Shocks and Technology Adoption: Evidence from Electronic Payment Systems *

Nicolas Crouzet† Apoorv Gupta‡ Filippo Mezzanotti§

First version: September, 2018
This version: November, 2020

Abstract

Theories of coordination failures in technology adoption have been influential in economics, but empirical evidence on their importance is limited. This paper studies the role of this friction in the adoption of digital payments systems, using data from the largest provider of electronic wallets in India during the 2016 Demonetization. Our empirical strategy exploits variation in the intensity with which Indian districts were exposed to the cash contraction induced by the Demonetization. Consistent with a dynamic technology adoption model with complementarities, we show that the rate of adoption of the technology increased persistently in response to the large but temporary cash contraction. Estimates of the model indicate that the 6-month adoption response would have been 60% lower absent adoption complementarities. This suggests that large but temporary policy interventions can resolve coordination failures in technology adoption, though we highlight an important limitation of this logic: temporary interventions can also exacerbate initial differences in adoption across regions or markets.

Keywords: Complementarity, Externalities, Technology Diffusion, Fintech, Demonetization.
JEL Classification: O33, G23, L86, E65

*We thank Pat Akey, Anthony DeFusco, Sabrina Howell, Ravi Jagannathan, Dean Karlan, Josh Lerner, Konstantin Milbradt, Avri Ravid, George Siddharth, Tim Simcoe, Jon Zinman and participants at the Kellogg Brown Bag, NBER Innovation & Entrepreneurship, Adam Smith Workshop in Finance, UNC Junior Finance Conference, UNC-Duke Entrepreneurship Conference, NY FED Fintech Conference, 2nd Toronto Fintech Conference, SFS Cavalcade, and Fintech & Digital Finance at SKEMA, Berkeley Finance seminar, NYU Finance seminar, CAFRAL India, NYU Indian Conference, HEC Entrepreneurship Conference, AFA 2020, and BYU Red Rock Conference for helpful comments and discussions. We thank in particular the discussants Johan Hombert, Roger Loh, Adrian Matray, Rafael Matta, Abhiroop Mukherjee, Amit Seru, Xu Ting, Guillaume Vuillemey, and Constantine Yannelis for their insightful comments. 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 in the study and the views expressed herein are solely of the authors.

† Northwestern University - Kellogg School of Management; contact: n-crouzet@kellogg.northwestern.edu.
‡ Dartmouth College; contact: apoorv.gupta@dartmouth.edu.
§ Northwestern University - Kellogg School of Management; contact: filippo.mezzanotti@kellogg.northwestern.edu.

1 Introduction

A rich literature in economics has argued that coordination failures could be an important obstacle to the adoption of new technologies (Rosenstein-Rodan, 1943; Carlton and Klamer, 1983). Coordination failures can arise when decisions to adopt a new technology are complements across users — that is, when the private value of adoption for each single user depends positively on adoption by other users (Katz and Shapiro, 1985, 1986).[1] In these situations, expectations of low adoption can become self-fulfilling. While the possibility of coordination failures is theoretically well understood (Murphy et al., 1989; Matsuyama, 1995), direct evidence of their importance is scarce. Using data on the adoption of a digital wallet technology during the 2016 Indian Demonetization, our paper provides novel evidence on coordination failures in technology adoption, and studies the role that policy can play in addressing them.

There are two reasons why documenting the role of coordination failures in adoption of this technology is useful. First, this product provides a clean test case for the more general proposition that coordination failures can slow down technology adoption. Digital wallets are network goods; this makes adoption decisions complements across users, and creates scope for coordination failures (Katz and Shapiro, 1994; Rysman, 2007). More importantly, relative to other network goods, digital wallets are generally cheap and simple to adopt, which helps isolate the role of coordination problems. Second, digital wallets are a canonical example of financial technology ("fintech") products. The rapid diffusion of information technology over the past two decades has raised expectations about the potential for fintech to improve financial inclusion, particularly in developing countries, where fostering access to financial services remains a key goal for policymakers.[2] Understanding the obstacles to their adoption is therefore also relevant to policy.

To examine the role of these coordination failures, this paper studies the adoption of a digital payment technology by merchants after the 2016 Indian Demonetization. This unexpected policy shock resulted in a large but temporary reduction in the availability of cash, leading to a temporary incentive to adopt the technology. Our analysis is organized in three parts. First, we develop a dynamic model of technology adoption with complementarities and use it to characterize the key features of the response of adoption of digital payments to a temporary shock to the availability of traditional means of payment. Second, we use merchant level data from the leading fintech payment system in India and quasi-exogenous variation in exposure to the Demonetization to test the model's predictions. Third, we quantify the contribution of complementarities to the overall adoption response by structurally estimating our model.

[1] Katz and Shapiro (1985, 1986) highlight how externalities can arise both directly — when the number of users affects the quality of the product — or indirectly — in situations where the number of users affect the value of other add-on products due to compatibility (e.g. hardware/software) or the type of post-purchase services (e.g. cars).

[2] For the benefits of fintech in payment systems, see Yermack (2018); Suri and Jack (2016); Jack and Suri (2014); Beck et al. (2018); Agarwal et al. (2019). For the role of fintech in funding markets, such as mortgage and small business lending, see Bartlett et al. (2018); Buchak et al. (2018); Fuster et al. (2018); Howell et al. (2018).

Our main findings are the following. First, the Demonetization caused an adoption wave among merchants, characterized by three features: a persistent increase in the *size* of the platform, that is, the total number of merchants using it; a persistent increase in the platform's *adoption rate*, that is, the number of *new* merchants adopting the platform each month; and *state-dependence* in adoption, meaning that the long-run adoption response depends on the initial (pre-shock) strength of complementarities. The latter two features are important, as we show that they are distinctive predictions of the model when complementarities are present. Second, our quantitative estimation of the model shows that complementarities were not only present, but played a large role: they account for approximately 60% of the long-run adoption response. Taken together, our results suggest that coordination problems — as opposed to pecuniary adoption costs or transaction costs — could be an important obstacle to the diffusion of fintech.

In this context, a central lesson from both the model and our data is that large-scale but temporary interventions can, in principle, be used to overcome coordination problems. However, on this point, our analysis offers an important caveat: interventions that are too short-lived can also exacerbate long-run differences in adoption across markets or regions. In fact, the state-dependence created by adoption externalities is key to this insight: markets with high initial adoption rates will experience persistent long-run adoption effects, while in other markets, the effects will not persist. We find evidence for this mechanism in the data, and explore its policy implications using counterfactual experiments in our estimated model. These experiments suggest that the length of a policy intervention may have a first order economic effect on both the level and the dispersion of the adoption response.

The empirical setting of the paper is the Indian Demonetization of 2016. On November 8th, 2016, the Indian 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 the total cash in circulation. The public was not given advanced warning, and the bills were voided effective immediately. A two-month 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, combined 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 harder to conclude, but more funds were available for use in electronic payments. Importantly, though the shock was very large, it was also temporary, as things normalized for the most part by February 2017.

In Section 2, we start by showing that the Demonetization led to a large aggregate 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 customers to pay at stores using funds deposited in their bank accounts. Payment is

then transferred to merchants' bank accounts via the app. The economic costs associated with the adoption of this technology for merchants are small; in fact, 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 were very common in India by 2016 (Agarwal et al., 2017). Aggregate activity on the platform doubled several times over during the two months following the Demonetization announcement. Additionally, this adoption response was persistent, though the shock was not. There was no significant mean-reversion in the aggregate number of merchants using the technology or in aggregate transaction volumes once cash withdrawal constraints were lifted.

The aggregate evidence thus suggests that the temporary contraction in cash led to a persistent increase in adoption of fintech payments. However, this finding alone is not sufficient to establish that complementarities played a role. First, theoretically, economic mechanisms other than complementarities may also generate persistent responses to transitory shocks. Second, empirically, long-term responses in aggregate event studies are potentially confounded by subsequent aggregate shocks.

In order to address the first issue, in Section 3, we characterize further the testable implications of adoption complementarities by studying a dynamic technology adoption model. The model builds on the framework of Burdzy et al. (2001). Firms face a choice between two technologies, one of which (the "platform") is subject to positive adoption externalities — the flow profits from operating under this technology increases with its rate of use by other firms.[3] Moreover, the relative benefit of adopting the platform is subject to aggregate shocks. The presence of these common shocks helps eliminate the potential equilibrium multiplicity arising from complementarities in adoption decisions.

The model predicts that following a large, temporary shock, the total number of firms using the platform increases persistently, consistent with the aggregate evidence of Section 2. However, it delivers two additional predictions. First, the presence of complementarities — on top of increasing the size of the platform — also generates a persistent increase in the adoption rate. In other words, the number of *new* firms joining the platform every period remains high even after the shock has dissipated. The reason is that, with complementarities, the initial adoption triggered by the shock, by temporarily raising the size of the platform, increases the relative future value of adoption for other firms. This "snowball" effect can generate endogenous persistence in the increase in adoption rates. Second, the model predicts that adoption responses exhibit state-dependence: the long-run adoption response depends on the pre-shock adoption rates. The intuition for this result is simple: initially stronger complementarities, in the form of a higher initial adoption rate, make it easier to reach the tipping point beyond which the platform has sufficient critical mass to continue growing even if the initial shock dissipates.

[3] Different mechanisms could account for this relationship: for instance, the more merchants are on the platform, the more valuable it is for consumers to use it, which in turn increases new merchants' incentive to join the platform. We discuss microfoundations in 2, and provide an example with a two-sided market in Appendix B.2.

In Section 4, we then show that the empirical predictions of the model with complementarities highlighted above are consistent with the adoption responses observed in the data after the Demonetization. In order to do this, we provide an empirical design to estimate the causal impact of the cash contraction on adoption. Our empirical design exploits variation across districts in the importance of chest banks — local bank branches in charge of the distribution of new currency — to identify variation in exposure to the shock. This design allows us to isolate the effect of the cash contraction from other effects of the Demonetization, therefore overcoming the limitations of the aggregate evidence. We show that the districts that were more exposed to the cash crunch also experienced a larger and more persistent increase in total adoption following the Demonetization, the first prediction of the model. Crucially, higher exposure also predicts a larger increase in the number of new firms joining the platform, even after restrictions on cash withdrawals are lifted — the second prediction of the model.[4] Finally, we find robust evidence consistent with state-dependence, the third prediction of the model with complementarities. Districts where the pre-shock value of joining the platform is higher — either because pre-shock adoption was high, or because they were located close to other high adoption cities — are characterized by a higher average adoption response. The same pattern holds at a disaggregated level: a firm's choice to use the technology is positively affected by the rate of adoption of firms in the same local industry. This latter result does not simply capture variation across locations and industries, since it holds conditional on these fixed effects interacted with time.

Altogether, this reduced-form evidence shows that a model with adoption complementarities can account for the *qualitative* features of the adoption response caused by the Demonetization. However, it is silent about the quantitative contribution of complementarities to the adoption response. In order to address this issue, in Section 5, we estimate the dynamic adoption model of Section 3 via simulated method of moments, using our data on fintech payments. The key parameter of interest is the size of adoption complementarities. Following the intuition described above, we show that this parameter can be identified using the difference between short- and long-run adoption rates following the shock.

Using the estimates of the model, we provide two main results. First, we show that complementarities are quantitatively important in understanding the total adoption response: they account for approximately 60% of the total response of adoption to the Demonetization, in the sense that the medium-run adoption rate would have been 60% lower (and declining), had the technology featured no complementarities in adoption. Second, we show that the persistence of the shock is crucial to understanding its effects, both in terms of average adoption, and for the variance of adoption across regions. As discussed earlier, temporary interventions may increase overall adoption. However, because of state-dependence, they can also exacerbate

[4] We provide a number of robustness tests that confirm the causal interpretation of these results. Among other things, we use our empirical design to show that consumption also temporarily declined following the shock. This evidence helps to reinforce the notion that our results capture the effects of a temporary shock to cash, rather than the effects of a demand shock.

initial differences in adoption. Consistent with this intuition, we show that, keeping the present value of the decline in cash constant, a cash crunch with a smaller initial magnitude (by around 40%) but a longer half-life (by a factor of 2), would have led to higher long-run adoption rates (by about 10%) and lower dispersion. Thus, an implication of our model is that policymakers with a preference for uniform adoption across regions or industries should generally favor smaller but more persistent interventions.

Contribution to the literature We contribute to three existing areas. First, our paper relates to the growing literature on fintech (Bartlett et al., 2018; Buchak et al., 2018; Fuster et al., 2018; Howell et al., 2018; Alvarez and Argente, 2019). Our contribution is to establish that externalities can be a quantitatively important obstacle to adoption of digital payments systems, beyond traditional pecuniary costs, such as setup or transaction fees, which are virtually absent for the technology we study. This finding is important because of the benefits of electronic payment systems documented in the literature (Yermack, 2018; Jack and Suri, 2014; Beck et al., 2018).[5] In general, the idea that large, temporary events may be instrumental in generating a large shift in technology adoption is not new in the policy discussion about electronic payments.[6] For instance, a well-known example of the rapid adoption of an electronic payment technology is M-PESA in Kenya, which grew by 70% between 2007 and 2011 and was used by 97% households in 2014 (Suri and Jack, 2016). In this context, the large-scale political unrest in early 2008 — which lasted for two months and led to shutdown of the traditional financial services — is commonly believed to have played a substantial role in driving the initial wave in adoption (Mas and Morawczynski, 2009). Similarly, in the Lake Kivu region in Rwanda, the disruption following the February 2008 earthquake is considered to have contributed to a significant increase in the use of the credit held on mobile phones (Blumenstock et al., 2016). Despite the frequency of such events, there is little work systematically documenting their implications for adoption dynamics. In this paper, we go beyond qualitative evidence and conduct a quantitative analysis to show that temporary shocks can indeed induce large and persistent increases in electronic payments use.

Second, our work relates to the literature studying the diffusion of technologies across firms.[7] We focus on the role of coordination problems for technologies featuring adoption complementarities (Arthur, 1989; Katz and Shapiro, 1985; Farrell and Saloner, 1986; Sakovics and Steiner, 2012). Relative to existing work on the

[5]Relatedly, we show that traditional payment technologies — such as credit or debit cards — where not widely adopted by new users during the Demonetization, though they were more actively used by existing users. Our paper thus also relates to the literature on debit cards and household behavior (Bachas et al., 2017; Schaner, 2017).

[6]Related work by Higgins (2019) explores how a permanent increase in the availability of debit cards in Mexico affected both consumers and retailers. Relative to Higgins (2019), our work directly quantifies the importance of complementarities in explaining the adoption response. Additionally, we argue that while that permanent interventions may not be necessary to spur adoption, interventions that are too temporary may also exacerbate dipersion in adoption rates in the long-run.

[7]There is an extensive literature on slow adoption of new technologies (Hall and Khan, 2003; Rosenberg, 1972), which 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 constraints (Atkin et al., 2017) to slow social learning and information frictions (Munshi, 2004; Young, 2009; Conley and Udry, 2010; Gupta et al., 2020) to lack of financial development (Comin and Nanda, 2019; Bircan and De Haas, 2019). For a review of this literature, see Foster and Rosenzweig (2010).

topic, we provide direct evidence on key predictions of models with *dynamic* coordination problems: first, temporary shocks lead to persistent responses; second, temporary shocks can have heterogeneous long-run effects depending on the initial strength of externalities.[8] Electronic payment systems are a natural example of technology exhibiting adoption complementarities (Katz and Shapiro, 1994; Rysman, 2007; Gowrisankaran and Stavins, 2004), and for which coordination problems may be an important obstacle to adoption (Crowe et al., 2010). While we interpret the adoption complementarities as arising from the network nature of the technology we study, alternative mechanisms —— for instance, learning about the costs and benefits of the technology (Munshi, 2004; Suri, 2011) —— may also contribute to adoption complementarities. In section 4, we provide evidence that is inconsistent with learning being the primary mechanism. While this evidence reinforces the notion that network effects are important in our context, the key implications of our analysis do not depend on the specific mechanism that generates complementarities in adoption.

Third, we contribute to the literature on the impact of the Demonetization on the Indian economy.[9] We focus on the adoption of electronic payment systems, and argue that the adoption wave following Demonetization may be difficult to rationalize in a model where complementarities do not play an important role. Despite the potentially large benefits of adopting electronic payments, it is also important to point out that any positive impact should be weighted against the large costs that the Demonetization imposed on the real economy, as documented by Chodorow-Reich et al. (2019), and consistent with our findings in Appendix C.[10] Our analysis uses the Demonetization as a laboratory to identify the frictions that are relevant to the adoption of electronic payments; it does not aim to provide a welfare evaluation of this event.

The rest of the paper is organized as follows. Section 2 provides some background on the Demonetization and documents aggregate adoption effects. Section 3 analyzes our dynamic adoption model and derives key predictions. Section 4 tests these predictions in the electronic wallet data. Section 5 estimates the model and provides counterfactuals. Section 6 concludes.

[8] Other papers providing related evidence are Björkegren (2015), who quantifies the impact of externalities in determining the cost of a tax on network goods; Saloner et al. (1995) who examine the importance of the potential size of a network in the decision to adopt a new technology (ATM); Tucker (2008), who studies how different types of adopters may influence the expansion of the network; and Ryan and Tucker (2012), who study the adoption of video calling by firms. A related empirical analysis of dynamic coordination problems is Foley-Fisher et al. (2020), who study self-fulfilling runs in the US life insurance market. Their analysis uses a different framework, and largely abstracts from the persistence of responses to temporary shocks. Related work by Buera et al. (2020) studies how coordination failures in technology adoption can amplify the effects of other negative distortions in a general equilibrium setting, but also how changes in distortions can spur adoption.

[9] This literature includes, among others, Banerjee et al. (2018), Agarwal et al. (2019), Dharmapala and Khanna (2018), and Lahiri (2020).

[10] Chodorow-Reich et al. (2019) quantifies the economic effects of the Demonetization using a cash in advance model that abstracts from externalities and assumes that cash and electronic payments have a fixed elasticity of substitution that is smaller than unity. Our paper complements this analysis by studying the mechanism explaining the adoption increase. In particular, in our framework the elasticity of substitution between means of payment is endogenous and reflects changes in the strength of externalities. Additionally, we provide a different research design for identifying quasi-exogenous exposure to Demonetization, which can be easily replicated using publicly available data.

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 unexpected 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 they 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 30, 2016. Two new bank notes, of Rs.500 and Rs.2,000, 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 the new and old currency was not immediate, and the public was unable to withdraw cash at the same rate as they were 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, it declined by almost 50% during November and continued declining in December.

This cash crunch partly reflected limits on cash withdrawals put in place by the RBI in order to manage the transition.[13] But it was also driven by the difficult logistics of the swap itself. In order to ensure that the policy remained undisclosed prior to its implementation, the RBI had not printed and circulated large amounts of new notes beforehand. This caused many banks to be unable to meet public demand for cash, even under the withdrawal limits.[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 noncash payment, such as cards, checks, or any other electronic payment method; cash transactions were the only ones to be specifically impaired.

[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.

[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] See Appendix Section A.1 for a more detailed discussion of the effects of the shock on cash availability. Banerjee et al. (2018) argue that the uncertainty surrounding the withdrawal limits may have exacerbated the cost of the intervention during this transition period.

[14] See also Krishnan and Siegel (2017). Section A.1 in Appendix contains a more extensive discussion of the effects of the Demonetization on cash in circulation.

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 finally able to withdraw cash from banks (see Figure 2). Furthermore, by January 30th, 2017, the Government lifted most of the remaining limitations on cash withdrawals, in particular removing any ATM withdrawal limit from current accounts.[15]

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 the notes had been deposited. Third, it was relatively short-lived: the cash shortage was particularly acute in November and December, it improved with the new year, and generally normalized with February.

2.2 Fintech payment systems during the Demonetization

Overall, the Demonetization was associated with a large uptake in electronic payments. We start by illustrating this fact using data from the leading digital-wallet company in the country. The company allows individuals and businesses to undertake transactions with each other using only their mobile phone. 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.[16]

Importantly, adoption and usage costs of this technology are low, in particular relative to traditional electronic payments. No investment in a point of sale (POS) terminal is required; the retailer simply needs a cellphone and a bank account, which are both very common in India (Agarwal et al., 2017). Furthermore, for small and medium-sized merchants — who make up the bulk of our data —, transactions using the technology do not involve fees.[17]

Figure 3 reports data for the total number and total value of transactions executed by merchants using this technology. 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, after the Demonetization, the shift towards this payment method was dramatic. In particular, in the first week after the shock, the number of transactions grew by more than 150%, and the value of transactions increased by almost 200%. For the first

[15] Consistent with the view that cash availability was most impaired during the November-January period, we collect data on Google search of key words that are related to the Demonetization events, and find that the perception of the negative consequences of the Demonetization on means of payment significantly improved with the new year. These results are reported in Figure E.1, and discussed more in detail in Appendix A.1.

[16] In Appendix A.3, we provide a more detailed description of the technology. Notice that a smart-phone with access to Internet is not necessary: in 2016, the company introduced a new service that allows customers to make payments using text messaging only.

[17] Finally, the time for retailer to be verified and approved are typically short. See the discussion of the features of the wallet in Appendix A.3.

month after the shock, weekly growth rates were consistently around 100%.

Crucially, this initial positive effect of the Demonetization on adoption did not dissipate, even after the cash-availability constraints were relaxed. The data show a slow-down in aggregate growth starting in January, which is when the limits on the circulation of new cash started to be lifted. However, after a small negative adjustment in early February, the average growth rate over the next two months remained positive, indicating that users did not abandon the platform as cash became widely available again.[18] In other words, a temporary decline in the availability of cash led to a permanent increase in the usage of the platform.

The data shared with us by the electronic wallet company end in June 2017. However, it is important to point out that the increase in electronic-wallet technologies in India also persisted after this period, and this is true despite the introduction of a new "Goods and Services Tax" (GST) in July 2017, which was generally thought to have increased the incentive to use cash for transactions to avoid this new tax. According to the official estimates by the Reserve Bank of India, monthly mobile-wallet transactions increased from 75 million to over 300 million between September 2016 and March 2017, which is the central period in our analysis. In March 2019, monthly transactions still totaled around 385 million, suggesting that adoption has not dissipated since the Demonetization. While our mechanisms may explain this persistent increase in adoption, it is also important to recognize that over such a long period of time total adoption will likely be affected by a variety of other factors that are outside the scope of our paper.

2.3 Traditional electronic payment systems during the Demonetization

Aside from this fintech platform, more traditional electronic payment technologies were also available to the public. We collected publicly available data on monthly debit and credit card activity aggregated at the national level by the RBI. Figure 4 presents these data. The first panel reports the growth in the number of transactions for both credit and debit cards, across ATMs and points of sale (stores). The second panel reports the growth in the number of cards, again divided between 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 growth rate of transactions at point of sales increases dramatically in both November and December, before returning to levels similar to the pre-shock period. This suggests that the Demonetization also led to a permanent increase in debit card transactions. Second, the short-run increase is completely driven by the intensive margin, unlike with the electronic wallet. In other words, the overall volume of debit card transactions increases only because debit-card holders start to use them more frequently. In fact, the second panel of Figure 4 shows no clear growth rate in the number

[18] We believe that this decline may be related to the announcement of a small fee in February (which was later canceled), an increase in competition, and entrance by other electronic payment companies.

	No frictions ($C = 0$, $\kappa = 0$)	Fixed costs ($C = 0$, $\kappa > 0$)	Complementarities ($C > 0$, $\kappa = 0$)
P1: Persistent increase in size of user base	✗	✓	✓
P2: Persistent increase in adoption rate	✗	✗	✓
P3: Positive dependence on initial adoption	✗	✗	✓

Table 1: Predictions across versions of the model.

of new cards during either November and December.

These findings speak to the differences between traditional and fintech electronic payments. Relative to the electronic wallet technology, for retailers, cards involve both larger fixed adoption costs (point of sales terminals) and flow use costs (transaction fees). The former, in particular, could explain why the extensive margin response was more limited for traditional must wait electronic payment methods.[19]

3 Theory

In this section, we analyze a dynamic model of technology diffusion. This model makes a number of specific empirical predictions, which depend on whether the technology features adoption complementarities, and are summarized in Table 1.[20] In Section 4, we test these predictions using the adoption of the electronic wallet technology after the Demonetization. In section 5, we use the model to quantify the contribution of complementarities to the adoption response and to produce counterfactual responses to the Demonetization shock.

3.1 Model

Economic environment Time is discrete: $t = 0, \Delta t, 2\Delta t, \ldots$. There is a collection of infinitely-lived firms, indexed by $i \in [0, 1]$, that discount the future at rate e^{-r}. We refer to these firms as a "district", by analogy with our empirical analysis in the following section. 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 "e-money". Flow profits are:

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

[19] In this respect, our setting differs substantially from Higgins (2019), who studies a technology — debit cards — which requires merchants to pay a large set-up cost as well as regular fees.

[20] The model is a variant of Frankel and Pauzner (2000) and Burdzy et al. (2001), with fixed costs of adjusting technology. It 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 of aggregate (public) signals in environment where agents' actions are complements (Morris and Shin, 2002). The two key differences of this framework with global games models is that (a) firms have no private information on the returns to adoption; (b) firms solve a dynamic coordination problem, instead of a one-shot static model. The latter difference is important, as it allows us to distinguish between short- and long-run effects of the shock.

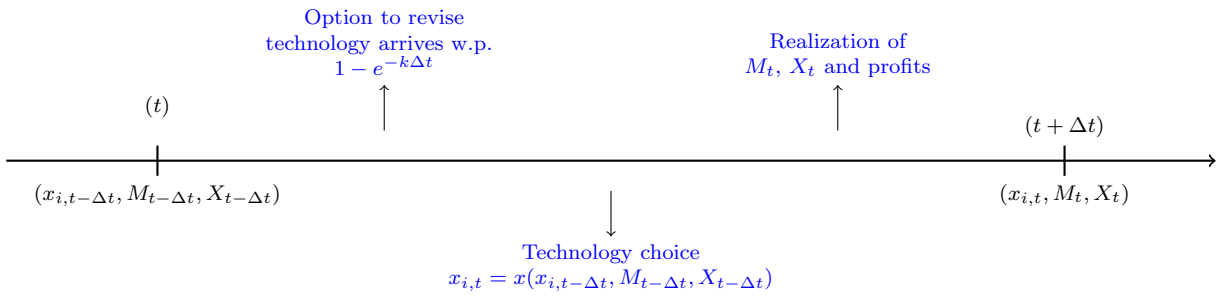


Figure 1: Timing of actions and events during a period.

where $M^e > 0$ and $C \geq 0$ are parameters, and $X_t \equiv \int_{i\in[0,1]} \mathbf{1}\{x_{i,t} = e\}\, di$. Since $C \geq 0$, flow profits to technology e are increasing in the number of other firms using e. The magnitude of C controls the strength of this effect. Below, we discuss what could explain these positive external returns in the case of the wallet technology, and provide a simple microfoundation in a two-sided market with consumers.

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

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

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

There are two frictions that might prevent switching between technologies. First, during each increment of time Δt, $1 - e^{-k\Delta t} \in [0,1]$ firms receive a "technology adjustment" shock and are able to change their technology adoption. This shock is purely idiosyncratic, and it arrives independently of 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 technology e. Specifically, a firm must pay a fixed cost κ if it decides to revise its technology from c to e. There is no cost of switching from e to c and no cost of staying with the same technology.

The timing of actions within period t is depicted in Figure 1. Note that firms make their technology choice at the beginning of period t, before either cash M_t or the current fraction of adopters X_t are determined. Their choice is thus conditioned on $\{x_{i,t-\Delta t}, M_{t-\Delta t}, X_{t-\Delta t}\}$.

