NBER WORKING PAPER SERIES

WHEN LESS IS MORE:
EXPERIMENTAL EVIDENCE ON INFORMATION DELIVERY DURING INDIA'S DEMONETIZATION

Abhijit Banerjee
Emily Breza
Arun G. Chandrasekhar
Benjamin Golub

Working Paper 24679
http://www.nber.org/papers/w24679

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2018

We thank Mohammad Akbarpour, George Beknazar-Yuzbashev, Chris Blattman, Pascaline Dupas, Matt Jackson, Rachel Kranton, Michael Kremer, Ted Miguel, Ben Olken, Evan Sadler and many seminar participants for helpful comments and conversations. We are grateful to Ruben Enikolopov for a very helpful discussion. We thank Pooja Suri, Arnesh Chowdhury, Piyush Tank, Tithee Mukhopadhyay, Anushka Chawla, Devika Lakhote, Alokik Mishra, and Shambhavi Priyam for exceptional research assistance. We thank the National Science Foundation for financial support under grant SES-1559469. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

When Less is More: Experimental Evidence on Information Delivery During India's Demonetization
Abhijit Banerjee, Emily Breza, Arun G. Chandrasekhar, and Benjamin Golub
NBER Working Paper No. 24679
June 2018
JEL No. D8,G41,O33

ABSTRACT

How should policymakers disseminate information: by broadcasting it widely (e.g., via mass media), or letting word spread from a small number of initially informed "seed" individuals? While conventional wisdom suggests delivering information more widely is better, we show theoretically and experimentally that this may not hold when people need to ask questions to fully comprehend the information they were given. In a field experiment during the chaotic 2016 Indian demonetization, we varied how information about demonetization's official rules was delivered to villages on two dimensions: how many were initially informed (broadcasting versus seeding) and whether the identity of the initially informed was publicly disclosed (common knowledge). The quality of information aggregation is measured in three ways: the volume of conversations about demonetization, the level of knowledge about demonetization rules, and choice quality in a strongly incentivized decision dependent on understanding the rules. Our results are consistent with four predictions of a model in which people need others' help to make the best use of announced information, but worry about signaling inability or unwillingness to correctly process the information they have access to. First, if who is informed is not publicized, broadcasting improves all three outcomes relative to seeding. Second, under seeding, publicizing who is informed improves all three outcomes. Third, when broadcasting, publicizing who is informed hurts along all three dimensions. Finally, when who is informed is made public, telling more individuals (broadcasting relative to seeding) is worse along all three dimensions.

Abhijit Banerjee
Department of Economics, E52-540
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and NBER
banerjee@mit.edu

Emily Breza
Harvard University
Littauer Center, M28
1805 Cambridge Street
Cambridge, MA 02138
and NBER
ebreza@fas.harvard.edu

Arun G. Chandrasekhar
Department of Economics
Stanford University
579 Serra Mall
Stanford, CA 94305
and NBER
arungc@stanford.edu

Benjamin Golub
Harvard University
Littauer Center, 308
1805 Cambridge Street
Cambridge, MA 02138
ben.golub@gmail.com

A randomized controlled trials registry entry is available at
https://www.socialscienceregistry.org/trials/2881

1. Introduction

How should a policymaker deliver information to a community? In practice, there are two commonly used strategies that are very different from each other: (1) broadcasting information widely to all (e.g., radio, television, newspaper, or a Twitter feed) and (2) delivering information to a select few "seed" individuals and relying on subsequent diffusion (which we see in viral marketing, agricultural extension services, or the introduction of microcredit).[1] Finding an effective way to provide information can be very important in getting people to make the right choices, and in extremis—for instance, during an epidemic or a natural disaster—can lead to lives being saved.

This paper is about the choice of a dissemination strategy. There are, of course, many factors that inform this choice: broadcasting may be infeasible in some contexts, say because there are no local TV channels, or seeding may be too expensive because there is not enough information about who the right seeds should be. Here we focus on one key factor, which is how the dissemination strategy affects engagement in social learning, and in particular people's willingness to ask questions—an issue that has received little attention in the social learning literature.

The point of departure of this paper is that people often need to ask questions in order to process the information they were given. At the same time, they may be hesitant to ask because of worries about what asking may imply about them—for example, that they are unable or unwilling to use information they already have access to. A recent survey by Chandrasekhar, Golub, and Yang (2017) asked 122 villagers in India about their willingness to ask questions of other community members about farming, health or finance: 88% of respondents felt constrained in terms of how many times they could seek advice from someone else in their community. In 64% of the cases where they felt limited in their capacity to ask for advice, the respondents said they refrained from seeking out information because they did not want to appear weak or uninformed. Chandrasekhar et al. (2017) goes on to develop a signaling model that captures this idea and finds support for its predictions in a lab experiment.[2]

The idea that people's signaling and reputational concerns affect their decisions to ask questions has important consequences for the design of public communication. In particular, both how many people are informed and what they are told about

[1]See, e.g., Leskovec et al. (2007); Ryan and Gross (1943); Conley and Udry (2010); Miller and Mobarak (2014); Banerjee et al. (2013); Beaman et al. (2016); Cai et al. (2015).

[2]Bursztyn, Egorov, and Jensen (2016) provides evidence that signaling concerns strongly influence the choices of high school students, potentially to the serious detriment of their educational careers.

what others know (meta-knowledge) will influence whether questions get asked and therefore how many people learn information accurately. For example, if a decision-maker knows that others believe he has access to certain information—say, because it was widely broadcast—he may be less willing to question others about it to avoid the suspicion that his comprehension is deficient. In contrast, if he is unlikely to have been informed and everyone knows it, no negative signal would be associated with asking questions.[3] Thus, while broadcasting has an obvious advantage over seeding in that information immediately reaches more people, it also has a potential disadvantage: it may activate reluctance to seek information and thus harm the social aspect of learning.

Motivated by this tension, we study, in theory and in an experiment, an environment in which we can vary both (i) how many people are given information (seeding a few versus broadcasting to everyone) and (ii) meta-knowledge about who is informed (making the set of informed people common knowledge or not). A simple theory of how the two dimensions affect incentives to seek leads to four predictions. (1) If it is *not* common knowledge who was informed, then broadcasting should increase engagement in social learning relative to seeding. In the broadcast case, people learn about at least the existence of important information, while the absence of common knowledge (as we show) limits signaling concerns. (2) With seeding, making it common knowledge who is informed should lead to more engagement in social learning: it helps people find those who are informed and ask them questions, again without exposing the seekers to signaling concerns. On the other hand, (3) broadcasting with common knowledge should generate *less* participation in social learning than broadcasting without common knowledge, because with common knowledge, it is understood that someone seeking information had access to it already. By the same token, (4) if it is common knowledge who is informed, then going from seeding to broadcasting should *reduce* engagement in social learning, because it is deterred by signaling concerns.

The implications of these results for the quality of information aggregation are less obvious. The fact that broadcast generates fewer conversations than seeding under common knowledge need not imply that people learn less—because the *direct*

[3]The potential negative signal does not necessarily have to be about lack of comprehension. The signal could concern one's trade-off between one's own and others' effort: "Why didn't you pay attention then instead of wasting my time now?" The implications for communication design are likely to be similar. Of course, in some contexts asking questions may be a positive signal. The direction of the effect in any specific case is an empirical matter.

effect of broadcasting is to inform more people.[4] Nevertheless, the theory opens up the possibility of reversing the ordering one would expect from the "infection-type" models often used to study information transmission (Bass, 1969; Bailey, 1975; Jackson, 2008; Jackson and Yariv, 2011; Aral and Walker, 2012; Akbarpour et al., 2017). In those models, people pass on the information with some probability without being asked, or ask questions without any strategic motive if they don't have the information (again with some probability).[5] In such a world, more information is always better: broadcasting is better than seeding whether or not there is common knowledge.[6] Under the theory proposed here, these conclusions may not hold—seeding can be better than broadcast if there is common knowledge, and common knowledge can be strictly worse than no common knowledge when the information is broadcast.

Motivated by the possibilities opened up by our theory,[7] we conducted a field experiment in India approximately six weeks after Prime Minister Narendra Modi announced the demonetization of all Rs. 500 and Rs. 1000 notes. The policy was unexpected and far-reaching, affecting 86% of India's currency. While there was near-universal awareness of the broad outlines of the policy in India, its chaotic implementation, with over 50 rule changes in a seven week period, led to widespread confusion and misinformation (see Appendix A). For example, in our sample, 25% did not understand that demonetized currency could only be deposited into a bank account, not be exchanged for new bills over the counter; 15% thought that the Rs. 10 coin was also being demonetized. As a result, the period following the policy announcement offered ample scope for trying out different strategies for informing the populace about the actual rules.

In over 200 villages, in the ten days (starting on December 21, 2016) leading up to the date when the old Rs. 500 and Rs. 1000 bills stopped being accepted for deposit into bank accounts, we randomly varied how we provided information to villages. The experiment varied at the village level (1) whether information was provided to all households or to just five seed households; (2) whether who was informed within the village was made common knowledge; (3) the number of facts provided, which

[4]Thus, further conditions are needed to predict less learning, and we discuss these when we present our theory.

[5]We are assuming, of course, that not everyone gets the message even when the information is broadcast.

[6]In general, these models do not discuss meta-knowledge, but more information about where signals are delivered cannot hurt. In a simple discussion of a theoretical benchmark, we formalize the prediction that without incentive frictions, more information cannot be worth less to the community.

[7]It is worth emphasizing that the theory in Chandrasekhar, Golub, and Yang (2017) existed in the public domain prior to the experiment.

could be either 2 or 24. The information we provided always consisted of a list of facts in a short pamphlet; the same pamphlet was provided to all households who received information in that village. The facts came directly from the Reserve Bank of India's circulars (released as of December 19th, 2016), and thus represent the information that the policymakers themselves chose to communicate to the public. We then returned to the villages to collect our outcome data, approximately three days after the intervention. Importantly for our experiment, the information contained in the pamphlets was unlikely to cover everything villagers needed to know about the policy. First, of course, in half the villages we only provided two facts; even 24 facts was short of a full description of all relevant aspects of the policies. Second, the facts conveyed in the RBI circulars involved terms that were not necessarily familiar to the recipients, and it would not have been clear to everyone how the facts applied to their decisions. As a result, communication was probably beneficial even for those who received the pamphlets; indeed, our hope was in part that the pamphlets would make the villagers realize that there was hard information to be had, and encourage the sharing of information, including information about topics that were not in the pamphlets.

The outcomes we measure are: engagement in social learning, policy knowledge, and choice in an incentivized decision. For engagement in social learning, we asked how many conversations villagers had had about demonetization over the prior three days. For knowledge, we asked questions about the demonetization rules. For incentivized choice, we asked the subjects to select one of the following three options: (a) same-day receipt of a Rs. 500 note (worth 2.5 days' wage) in the old currency, which was still legal for depositing in the bank; (b) an IOU for Rs. 200 in Rs. 100 notes (unaffected by demonetization) redeemable 3-5 days later; and (c) an IOU for dal (pigeon peas) worth Rs. 200, again redeemable 3-5 days later. At the time of the choice, subjects still had time to deposit the Rs. 500 note at the bank, no questions asked, for the cost of going to the bank.

As suggested above, in an infection model with mechanical transmission (i.e., no endogenous decision of engagement), providing information to all households in a village (a 10-fold increase), providing more meta-information (common knowledge), and providing more information in the each pamphlet (a 12-fold increase) should all lead to more knowledge being aggregated, and better decision-making on the part of our subjects. The reasons are simple: more information leads to a diffusion process with the same local dynamics but more starting points. Common knowledge should always increase the number of conversations, because it makes it easier to find those who know. Finally, more information in the pamphlets increases the number of facts that are mechanically being transmitted. Under the alternative theory sketched above, these simple intuitions can be reversed.

In addition our theory has the interesting prediction that the ranking of broadcast and seeding in terms of generating conversations, may be reversed by adding common knowledge. Without common knowledge, broadcast will tend to generate more conversations because it makes it easier to find someone to ask, but if there is common knowledge the fact people shy away from asking questions under broadcast but not under seeding, can lead to a reversal. The infection model has no clear prediction about whether seeding or broadcast generates more conversations: broadcast makes it easier to find someone to ask but it also makes it likelier that they already have enough information to make it unnecessary to talk.[8] However there is no reason why there would be reversal when we bring in common knowledge into the infection model.

We find strong evidence for all of these kinds of reversals. Our core results are as follows. First we look at endogenous participation in social learning.[9]. Adding common knowledge to a seeding strategy makes for more conversations: going from (Seed, No CK) to (Seed, CK) increases the number of conversations by 103% ($p = 0.04$) but among broadcast strategies we find the reverse: (Broadcast, CK) generates 63% fewer conversations ($p = 0.02$) than (Broadcast, No CK). Furthermore, (Broadcast, CK) leads to 61% *less* conversations (p = 0.029) than (Seed, CK) but going from (Seed, No CK) to (Broadcast, No CK) increases the number of conversations by 113% (p = 0.048).The fact that common knowledge has opposite effects across seeding and broadcast strategies and reverses the ranking of seeding and broadcast in terms of the number of conversations generated, are all consistent with the endogenous participation model sketched above and unlikely to obtain in models where there is no strategic motive behind participation in social learning.

Second, we turn to whether the changes in endogenous participation in learning correspond to changes in knowledge. Going from (Seed, No CK) to (Seed, CK) reduces the error rate on our knowledge survey by 7.3% ($p = 0.0142$). On the other hand going from (Seed, CK) to (Broadcast, CK) leads to a 5.6% *increase* in the error rate on our knowledge survey ($p = 0.062$). This shows that even though all households are

[8]Here we have in mind a version of the infection model where conversations are started by the uninformed non-strategically asking questions of their network neighbors

[9]Niehaus (2011) emphasizes a different aspect of endogenous participation. In his model the informed party decides whether or not to reveal what they have learnt

given signals instead of merely five, the amount of knowledge for a random household is less, not more, suggesting an important role for social learning. The exact opposite happens when going from (Seeding, No CK) to (Broadcast, No CK), corresponding to a 6.22% decline in the error rate ($p = 0.053$). Within broadcast, (Broadcast, CK) has a 4.6% higher error rate than (Broadcast, No CK), though the effect is not statistically significant ($p = 0.17$).

Third, when we look at whether subjects choose the Rs. 500 note, which at that time was still accepted for deposit by banks, or an IOU worth Rs. 200 in cash or in kind to be paid in 3-5 days, we again see a similar pattern. Going from (Seed, No CK) to (Seed, CK) leads to a 81% increase in the probability of choosing the Rs. 500 note ($p = 0.037$). Going from (Seed, CK) to (Broadcast, CK) leads to a 38.5% decline in the probability of choosing the Rs. 500 note ($p = 0.104$). I contrast, there is a 114% increase in the probability of choosing the note when going from (Seed, No CK) to (Broadcast, No CK) ($p = 0.014$) and going from (Broadcast, No CK) to (Broadcast, CK) leads to a 48% decline in the probability of choosing the Rs. 500 note ($p = 0.041$).

The results from the choice exercise, reassuringly, mirror what we find with the knowledge measures and conversations. Taken together, they provide clear evidence of the various ways in which more can be less, and more generally, make a strong case for taking strategic participation seriously when designing information campaigns.

The remainder of the paper is organized as follows. Section 2 describes the setting and experimental design. Section 3 presents our theoretical framework where agents endogenously choose to participate in social learning. There we study the equilibria of our model and show how social learning varies as we change the environment along the main dimensions. We also compare the predictions to those of models in which signaling-based seeking frictions are not present. We present our empirical results in Section 4. Section 5 provides a discussion.

2. Experiment

2.1. **Demonetization.** On November 8, 2016, Indian Prime Minister Narendra Modi announced a large-scale demonetization. At midnight after the announcement, all outstanding Rs. 500 and Rs. 1000 notes (the "specified bank notes" or SBNs) ceased to be legal tender. Demonetization affected 86% of circulating currency in terms of value, and individuals holding SBNs had until December 30, 2016 to deposit them in a bank or post office account. Modi intended for the surprise policy to curb "black money" and more broadly to accelerate the digitization of the Indian economy. The policy affected almost every household in the country, either because they held the SBNs, or through the cash shortages that resulted from problems in printing and distributing enough new bills fast enough.

The implementation of the policy was chaotic. The initial rollout revealed a number of ambiguities, loopholes, and unintended outcomes. As a result, the government changed the rules concerning demonetization over 50 times in the seven weeks following the announcement. The rule changes concerned issues such as the time frame for over-the-counter exchange of SBNs, the cash withdrawal limit, the SBN deposit limit, and various exemptions—e.g., for weddings, which tend to be paid for in cash. See Appendix A for a timeline of these rule changes.

2.2. **Setting.** Our study took place in 225 villages across 9 sub-districts in the state of Odisha, India. The baseline was conducted starting December 21, 2016, the intervention on December 23, 2016, and the endline ran from December 26 to 30, 2016. It is important to note that the last day to legally deposit SBNs at bank branches was December 30, 2016.

All of our study villages have two or more hamlets, each dominated by a different caste group. Typically one hamlet consists of scheduled caste and/or scheduled tribe individuals (SCST), commonly referred to as lower caste, and the other hamlet consists of general or otherwise-backwards caste (GMOBC) individuals, commonly referred to as upper caste. The two hamlets are typically 1/2 to 1 km apart. While the primary occupation differs by caste, the majority of the people across the villages in our sample are involved in agriculture and agriculture-related activities. Given the hamlet structure of the study area, all of our treatments and and outcomes were focused in only one randomly-chosen hamlet in each village. Basic sample statistics are provided in Tables 1 and 2. 89% of individual respondents in the sample had some kind of formal bank account, 80% of respondents were literate, and major occupations included being a casual laborer (21%), domestic worker (16%), landed farmer (16%) and share-cropper (9%).

2.3. **Baseline knowledge of demonetization rules.** Using responses from our baseline survey, we first explore the beliefs of villagers about the rules prior to our intervention. While villagers almost universally understood that the Rs. 500 and Rs. 1000 notes were being taken out of circulation, we document in Table 3 that many households had inaccurate beliefs about other aspects of the policy. For example, approximately 15% of the population thought (inaccurately) that the Rs. 10 coin was also being taken out of circulation with the policy;[10] 25% of villagers believed (falsely) that, at the time of our baseline survey, they could still exchange notes at the bank without first depositing them into an account. Moreover, only a handful of respondents could accurately tell us the deadline for being able to exchange the demonetized notes and only 50% of respondents could tell us that the notes could be deposited at post offices/RBI offices/village government offices. Our subjects were particularly uninformed about some of the economically important details, such as the weekly withdrawal limits from banks. 33% of respondents reported that they did not know the limit, and in total, only 22% of respondents could tell us the correct answer (Rs. 24,000). Respondents also had very poor knowledge about limits on ATM withdrawals (10% accuracy) and withdrawal limits on the low documentation *Jan Dhan* accounts used by the poor (13% accuracy). It is also important to note that the low levels of knowledge are not due to limits to financial inclusion in the study setting. As noted before, in our sample, 89% of respondents' households had bank accounts (Table 2).

One might ask if it is important for poor rural farmers with limited formal savings to understand minute details of the policy. However, one important implementation problem associated with demonetization was that there simply were not enough notes to meet demand, which ended up affecting the lives of most people. For example, employers were not able to pay cash wages on time, microfinance borrowers were not able to service their loans, and demand for cash purchases at small shops fell. Even for individuals without bank accounts, understanding the rules would have been useful for deciding whether to accept an IOU from an employer or customer, for example, or how much new supplies to order. Measuring these types of effects is beyond the scope of our study.

