that the adoption rule is *increasing* in $X_{t-\Delta}$. By contrast, when $C = 0$, the adoption rule is flat and independent of $X_{t-\Delta}$.[24] The slope is positive because adoption benefits depend positively on the current value of the number of users of e, $X_{t-\Delta}$. In turn, this is because, when adoption is sluggish ($k < +\infty$), the number of users of e displays some persistence. Firms re-optimizing their technology choice when $X_{t-\Delta}$ is currently high can expect it to stay high, at least in the near future. This raises the incentive to adopt e, so that the level of $M_{t-\Delta}$ must be higher in order to dissuade firms from moving to e.

The dynamics implied by this adoption rule are illustrated in the left panel of Figure 6. This panel plots the adoption threshold $\Phi(.)$ as well as two different trajectories, one (in red) for a district which starts from a low number of firms using technology e, and another (in blue) for a district which starts from a higher number of firms using technology e.

When the number of users is initially low (red line), the economy jumps from point A to point B as the negative shock to M_t occurs. Firms then start switching from c to e. But eventually, the economy reaches point C, on the adoption threshold. The economy then moves to the region in which abandoning e is optimal. Eventually, the economy converges back to point A. In this instance, the shock thus only has a temporary effect on technology choices.

On the other hand, if the initial number of firms using technology e, $X_{-\Delta}$ is sufficiently high, it can be

[24] See appendix A for a discussion of the frictionless case ($C = 0$ and $\kappa = 0$).

the case that X_t does not converge back to initial level, but instead, converges to 1. This is illustrated in the blue trajectory in Figure 6. On that trajectory, once the shock has taken place, the district permanently remains below the adoption threshold. In this case, the number of firms using e increases *permanently*, despite the fact that the shock is transitory. Importantly, firms that obtain the possibility of revising their technology choice always opt for e, even long after the shock has dissipated. As a result, the likelihood of switching also increases permanently. Thus, the model features *positive* state-dependence with respect to initial adoption rates.

We next provide an illustration of the quantitative properties of the model, by simulating the response of a large number of districts to the shock. These districts are assumed to have heterogeneous exposures to the aggregate shock S; namely, district d's shock is given by:

$$M_{d,0} = (1 - e^{-\theta\Delta})M^c + e^{-\theta\Delta}M_{d,-\Delta} - e^{\epsilon_d}S, \quad \epsilon_d \sim N(-\sigma_D^2/2, \sigma_D^2).$$

The average path of cash is reported in Figure 7. Districts are otherwise identical, save for their initial conditions $(M_{-\Delta,d}, X_d)$, which reflect the ergodic distribution of the model prior to the shock.[25]

The top row of Figure 8 shows the average response across districts. Consistent with the aggregate data discussed in the previous section, the number of firms using e increases permanently (top left panel of Figure 8). Moreover, the likelihood of switching also increases permanently (top right panel of Figure 8) .

This average response masks substantial heterogeneity across districts. First, districts which (all other things equal) experience a larger decline in M (that is, have a higher exposure e^{ϵ_d}) are more likely to remain in the adoption region in the long-run. Indeed, quantitatively, the model predicts that the long-run response of the number of users of e (the left panel of Merchant 9) is increasing in the exposure of the district to the shock, e^{ϵ_d}.

Second, districts with different initial conditions will also experience different long-run adoption dynamics (for a given exposure level). As discussed above, we should expect districts with high initial adoption to respond more to the shock, all other things equal. That is, the long-run response should be state-dependent, where the word "state" here refers to the endogenous state variable of the district, the initial number of users of e, $X_{0,d}$. The numerical simulations confirm this. The right panel of Figure 9 shows that the long-run response of both the number of users of e is increasing in the level of initial adoption, $X_{0,d}$. This results highlights the broader idea that long-run adoption dynamics are determined by the initial strength of complementarities.

[25] Appendix A reports the details of our numerical solution method for the model; it uses iterated deletion of strictly dominated strategies, leveraging the proof of lemma 1 by Frankel and Pauzner (2000).

3.2.2 Fixed costs ($C = 0$ and $\kappa > 0$)

In the model with only fixed costs, firms' technology choice follows a simple (S, s) rule. Two boundaries, $\underline{M}$ and $\overline{M}$, fully characterize technology choices: a firm chooses to switch to e if $M_t < \underline{M}$, to switch to c if $M_t > \overline{M}$, and the status quo when $\underline{M} < M_t < \overline{M}$. As illustrated in the left panel of Figure 10, a large shock moves the economy from its initial state (point A) to the adoption region (point B); but in finite time, the economy reaches the boundary $\underline{M}$ again (at point C). At that, point adoption ceases, but firms that receive the technology adjustment shock choose inaction, so that the fraction of users of e stays constant. Thus, large temporary shocks can have permanent effects on the number of users, just as in the model with complementarities. However, differently from the model with complementarities, the likelihood of switching does not increase permanently: it goes to zero as the shock dissipates.

An additional feature of the model with fixed costs is that it features *negative* state-dependence of adoption with respect to initial conditions, $X_{0,d}$. The expected time to go from point B (the point to which the economy is brought after the shock) to point C (the point at which the inaction region is reached again) does not depend on the initial number of users of e. Because the law of motion for X_t, from B to C, is simply $\Delta X_t = (1 - e^{-k\Delta})(1 - X_{t-\Delta})$, the cumulative change in X_t is a decreasing function of the initial number of users, X_0. This negative state-dependence is a consequence of the assumption that the total number of firms is fixed. Nevertheless, it stands in contrast to the model with complementarities.

Figures 11 and 12 further highlight these differences with respect to the complementarities model. Figure 11 reports the average response of the economy to the shock. The number of users increases permanently, but the likelihood of switching goes to zero after the shock has dissipated. Consistent with the long-run response of the number of users overall, across districts with different exposures to the shock, the long-run response of the number of users is positively related to shock exposure (left panel of Figure 12). However, as reported on the right panel of Figure 12, the long-run response of the number of users is negatively related to initial conditions, instead of the positive relationship predicted by the model with complementarities.

3.2.3 Shock persistence

This discussion of the complementarities case ($C > 0$ and $\kappa = 0$) focused on versions of the model where $\theta > k$, that is, the speed at which firms may adjust their technology choice is slow relative to the speed of mean-reversion of the shock. Under the alternative assumption ($\theta < k$), the pure complementarities model tends to generate a *stronger* permanent switch to e after the shock, but a *weaker* (and, in fact, *negative*) relationship between initial conditions and subsequents increases in the number of users, e.

The first part of this claim is illustrated in the bottom panel of Figure 13, which reports the adoption dynamics in a version of the model where $\theta < k$. The average fraction of firms using technology e rapidly converges to 1 after the shock, reflecting the fact that firms frequently receive the technology adjustment shock. As a result, adoption, on average, converges to 1, and the likelihood of switching also permanently increases, as illustrated in Figure 14.

Importantaly, this occurs independently of whether the initial adoption rate is high or not. As a result, there is little dependence on initial conditions — all districts tend to converge to $X_\infty = 1$ in this case. The right panel of Figure 15 illustrates this. There is a weak negative relationship between the change in the number of users and initial conditions when $\theta < k$, instead of the strong positive one when adjustment is more sluggish ($\theta > k$).

This interaction between shock persistence and state-dependence of responses has implications for policy. One might have thought that the presence of complementarities gives policymakers unsually strong powers in triggering technology adoption: temporary interventions can indeed have permanent effects. But in fact, the model indicates that this only occurs when the shock is sufficiently persistent, that is, when θ is below k. Instead, the more temporary the intervention is (that is, the higher θ is, in particular relative to the speed of adjustment of firms, k), the more likely it is that it will have heterogeneous effects across districts. Moreover, these effects will increase with initial adoption rates. Thus, very temporary interventions will do nothing more than accentuate differences in initial technology adoption. A policymaker with a preference for temporary interventions will therefore face a trade-off between the persistence of the shock and its distributional effects. This implication of the model also reinforces the importance of documenting state-dependence in the data.

There are three main take-aways from the analysis of the model. First, both the fixed cost and the complementarities model can generate persistent adoption following the temporary shock. Second, only the complementarities model can generate long-run increases in the likelihood of switching. Third, the complementarities model is characterized by a positive relationship between long-run adoption and the initial strength of complementarities. The latter two predictions, which are likely to survive in a more general model where *both* fixed costs *and* complementarities are active ($C > 0$ and $\kappa > 0$), can thus help us identify the presence of complementarities in the data. This is the question we turn to in the next section.

4 Adoption Dynamics

So far, section 2 showed that, in aggregate, the demonetization was followed by a wave of adoption of electronic payment systems. Moreover, the increase in the use of eletronic payment systems was persistent, despite the fact that the demonetization's effects on the supply of cash were relatively short-lived. Section

3 then analyzed a model of technology adoption which can rationalize these observations using either one of two potential mechanisms: fixed costs of adoption, or complementarities across firms in adoption decisions.

The goal of this section is to further test the model's predictions, and in particular those (reported in table 1) that are specific to complementarities. The two key predictions that characterize complementarities can be summarized as follows: first, the shock causes an increase in the likelihood that firms will switch to the platform, which persists beyond the shock itself; second, the size of the long-run change in adoption caused by the shock is positively related to initial adoption rates, and to the initial strength of complementarities more generally, a prediction which we referred to as "state-dependence."

We start by briefly describing the data we use in our analysis. We then develop a novel quasi-experimental framework to estimate the causal impact of the cash-contraction across Indian districts. Using this method, we show that the shock not only caused a persistent increase in overall adoption, but also led to an increase in the overall rate at which new firms switch to the platform, and that this increase persisted substantially beyond the cash crunch. We then use three separate empirical settings to test for state-dependence, and show that in all three, adoption following the shock was higher in situations where the initial strength of complementarities was stronger.

4.1 Data

The main data source we use in our analysis are merchant-level transactions from one of the leading digital wallet companies in the country.[26] We observe sales amount and number of transactions through QR code on a weekly basis for anonymized merchants between May 2016 and June 2017. For each merchant, we also know the location of the shop at the city level, as well as the store's detailed industry. For a random sub-sample of shops, the location is provided at the more detailed level of 6-digit pincode.[27] 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 transactions are geo-localized, therefore allowing us to aggregate them up at the same level as other data sources used in this study.

We obtain data on district-level banking information from the Reserve Bank of India (RBI). This includes three pieces of information: first, number of bank branches and banks operating these branches; second, information on the number of the currency chests by district, and the banks operating the chests; third, quarterly bank deposits at the bank-group level in each district.

[26] The specifics of the payment technology offered by the company are described in the introduction as well as in section 2. 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".

[27] A pincode in India is the approximate equivalent of a five-digit zip-code in the US. Pincodes were created by the postal service in India. India has a total of 19,238 pincodes, out of which 10,458 are covered in our dataset.

Finally, we complement these data with information from the Indian Population Census of 2011 to calculate a large set of district-level characteristics. These characteristics include: population, banking quality (share of villages with ATM and banking facility, number of bank branches and agricultural societies per capita), measures of socio-economic development (sex ratio, literacy rate, growth rate, employment rate, share of rural capital) and other administrative details including distance to the nearest urban center.

4.2 Heterogeneous shock exposure

4.2.1 Measuring exposure at the district level

In the first part of the paper we have shown that the demonetization was associated with a large increase in the use of electronic payment systems. However, the model from Section 3 makes a stronger prediction, showing that the increase of adoption should also be positively related to the size of the shock at local level, as highlighted in the left panels of Figures 9 and 12. Building on this idea, this section will develop a novel empirical strategy to estimate the causal effect of the cash contraction across districts.[28]

To identify heterogeneity in the exposure to the cash-contraction at the local level, we construct a shock measure that exploits heterogeneity across districts in the relative importance of chest banks in the local banking market. In the Indian system, currency chests are branches of commercial banks that are entrusted by the RBI with cash management tasks in the district. Currency chests receive new currency from the central bank and are in charge of distributing it locally. While the majority of Indian districts have at least one chest bank, districts differ in the total number of chest banks, as well as in chest banks' share of the local market.

Consistent with anecdotal evidence, we expect that districts with in which chest banks account for a larger share of the local banking market should experience a smaller cash crunch during the months of November and December.[29] On some level, this relationship is mechanical. Chest banks were the first institutions to receive new notes, so that, when chests account for a larger share of the local banking market, a larger share of the population can access the new bills. Furthermore, the importance of chest banks may be even a more salient determinant of the access to cash if these institutions were biased toward their own customers

[28] Moving to a cross-sectional analysis has also several advantages in terms of identification. An event-study framework, analogous to the results presented in section 2, is well-suited to examine the immediate reaction to a large shock, but it can have limitations when looking at medium-run responses. Estimates of medium-run effects will be confounded by any other aggregate shocks that may affect the outcomes of interest. In this context, there may be several aggregate factors that may play a role. Among other potential confounding factors, we note a change in marketing strategy of the electronic wallet company (e.g. announcement of a fee in February), changes in the competitive environment of the electronic payment industry starting in March 2017, and a series of discretionary RBI directives to encourage the use of electronic payments during the early months of the demonetization (for a description of the latter, see RBI, 2017). Additional aggregate macroeconomic events potentially confounding the demonetization include the US election, which occurred on the same day as the demonetization, a better than usual monsoon, and the launch of the Goods and Services tax, a VAT tax, on July 1, 2017.

[29] In the popular press, several articles argue that proximity — either geographical or institutional — to chest banks contributed to the public's ability to have early access to new cash. For instance, see https://www.thehindubusinessline.com/opinion/columns/all-you-wanted-to-know-about-currency-chest/article9370930.ece

or partner banks. Indeed, concerns of biases in chest bank behavior were widespread in India during the demonetization.[30]

To measure the local importance of chest banks, we combine public data on the location of chest banks with information on overall branching in India and data on bank deposits in the fall quarter of the year before demonetization (2015Q4). Ideally, we want to measure the share of deposits in a district held by banks operating currency chests in that district. However, data on deposits are not available at the district level for each bank. Instead, the data is only available at the bank-type level (G_d).[31] Since we have information on the number of branches for each bank at the district level, we can proxy for the share of bank deposits of each bank by scaling the total deposits of the bank-type in the district, by the banks' share of total branches in that bank-type and district.[32] We can then can compute our score as:

$$\text{Chest}_d = \frac{\sum_{b \in C_d} \sum_j D_{jbd}}{\sum_{b \in B_d} \sum_j D_{jbd}} \approx \frac{1}{D_d} \left(\sum_{g \in G_d} \left(D_{gd} \times \frac{N^c_{gd}}{N_{gd}} \right) \right)$$

where D_d is the total amount of deposit in the district d, D_{gd} and N_{gd} are respectively the amount of deposits and the number of branches in bank-type g and district d, and N^c_{gd} is the number of branches in the district for a bank with at least one currency chest in the area.[33] Since we want to interpret our instrument as a measure of the strength of the shock, our final score Exposure_d is simply the inverse of above chest measure *i.e.* $\text{Exposure}_d = 1 - \text{Chest}_d$. Figure 16 plots the distribution of this exposure score across districts. The figure shows a very smooth distribution centered on a median around 0.55, with large variation at both tails (SD is about 0.18).

Intuitively, areas where chest banks are less prominent — or high-exposure according to the index — should have experienced a higher cash contraction during the months of November and December. Given the data publicly available, this hypothesis cannot be tested directly. However, as discussed in section 2, the cash crunch occurred because old notes had to be deposited by the end of the year, but withdrawals were severely limited. Therefore, the growth in deposits during the last quarter of 2016 proxies for the cash contraction in the local area. Figure 17 provides evidence that is consistent with this intuition. Here, we

[30] In a report in December, the RBI has discussed this issue extensively. In one comment, they report how "these banks with currency chests are, therefore, advised to make visible efforts to dispel the perception of unequal allocation among other banks and their own branches." See https://economictimes.indiatimes.com/news/economy/finance/banks-with-currency-chest-need-to-boost-supply-for-crop-rbi/articleshow/55750835.cms?from=mdr.

[31] The RBI classifies banks in six bank groups: State Bank of India (SBI) and its associates (26%), nationalized banks (25%), regional rural banks (25%), private sector banks (23%) and foreign banks (1%).

[32] A simple example may help. Assume we are trying to figure out the local share of deposit by banks A and B, both rural banks. We know that rural banks in aggregate represents 20% of deposits in the district, and we know that bank A has 3 branches in the district, while bank B has only one. Our method will impute bank A's share of deposits to be 15%, while bank B will be 5%.

[33] In practice, this approximation relies on the assumption that the amount of deposits held by each bank is proportional to the number of branches, within each district. The strength of our first-stage analysis suggests that this approximation appears to be reasonable.

plot distribution of deposit growth across Indian districts in the last quarter of 2016 versus 2015. In normal times (2015), we see that the distribution is relatively tight around a small positive growth. During the demonetization, the histogram looks very different. First, almost no district experienced a contraction in deposits. Second, the median increase in deposits was one order of magnitude larger than during normal times. Third, the data overall shows much more dispersion across districts, suggesting that the effect of the demonetization were likely not uniform across Indian districts.[34]

Building on this intuition, we then show that our exposure measure indeed proxies for the severity of the cash crunch at the local level. Figure 18 shows that there is a strong relationship between district-level exposure to the shock and deposit growth. The same relationship holds using different measures of deposit growth and when including district-level controls, as shown in Table 3. More importantly, Table C.1 also shows that this strong relationship is unique to the demonetization quarter.[35]

4.2.2 Results

Using this measure of exposure, we are going to estimate the following difference-in-difference model:

$$\log\left(y_{d,t}\right) = \alpha_t + \alpha_d + \delta\left(\text{Exposure}_d \times \mathbf{1}_{\{t\geq t_0\}}\right) + \Gamma_t' Y_d + \epsilon_{d,t}, \tag{7}$$

where t is time (month), and d indexes the district, t_0 is the time of the shock (November 2016) and Exposure_d is the measure of district's exposure constructed with chest bank data, as explained above. The equation is estimated with standard errors clustered at district level, which is the level of the treatment (Bertrand et al., 2004). Lastly, in the main regression we include data between May 2016 and February 2017. [36]

Importantly, the specification is also augmented with a set of district-level controls (Y_d), which are measured before the shock and interacted with time dummies. The presence of controls is important for the causal interpretation of our results, because chest exposure is clearly not random. Table 2 examines this issue, by showing the difference across characteristics for districts characterized by different exposure. In general, exposure to chest banks is actually uncorrelated with several district-level demographic and economic characteristics, but not all of them. In particular, higher exposure (lower density of chest banks) is found in districts with a smaller deposit base, smaller population, and a larger share of rural population.

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

[35] In particular, the Table uses data since 2014 and shows that – in normal time – the relationship between these two quantities is small and generally insignificant. In the only case in which this relationship is positive, we find that the relationship is several order of magnitude smaller than in 2016Q4.

[36] We exclude sparsely populated north-eastern states and union territories from the analysis due to missing information on either district-level characteristics or banking variables. The seven north-eastern states include Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, Tripura while union territories include Anadaman and Nicobar Islands, Chandigarh, Dadra and Nagar Haveli, Daman and Diu, Lakshadweep and Pondicherry. All together these regions account for 1.5% of the total population of the country.

However, most of the variation in exposure is absorbed once we control for two simple determinants of local banking market, size of the deposit base in the quarter before the shock and percentage of village with ATM (last columns, Table 2). Taking a more conservative approach, our controls will include the log of deposit in the quarter before the demonetization, the percentage of villages with ATM, the log of population, the share of villages with banking facility, and the share of rural population.

Using this model, we now examine the causal effect of the cash contraction on technology adoption, mapping our results to the predictions of the model. After presenting the results, we will also discuss further evidence on the causal interpretation of our analyses.

Exposure and total adoption In Table 4, we start by showing that the exposure to the cash-contraction predicts the response across districts. In the first two columns, we examine the effect on the amount of transactions undertaken using the wallet technology. We find that districts which were more exposed to the shock saw a larger increase in the amount of usage of electronic payments in the months following the demonetization. This result is both economically and statistically significant. Comparing two districts one-standard deviation apart in terms of exposure, we find that the more exposed one experienced an increase in electronic wallet use that is 45% and 55% larger than average. The same effect — with comparable magnitudes — also appears for our main measure of adoption, which is the number of firms operating in the platform.

The same results can be also examined dynamically. In other words, we can estimate how the relative growth in technology adoption changed across districts characterized by different level of exposure in the months around the 2016 demonetization.[37] Results are reported in Figure (19), where the effects are estimated relative to the month before the shock (October). This figure highlights three main findings. First, we confirm that our main effect is not simply driven by differential trends between high- vs. low-affected regions. Second, we find that the shift in adoption across districts happened as soon as November. However, the effect is in general larger in December and January. Third, we find that the difference in the response persists also after the cash returns to circulate. In particular, the effects are still large and significant after the month of February. We find consistent results across both the amount of transactions and the number of firms using the technology. In line with the initial aggregate evidence, this analysis confirms that the temporary cash contraction led to a permanent increase in the adoption of payment technologies.