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

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

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

$$B(M_{t-\Delta t}, X_{t-\Delta t}) = \mathbb{E}_{t-\Delta t}\left[\left(\Pi_t^e - \Pi_t^c\right)\Delta t + e^{-(r+k)\Delta t}B(M_t, X_t) + e^{-r\Delta t}(1-e^{-k\Delta t})g(B(M_t, X_t))\right] \quad (2)$$

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 flow adoption benefits $\Pi_t^e - \Pi_t^c$. With fixed costs, $g(B) \geq 0$; in that case, $g(B)$ captures the option value of technology e. The resulting technology adoption rule for adjusting firms is:

$$x(x_{i,t-\Delta t}, B_{t-\Delta t}) = \begin{cases} c & \text{if } B_{t-\Delta t} \leq 0 \\ x_{i,t-\Delta t} & \text{if } B_{t-\Delta t} \in [0, \kappa] \\ e & \text{if } B_{t-\Delta t} > \kappa \end{cases}, \quad (3)$$

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

$$\Delta X_t = \left(1 - e^{-k\Delta t}\right)(1 - X_{t-\Delta t})a_{c\to e,t} - \left(1 - e^{-k\Delta t}\right)X_{t-\Delta t}a_{e\to e,t}. \quad (4)$$

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 (2)-(3) when X_t follows the law of motion given by (4), and cash follows the law of motion given by (1).

Discussion of key assumptions We make two main assumptions in this model. First, the technology e features positive external returns with respect to adoption by other firms in the district, that is, $C \geq 0$. 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. Appendix B.2 provides a more precisely microfounded version of the model, with a two-sided market, and which, additionally, the e-money technology is an add-on (so that firms that have adopted it still accept cash). The microfounded version is isomorphic to the baseline model described here, as discussed in the appendix. Alternatively, external returns could arise from spillovers across firms in learning how to use the technology. In either case, the main testable

predictions of the model highlighted below would be qualitatively identical. We discuss this issue in more detail in Section 4.

The second 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. This assumption captures the possibility that firms have heterogeneous (unobservable) abilities to adjust to market conditions as they change, because of behavioral or informational frictions that we leave unmodelled.[21] This friction affects technology choices symmetrically, not just the decision to adopt e, by contrast with the cost κ. It makes technology adjustment sluggish and allows for persistent deviations from the optimal technology choice even if fixed pecuniary costs of adoption are small, which we have argued is likely the case for the technology we study.

3.2 The effects of a cash crunch

We now discuss the effects of a sudden, unanticipated reduction in M_t of size S at date 0:

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

We start by discussing its effects in a version of the model where there are no fixed adoption costs, and complementarities are the only potential barrier to adoption ($C > 0$ and $\kappa = 0$). We highlight three key testable implications: a persistent effect of the shock on the *size* of the user base; a persistent effect on the *adoption rate*, that is, the number of new users joining the platform each period; and a dependence of long-run responses on initial adoption rates. We then contrast these predictions to other versions of the model in which complementarities are absent.

3.2.1 Predictions in the model with complementarities

With complementarities, technology choices depend on firms' expectations about how the number of users of e will evolve. In principle, this could lead to equilibrium multiplicity, with self-fulfilling expectations. However, with common shocks ($\sigma > 0$), Frankel and Pauzner (2000) show that there is a unique equilibrium, characterized by a frontier $\mathbf{\Phi}(.)$ such that firms adopt e, if and only if $M_{t-\Delta t} \leq \mathbf{\Phi}(X_{t-\Delta t})$.[22]

A key feature of the equilibrium is the fact that when $C > 0$, the adoption rule is *increasing* in $X_{t-\Delta t}$. The slope is positive because adoption benefits depend positively on the current value of the number of users

[21] From a theoretical standpoint, sluggishness helps neutralize the potential for complementarities to generate multiple equilibria, as emphasized by Frankel and Pauzner (2000).

[22] Frankel and Burdzy (2005) generalize the results of Frankel and Pauzner (2000) to the case of mean-reverting shocks, and provide technical restrictions on the stochastic shock process so that unicity is guaranteed. Our constant rate of mean reversion falls under the restrictions formulated in assumption A2 of their paper. The working paper version of Guimarães and Machado (2018) also discusses this issue. We thank Bernardo Guimarães for clarifying this point.

of e, $X_{t-\Delta t}$. 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 $X_{t-\Delta t}$ 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 $M_{t-\Delta t}$ must be higher in order to dissuade firms from moving to e. The dynamics implied by this adoption rule are illustrated in Figure 5. This panel plots the adoption threshold $\mathbf{\Phi}(.)$ as well as two different trajectories after the shock, 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. We next discuss qualitatively three key features of these dynamics.[23]

P1: persistent increase in the size of the user base The negative shock to M_t can have persistent effects on the total number of firms using the technology. If the initial number of adopters e is sufficiently high, it can be the case that, after the shock, X_t does not converge back to initial level, but instead, converges to 1. This is illustrated in the blue trajectory in Figure 5. 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.

P2: persistent increase in adoption rate Importantly, on the blue trajectory, firms with the option of changing technology always opt for e, even long after the shock has dissipated. Thus, with complementarities, the shock should lead not only to a persistent increase in size of the platform, but also in its adoption rate, that is, the flow of new users into the platform each period.

P3: positive dependence on initial adoption Finally, these adoption effects are stronger and more persistent, the higher the initial level of adoption. For instance, if 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. Thus, the model features *positive* state-dependence with respect to initial adoption rates.[24]

[23] Appendix B.4 establishes these properties more formally, using Monte-Carlo simulations.

[24] This latter feature also differentiates the model from a standard S-curve adoption models, in which there is no state-dependence in adoption. Appendix B.4.1 shows that state-dependence can be thought of as generating a "conditional" S-curve: the long-run adoption response to the shock is an increasing, S-shaped function of the initial level of adoption, as reported in Appendix Figure E.10. The analogy to standard S-curve is only superficial, though, as Figure E.10 traces out the long-run response as a function of initial adoption rates, while traditional S-curves refer to the dynamics of the total adoption level over time.

3.2.2 Other versions of the model

To what extent do the empirical predictions we highlighted characterize complementarities? Appendix B.4 discusses alternative versions of the model in more detail, and Table 1 summarizes the findings.

The frictionless model In the frictionless case ($C = 0$ and $\kappa = 0$), it is straightfoward to see that while the cash crunch causes a short-run spike in adoption, firms revert back to cash as the shock dies out. Thus, the frictionless model does not generate a persistent increase in adoption in response to a temporary shock.

Fixed costs In the model with fixed costs ($C = 0$ and $\kappa > 0$), firms' technology choice follows a simple (S, s) rule. Two boundaries, $\underline{M}$ and $\overline{M}$, fully characterize the policy: a firm chooses to switch to e if $M_t < \underline{M}$, to switch to c if $M_t > \overline{M}$, and inaction when $\underline{M} < M_t < \overline{M}$. As illustrated in Appendix Figure E.5, 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). There, 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 persistent effects on the user base, just as in the model with complementarities; but the adoption rate of the platform does not increase persistently, and instead it goes to zero as the shock dissipates. In other words, in a model characterized only by fixed costs, the increase in adoption generated by the shock is entirely due to firms switching at the height of the cash crunch, with no impact on new adopters outside this initial period. Additionally, the model with fixed costs does not feature state-dependence with respect to baseline adoption rates, since the initial number of users is irrelevant to adoption choices of other firms.[25]

Shock persistence The externalities due to complementarities in adoption may seem to give policymakers unusually strong powers in triggering technology adoption: temporary interventions can indeed have permanent effects. However, an important point to note is that an increase in average adoption can also be accompanied by more heterogeneity in adoption rates across districts, depending on initial differences in baseline adoption rates, consistent with the state-dependence of responses discussed above.

The strength of this effect crucially depends on the degree of persistence of the shock. At the extreme, very temporary interventions ($\theta \to \infty$) will do nothing more than accentuate differences in initial technology adoption. On the other hand, a more persistent shock is likely to lead to permanent adoption even in districts

[25] The total number of users in fact features slightly negative state dependence. In Appendix Figure E.5, 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. However, because the law of motion for X_t, from B to C, is simply $\Delta X_t = (1 - e^{-k\Delta t})(1 - X_{t-\Delta t})$, the cumulative change in X_t is a decreasing function of the initial number of users, X_0. Appendix B.6 uses numerical simulations of the model to further contrast the response of the long-run adoption rate and its state-dependence, with respect to the model with $C > 0$ and $\kappa = 0$.

with low baseline adoption rates. More specifically, state-dependence in the adoption response can occur whenever $\theta > k$ 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 model to generates a *stronger* permanent switch to e after the shock, but a *weaker* relationship between initial conditions and subsequent increases in the number of users. This is discussed in Appendix B.7. The strong state-dependence we document in the analysis of the next section indicates that $\theta > k$ is the empirically relevant case for the Demonetization (something our structural estimates of the model also confirm).[26]

Policymakers may therefore face a trade-off between the persistence of the shock and its distributional effects. The following section argues that there is strong state-dependence in the data, so that this trade-off is relevant empirically; Section 5 uses the model to study this trade-off quantitatively.

4 Adoption dynamics

This section uses micro data from the leading electronic wallet in India to test the three empirical predictions of the model with externalities highlighted in Section 3. We test the first two predictions, on the long-run increase in both the *size* of the platform and in its *adoption rate*, by using quasi-random variation in the exposure to the shock. Additionally, we provide both district- and firm-level evidence consistent with the third prediction, the positive dependence of adoption responses on baseline adoption rates.

4.1 Data

The main data we use in our analysis are merchant-level transactions from the leading digital-wallet companies in the country.[27] We observe weekly level data on the sales amount and number of transactions happening on the platform for anonymized merchants between May 2016 and June 2017.[28] For each merchant, we also observe the location of the shop at the district 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.[29] There are two key features of these data. First, the information is relatively high frequency, since we can aggregate the data at weekly or monthly levels. 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.

[26] For state-dependence to also generate dispersion in long-run adoption rates, initial conditions must also be sufficiently heterogeneous across districts. As we discuss in the next section, this is the case in the data we study.

[27] During the period of our study, this mobile platform was the largest provider of mobile transaction services. However, since March 2017, a few other public platforms have emerged, in part as a result of the government's "cashless economy" initiative.

[28] The company shared with us information on the top million firms using the platform during this period. This sample represents the quasi-totality of transactions — in both number and value — conducted using the off line technology. See Appendix Section A.3 for more details on the technology.

[29] 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.

We obtain data on district-level banking information from the RBI. This includes three pieces of information at the district level: first, the number of bank branches; second, information on the number of currency chests by district and the banks operating the chests; third, quarterly bank deposits at the bank-group level in each district. Finally, we complement this data with information from the Indian Population Census of 2011 to obtain a number of district-level characteristics, including: population, quality of banking services (share of villages with an ATM and banking facility, number of bank branches and agricultural societies per capita), socioeconomic 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 The effects of Demonetization on adoption

Next, we test the first two predictions of the model: the long-run increase in both the *size* of the platform and its *adoption rate*. The aggregate event study evidence discussed in Section 2 is qualitatively consistent with these predictions. At the same time, this aggregate event study evidence may not properly capture the long-run causal response of adoption to the shock. One particularly important confounding factor are national government policies that may have affected the subsequent adoption of electronic payments for reasons unrelated to externalities, as we describe in Appendix Section A.2. We overcome this concern by using quasi-random variation across different districts in exposure to the cash contraction. This approach allows us to recover the causal effect of the temporary cash contraction on adoption of electronic payments independently of any other aggregate shocks after the Demonetization.

Exposure measure To identify heterogeneity in the exposure to the cash contraction, we exploit the heterogeneity across districts in the relative importance of chest banks — defined as banks operating a currency chest in the district — 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 the chest banks, as well as in chest banks' share of the local deposit market.

Consistent with anecdotal evidence, we expect that districts where 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.[30] On some level, this relationship is mechanical. Chest banks were the first institutions to receive new notes, so in districts where chests account for a larger share of the local banking market, a larger

[30] 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.

share of the population can access the new bills faster. Furthermore, the importance of chest banks may be an even more salient determinant of access to cash if these institutions were biased toward their own customers or partners. Indeed, concerns of bias in chest-bank behavior were widespread in India during the Demonetization.[31] In any case, we will show that this connection between chest bank presence and the cash contraction is supported by data.

To measure the local importance of chest banks, we combine 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 are only available at the bank-type level (G_d).[32] 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.[33] 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 of banks of type g with at least one currency chest in the district.[34] Since we want to interpret our instrument as a measure of the strength of the shock, our final score Exposure_d is simply the converse of the above chest measure *i.e.* $\text{Exposure}_d = 1 - \text{Chest}_d$. The score is characterized by a very smooth distribution centered on a median around 0.55, with large variation at both tails (Figure E.15). Overall, exposure appears to be evenly distributed across the country, as very high and very low exposure districts can be found in every region (Figure E.16). Consistent with this idea, in the robustness section we show that results do not depend on any specific part of the country.

According to the logic of our approach, we expect areas where chest banks are less prominent — or have

[31] 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.

[32] 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%).

[33] A simple example may help. Assume we want to figure out the local share of deposit by two rural banks A and B. From the data, we know that rural banks in aggregate represents 20% of deposits in the district, and that bank A has 3 branches in the district, while bank B only has one. Our method will impute the share of deposits to be 15% for bank A, and 5% for B.

[34] 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.

higher exposure according to the index — to have experienced a higher cash contraction during the months of November and December. While we cannot directly observe the cash contraction at the local level, we can use deposit data to proxy for it. Cash declined 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 should proxy for the cash contraction in the local area. Figure E.17 provides evidence consistent with this intuition by plotting deposit growth across districts for the last quarter of both 2016 and 2015. In normal times (2015), the growth distribution is relatively tight around a small positive growth. During the Demonetization, the distribution looks very different. First, almost no district experienced a reduction in deposits. Second, the median increase in deposits was one order of magnitude larger than during normal times. Third, there is a lot of dispersion across districts, suggesting that the effect of the Demonetization was likely not uniform across Indian districts.[35]

Using this proxy for the cash crunch, we can provide evidence that supports the intuition behind our identification strategy. Figure 6 shows that there is a strong relationship between district-level exposure to the shock and deposit growth. The same relationship holds when using different measures of deposit growth and including district-level controls, as shown in Table E.1. Importantly, Table E.2 also shows that this strong relationship only holds during the Demonetization quarter, therefore further validating our approach.[36]

Econometric model Using this measure of exposure, for different outcome variables of interest, y, we estimate the following difference-in-difference model:

$$\log\left(y_{d,t}\right) = \alpha_t + \alpha_d + \delta\left(\text{Exposure}_d \times \mathbf{1}_{\{t \geq t_0\}}\right) + \Gamma_t' Y_d + \epsilon_{d,t}, \tag{5}$$

where t is time (month), d indexes the district, t_0 is the time of the shock (November 2016), and Exposure_d is the measure of the district's exposure constructed with chest-bank data, as explained above. The equation is estimated with standard errors clustered at the district level, which is the level of the treatment (Bertrand et al., 2004). Lastly, the specification is based on the data between May 2016 and June 2017.[37]

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,

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

[36] In particular, the Table uses data since 2014 and shows that, in normal times, the relationship between these two quantities is small and generally insignificant. In the only other case in which this relationship is positive, this effect is one-fourth of the magnitude of 2016Q4.

[37] We exclude sparsely populated northeastern 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, and Tripura while union territories include Anadaman and Nicobar Islands, Chandigarh, Dadra and Nagar Haveli, Daman and Diu, Lakshadweep and Pondicherry. Altogether these regions account for 1.5% of the Indian population.

because chest exposure is clearly not random. Table 2 examines this issue, by showing the difference across characteristics for districts characterized by different exposure. Exposure to chest banks is uncorrelated with several district-level demographic and economic characteristics, but not all of them. In particular, higher exposure is found in districts with a smaller deposit base, a smaller population, and a larger share of rural population. However, most of the variation in exposure is absorbed once we control for two observables: the size of the deposit base in the quarter before the shock and the percentage of villages with an ATM (last columns, Table 2). Taking a more conservative approach, our controls include the log of deposits in the quarter before the Demonetization, the percentage of villages with an ATM, the log of population, the share of villages with a banking facility, and the share of rural population.

Results Table 3 shows that districts more exposed to the cash contraction also experienced higher adoption of electronic payments. Column 1 shows that districts that were more exposed to the shock saw a larger increase in the amount transacted on the platform in the months following the Demonetization. This result is both economically and statistically significant. Districts with one standard deviation higher exposure experienced a 55% increase in the amount transacted on the platform relative to the average. Similarly, the number of firms operating on the platform — our main measure of adoption — increased by 20% more in districts with one standard deviation higher exposure to the shock (Column 2).[38]

In Figure 7 (first two panels) we plot the dynamics of the main effect, i.e. the month-by-month estimates of how districts characterized by different levels of exposure responded to the shock.[39] This figure highlights three main findings. First, it confirms that our main effect is not driven by differential trends across high- vs. low-affected areas. Second, the shift in adoption across districts happened as early as November. Third, the difference in the response also persists after the cash availability returns to normal level. In particular, the effects are still large and significant after the month of February. These findings, taken together with the aggregate-level evidence in Section 2, confirm that the temporary cash contraction led to a persistent increase in size of the user base of the electronic payment technology, consistent with the first prediction of the model.

Next, we test the second prediction of the model, which is that the shock led to a persistent increase in the adoption rate, that is, the flow of new users in the platform. We empirically test this by analyzing whether districts more affected by the shock witnessed a more persistent increase in new adopters. We define new adopters at time t as the firms using the technology for the first time at time t. The third panel of Figure 7 shows that districts experiencing a larger contraction in cash saw a larger increase in new adopters

[38] To be conservative, we measure the number of active firms in the platform as firms with at least 50 Rs. of transactions in the period. We obtain similar results when we use different transaction thresholds (including a threshold of zero).

[39] The specification is $\log\left(y_{d,t}\right) = \alpha_t + \alpha_d + \delta_t\left(\text{Exposure}_d\right) + \Gamma_t' Y_d + \epsilon_{d,t}$, and October is the base month.

joining the platform as early as on November 2016. Crucially, the relative increase in the number of new adopters continued even after January 2017, the last month during which cash withdrawal was constrained, and persisted for the whole of spring 2017. This persistent increase in new users is consistent with the second prediction of the model.

Robustness As stated above, we argue that the relationship between exposure to the cash contraction and adoption of electronic payments is causal. Consistent with this interpretation, we have shown that, conditional on covariates, more exposed areas do not look different than less exposed regions in pre-shock levels. Additionally, our effects are not driven by pre-trends across affected districts. As a further robustness check, we note that our main results are not driven by the response of any particular region in the country: our effects are stable when excluding any of the Indian states from our analysis (Appendix Figure E.18).

Given these results, one remaining concern to rule out is the presence of a contemporaneous demand shock that is correlated with our exposure measure but it is unrelated to the cash scarcity. We provide two tests to rule this out, which Appendix C expands on. First, we show that the same highly affected districts also experienced a larger decline in consumption during this period. In particular, using the same empirical model and a panel of almost 100k households in India around the Demonetization, we document that exposure to the cash contraction led to a temporary contraction in total consumption. This effect is mostly driven by a reduction of non-essential consumption items (e.g. recreational expenses). This result is interesting on its own, but also helps rule out the possibility that unobserved demand shocks could explain our results. In fact, a demand-side explanation of the increase in electronic payments would require that highly exposed districts receive a positive demand shock. Our findings reject this hypothesis and actually find that — consistent with a supply-side interpretation — highly affected areas saw a reduction in consumption.[40] Second, Appendix C also presents a full set of placebos that exploit the longer-panel dimension in the consumption data and confirm the quality of our empirical strategy.

4.3 State-dependence in adoption

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 a permanent shift in adoption, but that the increase in adoption will not be uniform across regions: it will crucially depend on the initial strength of complementarities in the area. In this section, we use the data on electronic payments to present evidence that are consistent with this prediction. Conceptually, the objective of these tests is not to causally identify a relationship between variables, but rather to generate empirical regularities that would support

[40] Appendix C discusses more in detail these tests as well as the possible alternative demand channels we rule out.

the importance of state-dependence in explaining the data. To do so, we will try to isolate the role of the state-dependence mechanism from alternative economic forces that might also be consistent with the same result. While none of the tests will be perfect, we believe that all of them together provide convincing evidence on the empirical importance of state-dependence.

Initial size and adoption In the model, the strength of complementarities in a district is completely captured by the pre-shock size of the user base. Specifically, the model predicts that the initial size of the platform in a district should amplify the adoption response to the shock. In Table E.4, we provide evidence that the data are consistent with this prediction.[41] In particular, 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 simple evidence is thus consistent with state-dependence as defined in the model. However, it suffers from two shortcomings. First, in reality, 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 or not adoption at the district level is the correct way to proxy for their initial strength. The second is the standard reflection problem (Manski, 1993; Rysman, 2019). Past adoption decisions by firms in the district may reflect unobservable heterogeneity across firms in a district that are unrelated to the strength of complementarities, but correlated to subsequent adoption decisions. To overcome these limitations, in what follows, we propose and implement two alternative tests for state-dependence, one involving district-level data, and the second involving firm-level data.

District level test First, we test how the increase in adoption differs depending on the distance between a district and areas in which the usage of the electronic wallets 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 also be 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 Comin et al. (2012).[42]

[41] Specifically, we estimate:

$$X_{d,t} = \alpha_t + \alpha_d + \delta \left(I_d \times \mathbf{1}_{\{t \geq t_0\}} \right) + \Gamma_t' Y_d + \epsilon_{d,t}, \tag{6}$$

where I_d is initial adoption levels in a district, measured either by dummy for whether a district had a positive adoption level before the Demonetization, or by the total amount of transactions in the district before the Demonetization.

[42] In particular, we define a district to be an electronic payment hub if there were more than 500 active firms pre-Demonetization (September 2016). The results are essentially identical if we use a threshold of 1,000 firms to define the hub districts. The nine hubs are spread evenly across the country. In particular, these districts are: Delhi, Chandigarh and Jaipur (North), Kolkata (East); Mumbai and Pune (West); Chennai, Bangalore and Rangareddy (South). The distance to the hub is defined as the minimum of the distance between the district and all the hubs.

We implement this test by running a simple difference-in-difference model where we compare the usage of wallet technologies around the Demonetization period across districts that are differentially close to a digital wallet hub. Despite the clear advantages relative to the naive evidence presented above, there still are two concerns with this approach. First, by sorting on distance we might capture variation coming 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 most important cities in the country, we should expect that being located close to them will have benefits that go beyond the effect of complementarities.[43]

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. 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 the effect of being located close to a large city. Third, as in the previous analyses, we augment the specification with a wide set of district-level covariates interacted with the time dummies. This implies a specification of the following form:

$$X_{d,s,t} = \alpha_{st} + \alpha_d + \delta \left(D_d \times \mathbf{1}_{\{t \geq t_0\}} \right) + \gamma \left(\tilde{D}_{d,s} \times \mathbf{1}_{\{t \geq t_0\}} \right) + \Gamma_t' Y_d + \epsilon_{d,t}, \tag{7}$$

where t indicates time, defined at the monthly level in this analysis, d indexes the district and s identifies the state of the district. D_d is the district's distance to the nearest electronic-wallet hub and $\tilde{D}_{d,s}$ is the district's distance to the capital district of the state. As before, standard errors clustered at district level.[44] The main coefficient of interest is δ — which provides the difference in the 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 4. Across all the outcomes — the amount of transactions, number of operating firms and number of new adopters — we find that the districts farther away from major hubs experienced a lower increase in the post-Demonetization period. The most conservative of the estimates indicates that a 50km increase in distance translates into a 19% lower increase in the amount of transactions. Importantly, these effects are not driven by differential trends in adoption between areas that are closer and further from hub cities (Figure 8). In general, distance does not matter before November, but it predicts differential responses starting in December.[45]

[43] A third concern is that distance may simply capturing variation in exposure to the shock, as defined before. However, we actually find that the two treatment variables are uncorrelated.

[44] It is important to point out that we remove major digital wallet hubs from this analysis. Notice that this exclusion does not affect our results; the results that includes the hubs are, if anything, stronger.

[45] In Table E.5, we show that we find similar results when we use a dichotomous definition of the treatment. In particular, we consider several alternatives, going from 400km down to 200km. Across all these tests, the results are stable and significant.

Firm level test Our second test examines the role of state-dependence using firm-level data. In the model, analogously to the district-level prediction, state-dependence implies a positive relationship between a firm's use of the technology and the overall use by other firms in the same area.[46] Using firm-level data, we can directly test this prediction, following an approach that is consistent with the empirical literature on spillover (e.g. Munshi 2004, Goolsbee and Klenow 2002). Furthermore, the use of firm-level data allows us to control directly for several dimensions of heterogeneity that may explain adoption decisions for reasons unrelated to externalities.

For each firm, we measure the total use of the technology by other companies located in the same geographical area and operate in the same industry. We choose this reference group because we believe that complementarities should be strongest among firms in the same area and industry. For instance, we 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. [47] In particular, we estimate:

$$x_{i,p,k,t} = \alpha_i + \alpha_{p,t} + \alpha_{k,t} + \rho x_{i,p,k,t-1} + \gamma X_{p,k,t-1} + \epsilon_{i,p,k,t}. \tag{8}$$

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 in the period May 2016-June 2017).[48] 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.[49] The variable $X_{p,k,t-1}$ is a measure 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 we used as the 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.[50] Standard errors are clustered at the location level (pincode).

Results reported in Table 5 provide evidence consistent with state-dependence. Across several specifications, we find that a higher volume of electronic transactions by firms in the same reference group strongly predicts more transactions for the firm itself in the following week. For instance, in our baseline we have that 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

[46] Consistent with this discussion, Appendix B.8 shows, using Monte-carlo evidence, that this positive relationship arises in the model when there are externalities, and is absent otherwise.

[47] Our results also hold when using alternative definitions of the reference group. For instance, in Table E.6 in the Appendix we define the relevant market as any firm in the same location (pincode), irrespective of the industry.

[48] We use pincode to identify firms' locations because we want to use the narrowest definition of location that is available in the data. Our main results also hold using districts (Table E.7 in Appendix).

[49] 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%).

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

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.

Overall, the main concern in this analysis is that past decisions by firms in the reference group may correlate with an individual firm's behavior because of unobservable heterogeneity across firms which are unrelated to the strength of complementarities — the reflection problem. To assuage this problem, we show that results still hold once we augment the baseline with firm fixed-effects (column 2), pincode-by-week fixed-effects (column 3), and industry-by-week fixed-effects (column 5) altogether. Relative to the baseline specification specification (column 1), the addition of these fixed-effects will allow us to keep constant in the model any characteristics of the area — even to the extent that these characteristics have a differential effect over time — and also adjust the estimates for changes in adoption rates in the same industry.[51]

We conclude by repeating the same analysis as before, but allowing for month-specific parameters for each of our outcomes (Figure 9).[52] Across the three outcomes, there are two key findings. 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 a function of the shock but a feature of technology choices in any scenario. Second, the effect of adoption in the reference group is much higher in the months of the Demonetization, relative to the preceding and succeeding months.

4.4 Alternative mechanisms and discussion

Overall, the evidence suggests that the Demonetization caused an adoption wave with features that are *qualitatively* consistent with three predictions of the model with externalities: (i) a persistent increase in the *size* of user base; (ii) a persistent increase in the *adoption rate*, that is, the flow of new users into the platform; and (iii) state-dependence in responses, that is, a positive relation between initial adoption rates and the initial strength of adoption externalities, broadly defined. In the context of our model, these predictions are *specific* to the presence of externalities, so these reduced-form results support the notion that externalities played a key role in shaping the adoption response following the shock.