2.4. **Experimental design.**

Sample. We enumerated an initial list of 276 villages which were assigned to treatments. We conducted our experiment in one hamlet in each village in that sample; half of the villages were randomly assigned to have their GMOBC hamlet in our

[10]This specific rumor spread across much of the country and was reported in the Indian press (e.g., http://www.thehindu.com/news/national/tamil-nadu/Rs.10-coins-pile-up-as-rumours-take-toll/article16966261.ece).

experiment and the other half to have their SCST hamlet in our experiment. We randomized villages to treatments before we verified that each village met our criteria. Any hamlet that had fewer than 20 households was dropped from the sample, yielding a set of 221 villages. Sixteen villages were then added in a new subdistrict to increase the sample to 237.[11] A baseline survey was administered only in the chosen hamlets described above. Given the rush of implementing 200+ interventions in a matter of days, some field errors were made. Endline data was not collected in 6 villages and the intervention did not happen in 5 villages (we also did not collect endline data there). In two villages, the elders refused entry to our surveyors. Ultimately, we have a sample of 225 villages that were treated and received endline surveys.[12]

Before we describe the treatments, it is important to note that the baseline survey also contained a module based on Banerjee et al. (2016) ("the gossip survey") to identify the individuals in each treatment hamlet that were assessed by others to be good at spreading information.[13,14]

Treatments. All of our experimental treatment arms involved distributing pamphlets with information about demonetization to the study villages. Our goal was to spread the official policy rules, and thus all information came from the RBI circulars released up until December 19th, 2016. We took this official information, published by the central bank, and subdivided it into 30 distinct policy rules. As we implemented our experiment over the last week before the December 30 deadline, the rules that we provided did not change over the course of our experiment. Through informal conversations in pilot villages, we also identified the 10 most useful rules for a typical

[11]Online Appendix K repeats our main analysis dropping these new villages and shows that our conclusions remain the same.

[12]Unfortunately, also due to the intense time pressure, in 16 of the villages our field team administered the intervention and endline to the wrong hamlet. While this should be idiosyncratic and orthogonal to treatment, we collected outcome data in the right hamlet and we redo our estimation using treatment assignment as instruments for treatment in Online Appendix J. All our results look nearly identical.

[13]We asked each individual "If we want to spread information about the money change policy put in place by the government recently, whom do you suggest we talk to? This person should be quick to understand and follow, spread the information widely, and explain it well to other people in the village. Who do you think are the best people to do this for your hamlet?" and we allowed them to nominate anywhere from 0 to 4 individuals. The results reported in Banerjee et al. (2016) show that this methodology identified the best people in the village to spread information—informing gossips led to three times as many people being reached as informing random people or informing prominent people.

[14]13 villages were dropped before information was even delivered because they were inaccessible to the survey staff. We show in Online Appendix M that this was not differential by treatment status.

villager in the study area.[15] Our experimental protocol involved giving a randomly-selected set of facts to each village—below we describe exactly how the selection is done. All individuals receiving lists of facts in a village received the same list.

Our core design is a 2 $\times$ 2 that varies how many people got information as well as whether there was common knowledge. Because another important dimension for information policies is the volume of information given to each individual, we added an arm varying whether individuals received long or short lists of facts. Prior work has shown that more information can overwhelm individuals and harm learning and choice quality (Carvalho and Silverman (2017), Beshears, Choi, Laibson, and Madrian (2013), Abaluck and Gruber (2011)), so we wanted to examine whether similar effects would be present in our social learning setting.

Thus, the treatments are:[16]

(1) Information dissemination:
- *Broadcast:* information was provided to all households in the hamlet.[17]
- *Seed:* information was provided to 5 seed households in the hamlet, chosen via the gossip survey.[18]

(2) Common knowledge:
- *No common knowledge:* we did not tell any subject that we were providing information to anyone else in the community.
- *Common knowledge:* we provided common knowledge of the information dissemination protocol. In "Broadcast" treatments in arm (1), every pamphlet contained a note that all other households received the same pamphlet. (Thus, if subjects understood and believed us, they had common knowledge of the pamphlet's distribution.) In the "Seed" treatments, every household received a notification that five individuals in their community (who were identified) were provided information about demonetization by us, and that the seeds were informed that we would inform everyone. Figure 1 summarizes the design.

[15]For example, one rule explained how foreigners could exchange their SBNs. This was not one of the "useful" facts on our list.

[16]We also attempted to get data from 30 villages where we did not intervene whatsoever and instead only collected endline data. We call these "status quo" villages. Unfortunately, these villages are not entirely comparable to our core set due to implementation failures that led to violations of randomization. We detail this in Online Appendix L.

[17]Pamphlets were dropped off at every household.

[18]Pamphlets were dropped off at each of these households. Households were not told that they were chosen for any particular reason.

(3) Information volume:
- *Long:* 24 facts were provided.
- *Short:* 2 facts were provided.

  The Short lists of facts contained one of the 10 "useful" facts, drawn uniformly at random, and a second fact drawn uniformly at random from the remaining 20, while the Long lists of facts were drawn uniformly.[19]

Appendix B provides the total list of facts from which we selected the list for each pamphlet, and Appendix C provides examples of the pamphlets we handed out.[20]

2.5. **Outcomes.** We have three main outcomes of interest at endline: engagement in social learning; general knowledge about facts surrounding the demonetization; and whether the respondent selected the demonetized Rs. 500 note as opposed to an IOU payable in 3-5 days for either Rs. 200 in non-demonetized notes or Rs. 200 in *dal*, a staple commodity.

First, we collected data on the volume of conversations about demonetization. This allows us to see whether engagement in social learning increased or decreased based on the signal distribution and knowledge structure provided in the treatment arm.

Second, we assess knowledge of facts surrounding demonetization. We survey the respondent on 34 facts and create a simple metric of knowledge.

Third, we offered subjects a choice between: (a) a Rs. 500 note; (b) an IOU to be filled in 3-5 days for Rs. 200 in two Rs. 100 notes; (c) an IOU to be filled in 3-5 days for Rs. 200 worth of dal. With a 1/6 probability, subjects actually received the item they chose. To implement the payment, we returned to each household in the sample before exiting the village, rolled the die, and provided either the Rs. 500 or the IOU notice.[21] The reason for using the IOU, which obviously relied on the villagers trusting us, was to make sure that the villagers did not go for the lower amount because they could get it right away, rather than after going to the bank. We nevertheless worried about the cost of going to the bank and depositing the 500 rupee note into an account. As noted already, 89% of respondents had bank accounts. We also collected data about the actual cost of going to the bank (see Table 2): based

[19]Thus, on average, in the Long treatment, 8 facts were useful. In the Short treatment, at least one fact was always useful, and the additional fact was useful with probability 1/3.

[20]Appendix G contains a version of our main analysis, looking separately at the endline knowledge of useful facts, facts that were reported in that particular village, and facts that were omitted from that village's pamphlets.

[21]In practice, we surprised the respondents by paying the cost of going to the bank for them by giving them the value in non-demonetized notes (Rs. 100 notes). Note that this was our last action before we exited the village; it occurred after each subject had already locked in their responses.

on the data we collected, the median wait times at banks was 10 minutes in the area and the median village in our sample was at about 20 minutes of a bank by foot.[22] At the time of our experiment, depositing the bill required no documentation of the source of the cash. Thus, selecting Rs. 200 or the equivalent was giving up more than one day's wages, even accounting for the travel to and time at the bank. We argue that this is evidence of confusion and measures a willingness to pay to avoid holding on to the demonetized note in a period where it was both legal and easy to convert. Further, we asked respondents who did not choose the Rs. 500 to provide an open-ended justification for their choice at the end of the survey module. Figure 3 shows that most individuals who did not choose the Rs. 500 note believed, mistakenly, that the deposit deadline had already passed. The choice between 200 rupees and the equivalent in *dal* was intended to capture general trust in paper currency and confusion about whether the 100 rupee bills had also become demonetized. Taking the money offered more flexibility, since *dal* was easy to buy in village stores.[23]

3. Endogenous Participation in Social Learning

Our focus is mainly on the the demand for information: the decisions people make about whether or not to engage in conversations about a topic of interest. As discussed in the introduction, this focus is motivated by our pilot surveys, in which many villagers reported that they were reluctant to seek information from others. The model we develop here is related to recent work by Bursztyn et al. (2016) and especially Chandrasekhar et al. (2017) on how reputational or signaling incentives affect information-seeking. Indeed, the formal framework we use is based on the latter paper. The idea of the model is very simple: high-ability people are better at interpreting any informational signals they get than low-ability people. As a result, conditional on already having a signal, they value additional conversation and clarification on that topic less than the low-ability people. Suppose now that it is common knowledge that most people received a signal. Then seeking information increases the likelihood in the public eye that a seeker has low ability, and that makes it costly for some people to ask questions. As a result, in that situation, fewer people will seek out information, which, *ceteris paribus*, reduces the quality of information aggregation.

[22]At this time, there were still news reports of very long queues at banks and ATMs in other, more urban parts of the country. In our study area, the waits had become much more manageable compared to the weeks following the policy announcement. Nevertheless, we were concerned that the villagers' perceived wait time could be very large. Our survey data showed that this was not the case—the median perceived wait time was 15 minutes, which was consistent with the reality.

[23]We explore this further in Online Appendix G.

Conversely, when most people are not expected to have a signal, seeking information is not informative about the seeker's ability, and therefore there will be more seeking.

After presenting the basic theory, we interpret our experimental treatments through the lens of this model and derive predictions about how the treatments affect engagement and learning. For the simplest analysis of the incentives to seek, we focus on one agent's decision of whether or not to seek information. In this two-person model, many important considerations—such as information aggregation across multiple individuals—are not modeled explicitly. In Section 3.6, we consider some alternative models to show why the phenomena we highlight are difficult to explain without endogenous seeking decisions, even when we do allow for the complexities that an explicit model of network structure entails.

In addition to the demand for information, it is natural also to consider the supply side: how much effort the initially informed invest in communicating. We discuss such considerations—which can be analyzed using standard public goods or free-rider models—in Appendix E.1. There, the associated predictions are briefly presented and compared with those of this section.

3.1. **The environment.** Consider a set N of agents (the village). The model focuses on the choice of a single decision-maker, $\mathrm{D} \in N$ of whether to seek or not.

3.1.1. *Timing.* The timing of the interaction is as follows:

(1) (a) The policymaker privately chooses a breadth of dissemination

$$\mathfrak{b} \in \{\text{Broadcast}, \text{Seed}, \text{None}\}.$$

The prior probability of breadth $\mathfrak{b}$ is $\beta_{\mathfrak{b}} \in (0,1)$. Conditional on $\mathfrak{b} =$ Broadcast, all members of the village N receive facts. Conditional on $\mathfrak{b} =$ Seed, a nonempty, proper subset S of individuals is randomly drawn to be informed.

(b) The policymaker sends a public signal (which reaches all members of N)

$$\mathfrak{p} \in \{\text{CK:Broadcast}, \text{CK:Seed}, \text{No CK}\}.$$

When a "CK:$\mathfrak{b}$" announcement is made, it is always the case that the breadth is in fact $\mathfrak{b}$. If no "CK:$\mathfrak{b}$" signal is sent, that is necessarily common knowledge; we call that outcome the No CK signal, which, practically, is an absence of such a public announcement. Under breadth $\mathfrak{b}$, the probability of a CK:$\mathfrak{b}$ announcement is $\chi_{\mathfrak{b}} \in (0,1)$.

(2) If $\mathfrak{b} \in \{\text{Broadcast}, \text{Seed}\}$, then with certainty the facts mechanically reach the Town Square.

(3) The decision-maker, $\mathrm{D} \in N$, privately learns his incremental value of getting additional information beyond the facts he received. He then decides whether to go to the Town Square to seek information about the facts delivered. D's decision is denoted by

$$d \in \{\text{NS (Not Seeking)}, \text{S (Seeking)}\}.$$

(4) An Observer in the Town Square sees whether D has come to seek information, and updates his belief about D's type.[24]

A treatment in our experiment may be summarized by a pair $\mathfrak{t} = (\mathfrak{b}, \mathfrak{p})$, the breadth of dissemination and the public signal.

The interpretation of the Town Square is that there are locations in the village (a store, tea shop, etc.) where exchange of information takes place and where the local news of the day can be accessed. There, individuals interested in learning about an issue can participate in conversations about it.

This model abstracts from important forces, such as social learning outside the Town Square and the dependence of learning and signaling on others' seeking decisions. To some extent such forces can be captured in parameters of this simple model; for instance, the extent of social learning may affect the probability that information is in the Town Square. In Section 3.6, we consider some models with richer social learning.

3.1.2. *Types and payoffs.* The payoff that D experiences from seeking depends on (i) what information there is to gain by going to the Town Square, compared to the information D already has; (ii) non-learning costs and benefits of going to the Town Square, such as the cost of time or the possibility of running into a friend; (iii) reputational payoffs depending on what people may infer about D based on his decision to go to the Town Square. This subsection introduces the primitives we use to model these considerations.

We posit that D has a privately known ability type $a \in \{\mathrm{H}, \mathrm{L}\}$, with prior probabilities $\alpha_{\mathrm{H}}, \alpha_{\mathrm{L}} \in (0, 1)$, respectively.[25] We will assume these are generic.[26] Let $I_{\mathrm{D}} \in \{0, 1\}$

[24]We will discuss beliefs about D's type more below.

[25]The ability random variable is independent of all others in the model except those defined below that explicitly condition on it.

[26]That is, drawn from a measure absolutely continuous with respect to the Lebesgue measure.

denote whether D has received facts from the policymaker. This occurs if $\mathfrak{b}$ = Broadcast or if $\mathfrak{b}$ = Seed and $\mathrm{D} \in S$. Let $I_{\mathrm{T}} \in \{0, 1\}$ denote whether there is information in the Town Square. The information is present ($I_{\mathrm{T}} = 1$) when $\mathfrak{b}$ = Broadcast or Seed, and absent otherwise.

With this notation in hand, we introduce quantities capturing (i) and (ii) above: the direct (i.e., non-reputational) payoffs of Seeking and Not Seeking. The random variable $V^{(I_{\mathrm{D}},I_{\mathrm{T}})}(\mathrm{S})$ is the direct payoff of Seeking when the informational states are $(I_{\mathrm{D}}, I_{\mathrm{T}})$, while $V^{(I_{\mathrm{D}})}(\mathrm{NS})$ is the direct payoff (which can be positive or negative) of not seeking when the seeker's information is I_{D}. The realizations of these V quantities for all their arguments—$\{V^{(I_{\mathrm{D}},I_{\mathrm{T}})}(\mathrm{S})\}_{I_{\mathrm{D}},I_{\mathrm{T}}}$ and $\{V^{(I_{\mathrm{D}})}(\mathrm{NS})\}_{I_{\mathrm{D}}}$—are known to D at stage (4), the time he makes his decision.

The following random variable, whose prior distribution we call $F_a^{(I_{\mathrm{D}},I_{\mathrm{T}})}$, represents the incremental direct payoff gain to seeking:

$$\Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})} := V^{(I_{\mathrm{D}},I_{\mathrm{T}})}(\mathrm{S}) - V^{(I_{\mathrm{D}})}(\mathrm{NS}) \sim F_a^{(I_{\mathrm{D}},I_{\mathrm{T}})}. \tag{3.1}$$

Crucially, the V random variables, and hence the random variable $\Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})}$, have distributions that depend on D's ability type. Because of this, if seeking decisions provide information about $\Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})}$, they can signal D's ability.

Perception payoffs. In addition to the direct payoff, D receives a reputational, or *perception*, payoff. If D chooses to seek and goes to the Town Square, this choice will be observed by some other villagers, who may make inferences about D's ability.

For a simple model of how D values others' assessment of him, we posit that, in the Town Square, there is an agent called the Observer (O), drawn uniformly at random from the village. This Observer sees D's decision of whether to seek or not. Because this person is also in the village, she has her own information, a realization I_{O}. (Thus, for example, when a broadcast has disseminated information to everyone, the Observer has received the information, too.) We assume D does not know in advance who may observe his decision to seek, and therefore does not condition the seeking decision on the realized identity of the Observer. The perception payoff enters D's utility function additively, as a term

$$\lambda \mathbb{P}(a = \mathrm{H} \mid d, \mathfrak{p}, I_{\mathrm{O}}),$$

where λ is a positive number. Note that the Observer is conditioning on everything she knows: the decision he observes D taking, the public signal, and the Observer's own information about the state. The idea behind the perception payoff is that D is better off when other villagers assess D's ability to be high—for example, because in that case those villagers are more likely to collaborate with D later.[27]

D's total payoff given seeking decision d is, therefore,

$$u^{(I_{\mathrm{D}},I_{\mathrm{T}})}(d) = V^{(I_{\mathrm{D}},I_{\mathrm{T}})}(d) + \lambda \mathbb{P}(a = \mathrm{H} \mid d, \mathfrak{p}, I_{\mathrm{O}}). \tag{3.2}$$

It will be useful to write the difference

$$u^{(I_{\mathrm{D}},I_{\mathrm{T}})}(\mathrm{S}) - u^{(I_{\mathrm{D}},I_{\mathrm{T}})}(\mathrm{NS}) = \Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})} - \lambda\Pi \tag{3.3}$$

where $\Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})}$ is defined in (3.1) and

$$\Pi = \mathbb{P}(a = \mathrm{H} \mid d = \mathrm{NS}, \mathfrak{p}, I_{\mathrm{O}}) - \mathbb{P}(a = \mathrm{H} \mid d = \mathrm{S}, \mathfrak{p}, I_{\mathrm{O}}). \tag{3.4}$$

D will take expectations over the perception payoffs in making his decision. In turn, the posterior belief that other villagers have about ability is endogenous: it depends on the seeking behaviors for both types, which depend on their payoffs. This leads us to an examination of the equilibria of the game.

3.2. **Equilibrium: Definition and basic observations.** We study a Bayesian equilibrium of this game. A strategy of D determines beliefs of the Observer—i.e. $\mathbb{P}(a = \mathrm{H} \mid d, \mathfrak{p}, I_{\mathrm{O}})$—for both values $d = \mathrm{S}, \mathrm{NS}$.[28] That, in turn, determines D's incentives, since he cares about perceptions.

A strategy for D is a map that gives a decision d as a function of the tuple of all realizations D knows at the time of his decision—ability a, public signal $\mathfrak{p}$, own information state I_{D}, and the values $V^{(I_{\mathrm{D}},I_{\mathrm{T}})}(d)$ across decisions d and pairs $(I_{\mathrm{D}}, I_{\mathrm{T}})$. However, the decision can actually be simplified: in any rational strategy, D will seek if and only if his expectation of his direct gain $\Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})}$ exceeds his expectation of the perception benefit of not seeking, Π, which in equilibrium is a known number.[29]

An equilibrium strategy is characterized by these conditions: (i) D seeks if and only if his expectation of $\Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})}$ is at least his expectation of $\lambda\Pi$; (ii) the beliefs about ability a in (3.4) are consistent with (i) and Bayes' rule.

If each $F_a^{(I_{\mathrm{D}},I_{\mathrm{T}})}$ has no atoms—an assumption we will maintain—then an equilibrium can be described essentially completely by specifying a cutoff for D to seek:

[27]Foundations for this assumption are discussed in Chandrasekhar et al. (2017).

[28]As usual, the equilibrium can be given a population interpretation: there is a population of D's, who have different draws of private information, and the Observer is inferring the attributes of a particular D in view of the population's behavior.

[29]D's decision does not depend on his private ability type a. The reason is as follows: Given $\Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})}$, D's ex post direct gain to seeking, (3.3), does not depend on his private ability type. Because his ability type is unobservable, the reputational payoff cannot depend on it, either.

how high D's expected value of $\Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})}$ has to be in order to choose $d = \mathrm{S}$. The cutoff, which we call $\underline{v}(\mathfrak{p}, I_{\mathrm{D}})$ only depends on the public signal $\mathfrak{p}$ and on I_{D} and, as a function of these, it is commonly known in equilibrium.[30]

3.3. **Assumptions.**

3.3.1. *Payoffs.* We now discuss assumptions on the distribution of $\Delta(\cdot,\cdot)$. First, for technical convenience, we will maintain the assumption that the support of $F_a(I_{\mathrm{D}}, I_{\mathrm{T}})$ includes the positive reals, for all values of a and $(I_{\mathrm{D}}, I_{\mathrm{T}})$.

Next, we make assumptions on how different abilities value information.

P1 (a) For any $(I_{\mathrm{D}}, I_{\mathrm{T}})$, the distribution $F_{\mathrm{L}}^{(I_{\mathrm{D}},I_{\mathrm{T}})}$ first-order stochastically dominates $F_{\mathrm{H}}^{(I_{\mathrm{D}},I_{\mathrm{T}})}$.