[37] To be precise, we estimate the following equation:

$$\log\left(y_{d,t}\right) = \alpha_t + \alpha_d + \delta_t \left(\mathbf{1}_{\{d\in\mathcal{T}\}} \times \mathbf{1}_{\{t\geq t_0\}}\right) + \Gamma_t' Y_d + \epsilon_{d,t}. \tag{8}$$

Exposure and new adopters Next, we examine how the shock affected the initial decision of firms to use the technology. Since the shock led to sharp increase in the total number of firms using the technology, we expect to find a similar effect in new firms joining the platform right around the shock. The key question is whether this relative increase will persist also when the situation normalizes. As we have shown with the model in section 3, this persistent increase in new firms starting to use the electronic wallet is a unique feature of the model characterized by externalities. Without externalities, switching into the platform will only happen at the height of the shock.

Results can be visualized directly in Figure 19.[38] As expected, we find that districts experiencing a larger contraction in cash saw a larger increase in new firms joining the platform as soon as on November 2016. But firms in highly affected areas kept joining the platform at a higher rate after January 2017, the last month during which cash withdrawal was constrained. The relative increase in highly exposed districts appears to persist for the whole spring 2017.

Thus, districts experiencing higher cash contractions saw a larger and persistent increase in the usage of electronic payments. Additionally, consistent with a model characterized with externalities, this effect is partly explained by the fact that the shock led to persistent increase in the number of firms joining the platform. Overall, to argue for the causal interpretation of these results, we have mostly leveraged the (conditional) balancing and the lack of pre-trends in our analysis. One remaining concern is that our exposure measure may be correlated with unobservable demand shocks that are contemporaneous to the demonetization but not necessarily related to the cash scarcity. To assuage this concern, it is important to highlight that later in the paper (in Section 5), we will show that the same highly affected districts also experienced a larger decline in consumption during this period. This joint effect on electronic payment and consumption can be easily explained by the cash contraction, but it would be inconsistent with any demand side explanation.[39] Lastly, to further bolster the identification, Section 5 will present a full set of placebo that exploit the longer panel dimension in the consumption data and confirm the quality of our empirical strategy.

4.3 State-dependence in adoption dynamics

One of the key predictions of the model with complementarities is the state-dependence of adoption. In particular, the model suggests that a temporary shock may lead to permanent adoption, but that the

[38] To be conservative, we define a firm as new if it uses the platform for the first time in a month and it undertakes transactions amounting to more than Rs.50. Results are consistent with alternate thresholds of Rs.0, Rs.10 and Rs.100.

[39] One story is the following. The demonetization clearly also increased overall policy uncertainty in the country, and this increase in uncertainty may hurt different regions differently because of some heterogeneity in the local industry mix. If our exposure measure is negatively correlated with the local effect of uncertainty, we may find a relative increase in electronic payment simply because economic activity in our exposed region contracted less than less exposed region. While this correlation structure could explain the effect on electronic payment adoption, it would then be inconsistent with the consumption effect.

increase in adoption will not be uniform across areas: it will crucially depend on the initial strength of complementarities in the area. For instance, in the simulations we have shown that there is strong positive relationship between the initial number of users and the increase in adoption (see the right panel of Figure 9). In this Section, we use the data on electronic payments to present three pieces of evidence that are consistent with this prediction. This evidence of state-dependence in adoption is crucial for two things. First, the results — together with the previous finding on new adopters — will confirm that a model without complementarities fails to capture crucial features of the adoption dynamics. Second, the analysis will show that the state-dependence is not just a theoretical feature of the model, but an economically relevant force that explains the heterogeneity in adoption during the demonetization.

4.3.1 Firm-level evidence

We start by examining whether the existing user base in local markets plays a role in determining individual firms' decision to use the platform. In the model, we can write the use of the technology by a firm i at time t in the area d as a function of its own use of the technology in the previous period, the level of the aggregate shock, and the level of adoption by other firms in the same market.[40] Under the assumption that the technology is characterized by positive externalities ($C > 0$), the level of adoption by other firms in the same area will positively predict the adoption by the firm (Column 2 of Table 6). The intuition for this result is simple: an increase in the use of the technology will increase the value of the technology itself, which in turn will then positively affect adoption by firms. Importantly, the same relationship will not hold without externalities (Column 1 of Table 6).

Therefore, the model implies a positive relationship between a firm use of the technology and the overall use by other firms in the same area. This idea is actually consistent with the approach used in other settings to test for the presence of spill-overs in behavior. For instance, Munshi (2004) has used a similar methodology to explore the role of social learning in agriculture in rural India. In our context, this relationship will create state-dependence, because differences in the initial level of adoption across markets will endogenously affect the pattern of adoption in the future.

We leverage on the granularity of the data to test for the presence of this relationship in our setting. For each firm, we measure the total use of the technology by firms that are located in the same geographical area and that operate in the same industry. We choose this set of firms as reference group because we believe that complementarities should be strongest among firms in the same area and industry. In particular, we

[40] To be precise, we could estimate a firm-level regression in the simulated data of the following form:

$$x_{i,d,t} = \alpha + \rho x_{i,d,t-\Delta} + \beta M_{d,t-\Delta} + \gamma X_{d,t-\Delta} + \epsilon_{d,t} \tag{9}$$

expect to find the largest overlap in customers for companies within the same area and industry, as well as the largest spillovers in learning about the value of the technology. However, later we will also consider alternative definitions of the reference group in robustness checks.

We then test whether technology use, for one firm, is indeed positively related to the use of the technology by other firms in its reference group in prior periods. In particular, we estimate:

$$x_{i,p,k,t} = \alpha_i + \alpha_{p,t} + \alpha_{k,t} + \rho x_{i,p,k,t-1} + \gamma X_{p,k,t-1} + \epsilon_{i,p,k,t}. \tag{10}$$

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, this measure could be a dummy for whether the firm used the platform, or it could be the amount of activity of the firm on the platform.[41] The variable $X_{p,k,t-1}$ is a measure of 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, and 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.[42] The model is estimated using weekly data from our electronic wallet company.[43] 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.

We start by estimating this equation without any fixed-effects in the first column of table 7. We find that a volume of electronic transactions by firms in the same reference group strongly predicts more transactions for the firm itself in the following week. This effect is not only statistically significant but also quantitatively large. In particular, a one-standard-deviation increase in transactions by firms in the reference group leads to a 40% increase in the amount of transactions for the firm, which corresponds to 18% of the standard-deviation of the outcome variable. The same results hold — with similar magnitude — when we look at the number of transactions, or at whether the firm was active on the platform.

In this analysis, the main concern is that past decisions by firms in the reference group may be correlated with an individual firm's behavior simply because they proxy for some unobservable heterogeneity across firms that is unrelated with the strength of complementarities. For instance, a certain area may have

[41] We classify firms into 14 broad industries: 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%).

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

[43] The sample is a balanced panel of all the firms that used the wallet between May 2016 and June 2017 and that have information on location (pincode). We use pincode to identify firms' locations 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.

on average more educated workers, who may then be more likely to adopt the platform, irrespective of complementarities. In this case, past adoption by other firms may simply proxy for the effect of education. In other words, the results could be driven by an omitted variable that is correlated with adoption by other firms but orthogonal to the strength of complementarities.

To assuage this concern, we proceed in three steps. First, we augment the specification with a firm-fixed effect. Since we are now exploiting only within-firm variation, the type of omitted variable that could explain our result would also need to be time-varying. Second, in the third column of Table 7, we add pincode-by-week fixed-effects. This fixed-effect will allow us to keep constant in the model any characteristic of the area, even to the extent that these characteristics have a differential effect over time. Third, we also add a detailed set of industry-by-week fixed-effects (column four of Table 7). Relative to the previous framework, in this specification we not only compare firms within the same location, but we also adjust the estimates for changes adoption rates in the same industry. Across all these specifications, we consistently find that the adoption intensity by firms in the same reference group is a strongly positive predictor of a firm's use of the platform.

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 importance of complementarities to be different in the three periods. The predictions of the model are actually consistent with this intuition. In fact, using the simulated data from the model, we can show that the importance of the adoption by other firms is particularly salient in the shock period. In other words, the model suggests that the role of complementarities in individual adoption decisions is particularly importance during the transition.

To explore this hypothesis, we repeat the same analyses as before, but rather than estimating one single parameter for the effect of externality, we estimate a month-specific parameter for each of our outcomes. The results are plotted in the three panels of Figure 20. Across the three outcomes, we draw two main conclusions. First, the positive effect documented before is always present in the data, both before and after the policy shock. This is reassuring, since the state-dependence induced by complementarities is not generated by the shock but should be a feature of technology choices in any scenario. Second, we find that the effects of adoption in the reference group is much higher in the months of the demonetization, relative to the preceding and succeeding months. As previously discussed, this result is also consistent with the model.

Two robustness tests are worth highlighting. First, the results also hold if we define the area of the reference group as the district (Table C.2 in Appendix).[44] Second, results are robust when we define the relevant market in a different way. For instance, in Table C.3 in the Appendix we define the relevant market as any firm in the same location (pincode), irrespective of the industry. The results are also in this case

[44] One key advantage of this approach is that we have the location based on district for the whole data.

qualitatively identical.[45]

Overall, these results highlight the importance of state-dependence in adoption at the firm-level. Consistent with the model's implications, we have shown that the use of technologies by neighboring firms positively correlates with adoption at the firm level, and that this effect is particularly strong during the shock period. Unlike the other results in this paper, these analyses are not to be interpreted in a causal sense. In particular, the objective here is to document how the presence of complementarities generates an endogenous time-dependence between adoption decisions of different firms, even when keeping other sources of variation constant. Despite the limitations, we believe that the tests provided are able to exclude alternative interpretation of the results.

4.3.2 District-level evidence

In the theory section, we modeled state-dependence by showing how the strength of the complementarities before the shock predicts the path of a district in response to the shock (see the right panel of figure 9). In this sub-section, we provide evidence that is closely related to that result.

In particular, the model directly predicts that the initial level of adoption in a district should amplify the adoption response to the shock. The intuition is the usual: a larger number of initial adopters increases the benefit of switching to the technology, and will therefore increase the likelihood of moving into a higher adoption equilibrium (for a given size of the shock). In Table C.4, we provide evidence that the data is consistent with this prediction.[46] Across two specifications, we find that a high initial level of adoption at the district level tends to be correlated with a higher change in adoption after the shock.

This *within-district* evidence is thus consistent with state-dependence as defined in the model. However, we think that this approach has several shortcomings. First, the model has obvious limitations since we are effectively using a function of the outcome to define the treatment. This feature has clear implications for the endogeneity of the parameter estimated, as well as for its interpretation. Second, the scope of complementarities may extend beyond the district. For example, if complementarities are due to a shared customer base, then it is unclear whether adoption at the district-level is the correct way to proxy for their initial strength.

To address these concerns and provide further evidence on state-dependence at the district-level, we develop an alternative test which exploits variation across districts. In particular, we test how the increase in adoption differs depending the distance between a district and areas in which the usage of electronic wallets

[45] Clearly, with this alternative approach we cannot control for location-by-time fixed effects.

[46] Specifically, we estimate:

$$X_{d,t} = \alpha_t + \alpha_d + \delta \left(I_d \times \mathbf{1}_{\{t \geq t_0\}} \right) + \Gamma_t' Y_d + \epsilon_{d,t}, \tag{11}$$

was large *prior* to November (*hubs*). The mapping between the strength of complementarities and distance to the electronic payment hub is intuitive. In the model, the heterogeneity in the strength of complementarities is completely determined by the number of users in the same area. In reality, individuals move across districts and therefore the size of adoption in neighboring districts will be also important. Therefore, being located close to a large hub — a center where electronic payment use is relatively common — may significantly increase the benefits of adoption.

Specifically, we run a simple difference-in-difference model analogous to the one used when studying the effect of exposure to the shock (equation 11) where we compare the usage of wallet technologies around the demonetization period across districts that are differentially close to a digital wallet hub.[47] The main coefficient of interest in the one on the interaction between a post-dummy and the measure of distance at the district level.[48]

Despite the clear advantages relative to the naive *within-district* model, there are two main concerns with this model. First, by sorting on distance we may get identification from areas that are located in more extreme or remote parts of the country. Second, since the electronic hubs are some of the largest and more important cities in the country, we should expect that being located close to them will have benefits that go beyond the effect of complementarities. In other words, distance may capture other forms of heterogeneity that may affect the response to the shock independently from complementarities.

We deal with these limitations in three ways. First, we limit the comparison to districts that are located within the same state, adding state by month fixed-effects. In this way, we only exploit distance variation between areas that are already located in similar parts of the country. The results also hold with this set of fixed-effects. Second, we also control for the distance to the capital of the state, also interacted with time effects. This control allows us to isolate the effect of the distance to a major electronic payment hub from a more general distance to a large urban area. Third, we will augment the specification with the same set of district-level covariates used in the previous analysis.

The general specification is therefore a difference-in-difference model of the following form:

$$X_{d,s,t} = \alpha_{st} + \alpha_d + \delta \left(D_d \times \mathbf{1}_{\{t \geq t_0\}} \right) + \gamma \left(\tilde{D}_{d,s} \times \mathbf{1}_{\{t \geq t_0\}} \right) + \Gamma_t' Y_d + \epsilon_{d,t}, \tag{12}$$

where t indicates time, defined at monthly level in this analysis, d indexes the district and s identifies

[47] In particular, we define a district to be an electronic payment hub if there were more than 500 of active firms pre-demonetization (September 2016). Results are essentially identical if we use a threshold of 1,000 firms to define the hub districts. The distance to the hub is defined as the minimum of the distance between the district and all the hubs.

[48] There are nine electronic payment hubs with more than 500 active firms pre-demonetization and are spread evenly across the country: Delhi, Chandigarh and Jaipur (North), Kolkata (East); Mumbai and Pune (West); Chennai, Bangalore and Rangareddy (South). These five districts are also among top metropolitan cities by population in India, suggesting that the penetration of electronic payment was limited to main urban centers in India.

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. It is important to point out that we exclude from this analysis the districts defined as major digit wallet hubs.[49] 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 wallet hubs.

These results are reported in Table 5. In the columns (1) and (4), we report the baseline regression where we only control for the distance to the capital district as well as the other control at district-level.[50] Across both outcomes — amount and number of firms — we find that the districts farther away from major hubs experienced lower increases in transactions in the post-demonetization period. The same result holds when adding the state-by-month fixed-effects (columns 2 and 5) and the covariates interacted with the post-dummy (column 3 and 6). If we take the most conservative of the estimates, we find that 50 km increase in distance translates into a 25% lower increase in the amount of transactions. Lastly, we find similar results if we use a dichotomous definition of the treatment (Table C.5).[51]

Lastly, Figure 21 allows us to explore the dynamics 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.[52] The dynamic effects confirm that our results do not simply pick 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. Furthermore, this figure also reveals that some of the initial kick in adoption is caused by being located close to a large hub and does not disappear as cash withdrawal constraints are relaxed.

Overall, these results help 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 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 strong state-dependence in adoption.

[49] Notice that this exclusion does not affect our results; results including the hubs are, if anything, stronger.

[50] Similar to before, we determine our controls by examining which of the characteristics in the balancing test are actually correlated with distance. As a result, we include here employment rate, share of rural population and log of total population, on top of distance to state capital.

[51] In particular, we consider several alternatives, going from 400km down to 200km. Across all these tests, results are stable and significant.

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

4.4 Discussion

There are two key take-aways from this empirical analysis: first, shock exposure predicts both the short- and long-run strength of adoption; second, there is pervasive evidence that adoption was state-dependent, in the sense that stronger initial benefits to adoption – proxied by either adoption in a firms' area and industry, by initial adoption in the district, or by distance to hubs – predict a a higher long-run adoption response.

It is worth mentioning that these results do not rule out the presence of fixed adoption costs. As highlighted by the analysis of the model, fixed adoption costs can in fact generate the set of first findings regarding heterogeneous shock exposure, though it can generate neither the long-run increase in new adoption, nor the state-dependence. Thus, the empirical results require a positive amount of complementarity in adoption in order to be rationalized. However, we should also point out that ex-ante the costs of adoption in this specific context are likely to be very low.[53] As a result, a mechanism relying in large part on fixed-cost is not very likely to explain these results. In fact, in the pure fixed cost model the benefit of the new technology stays constant as adoption increases.[54]

Additionally, documenting state-dependence is useful over and above the fact that it provides evidence in favor of adoption complementarities. Our evidence suggests that state-dependence explains a substantial portion of the heterogeneity in adoption responses to the shock. From the standpoint of a policymaker, state-dependence should thus be an important concern when considering the potential effects of a temporary intervention targeting technology adoption. To be more specific, the state-dependence of (long-run adoption) responses arise when the model has complementarities *and* relatively short-lived shocks, at least when compared to the typical adjustment speed of firms. In that case, a key prediction of the model is that the shock could even *widen* the differences in adoption in the long-run, as districts with large initial adoption rates will tend to convert to full adoption, while districts with low initial adoption rates will not shift substantially. The evidence in this section suggests that this is not a second-order concern, but very much a feature of adoption responses in the aftermath of the demonetization. If the inequality of technology adoption is a concern for policy-makers, then a more long-term intervention may - as suggested by the model - represent a better solution.

[53] For instance, in our specific example, the electronic wallet does not require any monetary cost, and the set up only requires a phone and a bank account and can be completed in few minutes.

[54] To be precise, the fixed-cost model would explain a large and persistent increase in adoption only if the net benefit of using the technology is positive (so that firms keep using it after the shock), but also too small to justify adopting it in the pre-period. The small adoption costs thus put a sharp restriction on the size of the net benefits to using the platform.

5 Demonetization and real economic activity

Results from previous sections provide evidence that the Indian demonetization led to a widespread and persistent rise in electronic payments. Combining data and a model, we have documented the importance of complementarity in adoption in explaining this increase. Our main results leverage on data from the main electronic wallet company in India. However, as discussed in Section 3 the same pattern appears to hold - with some caveat - also in traditional electronic payments.

Given the size and speed of these responses, one natural question is whether the rise in electronic money was indeed sufficient to shield the real economy from the cash crunch. In this section, we use household consumption data to show that the cash contraction indeed negatively affected the real economy. However, these effects were somehow limited to the most acute period of the demonetization. Furthermore, the cut was larger in less necessary goods, like recreational expense. This result confirms that the rise in electronic payment was not sufficient to compensate for the temporary cash reduction. Furthermore, given that the development of financial technology may be a desirable objective for a government (e.g. Yermack (2018), Rogoff (2017)), this study also helps to characterize the possible negative effect of the shock. Lastly, consistent with the previous discuss, this result also helps validating our main results.

5.1 Empirical setting

In this Section, we examine how household consumption responded to the cash swap using the same identification strategy used in Section 4. In other words, we will compare behaviors across districts that were characterized by different level of presence of chest banks before the demonetization. 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).[55]

The data set provides a representative sample of Indian households, where households selected to be representative based on 371 "homogeneous regions" across India. The survey has information on the monetary amount of household expenses across different large categories, and some other background information on the members of the households.[56] Overall, the data quality is considered high, in particular since CMIE collects the data in person, using specialized workers.[57] In particular, each household is interviewed every

[55]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.

[56]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.

[57]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-response 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

four months, and in the interview the person is asked about their behavior in the previous period.[58]

The main difference with the analyses in Section 4 is the timing. In the previous analyses, the district-level data was measured at monthly level. For this household data, the survey procedure is such that households belonging to different waves of interviews are asked about the same month at different point in times. Therefore, the reporting on November 2016 – the first month of the shock – is generally clustered together with a different group of months depending on the wave.[59] This feature is actually quite common among consumer surveys and it is similar to the Consumer Expenditure Survey in US.[60] Following the literature in this area (e.g. Parker et al. (2013)), we deal with this feature by organizing the data in event-time. In other words, for each household we aggregate data at the wave-level and we define the time of each wave relative to the wave containing November 2016.[61]

With this data set of about 95,000 households, we then estimate the following household-level difference-in-difference model:[62]

$$\log\left(y_{h,d,t}\right) = \alpha_t + \alpha_h + \delta_t \left(\text{Exposure}_d \times \mathbf{1}_{\{t \geq t_0\}}\right) + \Gamma_t' Y_{h,d} + \epsilon_{h,d,t}, \tag{13}$$

where $y_{h,d,t}$ are consumption measures for household h in district d and survey-time t . α_t and α_h are event-time and household fixed effects, $Exposure_d$ is the district's exposure as described in section 3, which is interacted with dummies for the survey-time post-demonetization, $Y_{h,d}$ are controls, that are either at district or individual level. For controls in the regression, we use the same district-level covariates as in the previous set of analysis along with addition of household-level controls including the age of head of the household and log of household income, both measured as in the last survey before the shock. 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 wave of interview after the shock relative to the pre-period, which corresponds to the three waves of interview before the shock (12 months). However, we also present the results dynamically, allowing each wave to have an independent effect. After presenting the main results, we will come back and discuss the causal interpretation of these coefficients and

of non-responding households and allow the sample to cover households with low socio-economic status.