In this section, we discuss three aspects of our analysis. First, we examine the extent to which factors that are not in the scope of our model (i.e. competition among platforms, fees and marketing) may nevertheless

[51] The inclusion of individual fixed-effects in a dynamic model may bias the main parameters in the model, as first discussed in Nickell (1981). However, there are two important things to highlight about our application. First, the presence of fixed-effect is not necessary to obtain the desired result, since we still find the same effect without any fixed-effects (column 1). Second, the Nickell bias is a feature of models characterized by short panels, as the bias converges to zero as $T-1$ increases, where T is the time-dimension in the panel. In our case, T is relatively large - data is at weekly level and the time span is almost a year - and therefore the bias will be small in magnitude. In particular, since our main prediction is on the direction of the relationship rather than on the exact magnitude, this issue will not affect the conclusion of this study.

[52] In fact, the model suggests that the effect estimated should actually be different across time. In particular, using the simulated data from the model, we can show that the importance of the adoption by other firms is particularly large in the shock period.

influence our reduced-form results. Second, we discuss which mechanism is most likely to explain the presence of externalities among retailers in their decision to adopt, comparing the role of pure network effects to learning among retailers. Lastly, we provide a final interpretation of our reduced form evidence.

To start, we do not believe that competition between platforms plays a primary role in our analysis. Within fintech, our partner firm was the largest provider in India, and could be considered the de facto monopolist for most of the sample period. At the same time, other traditional electronic payments did not experience any increase in new adopters at the time of the Demonetization (as discussed in Section 2). Furthermore, the nature of our data also implies that competition between platforms should increase measurement error and therefore – if anything – this would bias our analyses towards finding no effects.

Furthermore, we also want to stress how marketing efforts and pricing strategies by the platform should not be important confounding factors. If local marketing spending by the partner company is correlated with our district-level treatment, then our effects would capture responses to such marketing efforts. However, our partner company organizes customer acquisition through national campaigns, and there was no program targeting specific local areas. At the same time, pricing strategies to overcome coordination failure — as suggested in Rochet and Tirole (2006) and Weyl (2010) – do not play an essential role in our analysis as the fees to join the platform were zero during our sample period.[53]

While our results strongly support the idea that complementarities are key to explain the increase in adoption, we have been so far silent about the exact sources of complementarities in our context. The model of Section 3 does not take a clear stand on this; instead, it captures complementarities in reduced form, by assuming that the returns to adoption increase with the number of other adopters. As discussed there, complementarities in the decision to adopt can arise because of the network nature of electronic payment platforms (Katz and Shapiro, 1994; Rysman, 2007). At the same time, learning about the costs and benefits of the technology — either through social interactions or by observing the experiences of peers — could also make adoption decisions complements (Munshi, 2004; Young, 2009; Bailey et al., 2019).[54]

While some features of the adoption response may differ depending on the source of externalities, the main implications of our study do not depend on the nature of the underlying mechanisms. In particular, persistence in adoption after a temporary shock and state-dependence does not depend on the specific mechanism generating complementarities. While a quantification of the relative importance of these different mechanisms is outside the scope of this paper, we think that shedding more light on these mechanisms is important and more feasible empirically. We next provide novel evidence suggesting that learning is unlikely

[53] It is also important to highlight that our period of analysis ends before the introduction of the GST tax in India, which took place in July 2017.

[54] Notice that this is different than simply saying that people became aware of the technology after Demonetization. This simpler mechanism is akin to a fixed cost and therefore cannot explain the whole set of results, as discussed earlier in the paper.

the key driver of our results.

First, to the extent that social learning is the main driver of our results, we should expect to find stronger effects of cash contraction in districts where social learning among firms is more prevalent. We explore this idea by constructing a measure of social learning at district level and augmenting our baseline specification with additional interactions of our measure of cash contraction and proxy for social learning. We use the degree of language concentration in a district, since the social interactions and size of social networks in a region arguably increases when there are more individuals speaking the same language in that region (Michalopoulos, 2012).[55] Column (1) of Table E.8 shows that our main effects remain unchanged and that districts with higher language concentration do not have different adoption responses in the post-Demonetization period. In Column (2) we also do not find any evidence that the responses from cash contraction is different in districts with more language concentration. Since learning could also occur from peers that adopted the technology (Banerjee, 1992), we also test whether we find stronger effects when there are more similar firms in a region (Munshi, 2004).[56] However, also in this case, we find evidence that is inconsistent with this hypothesis (columns 3 and 4, Table E.8).

Lastly, since learning cannot be undone (at least within a few months), there should be more limited reversal after the shock if this feature were the key source of complementarities. Places that reached high adoption in January should remain at elevated levels afterwards.[57] In general, the data seem at odds with this scenario. In fact, despite the average increase in adoption, slightly more than a quarter of our district-month pairs experienced some negative growth after the Demonetization. Related to this point, we also find that firms that were early adopters — defined as firms already using the electronic wallet in October 2016 — experienced a substantial increase of about 100% in number of transactions between October 2016 to May 2017. If our effects were driven by learning, then there should have been very limited response from early adopters. Altogether, while we cannot definitively rule out that learning did not play a role in generating complementarities in adoption, these tests provide strong support that our estimates are not simply capturing the effects of learning.

The key take-away from this section is that complementarities in adoption decisions are necessary to rationalize simultaneously the persistence and the state-dependence in adoption documented in the data.

[55] We define language concentration in a district as: Language Concentration$_d = 1 - \sum_l s_{dl}^2$ where s_{dl} is the share of district d population speaking language l. We obtain the information on language distribution among population using Census of India, 2011.

[56] To examine this dimension, we use 4-digit industry concentration (based on share of employment) within a district as a proxy for observational learning, since the possibility of learning through peers in a region is arguably higher when there are more firms within the same industry in that district.

[57] We acknowledge that this difference in predictions between complementarities based on learning versus network externalities is beyond the scope of our simple model. In our simple model, state-dependence implies there can be long-run mean-reversion if initial adoption is sufficiently weak or if the initial shock is sufficiently small. Our point here is that in richer model where complementarities are explicitly modelled via learning, mean-reversion should, in principle, be weakened. Furthermore, this evidence also confirms our initial modeling assumption.

However, these results leave two related questions open. First, they do not indicate how large these complementarities are in the data — that is, in terms of the model of Section 3, how large C is. Second, from these reduced-form results, no conclusions can be drawn about the effects of alternative shocks (say, a more transitory cash contraction). By the same token, these results are silent regarding how a policymaker seeking to promote adoption may decide to structure an intervention. These two issues are connected since they both require some knowledge of the strength of externalities in this setting. In the next section, we address both by structurally estimating the model using the data on electronic wallet adoption and studying the model's counterfactual and policy implications.

5 Quantifying the role of complementarities

5.1 Estimation

We use the simulated method of moments to estimate the key parameters of the model. We start by describing briefly our approach, focusing on the intuition for how specific moments help identify different model parameters. We then report the results and discuss model fit.

Methodology and identification We calibrate two parameters. First, we set $r = -\log(0.90)/12$, corresponding to a time discount rate of 0.90 per year.[58] Second, we set $\theta = -\log(1-0.90)/(82/30)$, where θ is the (inverse of) the persistence of innovations to the money stock.[59] Additionally, and without loss of generality, we normalize the long-run mean of $M_{d,t}$ to $M^c = 1$.

We estimate the remaining $N_p = 5$ parameters of the model, $\Theta = (S, C, k, \sigma, M_e)$. They are, respectively, the size of the Demonetization shock (S), the strength of complementarities in adoption (C), the Poisson arrival rate of the technology switching shock (k), the standard deviation of normal innovations to the money stock (σ), and the profits associated with the electronic payments technology when there is no adoption (M^e). In order to estimate those parameters, we use the following set of regressions, on a balanced panel of districts:

$$\begin{aligned}
\Delta_{t_0} X_{d,t} &= \beta + \gamma \mathbf{1}\{t \geq t_0 + 3\} + \delta X_{d,t_0} + \zeta\left(\mathbf{1}\{t \geq t_0 + 3\} \times X_{d,t_0}\right) + \epsilon_{d,t}, \\
\hat{var}_t(\Delta_{t_0} X_{d,t}) &= \eta + \kappa \mathbf{1}\{t \geq t_0 + 3\} + \mu_t, \\
\hat{var}_d(\Delta_{t_0} X_{d,t}) &= \nu + \omega_d,
\end{aligned} \tag{9}$$

and we additionally estimate the average of the squared residuals $\hat{\epsilon}^2_{d,t}$ from the first regression in (9), through

[58] As described in appendix B, the model is solved using a discrete time approximation where $\Delta t = 1/10$, so that a time period is approximately one tenth of one month.

[59] This choice ensures it takes on average 82 days for the aggregate shock to be 90% dissipated. The choice of 82 days corresponds to the time which elapsed between the announcement of the cash swap (November 8th, 2016) and the date at which the government lifted most remaining restrictions on cash withdrawals (January 30th, 2017).

$\hat{\epsilon}^2_{d,t} = \xi + \omega_{d,t}$. In these regressions, d indexes the 512 districts included in our analysis, and t indexes months. The month t_0 is October, 2016 (the last month observed prior to the Demonetization shock), and $\Delta_{t_0} X_{d,t}$ is the cumulative change in adoption rates: $\Delta_{t_0} X_{d,t} = X_{d,t} - X_{d,t_0}$. We use the 8 months running from November, 2016 to June, 2017.[60] We compute the participation rate in each district, $X_{d,t}$, as the ratio of the number of monthly users active on the platform during month t, divided by the number of retailers with less than three employees, which we obtain from the 2014 Census.[61] Finally, $\hat{var}_t(.)$ denotes cross-sectional variances, while $\hat{var}_d(.)$ denotes within-district variances.

In order to estimate our 5 data parameters, we use $N_m = 8$ data moments from the regressions above: $\hat{\Xi} = (\hat{\beta}, \hat{\gamma}, \hat{\delta}, \hat{\zeta}, \hat{\xi}, \hat{\eta}, \hat{\kappa}, \hat{\nu})$. Appendix D reports the details of the estimation procedure. We use the bootstrap, clustering by district, in order to construct the variance-covariance matrix of data moments. Here, we focus instead on the intuition for why the chosen data moments help identify the five estimated parameters.

Our main parameter of interest is the strength of complementarities, C. Consistent with our earlier discussion of the model, this parameter is primarily identified by the difference between the short and medium-run response of adoption to the shock, $\hat{\gamma}$.[62] Without adoption complementarities ($C = 0$), the short-run adoption wave triggered by the shock has no bearing on the adoption decision of firms further down the road. As a result, once the shock is dissipated, there should be no further adoption by new firms. The model would then predict that $\hat{\gamma} = 0$. By contrast, when adoption complementarities are present ($C > 0$), the short-run adoption wave raises the value of future adoption for other firms, and so new firms continuing adopting even once the shock is dissipated, leading to positive values of $\hat{\gamma}$. Additionally, as discussed earlier in the paper, the dependence of the response to the shock on initial conditions $(\hat{\delta}, \hat{\zeta})$ also helps pin down the strength of adoption complementarities.

The rate at which firms reset their technology choice, k, is identified using estimates of the between-district variance of the change in adoption, $\hat{\eta}$ and $\hat{\kappa}$. The medium-run variance, $\hat{\kappa}$, is particularly informative about k. As highlighted in our ealier discussion, if firms reset their technology quickly relative to the persistence of the shock (i.e. k is sufficiently high relative to θ), then all districts will rapidly converge to full adoption, thus leading to lower cross-sectional variance in adoption rates in the medium-run.

Finally, the size of the shock, S, is primarily identified by the short-run adoption caused by the shock, which is $\hat{\beta}$. Absent an aggregate shock, $\hat{\beta}$ is not statistically different from 0, and the magnitude of the coefficient increases with the size of the shock, independent of the existence of complementarities. The

[60] We substract the initial adoption rate in order to eliminate district-specific fixed effects, but results either in levels or adding explicit fixed effects in the estimation of (9), are similar.

[61] Additionally, we re-normalize the Census retail counts so that the five districts by adopter share reach full adoption. Appendix D discusses this normalization in more detail, and shows that it does not materially affect our results.

[62] Here, we define "medium-run" as three months after the shock; by then, in the data, cash circulation had returned to pre-shock levels; and, in the model, the aggregate shock is more than 90% dissipated.

standard deviation of idiosyncratic innovations to districts, σ, is identified using the variance of residuals from the first equation in (9). The residual variation in adoption, after controlling for initial condition, should be driven by district-level shocks. The rate of profits associated with the electronic payments technology when there is no adoption, M^e, is identified using the variance of within-district adoption rates. Even when there are no complementarities, a lower level of M^e is associated with shorter adoption spells, and therefore lower overall volatility of adoption rates.

Results Table 6 reports estimates of the five structural parameters. The point estimate for the size of the shock, S, is 24.6% (with a 90% coverage interval of $[15.0\%, 34.1\%]$). The parameter S expresses the decline in profits associated with cash-based transactions, relative to their long-run mean. There are two numbers with which this estimate could be compared. First, recall that the cash denominations which were voided by the shock represented 86.4% of the total currency in circulation. The shock size we estimate is much smaller than this, but not all of the voided currency was actively used in transactions prior to shock (though it is difficult to measure exactly what fraction was). Second, Chodorow-Reich et al. (2019) estimates that the general equilibrium response of output to the shock was approximately 3 percentage points. Aside from being a general equilibrium estimate, this figure expresses the response of value added (not profits), includes the potential effects of substitution into electronic payments technologies, and encompasses all sectors of the economy. For these reasons, it is likely a lower bound on the size of the shock. Our point estimate however has a reasonable magnitude compared to theirs: for instance, assuming a labor share of 70% in retail, and no adjustment of labor or hours in the short-run, the implied decline in profit rates in retail using the 3% figure is $1/0.3 \times 3\% = 9\%$, or approximately one third of our point estimate.

The magnitudes of the point estimates for the level and the slope of the switching frontier are difficult to interpret explicitly, but it is worth making two points about them. First, the point estimate of C is 0.063, with a 90% coverage interval of $[0.056, 0.070]$. Our findings therefore reject the null of no adoption complementarities. Second, the point estimates imply that relative to cash, profits under the electronic technology are on average 3.0% lower if there are no other adopters, and 3.3% higher if there is full adoption. Together with other parameters, these differences imply that the equilibrium switching frontier is such that cash-based profits $M_{d,t}$ must fall by 12.6% in a district with $X_{d,t} = 0$ adoption, or about three standard deviations, in order for adoption to start. The estimated size of the shock substantially exceeds this threshold.

Finally, the point estimate of the rate of technology resetting implies that, on average, firms receive the option to adjust their technological choice every 6.1 months, with the 90% coverage interval of the arrival rate corresponding to frequencies between 4.3 and 10.6 months. The estimate of k is fairly imprecise, but it implies that arrival rates higher than 3 months can be rejected at the 1% level. As discussed earlier, this

relatively slow technological adjustment rate may reflect learning or cognitive costs associated with the use of the technology. Additionally, we can reject, at the 1% level, the null that $k > \theta$, so that at its point estimate, the model generates state-dependence in impulse responses.

Table 7 reports measures of goodness of fit. The first column reports the empirical value of the moments used in the estimation. The second column provides average values, standard deviations, and one-sided p-values obtained from $S_{CI} = 2000$ simulations of the model with structural parameters set to their estimated values, i.e. $\Theta = \hat{\Theta}$. We can reject equality of the empirical and simulated moments at the 1% for two of the eight moments, and overall, the over-identification test cannot reject the null that the model is correctly specified at the 1% level.[63] The model matches the empirical moments closely, particularly the short- and medium-run average adoption response, the two most precisely estimated moments in our data.[64]

5.2 Counterfactuals

Next, we use the estimated model to construct the quantitative answer to three questions about the effects of the shock, and the role played by complementarities in the adoption process.

How would adoption have responded, in the absence of complementarities? Figure 10 reports empirical and model-based paths of average adoption across districts, in the aftermath of the shock. At the point estimates reported in Table 6, adoption rises by approximately 4p.p. by the end of December, and 7p.p. by the end of May, in line with the empirical estimates. This result is not surprising, since these moments were explicitly targeted. The figure also reports a counterfactual path of adoption rates, under the assumption that there are no complementarities, that is, when $C = 0$. With respect to the data, and to our baseline estimate, the adoption path is similar during the first three months, when the cash crunch is still ongoing. After that, it diverges from the data and from the model with complementarities, declining in the medium-run. The gap is fairly substantial: the predicted increase in adoption rates without complementarities would have been 4p.p. (or approximately 60%) lower than observed. Thus, the model attributes a important share of the response of adoption rates to complementarities.

[63] The moment that the model has the most trouble matching is the effect of initial adoption on the long-run response. The reason is primarily that this moment is imprecisely estimated in our data, given the limited variation in intial adoption rates, so that it receives low weight in the estimation objective. A stronger degree of complementarities than the point estimate for C implies would help bring model and data closer in this respect.

[64] The moment with the worse fit is the interaction between the medium-run effect of the shock and the pre-shock adoption level. This moment depends strongly on the variance of innovations to the idiosyncratic shocks to districts, which we primarily identify using the average squared residuals from equation (9). A higher variance of idiosyncratic innovations, in the model, will mask the dependence on initial conditions; but it may be necessary to match large squared residuals in empirical estimates of (9), which themselves may reflect unmodelled sources of variation in adoption rates.

What if the cash swap had been completed more quickly? Figure 10 also reports counterfactual adoption paths which speak to the role of the size and persistence of the shock. We first construct adoption paths under the assumption that a 90% decay rate of the shock is two weeks, instead of three months; this captures an alternative world in which the cash swap would have been executed as rapidly as initially intended. Under this scenario, adoption would only have risen by approximately 2p.p., and the increase in the dispersion of adoption would have been negligible. Figure 10 also indicates that, if the shock had been smaller in magnitude — which could capture a situation in which only one denomination would have been replaced, for instance — the long-run response would have been smaller. With a shock half as large, the average adoption rate only rises by approximately 4p.p., versus 7p.p. in the baseline case. The model thus suggests that the persistence and size of the cash crunch might have had substantial, though unintended, positive effects on adoption overall.

What sort of intervention maximizes long-run adoption? We next use the model to ask whether a hypothetical policymaker could have achieved higher long-run changes in adoption rates by implementing the cash swap differently. In order to answer this question, we first define the cost of the cash swap as the present value of the decline in cash-based demand:

$$C(S,\theta) = \mathbb{E}_{t_0}\left[\sum_{n=0}^{+\infty} e^{-r\Delta n}\left\{M^c - M_{t_0+\Delta n}\right\}\right] = \frac{S}{1-\exp(-(r+\theta))}. \tag{10}$$

We next consider the following maximization problem for the hypothetical policymaker:

$$\arg\max_{S,\theta} \quad \mathbb{E}_{t_0}\left[\Delta_{t_0} X_{d,t_0+T}\right] - \frac{g}{2} var_{t_0}\left[\Delta_{t_0} X_{d,t_0+T}\right] \qquad \text{s.t. } C(S,\theta) \leq C(\hat{S},\theta_0) \tag{11}$$

where $\hat{S}$ is the value reported in Table 6, $\theta_0 = \log(1-0.90)/(82/30)$ is the calibrated shock persistence, and g is an arbitrary positive number.

This is the problem facing the hypothetical policymaker who chooses the size and persistence of the shock to cash-based demand, aims to maximize average adoption at horizon $T = 3$ years, and possibly exhibits some aversion to dispersion in adoption rates (when $g > 0$). The aversion to dispersion could capture a preference of policymakers toward broad-based adoption. Furthermore, we assume this policymaker is constrained in the total cost of the intervention, and we use the empirically estimated cost of the Demonetization shock as the maximum cost the policymaker can incur.

Table 8 reports the numerical solution to problem (11), under different values of g. Additionally, the first column reports the model estimates of the size and persistence of shocks, and the implied long-run first and second moments of the change in adoption rates.

Results for the first column, $g = 0$, show that the "constrained optimal" plan, for a policymaker that does not care about long-run dispersion in adoption rates across districts, involves choosing a shock that is more persistent but smaller than what we estimated. Thus, the model indicates that, given the total cost of the intervention implied by the model estimates, a policymaker seeking to maximize long-run adoption could have done better than the observed outcome, by making the shock both more persistent and smaller. The difference with respect to the estimated shock, however, is not large (the shock half-life is approximately one month, instead of 0.8 month in the baseline case, and the shock would have been only 20% smaller in size).

Because the "constrained optimal" shock is smaller, it also leads to more dispersion in adoption rates in the long run. The intuition for this is that, with a smaller shock, a higher initial adoption rate is required for the district to enter the adoption region. The initial differences between districts are then exacerbated. As a result, long-run dispersion in the "constrained optimal" plan when $g = 0$ is higher than in the model estimates, as indicated in Table 8.

However, as aversion to dispersion increases (that is, as g increases), the "constrained optimal plan" progressively involves smaller and more persistent shocks. As discussed in Section 3, a more persistent shock (of a given size) tends to reduce long-run dispersion in outcomes, because it reduces the degree of state-dependence of adoption rates. In the limit where the shock is permanent, all districts whose initial adoption rates are sufficiently high that the shock triggers *some* adoption in the short-run, will also converge to full adoption in the long-run. A policymaker who cares about dispersion thus has a motive to *further* increase the persistence of the intervention; at the point estimates of the model, this effect dominates the reduction in dispersion that a larger shock might generate. In particular, the "constrained optimal" plan with an aversion to dispersion of $g = 0.5$ leads to comparable long-run dispersion than in the estimated model, but a lower average adoption rate.

Overall, while the size and persistence of the shock had positive effects on long-run adoption — as discussed above —, the model also suggests that if the objective of the policy had been to increase long-run adoption while minimizing the dispersion in outcomes across districts, a more persistent but smaller intervention would have been preferable. That said, long-run adoption gains under these alternative policies are relatively mild, in the order of 10% to 15% of the long-run adoption increase implied by the estimates.

The analysis of this section has shown that the simple model of Section 3 can account well for key moments of the data. Counterfactuals suggest that complementarities account for 60% of the medium-run response of adoption, and that a smaller, but more persistent intervention may have led to a larger increase in long-run adoption rates, along with a lower long-run dispersion in adoption across districts.

6 Conclusion

An increasing number of new technologies feature network externalities. When this is the case, the technology's ability to grow and scale is subject to coordination frictions. How can this coordination friction be overcome? Furthermore, how can a policy intervention help to foster adoption? In this paper, we used the Indian Demonetization of 2016, and its subsequent effect on the adoption of electronic wallets, as a laboratory to study these questions.

We started by showing that the Demonetization led to a large and persistent increase in the overall use of this technology, even though the Demonetization shock itself was temporary. We argued that this large and persistent increase is consistent with a dynamic technology adoption model with externalities, and we derived some additional testable predictions unique to externalities. In particular, we showed that in this model, a temporary shock can cause a persistent increase in the adoption rate of the platform (as opposed to only its size), and that the response of adoption rates depends positively on initial adoption levels.

Using micro data on electronic payments, we then showed that these additional testable predictions are supported by the data. At the the district level, we proposed a novel identification strategy based on heterogeneity in the presence of chest banks to estimate the causal impact of the cash crunch. We showed that the cash crunch caused a persistent increase in the adoption rate by firms of electronic wallets. Additionally, the adoption responses are characterized by positive state-dependence, both at the district and the firm level. Finally, we provided a structural estimation of our dynamic model. This estimation suggests that about 60% of the total adoption response is due to complementarities.

Our analysis also highlighted some of the challenges faced by policymakers in environments with complementarities. In those environments, large, punctual interventions can have permanent effects on adoption because they effectively act as coordinating devices that help firms overcome coordination frictions. However, because of state-dependence, an intervention that is too brief can also exacerbate inequality in adoption rates. Policymakers may therefore face a trade-off between the length the intervention and how much it will exacerbate initial difference in adoption rates. These results are important for several economic areas, as the presence of externalities is common among technologies in the new economy.[65]

Our work suggests two avenues for future research. First, we highlighted some general testable predictions of dynamic adoption models with externalities, that could be tested in contexts other than the adoption of payments technology. Second, future work should study the strategic changes in firms' behavior in response to the adoption of electronic payments.

[65] On top of the network-based fintech sector already discussed, the presence of complementarity in adoption can be also generated by the type of social data acquisition that is typical of many online services (Bergemann et al., 2020).

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., P. Ghosh, J. Li, and T. Ruan (2019). Digital payments induce over-spending: Evidence from the 2016 demonetization in india.

Agarwal, S., W. Qian, B. Y. Yeung, and X. Zou (2019). Mobile wallet and entrepreneurial growth. In *AEA Papers and Proceedings*, Volume 109, pp. 48–53.

Alvarez, F. and D. Argente (2019). Consumer surplus of alternative payment methods: Paying uber with cash. *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2019-96).

Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal 99*(394), 116–131.

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.

Bachas, P., P. Gertler, S. Higgins, and E. Seira (2017). How debit cards enable the poor to save more. Technical report, National Bureau of Economic Research.

Bailey, M., D. M. Johnston, T. Kuchler, J. Stroebel, and A. Wong (2019). Peer effects in product adoption. Technical report, National Bureau of Economic Research.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics 107*(3), 797–817.

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. *Working Paper*.

Bartlett, R., A. Morse, R. Stanton, and N. Wallace (2018). Consumer-lending discrimination in the era of fintech. *Unpublished working paper. University of California, Berkeley*.

Beck, T., H. Pamuk, R. Ramrattan, and B. R. Uras (2018). Payment instruments, finance and development. *Journal of Development Economics 133*, 162–186.

Bergemann, D., A. Bonatti, and T. Gan (2020). The economics of social data.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics 119*(1), 249–275.

Bircan, C. and R. De Haas (2019). The limits of lending? banks and technology adoption across russia. *The Review of Financial Studies*.

Björkegren, D. (2015). The adoption of network goods: evidence from the spread of mobile phones in rwanda. *Review of Economic Studies*.

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.

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

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.

Buera, F., H. Hopenhayn, Y. Shin, and N. Tracther (2020). Big Push in Distorted Economies.

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*, 989–1018.

Carlton, D. W. and J. M. Klamer (1983). The need for coordination among firms, with special reference to network industries. *The University of Chicago Law Review 50*(2), 446–465.

Chodorow-Reich, G., G. Gopinath, P. Mishra, and A. Narayanan (2019). Cash and the Economy: Evidence from India's Demonetization. *The Quarterly Journal of Economics 135*(1), 57–103.

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

Comin, D. A., M. Dmitriev, and E. Rossi-Hansberg (2012). The spatial diffusion of technology. *Working Paper*.

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

Crowe, M., M. Rysman, and J. Stavins (2010). Mobile payments at the retail point of sale in the united states: prospects for adoption. *Review of Network Economics 9*(4).

Dharmapala, D. and V. S. Khanna (2018). Stock market reactions to india's 2016 demonetization: Implications for tax evasion, corruption, and financial constraints.

Farrell, J. and G. Saloner (1986). Installed base and compatibility: Innovation, product preannouncements, and predation. *American Economic Review*, 940–955.

Foley-Fisher, N., B. Narajabad, and S. Verani (2020). Self-fulfilling runs: Evidence from the us life insurance industry. *Forthcoming*.

Foster, A. D. and M. R. Rosenzweig (2010). Microeconomics of technology adoption. *Annual Review of Economics 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.

Frankel, D. M. and K. Burdzy (2005). Shocks and business cycles. *Advances in Theoretical Economics 5*(1).

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

Goolsbee, A. and P. J. Klenow (2002). Evidence on learning and network externalities in the diffusion of home computers. *The Journal of Law and Economics 45*(2), 317–343.

Gowrisankaran, G. and J. Stavins (2004). Network externalities and technology adoption: Lessons from electronic payments. *The RAND Journal of Economics 35*(2), 260–276.