A low-ability D always has at least as much to gain from seeking as a high-ability one, all else equal.

(b) For all values of I_{T}, the ratio $\frac{1-F_{\mathrm{L}}^{(I_{\mathrm{D}},I_{\mathrm{T}})}(v)}{1-F_{\mathrm{H}}^{(I_{\mathrm{D}},I_{\mathrm{T}})}(v)}$ is strictly increasing in I_{D} for any v.

For any cutoff, having a value of information above that cutoff signals low ability more when D is informed ($I_{\mathrm{D}} = 1$) than when D is not informed ($I_{\mathrm{D}} = 0$).

Assumption P1(a) reflects that a low-ability D needs more help to figure out the content of information. It ensures that seeking is (weakly) a signal of low ability, because for any cutoff D uses, the low-ability type is (weakly) more likely to exceed it. Assumption P1(b) imposes some structure on that signal, as described above.

Our next assumption imposes structure on how the informational states of D and of the Town Square affect the payoffs of seeking.

P2 (a) $F_a^{(I_{\mathrm{D}},1)}(v) < F_a^{(I_{\mathrm{D}},0)}(v)$ for all $v \geq 0$ and all values of a and I_{D}.

Regardless of ability and own signal, seeking is (in the stochastic sense) strictly more beneficial when there is information in the Town Square.

(b) $F_a^{(0,1)}$ first-order stochastically dominates $F_a^{(1,1)}$ for both values of a.

The direct benefit of seeking is weakly greater when one is uninformed, assuming there is information in the Town Square.

Our final assumption is for technical convenience.

P3 For any $(I_{\mathrm{D}}, I_{\mathrm{T}})$, the ratio $\frac{1-F_{\mathrm{H}}^{(I_{\mathrm{D}},I_{\mathrm{T}})}(v)}{1-F_{\mathrm{L}}^{(I_{\mathrm{D}},I_{\mathrm{T}})}(v)}$ is strictly decreasing in v for all $v \geq 0$.

[30]We make the innocuous tie-breaking assumption that the seeker seeks if and only if $\Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})} \geq \underline{v}(\mathfrak{p}, I_{\mathrm{D}})$.

This is a regularity condition on the distribution of values of seeking which is satisfied if, for example, F_{L} and F_{H} are stochastically ordered normal distributions centered to the left of zero. Economically, this means that the higher is the cutoff for seeking, the worse is the inference about D's ability if D chooses to seek. This condition is useful because it enables us to use the techniques of monotone comparative statics to study how $\underline{v}(\mathfrak{p})$, the cutoff for seeking, varies across treatments.

3.3.2. *Beliefs.* In our description of the timing of the game, we did not make any assumptions about how S, the set of seeded individuals, is drawn. We now make two assumptions on individuals' beliefs that restrict this distribution, which we will need in some, but not all, of our results.

- B1 For any $i \in N$, the probability $\mathbb{P}(i \in S)$ is between $1/n$ and $\overline{k}/n$ for some constant $\overline{k}$.
- B2 For any two individuals i and j, there is a constant C so that the conditional probability $\mathbb{P}(i \in S \mid j \in S, \mathfrak{b} = \text{Seed})$ is at most $C\mathbb{P}(i \in S \mid \mathfrak{b} = \text{Seed})$.

These assumptions say that there are not too few or too many seeds, and from the perspective of any j, individual i's membership in the seed set S is not too correlated with j's own.

3.4. **Dependence of seeking rates on treatment.** In general the model may have multiple equilibria.[31] However, under our assumptions (the key one being P3) the game has some nice structure. In particular, as the cutoffs[32] $\underline{v}(\mathfrak{p}, I_{\mathrm{D}})$ increase, incentives to seek decrease monotonically for all realizations of private information. (This occurs because, loosely speaking, seeking becomes a worse signal.) Because the resulting game of incomplete information then has a supermodular structure, we can identify an equilibrium that has maximum seeking in a strong sense: for every realization of D's private information, there is more seeking in that equilibrium than in any other. This equilibrium will always be stable under best-response dynamics, and call this the *maximum equilibrium.*[33]

Let $s(\mathfrak{t})$ be the probability, in the maximum equilibrium, that D chooses $d = \mathrm{S}$ (Seeking) in treatment $\mathfrak{t} = (\mathfrak{b}, \mathfrak{p})$—for example $\mathfrak{t} =$ (Seed, CK:Seed). This is an ex ante probability: we integrate over all ability types, information realizations, etc. We

[31]For more on this multiplicity, see Chandrasekhar et al. (2017).

[32]Introduced in Section 3.2 above.

[33]Making another selection, such as the *minimum* equilibrium, which also exists, would not change the analysis or the results. Of course, this selection point is moot if equilibrium is unique; conditions for uniqueness are available upon request.

focus on this statistic because it is one that is observed in our experiments. Now we can state the two main propositions yielding our predictions.

PROPOSITION **1.** Under the assumptions of Section 3.3.1:

(a) $s(\text{Broadcast, No CK}) > s(\text{Broadcast, CK})$;

(b) $s(\text{Seed, CK}) > s(\text{Broadcast, CK})$.

The proof of this and all other propositions appears in Section D.1 of the Appendix. We give the key ideas of the argument in the next subsection.

The second proposition relies on assumptions about beliefs, ranking the amount of communication in the Seed treatments.

PROPOSITION **2.** Under the assumptions of Sections 3.3.1 and 3.3.2, and assuming $\overline{k}/n$ is small enough, it holds that $s(\text{Seed, CK}) > s(\text{Seed, No CK})$.

Finally, the prediction that requires the most structure is:

PROPOSITION **3.** Take the assumptions of Sections 3.3.1 and 3.3.2 and, fixing all other parameters, suppose the following three quantities are small enough: (i) $\overline{k}/n$; (ii) β_{Seed}; and (iii) $\frac{1-\chi_{\text{Broadcast}}}{(\overline{k}/n)^2}$. Then $s(\text{Broadcast, No CK}) > s(\text{Seed, No CK})$.

3.4.1. *Intuition behind the Propositions.* We now explain the key forces behind each of the main predictions entailed in the propositions above.

Proposition 1

(a) (Broadcast, No CK) has more seeking than (Broadcast, CK). In both cases, D's assessment of direct payoffs is the same: since $I_{\text{D}} = 1$, D knows that $I_{\text{T}} = 1$. In the (Broadcast, CK) treatment, O is certain that D is informed, and D knows this. It is in that case that signaling concerns are the strongest they could be, by Assumption P1(b). In (Broadcast, No CK) the signaling effect is weaker, because some probability is placed on D not being informed. Thus, there is more seeking under (Broadcast, No CK).

(b) (Seed, CK) has more seeking than (Broadcast, CK):

Considering the signaling contribution to payoffs: for any given cutoffs, we can write the beliefs of the Observer conditional on $d = \text{S}$ (given either value of $\mathfrak{p}$) as a convex combination over values of I_{D}. The term corresponding to $I_{\text{D}} = 1$ is the same across the two treatments. This is the only term with a positive weight in the (Broadcast, CK) treatment. The term corresponding to $I_{\text{D}} = 0$ involves a weakly greater posterior that $a = \text{H}$ by Assumption P1. Thus, signaling concerns are smaller in (Seed, CK).

Turning now to the direct payoffs, $I_{\mathrm{T}} = 1$ is known in both cases. By Assumption P2(b), the value of seeking is greater for the uninformed, who are at least as prevalent in the Seed treatment. Thus, direct payoffs are greater there.

Proposition 2

First, under (Seed, CK), D is certain that information is in the Town Square, which by P2 shifts up the expected direct value of seeking relative to (Seed, No CK) by at least some positive amount. Now we turn to signaling concerns. Condition on $I_{\mathrm{D}} = 0$ (which is the case with high probability under Seed, since $\overline{k}/n$ is small by assumption). In this case, D is nearly certain that O is uninformed. Conditioning on $I_{\mathrm{O}} = 0$, by the same token, O is nearly certain that D is uninformed. Thus, signaling concerns are very similar to the case in which it is common knowledge that D is uninformed.

Proposition 3

For the argument behind Proposition 3, we need a lemma, which we state somewhat informally. It follows immediately from Bayes' rule.[34]

LEMMA **1.** Under the assumptions of Section 3.3.2, suppose that $(1 - \chi_{\text{Broadcast}})$ is small enough relative to $(\overline{k}/n)^2$. Then conditional on $\mathfrak{p} =$ No CK and any realizations of I_{D} and I_{O}, the probability that $\mathfrak{b} =$ Broadcast is negligibly small.

Now we can establish the proposition. Concerning the direct benefit: in (Seed, No CK), when D receives no information ($I_{\mathrm{D}} = 0$), the fact that β_{Seed} is small means that his expectations approximate those when $I_{\mathrm{T}} = 0$. In contrast, in (Broadcast, No CK), given that $I_{\mathrm{D}} = 1$, the breadth $\mathfrak{b}$ is in $\{\text{Broadcast}, \text{Seed}\}$ (i.e., not equal to "None") and information is certain to be in the Town Square ($I_{\mathrm{T}} = 1$). By Assumption P2, seeking is more valuable in this case.

Turning now to signaling concerns, the key step is to rule out the possibility that the observer under (Broadcast, No CK) assumes that since he has a signal, so does everyone else (i.e. the state is Broadcast). This is where we make use of the face that because there is no public announcement, by Lemma 1, O will be nearly certain that $\mathfrak{b} \neq$ Broadcast. Because $\overline{k}/n$ is small, he will also be nearly certain that D is not a seed. To sum up, O will believe I_{D} holds with high probability. Thus, signaling concerns are therefore almost the same in the two cases.

[34]Consider an observer who knows that $I_{\mathrm{D}} = I_{\mathrm{O}} = 1$ and that $\mathfrak{p} =$ No CK. His posterior likelihood ratio that $\mathfrak{b} =$ Broadcast has occurred versus $\mathfrak{b} =$ Seed is of order $(1 - \chi_{\text{Broadcast}})/(\overline{k}/n)^2$. Thus if this is small, then even this observer will consider Broadcast unlikely.

The proof formalizes these ideas using monotone comparative statics.

Comments on modeling choices. We close this subsection with some brief comments on our modeling choices. One choice we make is to assume that the Observer is not the source of the information that is available in the Town square. An alternative would have been to have the person asked for information to also be the Observer, thus merging the roles of the source T and O. However, this raises a variety of challenging modeling decisions: do we explicitly model the aggregation of information by this person? What if she herself is unable to process the signal she received? How are signaling concerns affected by the fact that she may be able to infer, based on the number of people coming to her, what the $(\mathfrak{b}, \mathfrak{p})$ realization is? Another direction would be to more realistically model a Town Square where there are many different people, and now the information D gets is obtained by talking to a member of this population, drawn according to some distribution. Aggregation of information in the Town Square would now have to be modeled explicitly, which presents considerable complications; there will also be potential for bilateral signaling, both by Seekers and Advisers. Our modeling abstracts from these complications to get at what we believe are the essential phenomena, though models addressing these richer concerns may be interesting in their own right.

3.5. **Knowledge and choice quality in equilibrium.** Propositions 1 and 2 focus on the rates of seeking—which, in the experiment, we measure by the amount of conversation. But our experiments also consider other outcomes: knowledge about demonetization and choice quality. To study these using our theory, we analyze the expected direct payoff

$$p(\mathfrak{t}) = \mathbb{E}[V^{(I_{\mathrm{D}}, I_{\mathrm{T}})}(d) \mid \mathfrak{t}]$$

in a given treatment $\mathfrak{t}$. This is the value of information gross of signaling concerns. Again, it is pooled over ability types and information realizations. Consider the comparisons of Propositions 1 and 2. When I_{D} is held fixed, the rankings are just as in that proposition:

Corollary **1.** Under the conditions of Proposition 2,

(a) $p(\text{Broadcast, No CK}) > p(\text{Broadcast, CK})$
(b) $p(\text{Seed, CK}) > p(\text{Seed, No CK})$

Note that in both (a) and (b), D's information endowment is the same. In (a), the proof of Proposition 1 shows that the direct value is the same on both sides of the inequality, while the signaling concerns are smaller on the left-hand side, furnishing the conclusion. In (b) the proof of Proposition 2 shows that the signaling concerns are no greater while the incremental value of information is appreciably higher.

When the comparison of two given treatments also involves changes in I_{D}, the comparisons are not as immediate. However, we will now discuss, somewhat informally, what is needed for the remaining rankings of knowledge and decision quality to parallel those that were derived for s above:

- $p(\text{Seed, CK}) > p(\text{Broadcast, CK})$
- $p(\text{Broadcast, No CK}) > p(\text{Seed, No CK})$ under the assumptions of Proposition 3.

For the first item, let us consider how the inequality could possibly be reversed relative to the corresponding item in Proposition 1. For a reversal, it would have to be that the base level of knowledge possessed by agents in (Broadcast, CK) is enough to make them better off even if signaling concerns deter seeking. The reversal would therefore *not* happen if we assume: (a) low-ability types who don't seek make decisions approximately as if they were uninformed, and (b) there are enough low-ability types. In that case, seeking rates become pivotal to the welfare of enough of the population; knowledge and choice quality then move in tandem with seeking rates.

The condition needed for the second ranking is similar. If we assume that β_{Seed} is small, then, as we argued in Proposition 3, the expected incremental direct benefit of seeking ($\Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})}$) is very close to its expectation under $I_{\mathrm{D}} = I_{\mathrm{T}} = 0$. Under (Broadcast, No CK), it is much higher, while signaling concerns are very similar across the two cases. Thus equilibrium welfare must also be higher for those types who need to seek in order to do better than their uninformed welfare.

3.6. **Benchmarks from models without seeking frictions.** We close this section by arguing that the predictions coming out of our endogenous social learning model above are not consistent with some benchmark network communication models that do not feature endogenous seeking frictions.

3.6.1. *Tagged information transmission.* The first network learning model we look adapts those of Acemoglu et al. (2014), Möbius et al. (2015), and others. We present it informally here and defer the details to Appendix E.2. In brief, there is a network of communication opportunities. Initially, agents are endowed with some information—their understanding of the facts we give them, and any information about demonetization they may have otherwise. Each time period, they have opportunities to talk to others, realized randomly. When they talk, they convey a message and its original source: this is the essence of the tagging model, where the deck is stacked in terms of aggregating information correctly. This extreme assumption abstracts away from the complex issues of how players might make inferences from reports that did not track source information. (We reconsider this simplification below when we discuss another class of models.)

Importantly, in the tagging models, information aggregation at any given moment needn't be complete. Because of randomness in communication opportunities and dropped messages, a given individual may not have access to all signals received in the community, or even in his neighborhood. However, the following is a general result. Suppose initial endowments of information improve, in the sense that they become Blackwell more informative about the state of interest. Then, after the aggregation process, each individual has better information. In particular, each individuals' decisions about anything determined by the state will be better in expectation after the change.

In terms of interpretation, this means that making more agents informed, or increasing the amount of information given to each individual, can only improve aggregate outcomes. Common knowledge had no role to play in the story above. To look at the case where it *can* have such a role, take the model of Acemoglu et al. (2014), which is essentially the tagged model along with endogenous decisions of whether to drop out of the social learning process or stay engaged in hopes of learning more. There, social learning is *improved* by making it public that many agents are informed, because it increases the amount of information that any one of them can expect to receive by a given time. The essential reason is the strategic complementarity between the engagement of different agents.

To summarize, a standard class of models without aggregation frictions predicts that endowing the community with more information will be reflected in better individual decisions, and that common knowledge should also help.

Evidence. We document in Table 4 that, contrary to the predictions of the benchmark model sketched above, there is no detectable beneficial effect of informing more people, or giving them more information. Panel A shows that more information per pamphlet does not lead to more conversations or better outcomes. Providing a 12-fold increase in the number of facts leads to a 26% decline in the number of conversations, no change in knowledge, and no change in the probability of picking Rs. 500. Panel B shows that broadcasting information to 100% of households instead of 10% leads to no change in either the number of conversations, knowledge, or in the probability of picking Rs. 500.

Thus providing a greater amount of information to each person does not lead to greater knowledge in the population.[35] More strikingly, when we provide information to ten times the number of people, we do not see the expected in knowledge and or an improvement in quality of decisions made. This is despite the fact that there are low levels of knowledge on average, even among seeds, which suggests that there is considerable scope for improvement in learning in these communities.

3.6.2. *Herding models.* An extreme assumption in the types of models discussed in the previous section is that agents transmit the original sources of all the pieces of information they convey (or at least a sufficient statistic). Relaxing this assumption raises the issue of how agents make inferences from coarsened observations that do *not* track their sources of the information. A tractable way to study these difficulties is to use a sequential social learning model, which seems reasonable in our setting as agents are not likely to engage in information exchange on too many distinct occasions (as we verify in our survey data).

In general, characterizing learning quality exactly in herding models can be very difficult. However, an approach of Lobel and Sadler (2015), which applies to sequential learning in arbitrary conversation networks, can be used to argue why phenomena such as those our main model produces are unlikely to be explained by standard sequential models. We flesh out the details of the argument here in E.3.

Consider a binary decision (such as the one we gave our agents) about whether or not to accept certain denomination of currency. Individuals form opinions about this. Differences in private information lead to heterogeneity in the strengths of their beliefs. In particular, the messages an individual has received affect the strength of his posterior belief about the right action to take.

Lobel and Sadler (2015) show that in equilibrium, most agents' decisions are at least as good as those decisions taken by those who are "experts"—very sure of the right answer based on private information (i.e. their own understanding) alone. The intuition can be most easily seen in a model where all predecessors are observed: if decisions were substantially worse than the expert benchmark for arbitrarily late movers, then the well-informed would speak against the prevailing view, revealing

[35] This is consistent with Carvalho and Silverman (2017), who argue that complexity can lead to worse decision-making and can lead to individuals taking dominated options. They study this issue in the context of portfolio choice.

their superior information and persuading others. Remarkably, the same remains true even when agents observe only a few predecessors, under some conditions. The main substantive one is that the network is connected enough, with everyone having indirect access to many others.

It can be deduced from this that improving information endowments can only hurt learning if it was already quite good. In other words, the known forces from herding or information cascades will have difficulty explaining how adding information can lead to outcomes in which most people do worse than the individual decisions of the "well-informed" individuals.

4. Results

4.1. **Endogenous participation in social learning.**

4.1.1. *Volume of Conversations.*

We begin by looking at which delivery mechanisms led to more or less endogenous participation in social learning, measured by the number of conversations the subject had over the prior three days about demonetization.

Table 5 presents regressions of the number of conversations on the various treatments.[36] In each regression, (Seed, No CK) is the omitted treatment arm. The coefficients are additive, so to compare (Broadcast, Common Knowledge) to the omitted category, it is necessary to add the coefficients: CK, Broadcast, and Broadcast $\times$ CK. In each regression specification, we present the p-values throughout, with standard errors clustered at the village level, and for two additional key comparisons. The test (CK + Broadcast $\times$ CK = 0) allows us to compare (Broadcast, CK) to (Broadcast, No CK). The test (Broadcast + Broadcast $\times$ CK = 0) allows us to compare (Broadcast, CK) with (Seed, CK).

The outcome variable in column 1 is the number conversations about the demonetization that the respondent was a part of over the last three days. Going from (Seed, No CK) to (Seed, CK) increases the number of conversations by 103% (0.65 more conversations, $p = 0.04$). This is consistent with the model described above: a typical villager now knows that there is no expectation that they have the information because it is common knowledge that they did not receive signals and because the seeds are known, which emboldens them to seek information.

[36]For all of our main results, we focus on our core 2 $\times$ 2 treatment design, pooling across the Long and Short lists of facts. Appendix F provides the analysis separately for Long and Short information and also discuss how one might interpret the length of the fact list through the lens of the model.