[58] Thus, about 39,500 households are surveyed every month.

[59] For example, 25% percent of households will be asked about August-November 2016 consumption in December 2016, 25% percent will be asked about September-December 2016 consumption in January 2017 and so on. Thus, November 2016 consumption will be recorded with other months depending on the month it was surveyed between December 2016-March 2017.

[60] The main difference is that the Consumer Expenditure Survey is ran every three months rather than four months.

[61] Therefore, the time in the panel will be one for the wave in which a household was interviewed about November, and it is zero for the wave that happened four months before the one in which November 2016 was contained and one for the one that happened four months after.

[62] 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.

robustness.

5.2 Main results

Table 8 shows results for consumption responses based on exposure to the shock. Column (1) shows that relative to the pre-period, total consumption was cut more for households located in highly affected district. The effect is sizable: consider a one-standard deviation increase in the chest bank score corresponds to about a 3.6% relative decline in total consumption. The same holds when using a dichotomous version of the shock: in this case, the highly affected households (top quartile) saw a relative drop of about 5.7%.

Importantly, the impact of the shock was temporary. Looking at the interaction between the treatment and dummies identifying the next 3 waves in which the household was interviewed, we consistently find a small and non-significant coefficients. If anything, the coefficient is positive during some of the post-periods, but generally this result is non-significant. This effect suggests that the cash contraction only significantly impacted household behavior during the months immediately after the demonetization and did not lead to a permanent change in consumption behavior. This evidence is consistent with the idea that the shock was really binding only between November and January.[63]

To understand better the effect on consumption dynamics, we repeat the same analysis by dividing aggregate consumption into finer categories (Table 9).[64] As a first step, we divide consumption between necessary and unnecessary consumption. With this measure, we want to identify the subset of the consumption basket that is harder to cut for households, and therefore for which we expect to see lower responses. Specifically, we consider necessary consumption in this setting all consumption expense for food, rent and bills, and utilities (power and gas). In particular, we find that consumption was cut extensively across the two dimensions. However, the effect for unnecessary consumption is economically larger. Specifically, on average the effect for this group is about 22% higher. The same difference holds also when looking using a dichotomous treatment (Appendix Table C.6). In this case, highly affected households cut necessary by 4%, while unnecessary by about 8%.

This result crucially depends on our categorization across necessary and unnecessary, which is arguably arbitrary. However, the same result holds when we look at specific consumption groups. Always in Table 9, we consider three consumption categories: rent and bills, food, and recreational expenses. For the first group - rent and bills - we find essentially no effect of the demonetization.In our mind, this result can be considered like a placebo test, because it is ex-ante unlikely that a temporary shock would lead to a change

[63]Because of the data structure, the timing of the second wave is different across households, depending on the group in which they belong. However, for three-quarter of the sample, the second wave starts on or after January 2017, which is the month in which we saw a net influx of cash back in the economy.

[64]One concern is that, as we move to more dis-aggregated consumption responses, we may incur on an increase in noise in the data.

in rent-related expenses, which are generally fixed in long-term contract. For food, the effect is still negative and significant. In particular, a one standard-deviation increase in exposure led to about 3% decline in food expenditure. However, this effect on food dwarfs relative to the cut on recreational expenses. For this category, we find that a standard-deviation increases led to more than a 15% cut in consumption.

At face value, these results show that the cash contraction had real effect on individual behavior. This confirms that the shift toward electronic payment documented in the paper was not sufficient to replace cash on transactions. Therefore, while the policy had clear positive effect in terms of fostering the adoption of electronic money, it also negatively impacted the welfare of households. However, to better evaluate this negative welfare effects, we need to highlight the other two key results. First, the contraction in consumption was highly temporary. As soon as cash came back and limitations for circulations were lifted - on average after January- consumption converged back to pre-shock levels. Second, the response to the shock was on average larger on less-necessary goods (e.g. recreation).

5.3 Robustness

In this Section, we present a set of robustness test that helps validating the causal interpretation of the results presented before. Together with the discussion in Section 4, these tests provide support to the idea that district's exposure to chest bank represent a credible empirical method to explore the effect of the demonetization.

In Figure 22, we confirm that the consumption result is not driven by differences in pre-trend between affected districts. Echoing the results on electronic payment in Section 4, we find that the exposure measure constructed based on chest banks do not appear to be correlated with differential trends in consumption before the shock. However, it does predict a different response during the demonetization —- as previous discussed.[65]

One residual concern is that districts with high exposure to chest banks are regions that are particularly sensitive to business cycle fluctuations. The pre-trend analysis partially helps with this concern, but it cannot rule this out completely because it focuses on one specific economic environment. Therefore, to bolster our identification further, we construct a large set of placebo tests, in which we repeat our main analysis centering it in periods in which there was no contraction in cash. In particular, to keep our approach general enough, we consider placebo shocks happening every months between February-2015 and February-2016. We then replicate our main specification, testing for the presence of differential response across households in the

[65] This analysis shows a positive and borderline significant effect in consumption two quarters after the demonetization. One interpretation is that households have shifted some consumption in the future. Consistent with this interpretation, we actually find that the effect is driven completely by unnecessary consumption, which is a category that contains also durable expenditure. However, we also want to point out that this positive result is statistically weak and it does not replicate using alternative treatment specifications (e.g. using top quartile).

wave of the placebo shock relative to the previous one.[66]

The result of this set of placebo tests are reported in Figure B.4. The general finding is that - in normal time — there is essentially no statistical difference in the change in total consumption between households in districts with different chest bank exposure. Together with the pre-trend analysis, this test excludes that differential exposure to business cycle may explain our results. More broadly, this test provide new evidence on the quality of our empirical specification.

6 Conclusion

In this paper, we study adoption dynamics for a technology characterized by externalities. In particular our main focus is on electronic payment systems and we use the Indian demonetization of 2016 as a laboratory to study how a large but temporary aggregate shock can affect adoption in this context. 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 succeeded into fostering large scale adoption of electronic payments because of role complementarity. In particular, the shock served as a coordinating device that allowed firm to overcome coordination frictions in the adoption of the technology.

We open the paper by documenting a large and persistent increase in the aggregate adoption of electronic wallet technologies by firms in India during the demonetization period. Furthermore, we find similar results on more traditional electronic payment. However, consistent with the idea that traditional technologies are characterized with higher adoption costs than fintech, we also document that in large part the effect on electronic payment is driven by an intensive margin, while the extensive margin effect (new cards or POS) is minimal.

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. At the same time, a key result from the model is the state-dependence. In particular, we show that the adoption response in a model with complementarity and temporary shocks is not necessarily uniform. In particular, the initial strength of complementarity - which is proxied in the model by the initial number of users - positively predicts the response in the long-run. This long-run state-dependence disappears when the shock is more persistent (or at the limit permanent). This result highlights the trade-off between length of the shock and distributional effects. Assuming that a policy-maker prefers a short-lived intervention to a longer one, the model suggests

[66] In our main result, there is essentially no difference whether we compare the effect on the previous wave - like in the Figure 22 - or the average of the previous three waves, like in Table 8. Here we choose to compare to the previous wave because this allows us to go more back in time with the placebo.

that – while the temporary shock can be successful to increase overall adoption – this policy shock may increase dispersion in adoption.

Following this model, we turn again to the disaggregated data on electronic payment. The objective is to test in the data the extent to which the dynamics of adoption observed during the Indian demonetization were consistent with the one predicted by the model. As a first step, we develop a novel identification model based on the heterogeneity in the presence of chest banks across India to estimate the causal impact of the cash contraction on adoption. Using this methodology, we confirm that the shock had a persistent effect in adoption also across districts. In part, this permanent increase in adoption is achieved because the shock also increased the entry of new firms into the platform. As predicted by the model with complementarity, this difference in entry persists also after the end of the cash scarcity.

After this analysis, we also show that patterns of adoption are consistent with the presence of state-dependence. In particular, using three separate tests, we document that the increase in adoption appears to be positively affected the initial strength of complementarity. First, we show that districts with higher level of early adoption on average increased technology adoption more. Second, we also show that the propagation of electronic payment technologies was stronger in areas that were closer the centers in which electronic payments were already widely used in the pre-shock period. Third, we show that firm-level decision to use the technology is influenced by the recent behavior of other firms that are likely to share the same customer base as the firm itself.

As a last step in this analysis, we also confirm that the large increase in electronic payment was overall insufficient to limit the impact of the cash contraction. Consistent with this result, we find that the cash contraction caused a sizable, negative effect on consumption of Indian households. In particular, using the same approach based on chest-bank and a sample of about 95k households, we show that a one-standard deviation increase in exposure to the shock led to more than a 3% decline in consumption. At the same time, we show that this effect was completely temporary and generally larger for unnecessary consumption.

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.

Banerjee, A. V., E. Breza, A. G. Chandrasekhar, and B. Golub (2018). When less is more: Experimental evidence on information delivery during india's demonetization.

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.

Comin, D. and R. Nanda (2019). Financial development and technology diffusion. *IMF Economic Review*.

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.

Gupta, A., J. Ponticelli, and A. Tesei (2019). Technology adoption and access to credit via mobile phones. *Working Paper*.

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

Higgins, S. (2019). Financial technology adoption.

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.

Krishnan, D. (2017, 21). Survey of the effects of demonetisation on 28 slum neighbourhoods in mumbai. *Economic and Political Weekly Vol. 52*(Issue No. 3).

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

Parker, J. A., N. S. Souleles, D. S. Johnson, and R. McClelland (2013). Consumer spending and the economic stimulus payments of 2008. *American Economic Review 103*(6), 2530–53.

RBI (2017). Macroeconomic impact of demonetisation: A preliminary assessment. Technical report, Reserve Bank of India.

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.

Rysman, M. (2007). An empirical analysis of payment card usage. *The Journal of Industrial Economics 55*(1), 1–36.

Suri, T. (2011). Selection and comparative advantage in technology adoption. *Econometrica 79*(1), 159–209.

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.

A Appendix to section 3

A.1 Derivations

A.1.1 Value functions

The value of a firm which is operating under technology $x_{i,t}$ in period t, *after* any potential technology revisions, but before the realization of the money shock M_t, is:

$$\begin{aligned} V(x_{i,t}, M_{t-\Delta}, X_{t-\Delta}) \quad &= \mathbb{E}_{t-\Delta}\Big[\Pi\left(x_{i,t}, M_t, X_t\right)\Delta \quad + \\ &\qquad e^{-r\Delta}\left\{\left(1-e^{-k\Delta}\right)V_R(x_{i,t}, M_t, X_t) + e^{-k\Delta}V(x_{i,t}, M_t, X_t)\right\}\Big]. \end{aligned}$$

Here, $V_R(x_{i,t}, M_t, X_{t-\Delta})$ denotes the value of a firm that receives the option to revise its technological choice early on in period $t+\Delta$ (and has entered that period with technology choice $x_{i,t}$). This value is given by:

$$V_R(x_{i,t}, M_t) = \begin{cases} V(e, M_t, X_t) - \kappa & \text{if } x_{i,t} = c \text{ and } V(e, M_t, X_t) - V(c, M_t, X_t) \geq \kappa \\ V(c, M_t, X_t) & \text{if } x_{i,t} = c \text{ and } V(e, M_t, X_t) - V(c, M_t, X_t) < \kappa \\ V(e, M_t, X_t) & \text{if } x_{i,t} = e \text{ and } V(e, M_t, X_t) - V(c, M_t, X_t) \geq 0 \\ V(c, M_t, X_t) & \text{if } x_{i,t} = e \text{ and } V(e, M_t, X_t) - V(c, M_t, X_t) < 0 \end{cases}$$

(Note that this assumes that κ is a fixed cost that does not scale with the size of the time period, Δ. So it should be interpreted in units of firms value.) Denote by:

$$B(M_{t-\Delta}, X_{t-\Delta}) = V(1, M_{t-\Delta}, X_{t-\Delta}) - V(0, M_{t-\Delta}, X_{t-\Delta}).$$

This is the value of a firm which has the electronics payment in place, relative to one that doesn't. Straightforward computation then shows that the gross adoption benefits follow (3).

A.1.2 The relative value of adoption in complementarities model ($C > 0$ and $\kappa = 0$)

The conditional distribution of $M_{t+\Delta n}$, $n \geq -1$, given initial conditions $M_{t-\Delta}$ is:

$$M_{t+\Delta n}|M_{t-\Delta} \sim N\left((1-e^{-(n+1)\theta\Delta})M^c + e^{-(n+1)\theta\Delta}M_{t-\Delta}, \frac{1-e^{-(n+1)\theta\Delta}}{1-e^{-\theta\Delta}}\Delta\sigma^2\right).$$

The net benefits of adoption can be written as:

$$B(M_{t-\Delta}, X_{t-\Delta}) \quad = \mathbb{E}_{t-\Delta}\left[\sum_{n\geq 0} e^{-(r+k)\Delta n}\left(M^e + CX_{t+\Delta n} - M_{t+\Delta n}\right)\Delta\right]$$

We need to compute:

$$\begin{aligned} PVM_{t-\Delta} = \mathbb{E}_{t-\Delta}\left[\sum_{n\geq 0} e^{-(r+k)\Delta n} M_{t+\Delta n}\Delta\right] \quad &= \sum_{n\geq 0} e^{-(r+k)\Delta n}\left\{\left(1-e^{-(n+1)\theta\Delta}\right)M^c + e^{-(n+1)\theta\Delta}M_{t-\Delta}\right\} \\ &= \sum_{n\geq 0} e^{-(r+k)\Delta n}\left\{\left(1-e^{-(n+1)\theta\Delta}\right)M_c + e^{-(n+1)\theta\Delta}M_{t-\Delta}\right\}\Delta \\ &= \frac{\Delta}{1-e^{-(r+k)\Delta}}M^c + \frac{e^{-\theta\Delta}\Delta}{1-e^{-(r+k+\theta)\Delta}}\left(M_{t-\Delta} - M^c\right) \end{aligned}$$

Finally, we need to compute:

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

The dynamics of the adopter share are:

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

Thus we have:

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

Moreover,

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

So:

$$\begin{aligned}
PVX_{t-\Delta} &= \frac{1-e^{-k\Delta}}{1-e^{-(r+2k)\Delta}}\sum_{n=0}^{+\infty} e^{-(r+k)\Delta n}\mathbb{E}_{t-\Delta}\left[a_{t+\Delta n}\Delta\right] + \frac{\Delta e^{-k\Delta}}{1-e^{-(r+2k)\Delta}} X_{t-\Delta} \\
&= \frac{1-e^{-k\Delta}}{1-e^{-(r+2k)\Delta}} PVA_{t-\Delta} + \frac{\Delta e^{-k\Delta}}{1-e^{-(r+2k)\Delta}} X_{t-\Delta},
\end{aligned} \tag{14}$$

where:

$$PVA_{t-\Delta} = \sum_{n=0}^{+\infty} e^{-(r+k)\Delta n}\mathbb{E}_{t-\Delta}\left[a_{t+\Delta n}\Delta\right]. \tag{15}$$

Therefore,

$$\begin{aligned} B_{t-\Delta} &= \frac{\Delta}{1-e^{-(r+k)\Delta}}\left(M^e - M^c\right) + \frac{\Delta e^{-\theta\Delta}}{1-e^{-(r+k+\theta)\Delta}}\left(M^c - M_{t-\Delta}\right) \\ &\quad + \left\{ \frac{\Delta e^{-k\Delta}}{1-e^{-(r+2k)\Delta}} X_{t-\Delta} + \frac{1-e^{-k\Delta}}{1-e^{-(r+2k)\Delta}} PVA_{t-\Delta} \right\} \times C. \end{aligned}$$

This shows, in particular, that the value of adoption depends positively on the current level of adopters, so long as $k < +\infty$. This is the reason for the positive slope in the adoption frontier $\Phi(.)$.

A.2 Numerical solution method

In what follows we describe the numerical method for constructing the function $\Phi(.)$ that characterizes equilibrium adoption strategies in the model with complementarities.

First, given a mapping $\Phi(.) : [0,1] \to \mathbb{R}$, define the functions:

$$\begin{aligned} PVA(M_{t-\Delta}, X_{t-\Delta}; \Phi) &= \sum_{n=0}^{+\infty} e^{-(r+k)\Delta n} \mathbb{E}_{t-\Delta}\left[\mathbf{1}\left\{ M_{t+\Delta(n-1)} \geq \Phi(X_{t+\Delta(n-1)}) \right\} \Delta \right] \\ B(M_{t-\Delta}, X_{t-\Delta}; \Phi) &= \frac{\Delta}{1-e^{-(r+k)\Delta}}\left(M^e - M^c\right) + \frac{\Delta e^{-\theta\Delta}}{1-e^{-(r+k+\theta)\Delta}}\left(M^c - M_{t-\Delta}\right) \\ &\quad + \left\{ \frac{\Delta e^{-k\Delta}}{1-e^{-(r+2k)\Delta}} X_{t-\Delta} + \frac{1-e^{-k\Delta}}{1-e^{-(r+2k)\Delta}} PVA(M_{t-\Delta}, X_{t-\Delta}; \Phi) \right\} \times C. \end{aligned}$$

In the definition of the function $PVA(M_{t-\Delta}, X_{t-\Delta}; \Phi)$, the sequence $X_{t+\Delta(n-1)}$, in particular, is assumed to follow:

$$X_{t+\Delta n} = e^{-k\Delta} X_{t+\Delta(n-1)} + (1-e^{-k\Delta}) \mathbf{1}\left\{ M_{t+\Delta(n-1)} \geq \Phi(X_{t+\Delta(n-1)}) \right\},$$

starting from $(X_{t-\Delta}, M_{t-\Delta})$.

With these definitions, the algorithm proceeds as follows:

- **Initialization:** We derive a threshold rule $\underline{\Phi}(.)$ such that adoption of electronic money ($a_{e,t} = 1$) is a strictly dominant strategy, if and only if, $M_{t-\Delta} \leq \underline{\Phi}(X_{t-\Delta})$. For adoption of electronic money to be a strictly dominant strategy it must be that $B_{t-\Delta} \geq 0$ even if the firm expects no adoption at all by other firms, so that $PVA_{t-\Delta} = 0$. In that case:

$$B_{t-\Delta} = \frac{\Delta}{1-e^{-(r+k)\Delta}}\left(M^e - M^c\right) + \frac{\Delta e^{-\theta\Delta}}{1-e^{-(r+k+\theta)\Delta}}\left(M^c - M_{t-\Delta}\right) + \left\{ \frac{\Delta e^{-k\Delta}}{1-e^{-(r+2k)\Delta}} X_{t-\Delta} \right\} \times C,$$

  and so $B_{t-\Delta} \geq 0$, if and only if:

$$\begin{aligned} 0 &\leq M^e - M^c + \frac{e^{-\theta\Delta}(1-e^{-(r+k)\Delta})}{1-e^{-(r+k+\theta)\Delta}}\left(M^c - M_{t-\Delta}\right) + \left\{ \frac{e^{-k\Delta}(1-e^{-(r+k)\Delta})}{1-e^{-(r+2k)\Delta}} X_{t-\Delta} \right\} \times C \\ M_{t-\Delta} &\leq \underline{\Phi}(X_{t-\Delta}) = M^c - \frac{1-e^{-(r+k+\theta)\Delta}}{e^{-\theta\Delta} - e^{-(r+k+\theta)\Delta}}(M^c - M^e) + \frac{e^{-k\Delta} - e^{-(r+2k+\theta)\Delta}}{e^{-\theta\Delta} - e^{-(r+2k+\theta)\Delta}} C X_{t-\Delta} \end{aligned}$$

  Following similar steps, the upper threshold for $M_{t-\Delta}$ above which adoption of cash is a strictly dominant strategy is:

$$M_{t-\Delta} \geq \overline{\Phi}(X_{t-\Delta}) = \underline{\Phi}(X_{t-\Delta}) + \frac{e^{-k\Delta} - e^{-(r+2k+\theta)\Delta}}{e^{-\theta\Delta} - e^{-(r+2k+\theta)\Delta}} \frac{1-e^{-k\Delta}}{e^{-k\Delta} - e^{-(r+2k)\Delta}} C.$$

  Given these functions, we set $\underline{\Phi}^{(0)} = \underline{\Phi}$ and $\overline{\Phi}^{(0)} = \overline{\Phi}$.

- **Iteration:** At step n, given two functions $\underline{\Phi}^{(n)}$ and $\overline{\Phi}^{(n)}$, we compute their iterates as the solutions to:

$$B(\overline{\Phi}^{(n+1)}(X_{t-\Delta}), X_{t-\Delta}; \overline{\Phi}^{(n)}) = 0,$$

$$B(\underline{\Phi}^{(n+1)}(X_{t-\Delta}), X_{t-\Delta}; \underline{\Phi}^{(n)}) = 0.$$

These iterates are constructed on a linear grid for X.