Guimarães, B. and C. Machado (2018). Dynamic coordination and the optimal stimulus policies. *The Economic Journal 128*(615), 2785–2811.

Gupta, A., J. Ponticelli, and A. Tesei (2020). Access to information, technology adoption and productivity: Large-scale evidence from agriculture in india. *Working Paper*.

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

Hennessy, C. A. and T. M. Whited (2005). Debt dynamics. *Journal of Finance 60*(3), 1129–1165.

Hennessy, C. A. and T. M. Whited (2007). How costly is external financing? evidence from a structural estimation. *Journal of Finance 62*(4), 1705–1745.

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.

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

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. *American Economic Review 75*(3), 424–440.

Katz, M. L. and C. Shapiro (1986). Technology adoption in the presence of network externalities. *Journal of Political Economy 94*(4), 822–841.

Krishnan, D. and S. Siegel (2017). Effects of demonetization: Evidence from 28 slum neighborhoods in mumbai.

Lahiri, A. (2020). The great indian demonetization. *Journal of Economic Perspectives 34*(1), 55–74.

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

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies 60*(3), 531–542.

Mas, I. and O. Morawczynski (2009). Designing mobile money services lessons from m-pesa. *Innovations: Technology, Governance, Globalization 4*(2), 77–91.

Matsuyama, K. (1995). Complementarities and cumulative processes in models of monopolistic competition. *Journal of Economic Literature 33*(2), 701–729.

Michaelides, A. and S. Ng (2000). Estimating the rational expectations model of speculative storage: A monte carlo comparison of three simulation estimators. *Journal of Econometrics 96*(2), 231–266.

Michalopoulos, S. (2012). The Origins of Ethnolinguistic Diversity. *American Economic Review 102*(4), 1508–39.

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.

Murphy, K. M., A. Shleifer, and R. W. Vishny (1989). Industrialization and the big push. *Journal of Political Economy 97*(5), 1003–1026.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, 1417–1426.

Pakes, A. and D. Pollard (1989). Simulation and the asymptotics of optimization estimators. *Econometrica*, 1027–1057.

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.

Rochet, J.-C. and J. Tirole (2006). Two-sided markets: a progress report. *The RAND Journal of Economics 37*(3), 645–667.

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

Rosenstein-Rodan, P. N. (1943). Problems of industrialisation of eastern and south-eastern europe. *The Economic Journal 53*(210/211), 202–211.

Rust, J. (1994). Structural estimation of markov decision processes. *Handbook of Econometrics 4*, 3081–3143.

Ryan, S. P. and C. Tucker (2012). Heterogeneity and the dynamics of technology adoption. *Quantitative Marketing and Economics 10*(1), 63–109.

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

Rysman, M. (2019). The reflection problem in network effect estimation. *Journal of Economics & Management Strategy 28*(1), 153–158.

Sakovics, J. and J. Steiner (2012). Who matters in coordination problems? *American Economic Review 102*(7), 3439–61.

Saloner, G., A. Shepard, et al. (1995). Adoption of technologies with network effects: An empirical examination of the adoption of teller machines. *RAND Journal of Economics 26*(3), 479–501.

Schaner, S. (2017). The cost of convenience? transaction costs, bargaining power, and savings account use in kenya. *Journal of Human Resources 52*(4), 919–945.

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.

Taylor, L. A. (2010). Why are ceos rarely fired? evidence from structural estimation. *Journal of Finance 65*(6), 2051–2087.

Tucker, C. (2008). Identifying formal and informal influence in technology adoption with network externalities. *Management Science 54*(12), 2024–2038.

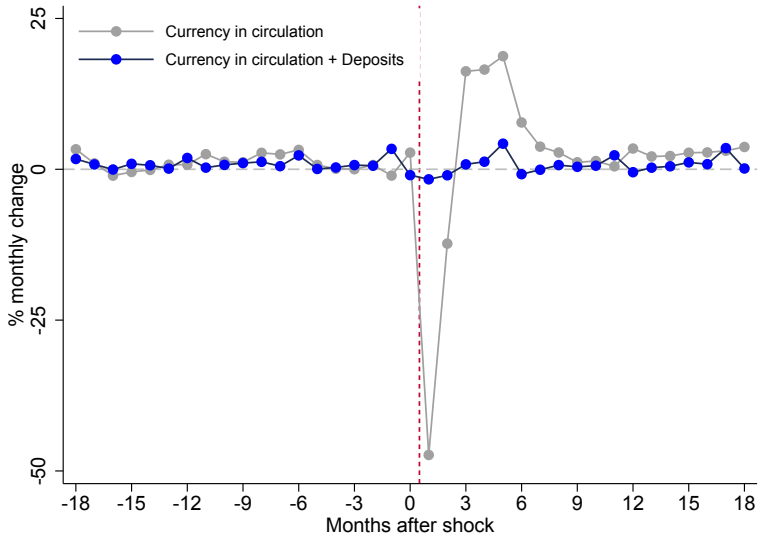
Weyl, E. G. (2010). A price theory of multi-sided platforms. *American Economic Review 100*(4), 1642–72.

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

Young, H. P. (2009). Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning. *American Economic Review 99*(5), 1899–1924.

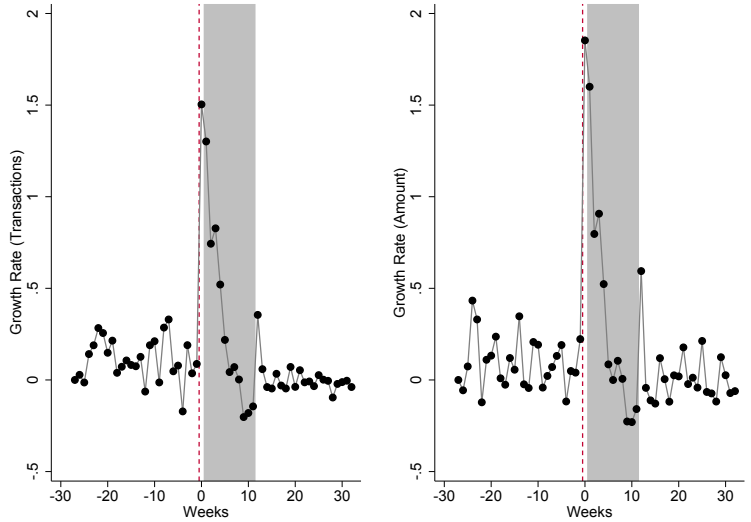
Figures and tables

Figure 2: Change in nominal value of currency in circulation



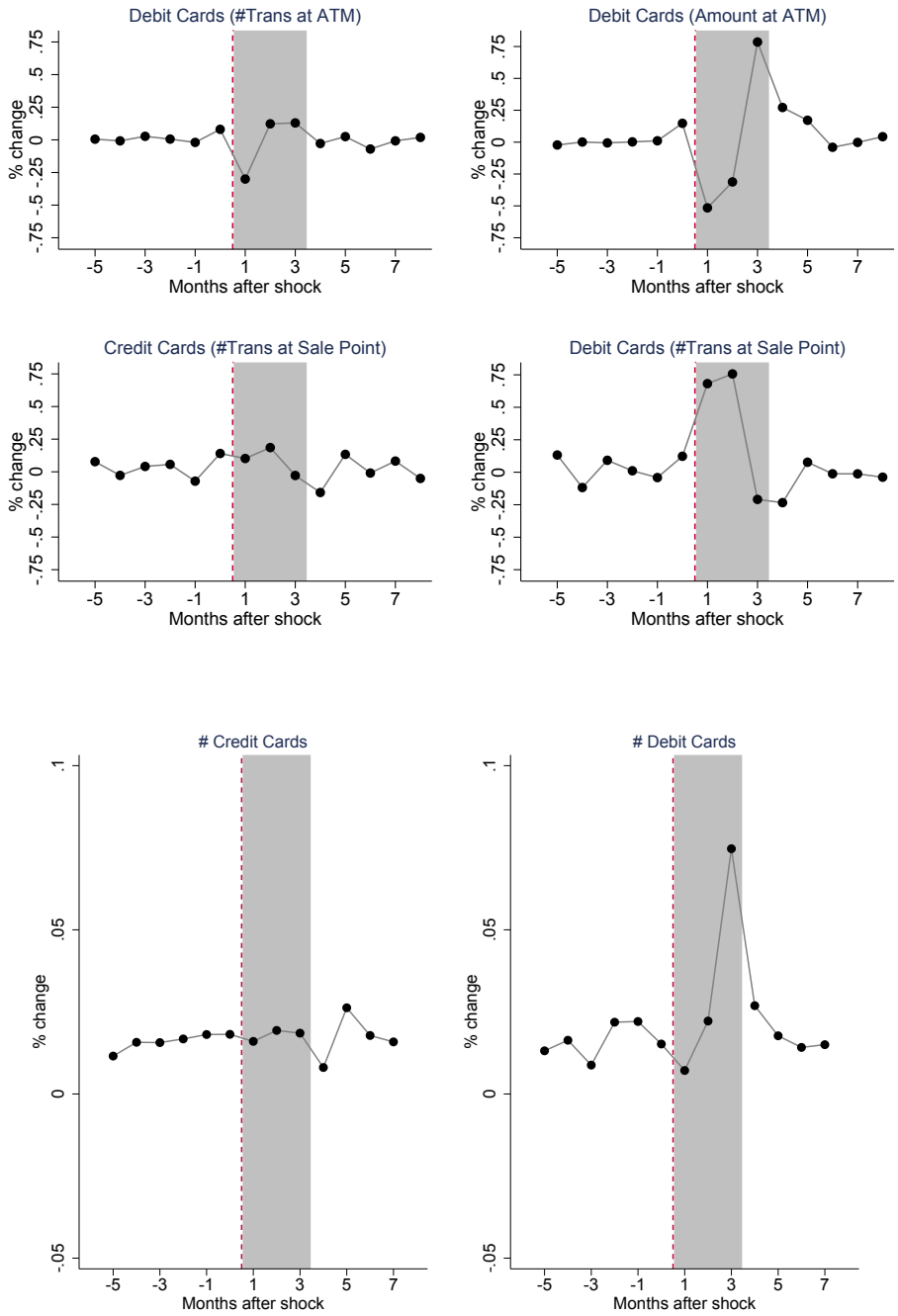
Notes: The figure shows the change in the nominal value of the stock of currency in circulation (in grey) and change in the value of the total money supply (in blue) in India. Month 0 is the month of October 2016; the figures are end-of-month estimates. Source: Reserve Bank of India.

Figure 3: Amount and Transaction Growth on Mobile Payment Platform



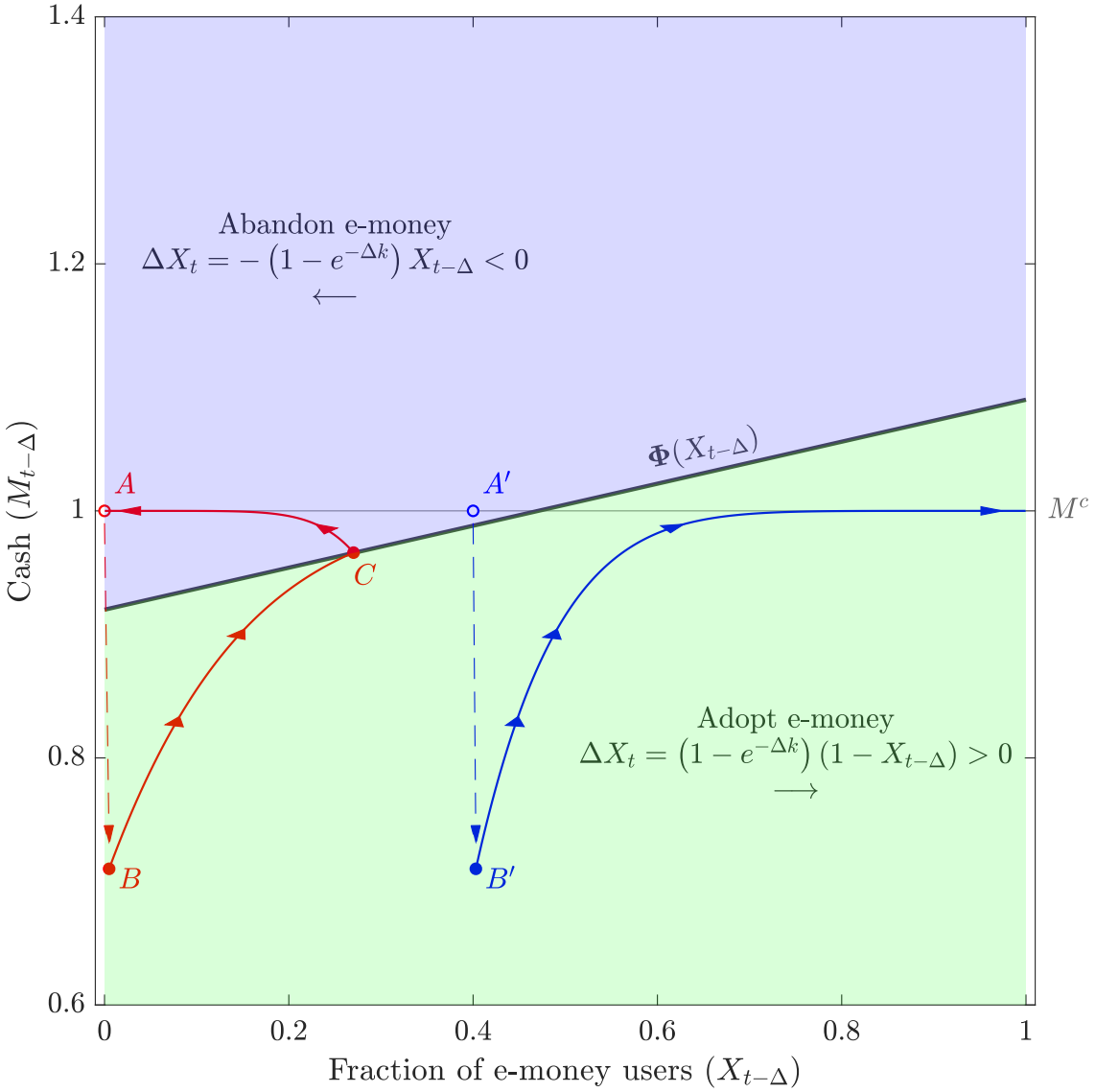
Notes: Week-over-week growth rate in the number of transactions (left panel) and total amounts (right panel) on the electronic wallet platform. The dashed red line indicates the week of November 8th, 2016. Data is discussed in Appendix Section A.3.

Figure 4: Changes across alternate electronic payment systems



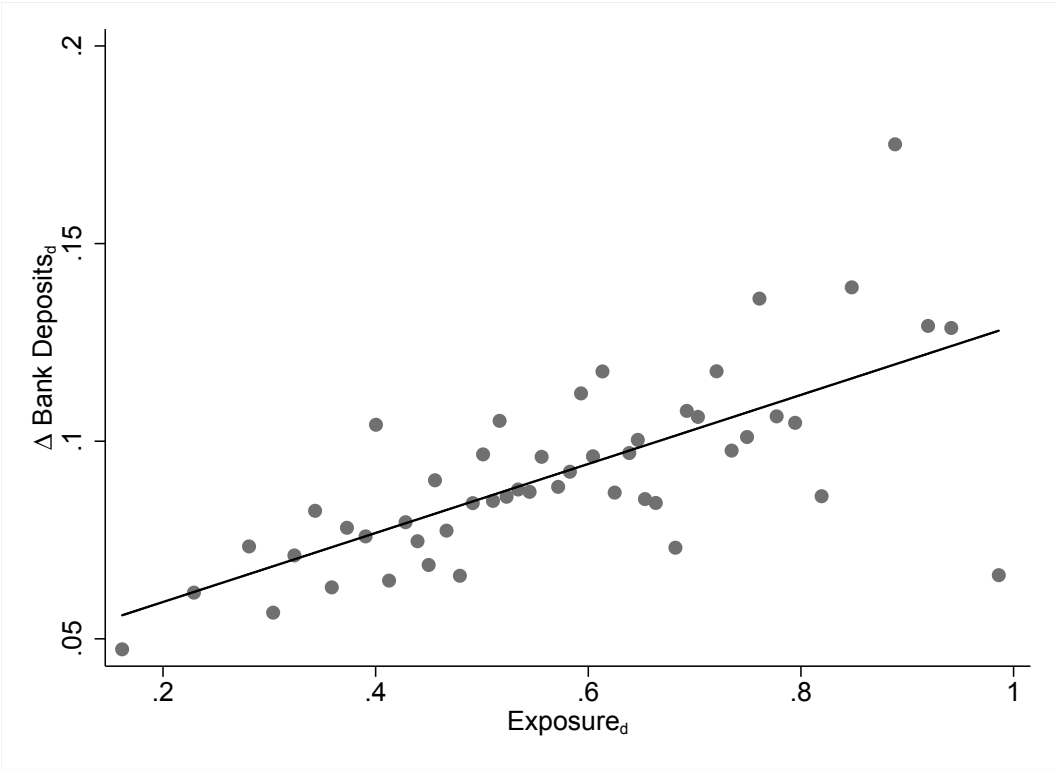
Notes: Change in the use of other electronic payment systems for credit cards and debit cards around the period of the shock. The top panel reports the measures of intensive margin use, and the bottom panel reports the measures of adoption. All the data are monthly and aggregated at the national level. The x-axis represents the month, where October 2016 is normalized to be zero. Source: Reserve Bank of India.

Figure 5: Adoption dynamics in the model with complementarities ($C > 0$ and $\kappa = 0$).



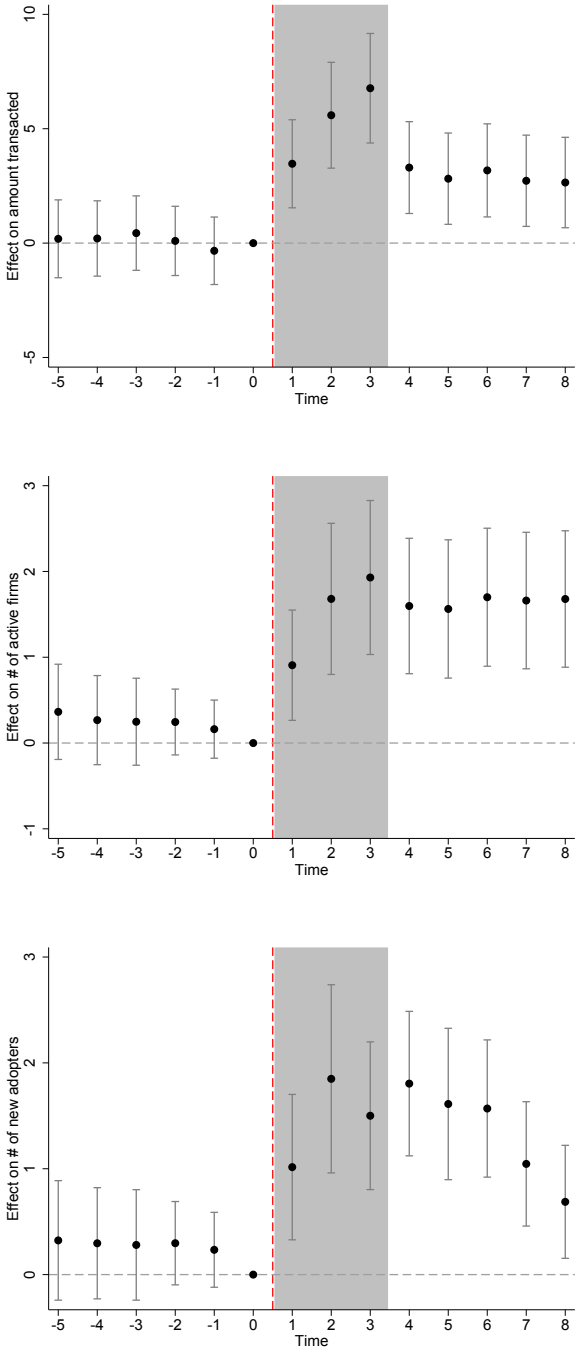
Notes: The model illustrated here corresponds to the case $\theta > k$ (the shock is transitory relative to the adjustment speed of firms). The red line shows the path of a district that starts with a low adoption level $X_{d,0} = 0$. The blue line shows the path of a district that starts with a high 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 6: Relation between Exposure and 2016 Q4 deposit growth



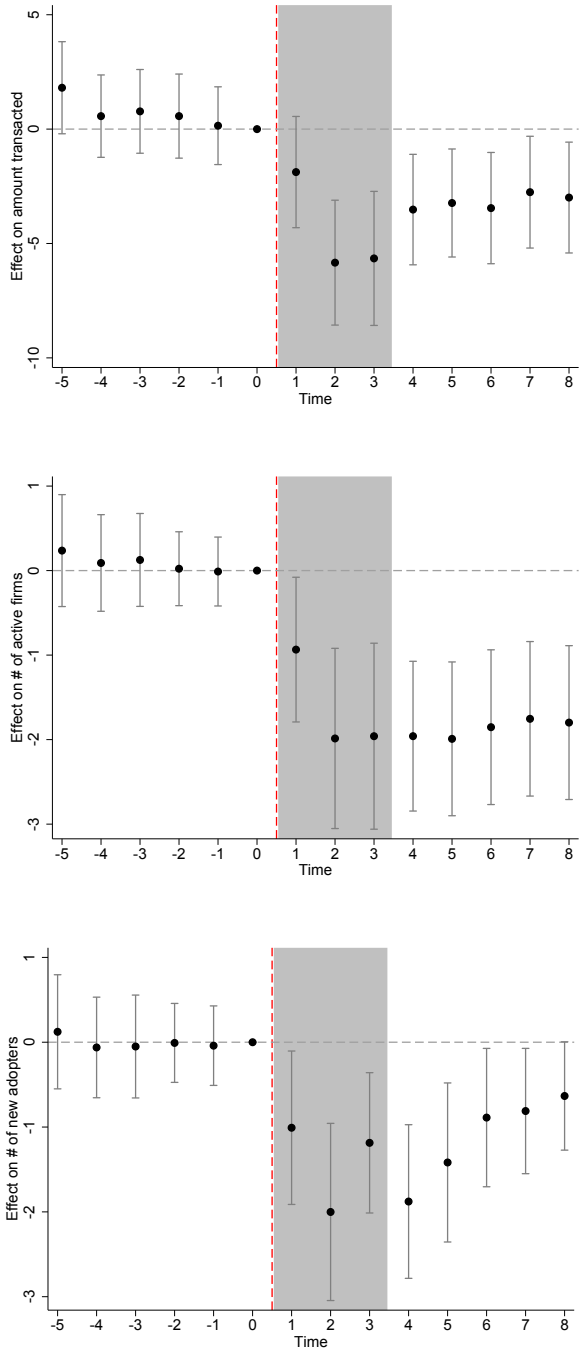
Notes: The figure shows the relation between our measure of Exposure$_d$ (as described in Section 4) and the change in bank deposits in the district between September 30, 2016 and December 31, 2016 *i.e.* during the quarter of demonetization. Source: Reserve Bank of India.

Figure 7: District adoption dynamics in electronic payments data based on exposure to shock



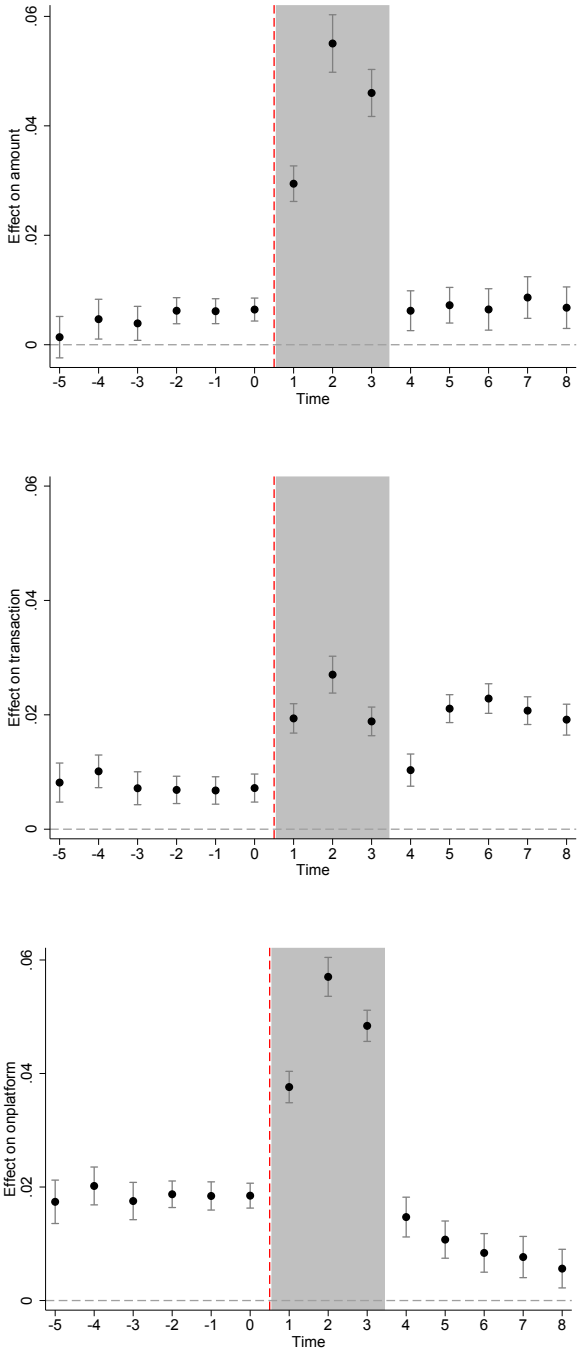
Notes: The figure plots the dynamic treatment effects of the demonetization shock on technology adoption of electronic payment systems. The graphs report the coefficients δ_t from specification 5; the top panel reports the effects for the total amount of transactions (in logs), the middle panel reports the effects for the total number of active firms on the platform (in logs), and the bottom panel reports the effect for the total number of new firms on the platform (in logs). The x-axis represents the 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 8: District adoption dynamics in electronic payments data based on distance to electronic hub



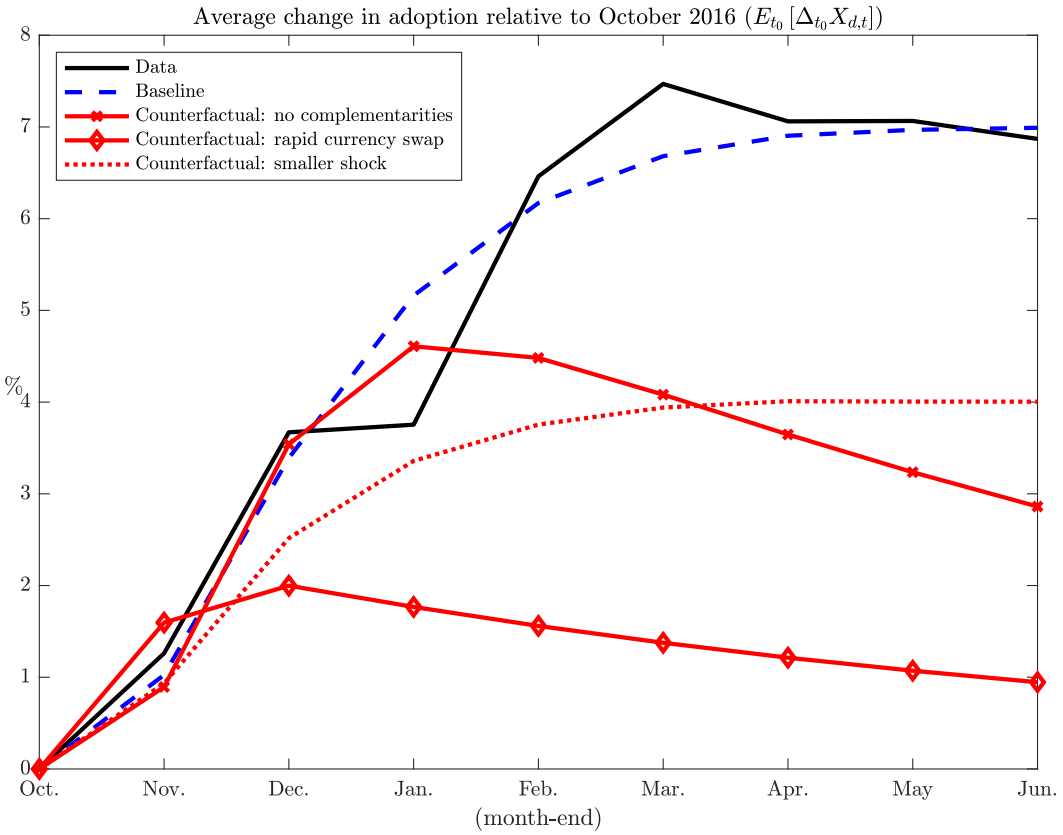
Notes: The figure plots the dynamic effects of adoption across districts based on a district's initial adoption rates as proxied by the distance of that district to the closest district with more than 500 active firms before demonetization. The specification we estimate δ_t in the dynamic version of equation 7. The top panel reports the effects for the total amount of transactions (in logs), the middle panel looks at the total number of firms, while the bottom panel reports the effects for the 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 9: Firm adoption dynamics in electronic payments data based on existing 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 8 in which each coefficient is interacted with a weekly dummy; we reported the monthly estimates of the coefficient γ. The top panel reports the effects when x is the total amount of transactions, the middle panel reports the effects when x is the total number of transactions, and the bottom panel reports the 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 10: Counterfactual paths of average adoption rates across districts.