At face value, it is also consistent with a simpler model where there is no strategic seeking and the effect is driven by the fact that the villagers know whom to ask, as well as the possibility that seeds have more of a motivation to spread information. However we do not think this is the case for two reasons. First, in Online Appendix H, we show the same results split by whether the household was a seed or not and demonstrate that our results are primarily driven by an increase in conversation volume among non-seed households, rather than by non-seed households seeking out seeds or vice versa.[37] As Table H.1 shows there is a (noisily estimated) increase of 1.3 in the conversation count for a Seed in CK relative to No CK ($p = 0.39$). If every seeded household gained 1.3 conversations, then this explains 6.5 more conversations, which is only 28 percent of the 23 extra conversations we find a village of 50 households. (Even if we assume that the true number of seed conversations is double the number implied by the coefficient—13 conversations—this at best would only explain 56% of the increase in conversations.) Second, we collected data about the nature of the conversations—whether they were the result of a directed question or statement about demonetization (purposeful) or merely something that came up in a broader conversation (incidental). These are reported in Subsubsection 4.1.2, below. They make it clear that most of the increase came from incidental conversations–in other words not from people going out to ask questions from seeds or seeds coming to deliver a message.

Next we look at what happens when we compare strategies that employ common knowledge. Going from (Seed, CK) to (Broadcast, CK), which corresponds to a 10-fold increase in the number of households informed (from 5 households to 100% of households), leads to a 61% *decline* in the volume of conversations (0.78 fewer conversations, $p = 0.029$). Again this is consistent with the model, though it could also be that because everyone is informed, there is less need for conversations. However, given how little people know (even in (Broadcast, CK) villages (see below), this seems unlikely.

When we look at (Broadcast, No CK) versus (Seed, No CK), we are comparing a situation where we provided signals to all versus just a few, but in either case no agent knows whether or not any other agent has necessarily received a signal. In sharp contrast, we find that a 10-fold increase in the number of households informed leads to an increase in the volume of conversations by 113% (0.708 more conversations,

[37]We remind the reader that every village had "seed" households selected by the same process ex ante, but in Broadcast treatments all households were treated. In Online Appendix H, Table H.1, shows that all our main results hold for the households that are not seeds.

$p = 0.048$). This makes intuitive sense: essentially with (Seed, No CK) a typical household doesn't even know that there is something to converse about, whereas that is not true with (Broadcast, No CK). Note however that this also goes against the idea that the reason why there is less seeking with (Broadcast, CK) than with (Seed, CK) is that people already have enough information. They seem to act as if they need information as long as they can hide that fact from others.

The sharpest test of our model is when we go from (Broadcast, No CK) to (Broadcast, CK). This leads to a 63% *decline* in the volume of conversations (0.84 fewer conversations, $p = 0.02$). Making broadcast common knowledge should not reduce conversations unless signaling concerns are very powerful: in that world, the fact that it is not common knowledge that one received a signal allows the agent to ask questions about demonetization more freely.

In sum, our results show that common knowledge affects considerably endogenous participation in social learning. When only a few individuals are seeded, it greatly increases aggregate conversations. We have also shown evidence for two non-monotonicities consistent with our model: first, adding common knowledge to a broadcast delivery mechanism can discourage conversations; and, second, if there is common knowledge, going from only 10% to 100% of the population being informed actually discourages conversations. As one may have expected, if there is no common knowledge, increasing the number informed increases conversations, in contrast.

4.1.2. *Types of Conversations: Purposeful and Incidental.* As mentioned above, we collected information both on the number of conversations and then the number of conversation by type: purposeful and incidental. Purposeful conversations were initiated with the sole purpose of talking about demonetization, while incidental conversations were initiated for some other purpose but then touched on the topic of demonetization. Columns 2 and 3 of Table 5 break up the number of conversations that the subject participated in by whether they were incidental (column 2) or purposeful (column 3). Incidental conversations comprise the vast majority, 78%, of reported conversations. As columns 2 and 3 make clear, our core results broadly go through for each type of conversation, but significantly more of the impact of the interventions comes from the incidental conversations. Consistent with that, column 3 of Appendix Table H.1 shows that the increase in conversations where seeds are CK do not appear to be driven by the seed actively going out to explain the information to others, nor others actively seeking out the seeds. The primary driver of information aggregation here is conversations among non-seeds, and we see no evidence of an effort by seeds to coordinate conversations about the topic.

The fact that individuals largely discuss demonetization via incidental conversations is consistent with the metaphor used in our theory–that information aggregation occurs in the Town Square. Engaging in conversations about demonetization in public places while going about one's business can be viewed as tapping into the information aggregated there.

4.2. **Information aggregation and choice.** We present regressions in Table 6 which show how knowledge of the demonetization rules and incentivized choice behavior depend on treatment cell. Recall that the quality of the choice depended on the respondent's understanding of the demonetization rules.

In column 1, we turn to whether the changes in endogenous participation in learning correspond to changes in knowledge. This is primarily an empirical question, though we provide sufficient conditions for it to happen in section 3.5. To see why an increase in conversations may not lead to an increase in learning, note, for example that even though there are fewer conversations happening in (Broadcast, CK) as compared to (Seed, CK), 10-times the number of households received information under broadcast treatments, so it is entirely possible that they still learned more. Therefore the finding that (Broadcast, CK) generates less learning than (Seed, CK) is a more powerful test of our theory than the fact that there are more conversations in (Seed, CK). If the reason why there were fewer conversations in (Broadcast, CK) is that people got enough information from their signals so that they did not need to ask questions (rather than they chose not to ask questions for fear of revealing their type, as in our theory), we would expect (Broadcast, CK) to out-perform (Seed), CK) in terms of knowledge, even if there are fewer conversations. It is only when there are important strategic reasons for not asking questions *and* social learning is an important part of gathering information, that we would expect a reversal of the ”natural” ordering, where broadcasting does better than seeding.

We find evidence for the strong reversals that our model predicts. The outcome variable is the error rate on our knowledge metric. In our metric, the (Seed, no CK) mean is 0.434. Going from seeding to broadcast leads to a 5.6% *increase* in the error rate on our knowledge survey ($p = 0.062$). This is striking and shows that though 100% of households receive information instead of 10%, the amount of aggregated information that a random household has at the end of the day is actually less, not more. Also, turning to broadcast strategies, adding common knowledge leads to a 4.6% *increase* in the error rate, though the effect is not quite statistically significant ($p = 0.174$). In addition, going from (Seed, No CK) to (Seed, CK) decreases the error rate on our knowledge survey by 7.3% ($p = 0.0142$) and going from (Seed, No CK) to (Broadcast, No CK) actually makes people better informed and reduces the error rate by 6.4% ($p = 0.05$). It is worth noting that we see reductions in knowledge exactly where we see conversations declining which strongly suggests that people do learn from each other and the reduction in conversations results in a reduction in knowledge.

In column 2, we turn to the impact of our experimental treatments on incentivized choice. We look at whether subjects choose the Rs. 500 note on the spot, which they could still deposit in their accounts, or an IOU worth Rs. 200 to be paid in 3-5 days, taking a loss of about 1.5 days wages. The probability of selecting the Rs. 500 note in the omitted category (Seed, No CK) is only 5.92%.Going from seeding to broadcast, conditional on common knowledge, leads to a 38.5% or 4.13pp *decline* in the probability of choosing the Rs. 500 note ($p = 0.104$). Looking at broadcast strategies, adding common knowledge leads to a 48% *decline* in the probability of choosing the Rs. 500 note ($p = 0.041$). In addition, going from (Seed, No CK) to (Seed, CK) leads to a 4.8pp or an 81% increase in the probability of choosing the Rs. 500 note ($p = 0.037$) but going from (Seed, No CK) to (Broadcast, No CK) corresponds to a 6.77pp or 114% increase in the probability of choosing the Rs. 500 note ($p = 0.014$).

Taken together, our results clearly demonstrate that broadcasting information is better than seeding in a world without common knowledge. However, increasing the number of informed households has opposite effects, depending whether there is common knowledge or not. In a world without common knowledge, the conventional wisdom holds: increasing the number informed encourages more conversations and better decision making. However, under common knowledge, broadcasting information actually backfires, leading to worse outcomes across the board. These results are consistent with our framework of endogenous communication. One bottom line result is that seeding just five households combined with common knowledge makes the outcomes indistinguishable from (Broadcast, No CK), where ten times as many people were seeded. And finally, and perhaps more strikingly, either holding common knowledge fixed and moving from seed to broadcast or holding broadcast fixed and moving from no common knowledge to common knowledge actually reduces conversation volume, knowledge, and quality of choice

5. Conclusion

Social learning happens in part through choices by the participants about whether to ask questions. We show that, consistent with prior lab-in-field research by a subset of us, Chandrasekhar et al. (2017), the number of signals and the structure of common knowledge matter considerably for the extent of participation in social learning. In particular we find evidence for a set of clear reversals that are consistent with our model and prima facie inconsistent with a model where there is no endogenous participation. When looking at targeted seeding, going from no common knowledge to common knowledge increases conversations but the exact opposite is true for broadcast strategies. Moreover conversations actually decline when, holding common knowledge fixed, more people are provided information. Furthermore in our setting, this increase or decline in conversation volume is met with a corresponding increase or decline in knowledge about the rules as well as quality of choice. Thus, the success of an information intervention depends crucially on the details of the design and how it affects endogenous communication.

Of the full set of experimental interventions, two consistently perform well along all the dimensions—conversations, knowledge, and choice—and have comparable benefits to one another: seed with common knowledge and broadcast without common knowledge. Note, however, that broadcast, no common knowledge is not easy to implement in a non-experimental setting. Most, if not all, broadcast technologies such as radio, television, newspaper, or the village crier intrinsically contain a common knowledge component. Moreover, it would be difficult to repeat a non-common knowledge broadcast strategy without it eventually becoming common knowledge.

The results have implications for how researchers and policymakers should think about the use of broadcast media versus extension to educate individuals, and how extension should be structured. The results indicate that the benefits of extension strategies can be magnified with common knowledge.

References

Abaluck, J. and J. Gruber (2011): "Choice inconsistencies among the elderly: evidence from plan choice in the Medicare Part D program," *American Economic Review*, 101, 1180–1210. 2.4

Acemoglu, D., K. Bimpikis, and A. Ozdaglar (2014): "Dynamics of information exchange in endogenous social networks," *Theoretical Economics*, 9, 41–97. 3.6.1

Akbarpour, M., S. Malladi, and A. Saberi (2017): "Diffusion, Seeding, and the Value of Network Information," Available at SSRN: https://ssrn.com/abstract=3062830. 1

Aral, S. and D. Walker (2012): "Identifying Influential and Susceptible Members of Social Networks," *Science*, 337, 337–341. 1

Bailey, N. (1975): *The Mathematical Theory of Infectious Diseases and Its Applications.*, Griffin: London. 1

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013): "The diffusion of microfinance," *Science*, 341, 1236498. 1

——— (2016): "Gossip: Identifying Central Individuals in a Social Network," *Working Paper*. 2.4, 13

Banerjee, A., L. Iyer, and R. Somanathan (2007): "Public action for public goods," *Handbook of Development Economics*, 4, 3117–3154. E.1

Banerjee, A. V. and E. S. Maskin (1996): "A Walrasian theory of money and barter," *The Quarterly Journal of Economics*, 111, 955–1005. G

Bass, F. M. (1969): "A new product growth for model consumer durables," *Management science*, 15, 215–227. 1

Beaman, L., A. BenYishay, J. Magruder, and A. M. Mobarak (2016): "Can Network Theory based Targeting Increase Technology Adoption?" *Working Paper*. 1

Beshears, J., J. J. Choi, D. Laibson, and B. C. Madrian (2013): "Simplification and saving," *Journal of economic behavior & organization*, 95, 130–145. 2.4

Bursztyn, L., G. Egorov, and R. Jensen (2016): "Cool to Be Smart or Smart to Be Cool? Understanding Peer Pressure in Education," . 2, 3

Cai, J., A. De Janvry, and E. Sadoulet (2015): "Social networks and the decision to insure," *American Economic Journal: Applied Economics*, 7, 81–108. 1

Carvalho, L. and D. Silverman (2017): "Complexity and Sophistication," . 2.4, 35

Chandrasekhar, A. G., B. Golub, and H. Yang (2017): "Signaling, Shame, and Silence in Social Learning," *Working paper, Stanford University*. 1, 7, 3, 27, 31, 5, E.1

Conley, T. G. and C. R. Udry (2010): "Learning about a New Technology: Pineapple in Ghana," *The American Economic Review*, 35–69. 1

Galeotti, A. and S. Goyal (2010): "The Law of the Few," *American Economic Review*, 100, 1468–1492. E.1

Jackson, M. and L. Yariv (2011): "Diffusion, strategic interaction, and social structure," *Handbook of Social Economics, San Diego: North Holland, edited by Benhabib, J. and Bisin, A. and Jackson, M.O.* 1

Jackson, M. O. (2008): *Social and Economic Networks.*, Princeton University Press. 1

Kiyotaki, N. and R. Wright (1989): "On money as a medium of exchange," *Journal of Political Economy*, 97, 927–954. G

Leskovec, J., L. A. Adamic, and B. A. Huberman (2007): "The dynamics of viral marketing," *ACM Transactions on the Web (TWEB)*, 1, 5. 1

Lobel, I. and E. Sadler (2015): "Information diffusion in networks through social learning," *Theoretical Economics*, 10, 807–851. 3.6.2, E.3

Miller, G. and A. M. Mobarak (2014): "Learning about new technologies through social networks: experimental evidence on nontraditional stoves in Bangladesh," *Marketing Science*, 34, 480–499. 1

Möbius, M., T. Phan, and A. Szeidl (2015): "Treasure hunt: Social learning int he field," NBER Working Paper No. 21014. 3.6.1

Niehaus, P. (2011): "Filtered Social Learning," *Journal of Political Economy*, 119, 686–720. 9

Ryan, B. and N. C. Gross (1943): "The diffusion of hybrid seed corn in two Iowa communities." *Rural sociology*, 8, 15. 1

Samuelson, P. A. (1958): "An exact consumption-loan model of interest with or without the social contrivance of money," *Journal of Political Economy*, 66, 467–482. G

Wallace, N. (1980): "The Overlapping Generations Model of Fiat Money, in Models of Monetary Economies (J. Kareken and N. Wallace, eds.), 49(82)," *Federal Reserve Bank of Minneapolis*. G

Figures

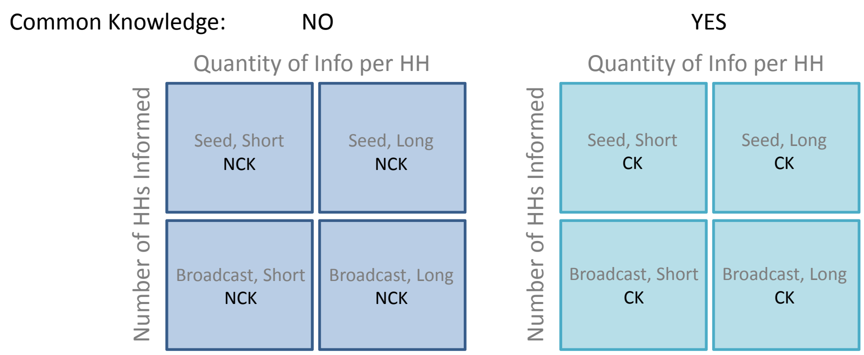


Figure 1. Experimental Design

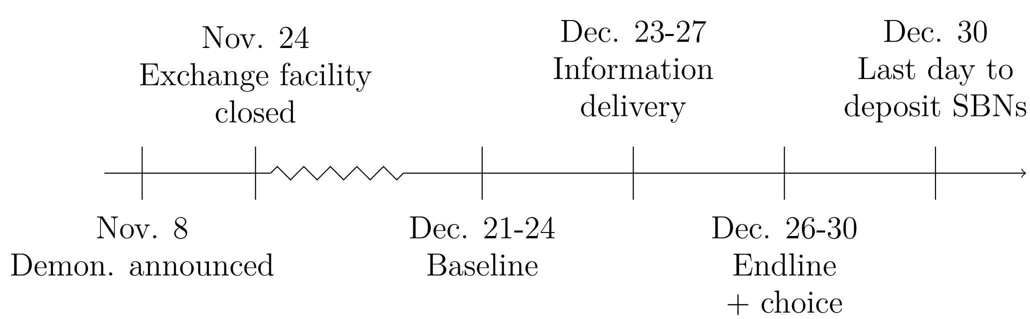


Figure 2. Intervention Timeline

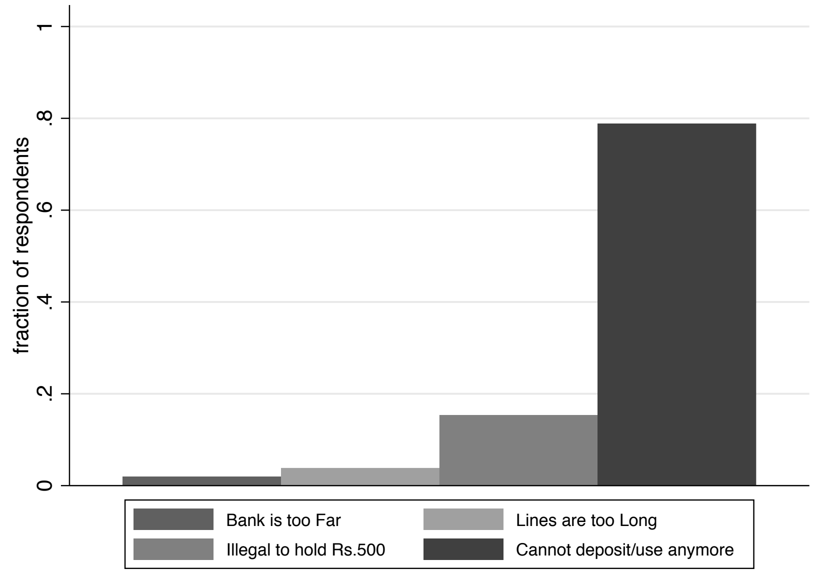
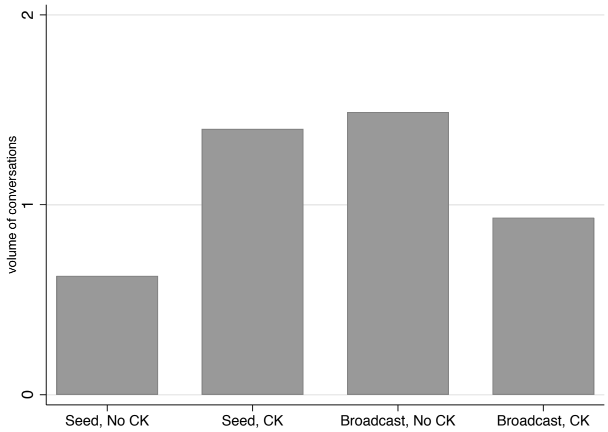
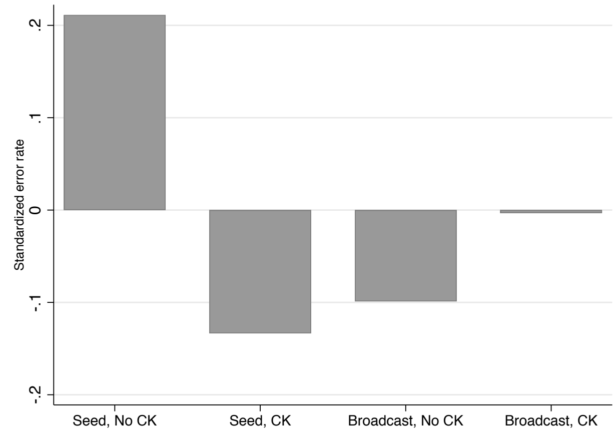


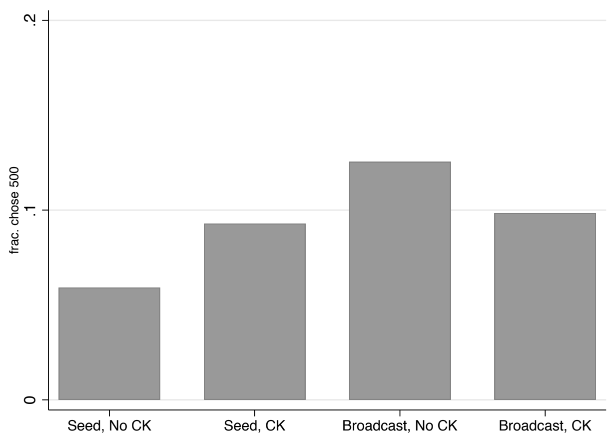
FIGURE 3. Why did you not choose 500?