- **Convergence:** We repeat the iteration step until $\max \left|\overline{\Phi}^{(n+1)}(.) - \overline{\Phi}^{(n)}(.)\right|$, $\max \left|\underline{\Phi}^{(n+1)}(.) - \underline{\Phi}^{(n)}(.)\right|$, and $\max \left|\overline{\Phi}^{(n+1)}(.) - \underline{\Phi}^{(n+1)}(.)\right|$ are below some threshold.

The only difficulties are in the computation of $PVA(M_{t-\Delta}, X_{t-\Delta}; \Phi)$, which in general has no closed form. To compute it, we use a Monte-Carlo approach: we simulate a large number of sample paths for the money stock starting at $M_{t-\Delta}$, and the implied path for $X_{t-\Delta}$ under the adoption rule $\Phi(.)$, and we then average across these sample paths. The threshold rule is interpolated linearly between the points of the grid for X.

A.3 The cash crunch in the frictionless model ($C = 0$ and $\kappa = 0$)

The left panel of figure B.1 reports the joint dynamics of (X_t, M_t) in the frictionless model. This graph is constructed under the assumption that $M^c > M^e$, so that on average, there are higher flow profits to technology c. The red line shows the average trajectory of a district which starts from point A, where $X_{-\Delta} = 0$ and $M_{-\Delta} = M^c$. At time 0, the shock shifts the economy from point A to point B. At point B, the stock of cash has fallen enough that the optimal technology choice of revising firms is to switch from c to e. As a result the number of firms using technology e, X_t, increases for a period of time. At the same time, the money stock reverts toward its long-run mean, M^c. After a certain time, it reaches the level $\underline{M}$ at which firms that revise their technology choice choose c over e.[67] In the long-run, the district will therefore converge back to point A.

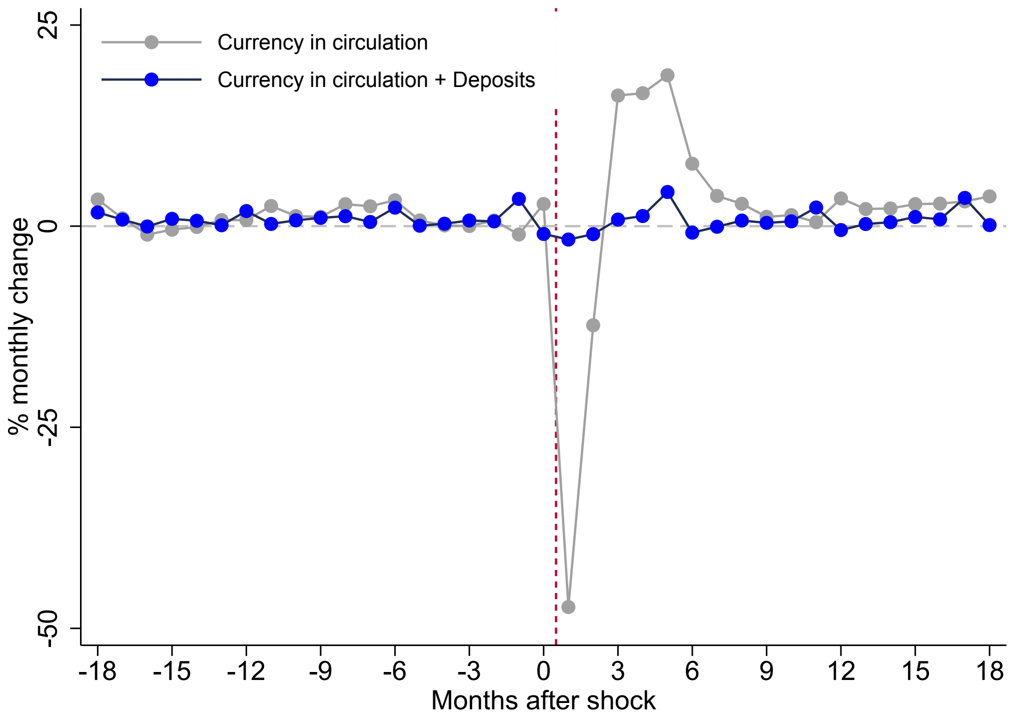
The top row of figure B.2 further illustrates this point. This graph plots the average response of a large number of districts to a common shock S. (The corresponding average path of M_t across the D districts is reported in figure 7.) On average across districts, the number of firms using technology e rises during the period when M_t is still substantially below its long-run mean, but thereafter rapidly returns to zero, since firms that revise their technology choice find it optimal to switch back to c once M_t is close enough to its long-run mean. Thus, the frictionless model cannot generate permanent increases in the number of firms using technology e out of a transitory shock to M_t. Consistent with this, the long-run response of districts is zero, and in particular, it is independent of their individual exposures, as reported on the left panel of figure B.3.

Additionally, the sequence of technology choices by firms in a district, following the shock, is independent of the initial fraction of firms already using technology e prior to the shock, $X_{-\Delta}$. The left panel of figure B.1 illustrates this, by also showing (in blue) the trajectory of a district starting from $X_{-\Delta} = 0.4 > 0$. In the long-run, this district also converges to zero adoption. For the same reasons as in the fixed cost model, the mechanical relationship between adoption level and adoption rate in the model then implies that the change in the number of users of e depends negatively on the initial number of users, as illustrated in the right panel of figure B.3.

[67] In the absence of complementarities ($C = 0$) or fixed costs ($\kappa = 0$), it is straightfoward to see (using equation 3) that the gross value of adoption, B_t, only depends on the level of cash, M_t. Therefore, the technology choice is entirely determined by the level of the aggregate shock, $M_{t-\Delta}$. One can then verify that, given the functional forms for flow profits, firms switch from c to e whenever $M_{t-\Delta} \leq \underline{M} = M^c - \frac{1 - e^{-(r+k+\theta)\Delta}}{e^{-\theta\Delta} - e^{-(r+k+\theta)\Delta}} (M^c - M^e)$. When shocks are purely transitory ($\theta = +\infty$), firms either always or never switch (depending on whether $M^e \gtrless M^c$), while when shocks are permanent $\theta = 0$, firms switch as soon a shock pushes M_t below the flow profits from technology e in the absence of complementarities, M^e.

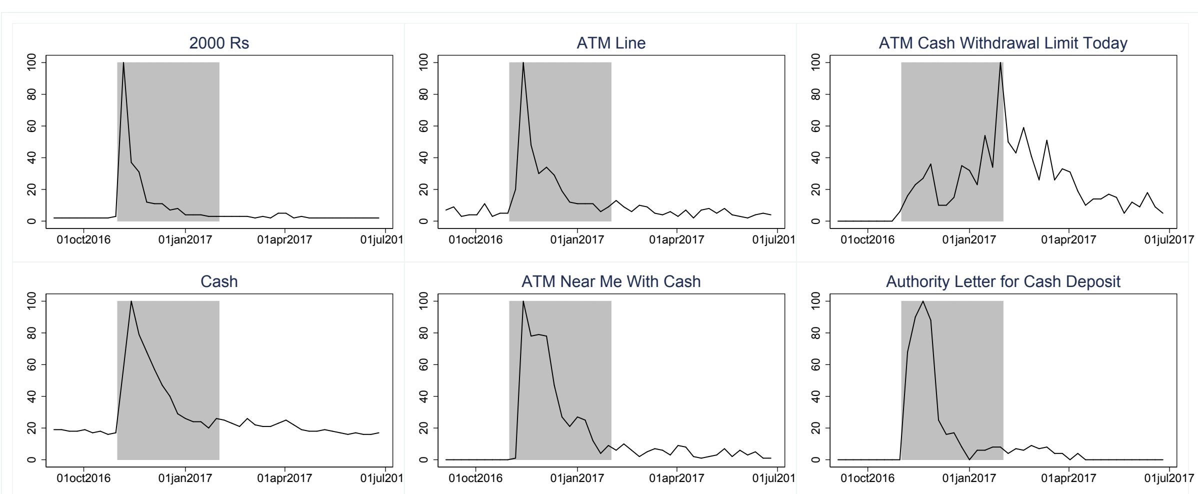
B Figures and tables

Figure 2: Change in nominal value of currency in circulation



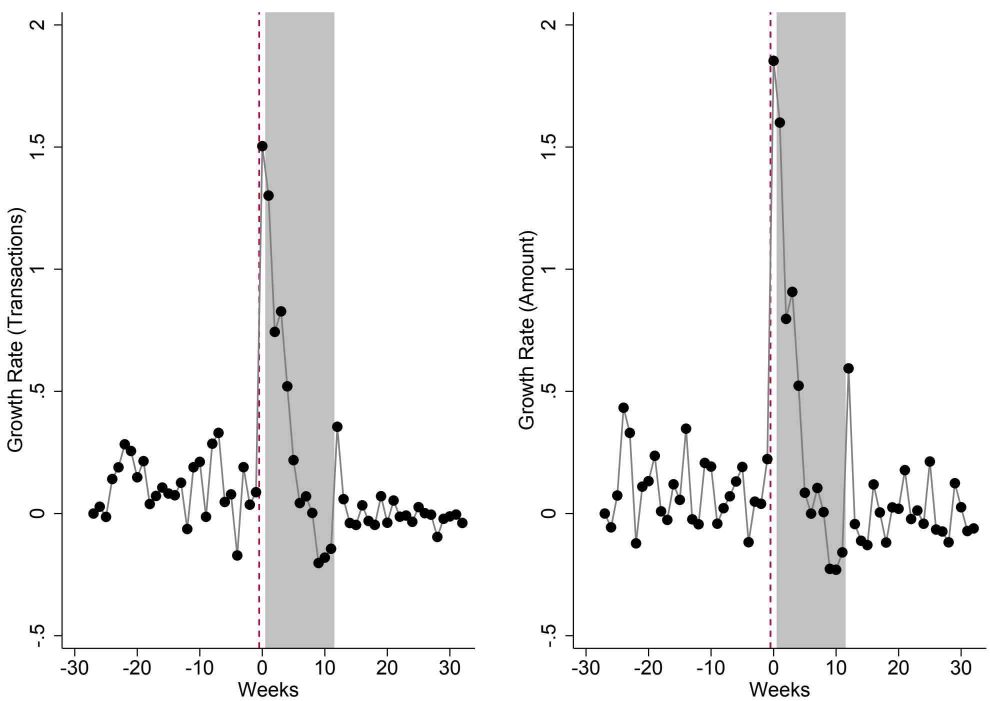
Notes: The figure shows the change in the nominal value of the stock of currency in circulation (in grey) and change in the value of total money supply (in blue) in India. Month 0 is the month of October 2016; the figures are end-of-month estimates. Source: Reserve Bank of India.

Figure 3: Evidence from Google Search Trends



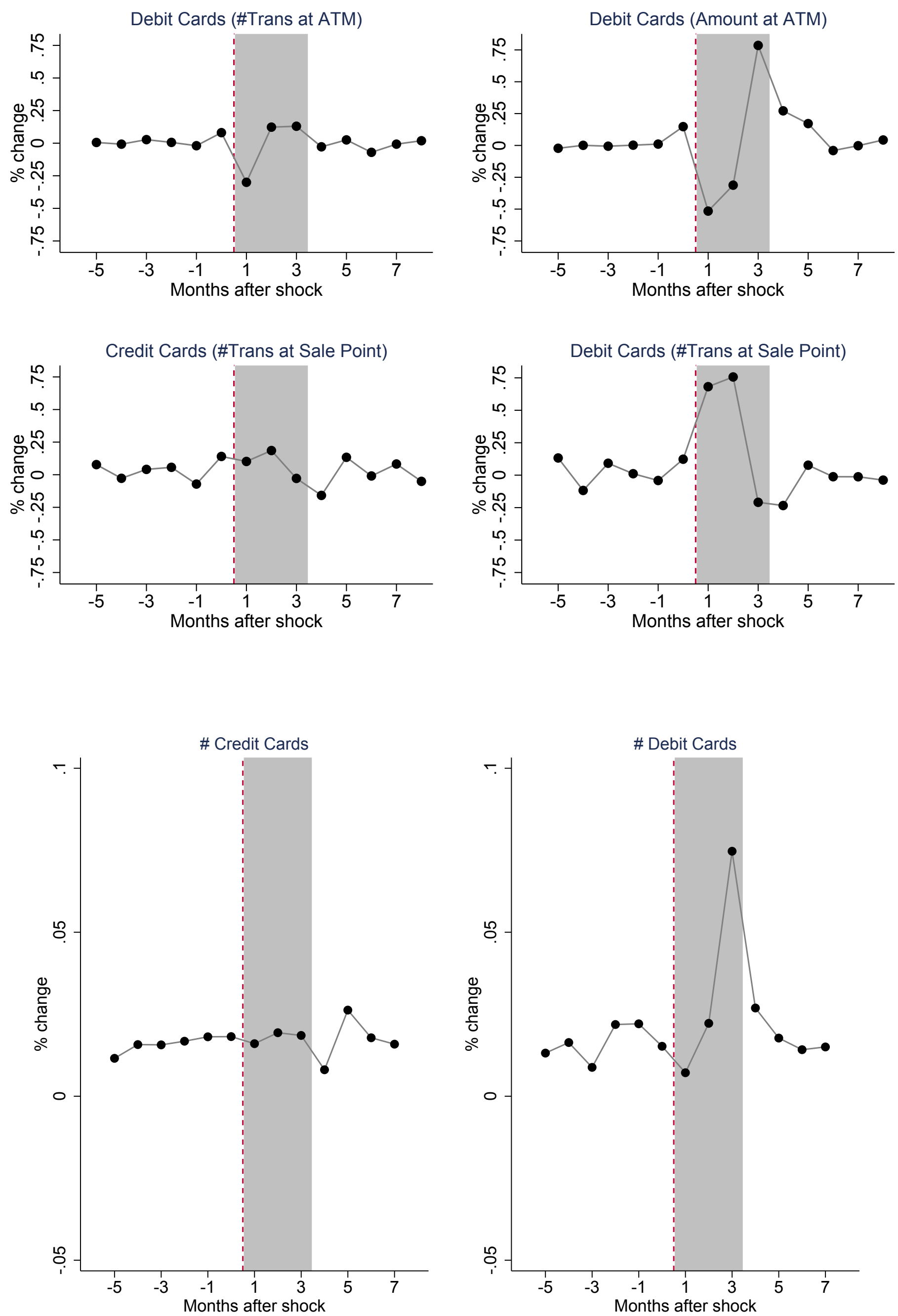
Notes: The figure reports the daily plot between September 2016 and July 2017 of Google Searches for several key words that could be representative of public actions and information associated with the demonetization shocks. Data is obtained by Google Trends, and the index is normalized by Google to be 0 to 100, with value of 100 assigned to the day with maximum searches made for that topic. Source: Google Search Index.

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: 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 aggregated at the national level. The x-axis represents month, where October 2016 is normalized to be zero. Source: Reserve Bank of India.

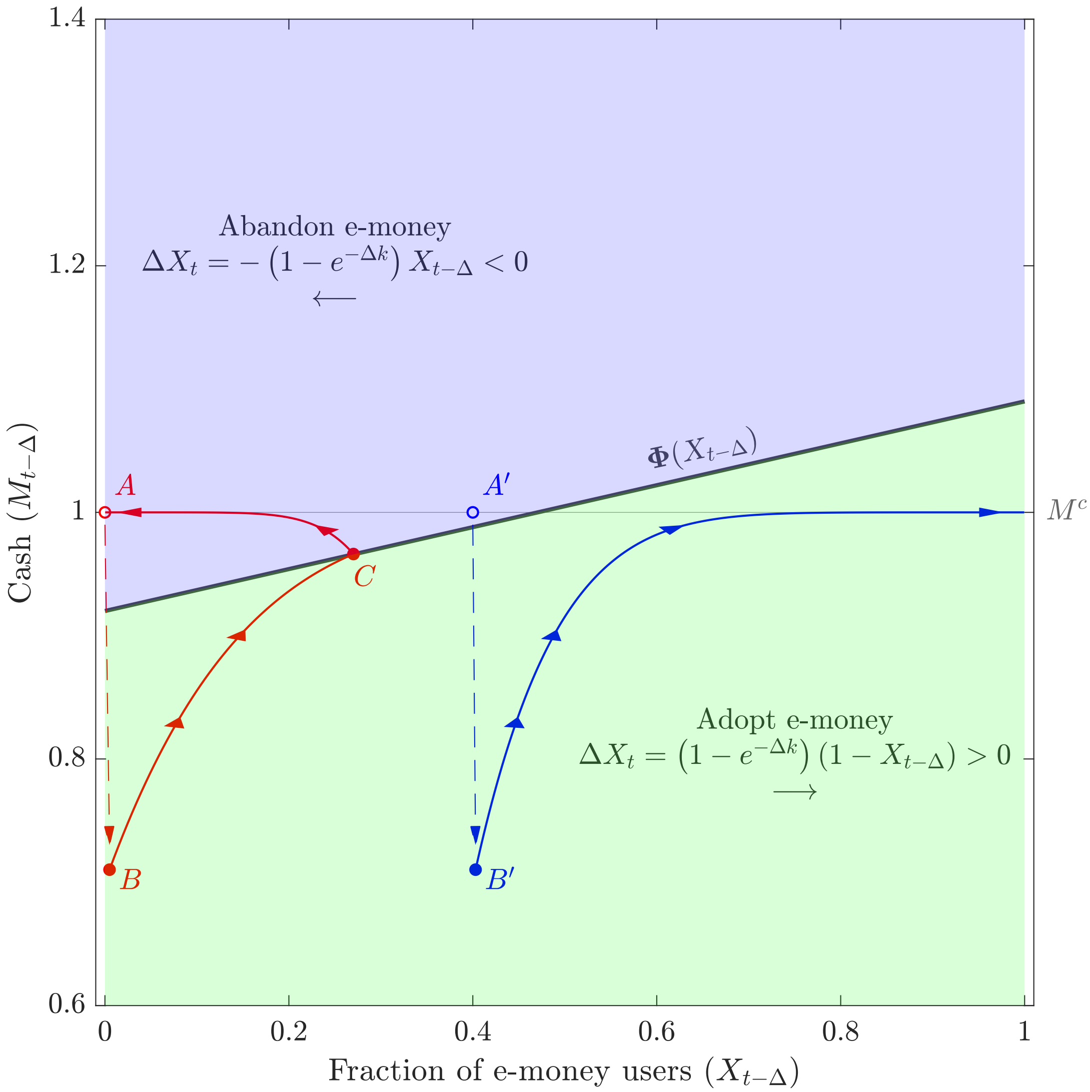


Figure 6: Adoption dynamics in response to a large decline in M_t in the complementarities model ($C > 0$ and $\kappa = 0$). The model illustrated here corresponds to the case $\theta > k$ (the shock is transitory relative to the adjustment speed of firms.) The red line shows the path of a district that starts with a low adoption level $X_{d,0} = 0$. The blue line shows the path of a district that start with a high adoption level, $X_{d,0} = 0.4$. The paths are constructed assuming that each district receives no other shock than the initial decline in M_t, i.e. that $\epsilon_t = 0$ for all $t > 0$.

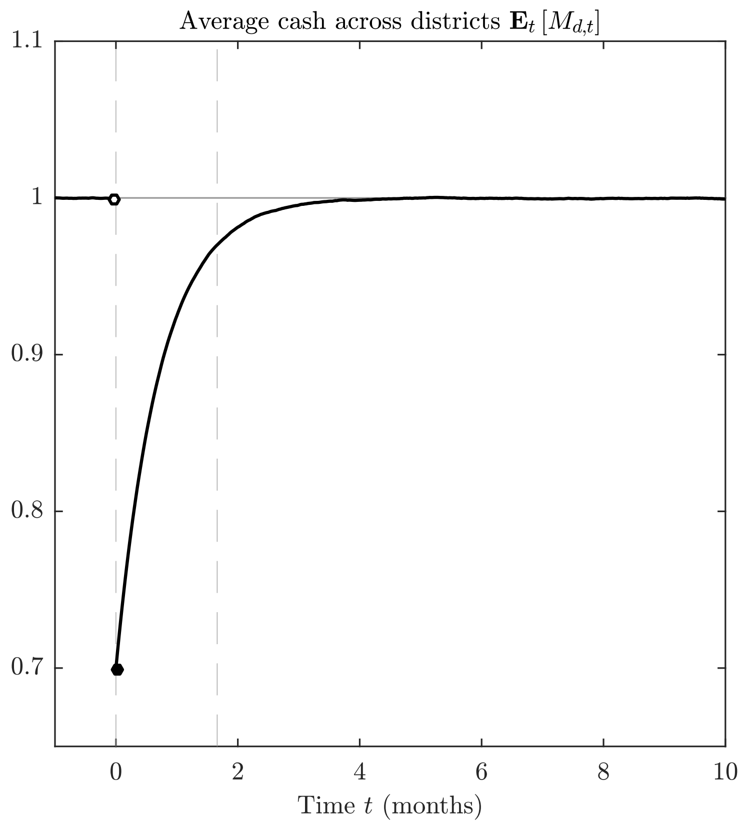


Figure 7: Path of the average level of cash across districts, $\mathbb{E}_t\left[M_{d,t}\right]$, after the cash crunch. The first grey dashed line indicates the date of the shock, and the second one indicates the date at which M_t is back to within 90% of its long-run value, $M_t = M^c = 1$. The model is simulated for $D = 10^4$ districts, with a burn-in period of 5 years. The persistence of the shock is $\theta = 1.38$, corresponding to a half-life of two weeks.