Notes: The black solid line reports the empirical change in average adoption rates across districts. The other lines report average changes in adoption rates constructed using $S = 100$ simulations from the model, each of a dataset of the same size as the actual data. The dashed blue line is the change in adoption rate obtained from the model evaluated at the point estimates reported in table 6. The solid crossed red line is the average change in adoption rate in the absence of complementarities, assuming that the switching frontier (which is flat without externalities) has the same level as the switching frontier with externalities when adoption is 0. The solid diamond red line is the change in adoption rate when $\theta = 4.6$, corresponding to a 90% decay time of 15 days. The dotted red line is the change in adoption rate when the shock has half the initial size as estimated in table 6.

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 the district-characteristics. The treatment variables is our measure of Exposure$_d$ as described in Section 4. 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 an ATM. Robust standard errors are reported in parentheses. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 3: Exposure$_d$ and adoption of digital wallet

	log(amount)	log(# users)	log(# switchers)
	(1)	(2)	(3)
(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 $\times$ 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 (5). 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 a 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 errors clustered at the district level are reported in parentheses. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 4: District adoption rate of digital wallet based on distance to the hubs

	log(amount)		log(# users)		log(# switchers)	
	(1)	(2)	(3)	(4)	(5)	(6)
(Distance to hub)$_d$ $\times$ $\mathbf{1}(t \geq t_0)$	-5.098***	-3.958***	-2.233***	-1.724***	-1.613***	-1.100***
	[0.936]	[1.190]	[0.468]	[0.497]	[0.361]	[0.387]
Observations	6,846	6,846	6,846	6,846	6,846	6,846
R-squared	0.852	0.886	0.871	0.912	0.821	0.871
District Controls $\times$ Month f.e.	✓	✓	✓	✓	✓	✓
State $\times$ Month f.e.		✓		✓		✓

Notes: Difference-in-differences estimate of the effect of initial conditions, using the distance to the nearest hub (defined as districts with more than 500 retailers in September 2016) as a proxy for the initial share of adopters. The specification estimated is equation 7. In Columns (1) and (2), the dependent variable is the log of the total amount (in Rs.) of transactions carried out using a digital wallet in district d during month t; in Columns (3) and (4), the dependent variable is the log of the total number of active retailers using a 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, level of population and distance to state capital. Standard errors clustered at district level are reported in parentheses. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 5: 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.		✓	✓	✓
Pincode × Week F.E.			✓	✓
Industry × 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 8 in which each coefficient is interacted with a weekly dummy; we reported the estimates of the coefficient γ. The top panel reports effects when x is the total value of the transactions, the middle panel reports the effects when x is the total number of transactions, and the bottom panel reports the effects when x is a dummy for whether the firm used the platform. Standard errors clustered at pincode level are reported in parentheses. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table 6: Point estimates and standard deviations for $\hat{\Theta}$.

Parameter		**Estimate**	**Standard error**
S	Size of aggregate shock	0.246	(0.047)
C	Adoption complementarities	0.063	(0.004)
k	Speed of technology adjustment	0.163	(0.041)
σ	Volatility of idiosyncratic innovations	0.039	(0.011)
M^e	Returns to electronic payments when $X_{d,t} = 0$	0.970	(0.004)

Notes: The parameters are estimated on a balanced panel with 512 districts and 8 months. The estimation procedure uses the simulated method of moments and is described in section 5. Standard errors are reported in parenthesis; they are computed using the bootstrap described in Appendix D.

Table 7: Model fit for the SMM estimation.

Moment		Emp. value	Sim. value	Std. error	p-value
$\hat{\beta}$	Short-run average effect	0.030	0.032	0.004	0.32
$\hat{\gamma}$	Medium-run average effect	0.038	0.035	0.003	0.06
$\hat{\delta}$	Short-run effect of initial adoption	0.081	0.080	0.005	0.40
$\hat{\zeta}$	Medium-run effect of initial adoption	0.027	0.007	0.004	0.00
$\hat{\xi}$	Mean squared residuals	0.083	0.093	0.004	0.02
$\hat{\eta}$	Short-run between-district variance	0.098	0.096	0.004	0.26
$\hat{\kappa}$	Medium-run between-district variance	0.102	0.092	0.007	0.08
$\hat{\xi}$	Within-district variance	0.045	0.050	0.003	0.00
OID statistic	**Degrees of freedom**	**p-value**			
4.7076	3	0.1945			

Notes: The second column shows the empirical values of the moments used in the estimation of the model, and described in section 5. The simulated values are computed using the point estimates reported in table 6. We simulate 2000 panels consisting of 512 districts, and sample data from each panel at the monthly frequency. We then use each panel to compute the moments described in equation (9) and used in the estimation of the model. The standard error reported is the simulated sample standard error. The p-values reported for each moment are one-sided: they are the fraction of observations for which the simulated moment is at least as far from the average simulated moment as the empirical moment is. In the estimation procedure, we use the square root of all second order moments; the table above reports these standard errors and not the variance. More details on the estimation procedure are reported in Appendix D.

Table 8: Alternative interventions.

	Baseline	Alternative interventions			
		$g = 0$	$g = 0.10$	$g = 0.25$	$g = 0.5$
Shock size (p.p.)	24.56	21.00	18.31	16.81	14.22
Shock half-life (months)	0.82	1.05	1.29	1.37	1.55
$\mathbb{E}_{t_0}\left[\Delta_{t_0} X_{d,t_0+T}\right]$ (p.p.)	7.22	8.34	8.32	8.23	7.82
$sd_{t_0}\left[\Delta_{t_0} X_{d,t_0+T}\right]$ (p.p.)	26.42	36.66	34.64	28.67	24.81

Notes: The column marked "Baseline" report the estimated shock size, the shock half-life, and the mean and standard deviation of long-run changes in average adoption rates; we use $T = 3$ years and $s = 100$ simulations to compute these moments. The other columns report these moments under alternative scenarios. For each value of g — the aversion to dispersion in the planner's objective function — we compute the value of the shock size and persistence which maximizes the objective described in equation (11).

Online Appendix for "Shocks and Technology Adoption: Evidence from Electronic Payment Systems"

This version: November, 2020

A Institutional background

A.1 The impact of the Demonetization

This Section provides some extra background on the Demonetization. This information complements the discussion that is provided in the main body of the paper.

As discussed before, the announcement of the Demonetization on November 8 2016 voided automatically about 86.4% of the total value of currency in circulation. Even if you had until the end of the year to deposit the old notes in the banking sector, the voided bills could not be used immediately after the announcement. At the same time, the new notes were not available right away, as the Government in November had not even finished printing all the necessary bills. Combining these two things together, India found itself with a shortage of currency in cash overnight.

Evidence of the scarcity of cash is abundant during this period. One way to see this issue is to focus on the disruption that characterized banks' operation during this period. In a survey of 214 households in 28 slums in the city of Mumbai, 88% of 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% of households reported waiting for more than 4 hours (Krishnan and Siegel, 2017). Another randomized survey conducted over nine districts in India 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.[66] This evidence confirms the presence of a large unmet demand for cash during the aftermath of the Demonetization.

For consumers, the generalized scarcity of cash was made potentially worse by the constraints on cash withdrawal that were put into place by the Government. For instance, 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 at Rs.20,000 per week, and ATM withdrawals were capped at Rs.4,000 per card per day, for the days following the announcement. However, 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 the withdrawal limits and exceptions, and argue that this uncertainty may have exacerbated the overall confusion during this transition period.

Despite its magnitude, the cash crunch was a temporary phenomenon. Overall, things significantly improved in January and essentially normalized in February. The cash in circulation grew significantly again in January 2017, suggesting that the public was able to withdraw cash from banks (see Figure 2). This evidence is consistent with the idea that banks in January had a sufficient inventory of new bills to meet the transactional demand for cash in the country. Indeed, by January 30th, 2017, the Government lifted most of the remaining limitations on cash withdrawals, in particular removing any ATM withdrawal limit from current accounts. In practice, limits started to be progressively relaxed after the announcement, as banks' reserve started to receive the new bills. After January, the only limitation left was on withdrawal from saving accounts. However, this limitation was relatively high - i.e. raised in February 2017 to Rs.50,000 per week -, not particularly binding (you could move money out saving account to withdraw). Moreover, by mid-March 2017, all limits on withdrawals had been removed.

This narrative on the timing of the Demonetization — large disruption in the short run, with relative normalization starting around February — is also consistent with data on internet searches, which may help us to elicit individuals' perception. In particular, figure E.1 reports the monthly plot (09/2016 to 07/2017) of Google searches for several key words that could be associated with the shock. For instance, we collect data on searches on the words "Cash" or "ATM line," and others. Data is obtained by Google Trends, and the index is normalized by Google to be from 0 to 100, with the value of 100 assigned to the day with the maximum number of searches made on that topic. Across all the panels, we find that Google searches that are related to 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 reached its 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. Altogether, this information is consistent with the relatively short-lived nature of the shock.

[66] https://economictimes.indiatimes.com/news/politics-and-nation/how-delhi-lost-a-working-day-to-demonetisation/articleshow/56041967.cms

A.2 The Demonetization and the policy response

As discussed in the paper, the November Demonetization led to a large contraction in cash. While the initial objective of the governments was not to foster a shift towards electronic payments, the increase in electronic payments that characterized the first weeks did not pass unnoticed. As a result, the Government decided to actively support the increase in electronic payments in the months following the Demonetization.

On top of generic announcements from top politicians, the Government and the RBI put into place several interventions in the area of electric payments in the months following the Demonetization. First, the Government actively supported the adoption of traditional electronic payment technologies by trying to lower the adoption costs of point-of-sales (POS) system, in particular for small businesses. One example of this type of program was the grants that were provided by the National Bank for Agriculture and Rural Development (NABARD) to support the acquisition of POS machines in small villages.[67] Second, the Government partnered with several other organizations to provide discounts on their products when payments are made electronically. The main discounts involved patrol and railroads. For instance, the Government partnered with Indian Oil Corp, Bharat Petroleum, and Hindustan Petroleum to give a 0.75% discount to the consumers if they paid electronically. For railroads, the incentives ranged from a small discount on ticket acquired with electronic payments – generally up to 0.5% - to free accidental insurance for travelers.[68]

This government response is important for us because it suggests that an aggregate event study may not be sufficient to understand the effects over the medium-run of the temporary shock. While the immediate response is still likely going to capture only the effect of the cash contractions, the response over time is going to capture a combination of the persistence of the temporary shock and the effect of the Government's policies rolled over in the meantime.

This motivated us to conduct most of our analyses at the disaggregated level, exploiting heterogeneity across districts in the exposure to the temporary cash contraction. Relative to the event study, this approach requires much weaker identification assumptions, since this specification differences out any aggregate change in policy during our period. In fact, at best of our knowledge, all the policies that were introduced by the Government were aggregate in nature and therefore they did not target specifically any district or subset of the country. From a detailed analyses of the policy changes, we found no evidence that any intervention was designed formally or informally to specifically target areas that were more affected by the cash contraction, which is what would be the issue in our case. In a way, for most policies the inability to target certain areas is intuitive (e.g. discount for railways were available from every city, not only those we categorize as highly affected). In general, the only policy that may have targeted a specific area is the provision of subsidy on POS. In this regard, it is important to point out that our setting — electronic wallet technology — did not require any POS. Furtermore, we found no evidence of targeting of POS subsidies in the law or media coverage.

Overall, our disaggregate approach seems to be well suited to examine the impact of the cash contraction, conditional on aggregate changes. In the empirical section, we discuss more carefully the identification assumptions for this model, for instance highlighting also the importance of district level controls for identification.

A.3 Electronic Wallet Technology

The main focus of the paper is on the adoption of one specific type of electronic payment option, which is the electronic wallet technology. This section provides further details on this technology, which are important to understand the context of the paper.

Our data comes from a company that – at the time of the Demonetization – was the largest player in the provision of electronic wallet payment services and the main fintech company active in India. The company allows individuals and businesses to undertake transactions with each other using only their mobile phone. To use the service, a customer would normally need to download an application and link their bank account to the application. However, in 2016 the company also established a new service that allows customers to make payments without the need of internet or a smart phone.

[67] https://indianexpress.com/article/business/business-others/demonetisation-nabard-cashless-economy-debit-cards-4417938/

[68] https://economictimes.indiatimes.com/industry/transportation/railways/soon-you-will-be-rewarded-for-cashless-booking-of-railway-tickets/articleshow/61954004.cms?from=mdr

To be more specific, in our context there are multiple ways to transact using the digital wallet. First, customers can scan the merchants' unique QR code in the application installed on their smartphones to complete the transaction. Second, instead of scanning the QR code, customers can enter the mobile number of the merchant. In this case, the merchant would receive a unique code from the company, which is then used by the customer to complete the transaction. Third, if a smartphone or mobile internet are not available, customers can call a toll-free number and ask the wallet company to make a transaction using the cell-phone number of the merchant. To use this feature, customers needed to be enrolled through a one-time verification process.

Money can then be transferred from the electronic wallet to a traditional bank account. Therefore, the technology is in many respect similar to a credit card or other more traditional electronic payment systems. However, relative to these other electronic payment technologies, adoption costs are much lower, since merchants and consumers can access the electronic wallet almost instantaneously, without the need of anything more than a phone and a bank account. In particular, from the standpoint of the retailers, the mobile wallet does not require the acquisition of a POS.

On top of fixed cost, variable costs of this technology are also very limited, in particular for small merchants. Merchants using the digital wallet are classified by the provider into three segments: small, medium and large. Small merchants have lower limits on the 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 the transfer amount as a fees. Our data only covers small and medium merchants. From our discussion with the company, large merchants tend to have more personalized contract, which could bundle different services and payment options together.

B Appendix to Section 3

B.1 Derivations

B.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:

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

Here, $V_R(x_{i,t}, M_t, X_{t-\Delta t})$ denotes the value of a firm that receives the option to revise its technological choice early on in period $t+\Delta t$ (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, Δt. So it should be interpreted in units of firms value.) Denote by:

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

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 (2).

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

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

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

The net benefits of adoption can be written as:

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

We need to compute:

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

Finally, we need to compute:

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

The dynamics of the adopter share are:

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

Thus we have:

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

Moreover,

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

So:

$$
\begin{aligned}
PVX_{t-\Delta t} &= \frac{1-e^{-k\Delta t}}{1-e^{-(r+2k)\Delta t}}\sum_{n=0}^{+\infty} e^{-(r+k)\Delta tn}\mathbb{E}_{t-\Delta t}\left[a_{t+\Delta tn}\Delta t\right] + \frac{\Delta t e^{-k\Delta t}}{1-e^{-(r+2k)\Delta t}} X_{t-\Delta t} \\
&= \frac{1-e^{-k\Delta t}}{1-e^{-(r+2k)\Delta t}} PVA_{t-\Delta t} + \frac{\Delta t e^{-k\Delta t}}{1-e^{-(r+2k)\Delta t}} X_{t-\Delta t},
\end{aligned} \tag{12}
$$

where:

$$
PVA_{t-\Delta t} = \sum_{n=0}^{+\infty} e^{-(r+k)\Delta tn}\mathbb{E}_{t-\Delta t}\left[a_{t+\Delta tn}\Delta t\right]. \tag{13}
$$

Therefore,

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

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 $\Phi(.)$.

B.2 Microfoundations

This appendix describes a version of the model with extended microfoundations. Relative to the baseline model, the model described here has two additional features. First, firms that have adopted the electronic payments technology can still accept payments in cash, so that the electronic payments technology is an add-on, not an alternative to cash. Second, the choice of consumers between cash and electronic payments is explicitly modelled. The main point of this appendix is that the model with extended microfoundations is isomorphic to the simpler model studied in the main text.

Consumers There is a continuum of mass 1 of identical households. Each period, households randomly meet with firms. Each household holds D units of deposits, where D is exogenous and fixed. Deposits can be used for payment in retail transactions, either by converting them to cash or by using them in electronic payments. Households can only withdraw up to L_t units of cash, where L_t is exogenous. Finally, they behave myopically: each period, after observing the number of firms that accept electronic payments,

$X_t \equiv \int_{i\in[0,1]} \mathbf{1}\{x_{i,t} = e\}\, di \in [0,1]$, they solve the following problem:

$$\max_{C_t^c, C_t^e, L_t^c, L_t^e} \quad X_t\left(\zeta C_t^e + (1-\zeta)C_t^c\right) + (1-X_t)C_t^c - \frac{1}{2\gamma}\left(\frac{L_t^e - L^e}{P_t}\right)^2$$
$$\begin{aligned} \text{s.t.} \quad L_t^c + L_t^e &\leq D \quad [\lambda_t] \\ L_t^c &\leq L_t \quad [\mu_t] \\ P_t C_t^c &\leq L_t^c \quad [\nu_t^c] \\ P_t C_t^e &\leq L_t^e \quad [\nu_t^e] \end{aligned}$$

Because meetings are random, the probability that a household meets a firm that accepts both electronic payments and cash is X_t. Upon meeting, the household and the firm decide on which means of payment to use in order to conduct the transaction. We assume that electronic money is chosen with probability ζ, and cash is chosen otherwise; the probability ζ is exogenous, constant, and strictly positive. Meeting a firm that accepts both electronic payments and cash thus yields expected utility $\zeta C_t^e + (1-\zeta)C_t^c$ to the household. If the household instead meets a firm that only accepts cash, the meeting yields utility C_t^c.

Additionally, there are quadratic utility costs associated with holding real balances of electronic means of payment away from an exogenous level L_e. Here, L_e could be arbitrarily small. This cost is non-pecuniary: it is a shorthand for modeling cognitive or, in this static framework, opportunity costs of adjusting real balances of electronic money. Finally, the household's problem is subject to two constraints that state that consumption using either type of payment cannot exceed real balances of each type.[69] We assume that prices of consumption goods are constant, and normalize them to $P_t = 1$. Eliminating the multipliers ν_t^c and ν_t^e, the necessary first-order conditions for optimality for this problem can be written as:

$$\begin{aligned} &\lambda_t + \frac{1}{\gamma}\left(L_t^e - L^e\right) = \zeta X_t \\ &\lambda_t + \nu_t = 1 - X_t + (1-\zeta)X_t \end{aligned} \tag{14}$$

along with four complementary-slackness conditions, $\lambda_t\left(D - L_t^c - L_t^e\right) = 0$, $\mu_t\left(L_t - L_t^c\right) = 0$, $\nu_t^c\left(L_t^c - C_t^c\right) = 0$, and $\nu_t^e(L_t^e - C_t^e) = 0$. The two state variables of the household's problem are X_t and L_t.

Firms The problem of each firm is identical to that described in Section 3, except for the definition of flow profits of each firms. Namely, we now assume that profits are now given by:

$$\Pi(x_{i,t}, C_t^c, C_t^e) = \begin{cases} (\mu-1)\left(\zeta C_t^e + (1-\zeta)C_t^c\right) & \text{if } x_{i,t} = e, \\ (\mu-1)C_t^c & \text{if } x_{i,t} = c. \end{cases}$$

where $\mu > 1$ is a constant markup over marginal cost. Each period, the firm meets a different household. If the firm has adopted electronic payments ($x_{i,t} = e$), its expected revenue is $\zeta C_t^e + (1-\zeta)C_t^c$. Otherwise, its revenue is C_t^c. The rest of the firms' problem is identical. Following the same steps as in the main text, net adoption benefits follow:

$$B_{t-\Delta t} = \mathbb{E}_{t-\Delta t}\left[(\mu-1)\zeta\left(C_t^e - C_t^c\right)\Delta t + e^{-(r+k)\Delta t}B_t + e^{-r\Delta t}(1-e^{-k\Delta t})g(B_t)\right], \tag{15}$$

implying that the state variables in each individual firm's problem are now $(x_{i,t-\Delta t}, C_{t-\Delta t}^e, C_{t-\Delta t}^c)$. The adoption rule, equation (3), and the law of motion for the number of adopters, equation (2), are defined in the same way as in the model in the main text.

Equilibrium There is a unique exogenous stochastic process in this model, L_t, the dynamics of which we leave unspecified for now. An equilibrium of this model is a technology choice rule, x, mapping $\{c,e\}\times\mathbb{R}^2 \to \{c,e\}$, a function for the gross adoption benefit, B, mapping $\mathbb{R}^3 \to \mathbb{R}$, and household choice rules C^c, C^e, L^c, L^e and their associated Lagrange multipliers, mapping $\mathbb{R}^2 \to \mathbb{R}$. These objects must be such that, given the processes for L_t, the technology choice rule and the gross adoption benefit solve the system of equations

[69] Because of the static nature of the household's problem, these are not, strictly speaking, "cash in advance" constraints.

(15) and (3) when X_t follows the law of motion given by (4), and the household choice rules satisfy the first order conditions (14), along with the associated complementary slackness conditions.

Isomorphism to baseline model Next, we show that the baseline model of the main is a particular case of the microfounded model described above. Specifically, we assume that deposits, D, are large relative to both cash in circulation and to potential demand for electronic payments $D \geq L_t + C^e + \gamma\zeta$ In this case, the unique equilibrium has the following features. First, $\lambda_t = 0$, since the deposit constraint is slack when deposits are sufficiently high. Second, when $X_t > 0$, the constraint $L_t^e = C_t^e$ binds, so that:

$$C_t^e = L^e + \gamma\zeta X_t.$$

Moreover, $\mu_t^c = \nu_t = 1 - X_t + (1-\zeta)X_t > 0$, so that $C_t^c = L_t$. Additionally, when $X_t = 0$, the solution is $C_t^e = C^e$ and $C_t^c = L_t$. The flow benefits of adoption are then given by:

$$\Pi_t^e - \Pi_t^c = (\mu - 1)\zeta\left(C_t^e - C_t^c\right) = (\mu-1)\zeta\left(L^e + \gamma\zeta X_t - L_t\right)$$

Thus, the microfounded model produces identical dynamics to the model in the main text so long as:

$$C = (\mu-1)\gamma\zeta^2, \quad M^e = (\mu-1)\zeta L^e, \quad M_t = (\mu-1)\zeta L_t.$$

where C, M^e and M_t are the exogenous parameters and processes described in Section 3. In this version of the model, the reduced-form parameter governing externalities, $C = (\mu-1)\gamma\zeta^2$, is large either when the slope of adjustment costs for electronic money, which is given by $1/\gamma$, is high (so that households adjust their holdings of electronic money rapidly in response to changes in X_t), or ζ is high, so that when a match between households using e-money and firms accepting it occurs, e-money is likely to be the medium of exchange chosen.

B.3 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:

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

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

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

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

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, $M_{t-\Delta t} \leq \underline{\Phi}(X_{t-\Delta t})$. For adoption of electronic money to be a strictly dominant strategy it must be that $B_{t-\Delta t} \geq 0$ even if the firm expects no adoption at all by other firms, so that $PVA_{t-\Delta t} = 0$. In that case:

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

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

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

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

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

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:

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

  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 $PVA(M_{t-\Delta t}, X_{t-\Delta t}; \Phi)$, 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 $M_{t-\Delta t}$, and the implied path for $X_{t-\Delta t}$ 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.

B.4 Predictions across versions of the model

We next discuss the qualitative predictions of the model emphasized in the main text, using Monte-Carlo simulations of the response of a large number of districts to an aggregate shock to cash-based demand M_t.

B.4.1 The cash crunch in the model with complementarities ($C > 0$ and $\kappa = 0$)

We first provide an illustration of the quantitative properties of the model with complementarities, 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:

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

The average path of cash is reported in Figure E.2. Districts are otherwise identical, save for their initial conditions $(M_{-\Delta tt,d}, X_d)$, which reflect the ergodic distribution of the model prior to the shock. The calibration of the model is otherwise that reported in table 6.

Figure E.3 shows the average response across districts. The number of firms using e increases permanently (left panel of Figure E.3). Moreover, the likelihood of switching also increases permanently (right panel of Figure E.3).

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 Figure E.4) 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 E.4 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 result highlights the broader idea that long-run adoption dynamics are determined by the initial strength of complementarities.

B.5 The cash crunch in the frictionless model ($C = 0$ and $\kappa = 0$)

The left panel of figure E.11 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 $X_{-\Delta t} = 0$ and $M_{-\Delta t} = M^c$. 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.[70] In the long-run, the district will therefore converge back to point A.

The top row of figure E.12 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 E.2.) 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 E.13.

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, $X_{-\Delta t}$. The left panel of figure E.11 illustrates this, by also showing (in blue) the trajectory of a district starting from $X_{-\Delta t} = 0.4 > 0$. 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 E.13.

B.6 The cash crunch in the model with fixed costs ($C = 0$ and $\kappa > 0$): Monte-Carlo illustration

Figures E.6 and E.7 highlight the differences between the fixed cost and the complementarities model using Monte-Carlo simulations. Figure E.6 reports the average response of the economy to the same shock as above in the model with fixed costs. 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 E.7). However, as reported on the right panel of Figure E.7, 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.

[70] In the absence of complementarities ($C = 0$) or fixed costs ($\kappa = 0$), it is straightforward to see (using equation 2) 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, $M_{t-\Delta t}$. One can then verify that, given the functional forms for flow profits, firms switch from c to e whenever $M_{t-\Delta t} \leq \underline{M} = M^c - \frac{1 - e^{-(r+k+\theta)\Delta t}}{e^{-\theta \Delta t} - e^{-(r+k+\theta)\Delta t}} (M^c - M^e)$. 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.7 Persistence and state-dependence in the model with complementarities ($C > 0$ and $\kappa = 0$)

The discussion in the main text focuses on versions of the model with complementarities in which $\theta > k$, 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* relationship between initial conditions and subsequents increases in the number of users.

The first part of this claim is illustrated in Appendix Figure E.8, which describes 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 converges to 1, and the likelihood of switching also increases, as illustrated in Appendix Figure E.9. Importantly, 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 Appendix Figure E.10 illustrates this further in numerical simulations of the model. 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 ($\theta > k$).