(A) Volume of conversations



(B) Knowledge error



(C) Chose old 500

FIGURE 4. Raw Data: Core Experiment Outcomes

Tables

Table 1. Summary Statistics

	mean	sd	obs
Female	0.32	(0.47)	1082
SC/ST	0.50	(0.50)	1082
Age	39.18	(11.88)	1079
Casual laborer	0.21	(0.41)	1082
Farmer: landed	0.16	(0.37)	1082
Domestic work	0.16	(0.37)	1082
Farmer: sharecropper	0.09	(0.29)	1082
Unemployed	0.02	(0.14)	1082
Bank account holder	0.89	(0.31)	1078
Literate	0.80	(0.40)	1047

Notes: This table gives summary statistics on the endline sample used for analysis.

Table 2. Bank Summary Statistics

	median	mean	sd	obs
Actual wait time at banks (mins)	10.00	11.86	(7.87)	51
Perceived wait time at banks (mins)	15.00	17.06	(22.13)	32
Nearest Bank (mins)	20.00	19.84	(9.88)	63

Notes: This table gives actual wait time at banks near our sample villages. We surveyed bank employees at 51 banks. It also gives perceived wait time and perceived time taken to reach the nearest bank by a sub-sample of the endline respondents.

Table 3. Baseline Error Statistics

	mean	sd	obs
10 rupees coin	0.15	(0.36)	965
General currency	0.17	(0.38)	965
Withdrawal limits on Jan Dhan accounts	0.87	(0.33)	965
Over-the-counter exchange	0.25	(0.44)	965
Weekly withdrawal limits from bank accounts	0.78	(0.41)	965
Daily withdrawal limits on ATMs	0.90	(0.30)	965
Exchange locations other than banks	0.50	(0.50)	966

Notes: This tables gives error rates on knowledge about demonetization in the baseline sample.

Table 4. Frictionless benchmark

Panel A: Short vs. Long

	(1)	(2)	(3)
	OLS	OLS	OLS
VARIABLES	Volume	Knowledge error	Chose 500
Long	-0.296	0.00692	-0.0183
	(0.250)	(0.00946)	(0.0180)
	[0.238]	[0.465]	[0.309]
Observations	1,078	1,082	1,067
Short Mean	1.136	0.417	0.0954

Panel B: Seed vs. Broadcast

	(1)	(2)	(3)
	OLS	OLS	OLS
VARIABLES	Volume	Knowledge error	Chose 500
Broadcast	-0.0399	-0.00236	0.0129
	(0.253)	(0.00936)	(0.0186)
	[0.875]	[0.802]	[0.490]
Observations	1,078	1,082	1,067
Seed Mean	0.998	0.418	0.0755

Notes: All columns control for randomization strata (subdistrict) fixed effects. They also control for date and time of entry into the village, caste category of the treatment hamlet and distance from the village to an urban center. Respondent-level controls include age, gender, literacy and potential seed status. Standard errors (clustered at the village-level) are reported in parentheses and p-values are reported in brackets.

Table 5. Engagement in social learning

VARIABLES	(1) OLS Volume of conversations	(2) OLS # incidental conversations	(3) OLS # purposeful conversations
CK	0.651	0.447	0.204
	(0.318)	(0.262)	(0.105)
	[0.0420]	[0.0901]	[0.0527]
Broadcast	0.708	0.520	0.188
	(0.356)	(0.320)	(0.127)
	[0.0477]	[0.106]	[0.142]
Broadcast × CK	-1.491	-1.113	-0.378
	(0.529)	(0.442)	(0.190)
	[0.00535]	[0.0125]	[0.0482]
Observations	1,078	1,078	1,078
Seed, No CK Mean	0.627	0.490	0.137
CK + BC × CK = 0 p-val	0.0211	0.0314	0.247
BC + BC × CK = 0 p-val	0.0292	0.0399	0.119

Notes: All columns control for randomization strata (subdistrict) fixed effects. They also control for date and time of entry into the village, caste category of the treatment hamlet and distance from the village to an urban center. Respondent-level controls include age, gender, literacy and potential seed status. Standard errors (clustered at the village-level) are reported in parentheses and *p*-values are reported in brackets.

Table 6. Knowledge and decision-making

	(1)	(2)
	OLS	OLS
VARIABLES	Knowledge Error	Chose 500
CK	-0.0318	0.0480
	(0.0129)	(0.0228)
	[0.0142]	[0.0368]
Broadcast	-0.0279	0.0677
	(0.0143)	(0.0272)
	[0.0525]	[0.0135]
Broadcast × CK	0.0506	-0.109
	(0.0193)	(0.0392)
	[0.00958]	[0.00583]
Observations	1,082	1,067
Seed, No CK Mean	0.434	0.0592
CK + BC × CK = 0 p-val	0.174	0.0409
BC + BC × CK = 0 p-val	0.0621	0.104

Notes: All columns control for randomization strata (sub-district) fixed effects. They also control for date and time of entry into the village, caste category of the treatment hamlet and distance from the village to an urban center. Respondent-level controls include age, gender, literacy and potential seed status. Standard errors (clustered at the village-level) are reported in parentheses and p-values are reported in brackets.

Appendix A. Timeline of Rule Changes

Nov-08	Rs. 500 and Rs. 1000 notes shall have their legal tender withdrawn wef midnight Nov 8
	Closure of ATMs from Nov 9th to Nov 11th
	All ATM free of cost of dispensation
	ATM machine withdrawal limit: Rs. 2000 per day per card (till Nov. 18th); Rs. 4000 thereafter
Nov-09	Re-Calibration of ATMs to dispense Rs. 50 and Rs. 100 notes
	Withdrawal of Rs. 2000 limit per day per card
	Cash withdrawals could be made from Banking Correspondents and Aadhar Enabled Payment Systems
Nov-10	Rs. 4000 or below could be exchanged for any denomination at banks
	Max deposit for an account without KYC: Rs. 40000
	Cash withdrawal per day: Rs. 10,000; with a limit of Rs. 20,000 in one week
Nov-13	Limit for over the counter withdrawal: Rs. 4500
	Daily withdrawal on debit cards: Rs. 2500
	Weekly withdrawal limit: Rs. 24,000
	Daily limit of Rs. 10,000: withdrawn
	Separate queues for senior citizens and disabled
Nov-14	Waivers of ATM customer charge
	Current account holders: Withdrawal limits Rs. 50,000 with notes of mostly Rs. 2000
Nov-17	Over the counter exchange of notes limited to Rs. 2000
	PAN card is mandatory for deposits over Rs. 50,000, or opening a bank account
Nov-20	Withdrawal of ATM: limit unchanged at Rs. 2500
Nov-21	Cash withdrawal for wedding: Rs. 2,50,000 for each party for wedding before Dec. 30th, for customers with full KYC
	60 day extra for small borrowers to repay loan dues
	Limit of Rs. 50,000 withdrawal also extended to overdraft, cash credit account (in addition of current account - Nov-14)
	Farmers can purchase seeds with the old Rs. 500 notes
Nov-22	Prepaid payment instruments: limit extended from Rs. 10,000 to Rs. 20,000 in order to push electronic payment systems
	For wedding payments: a list must be provided with details of payments for anyone to whom a payment of more that 10,000 is to be made for wedding purposes
Nov-23	SBNs not allowed to deposit money in Small Saving Schemes
Nov-24	No over the counter exchange of SBNs wef midnight Nov-24
	Only the old Rs. 500 notes will be accepted till Dec. 15th in the following places: government school or college fees, pre-paid mobiles, consumer co-op stores, tolls for highways
Nov-25	Weekly withdrawal limit: Rs. 24,000 (unchanged)
	Foreign citizens allowed to exchange Rs. 5000 per week till Dec 15th

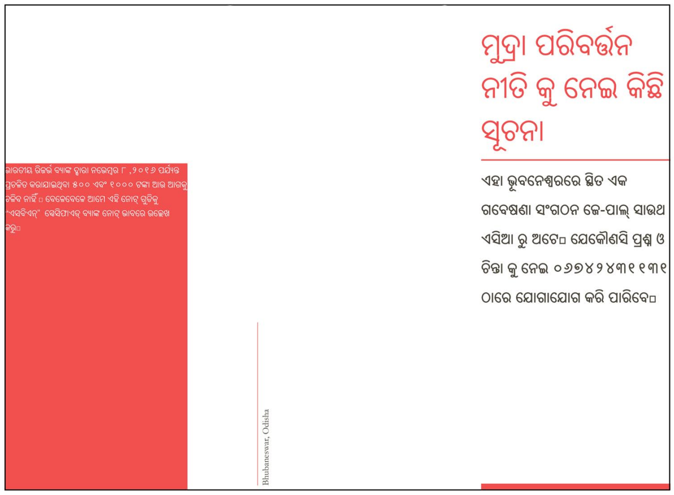
Nov-28	Relaxation in norms of withdrawal from deposit accounts of deposits made in legal tender note wef Nov-29
Nov-29	For account holders of Pradhan Mantri Jan Dhan Yojana: limit of Rs. 10,000 withdrawal per month for full KYC customers; Rs. 5000 with customers with partial KYC
Dec-02	Aadhaar-based Authentication for Card Present Transactions
Dec-06	Relaxation in Additional Factor of Authentication for payments upto Rs. 2000 for card network provided authentication solutions
Dec-07	Old Rs. 500 notes can only be used for purchase of railway tickets till Dec. 10th
Dec-08	OTP based e-KYC allowed
Dec-16	Pradhan Mantri Garib Kalyan Deposit Scheme Issued wef Dec 17
	Foreign citizens allowed to exchange Rs. 5000 per week till Dec 31st
	Merchant discount rate for debit card transactions revised
	No customer charges to be levied for IIMPS, UPI, USSD
Dec-19	SBNs of more than Rs. 5000 to be accepted only once till Dec 30th to full KYC customers
Dec-21	The limit of Rs. 5000 deposit not applicable to full KYC customers
Dec-26	60 day extra for short term crop loans
Dec-29	Additional working capital for MSEs
Dec-30	Closure of the scheme of exchange of Specified Bank Notes
	PPI guideline (issued Nov 22) extended
	ATM machine withdrawal limit: Rs. 4500 per day per card
Dec-31	Grace period for non-present Indians for SBN exchange at RBI
Jan-03	Allocation changes to cash in rural areas
	Foreign citizens allowed to exchange Rs. 5000 per week till Jan 31
Jan-16	ATM limit extended to Rs. 10,000 per day per card
	Current account withdrawal limits extended to 1,00,000

Appendix B. List of Facts

Chapter 1: DEPOSITING OR TENDERING SPECIFIED BANK NOTES	1. The old Rs. 500 and Rs.1000 notes will be accepted at bank branches until 30/12/2016. If you deposit more than Rs. 5,000 then you will have to provide a rationale for why you didnt deposit the notes earlier. 2. You will get value for the entire volume of notes tendered at the bank branches / RBI offices. 3. If you are not able to personally visit the branch, you may send a representative with a written authority letter and his/her identity proof with tendering the notes. 4. Banks will not be accepting the old Rs.500 and Rs. 1000 notes for deposits in Small Saving Schemes. The deposits canbe made in Post Office Savings accounts. 5. Quoting of PAN is mandatory in the following transactions: Deposit with a bank in cash exceeding Rs. 50,000 in a single day; Purchase of bank drafts or pay orders or bankers cheques from a bank in cash for an amount exceeding Rs. 50,000 in a single day; A time deposit with a Bank or a Post Office; Total cash deposit of more than Rs. 2,50,000 during November 09 to December 30th, 2016
Chapter 2: EXCHANGING SPECIFIED BANK NOTES	1. The over the counter exchange facility has been discontinued from the midnight of 24th November, 2016 at all banks. This means that the bank wont exchange the notes for you anymore. You must first deposit them into an account. 2. All of the old Rs.500 and Rs. 1,000 notes can be exchanged at RBI Offices only, up to Rs.2000 per person. 3. Until December 15th, 2016, foreign citizens will be allowed to exchange up to Rs. 5000 per week. It is mandatory for them to have this transaction entered in their passports. 4. Separate queues will be arrangedfor Senior Citizens and Divyang persons, customers with accounts in the Bankand for customers for exchange of notes (when applicable).
Chapter 3: CASH WITHDRAWAL AT BANK BRANCHES	1. The weekly limit of Rs. 20,000 for withdrawal from Bank accounts has been increased to Rs. 24,000. The limit of Rs. 10,000 per day has been removed. 2. RBI has issued a notification to allow withdrawals of deposits made in the valid notes (including the new notes) on or after November 29, 2016 beyond the current limits. The notification states that available higher denominations bank notes of Rs. 2000 and Rs. 500 are to be issued for such withdrawals as far as possible. 3. Business entities having Current Accounts which are operational for last three months or more will be allowed to draw Rs. 50,000 per week. This can be done in a single transaction or multiple transactions. 4. To protect innocent farmers and rural account holders of PMJDY from money launders, temporarily banks will: (1) allow account holders with full KYC to withdraw Rs. 10,000 in a month;(2) allow account holders with limited KYC to withdraw Rs.5,000 per month, withthe maximum of Rs.10,000 from the amount deposited through SBN after Nov 09,2016 5. District Central Cooperative Banks (DCCBs) will also facilitate withdrawals with the same limits as normal banks.
Chapter 4: ATM WITHDRAWALS	1. Withdrawal limit increased to Rs. 2,500 per day for ATMs that have been recalibrated to fit the new bills. This will enable dispensing of lower denomination currency notes for about Rs.500 per withdrawal. The new Rs. 500 notes can be withdrawn 2. Micro ATMs will be deployed to dispense cash against Debit/Credit cards up to the cash limits applicable for ATMs. 3. ATMs which are yet to berecalibrated, will continue to dispense Rs. 2000 till they are recalibrated.
Chapter 5: SPECIAL PROVISIONS FOR FARMERS	1. Farmers would be permitted to withdraw up to Rs. 25,000 per week in cash from their KYC compliant accounts for loans. These cash withdrawals would be subject to the normal loan limits and conditions. This facility will also apply to the Kisan Credit Cards (KCC). 2. Farmers receiving payments into their bank accounts through cheque or other electronic means for selling their produce, will be permitted to withdraw up to Rs.25,000 per week in cash. But these accounts will have to be KYC compliant. 3. Farmers can purchase seeds with the old bank notes of 500 from the State or Central Govenment Outlets, Public Sector Undertakings, National or State Seeds Corporations, Central or State Agricultural Universities and the Indian Council of Agricultural Research (ICAR), with ID proof.

	4. Traders registered with APMC markets/mandis will be permitted to withdraw up to Rs. 50,000 per week in cashfrom their KYC compliant accounts as in the case of business entities. 5. The last date for payment of crop insurance premium has been extended by 15 days to 31st December,2016.
Chapter 6: SPECIAL PROVISIONS FOR WEDDINGS	1. In the case of a wedding, one individual from the family (parent or the person themselves) will be able to withdraw Rs. 2,50,000 from a KYC compliant bank account. PAN details and self-declaration will have to be submitted stating only one person is withdrawing the amount. The girls and the boys family can withdraw this amount separately. 2. The application for withdrawal for a wedding has to be accompanied by the following documents: An application form; Evidence of the wedding, including the invitation card, copies of receipts for advance payments already made, such as Marriage hall booking, advance payments to caterers, etc.; A declaration from the person who has to be paid more than Rs. 10,000 stating that they do not have a bank account, anda complete list of people who have to be paid in cash and the purpose for the payment.
Chapter 7: OTHER DETAILS	1. In Odisha, Panchayat offices can be used for banking services in areas where banks are too far or banking facilities are not available. 2. You can use NEFT/RTGS/IMPS/InternetBanking/Mobile Banking or any other electronic/ non-cash mode of payment. 3. Valid Identity proof is any of the following: Aadhaar Card, Driving License, Voter ID Card, Pass Port, NREGA Card, PAN Card, Identity Card Issued by Government Department, Public Sector Unit to its Staff. 4. You may approach the control roomof RBI on Telephone Nos 022-22602201 22602944 5. The date for submission of annual life certificate has been extended to January 15, 2017 from November for all government pensioners 6. As of December 15, 2016, specified bank notes of only Rs. 500 can no longer be used for the following: Government hospitals and pharmacies, railway and government bus tickets, consumer cooperative stores, government and court fees, government School fees, mobile top-ups, milk booths, crematoria and burial grounds, LPG gas cylinders, Archaelogical Survey of India monuments, utilities, toll payments

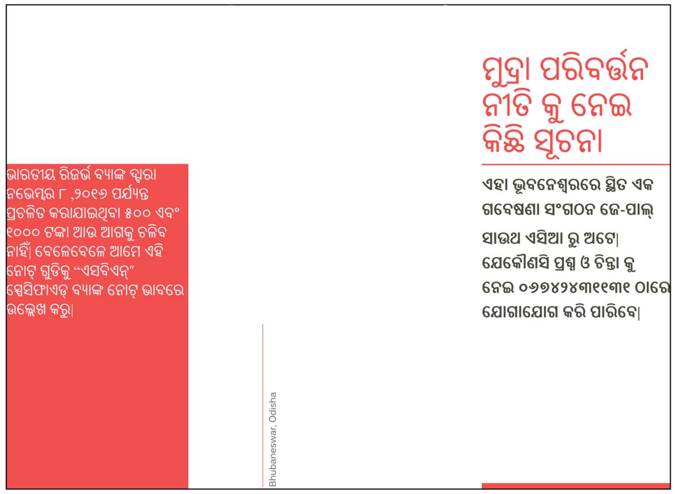
Appendix C. Example Pamphlet Excerpts



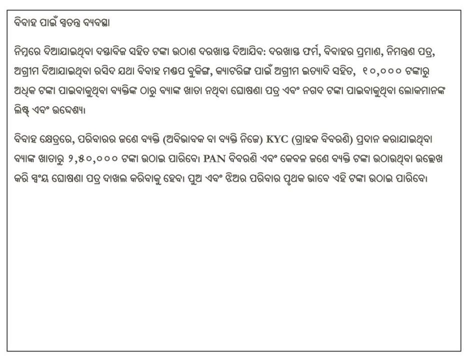
(A) Front

(B) Back

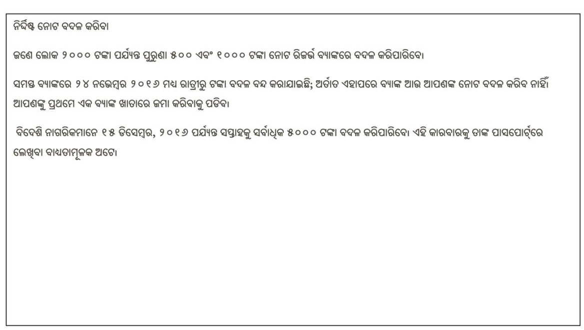
Figure C.1. Short pamphlet (2 facts)



(A) Front



(B) Page 1/8



(C) Page 2/8

Figure C.2. Long pamphlet (24 facts)

Appendix D. Technical details and proofs for Section 3

D.1. **Signaling model.**

D.1.1. *Preliminaries for Proof of Main Proposition.* Introduce an *index* $\omega \in (0,1)$ for the type of the decision-maker D. This index is drawn uniformly from [0,1]. By the assumption of no atoms, we can view $\Delta^{(I_\mathrm{D},I_\mathrm{T})}$ as a continuous increasing function $(0,1) \to \mathbb{R}$. Moreover, by P2, we may assume that, pointwise, $\Delta^{(I_\mathrm{D},1)}(\omega) > \Delta^{(I_\mathrm{D},0)}(\omega)$ and $\Delta^{(0,1)}(\omega) \geq \Delta^{(1,1)}(\omega)$. This uses the standard coupling for random variables ordered by stochastic dominance.