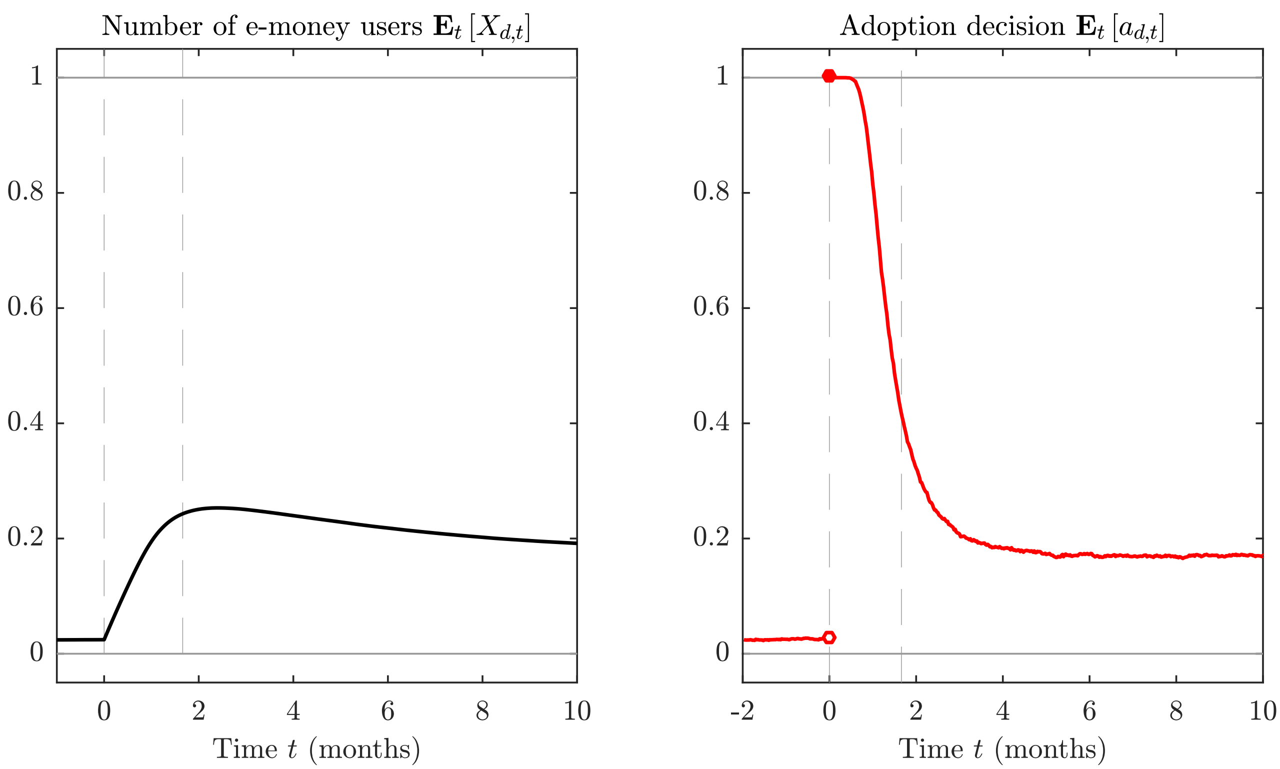


Figure 8: Average number of users ($\mathbb{E}_t\left[X_{d,t}\right]$, left column) and average adoption decision ($\mathbb{E}_t\left[a_{d,t}\right]$, right column) after the cash crunch in the complementarities model ($C > 0$ and $\kappa = 0$). The results reported here are generated using a version of the model where $\theta > k$ (the shock is transitory relative to the adjustment speed of firms.) Specifically, the model is solved with $k = 0.2$, corresponding to an average waiting time between technology resets of 5.0 months, while the persistence of the shock is $\theta = 1.38$, correspond to a half-life of two weeks.

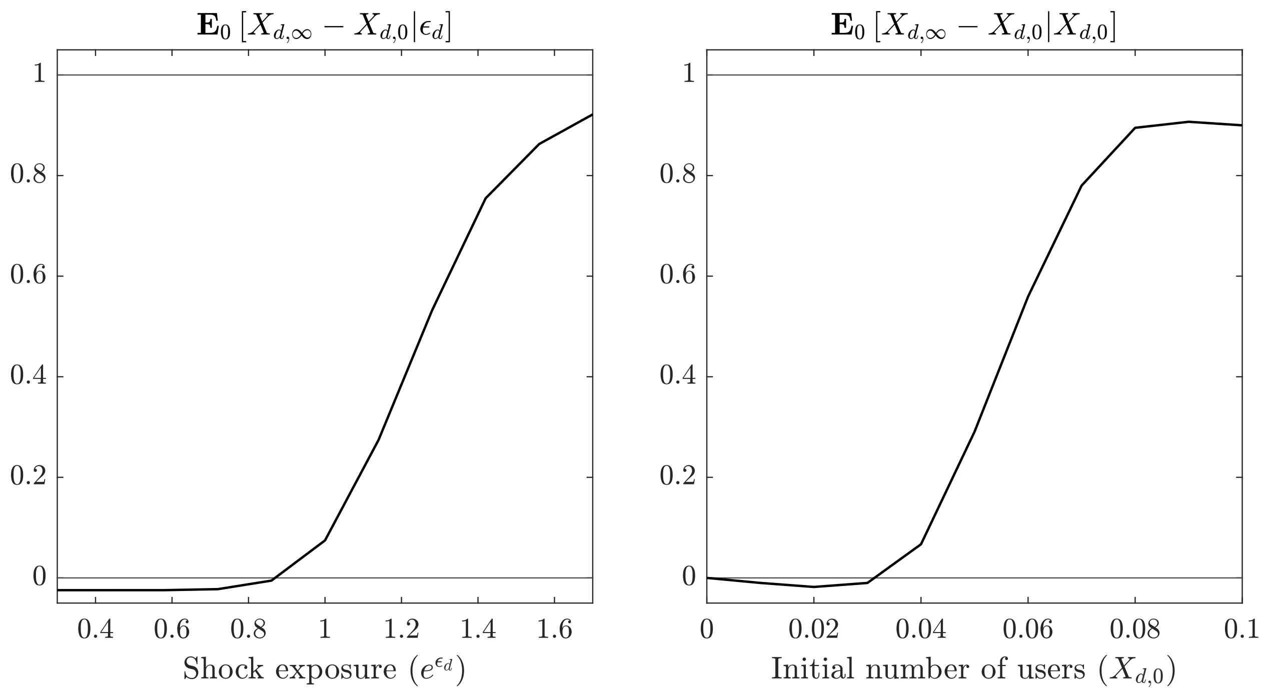


Figure 9: Conditional impulse responses in the complementarities model ($C > 0$ and $\kappa = 0$). These impulse responses are generated from the same type of simulations as figure 8. The left column reports the relationship between the district's exposure to the shock, proxied by e^{ϵ_d} (with a value of 1 indicating an average exposure to the shock), and the long-run change in the number of users after the shock. The right column reports the relationship between the initial number of users, $X_{d,0}$, and the long-run change in the number of users after the shock. The long-run number of users in the left panel is defined as $E_0\left[X_{d,\infty} - X_{d,0}|\epsilon_d\right] = \lim_{t\to+\infty} E_0\left[X_{d,t} - X_{d,0}|\epsilon_d\right]$ (and similarly for the right panel). For both columns, in each district, the adoption path is constructed by averaging across 10^3 draws. The limit as $t \to \infty$ is obtained by simulating the response of each district for five years, and using the end-of-simulation values.

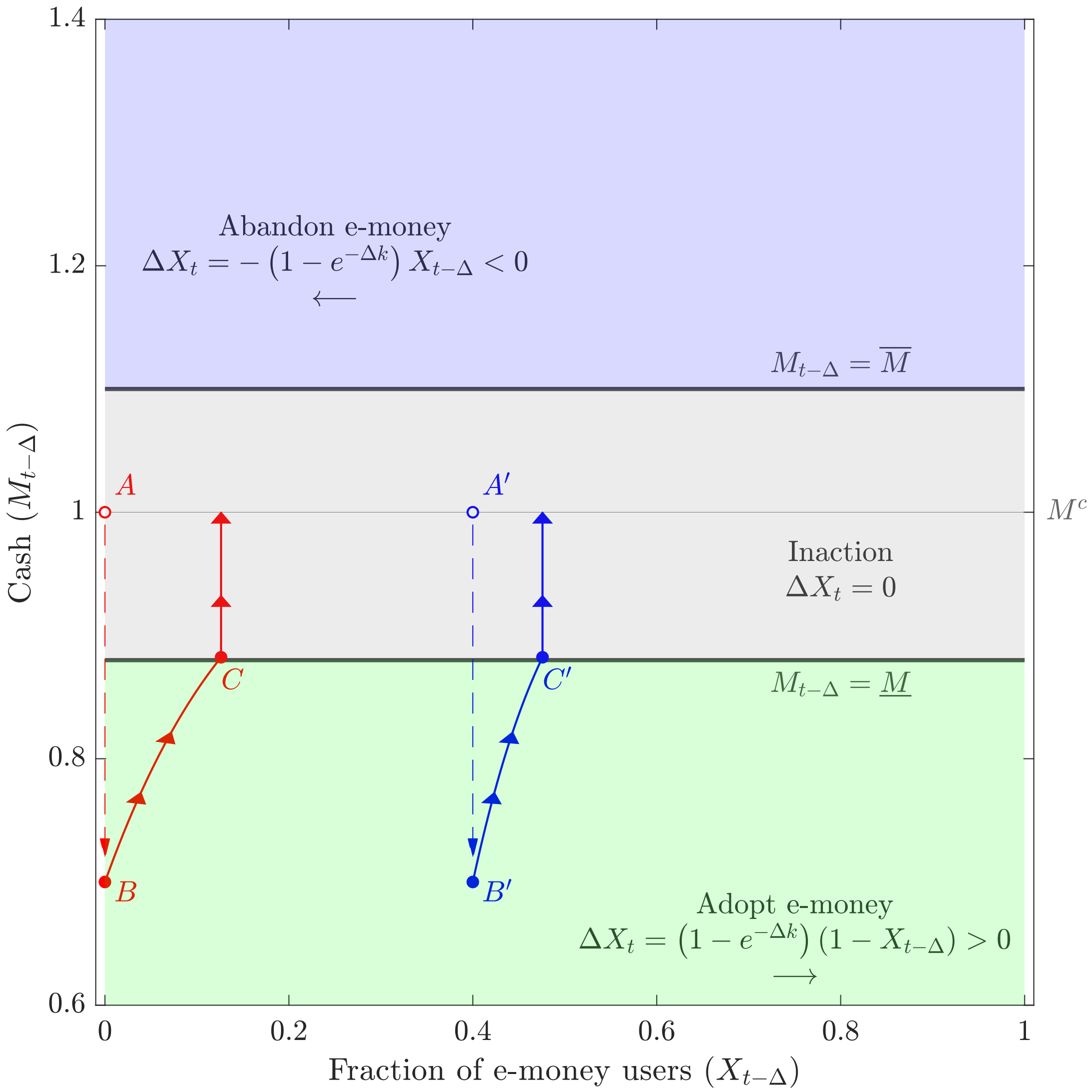


Figure 10: Adoption dynamics in response to a large decline in M_t in the fixed cost model ($C = 0$ and $\kappa > 0$). The red line shows the path of a district that starts with a low adoption level $X_{d,0} = 0$. The blue line shows the path of a district that start with a high adoption level, $X_{d,0} = 0.4$. The paths are constructed assuming that each district receives no other shock than the initial decline in M_t, i.e. that $\epsilon_t = 0$ for all $t > 0$.

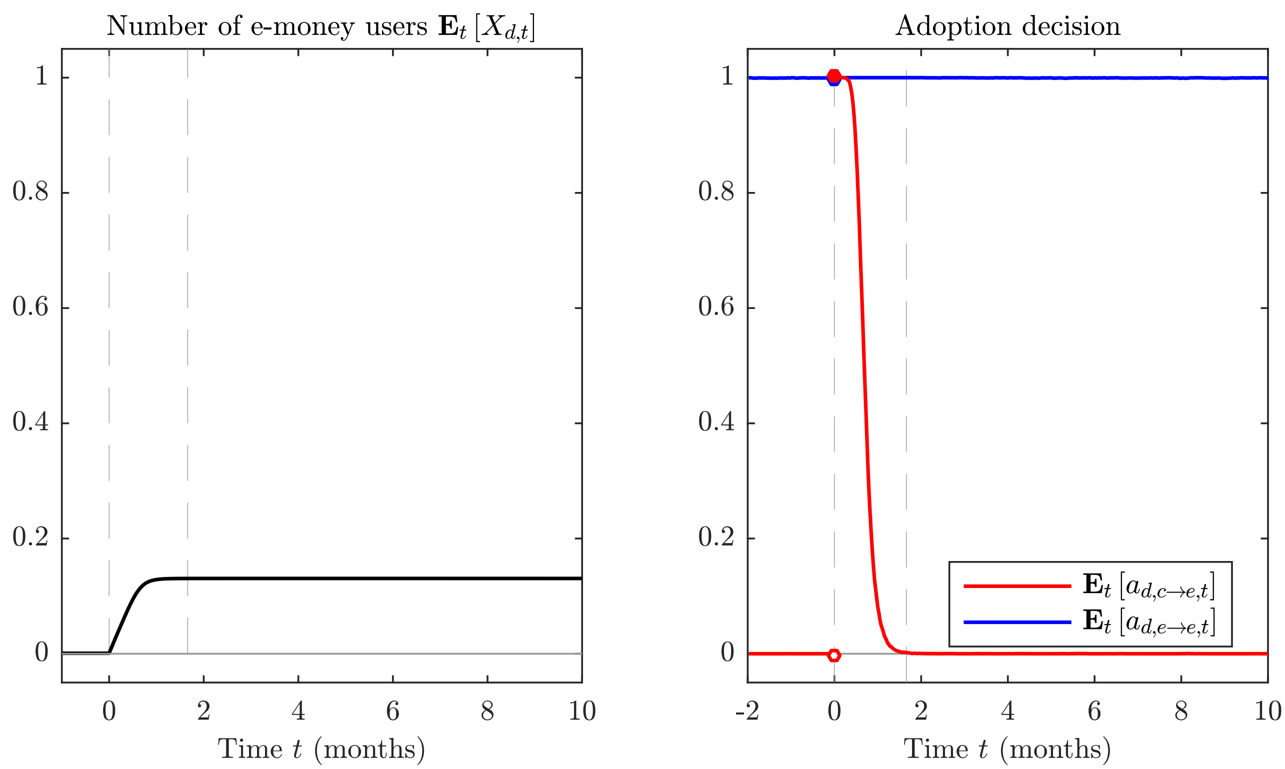


Figure 11: Average number of users ($\mathbb{E}_t\left[X_{d,t}\right]$, left column) and average adoption decisions (right column) after the cash crunch in the fixed cost model ($C = 0$ and $\kappa > 0$). The graph on the right panel reports separately the adoption decision of firms currently using cash and the adoption decision of firms currently using electronic money. The calibration assumes that $M^e > M^c$ and $k = 0.2$, corresponding to an average waiting time between technology resets of 5.0 months.

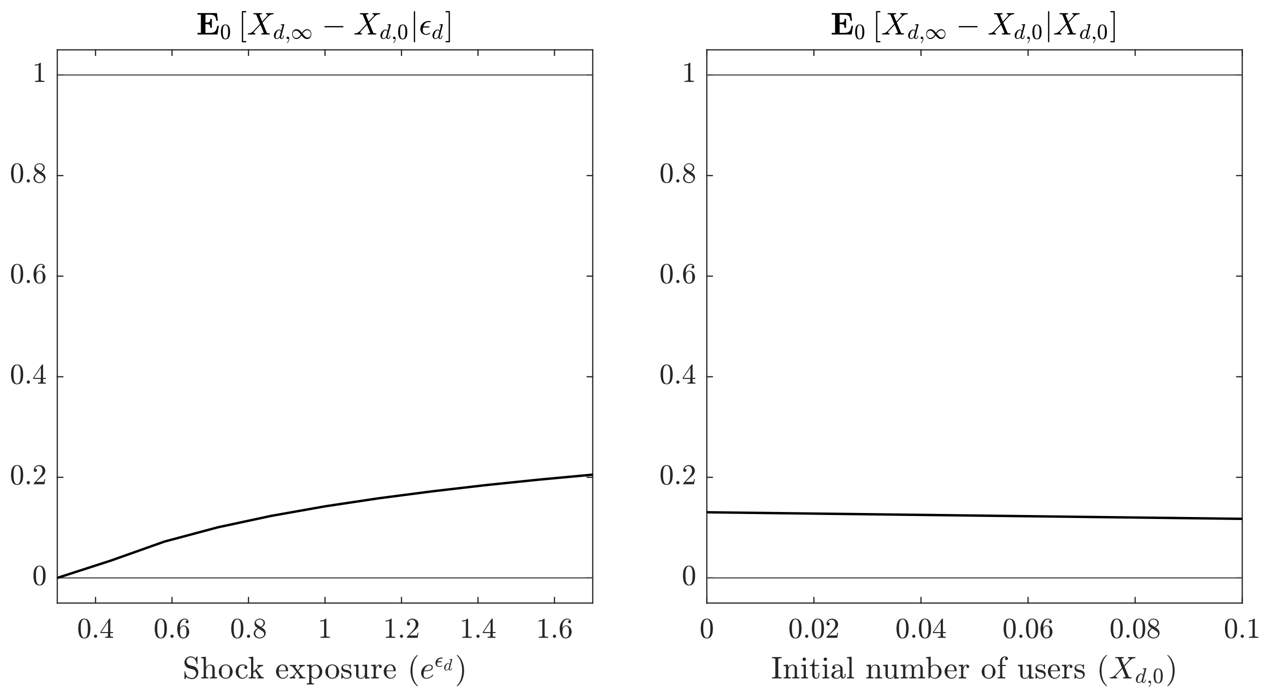


Figure 12: Conditional impulse responses in the fixed cost model ($C > 0$ and $\kappa = 0$). These impulse responses are generated from the same type of simulations as figure 11. The left column reports the relationship between the district's exposure to the shock, proxied by e^{ϵ_d} (with a value of 1 indicating an average exposure to the shock), and the long-run change in the number of users after the shock. The right column reports the relationship between the initial number of users, $X_{d,0}$, and the long-run change in the number of users after the shock. The long-run number of users in the left panel is defined as $E_0\left[X_{d,\infty} - X_{d,0}|\epsilon_d\right] = \lim_{t\to+\infty} E_0\left[X_{d,t} - X_{d,0}|\epsilon_d\right]$ (and similarly for the right panel). For both columns, in each district, the adoption path is constructed by averaging across 10^3 draws. The limit as $t \to \infty$ is obtained by simulating the response of each district for five years, and using the end-of-simulation values.