B.8 Firm-level state-dependence: Monte-Carlo illustration

Finally, we discuss the firm-level predictions of the model with respect to state dependence. Specifically, we show that the model predicts a positive relationship between firm-level adoption and the overall existing user base, conditional on the level of cash, only when there are positive complementarities ($C > 0$). To be precise, we simulate data from variants of the model and estimate a firm-level regression of the following form:

$$x_{i,d,t} = \alpha + \rho x_{i,d,t-\Delta} + \beta M_{d,t-\Delta} + \gamma X_{d,t-\Delta} + \epsilon_{d,t} \tag{16}$$

Table E.3 reports the results. 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 E.3). The intuition for this result is simple: an increase in the use of the technology will increase the value of the technology itself, which will in turn positively affect adoption by firms. Importantly, the same relationship does not hold without externalities (Column 1 of Table E.3).

Finally, in Figure E.14 we also report estimates of the coefficient γ_t in the monthly version of the regression above:

$$x_{i,d,t} = \sum_{t'} \left(\alpha_{t'} + \rho_{t'} x_{i,d,t'-\Delta} + \beta_{t'} M_{d,t'-\Delta} + \gamma_{t'} X_{d,t'-\Delta}\right) \mathbf{1}_{\{t=t'\}} + \epsilon_{d,t}. \tag{17}$$

The coefficients are the Monte-Carlo analogs of the point estimates reported in Figure 9. The simulated sample we use is the same as for the the estimation of Model (16) above. As in the data, we find that the effect of the adopter share is highest during the months that immediately follow the shock.

C Cash contraction and Consumption

In this Section, we examine how household consumption responded to the cash swap using the same identification strategy from Section 5. In other words, we compare behaviors across districts that were characterized by different exposure to chest banks before the Demonetization. The objective of this analysis is twofold. First, these tests can provide novel evidence on how the Demonetization affected the real economy. Results from previous sections provide evidence that the Indian Demonetization led to a widespread and persistent rise in electronic payments. Given the size and speed of these responses, a natural question is whether the rise in electronic money was indeed sufficient to shield the real economy from the cash crunch.

Second, as discussed briefly in Section 4, this evidence on consumption is useful because it provides further robustness on the quality of our empirical model to identify the supply side effect of a cash contraction. The intuition for this second aspect is simple. Our tests on electronic payment — in particular the sharp response right around the policy shock — provides very strong evidence regarding the fact that our estimates capture how electronic payment use was affected by the Demonetization. However, the Demonetization could have

affected the use of electronic payments in several ways, and not only because of a contraction in cash (supply shock). For instance, the Demonetization may have increased the overall uncertainty in the economy, which in turn may have reduced consumption. To the extent that our treatment captures this dimension — e.g. highly affected areas are places with lower increase in uncertainty — this alternative explanation may also be explaining our results.

The good news is that consumption response may help separating explanations based on cash contraction from alternative demand-side mechanisms. In particular, a demand side explanation would generally predict that the effects for consumption and electronic payments should go in the same direction. Instead, the opposite results — i.e. highly exposed areas experienced both higher increase in electronic payments and lower consumption — would be hard to rationalize by a demand mechanism, but easy to interpret as a supply side shock. In this sense, exploring the consumption response could provide useful evidence for our mechanism. In terms of robustness, the consumption data has a longer time series than the electronic payments. This will allow us to run several extra tests on the quality of our analysis.

C.1 Empirical setting

In this Section, we examine how household consumption responded to the cash swap using the same identification strategy from Section 4. In other words, we compare behaviors across districts that were characterized by different exposure to chest banks before the Demonetization.

To measure the changes in consumption behavior by Indian households, we use data from the Consumer Pyramids database maintained by the Center for Monitoring Indian Economy (CMIE). This dataset 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.

The data set provides a representative sample of Indian households, where households are selected to be representative of the population across 371 "homogeneous regions" across India. The survey has information on the monetary amount of the household expenditure across different large categories and some other background information on the members of the households. 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. Overall, the data quality is considered high, in particular since CMIE collects the data in person using specialized workers. Each household is interviewed every four months and is asked about their consumption pattern in the preceding four months. Thus, about 39,500 households are surveyed every month.

The data is organized in event-time around the month of the shock. 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. The final sample used in the analysis is constituted by about 95,000 households. We reach this count because we consider households for which the age of the head of household is between 18 and 75 years as of September 2016. To make the panel balance, we also only consider households with non-missing information between June 2016 and March 2017.

The main difference compared with the analyses in Section 4 is the timing. Before, the district-level data were measured at monthly level. For these household data, the survey procedure is such that households belonging to different waves of interviews are asked about the same month at different points in time. 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.[71] This feature is quite common among consumer surveys, and it is similar to the Consumer Expenditure Survey in the US.[72] Following the literature in this area (e.g. Parker et al. (2013)), we deal with this feature by organizing the data by 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.[73]

[71] 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.

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

[73] Therefore, the time in the panel is the one for the wave in which a household was interviewed about November, and it is

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

$$\log\left(y_{h,d,t}\right) = \alpha_t + \alpha_h + \delta_t \left(\text{Exposure}_d \times \mathbf{1}_{\{t \geq t_0\}}\right) + \Gamma_t' Y_{h,d} + \epsilon_{h,d,t}, \tag{18}$$

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, and $Y_{h,d}$ are controls, which 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 analyses along with the addition of household-level controls including the age of the 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.

C.2 Main results

Table E.9 shows the 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 the highly affected district. The effect is sizable: 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, these results are not driven by differences in pre-trends between affected districts (Figure E.19).[74]

Therefore, the cash contraction negatively affected household consumption. However, there are three important things to point out about this negative effect. First, 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 coefficient. 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 only binding between November and January.

Second, consistent with the idea that households were able to partially limit the impact of the shock, the contraction in consumption was larger for items that are less costly to cut for households. As a first step, we divide consumption into necessary and unnecessary items, where the former group contains expenses for food, rent and bills, and utilities (power and gas) while the latter contains the remaining part of the consumption basket. Table E.10 shows that, when consumption is split between the two baskets of goods, the effect on unnecessary consumption was economically larger (about 22% higher).[75]

This last result does not depend on the way we categorize consumption as necessary and unnecessary. In Columns (3)-(5) of Table E.10, 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. 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 in comparison 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.

Third, we also find direct evidence that electronic payments helped to partially limit the impact of the shock. While this evidence confirms that the rise in electronic payment was unable to undo the effects of the cash contraction, it may still be the case that electronic payments helped to partially limit the impact of the shock. To test this hypothesis, we examine the responsiveness to the shock across areas characterized by different levels of penetration of electronic payments in the pre-shock period. In particular, we focus on

zero for the wave that happened four months before the one that includes November 2016 and one for the one that happened four months after.

[74] This analysis shows a positive and borderline significant effect on consumption two quarters after the Demonetization. One interpretation is that households have shifted some consumption to the future. Consistent with this interpretation, we actually find that the effect is driven entirely by unnecessary consumption, which is a category that contains 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).

[75] The same difference also holds when looking using a dichotomous treatment (Appendix Table E.11): here necessary consumption is cut by 4%, while unnecessary consumption by about 8%.

the penetration of debit cards, which we proxy by the number of ATMs per million people in a district.[76] Our focus on traditional electronic payment is motivated by its relative size. In fact, debit cards represent the largest share of electronic transaction in India. Furthermore, while the issuance of new debit cards was overall modest, the Demonetization led to an increase in the amount of transactions, suggesting that debit cards were indeed used as a way to replace cash during the shock period.

The results of this analysis are presented in Table E.12. The key parameter in these regressions is the triple interaction between the time dummies, the measure of exposure to the shock, and a dummy that a value of one for districts that have an above-median number of ATMs per one million people. We repeat the same analysis using both the continuous (odd columns) and dichotomous (even columns) versions of the shock. Looking at total consumption (columns 1 and 2), we find consistently that the effect of the cash contraction was smaller in districts with a high penetration in electronic payments. Depending on the specification, districts with high penetration experienced a contraction in total consumption that is between 60% and 90% smaller than in low penetration areas.[77]

These results show that the cash contraction had a negative effect on individual consumption. However, the negative effects were somehow limited to the most acute period of the Demonetization. Furthermore, the cut was larger for unnecessary goods, like recreational expenses, and much more limited for food expense. Building on these patterns, we also show that the presence of a developed electronic payment infrastructure in a local market explains part of the variation in the response to the shock in the local market. This evidence suggests that – while electronic money was not sufficient to completely shield the economy from the contraction – its presence may have played a role in limiting the costs of the Demonetization.

Furthermore, this evidence is consistent with the interpretation of our specification as correctly capturing heterogeneity on the cash contraction. Consistent with this supply side interpretation, we find that our treatment predicts both lower consumption and higher use of electronic payments.

C.3 Placebo tests

In the body of the paper, we have also mentioned that the longer time series in the consumption data also allows us to run more detailed placebo tests on our treatment measure.

In general, before this test, 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 point in time. 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 month between February 2015 and February 2016. We then replicate our main specification, testing for the presence of a differential response across households in the wave of the placebo shock relative to the previous one.[78]

The results of this set of placebo tests are reported in Figure E.20. The general finding is that — in normal times — 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 the concern that differential exposure to business cycles may explain our results. More broadly, this test provide new evidence on the validity of our empirical specification.

[76] The underlying assumption is that districts with a high number of ATMs per person will also be characterized by the highest concentration of debit cards and POS machines. We focus on ATM rather than directly on cards or POS, since we cannot directly measure the number of debit cards or POS machines at the district level, but only in aggregate.

[77] Table E.12 also examines the same effect across types of consumption. In particular, the access to electronic payments helped to reduce the impact of the shock in necessary consumption (columns 3 and 4),the impact in explaining the effect for unnecessary consumption was minimal (columns 5 and 6). This heterogeneity between types of consumption is consistent with both demand and supply mechanisms. On the one hand, consumers facing a scarce access to electronic payments may be more likely to allocate a larger share of their electronic money to necessary consumption. On the other hand, for necessary consumption – in particular food – consumers are more likely to face the option to trade with retailers that are larger in size (e.g. grocery chains) relative to unnecessary consumption (e.g. restaurants).

[78] In our main result, there is essentially no difference when we compare the effect on the previous wave - as in Figure E.19 - or the average of the previous three waves, like in Table E.9. Here we choose to compare to the previous wave because this allows us to go further back in time with the placebo.

D Appendix to Section 5

Let Y and Z denote the dependendent and independent variables in the system of equations (9); we first construct the OLS estimate of the data moments, $\hat{\Xi} = (Z'Z)^{-1}Z'Y$. We then estimate the variance-covariance matrix of $\hat{\Xi}$ using the bootstrap. Specifically, we let:

$$var\left(\hat{\Xi}\right) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\Xi}_b - \hat{\Xi}\right)'\left(\hat{\Xi}_b - \hat{\Xi}\right),$$

where $\hat{\Xi}_b$ is the estimate obtained in replication b of the bootstrap. We use $B = 100$ and sample with replacement district by district.

The point estimate for the $N_p \times 1$ vector of parameters Θ is obtained by solving:

$$\hat{\Theta} = \arg\min\left(\hat{\Xi} - \frac{1}{S}\sum_{s=1}^{S}\Xi_{sim}\left(\Theta;\gamma_s\right)\right)' W \left(\hat{\Xi} - \frac{1}{S}\sum_{s=1}^{S}\Xi_{sim}\left(\Theta;\gamma_s\right)\right).$$

In this objective, S is the number of simulations, and $\Xi_{sim}(\Theta;\gamma_s)$ is the same vector of moments as above, estimated using data produced by simulation s. We use $S = 20$ simulations, in keeping with the recommendations of Michaelides and Ng (2000). Each simulation has the same size as the panel data; data is sampled monthly from model simulations. We simulate data with a burn-in period of 10 years for each district. Additionally, γ_s is a vector of random disturbances for simulation s, which we keep constant across values of Θ for which the objective is evaluated. We use Matlab's patternsearch routine to minimize the objective, with 20 randomly drawn starting points for Θ.

Following the litterature (Pakes and Pollard, 1989; Rust, 1994; Hennessy and Whited, 2005, 2007; Taylor, 2010), we use the optimal weighting matrix:

$$W = \frac{1}{N_m} var\left(\hat{\Xi}\right)^{-1}.$$

The variance-covariance matrix for $\hat{\Theta}$, the vector of estimated parameters, is obtained as:

$$\Omega = \left(1 + \frac{1}{S}\right)\left\{\left(\frac{\partial G}{\partial \Theta}(\hat{\Theta})\right)' W \left(\frac{\partial G}{\partial \Theta}(\hat{\Theta})\right)\right\}^{-1},$$

with:

$$G(\Theta) \equiv \hat{\Xi} - \frac{1}{S}\sum_{s=1}^{S}\Xi_{sim}\left(\Theta;\gamma_s\right).$$

We approximate the Jacobian of $G(.)$ using numerical differentiation. We also report the following test statistic for over-identifying restrictions:

$$J = \frac{S}{1+S} G(\hat{\Theta})' \left(var\left(\hat{\Xi}\right)^{-1}\right) G(\hat{\Theta}),$$

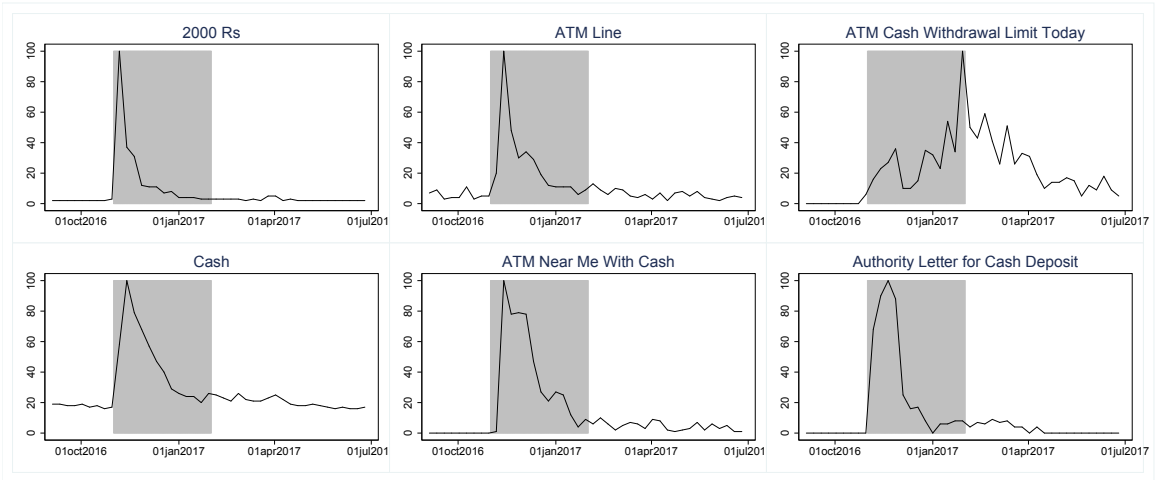
which is distributed as a χ-squared with $N_m - N_p$ degrees of freedom under the null that the over-identifying restrictions hold. Additionally, we use $S_{CI} = 2000$ simulations of the panel, with parameters set to $\hat{\Theta}$, to construct the standard errors and p-values reported in table 7.

In the data, we also re-normalize the Census retail counts so that at least $n \geq 0$ districts reach full adoption. Specifically, for all districts d, we define $X_{d,t} = \min(N_{d,t}/\bar{N}_d^{(n)}, 1)$, where $N_{d,t}$ is the number of adopters per district, and $\bar{N}_d^{(n)} = \frac{N_{d_n,t_0}}{N_{d_n}} N_d$, N_d is the Census count of retailers in district d in 2014, and d_n is a reference district. The reference district is defined as the district with the nth highest un-normalized maximum adoption rate, i.e. the nth highest value of $\max_t \frac{N_{d,t}}{N_d}$. We do this because it is unclear whether the Census counts properly measure the pool of potential adopters. We experimented with values ranging

from $n = 0$ (no normalization) to $n = 10$ (the 10 highest-adoption districts reach full adoption). In all cases, we can reject the null of no complementarities, and estimates of the contribution of complementarities to the long-run change in adoption are largely unchanged, ranging from 40% to 65%. We use $n = 5$ in the estimation that follows.

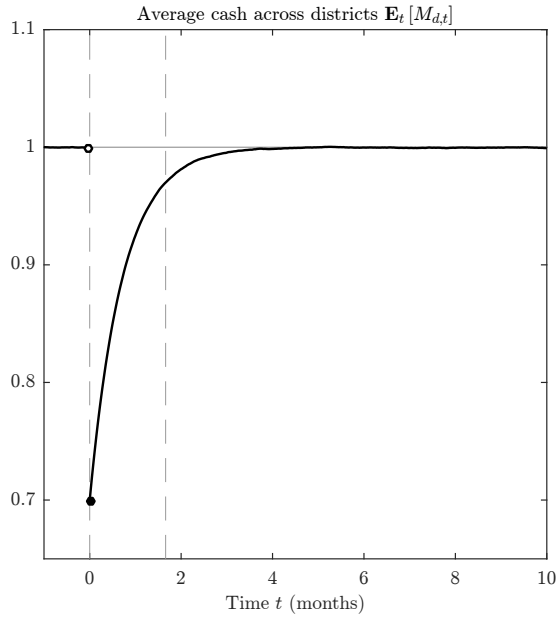
E Appendix figures and tables

Figure E.1: Evidence from Google Search Trends



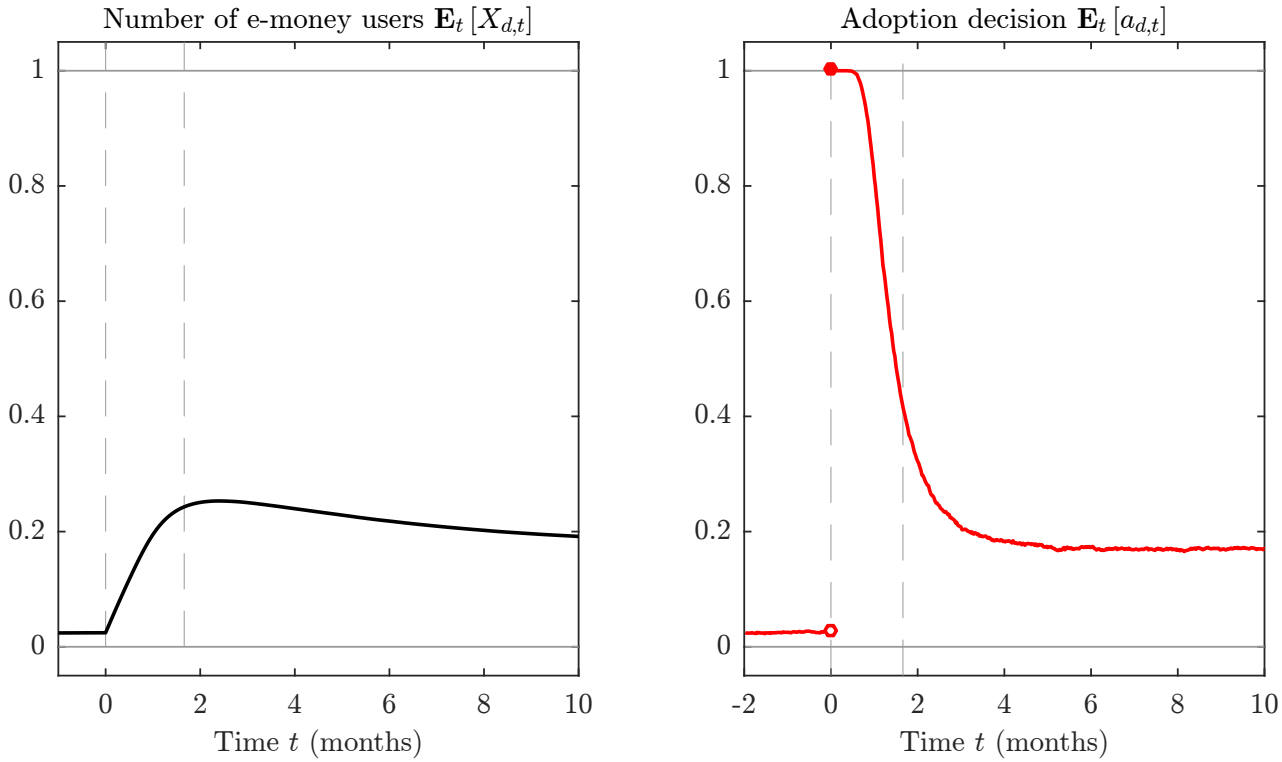
Notes: The figure reports the daily plot between September 2016 and July 2017 of Google searches for several key words that could be representative of public actions and information associated with the demonetization shocks. Data is obtained through Google Trends, and the index is normalized by Google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

Figure E.2: Path of the average level of cash across districts after the cash crunch.



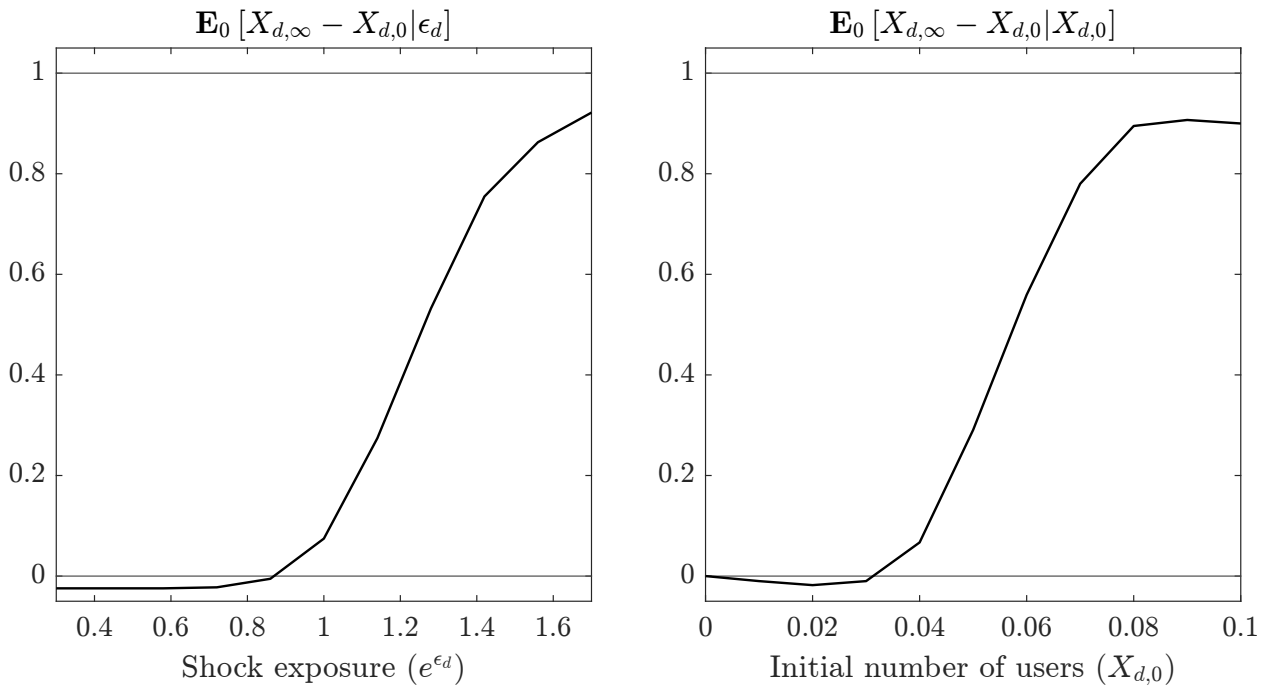
Notes : Path of the average level of cash across districts, $\mathbb{E}_t[M_{d,t}]$, after the cash crunch. The first grey dashed line indicates the date of the shock, and the second one indicates the date at which M_t is back to within 90% of its long-run value, $M_t = M^c = 1$. The model is simulated 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 E.3: Average number of users in the model with complementarities.



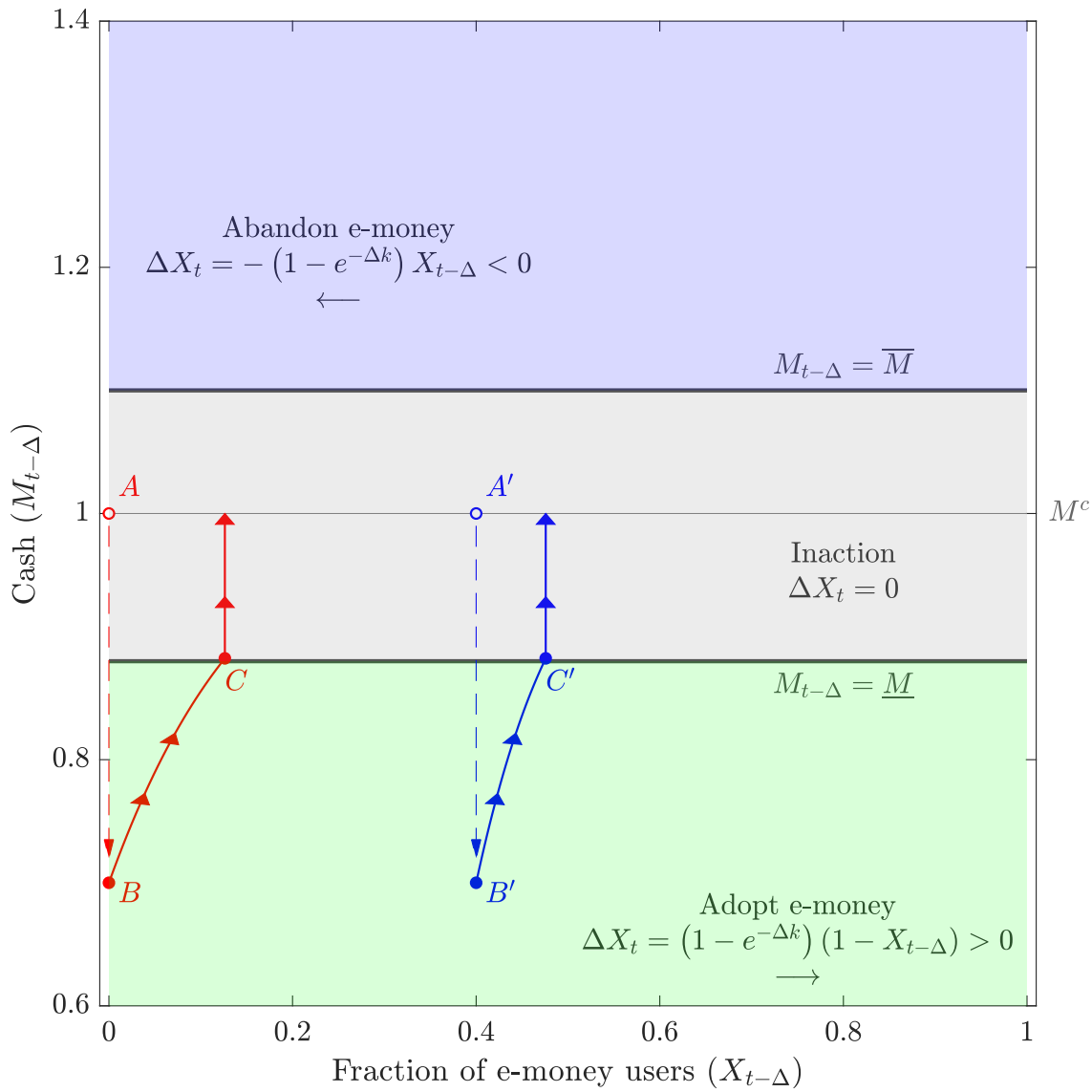
Notes : Average number of users ($\mathbb{E}_t\left[X_{d,t}\right]$, left column) and average adoption decision ($\mathbb{E}_t\left[a_{d,t}\right]$, right column) after the cash crunch in the complementarities model ($C > 0$ and $\kappa = 0$). The results reported here are generated using a version of 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 $\theta = 1.38$, corresponding to a half-life of two weeks.

Figure E.4: Conditional impulse responses in the model with complementarities.