Recall the payoff difference formula (3.3)

$$u^{(I_\mathrm{D},I_\mathrm{T})}(\mathrm{S}) - u^{(I_\mathrm{D},I_\mathrm{T})}(\mathrm{NS}) = \Delta^{(I_\mathrm{D},I_\mathrm{T})} - \lambda\Pi,$$

where Π is the signaling penalty. For any $\mathfrak{p}$, a strategy profile in which D is best-responding can be summarized by a vector of interior cutoffs $\boldsymbol{c} = (c(\mathfrak{p}, I_\mathrm{D}))_{I_\mathrm{D}}$ such that D seeks given I_D if his index ω is above $c(\mathfrak{p}, I_\mathrm{D})$, and does not seek if his index is below $c(\mathfrak{p}, I_\mathrm{D})$. (Interiority is guaranteed by the assumption that the distributions of Δ in each case have full support.)

We may now write the right-hand side of (3.3) as

$$W^{(I_\mathrm{D},I_\mathrm{T})}(\omega;\boldsymbol{c}) = \Delta^{(I_\mathrm{D},I_\mathrm{T})}(\omega) - \lambda\Pi(\boldsymbol{c}).$$

Here $\Delta^{(I_\mathrm{D},I_\mathrm{T})}(\omega)$ is increasing in ω and $\Pi(\boldsymbol{c})$ is increasing in $\boldsymbol{c}$ by P3.

Define $W^{(I_\mathrm{D},\mathfrak{p})}(\omega)$ to be the expectation of $W(\omega)$ given public signal $\mathfrak{p}$ and a realization of I_D. Define the analogous notation for Δ.

Because λ is a finite constant, cutoffs given both values of I_D are guaranteed to be in some compact subset $\mathcal{C} \subseteq (0,1)$ irrespective of strategies; so we will restrict attention to this subset from now on in studying equilibria.[38]

For each $\mathfrak{p}$ and each ω, the payoff advantage $W^{(I_\mathrm{D},\mathfrak{p})}(\omega)$ of seeking is monotone decreasing in the cutoff vector $\boldsymbol{c}$, so this is a supermodular game. In particular, a minimum equilibrium cutoff profile (which corresponds to maximum seeking) exists. We now state two results which follow from the supermodular structure of the game:

Fact **1.** The following hold:

SM1 If $W^{(I_\mathrm{D},\mathfrak{p})}(\omega;\boldsymbol{c})$ strictly increases for each $\omega, \boldsymbol{c} \in \mathcal{C}$ and I_D then the minimum cutoff $\boldsymbol{c}$ strictly decreases in each component.

[38]To show the cutoff does not get arbitrarily close to 0 in ω space, we can simply note that each function $\Delta^{(I_\mathrm{D},\mathfrak{p})}(\omega)$ is negative below some $\omega > 0$. Because $\Pi \geq 0$, cutoffs cannot occur in the region where W is negative.

SM2 Let $\iota_{\mathfrak{p}}$ be the ex ante probability of $I_{\mathrm{D}} = 1$ given $\mathfrak{p}$. Then, for each $\mathfrak{p}$, the maximum equilibrium cutoff $c(\mathfrak{p}, 0)$ is continuous in $\iota_{\mathfrak{p}}$ at $\iota_{\mathfrak{p}} = 0$ for generic priors $(\alpha_{\mathrm{H}}, \alpha_{\mathrm{L}})$.

The first part, SM1, is a standard monotone comparative statics fact. The second, SM2, is argued as follows. Define a reaction function $r_{\iota_{\mathfrak{p}}} : \mathcal{C}^2 \to \mathcal{C}^2$ mapping any cutoffs $\boldsymbol{c}$ to the best-response cutoffs when the Observer updates assuming the cutoffs $\boldsymbol{c}$. Because the distribution of $\Delta^{(I_{\mathrm{D}}, I_{\mathrm{T}})}$ has full support, inferences of the Observer depend arbitrarily little on the behavior of $I_{\mathrm{D}} = 1$ types as $\iota_{\mathfrak{p}} \downarrow 0$. Thus, the reaction functions $r_{\iota_{\mathfrak{p}}}$ may be bounded within an arbitrarily narrow band of the reaction functions r_0. Thus, for generic parameters (guaranteeing that r is transversal to the hyperplane $(x, y) \mapsto (x, y)$ at the equilibrium), the equilibrium will be continuous in $\iota_{\mathfrak{p}}$.

D.1.2. *Proof of Proposition 1.*

(a) (Broadcast, No CK) has more seeking than (Broadcast, CK). In both cases, $W^{(I_{\mathrm{D}}, \mathfrak{p})}(\omega)$: since $I_{\mathrm{D}} = 1$, D knows that $I_{\mathrm{T}} = 1$.

Now we turn to signaling concerns. Denote by $\mathcal{I}_{\mathrm{D}}$ all the information D has when making his decision. Write

$$\mathbf{E}^{\mathrm{D}}\left[\Pi(\boldsymbol{c}) \mid \mathcal{I}_{\mathrm{D}}\right] = \xi \mathbb{P}_{\boldsymbol{c}}(a = \mathrm{H} \mid d = 1, \mathfrak{p}, I_{\mathrm{D}} = 1) + (1 - \xi)\mathbb{P}_{\boldsymbol{c}}(a = \mathrm{H} \mid d = 1, \mathfrak{p}, I_{\mathrm{D}} = 0). \tag{D.1}$$

This says that D's interim expectation of perception payoffs can be written as a convex combination (involving a weight ξ that depends on $\mathcal{I}_{\mathrm{D}}$) of conditional probabilities of $a = \mathrm{H}$ *given* the value of I_{D}. The probabilities assessed by O depend on the cutoffs used, hence the subscripts $\boldsymbol{c}$. Note that under (Broadcast, CK), $\xi = 1$, while under (Broadcast, No CK), ξ is not 1 because the probability of Seeding is positive and the seed set S is a proper (strict) subset of N. Now, by P1(b), the first probability (the one being multiplied by ξ) is smaller than the second probability (the one being multiplied by $1 - \xi$), by P1(b). This formalizes the claim that signaling concerns could not be greater than they are in the (Broadcast, CK) case. Applying SM1 finishes the proof.

(b) (Seed, CK) has more seeking than (Broadcast, CK).

Considering the signaling contribution to payoffs: for any given cutoffs, just as in (a), we can write the update of the Observer (given either value of $\mathfrak{p}$) as a convex combination conditioning on values of I_{D}. The term corresponding to $I_{\mathrm{D}} = 1$ is the same across the two treatments, and the term corresponding to $I_{\mathrm{D}} = 0$ involves a strictly lower posterior that $a = \mathrm{H}$. Only the first term is nonzero in the (Broadcast, CK) treatment, while both contribute in the (Seed, CK) treatment. Turning now to the direct payoffs, $I_{\mathrm{T}} = 1$ is known in both cases. By Assumption P2(b), $\Delta^{(0,1)}(\omega) \leq \Delta^{(1,1)}(\omega)$ for every ω.

Applying SM1 to the two W functions gives the result.

D.1.3. *Proof of Proposition 2.* First, under (Seed, CK), D is certain that information is in the Town Square, while under (Seed, No CK) this probability is strictly less than 1 assuming $I_{\mathrm{D}} = 0$. Thus $\Delta^{(0,\text{CK:Seed})}(\omega)$ is pointwise strictly greater than $\Delta^{(0,\text{No CK})}(\omega)$. By compactness of $\mathcal{C}$, it is strictly greater for all $\omega \in \mathcal{C}$, by at least a positive quantity $\nu > 0$.

Now we turn to signaling concerns. Condition first on $I_{\mathrm{D}} = 0$. By the argument given in the main text, once $\overline{k}/n$ is small enough, in the decomposition of (D.1) the weight on the $I_{\mathrm{D}} = 1$ term under either value of $\mathfrak{p}$ is arbitrarily small. Thus, the difference between signaling payoffs under $\mathfrak{p} = \text{No CK}$ and under $\mathfrak{p} = \text{CK:Seed}$ is less than ν. Thus we see $W^{(0,\mathfrak{p})}$ strictly increases pointwise for each $\omega, \boldsymbol{c} \in \mathcal{C}$ when we move from $\mathfrak{p} = \text{No CK}$ to $\mathfrak{p} = \text{CK:Seed}$.

Because the realizations with $I_{\mathrm{D}} = 1$ become very unlikely (by smallness of $\overline{k}/n$), we can apply SM2 to finish the proof.

D.1.4. *Proof of Proposition 3.* We now state a formal version of Lemma 1, whose proof follows by Bayes' rule.

Lemma 1. Fix any $\epsilon > 0$. Then there is a δ (depending on this ϵ) so that if $(1 - \chi_{\text{Broadcast}}) < \delta(\overline{k}/n)^2$, then conditional on $\mathfrak{p} = \text{No CK}$ and any realizations of I_{D} and I_{O}, the probability that $\mathfrak{b} = \text{Broadcast}$ is at most ϵ.

Now, to prove the proposition in several steps. First, we will show that (Seed, No CK) has a level of seeking arbitrarily close to the one when it is common knowledge that $I_{\mathrm{T}} = 0$ and $I_{\mathrm{D}} = 0$.

Consider (Seed, No CK). Condition on $I_{\mathrm{D}} = 0$. When D receives no information ($I_{\mathrm{D}} = 0$), the fact that β_{Seed} is small means that his expectations approximate those when $I_{\mathrm{T}} = 0$. Thus, his direct benefits as a function of ω are arbitrarily close to $\Delta^{(0,0)}$ on the compact set $\mathcal{C}$. Moreover, in (Seed, No CK), conditioning on $I_{\mathrm{D}} = 0$, D is certain that $\mathfrak{b} \neq \text{Broadcast}$, and thus (because the probability of seeding is small) he believes that $I_{\mathrm{O}} = 0$ with high probability, and thus signaling concerns are uniformly bounded by an arbitrarily small number on $\mathcal{C}$. By the full support assumption on $\Delta^{(0,0)}$, it follows that for any cutoffs, there is an arbitrarily small measure of ω for which the decision differs from the case where Π is exactly zero. Finally, applying SM2 shows that the conclusion extends even when we take into account the $I_{\mathrm{D}} = 1$ realizations.

Now consider (Broadcast, No CK), every realized D is certain that $I_{\mathrm{T}} = 1$ and thus assesses the direct benefits to be greater than his $I_{\mathrm{D}} = 0$ counterpart, by an amount bounded away from 0, as in Proposition 2. Fourth, under (Broadcast, No CK), signaling concerns are negligible, as follows. By the lemma, conditional I_{D}, D is nearly certain that $\mathfrak{b} \neq$ Broadcast. The probability of $\mathfrak{b} =$ Seed is small. Putting these facts together, D is also nearly certain that $I_{\mathrm{O}} = 0$. Thus, in the decomposition of (D.1) the weight on the $I_{\mathrm{D}} = 1$ term under either value of $\mathfrak{p}$ is arbitrarily small. Continuing from that point just as in the proof of Proposition 2, we conclude that signaling concerns are negligible. Thus, seeking rates are as if it is common knowledge that $I_{\mathrm{T}} = 1$ and $I_{\mathrm{D}} = 0$.

By P2, there is more seeking when it is common knowledge that $I_{\mathrm{T}} = 1$ and $I_{\mathrm{D}} = 0$ than when it is common knowledge that $I_{\mathrm{T}} = 0$ and $I_{\mathrm{D}} = 0$ (this follows by a simple comparison of direct payoffs without any signaling concerns).

Appendix E. Alternative Models

E.1. **Supply Effects: Information as a Public Good.** The core model of Chandrasekhar, Golub, and Yang (2017) and its application to our setting focuses on seeking effort or endogenous participation in learning. A different kind of explanation focuses on the effort of those informed to understand, filter, and communicate the information in a useful way to others. The simplest framework to capture this is a model of public goods provision and free-riding. This class of model has been studied extensively in a development context, and we rely on arguments from Banerjee, Iyer, and Somanathan (2007) to explain why supply-side effects are unlikely to explain our results.

A robust point within such public goods models is that enlarging the set of people who are able to provide a public good should not, in equilibrium, reduce its aggregate provision. Indeed, if anything provision should slightly increase, which is contrary to our empirical results.

For a simple model, consider a situation where those initially given information have the opportunity to provide the public good of processing and disseminating it to others. There are n agents, and each of those informed believes that k in total are able to contribute. Every i who has information invests an effort $z_i \geq 0$ in transmitting. Their payoffs are given by

$$U_i(z_1, \ldots, z_n) = V\left(\sum_i z_i\right) - cz_i.$$

Here V is an increasing, smooth function with $V'(z)$ tending to 0 at large arguments z, and $c > 0$ is a cost parameter. Those who are unable to contribute are constrained to $z_i = 0$ and are passive. The key fact, which is formalized for instance by Banerjee, Iyer, and Somanathan (2007), is that at any equilibrium with some people contribution, for those contributing we have

$$V'\left(\sum_i z_i\right) = c, \tag{E.1}$$

so the aggregate level of contribution cannot depend on n or k. The intuition is simple: the free-riding problem is self-limiting, at least in the sense of aggregate (though not per-person) provision. If more agents try to free-ride, then others have more reason to provide the good. A similar force is present in the network model of Galeotti and Goyal (2010): there, endogenously, networks form so that only a few people provide the public good but everyone can access it, and a larger number of potential providers does not make for less provision.

If agents have a private benefit term in their utility function, $v_i(z_i)$, where v is increasing and $v'(z_i) > c$ for $z_i \in [0, \delta)$, then as long as there are sufficiently many agents who can provide the public good, the amount provided will be at least $k\delta$—a lower bound which is increasing in k. A similar argument applies if only some agents have such a v term.

Thus, natural public goods theories do not predict a decrease in the amount of overall provision, and thus in overall learning, as k (the number of potential providers) increases. One can, of course, elaborate these models with stochastic k and idiosyncratic c_i, but the basic intuition described above is quite robust.

One further supply-side effect to consider is one of social obligation. If the seeds are publicly "deputized," as they are in the CK treatment, each may face stronger incentives to provide information relative to a situation in which provision opportunities are diffuse. Though this is outside a basic public goods model, our evidence on seed effort does not support this hypothesis.

E.1.1. *Application to Experiment.* The number of people, k, who can contribute is either $k = 5$ or $k = n$. Under common knowledge, this matches up with the beliefs agents hold, so in this sense the simple model is faithful to the experiment. Thus, the basic public goods theory predicts (contrary to the demand-side theory) that holding CK fixed and moving from Seed to Broadcast should not hurt aggregate provision.

When common knowledge is not present, agents will have beliefs about k. But as long as their beliefs about k are reasonably consistent (e.g., agents have common priors about it), the essence of the above argument goes through: a stochastic version of (E.1) still holds, and changes in beliefs about k alone should not lead to large swings in provision.

This model is inconsistent with our empirical findings for several reasons. First, aggregate provision of effort cannot decline, as established above. If the number of people a typical subject in our random sample conversed with measures conversational effort, this means that the number of conversations for the average person must not decline. Column 1 of Table 5 shows that, conditional on common knowledge, going from $k = 5$ to $k = n$ corresponds to a 61% decline in the number of conversations ($p = 0.029$), which means that aggregate contribution to conversations must be decreasing.

Second, the model suggests that the amount of value being generated cannot decline, since after all otherwise a given individual would have an incentive to put in some more effort to gain more marginal benefit. Here, we can measure this either through knowledge or choice quality. Turning to Table 6, recall that columns 1 (for knowledge) and 2 (for choice) show robust declines in aggregate social learning and quality of choice when we go from $k = 5$ to $k = n$ under common knowledge ($p = 0.0621$ and $p = 0.104$).

E.2. **Tagged Information Aggregation.** There is an undirected graph $G = (N, E)$ of potential communication opportunities, corresponding to the social network with nodes N and edges E. At time 0, agents are endowed with certain information, the realization of a random variable S_i. (In our application, this represents one's degree of understanding of the information delivered in the intervention.) At each discrete time $t = 1, 2, \ldots$ a subset $E_t \subseteq E$ of agents who can communicate is realized randomly.[39] We make no assumptions on this process: it may involve arbitrary correlations, etc. If agents i and j are able to communicate at time t, they send each other messages, with the $i \to j$ message $m_{ij,t}$ reaching its destination with probability $p_{ij,t}$. Again, we make no assumptions on these numbers. Critically, information is "tagged." This means that at time t, agent i's information, $I_{i,t}$, consists of a set of signals labeled by their origin (formally, a set of pairs (k, S_k)). When agent i sends a message to j, the message reveals his whole information set I_t, which then is incorporated into j's information. Consider any improvement in initial information—making the profile of initial signal random variables $(S_i)_{i\in N}$ more informative in the Blackwell sense to obtain a new profile $(\widetilde{S}_i)_{i\in N}$. Then, holding fixed the parameters of the model, at any time t and for any agent i, the information $\widetilde{I}_{i,t}$ dominates $\widetilde{I}_{i,t}$.[40]

E.3. **Herding model.** We briefly review the notation of the Lobel and Sadler (2015) model, paraphrasing their Section 2. Agents, indexed by natural numbers n which correspond to the time they move, sequentially make choices $x_n \in \{0, 1\}$, which can be thought of making the correct choice or statement about the new currency. Agents receive a positive payoff from matching the state $\theta \in \{0, 1\}$, and zero otherwise. In contrast to the tagging model, this is a maximally coarsened mode of communication. Each individual, when acting, observes two things: a private signal $s_n \in \mathcal{S}$, and the actions of a set of predecessors $B(n)$, which may be drawn with randomness. This allows us to encode network structure into the model. Private signals are conditionally independent given the true state θ.

[39]We omit formal notation for the probability space in the background.

[40]Formally, if we order information sets by containment, then under this order $\widetilde{I}_{i,t}$ first-order stochastically dominates $I_{i,t}$.

Lobel and Sadler (2015) show that in equilibrium, the decisions of all sufficiently late-moving agents (those with high n) are at least as good as those decisions that would be made based on s_n alone, for the most informative possible realizations of s_n. To state this more formally, they define the private belief p_n as the belief about θ induced by n's signal, and define the strongest possible private beliefs to be the extreme points of the support of p_n, which they denote by $\underline{\beta}$ and $\overline{\beta}$. So, more formally, Lobel and Sadler (2015) show that the decisions of all sufficiently late-moving agents achieve essentially the utility that would be achieved by getting one of the strongest possible private signals. This requires some conditions on the network structure. The simplest of these (in their Theorem 1) is that individuals' neighborhoods are independent, and each late-moving agent has paths of observation leading back to arbitrarily many prior movers' choices.

Though in the sequential social learning model, equilibrium outcomes may be non-montonic in signal endowments, the Lobel-Sadler lower bound described above is monotonic in signal endowments: when we make everyone's initial information better, the $\underline{\beta}$ and $\overline{\beta}$ become more extreme (corresponding to stronger signals and better decisions) and the lower bound is strengthened.

Appendix F. Heterogeneity by length of information

We now look at the interaction of our core treatment cells with the amount of information in the pamphlet. Whether this should accelerate or dampen the effect of going to common knowledge in a given information delivery system depends on the details of the model and therefore becomes an empirical question.

To see why, consider the case of (Broadcast, CK) and now imagine comparing a world in which only two facts are given as compared to a world where a lengthy pamphlet of 24 facts is given. What matters is how the type-specific marginal value of information distributions, F_{H} and F_{L}, move when we go from a short set of facts to a long set of facts. Assume for the moment that the cost of figuring out which of the 24 facts are useful, or coordinating on the same topic of conversation out of the now 24 possibilities, is very high no matter if the individual is a high or low ability type. In this case, the scope for signaling reduces, and therefore going from (Broadcast, No CK, Long) to (Broadcast, CK, Long) should generate less of a reduction in endogenous participation in social learning than going from (Broadcast, No CK, Short) to (Broadcast, CK, Short). Now on the other hand, if it was very easy for high ability types to figure out what is useful, but the task was arduous for low ability types, then scope for signaling could actually increase.

Turning to seeding, observe that in seeding with or without common knowledge, the length of the information is not commonly known either way. So, long sets of facts should likely have no effect on endogenous participation.