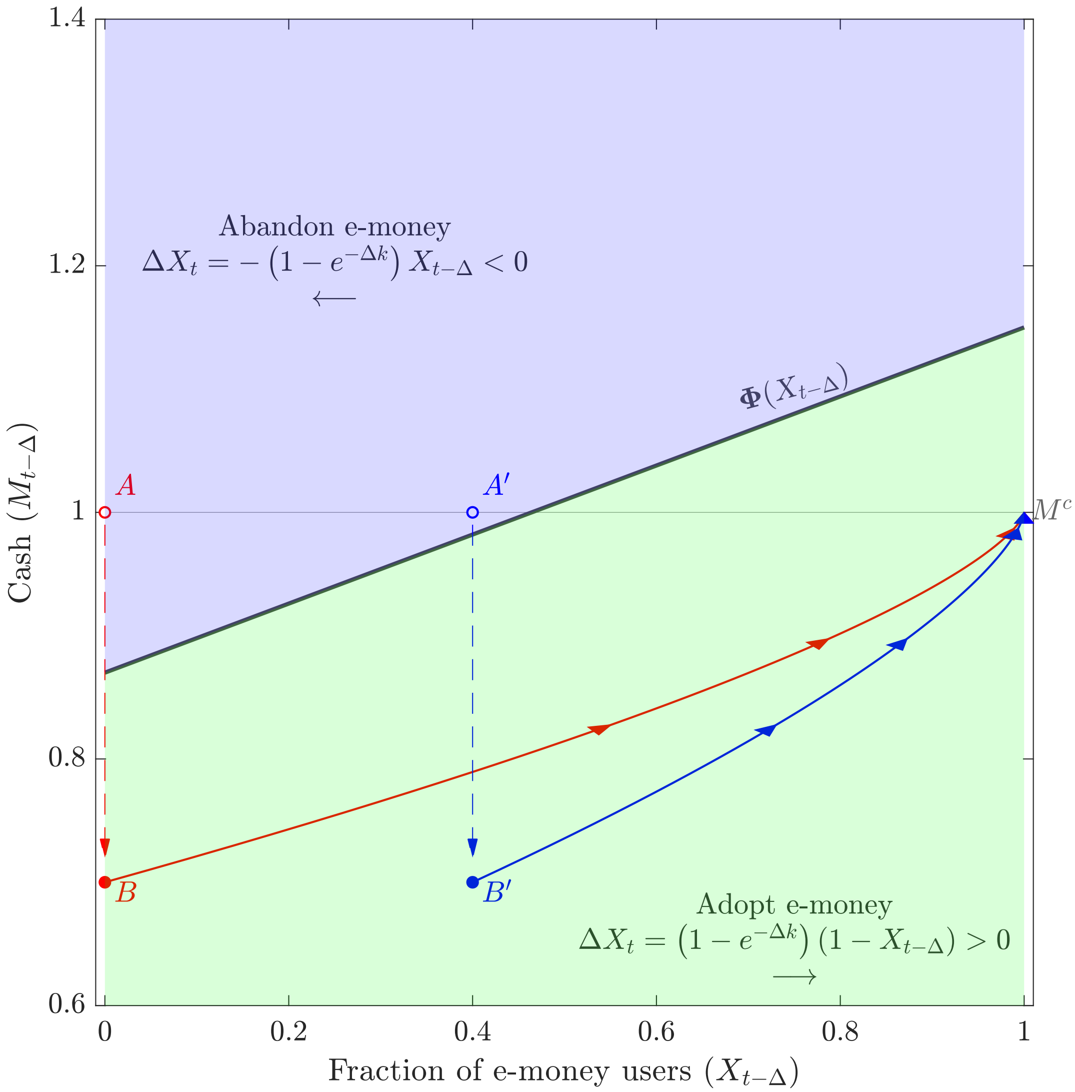


Figure 13: Adoption dynamics in response to a large decline in M_t in the complementarities model ($C > 0$ and $\kappa = 0$). The model illustrated here corresponds to the case $\theta < k$ (the shock is transitory relative to the adjustment speed of firms.) The red line shows the path of a district that starts with a low adoption level $X_{d,0} = 0$. The blue line shows the path of a district that start with a high adoption level, $X_{d,0} = 0.4$. The paths are constructed assuming that each district receives no other shock than the initial decline in M_t, i.e. that $\epsilon_t = 0$ for all $t > 0$.

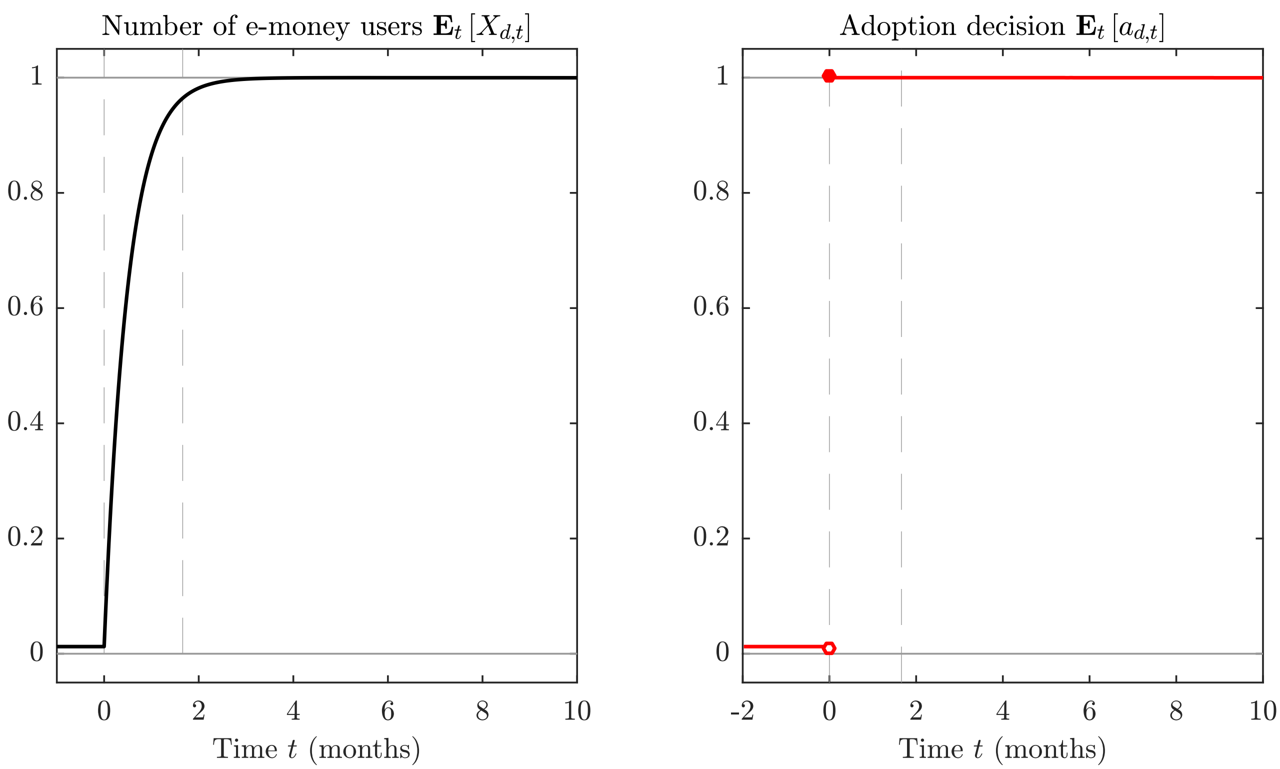


Figure 14: Average number of users ($\mathbb{E}_t\left[X_{d,t}\right]$, left column) and average adoption decision ($\mathbb{E}_t\left[a_{d,t}\right]$, right column) after the cash crunch in the complementarities model ($C > 0$ and $\kappa = 0$). The results reported here are generated using a version of the model where $\theta < k$ (the shock is persistent relative to the adjustment speed of firms.) Specifically, the model is solved with $k = 2$, corresponding to an average waiting time between technology resets of 2 weeks, while the persistence of the shock is $\theta = 1.38$, correspond to a half-life of two weeks.

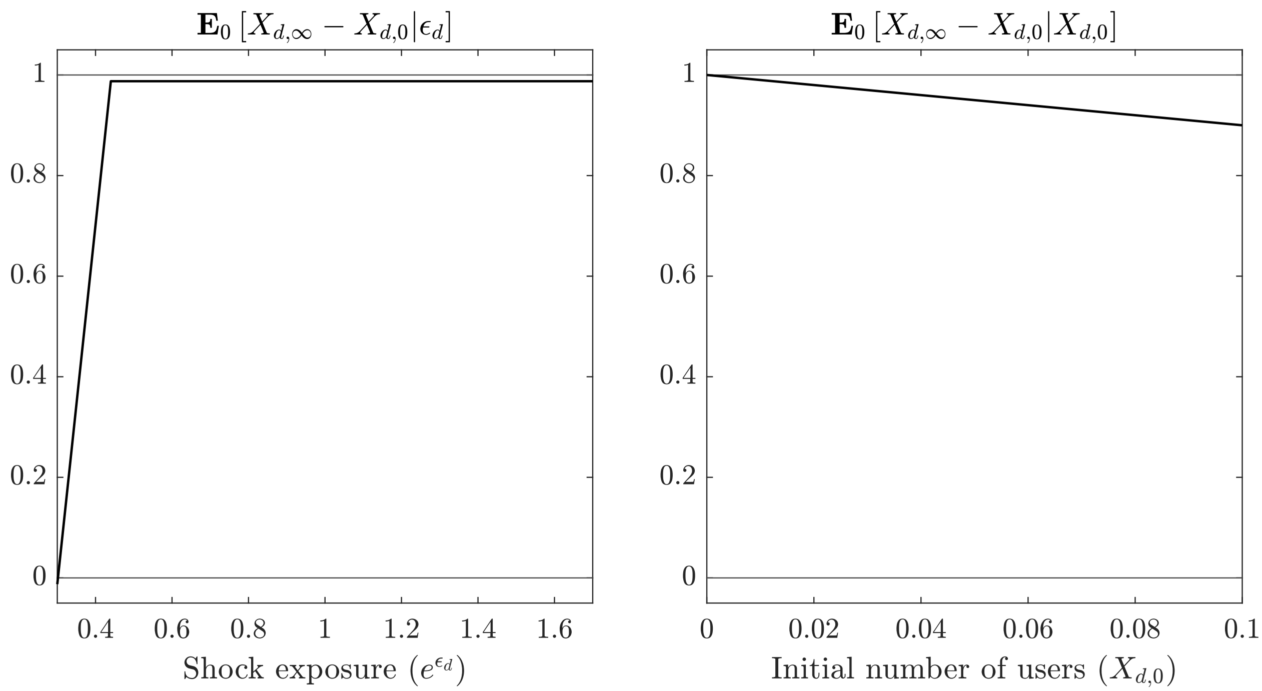
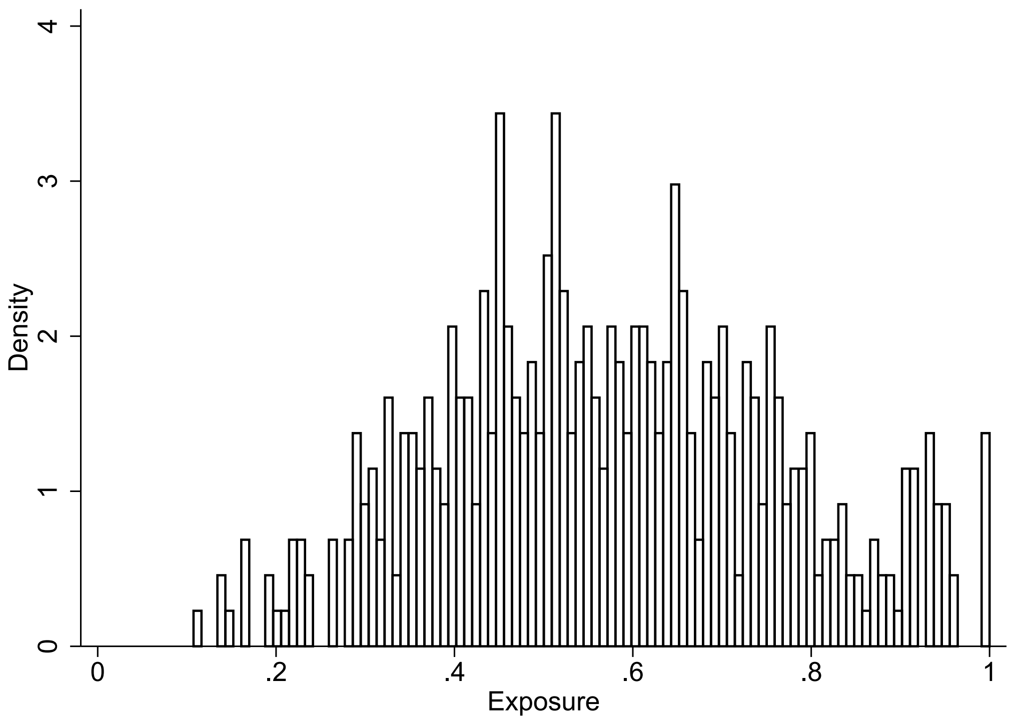


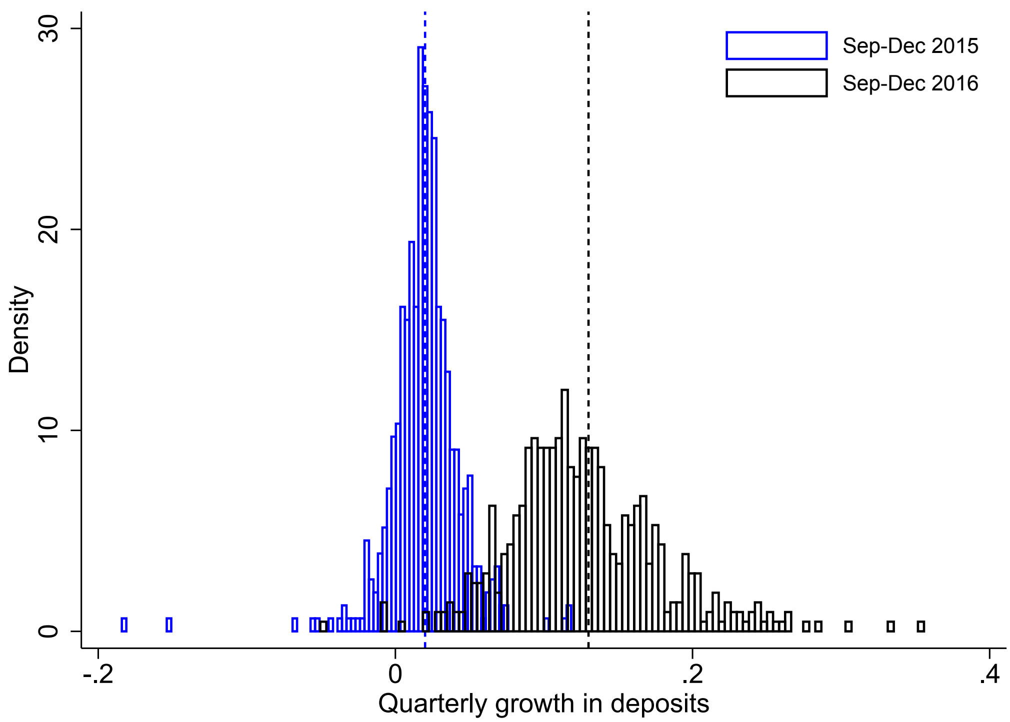
Figure 15: Conditional impulse responses in the complementarities model ($C > 0$ and $\kappa = 0$). These impulse responses are generated from the same type of simulations as figure 14, that is, the case where $\theta < k$ (the shock is persistent relative to the adjustment speed of firms.) The left column reports the relationship between the district's exposure to the shock, proxied by e^{ϵ_d} (with a value of 1 indicating an average exposure to the shock), and the long-run change in the number of users after the shock. The right column reports the relationship between the initial number of users, $X_{d,0}$, and the long-run change in the number of users after the shock. The long-run number of users in the left panel is defined as $E_0 [X_{d,\infty} - X_{d,0} | \epsilon_d] = \lim_{t \to +\infty} E_0 [X_{d,t} - X_{d,0} | \epsilon_d]$ (and similarly for the right panel). For both columns, in each district, the adoption path is constructed by averaging across 10^3 draws. The limit as $t \to \infty$ is obtained by simulating the response of each district for five years, and using the end-of-simulation values.

Figure 16: Distribution of Exposure_d across districts



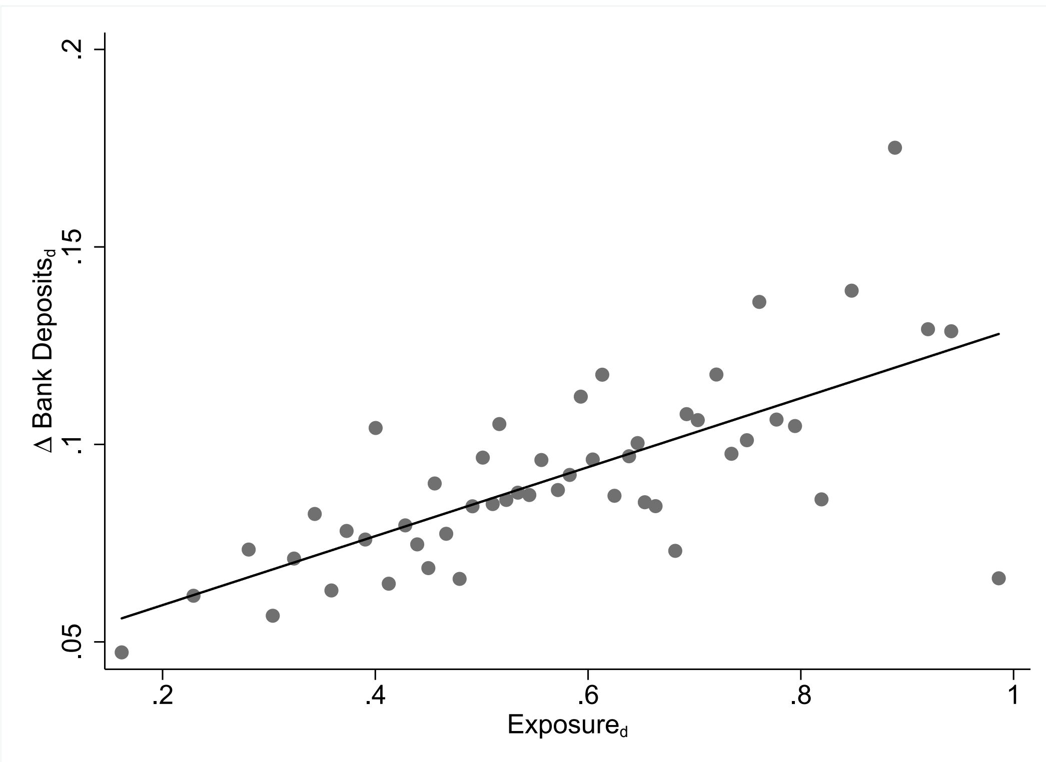
Notes: The figure shows the distribution of Exposure_d (as described in Section 2) across Indian districts. Source: Reserve Bank of India

Figure 17: Distribution of growth in deposits across districts



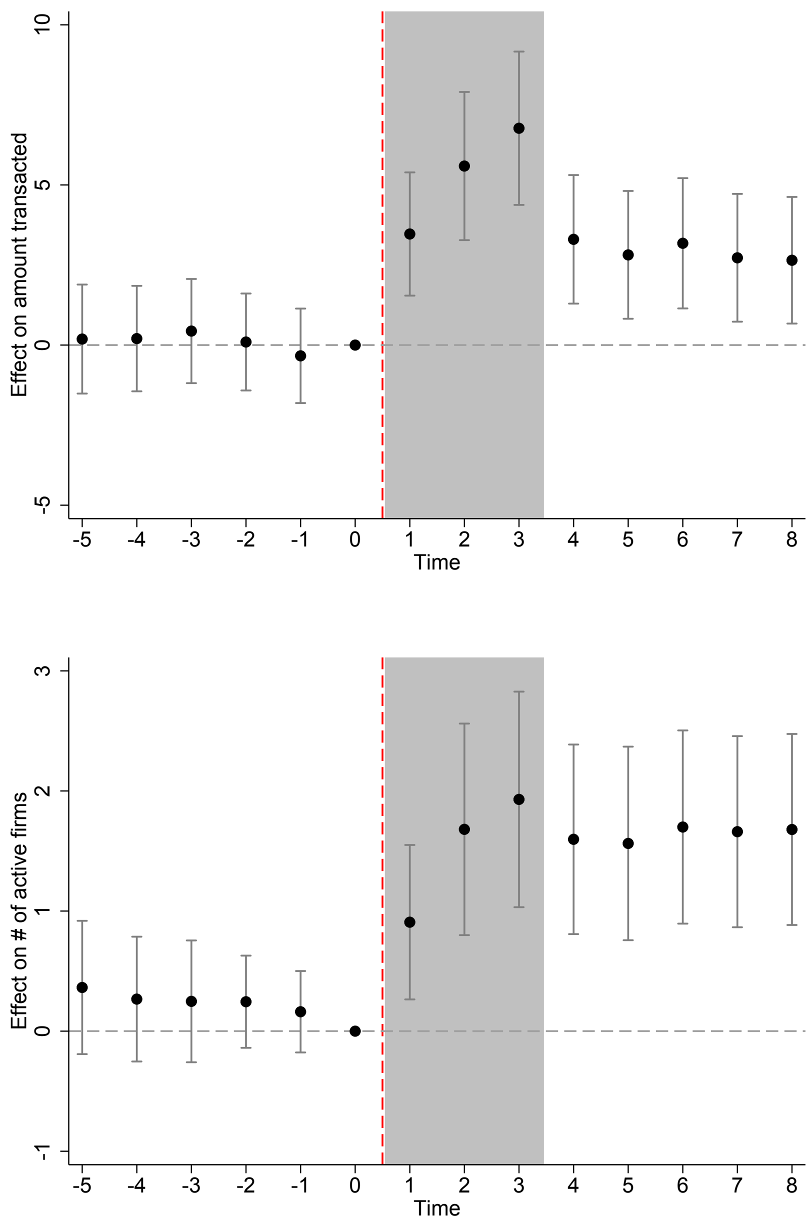
Notes: Distribution across deposits of the growth in total banking sector deposits from October to December during year 2015 (blue) and 2016 (black). The vertical dashed lines represent the corresponding mean deposit growth for these years. Source: Reserve Bank of India.