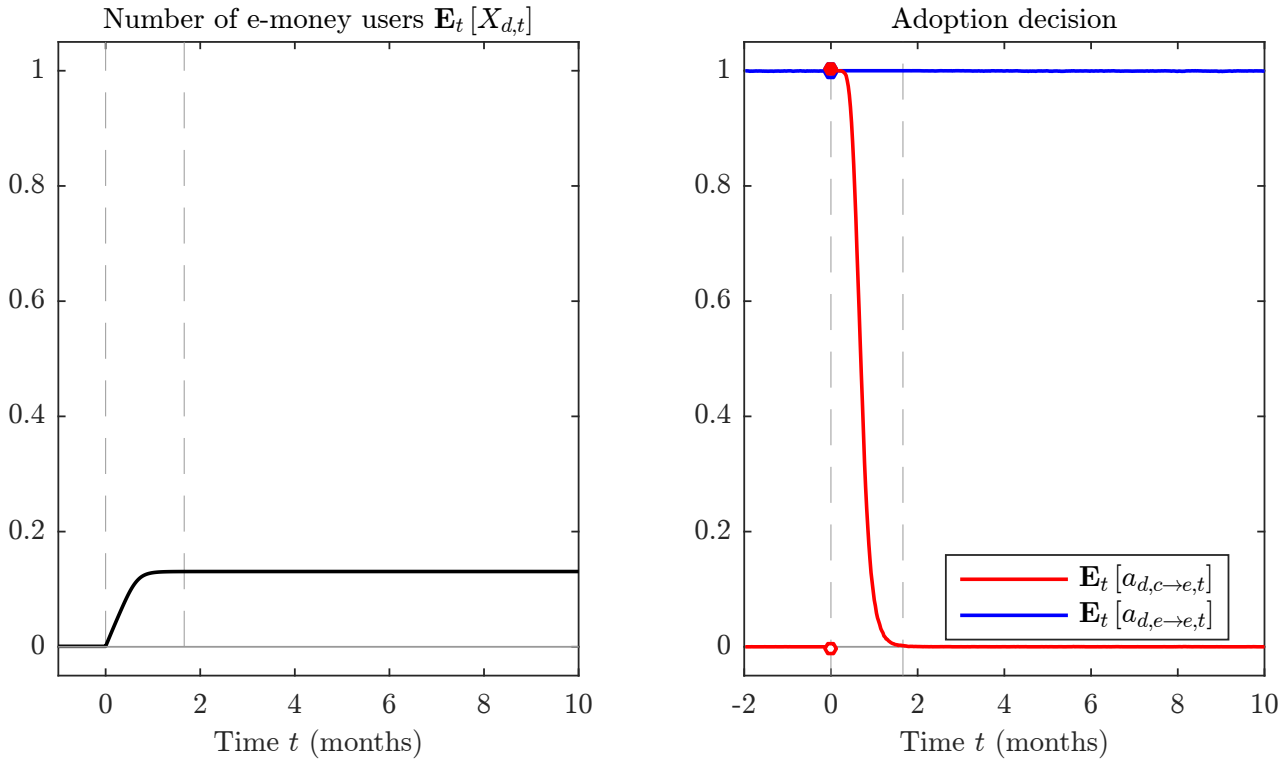
Notes : Conditional impulse responses in the complementarities model ($C > 0$ and $\kappa = 0$). These impulse responses are generated from the same type of simulations as in figure E.3. The left column reports the relationship between the district's exposure to the shock, proxied by e^{ϵ_d} (with a 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 E.5: Adoption dynamics in the fixed cost model ($C = 0$ and $\kappa > 0$).



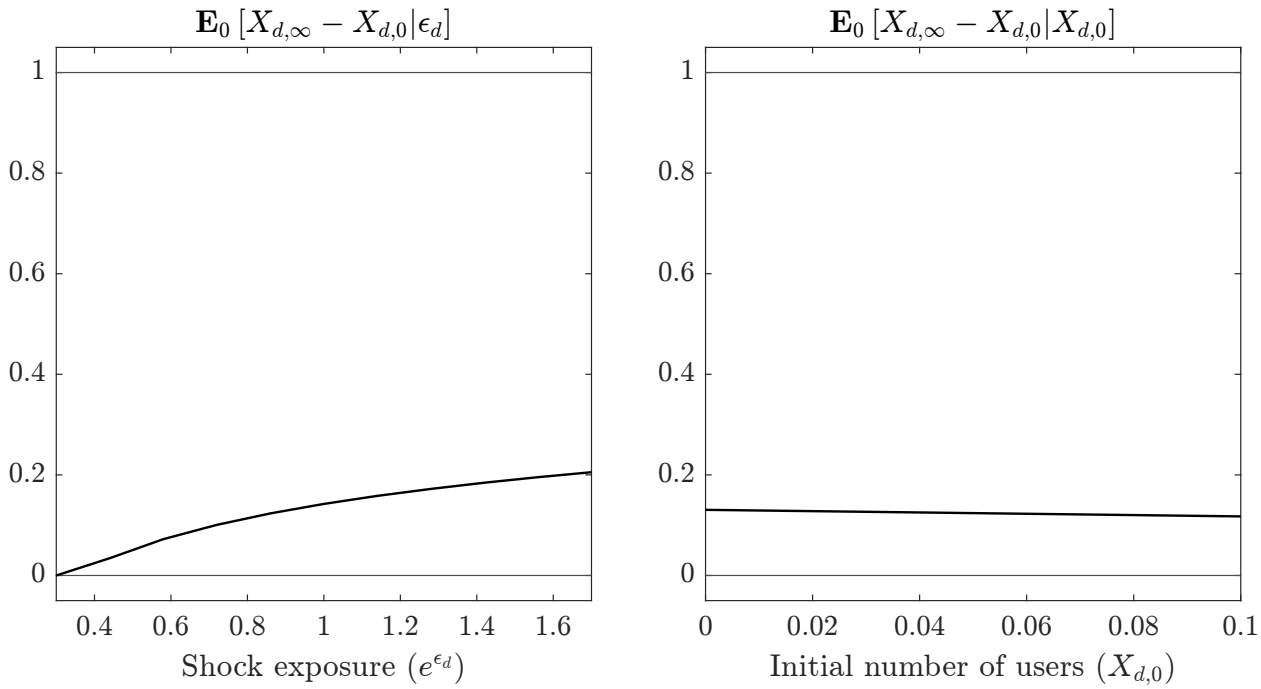
Notes: The red line shows the path of a district that starts with a low adoption level $X_{d,0} = 0$. The blue line shows the path of a district that starts with a high adoption level, $X_{d,0} = 0.4$. The paths are constructed assuming that each district receives no other shock than the initial decline in M_t, i.e. that $\epsilon_t = 0$ for all $t > 0$.

Figure E.6: Average number of users in the fixed cost model.



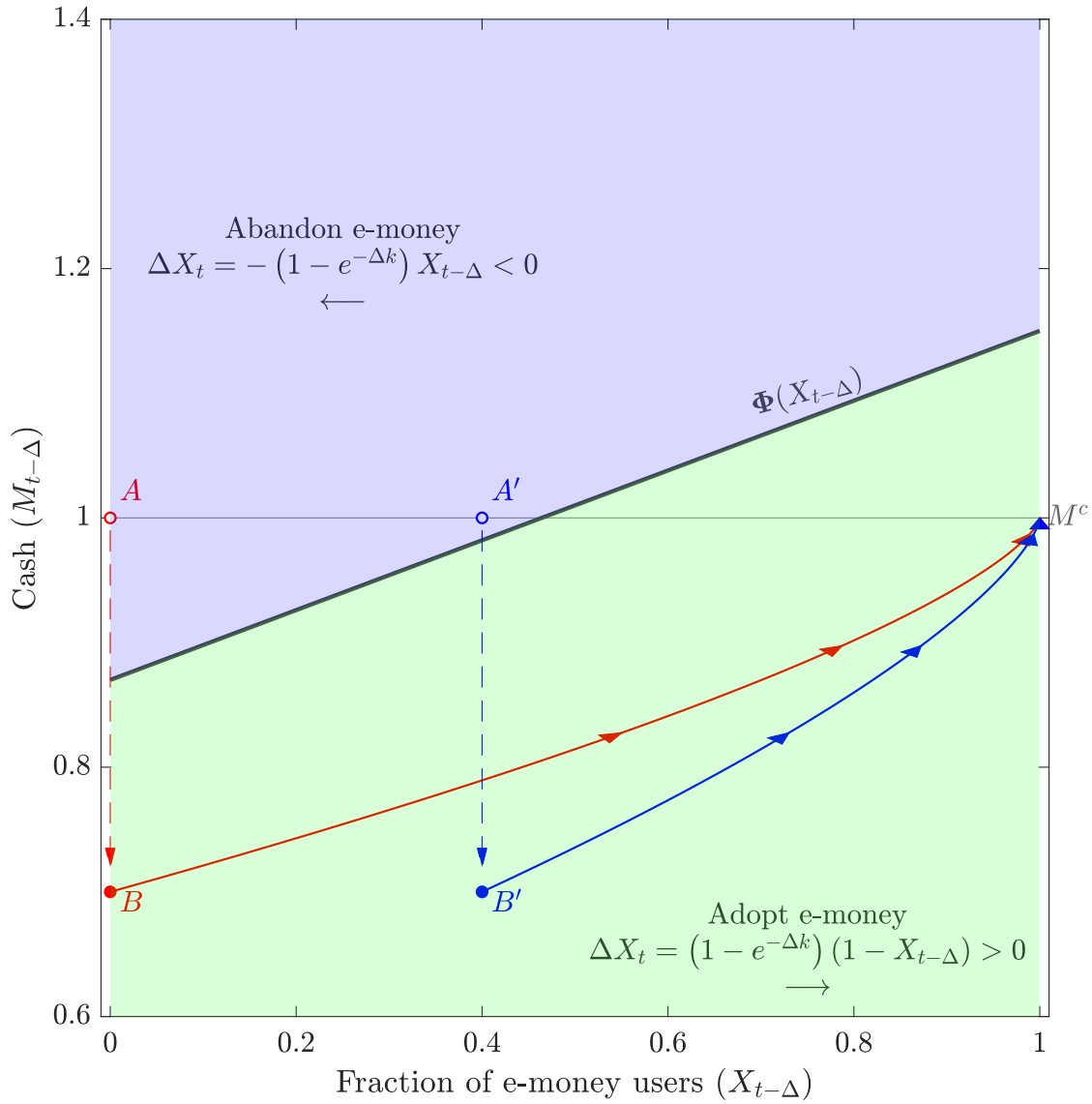
Notes : Average number of users ($\mathbb{E}_t\left[X_{d,t}\right]$, left column) and average adoption decisions (right column) after the cash crunch in the fixed cost model ($C = 0$ and $\kappa > 0$). The graph on the right panel reports separately the adoption decision of firms currently using cash 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 E.7: Conditional impulse responses in the fixed cost model.



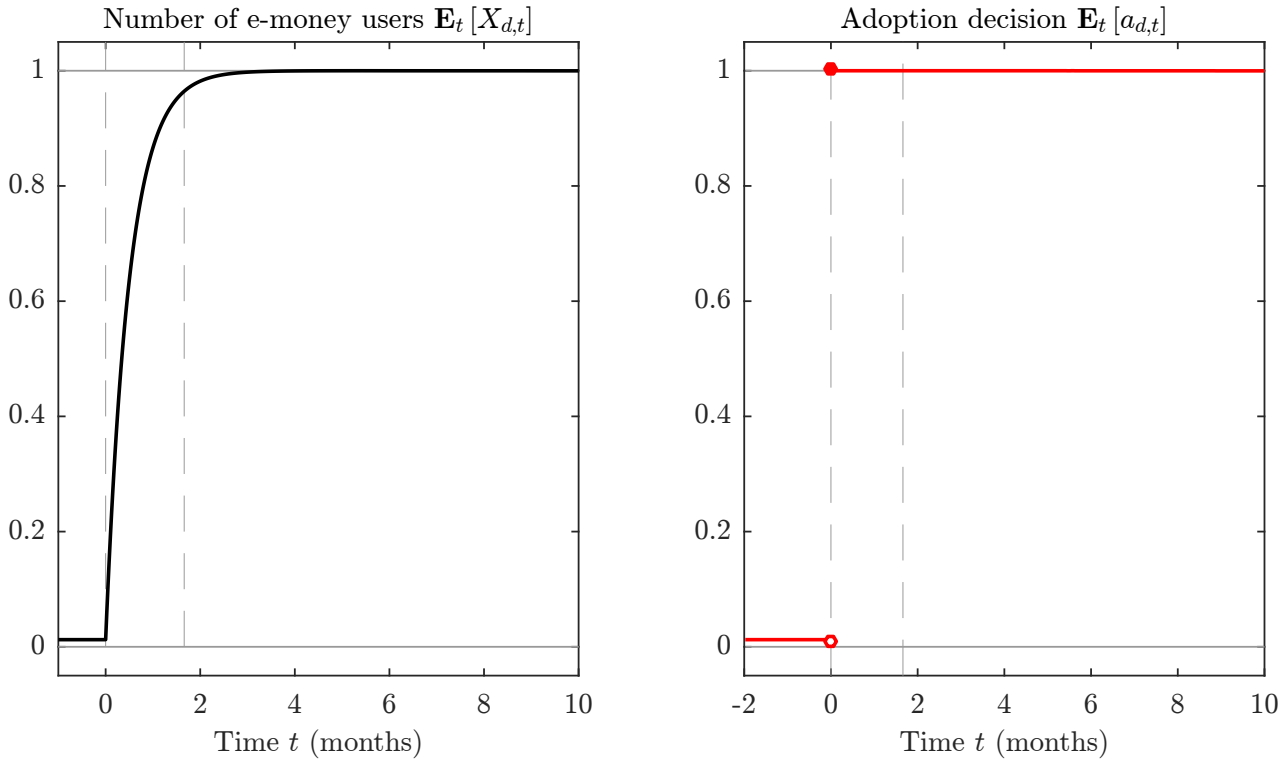
Notes : Conditional impulse responses in the fixed cost model ($C > 0$ and $\kappa = 0$). These impulse responses are generated from the same type of simulations as in figure E.6. The left column reports the relationship between the district's exposure to the shock, proxied by e^{ϵ_d} (with a value of 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 E.8: Adoption dynamics in the model with complementarities and persistent shocks.



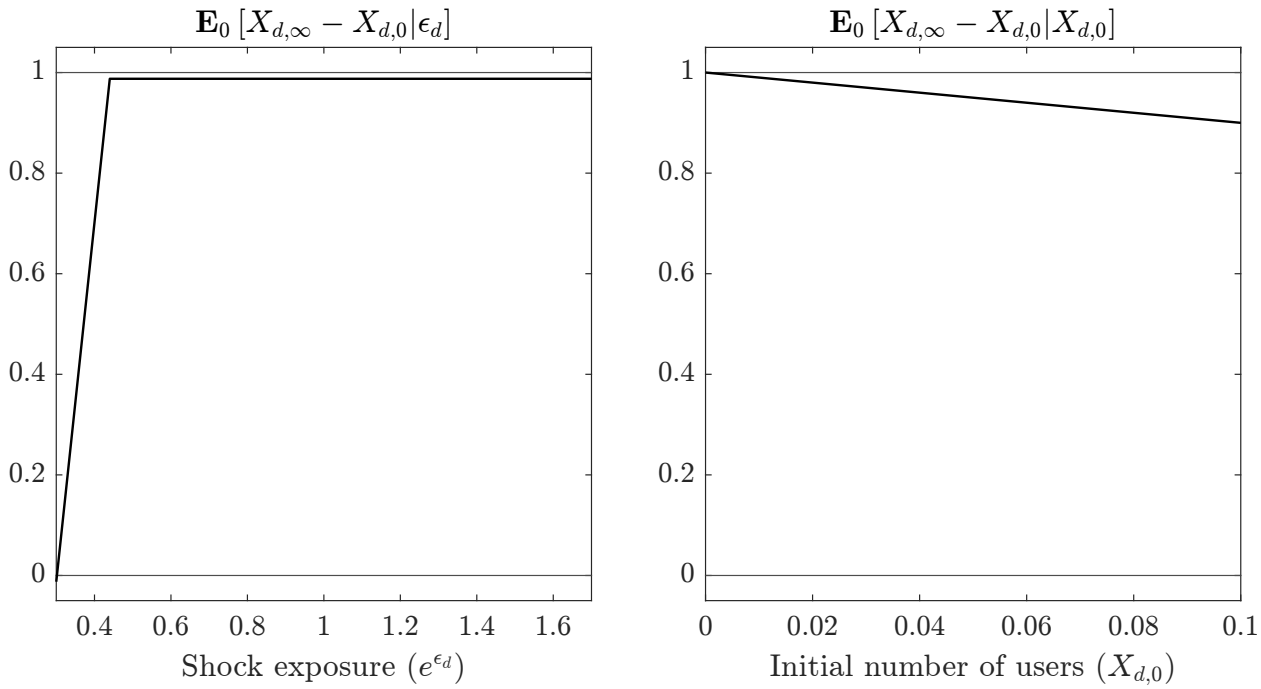
Notes : Adoption dynamics in response to a large decline in M_t in the complementarities model ($C > 0$ and $\kappa = 0$). The model illustrated here corresponds to the case $\theta < k$ (the shock is persistent relative to the adjustment speed of firms.) The red line shows the path of a district that starts with a low adoption level $X_{d,0} = 0$. The blue line shows the path of a district that starts 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 E.9: Average number of users in the model with complementarities and persistent shocks.



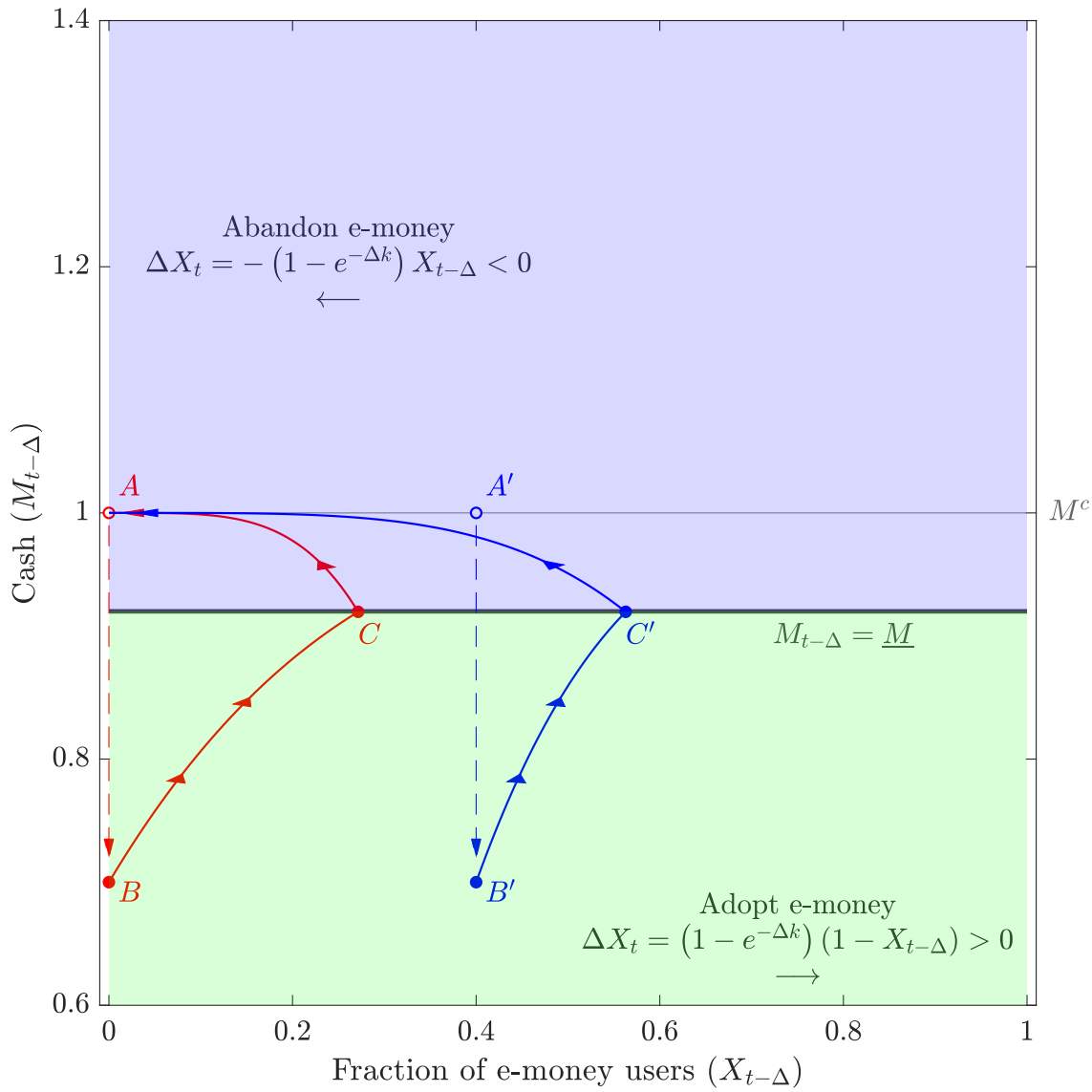
Notes : Average number of users ($\mathbb{E}_t\left[X_{d,t}\right]$, left column) and average adoption decision ($\mathbb{E}_t\left[a_{d,t}\right]$, right column) after the cash crunch in the complementarities model ($C > 0$ and $\kappa = 0$). The results reported here are generated using a version of 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 $\theta = 1.38$, corresponding to a half-life of two weeks.

Figure E.10: Conditional impulse responses in the model with complementarities and persistent shocks.



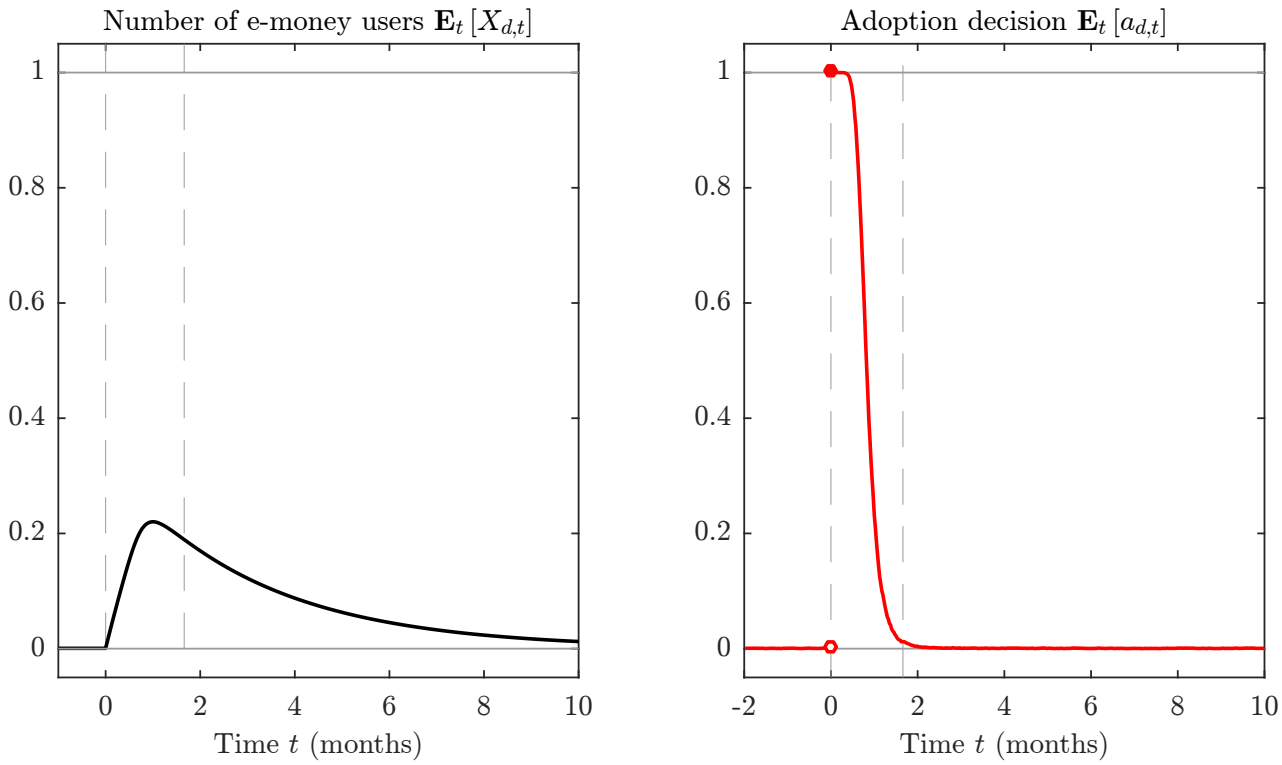
Notes : Conditional impulse responses in the complementarities model ($C > 0$ and $\kappa = 0$). These impulse responses are generated from the same type of simulations as in figure E.9, that is, the case where $\theta < k$ (the shock is persistent relative to the adjustment speed of firms.) The left 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). 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 E.11: Adoption dynamics in the frictionless model.



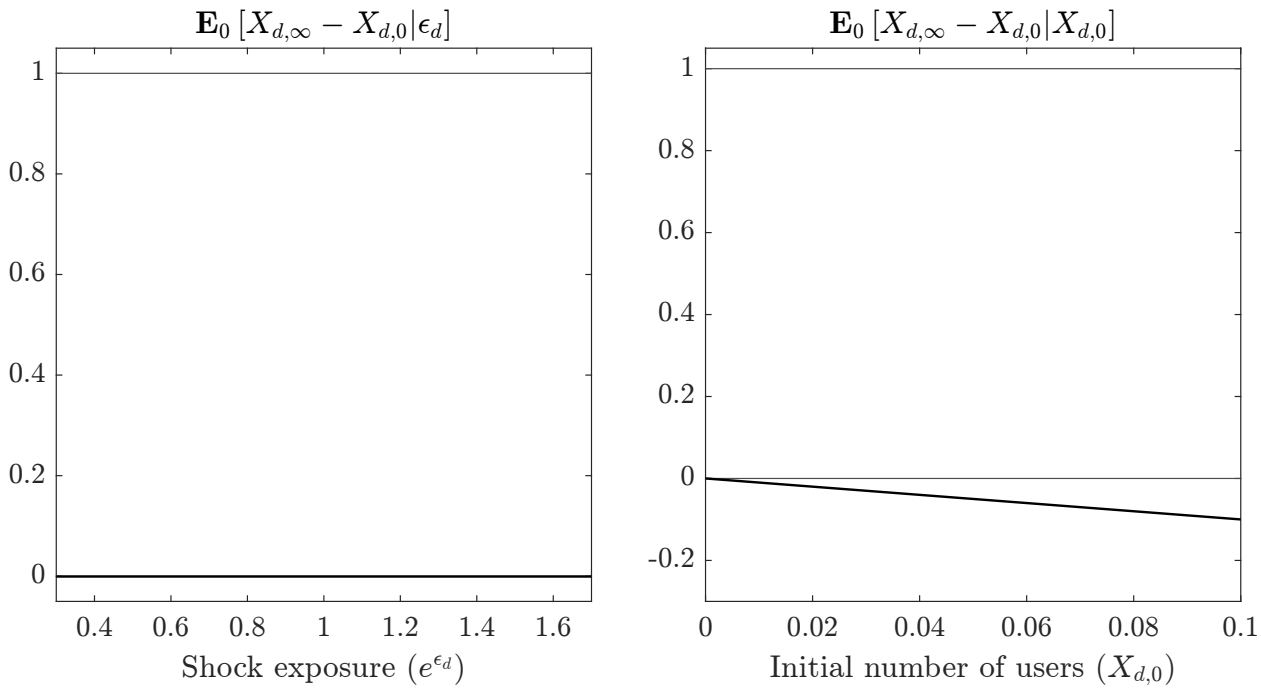
Notes : Adoption dynamics in response to a large decline in M_t in the frictionless model ($C = 0$ and $\kappa = 0$). The red line shows the path of a district that starts with a low adoption level $X_{d,0} = 0$. The blue line shows the path of a district that starts with a high adoption level, $X_{d,0} = 0.4$. 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 E.12: Average number of users in the frictionless model.