We now turn to the data in Table F.1 to look at how going from two to 24 facts differentially impacts the effects of interest. For the most part the effect is noisy, and there is no differential effect. The one plausible finding is that going from (Broadcast, No CK) to (Broadcast, CK) is less of a deterrent to purposeful conversations ($p = 0.15$) when the facts are long. If this is to be taken seriously, minding the caveat that for overall conversations this effect is not distinguishable from zero ($p = 0.251$), it is evidence in favor of the idea that sorting through the 24 facts or deciding which topic to coordinate on and converse about is costly enough for both ability types that the signaling motive is dampened by the longer list. Said differently, it is, if anything, consistent with the story that it is much less likely for someone to go ask about information when it is known that they have received two facts, than when it is known that they received a lengthy booklet of facts.

Table F.2 repeats the same exercise now turning to knowledge and choice. Of note is that a similar pattern is true here. There is mostly no detectable effect. But if we

Table F.1. Conversations: Length interactions

VARIABLES	(1) OLS Volume of conversations	(2) OLS # incidental conversations	(3) OLS # purposeful conversations
CK	0.825	0.693	0.132
	(0.496)	(0.412)	(0.163)
	[0.0982]	[0.0937]	[0.421]
Broadcast	0.963	0.665	0.297
	(0.545)	(0.481)	(0.219)
	[0.0787]	[0.168]	[0.175]
Long	-0.0939	-0.00127	-0.0926
	(0.372)	(0.330)	(0.130)
	[0.801]	[0.997]	[0.478]
Broadcast × CK	-2.212	-1.614	-0.599
	(0.735)	(0.626)	(0.264)
	[0.00296]	[0.0107]	[0.0244]
CK × Long	-0.372	-0.485	0.113
	(0.562)	(0.480)	(0.194)
	[0.508]	[0.313]	[0.560]
Broadcast × Long	-0.563	-0.319	-0.244
	(0.680)	(0.616)	(0.233)
	[0.408]	[0.605]	[0.295]
Broadcast × CK × Long	1.448	1.006	0.442
	(0.809)	(0.733)	(0.281)
	[0.0752]	[0.172]	[0.118]
Observations	1,078	1,078	1,078
Seed, No CK, Short Mean	0.523	0.385	0.138
CK + BC × CK = 0 p-val	0.00573	0.0275	0.0365
BC + BC × CK = 0 p-val	0.0170	0.0259	0.0602
Long + CK × Long + BC × Long + BC × CK × Long = 0	0.251	0.520	0.155

Notes: All columns control for randomization strata (subdistrict) fixed effects. They also control for date and time of entry into the village, caste category of the treatment hamlet and distance from the village to an urban center. Respondent-level controls include age, gender, literacy and potential seed status. Standard errors (clustered at the village-level) are reported in parentheses and p-values are reported in brackets.

had to guess, at $p = 0.5$ for both outcomes, it suggests that perhaps introducing CK to the broadcast cell has less of a detrimental effect on both knowledge and choice quality. This is extremely noisy, speculative evidence that suggests if anything, a stigma-like effect operates more when there are only two facts.

Table F.2. Knowledge and choice: Length interactions

VARIABLES	(1) OLS Knowledge Error	(2) OLS Chose 500
CK	-0.0215	0.0542
	(0.0162)	(0.0404)
	[0.185]	[0.181]
Broadcast	-0.0264	0.0804
	(0.0169)	(0.0361)
	[0.121]	[0.0269]
Long	0.0131	-0.00591
	(0.0174)	(0.0300)
	[0.451]	[0.844]
Broadcast × CK	0.0537	-0.144
	(0.0247)	(0.0556)
	[0.0312]	[0.0104]
CK × Long	-0.0167	-0.0144
	(0.0255)	(0.0508)
	[0.513]	[0.777]
Broadcast × Long	0.000655	-0.0284
	(0.0262)	(0.0548)
	[0.980]	[0.605]
Broadcast × CK × Long	-0.00862	0.0696
	(0.0383)	(0.0785)
	[0.822]	[0.376]
Observations	1,082	1,067
Seed, No CK, Short Mean	0.436	0.0374
CK + BC × CK = 0 p-val	0.0919	0.0141
BC + BC × CK = 0 p-val	0.120	0.133
Long + CK × Long + BC × Long + BC × CK × Long = 0	0.532	0.550

Notes: All columns control for randomization strata (subdistrict) fixed effects. They also control for date and time of entry into the village, caste category of the treatment hamlet and distance from the village to an urban center. Respondent-level controls include age, gender, literacy and potential seed status. Standard errors (clustered at the village-level) are reported in parentheses and p-values are reported in brackets.

ONLINE APPENDIX: NOT FOR PUBLICATION

Appendix G. Other choice and knowledge error metrics

Recall that because we randomized content, we have variation in whether the questions we ask about in the endline were actually provided to the villagers and also how relevant the information was. Table G.1 looks at whether facts are more likely to be known if (a) they were actually the ones provided in the information pamphlet to the village and (b) whether they were ex-ante deemed to be more useful to villagers. This would tell us whether there were complementarities and filtering occuring in the social learning process. The analysis is conducted on a person-fact level. Thus, it is a panel of the respondent's answers to each of the 34 facts asked to them.

In columns 1 and 2, for facts that were not provided and not useful respectively, we see that neither (Seed, CK) nor (Broadcast, No CK) are indistinguishable from (Seed, No CK). However, when we look at the effect on knowledge of facts that were provided during information delivery, adding Common Knowledge to the Seed treatment decreases error in knowledge by 17.1% (column 1, $p = 0.014$). Under no Common Knowledge, Broadcast decreases error in knowledge by 15% (column (1), $p = 0.0345$) relative to Seed. Similarly, in column 2 we see that holding useful facts fixed, (Seed, CK) decreases error in knowledge by 6.6% ($p = 0.008$) and (Broadcast, No CK) decreases error in knowledge by 6% ($p = 0.0345$), compared to (Seed, No CK). We can conclude that the core effects on aggregation are being driven by facts that were provided during information delivery and facts that were deemend useful.

Next we turn to the fact that even if the subject rejected the Rs. 500 in favor of a 3-5 day IOU for either Rs. 200 in non-demonetized notes or Rs. 200 worth of dal, we know which they picked. Table G.2 explores this. Column 1 looks at a regression where the outcome variable is a dummy for picking the dal option. We can see that relative to (Seed, No CK), adding common knowledge considerably reduces the probability of selecting dal which corresopnds to a 15.6% decline ($p = 0.135$). We also see a 14% decrease in the probability of selecting dal when going from (Seed, No CK) to (Broadcast, No CK) ($p = 0.138$). The interaction of broadcast with common knowledge has a large point estimate but is extremely noisy, however.

Note that the above says nothing about where the mass that moves away from dal ends up going. In columns 2 and 3, we present the results of a multinomial logit, where the omitted category is dal and the first column is Rs. 200 relative to dal

Table G.1. Heterogeneity in knowledge error

VARIABLES	(1) OLS Knowledge error (Told)	(2) OLS Knowledge error (Useful)
CK	0.0239	0.0352
	(0.0282)	(0.0669)
	[0.396]	[0.599]
Broadcast	0.0189	0.0325
	(0.0270)	(0.0658)
	[0.486]	[0.622]
Told/Useful	0.0840	-0.0750
	(0.0410)	(0.0488)
	[0.0419]	[0.126]
Broadcast × CK	-0.0160	-0.117
	(0.0390)	(0.0941)
	[0.682]	[0.216]
CK × Told/Useful Facts	-0.112	-0.0661
	(0.0596)	(0.0686)
	[0.0614]	[0.336]
BC × Told/Useful Facts	-0.0962	-0.0606
	(0.0575)	(0.0676)
	[0.0962]	[0.371]
BC × CK × Told/Useful Facts	0.125	0.163
	(0.0852)	(0.0975)
	[0.145]	[0.0957]
Observations	36,788	36,788
Seed, No CK, Untold/Not useful Mean	0.431	0.543
CK + CK × Told/Useful = 0 p-val	0.0140	0.00829
BC + BC × Told/Useful = 0 p-val	0.0345	0.0345

Notes: All columns control for randomization strata (subdistrict) fixed effects. They also control for date and time of entry into the village, caste category of the treatment hamlet and distance from the village to an urban center. Respondent-level controls include age, gender, literacy and potential seed status. Column (1) displays effects on error rate of if the fact being asked about was told during information delivery. Column (2) displays effects on error rate of if the fact being asked about is a useful fact or not. Standard errors (clustered at the village-level) are reported in parentheses and p-values are reported in brackets.

and the second is Rs. 500 relative to dal. We see that going to (Seed, CK) from (Seed, No CK) leads to a 3.4pp increase in the probability of selecting the IOU for Rs. 200 in cash instead of dal, relative to a mean rate of selection of Rs. 200 of 40.8% ($p = 0.285$). However we cannot detect any broadcast or broadcast interacted with common knowledge effects. When we compare the choice of Rs. 500 relative to dal, the resulting marginal changes in the probability of picking Rs. 500 look much like our main results: a 4.7pp increase when we move to (Seed, CK), a 6.9pp increase when we move to (Broadcast, No CK), and a relative decline of 4.1pp when going from (Broadcast, No CK) to (Broadcast, CK), all on a base rate of picking Rs. 500 at 5.9%.

Recall that we had two successful information dissemination strategies: (Seed, CK) and (Broadcast, No CK). We find that in the former, but not the latter, we also see movement away from dal in favor of Rs. 200 in cash. This suggests that at least some part of the misinformation involved decreased confidence in Rs. 100 notes as well, because otherwise Rs. 200 in cash should dominate dal.

Finally, because the dal, equivalent cash, and Rs. 500 are welfare-ordered, in that order, we have in column 4 an ordinal logit which shows again that (Seed, CK) and (Broadcast, No CK), relative to (Seed, No CK) improve outcomes in choice quality.

Table G.2. Other choice outcomes

VARIABLES	(1) OLS Chose dal	(2) Multinomial Logit Chose 200	(3) Multinomial Logit Chose 500	(4) Ordinal Logit Choice
CK	-0.0832	0.257	0.700	0.377
	(0.0554)	(0.241)	(0.357)	(0.208)
	[0.135]	[0.285]	[0.0496]	[0.0699]
Broadcast	-0.0756	0.124	0.932	0.398
	(0.0507)	(0.223)	(0.340)	(0.193)
	[0.138]	[0.578]	[0.00611]	[0.0396]
Broadcast × CK	0.0887	-0.117	-1.170	-0.523
	(0.0782)	(0.332)	(0.464)	(0.297)
	[0.258]	[0.724]	[0.0116]	[0.0780]
Observations	1,067	1,067	1,067	1,067
Seed, No CK Mean	0.533	0.408	0.059	
CK + BC × CK = 0 p-val	0.914	0.539	0.126	0.451
BC + BC × CK = 0 p-val	0.826	0.978	0.467	0.567

Notes: All columns control for randomization strata (subdistrict) fixed effects. They also control for date and time of entry into the village, caste category of the treatment hamlet and distance from the village to an urban center. Respondent-level controls include age, gender, literacy and potential seed status. Standard errors (clustered at the village-level) are reported in parentheses and p-values are reported in brackets.

Our study was certainly not designed to quantify the costs and benefits of demonitization in India. However, by studying misinformation and its remedies during the SBN deposit window, a few, more modest lessons emerge. First, we show that in the context of rural Orissa, while basic policy knowledge was near-universal, individuals still had a poor grasp on some of the most basic policy rules at baseline. This suggests that there was substantial room for improvement in the quality of outreach between the policy makers and villagers. Second, in our experiment, we show that decisions are impacted by the provision of information. Individuals in treatments that lead to better community wide knowledge of the policy do change their incentivized choices and are more likely to recognize that an old Rs. 500 note is more valuable than Rs. 200 in the days before the deadline. Moreover in the some treatment conditions associated with improved knowledge, namely (Seed, CK), individuals are more likely to choose currency over commodities of equivalent face value. This result suggests that a portion of the individuals preferring lentils over cash in our benchmark, non-intervention villages were likely doing so out of a loss of confidence in paper money. This observation relates back to the foundational macroeconomic literature on fiat money (Samuelson, 1958; Kiyotaki and Wright, 1989; Banerjee and Maskin, 1996; Wallace, 1980) and suggests that sowing confusion about the government's intervention in the currency undermines trust.

Appendix H. Heterogeneous Communication by Potential Seeds

Table H.1. How much more do potential seed households speak?

VARIABLES	(1) OLS Volume of conversations	(2) OLS # incidental conversations	(3) OLS # purposeful conversations
Seed HH	0.606	0.0724	0.533
	(0.857)	(0.411)	(0.479)
	[0.481]	[0.860]	[0.267]
CK	0.522	0.325	0.197
	(0.303)	(0.253)	(0.103)
	[0.0866]	[0.202]	[0.0560]
Broadcast	0.723	0.542	0.181
	(0.364)	(0.333)	(0.106)
	[0.0480]	[0.105]	[0.0906]
Broadcast × CK	-1.364	-1.058	-0.306
	(0.507)	(0.429)	(0.175)
	[0.00778]	[0.0146]	[0.0821]
Seed HH × CK	1.305	1.251	0.0540
	(1.499)	(1.156)	(0.619)
	[0.385]	[0.280]	[0.931]
Seed HH × BC	-0.505	-0.694	0.189
	(1.161)	(0.616)	(0.816)
	[0.664]	[0.261]	[0.817]
Seed HH × BC × CK	-0.917	0.0699	-0.986
	(1.874)	(1.514)	(0.898)
	[0.625]	[0.963]	[0.273]
Observations	1,078	1,078	1,078
Seed, No CK, Non-seed HH Mean	0.627	0.490	0.137
CK + BC × CK = 0 p-val	0.0168	0.0168	0.397
BC + BC × CK = 0 p-val	0.0435	0.0419	0.311

Notes: All columns control for randomization strata (subdistrict) fixed effects. They also control for date and time of entry into the village, caste category of the treatment hamlet and distance from the village to an urban center. Respondent-level controls include age, gender, literacy and potential seed status. Standard errors (clustered at the village-level) are reported in parentheses and p-values are reported in brackets.

Table H.1 looks at how the volume of conversations changed by treatment, and in particular whether there was differential conversation participation by "seed households" relative to the others. Specifically, this allows us to ask if part of the positive effect on communication in (Seed, CK) relative to (Seed, No CK) is coming from the seed household itself putting in more effort and having more conversations. We remind the reader that every village (even broadcast treamtents) has a set of "seed households." This is because the seeds were chosen using responses to the gossip survey that was conducted at baseline in each village.

In Table H.1, we see that our main results hold for the households that are not seeds: (1) adding common knowledge to seeding increases conversations, (2) broadcasting information to all households without common knowledge raises conversations relative to seeding, (3) broadcasting information to all households reduces conversations if there is common knowledge, and (4) adding common knowledge to broadcasting reduces conversations.

Turning to the seed households, there is a noisily estimated 1.3 increase in the conversation count for a Seed in CK relative to No CK ($p = 0.39$). If anything, this entirely comes from incidental conversations, and one cannot statistically reject an effect size of 0. Note that there is a 0.5 increase in conversations per random non-seeded households. This means that in a village of 50 households, there will be 23 extra conversations. If every seeded household gained 1.3 conversations, then this explains 6.5 or 29% of the increase in conversations. (Even if we assume that there are double the coefficient's number, so 13 conversations, this at best would only explain 56% of the increase in conversations.) Finally, note that by column 3, because the effect is not coming from purposeful seeking or advising behavior, any increase in seed conversations does not appear to be driven by the seed actively going out to explain the information to others, nor others actively seeking out the seeds. Taken together, this suggests that a primary driver of information aggregation here comes from decentralized conversations among non-seeds.

Appendix I. Randomization Balance

Table I.1 presents a balance table, comparing (Seed, No CK), (Seed, CK), (Broadcast, No CK), and (Broadcast, CK) across whether the village is very rural, peri-urban, time of entry for endline survey, date of entry, whether the village was reassigned, gender of subject, literacy of subject, whether the subject has a bank account, and age of subject.

Columns 1-4 present means by covariate in the four treatment cells aforementioned, in that order. Columns 5-10 present p-values of pairwise comparisons of differences in means across cells. Notably of the 54 pairwise comparions, only 11% have a p-value below 0.1 and only 5.5% have a p-value below 0.05. We can therefore see that there is reasonable balance.

Table I.1. Balance

	Means				Pairwise Differences p-values					
	(1) Seed, No CK	(2) Seed, CK	(3) Broadcast, No CK	(4) Broadcast, CK	(5) SNCK - SCK	(6) SNCK - BCNK	(7) SNCK - BCK	(8) SCK - BNCK	(9) SCK - BCK	(10) BNCK - BCK
Beyond 40kms of urban center	.14	.21	.1	.22	.39	.53	.35	.13	.93	.11
Within 5kms of urban center	.31	.4	.35	.31	.41	.73	1	.63	.39	.72
Standardized entry time	-.12	.1	.02	-.21	.23	.49	.65	.71	.13	.3
Survey date	3.55	3.64	3.7	3.76	.54	.26	.12	.64	.36	.63
New strata	.09	.07	.05	0	.83	.53	.05	.67	.05	.09
Female	.32	.25	.33	.39	.25	.91	.29	.17	.02	.29
Literate	.8	.8	.82	.78	.89	.75	.6	.66	.74	.41
Bank account holder	.91	.86	.85	.93	.27	.1	.56	.9	.16	.04
Age	40.01	40.06	38.27	38.24	.97	.12	.15	.14	.16	.98

Appendix J. Instrumenting for treatment assignment

Typically a village has one SCST hamlet and one GOBC hamlet. In conducting our intervention in a small sample of 16 villages, our field staff visited the wrong hamlet. However, we did an endline in these "missed" hamlets, which were intended to receive the treatment, as well though with a slightly smaller random sample. Here we present our main results where we only look at the set of hamlets originally that should have received treatments. We instrument for actual treatment assignment with intended treatment assignment.

Table J.1 and J.2 present versions of our main results with this IV strategy. We see that all our main results essentially go through.

Table J.1. Engagement in social learning

VARIABLES	(1) IV OLS Volume of conversations	(2) IV OLS # incidental conversations	(3) IV OLS # purposeful conversations
CK	0.681	0.464	0.217
	(0.328)	(0.270)	(0.107)
	[0.0380]	[0.0862]	[0.0430]
Broadcast	0.888	0.617	0.271
	(0.377)	(0.338)	(0.141)
	[0.0185]	[0.0679]	[0.0540]
BC × CK	-1.720	-1.236	-0.485
	(0.546)	(0.456)	(0.199)
	[0.00164]	[0.00672]	[0.0151]
Observations	1,068	1,068	1,068
Seed, No CK Mean	0.651	0.514	0.137
CK + BC × CK = 0 p-val	0.00478	0.0145	0.0846
BC + BC × CK = 0 p-val	0.0191	0.0305	0.0759

Notes: All columns control for randomization strata (subdistrict) fixed effects. They also control for date and time of entry into the village, caste category of the treatment hamlet and distance from the village to an urban center. Respondent-level controls include age, gender, literacy and potential seed status. Only outcomes from intended treatment hamlets are used. CK, Broadcast and BC×CK are instrumented with CK in intended hamlet, Broadcast in intended hamlet and BC×CK in intended hamlet. Standard errors (clustered at the village-level) are reported in parentheses and *p*-values are reported in brackets.

Table J.2. Knowledge and decision-making

VARIABLES	(1) IV OLS Knowledge Error	(2) IV OLS Chose 500
CK	-0.0427	0.0459
	(0.0127)	(0.0225)
	[0.000804]	[0.0409]
Broadcast	-0.0327	0.0653
	(0.0147)	(0.0277)
	[0.0261]	[0.0183]
BC × CK	0.0639	-0.110
	(0.0195)	(0.0396)
	[0.00107]	[0.00560]
Observations	1,073	1,057
Seed, No CK Mean	0.436	0.0557
CK + BC × CK = 0 p-val	0.128	0.0361
BC + BC × CK = 0 p-val	0.00887	0.0844

Notes: All columns control for randomization strata (sub-district) fixed effects. They also control for date and time of entry into the village, caste category of the treatment hamlet and distance from the village to an urban center. Respondent-level controls include age, gender, literacy and potential seed status. CK, Broadcast and BC×CK are instrumented with CK in intended hamlet, Broadcast in intended hamlet and BC×CK in intended hamlet. Only outcomes from intended treatment hamlets are used. CK, Broadcast and BC×CK are instrumented with CK in intended hamlet, Broadcast in intended hamlet and BC×CK in intended hamlet. Standard errors (clustered at the village-level) are reported in parentheses and *p*-values are reported in brackets.