Figure 18: Relation between Exposure and 2016 Q4 deposit growth



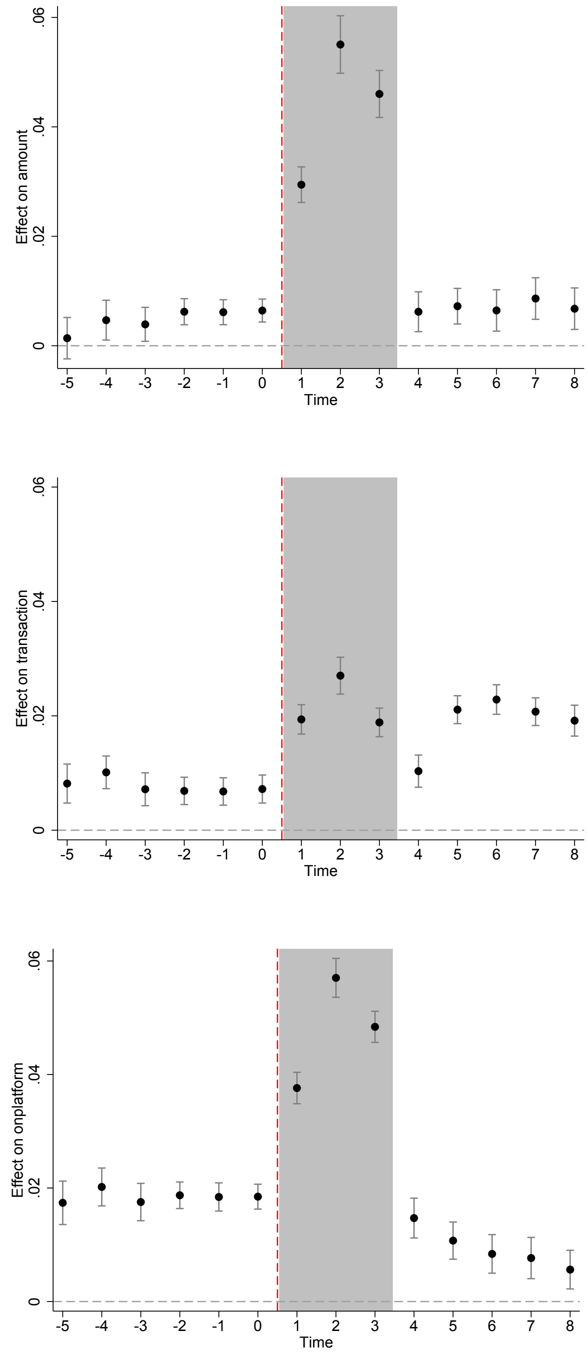
Notes: The figure shows the relation between our measure of Exposure$_d$ (as described in Section 2) and change in bank deposits in the district between September 30,2016 and December 31,2016 *i.e.* during the quarter of demonetization. Source: Reserve Bank of India

Figure 19: District adoption dynamics in electronic payments data based on exposure to shock



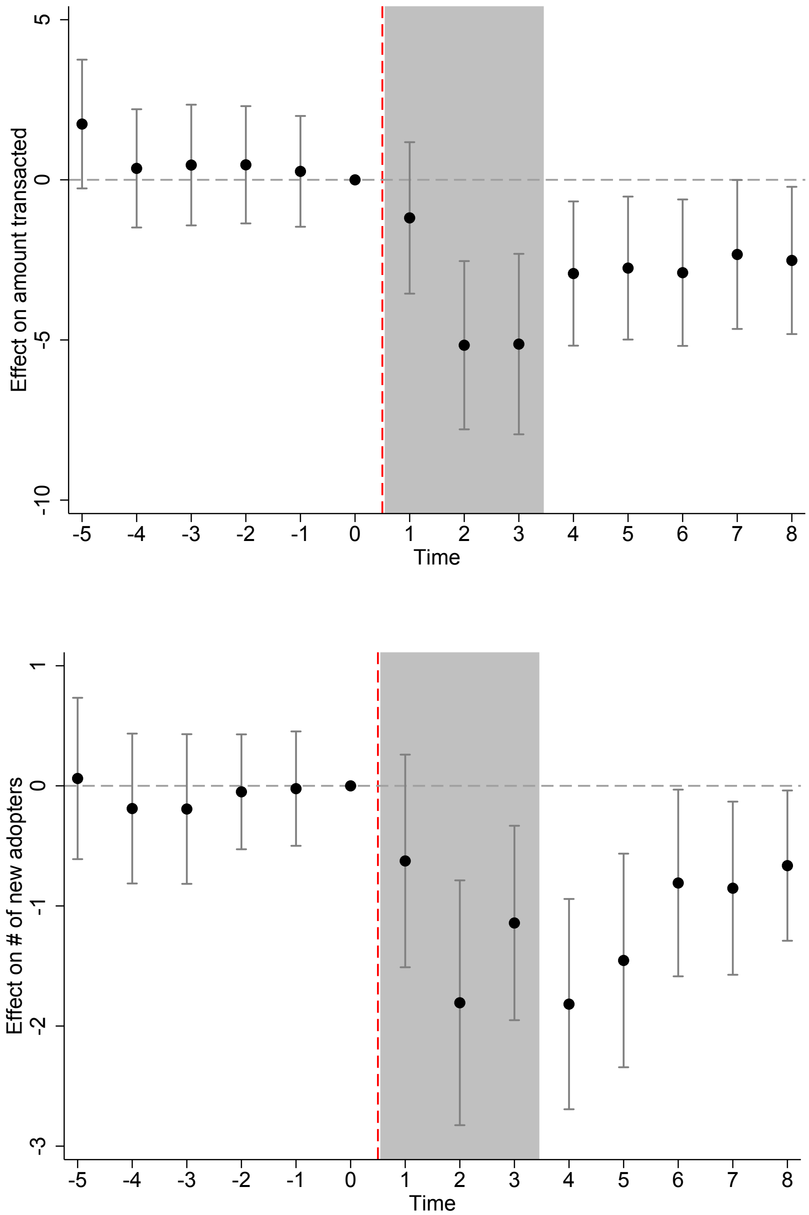
Notes: The figure plots the dynamic treatment effects of the demonetization shock on technology adoption of electronic payment systems. The graphs report the coefficients δ_t from specification 8; the top panel reports effects for the the total amount of transactions (in logs), and the bottom panel reports effects for total number of active firms on the platform (in logs). The x-axis represents month, where October 2016 is normalized to be zero. 95% confidence intervals are represented with the vertical lines; standard errors are clustered at the district level.

Figure 20: Firm adoption dynamics in electronic payments data based on existing share of adopters



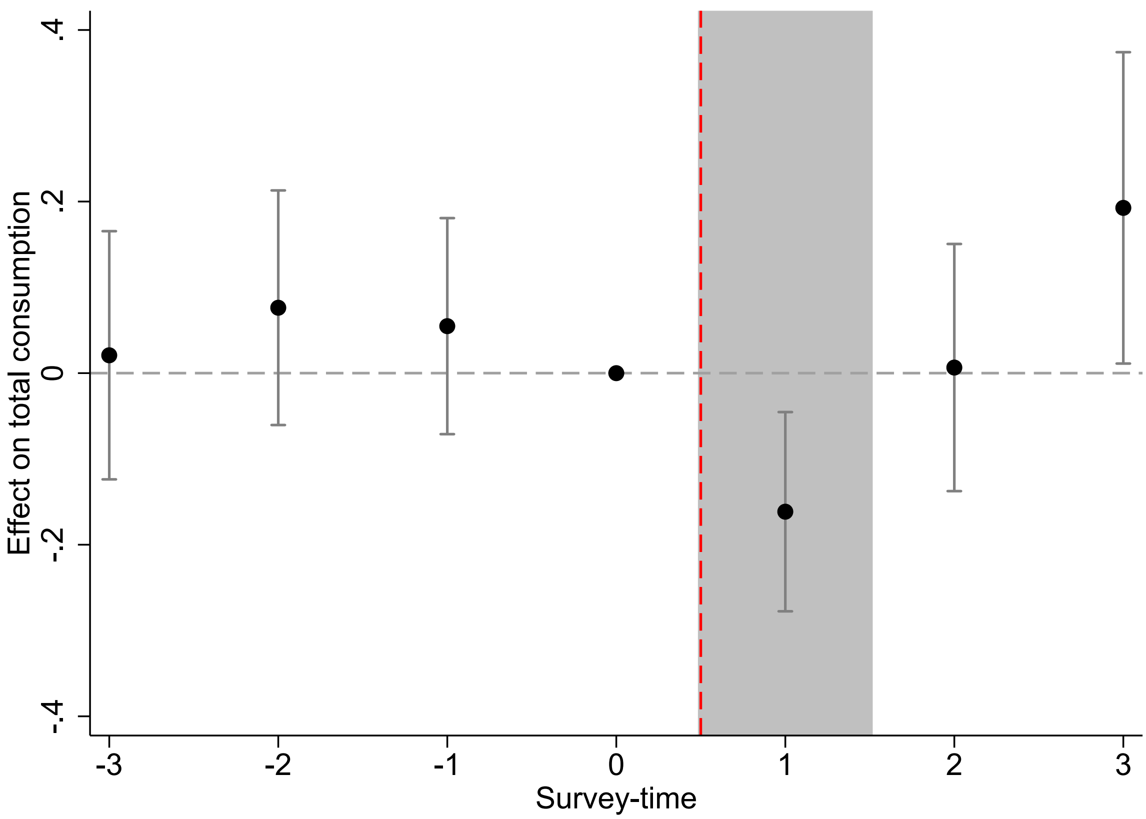
Notes: The figure plots month-by-month estimates of the dependence of firm-level adoption rates on the share of other adopters in the industry/pincode. The specification we estimate is a version of equation 10 in which each coefficient is interacted with a weekly dummy; we reported the monthly estimates of the coefficient γ. The top panel reports effects when x is the total amount 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 21: 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 as proxied by the distance of that district to the closest district with more than 500 active firms before demonetization. The specification we estimate δ_t in the dynamic version of equation 11. The top panel reports effects for the the total amount of transactions (in logs), and the bottom panel reports effects for total number of new firms transacting on the platform (in logs). The x-axis represents month, where October 2016 is normalized to be zero. 95% confidence intervals are represented with the vertical lines; standard errors are clustered at the district level

Figure 22: Consumption responses based on exposure to the shock



Notes: The figure plots survey month-by-survey month estimates of consumption responses depending on exposure to the shock (Exposure$_d$). The specification we estimate is a version of equation 13 in which each coefficient is based on the interaction of the treatment variable with a event-time dummy. We report the event-time estimates of the coefficient δ. Treatment is our measure of Exposure$_d$ as described in Section 2. The dependent variable on y-axis is the (log) total expense by households (as described in Section 4). 95% confidence intervals are represented with the vertical lines; standard errors are clustered at the district level. Source: CMIE Consumption Data.

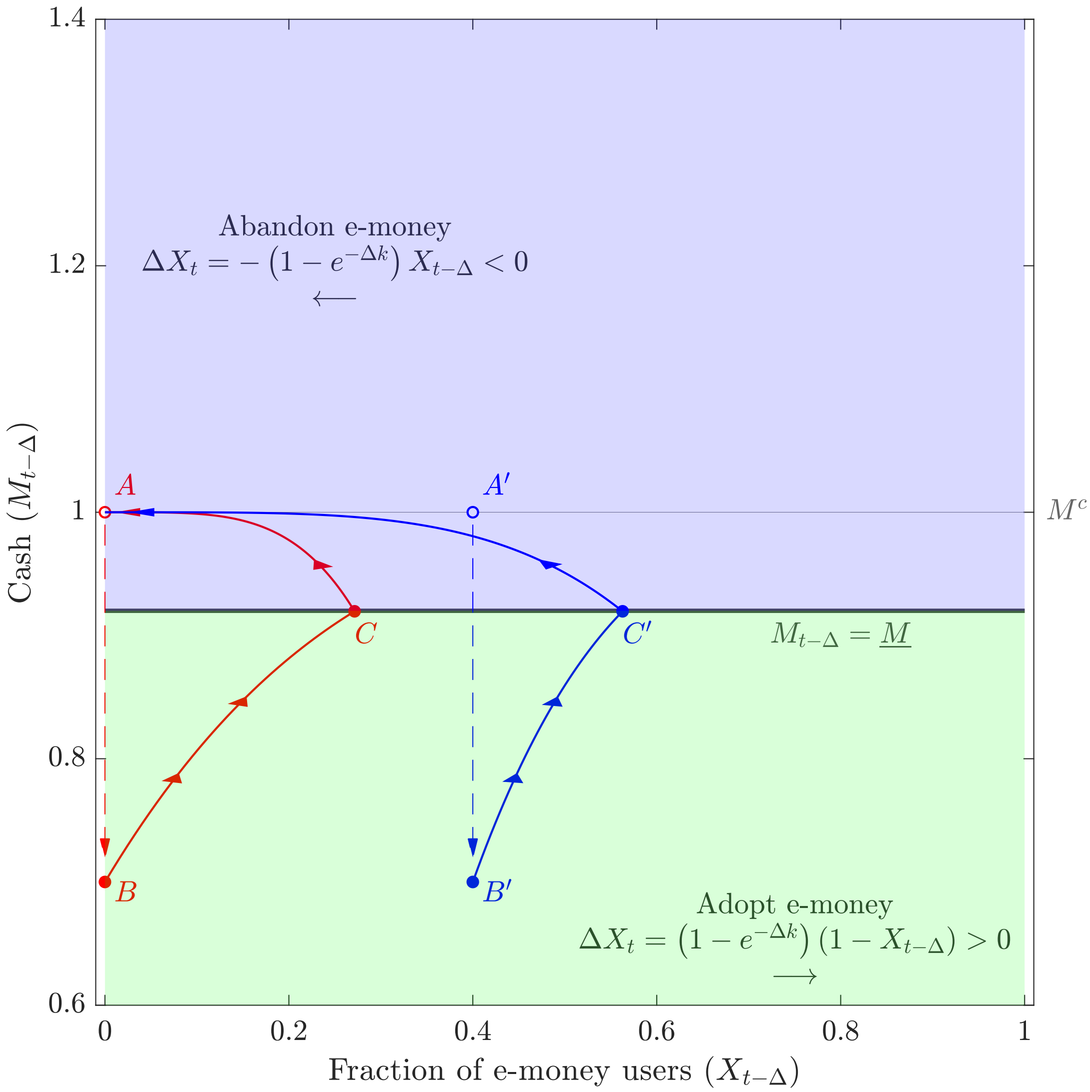


Figure B.1: Adoption dynamics in response to a large decline in M_t in the frictionless model ($C = 0$ and $\kappa = 0$). The red line shows the path of a district that starts with a low adoption level $X_{d,0} = 0$. The blue line shows the path of a district that start with a high adoption level, $X_{d,0} = 0.4$. The paths are constructed assuming that each district receives no other shock than the initial decline in M_t, i.e. that $\epsilon_t = 0$ for all $t > 0$.

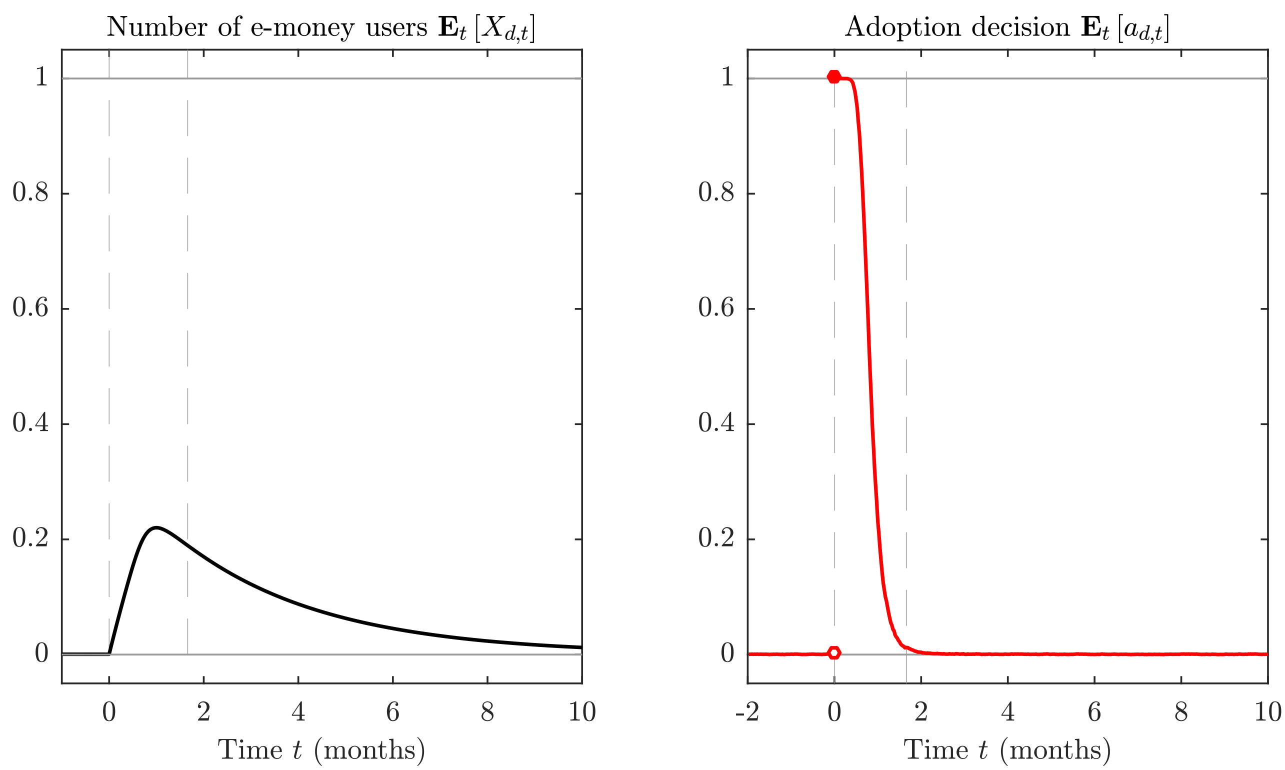


Figure B.2: Average number of users ($\mathbb{E}_t\left[X_{d,t}\right]$, left column) and average adoption decision ($\mathbb{E}_t\left[a_{d,t}\right]$, right column) after the cash crunch in the frictionless model ($C = 0$ and $\kappa = 0$).