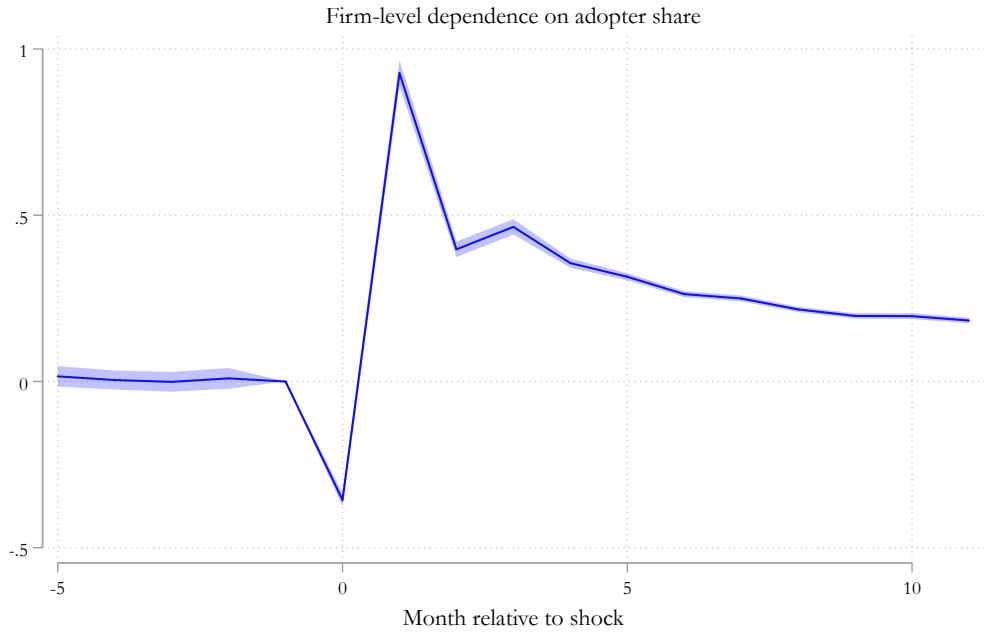
Notes : Average number of users ($\mathbb{E}_t\left[X_{d,t}\right]$, left column) and average adoption decision ($\mathbb{E}_t\left[a_{d,t}\right]$, right column) after the cash crunch in the frictionless model ($C = 0$ and $\kappa = 0$).

Figure E.13: Conditional impulse responses in the frictionless model.



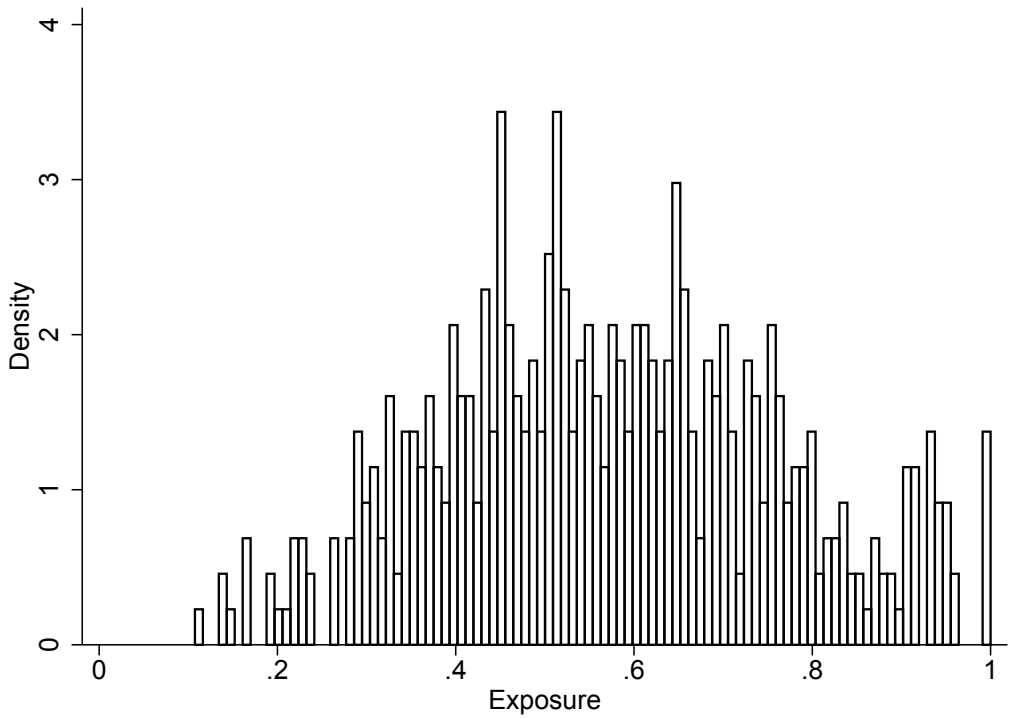
Notes: Conditional impulse responses in the frictionless model ($C = 0$ and $\kappa = 0$). These impulse responses are generated from the same type of simulations as in figure E.9, that is, the case where $\theta < k$ (the shock is persistent relative to the adjustment speed of firms.) The left column 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 E.14: Dynamic effect of adopter share on firm-level adoption in Monte-Carlo simulations.



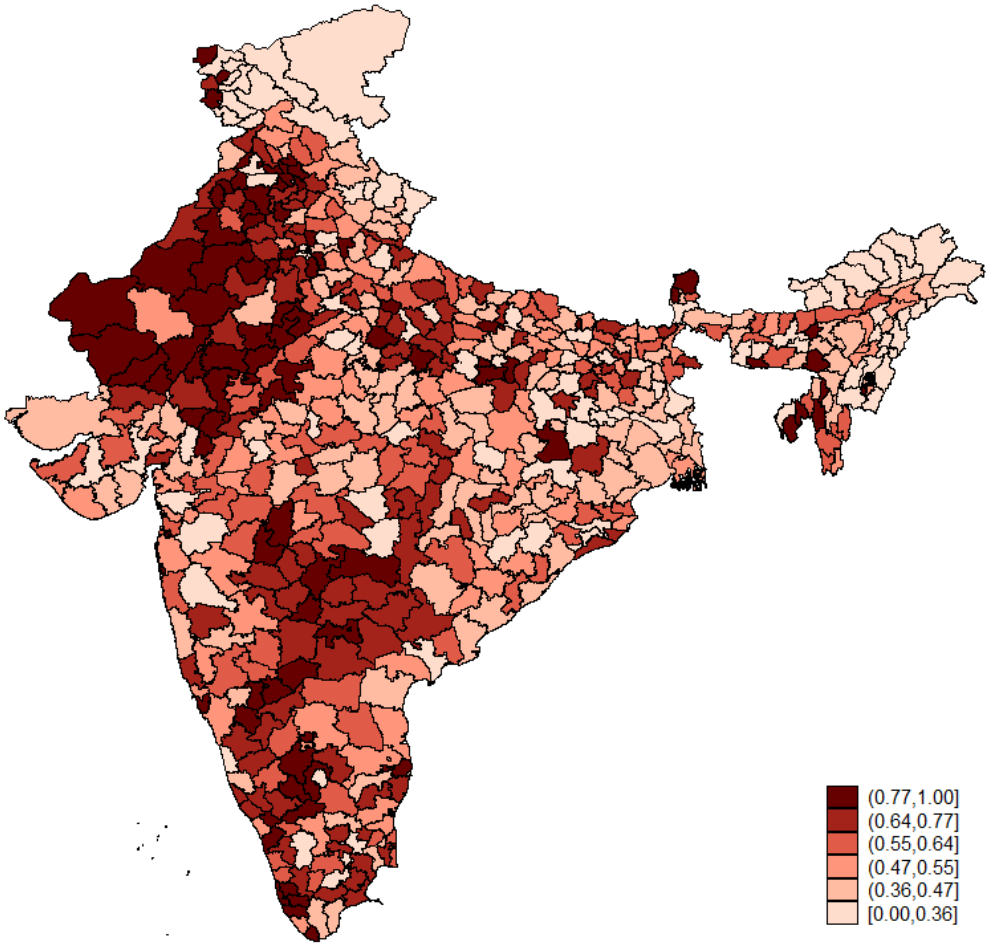
Notes: This graph reports estimates of the coefficient γ_t in Model (17). This regression model is the same as the one used on the data and reported in Figure 9. The coefficients reported capture the dependence on the existing share of adopters at month t, $X_{d,t}$, with the month-9 coefficient normalized to 0, along with 95% confidence bands. The simulated data is aggregated at the district level and sampled monthly; 512 districts (as in the data) and 1000 simulations are used, as in Table E.3.

Figure E.15: Distribution of Exposure$_d$ across districts



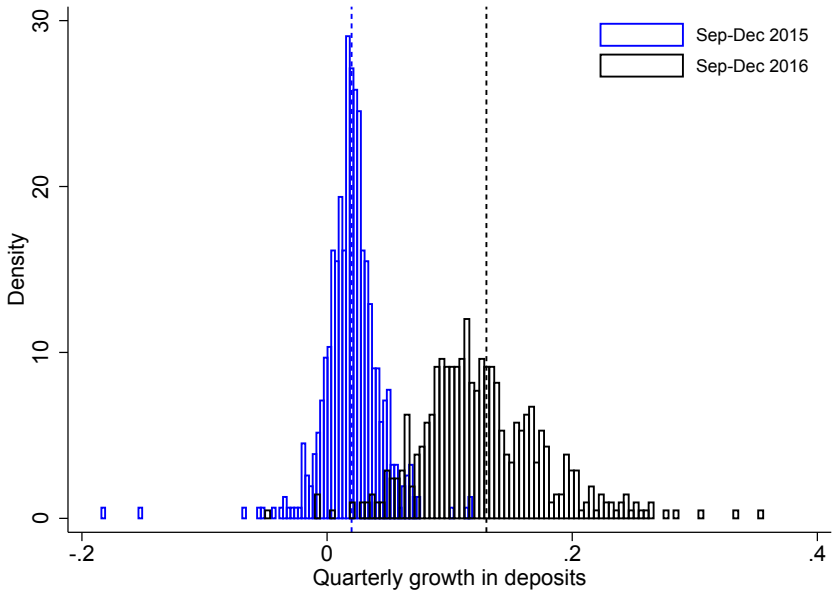
Notes: The figure shows the distribution of Exposure$_d$ (as described in Section 4) across Indian districts. Source: Reserve Bank of India.

Figure E.16: Map of the Distribution of Exposure$_d$



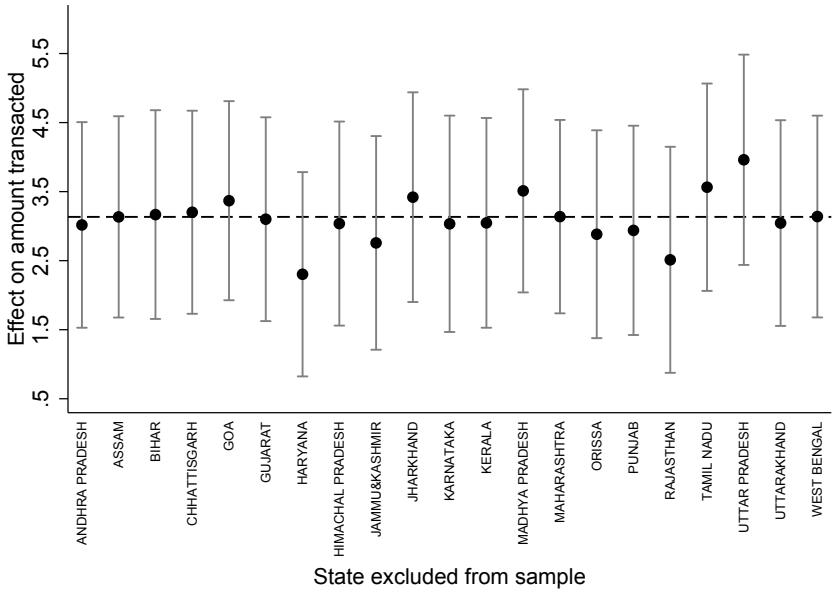
Notes: The figure maps the distribution of Exposure$_d$ (as described in Section 4) across Indian districts. Source: Reserve Bank of India

Figure E.17: Distribution of growth in deposits across districts



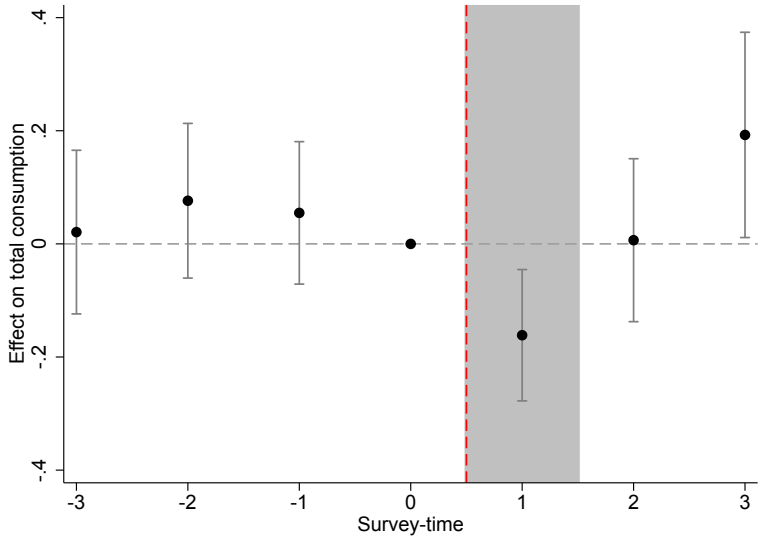
Notes: Distribution across deposits of the growth in total banking sector deposits from October to December during the year 2015 (blue) and 2016 (black). The vertical dashed lines represents the corresponding mean deposit growth for these years. Source: Reserve Bank of India.

Figure E.18: Robustness: one-state out



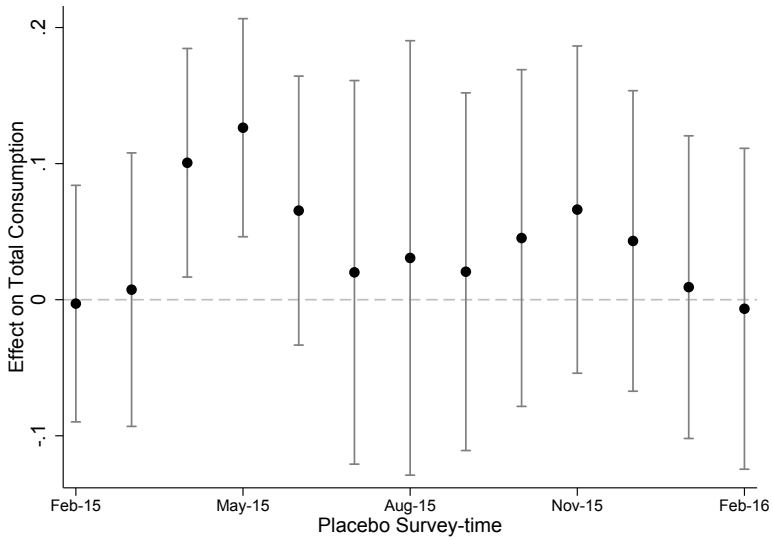
Notes: This figure reports a robustness in which we exclude from the main analysis one state at a time and we recalculate the main coefficient of interest. In particular, we consider the specification in which we look at amount of transactions as an outcome and we consider the coefficient on post multiplied to the chest exposure measure. Each bar reports the main coefficient for the specification excluding the state in the x-axis and the 95% confidence interval. The horizontal dashed line is the main coefficient from the main table of the paper, added for reference.

Figure E.19: Consumption responses based on exposure to the shock



Notes: The figure plots estimates of consumption responses depending on exposure to the shock (Exposure$_d$). The specification we estimate is a version of equation 18 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. The treatment is our measure of Exposure$_d$ as described in Section 4. The dependent variable on the y-axis is the (log) total expense by household (as described in Section C). 95% confidence intervals are represented with the vertical lines; standard errors are clustered at the district level. Source: CMIE Consumption Data.

Figure E.20: Consumption responses based on placebo shocks



Notes: The figure plots the 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 the x-axis. The specification we estimate is a version of equation 18 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 right after shock. The treatment variable is our measure of Exposure$_d$ for the district (as described in Section 4). The dependent variable $\log(y_{h,d,t})$ is the log of total consumption (as described in Section C). 95% confidence intervals are represented with the vertical lines; standard errors are clustered at the district level. Source: CMIE Consumption Data.

Table E.1: 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 reports the results from regression of the district-level deposit growth (between September 30, 2016 and December 31, 2016) on the measure of Exposure$_d$ for the district (as described in Section 4). Columns (1) and (2) use the measure of change in total deposits. Columns (3) and (4) uses the measure of abnormal growth in total deposits, which adjust for the normal deposit growth in the district across the last two years. Specifically, we subtract the mean deposit growth in the last 8 quarters from the growth in 2016Q4 deposits. Columns (5) and (6) uses the dependent variable of deposit growth that is 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. Robust standard errors are reported in parentheses; $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table E.2: 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 quarters 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 4). 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 E.3: 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)
Obs. per sim.	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 16 on simulated firm-level data. The coefficient ρ is the autocorrelation of a 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; 512 districts (as in the data) and 1000 simulations are used. The 95% Monte-Carlo confidence interval is reported in parentheses.

Table E.4: District adoption rates based on initial adoption in electronic payment data: OLS

	log(amount)		log(# users)		log(# switchers)	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}$ (Any Adopter)$_d$ $\times$ $\mathbf{1}(t \geq t_0)$	1.416***		1.751***		1.312***	
	[0.379]		[0.188]		[0.150]	
log(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 $\times$ Month f.e.	✓	✓	✓	✓	✓	✓

Notes: The table shows adoption dependence on initial conditions at the district level. The specification estimated is equation 6. 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) and (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) and (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 E.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.) $\times \mathbf{1}_{\{t \geq t_0\}}$	-1.309***	-1.129***	-1.113***	-0.537***	-0.499***	-0.482***	-0.358**	-0.357***	-0.360***
	[0.374]	[0.357]	[0.345]	[0.182]	[0.151]	[0.138]	[0.143]	[0.116]	[0.108]
Observations	6,846	6,846	6,846	6,846	6,846	6,846	6,846	6,846	6,846
R-squared	0.886	0.886	0.886	0.912	0.912	0.912	0.871	0.871	0.871
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 7, 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 (log) pre-shock banking deposits, share of villages with ATM facilities, share of villages with banking facility, share of rural population, level of population, distance to state capital and employment rate in the district. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table E.6: 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 in the week. Standard errors are clustered at the pincode level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table E.7: 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.		✓	✓	✓
District × Week F.E.			✓	✓
Industry × 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 8 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 in the week. Standard errors are clustered at the district level. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table E.8: Robustness: Firm adoption after inclusion of controls for language concentration and industry concentration

	log(# switchers)			
	(1)	(2)	(3)	(4)
$(\text{Exposure})_d \times \mathbf{1}(t \geq t_0)$	0.877***	0.894***	0.716**	0.688**
	[0.328]	[0.337]	[0.323]	[0.314]
$(\text{Lang. Conc.})_d \times \mathbf{1}(t \geq t_0)$	-0.034	-0.189		
	[0.057]	[0.206]		
$(\text{Exposure})_d \times (\text{Lang. Conc.})_d \times \mathbf{1}(t \geq t_0)$		0.284		
		[0.385]		
$(\text{Ind. Conc.})_d \times \mathbf{1}(t \geq t_0)$			0.247***	0.104
			[0.052]	[0.175]
$(\text{Exposure})_d \times (\text{Ind Conc.})_d \times \mathbf{1}(t \geq t_0)$				0.249
				[0.286]
Observations	6,454	6,454	6,552	6,552
R-squared	0.832	0.832	0.832	0.832
District f.e.	✓	✓	✓	✓
Month f.e.	✓	✓	✓	✓
District Controls × Month f.e.	✓	✓	✓	✓

Notes: The table reports estimates of the effect of cash contraction on the adoption of digital wallet, after controlling for and interacting with proxies of learning in a district. Language concentration in a district is defined as: Lang. Conc.$_d = 1 - \sum_l (\text{share of district } d \text{ population speaking language } l)^2$. Industry concentration in a district is defined as: Ind. Conc.$_d = 1 - \sum_i (\text{share of workers in district } d \text{ employed in industry } i)^2$. We obtain the information on language distribution among population using Census of India (2011), and information on share of employment across 121 industries from Economic Census (2013). Standard errors clustered at district level are reported in parentheses. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.

Table E.9: Consumption responses based on exposure to the shock

Exposure$_d$ **:**	log(Expense$_{\text{Total}}$)	
	Continuous measure	Top 25%
	(1)	(2)
(Exposure)$_d \times \mathbf{1}(t = t_1)$	-0.199***	-0.0577**
	(0.0637)	(0.0234)
(Exposure)$_d \times \mathbf{1}(t = t_2)$	-0.0337	-0.0199
	(0.0815)	(0.0296)
(Exposure)$_d \times \mathbf{1}(t = t_3)$	0.148	0.0146
	(0.102)	(0.0370)
(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 the difference-in-differences estimate for consumption responses for each event-time after the demonetization shock relative to the pre-period (four event-time). The specification estimated is equation 18. The 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 the log of total consumption as defined in Section C. District-level controls include (log) pre-shock banking deposits, share of villages with ATM facilities, share of villages with a 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 E.10: 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 the difference-in-differences estimate for consumption responses across various categories for each event-time after the demonetization shock relative the pre-period (four event-time). The specification estimated is equation 18. The treatment variable is our measure of Exposure_d for the district (as described in Section 4). The dependent variable $\log(y_{h,d,t})$ is either the log of consumption of necessary goods (Column (1)); the log of consumption of unnecessary goods (Column (2)); log of expenditure on bills and rent (Column (3)); the log of expenditure on food (Column (4)); the log of expenditure on recreation activities (Column (5)) as defined in Section C. District-level controls include (log) pre-shock banking deposits, share of villages with ATM facilities, share of villages with a 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 E.11: 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 18. Treatment variable takes the value of 1 if our measure of Exposure$_d$ for the district (as described in Section 4) 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 C. 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 E.12: Heterogeneous consumption responses by district's exposure to alternate payment system

	Total		Necessary		Unnecessary	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}(t = t_1) \times (\text{Exposure})_d$	-0.303***		-0.298***		-0.280**	
	(0.0771)		(0.0740)		(0.121)	
$\mathbf{1}(t = t_2) \times (\text{Exposure})_d$	-0.177*		-0.201**		-0.114	
	(0.0972)		(0.0889)		(0.157)	
$\mathbf{1}(t = t_3) \times (\text{Exposure})_d$	0.103		0.0199		0.275	
	(0.131)		(0.108)		(0.203)	
$\mathbf{1}(t = t_4) \times (\text{Exposure})_d$	0.121		-0.124		0.445	
	(0.212)		(0.182)		(0.319)	
$\mathbf{1}(t = t_1) \times \mathbf{1}(\text{ATM})_d$	-0.118*	-0.0506*	-0.127**	-0.0491**	-0.0917	-0.0413
	(0.0615)	(0.0298)	(0.0518)	(0.0241)	(0.0963)	(0.0463)
$\mathbf{1}(t = t_2) \times \mathbf{1}(\text{ATM})_d$	-0.148**	-0.0535*	-0.157***	-0.0528**	-0.116	-0.0438
	(0.0663)	(0.0302)	(0.0584)	(0.0253)	(0.105)	(0.0490)
$\mathbf{1}(t = t_3) \times \mathbf{1}(\text{ATM})_d$	-0.0431	0.0000	-0.0481	-0.0105	-0.0118	0.0356
	(0.0731)	(0.0315)	(0.0615)	(0.0282)	(0.118)	(0.0486)
$\mathbf{1}(t = t_4) \times \mathbf{1}(\text{ATM})_d$	0.117	0.0644	-0.0285	-0.00480	0.299*	0.145*
	(0.114)	(0.0604)	(0.0968)	(0.0516)	(0.174)	(0.0864)
$\mathbf{1}(t = t_1) \times (\text{Top 25\% Exposure})_d$		-0.106***		-0.104***		-0.102**
		(0.0318)		(0.0291)		(0.0510)
$\mathbf{1}(t = t_2) \times (\text{Top 25\% Exposure})_d$		-0.0782**		-0.0829**		-0.0666
		(0.0375)		(0.0335)		(0.0609)
$\mathbf{1}(t = t_3) \times (\text{Top 25\% Exposure})_d$		0.00985		-0.0126		0.0669
		(0.0478)		(0.0385)		(0.0761)
$\mathbf{1}(t = t_4) \times (\text{Top 25\% Exposure})_d$		0.0227		-0.0993		0.211
		(0.0914)		(0.0776)		(0.137)
$\mathbf{1}(t = t_1) \times (\text{Exposure})_d \times \mathbf{1}(\text{ATM})_d$	0.185*		0.217**		0.129	
	(0.102)		(0.0987)		(0.151)	
$\mathbf{1}(t = t_2) \times (\text{Exposure})_d \times \mathbf{1}(\text{ATM})_d$	0.232*		0.261**		0.167	
	(0.119)		(0.112)		(0.185)	
$\mathbf{1}(t = t_3) \times (\text{Exposure})_d \times \mathbf{1}(\text{ATM})_d$	0.0744		0.0724		0.0536	
	(0.142)		(0.116)		(0.226)	
$\mathbf{1}(t = t_4) \times (\text{Exposure})_d \times \mathbf{1}(\text{ATM})_d$	-0.139		0.0818		-0.454	
	(0.218)		(0.202)		(0.326)	
$\mathbf{1}(t = t_1) \times (\text{Top 25\% Exposure})_d \times \mathbf{1}(\text{ATM})_d$		0.0913**		0.114***		0.0479
		(0.0437)		(0.0423)		(0.0645)
$\mathbf{1}(t = t_2) \times (\text{Top 25\% Exposure})_d \times \mathbf{1}(\text{ATM})_d$		0.101*		0.116**		0.0628
		(0.0520)		(0.0470)		(0.0811)
$\mathbf{1}(t = t_3) \times (\text{Top 25\% Exposure})_d \times \mathbf{1}(\text{ATM})_d$		0.00388		0.0110		-0.0363
		(0.0600)		(0.0494)		(0.0933)
$\mathbf{1}(t = t_4) \times (\text{Top 25\% Exposure})_d \times \mathbf{1}(\text{ATM})_d$		-0.0622		0.0771		-0.289*
		(0.0994)		(0.0949)		(0.148)
Observations	554,894	554,894	554,899	554,899	554,899	554,899
R-squared	0.704	0.704	0.730	0.730	0.618	0.618

Notes: The table shows triple-difference estimate for consumption responses for each event-time post the demonetization shock relative the pre-period (four event-time), based on district's access to ATM facility. Treatment variable is our measure of Exposure_d for the district (odd columns) and takes the values of 1 if the measure of Exposure_d is in the top quartile of the distribution (even columns). $\mathbf{1}(\text{ATM})_d$ takes the values of 1 if the number of ATM per capita in district is above the median of the distribution. The dependent variable $\log(y_{h,d,t})$ is either the log of total consumption (Column 1-2); log of necessary consumption (Column 3-4); log of unnecessary consumption (Column 5-6). as defined in Section C. 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. $***: p < 0.01$, $**: p < 0.05$, $*: p < 0.1$.