Appendix K. Dropping villages from new subdistrict

From our original sample we added 16 new villages from a new subdistrict. Unfortunately, the reassignment was not randomly done, which we discuss at length in Online Appendix L. To deal with this, here we repeat our main results dropping the set of 16 villages that were assigned a new subidstrict. Tables K.1 and K.2 show that all of our main results go through.

Table K.1. Engagement in social learning

VARIABLES	(1) OLS Volume of conversations	(2) OLS # incidental conversations	(3) OLS # purposeful conversations
CK	0.602	0.401	0.201
	(0.333)	(0.275)	(0.111)
	[0.0722]	[0.147]	[0.0703]
Broadcast	0.689	0.495	0.193
	(0.364)	(0.327)	(0.132)
	[0.0601]	[0.131]	[0.146]
Broadcast $\times$ CK	-1.445	-1.065	-0.380
	(0.539)	(0.450)	(0.193)
	[0.00807]	[0.0191]	[0.0499]
Observations	1,020	1,020	1,020
Seed, No CK Mean	0.685	0.536	0.150
CK + BC $\times$ CK = 0 p-val	0.0224	0.0332	0.248
BC + BC $\times$ CK = 0 p-val	0.0387	0.0535	0.128

Notes: All columns control for randomization strata (subdistrict) fixed effects. They also control for date and time of entry into the village, caste category of the treatment hamlet and distance from the village to an urban center. Respondent-level controls include age, gender, literacy and potential seed status. Villages from newly added strata are not included in this sample. Standard errors (clustered at the village-level) are reported in parentheses and p-values are reported in brackets.

Table K.2. Knowledge and decision-making

VARIABLES	(1) OLS Knowledge Error	(2) OLS Chose 500
CK	-0.0372	0.0529
	(0.0130)	(0.0235)
	[0.00474]	[0.0256]
Broadcast	-0.0273	0.0734
	(0.0145)	(0.0275)
	[0.0608]	[0.00839]
Broadcast × CK	0.0539	-0.116
	(0.0194)	(0.0395)
	[0.00589]	[0.00360]
Observations	1,024	1,009
Seed, No CK Mean	0.438	0.0534
CK + BC × CK = 0 p-val	0.228	0.0366
BC + BC × CK = 0 p-val	0.0281	0.0947

Notes: All columns control for randomization strata (sub-district) fixed effects. They also control for date and time of entry into the village, caste category of the treatment hamlet and distance from the village to an urban center. Respondent-level controls include age, gender, literacy and potential seed status. Villages from newly added strata are not included in this sample. Standard errors (clustered at the village-level) are reported in parentheses and *p*-values are reported in brackets.

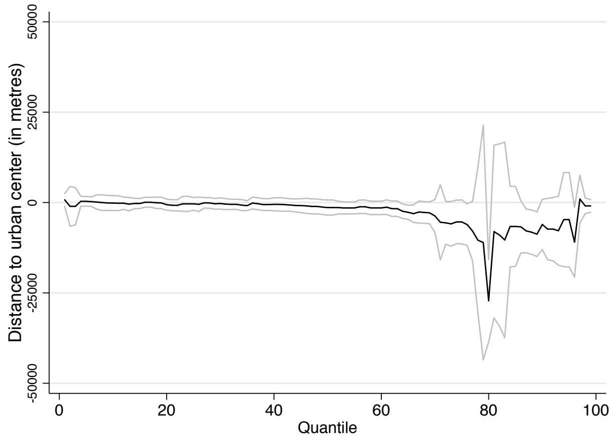
Appendix L. Status Quo Appendix

We also attempted to get 30 villages of data where we did not intervene whatsoever and instead only collected endline data. We call these the “status quo” villages. Unfortunately, these villages are not entirely comparable to our core set. “Status quo” villages are considerably more likely to be peri-urban/neighboring a city, larger in size, more educated, and due to survey logistics were surveyed much closer to the deadline. This was due to the following implementation failures: (1) mechanically, survey teams were less familiar with the “status quo” villages because no treatment was delivered, and unfortunately, they went to these villages after intervention villages. This both pushed the visits closer to the deadline and later in any given day; (2) a share of initially selected “status quo” villages were dropped and the replacements were not randomly drawn from a list of a villages in a subdistrict, placing them city-adjacent; (3) there was geographic imbalance in the initial randomization between “status quo” and intervention villages. Therefore, we do not include these along with the analysis.

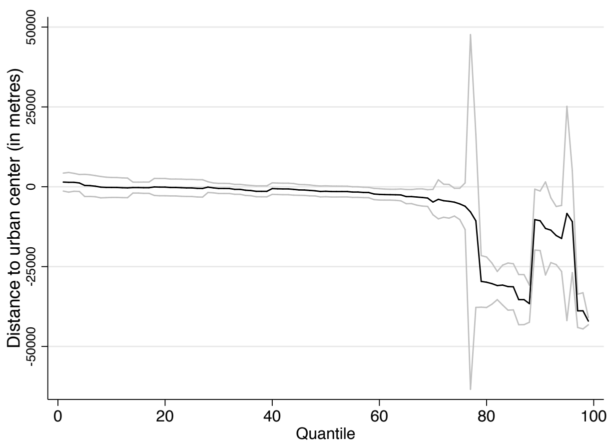
We can include “status quo” in a regression analysis to compare it to our other treatments, but we need to keep in mind that this is observational, and relies on controlling for the distribution of distance from cities, survey timing, etc. That means when we compare to “status quo” we should interpret it with caution. When we do this, we find suggestive evidence that the number of conversations between “status quo” villages and (Seed, No CK) is similar, while (Seed, CK) exceeds “status quo”. Our information and choice analysis have commensurate estimates, but results are noisier.

Recall that the goal of the paper is to understand how changes to the seeding structure affect endogenous participation and subsequent knowledge and choice. The “status quo” treatment cell is unnecessary for accomplishing this.

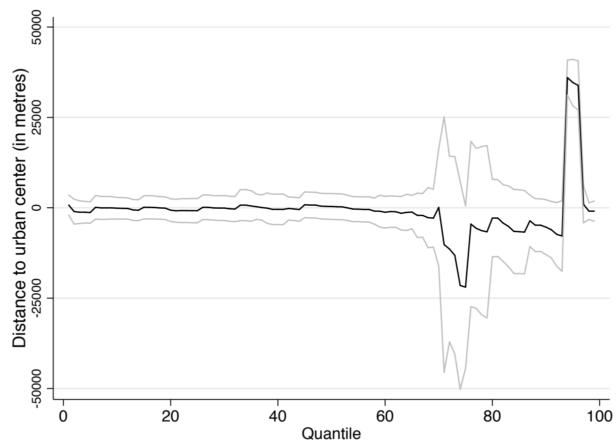
We begin by looking at the distance distributions for the “status quo” and intervention villages. Figure L.1, Panels A, B, and C present coefficients from a quantile regression of distance from urban center against “status quo”, conditional on caste of the hamlet. Panel A conditions on caste, and Panels B and C consider only data from GOBC and SC/ST, respectively. We see that “status quo” hamlets are much more likely to be considerably closer to an urban center particularly in the tail of the distribution.



(A) Controlling for hamlet caste



(B) Only General caste hamlets



(C) Only SC/ST hamlets

FIGURE L.1. Distance to urban center: status quo vs. treated

Table L.1. Imbalance: status quo vs. treated

VARIABLES	(1) OLS Beyond 40kms of urban center	(2) OLS Within 5kms or urban center	(3) OLS Standardized entry time	(4) OLS Survey day	(5) OLS New strata	(6) OLS Female	(7) OLS Literate	(8) OLS Has bank account	(9) OLS Age	(10) OLS Surveyed seed	(11) OLS Surveyed seed
Control	-0.106	0.137	0.312	0.214	0.0488	-0.0223	-0.0349	-0.0101	0.937	0.0326	0.0232
	(0.0508)	(0.105)	(0.175)	(0.109)	(0.0601)	(0.0574)	(0.0427)	(0.0409)	(0.972)	(0.0230)	(0.0104)
	[0.0380]	[0.193]	[0.0764]	[0.0511]	[0.417]	[0.699]	[0.414]	[0.805]	[0.336]	[0.158]	[0.0266]
Observations	1,242	1,242	1,248	1,241	1,248	1,248	1,209	1,244	1,239	1,248	1,248
Treated Mean	0.166	0.345	-0.0539	3.660	0.0536	0.323	0.800	0.890	39.18	0.0518	0

Notes: Columns (1) and (2) are covariates describing distance from the village to an urban center. Column (10) is a dummy for if respondent was a potential seed. Column (11) is a dummy for if respondent was a potential controlling for if the household being surveyed was a potential seed household. Standard errors (clustered at the village-level) are reported in parentheses and p-values are reported in brackets.

Table L.1 presents information analogous to our prior balance table, to show that "status quo" is often imbalanced. Column 1 shows that these villages are much less likely to be very rural, defined as beyond 40km from the nearest city: 6% instead of 16% ($p = 0.038$). Column 2 shows that these villages are 13.7pp likely to be peri-urban, within 5km of a city ($p = 0.193$). These distance imbalances come from several issues. In the original randomization, we were unlucky and had some imbalance. This was compounded by the "status quo" villages not being drawn randomly from a list of villages in the replacement subdistrict (10% of the sample fall into this category and were all within the 61th percentile of distance to an urban center in the treatment distance distribution).

Column 3 and 4 look at time of entry. We see that they were much more likely to be visited later in the day (0.312 standard deviations later, $p = 0.076$) and later during the study period (0.2 days later, $p = 0.05$). The time of day matters because it can affect the composition of which members of which households are home (for instance whether they are working in the field or in town or are home). Furthermore, status quo villages are much more likely to be done about half a day later than the treatment villages.

Columns 5 - 9 show no detectable difference in terms of likelihood of being replaced, a female subject being surveyed, a literate subject being surveyed, the subject having a bank account, nor age. Columns 10 and 11 do show that the respondent is more likely to be a seed, and conditional on interviewing a seed household, the seed himself is more likely to be interviewed.

Table L.2. Experiment Outcomes: status quo vs. treated

VARIABLES	(1) OLS Volume of conversations	(2) OLS # incidental conversations	(3) OLS # purposeful conversations	(4) OLS Knowledge error	(5) OLS Chose 500
Seed	0.00619	0.0483	-0.0421	0.0202	-0.0115
	(0.455)	(0.409)	(0.134)	(0.0183)	(0.0335)
	[0.989]	[0.906]	[0.753]	[0.272]	[0.732]
Seed × CK	0.688	0.342	0.346	-0.0303	0.0399
	(0.345)	(0.276)	(0.125)	(0.0146)	(0.0296)
	[0.0471]	[0.216]	[0.00600]	[0.0392]	[0.180]
Broadcast	0.519	0.352	0.167	-0.00244	0.0584
	(0.523)	(0.479)	(0.157)	(0.0160)	(0.0306)
	[0.323]	[0.464]	[0.289]	[0.879]	[0.0577]
Broadcast × CK	-0.854	-0.621	-0.233	0.0144	-0.0421
	(0.442)	(0.408)	(0.159)	(0.0155)	(0.0290)
	[0.0547]	[0.130]	[0.144]	[0.354]	[0.149]
Observations	1,190	1,190	1,190	1,194	1,179
Status Quo Mean	1.116	0.939	0.177	0.412	0.0793
Seed + Seed × CK = 0 pval	0.128	0.325	0.0231	0.478	0.370
BC + BC × CK = Seed + Seed × CK	0.00294	0.0167	0.00576	0.119	0.725

Notes: All columns control for randomization strata (subdistrict) fixed effects. They also control for date and time of entry into the village, caste category of the treatment hamlet and distance from the village to an urban center. Respondent-level controls include age, gender, literacy and potential seed status. Standard errors (clustered at the village-level) are reported in parentheses and p-values are reported in brackets.

Against this backdrop, Table L.2 presents the main regressions of our paper, bringing in the status quo villages as well, as the omitted category. We are controlling for entry time, survey date, flexibly for distance, caste of hamlet, whether it was replaced, and subdistrict fixed effects. We find similar results to our main results. In column 1 we look at total volume of conversations. As one would have thought, (Seed, No CK) is not appreciably different from status quo, since we only handed out 5 pamphlets and there was no common knowledge of this. Meanwhile, (Seed, CK) is statistically distinguishable from (Seed, No CK), and corresponds to a 0.688 increase in the number of people spoken to relative to status quo ($p = 0.128$). We see that going from status quo to (Broadcast, No CK) leads to a large increase in the number of people spoken to, though this is not statistically distinguishable from zero ($p = 0.323$). However, we can precisely say that adding common knowledge to broadcast reduces the conversation rate relative to (Broadcast, No CK) ($p = 0.055$). And we also see that conditional on common knowledge, going from seeding to broadcast reduces conversations ($p = 0.003$). These same patterns largely hold in columns 2 and 3 across incidental and purposeful conversations, as well as in columns 4 and 5 across knowledge and choice.

Taken together, the evidence suggests that when controlling for sources of imbalance and failures in execution, status quo mostly behaves like (Seed, No CK), whereas (Seed, CK) and (Broadcast, No CK) perform better on conversation and choice metrics.

Appendix M. Attrition

Table M.1 presents p-values from a regression at the village level, among the 237 villages in our baseline, of whether a village dropped out of the study before endline on treatment assignment. We conduct all pairwise comparisons among (Seed, No CK), (Seed, CK), (Broadcast, No CK), (Broadcast, CK), and Status Quo. We find there is no differential attrition of village by treatment assignment. The attrition rates respectively are 7.4%, 5.66%, 5.77%, 2.1%, and 6.25%.

Table M.1. Attrition

SNCK - SCK	SNCK - BNCK	SNCK - BCK	SCK - BNCK	SCK - BCK	BNCK - BCK	SNCK - SQ	SCK - SQ	BNCK - SQ	BCK - SQ
.72	.74	.2	.98	.35	.34	.91	.84	.93	.39

Notes: p-values listed from pairwise comparisons of attrition rates.

Appendix N. Effect on Joint Distribution of Conversations and Information Quality

Here we look at how the joint distribution of conversations and information quality move. Table N.1 presents multinomial logistic regressions. In column 1, the outcome variable takes on values of "Conversations and High Knowledge", "Conversations and Low Knowledge," "No Conversations and High Knowledge," and "No Conversations and Low Knowledge". Therefore we look at whether as we move across treatments, for instance from (Seed, No CK) to (Seed, CK), whether the mass moves towards the joint outcome of both conversations going up and quality of information going up. This provides suggestive evidence consistent with social learning. Column 2 repeats the exercise but where information quality in this case is measured by whether the respondent chose the Rs. 500 note. Figure N.1 presents the same results with raw data.

We find that going from (Seed, No CK) to (Seed, CK) leads to a large increase in the mass of respondents who both have more conversations and have higher information quality (measured by knowledge and choice). The same is the case when comparing (Seed, No CK) to (Broadcast, No CK). However, we see that (Broadcast, No CK) is differentially less likely to both increase knowledge and conversations together, and more likely to push mass into the no conversations cells. This is consistent with a story wherein (Seed, CK) and (Broadcast, No CK) both encourage engagement in social learning whereas (Broadcast, No CK) discourages social learning.

Table N.1. Joint distribution of conversations and information quality

	(1) Knowledge	(2) Rs. 500
Convo_Knowledge		
CK	1.603	1.682
	(0.330)	(0.799)
	[1.18e-06]	[0.0352]
Broadcast	1.648	1.963
	(0.416)	(0.867)
	[7.57e-05]	[0.0236]
Broadcast × CK	-2.351	-2.858
	(0.552)	(1.043)
	[2.02e-05]	[0.00614]
Convo_NoKnowledge		
CK	1.190	1.052
	(0.422)	(0.261)
	[0.00480]	[5.71e-05]
Broadcast	1.114	1.011
	(0.474)	(0.296)
	[0.0188]	[0.000640]
Broadcast × CK	-2.281	-1.661
	(0.667)	(0.405)
	[0.000622]	[4.06e-05]
NoConvo_Knowledge		
CK	0.775	0.350
	(0.279)	(0.362)
	[0.00542]	[0.333]
Broadcast	0.889	0.693
	(0.326)	(0.358)
	[0.00634]	[0.0530]
Broadcast × CK	-1.292	-0.791
	(0.439)	(0.532)
	[0.00324]	[0.137]
Observations	1,082	1,067
Convo, Knowledge: CK + BC × CK = 0 p-val	0.115	0.0342
Convo, Knowledge: BC + BC × CK = 0 p-val	0.0564	0.125
Convo, No Knowledge: CK + BC × CK = 0 p-val	0.0253	0.0503
Convo, No Knowledge: BC + BC × CK = 0 p-val	0.0113	0.0148
No Convo, Knowledge: CK + BC × CK = 0 p-val	0.130	0.288
No Convo, Knowledge: BC + BC × CK = 0 p-val	0.138	0.796

Notes: The table presents marginal effects from a multinomial regression on treatment. In each column the outcome variable consists of whether or not the participant had conversations about demonetization with a measure of information quality. In column 1 this measure is whether the participant has above average knowledge on our test. In column 2 this is whether the participant selected the Rs. 50 note. Standard errors (clustered at the village-level) are reported in parentheses and *p*-values are reported in brackets.

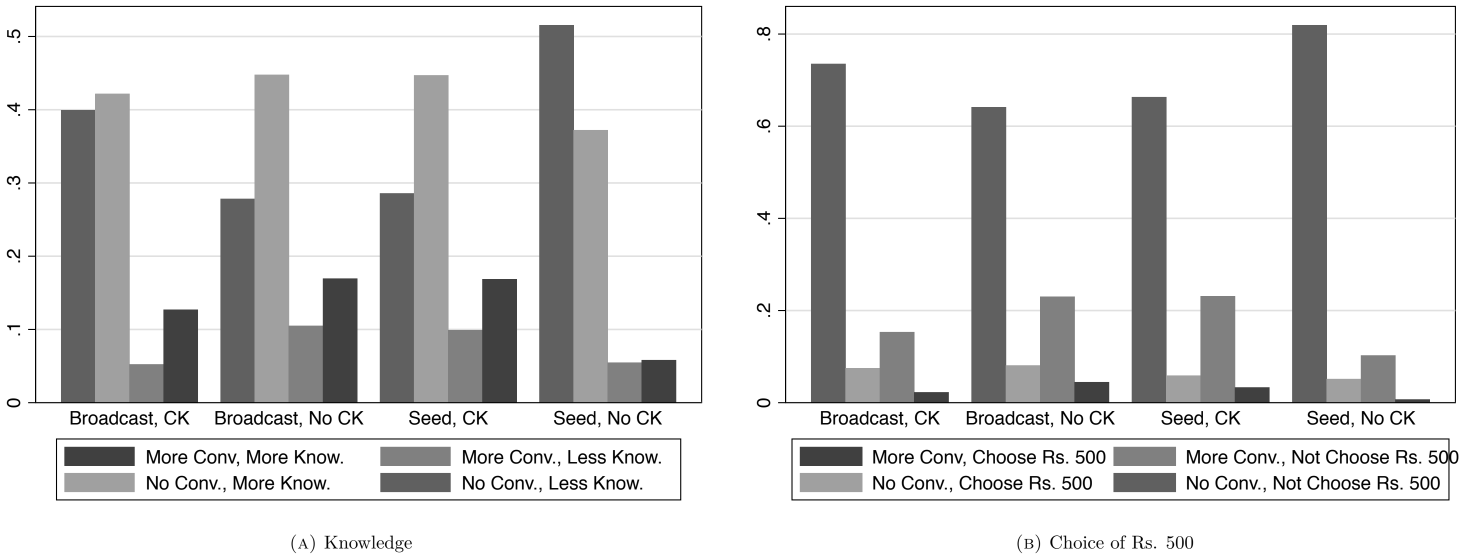


FIGURE N.1. Joint distribution of conversations and information quality