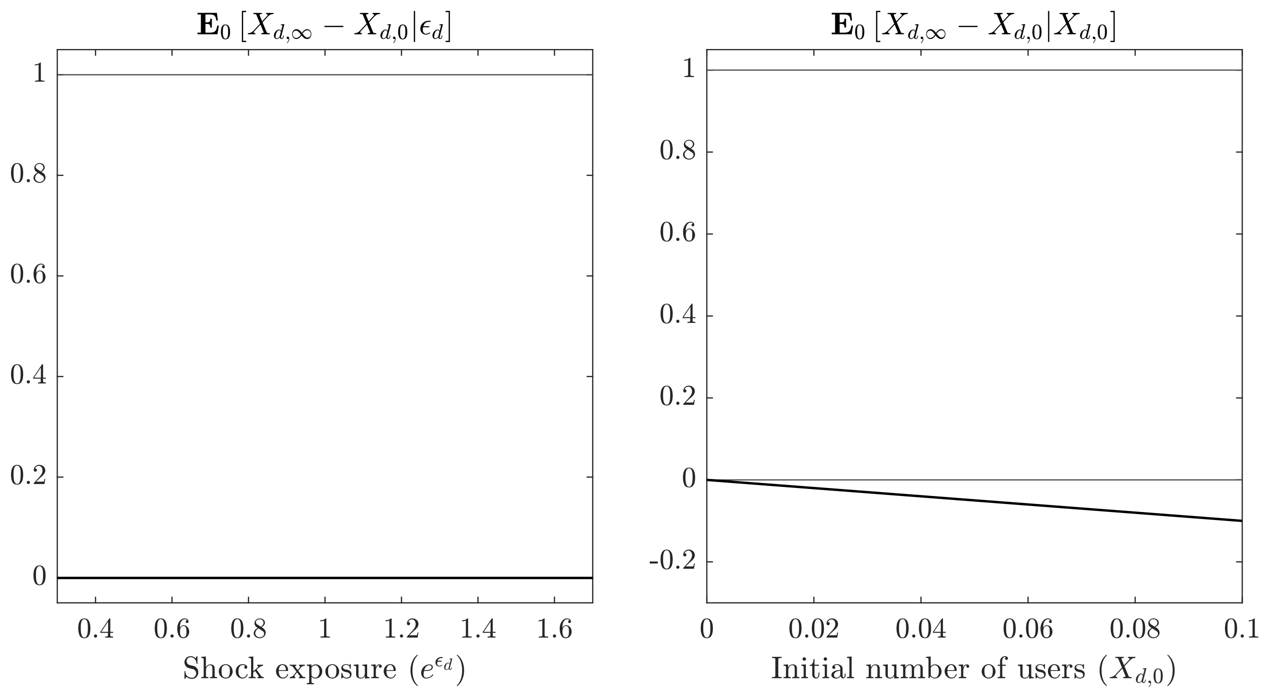
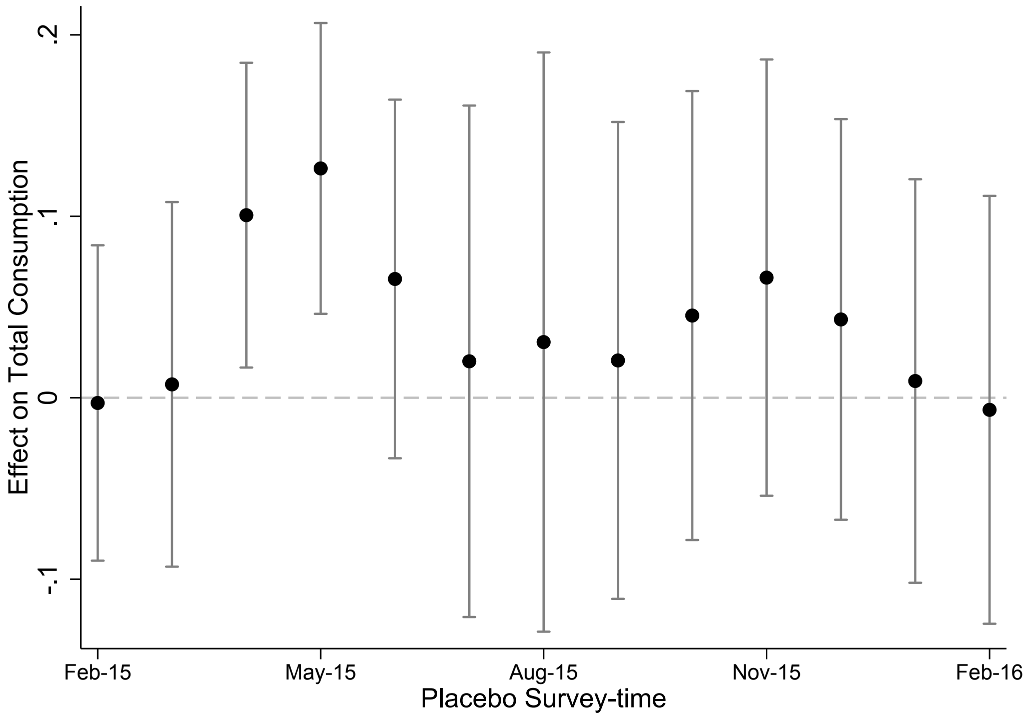


Figure B.3: Conditional impulse responses in the frictionless model ($C = 0$ and $\kappa = 0$). These impulse responses are generated from the same type of simulations as figure 14, that is, the case where $\theta < k$ (the shock is persistent relative to the adjustment speed of firms.) The left column reports the relationship between the district's exposure to the shock, proxied by e^{ϵ_d} (with a value of 1 indicating an average exposure to the shock), and the long-run change in the number of users after the shock. The right column reports the relationship between the initial number of users, $X_{d,0}$, and the long-run change in the number of users after the shock. The long-run number of users in the left panel is defined as $E_0\left[X_{d,\infty} - X_{d,0}|\epsilon_d\right] = \lim_{t\to+\infty} E_0\left[X_{d,t} - X_{d,0}|\epsilon_d\right]$ (and similarly for the right panel). In the left panel, the long-run change in the number of users is represented by the thick black line (which, for this version of the model, is constant and equal to 0). For both columns, in each district, the adoption path is constructed by averaging across 10^3 draws. The limit as $t \to \infty$ is obtained by simulating the response of each district for five years, and using the end-of-simulation values.

Figure B.4: Consumption responses based on placebo shocks



Notes: The figure plots survey month-by-survey month estimates of consumption responses depending on exposure to the shock where we assume the occurrence of a "fake" shock in each survey-time corresponding to each entry on x-axis. The specification we estimate is a version of equation 13 in which each coefficient is based on the interaction of the treatment variable (Exposure_d) with an event-time dummy. We report the coefficient δ for the event-time after shock. Treatment variable is our measure of Exposure_d for the district (as described in Section 2). The dependent variable $\log(y_{h,d,t})$ is the log of total consumption (as described in Section 4). 95% confidence intervals are represented with the vertical lines; standard errors are clustered at the district level. Source: CMIE Consumption Data.

Table 2: Exposure$_d$ and district characterisitics (Balance Test)

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	**mean**	**univariate OLS**		**baseline controls**	
		coeff.	R^2	coeff.	R^2
Log(Pre Deposits)	11.083	-1.290***	0.054		
	(0.048)	(0.273)			
% villages with ATM	0.036	0.090***	0.040		
	(0.004)	(0.023)			
# Bank Branches per 1000's	0.047	0.002	0.000	0.015	0.234
	(0.002)	(0.012)		(0.012)	
# Agri Credit Societies per 1000's	0.045	-0.016	0.001	0.016	0.062
	(0.004)	(0.027)		(0.022)	
% villages with banks	0.085	0.131***	0.033	0.058	0.580
	(0.006)	(0.036)		(0.036)	
Log(Population)	14.376	-0.501**	0.015	0.304	0.481
	(0.035)	(0.208)		(0.199)	
Literacy rate	0.622	-0.029	0.003	-0.001	0.227
	(0.005)	(0.025)		(0.025)	
Sex Ratio	0.946	0.008	0.001	-0.009	0.063
	(0.003)	(0.015)		(0.017)	
Growth Rate	0.208	-0.219	0.014	-0.232	0.021
	(0.016)	(0.139)		(0.171)	
Working Pop./Total Pop.	0.410	0.026	0.005	0.010	0.075
	(0.003)	(0.016)		(0.017)	
Distance to State Capital(kms.)	0.215	0.035	0.002	0.026	0.016
	(0.006)	(0.032)		(0.032)	
Rural Pop./Total Pop.	0.746	0.170***	0.034	0.046	0.464
	(0.008)	(0.047)		(0.039)	

Notes: The table tests for differences in observable district-characteristics and Exposure$_d$. Column 1 reports the mean of district-characteristics. The treatment variables is our measure of Exposure$_d$ as described in Section 2. Columns (2) & (3) report the coefficient of the univariate OLS regression of each variable on the treatment variable. Columns (4) & (5) report the coefficients after controlling for the pre-demonetization bank deposits in the districts (in logs) and share of villages with ATM. Robust standard errors reported in parentheses. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 3: Share of Chest Banks and Deposit Growth

	Δ log(deposits)		Δ log(deposits$^{adj.}$)		Δ log(deposits^{N})	
	(1)	(2)	(3)	(4)	(5)	(6)
Chest Exposure	0.094***	0.083***	0.085***	0.075***	1.821***	1.621***
	[0.013]	[0.012]	[0.013]	[0.012]	[0.257]	[0.238]
log(Pre Deposits)		-0.035***		-0.035***		-0.677***
		[0.003]		[0.003]		[0.063]
% villages with ATM		0.023		0.020		0.445
		[0.040]		[0.042]		[0.769]
% villages with banks		-0.051**		-0.051**		-1.000**
		[0.023]		[0.024]		[0.449]
Rural Pop./Total Pop.		-0.063***		-0.070***		-1.224***
		[0.016]		[0.017]		[0.317]
log(population)		0.036***		0.035***		0.707***
		[0.003]		[0.003]		[0.068]
Observations	512	512	512	512	512	512
R-squared	0.118	0.313	0.099	0.290	0.118	0.313
District Controls		✓		✓		✓

Notes: The table report results from regression of district-level deposit growth (between September 30,2016 and December 31,2016) our measure of Exposure$_d$ for the district (as described in Section 2). Columns (1) and (2) uses the measure of change in total deposits. Column (3) and (4) uses the measure of abnormal growth in total deposits, which adjust for the normal deposit growth by the growth in district-deposit in same quarter for the last two years. Column (5) and (6) uses dependent variable of deposit growth that normalized to have mean zero and standard deviation 1. Odd columns shows the correlation without any controls. Even columns include the district-level controls for (log) pre-shock banking deposits, share of villages with ATM facilities, share of villages with banking facility, share of rural population and level of population in the district. Standard error clustered at district level are reported in parentheses; $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 4: Exposure$_d$ and adoption of digital wallet

	log(amount)	log(# users)	log(# switchers)
	(1)	(2)	(3)
$(\text{Exposure})_d \times \mathbf{1}(t \geq t_0)$	3.134***	1.054**	0.851***
	[0.884]	[0.423]	[0.326]
Observations	6,846	6,846	6,552
R-squared	0.849	0.868	0.830
District f.e.	✓	✓	✓
Month f.e.	✓	✓	✓
District Controls × Month f.e.	✓	✓	✓

Notes: Difference-in-differences estimates of the effect of the shock on the adoption of digital wallet. The estimated specification is equation (7). In Column (1), the dependent variable is the log of the total amount (in Rs.) of transactions carried out using digital wallet in district d during month t; in Column (2), the dependent variable is the log of the total number of active retailers using digital wallet in district d during month t; in Column (3), the dependent variable is the log of the total number of new retailers joining the digital wallet in district d during month t. District controls include (log) pre-shock banking deposits, share of villages with ATM facilities, share of villages with banking facility, share of rural population and level of population in the district. Standard error clustered at district level are reported in parentheses. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 5: District adoption rate of digital wallet based on initial adoption rate

	log(amount)		log(# users)		log(# switchers)	
	(1)	(2)	(3)	(4)	(5)	(6)
$(\text{Distance to hub})_d \times \mathbf{1}\,(t \geq t_0)$	-5.483***	-3.682***	-2.539***	-1.423***	-1.835***	-0.893**
	[0.945]	[1.176]	[0.482]	[0.521]	[0.369]	[0.401]
Observations	6,846	6,846	6,846	6,846	6,846	6,846
R-squared	0.848	0.885	0.864	0.909	0.811	0.868
District f.e.	✓	✓	✓	✓	✓	✓
Month f.e.	✓	✓	✓	✓	✓	✓
District Controls × Month f.e.	✓	✓	✓	✓	✓	✓
State × Month f.e.		✓		✓		✓

Notes: Difference in differences estimate of the effect of initial conditions, using distance to the nearest hub (defined as districts with greater than 500 retailers in September 2016) as a proxy for the initial share of adopters. The specification estimated is equation 12. In Columns (1)-(2), the dependent variable is the log of the total amount (in Rs.) of transactions carried out using digital wallet in district d during month t; in Columns (3)-(4), the dependent variable is the log of the total number of active retailers using digital wallet in district d during month t; in Columns (5)-(6), the dependent variable is the log of the total number of new retailers joining the digital wallet in district d during month t. District-level controls include distance to state capital, employment rate, share of rural population and log of total population. Standard error clustered at district level are reported in parentheses. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

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 9 on simulated 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: 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 10 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 error clustered at pincode level are reported in parentheses. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 8: Consumption responses based on exposure to the shock

Exposure$_d$:	log(Expense$_{\text{Total}}$)	
	Continuous measure	Top 25%
	(1)	(2)
$(\text{Exposure})_d \times \mathbf{1}(t = t_1)$	-0.199***	-0.0577**
	(0.0637)	(0.0234)
$(\text{Exposure})_d \times \mathbf{1}(t = t_2)$	-0.0337	-0.0199
	(0.0815)	(0.0296)
$(\text{Exposure})_d \times \mathbf{1}(t = t_3)$	0.148	0.0146
	(0.102)	(0.0370)
$(\text{Exposure})_d \times \mathbf{1}(t = t_4)$	0.0252	-0.0187
	(0.141)	(0.0588)
Household f.e.	✓	✓
Survey-time f.e.	✓	✓
District Controls × Survey-time f.e.	✓	✓
Household controls × Survey-time f.e.	✓	✓
Observations	564,690	564,690
R-squared	0.707	0.706

Notes: The table shows difference-in-differences estimate for consumption responses for each event-time post the demonetization shock relative the pre-period (four event-time). The specification estimated is equation 13. Treatment variable is our measure of Exposure$_d$ for the district (Column (1)) and takes the values of 1 if the measure of Exposure$_d$ is in the top quartile of the distribution (Column (2)). The dependent variable $\log(y_{h,d,t})$ is either log of total consumption as defined in Section 5. District-level controls include (log) pre-shock banking deposits, share of villages with ATM facilities, share of villages with banking facility, share of rural population and level of population in the district. Household-level controls include pre-shock income and age of head of the household. Standard errors are clustered at the district level. $*** : p < 0.01$, $** : p < 0.05$, $* : p < 0.1$.

Table 9: Consumption responses across categories based on exposure to the shock

	Necessary	Unnecessary	Bills and Rent	Food	Recreation
	(1)	(2)	(3)	(4)	(5)
$(\text{Exposure})_d \times \mathbf{1}(t = t_1)$	-0.174***	-0.211**	0.250	-0.185***	-0.996**
	(0.0573)	(0.0987)	(0.268)	(0.0595)	(0.431)
Household f.e.	✓	✓	✓	✓	✓
Survey-time f.e.	✓	✓	✓	✓	✓
District Controls × Survey-time f.e.	✓	✓	✓	✓	✓
Household controls × Survey-time f.e.	✓	✓	✓	✓	✓
Observations	564,690	564,690	564,690	564,690	564,690
R-squared	0.731	0.622	0.700	0.684	0.460

Notes: The table shows difference-in-differences estimate for consumption responses across various categories for each event-time post the demonetization shock relative the pre-period (four event-time). The specification estimated is equation 13. Treatment variable is our measure of Exposure$_d$ for the district (as described in Section 2). The dependent variable $\log(y_{h,d,t})$ is either log of consumption of necessary goods (Column (1)); log of consumption of unnecessary goods (Column (2)); log of expenditure on bills and rent (Column (3)); log of expenditure on food (Column (4)); log of expenditure on recreation activities (Column (5)) as defined in Section 5. District-level controls include (log) pre-shock banking deposits, share of villages with ATM facilities, share of villages with banking facility, share of rural population and level of population in the district. Household-level controls include pre-shock income and age of head of the household. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table B.1: Exposure$_d$ and Deposit Growth (pre-shock quarters)

	(1) 201604	(2) 201603	(3) 201602	(4) 201601	(5) 201504	(6) 201503	(7) 201502	(8) 201501	(9) 201404	(10) 201403	(11) 201402	(12) 201401
Chest Exposure	1.621***	-0.404	0.476**	0.137	0.163	0.342	-0.040	0.315	0.345	-0.734***	0.165	0.012
	[0.238]	[0.260]	[0.236]	[0.234]	[0.268]	[0.255]	[0.231]	[0.240]	[0.291]	[0.280]	[0.257]	[0.269]
Observations	512	512	512	512	512	512	512	512	512	512	512	512
R-squared	0.313	0.027	0.026	0.162	0.020	0.054	0.044	0.061	0.017	0.037	0.100	0.124
District Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Regression of district-level deposit growth for all eleven 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. Treatment variable is our measure of Exposure$_d$ for the district (as described in Section 2). District-level controls include (log) pre-shock banking deposits, share of villages with ATM facilities, share of villages with banking facility, share of rural population and level of population in the district. Standard error in parentheses; $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table B.2: 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 10 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.3: 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.4: District adoption rates based on initial adoption in electronic payments data: OLS

	log(amount)		log(# users)		log(# switchers)	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}\,(\text{Any Adopter})_d \times \mathbf{1}\,(t \geq t_0)$	1.416***		1.751***		1.312***	
	[0.379]		[0.188]		[0.150]	
$\log(\text{pre-amount})_d \times \mathbf{1}\,(t \geq t_0)$		0.050		0.173***		0.127***
		[0.050]		[0.022]		[0.018]
Observations	6,846	6,846	6,846	6,846	6,552	6,552
R-squared	0.849	0.848	0.880	0.878	0.842	0.839
District f.e.	✓	✓	✓	✓	✓	✓
Month f.e.	✓	✓	✓	✓	✓	✓
District Controls × Month f.e.	✓	✓	✓	✓	✓	✓

Notes: The table shows adoption dependence on initial conditions at the district level. The specification estimated is equation 11. In the first row, I_d is a dummy if a district had a positive adoption level before the demonetization. In the second row, I_d the total amount of transactions before the demonetization. In Columns (1)-(2), the dependent variable is the log of the total amount (in Rs.) of transactions carried out using digital wallet in district d during month t; in Columns (3)-(4), the dependent variable is the log of the total number of active retailers using digital wallet in district d during month t; in Columns (5)-(6), the dependent variable is the log of the total number of new retailers joining the digital wallet in district d during month t. District-level controls include (log) pre-shock banking deposits, share of villages with ATM facilities, share of villages with banking facility, share of rural population and level of population in the district. Standard errors are clustered at the district level. $*** : p < 0.01$, $** : p < 0.05$, $* : p < 0.1$.

Table B.5: District adoption rates based on initial adoption: Alternative specification

	log(amount)			log(# users)			log(# switchers)		
	$\delta = 200$	$\delta = 300$	$\delta = 400$	$\delta = 200$	$\delta = 300$	$\delta = 400$	$\delta = 200$	$\delta = 300$	$\delta = 400$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(Distance To Hub $> \delta$ km.) $\times$ $\mathbf{1}_{\{t \geq t_0\}}$	-1.315***	-1.035***	-1.053***	-0.506***	-0.432***	-0.413***	-0.220	-0.252**	-0.276**
	[0.369]	[0.362]	[0.347]	[0.181]	[0.158]	[0.146]	[0.136]	[0.120]	[0.114]
Observations	6,846	6,846	6,846	6,846	6,846	6,846	6,552	6,552	6,552
R-squared	0.885	0.885	0.885	0.910	0.910	0.910	0.882	0.882	0.883
District f.e.	✓	✓	✓	✓	✓	✓	✓	✓	✓
Month f.e.	✓	✓	✓	✓	✓	✓	✓	✓	✓
District Controls $\times$ Month f.e.	✓	✓	✓	✓	✓	✓	✓	✓	✓
State $\times$ Month f.e.	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The table shows difference-in-differences estimate of the effect of initial conditions, using distance to the nearest hub (defined as districts with greater than 500 retailers in September 2016) as a proxy for the initial share of adopters. The specification estimated is equation 12, 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 or the log of the total nominal value of transactions; log of total number of active firms; log of total number of new firms on the digital wallet. District-level controls include distance to state capital, employment rate, share of rural population and log of total population. Standard errors are clustered at the district level. $*** : p < 0.01$, $** : p < 0.05$, $* : p < 0.1$.

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

	log(Expense)		
	Total	Necessary	Unnecessary
	(1)	(2)	(3)
$\mathbf{1}_{\{t=t_1\}} \times$ (Top 25% Exposure)$_d$	-0.0577**	-0.0427*	-0.0781**
	(0.0234)	(0.0230)	(0.0343)
$\mathbf{1}_{\{t=t_2\}} \times$ (Top 25% Exposure)$_d$	-0.0199	-0.0172	-0.0277
	(0.0296)	(0.0266)	(0.0454)
$\mathbf{1}_{\{t=t_3\}} \times$ (Top 25% Exposure)$_d$	0.0146	-0.00438	0.0519
	(0.0370)	(0.0307)	(0.0533)
$\mathbf{1}_{\{t=t_4\}} \times$ (Top 25% Exposure)$_d$	-0.0187	-0.0588	0.0374
	(0.0588)	(0.0580)	(0.0786)
Household f.e.	✓	✓	✓
Survey-time f.e.	✓	✓	✓
District Controls × Survey-time f.e.	✓	✓	✓
Household controls × Survey-time f.e.	✓	✓	✓
Observations	564,690	564,690	564,690
R-squared	0.706	0.731	0.622

Notes: The table shows difference-in-differences estimate for consumption responses for each event-time post the demonetization shock relative the pre-period (four event-time). The specification estimated is equation 13. Treatment variable takes the value of 1 if our measure of Exposure$_d$ for the district (as described in Section 2) is in the top 25% value of exposure. 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)) as defined in Section 5. District-level controls include (log) pre-shock banking deposits, share of villages with ATM facilities, share of villages with banking facility, share of rural population and level of population in the district. Household-level controls include pre-shock income and age of head of the household. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.