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

ABHIJIT BANERJEE⋆, EMILY BREZA§, ARUN G. CHANDRASEKHAR‡, AND BENJAMIN GOLUB†

Abstract. How should information be disseminated to large populations? The options include broadcasts (e.g., via mass media) and informing a small number of "seeds" who then spread the message. While it may seem natural to try to reach the maximum number of people from the beginning, we show, theoretically and experimentally, that information frictions can reverse this result when incentives to seek are endogenous to the information policy. In a field experiment during the chaotic 2016 Indian demonetization, we varied how information about the policy was delivered to villages along two dimensions: how many people were initially informed (i.e. broadcasting versus seeding) and whether the identities of the initially informed were 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. Under common knowledge, broadcasting performs worse and seeding performs better (relative to no common knowledge). Moreover, with common knowledge, seeding is the more effective strategy of the two. These comparisons hold on all three outcomes.

Date: This Version: October 21, 2019.
We gratefully acknowledge financial support from The Pershing Square Fund for Research on the Foundations of Human Behavior and The National Science Foundation under SES-1658940. We thank Mohammad Akbarpour, George Beknazar-Yuzbashev, Leo Burzstyn, 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 an insightful discussion. We thank Pooja Suri, Arnesh Chowdhury, Piyush Tank, Tithee Mukhopadhyay, Anushka Chawla, Devika Lakhote, Alokik Mishra, Shambhavi Priyam, Aneesha Parvathaneni, Vedika Vyas, and Xuan Zhu for excellent research assistance. Stanford IRB-35996.
⋆Department of Economics, MIT; NBER; JPAL.
§Department of Economics, Harvard University; NBER; JPAL.
‡Department of Economics, Stanford University; NBER; JPAL.
†Department of Economics, Harvard University.

1. Introduction

How should new information that is potentially valuable to a large population be dispersed? For example, during an epidemic of zika, there is a useful list of do's and don't's; how does a government or an NGO get that information to the relevant population? In practice, there are two commonly used strategies: (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] It may seem natural to try to reach the maximum number of people from the beginning as long as the incremental costs of doing so are not too high; indeed, conventional wisdom suggests that broader outreach is better. However, endogenous responses to an information policy are also important to consider, and may reverse this intuition. For instance, suppose people need to ask questions to comprehend the information they were given. This decision depends not only on knowledge but on the social effects of seeking information. If people believe asking basic questions despite being informed is potentially compromising, they may be less willing to ask necessary questions after broadcasts. As a result, they may end up learning less.

Prior work that motivated the current investigation suggests that this is real concern: in a survey conducted prior to our experiment Chandrasekhar et al. (2018) asked 122 villagers, also in India, about their willingness to ask questions of other community members about important practical decisions: 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. (2018) goes on to develop a signaling model of this friction and, in a laboratory experiment, finds that distortions in seeking behavior due to reputational concerns are present and economically significant. This is also related to Bursztyn, Egorov, and Jensen (2018), which provides evidence that signaling concerns strongly influence the choices of high school students, potentially to the serious detriment of their educational careers.[2]

[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] This is also related to a broader literature that uses shame or signaling to change behaviors ranging from tax compliance to voter turnout to vaccinations. See, for example, Butera et al. (2019), Perez-Truglia and Troiano (2018), Gerber et al. (2008), and Karing (2018)..

In other words, people do seem to strategically shy away from asking questions depending on perceptions. This raises the prospect of unintended consequences of informing more people. To investigate whether this is a consequential friction policy-relevant setting, we conducted a randomized experiment during the 2016 Indian demonetization, 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, its chaotic implementation, involving over 50 rule changes in a seven week period, led to widespread confusion and misinformation (see Appendix A). For example, in our sample, at baseline 15% thought that the Rs. 10 coin was also being demonetized though this was never a possibility; 25% did not understand that demonetized currency could only be deposited into a bank account (as opposed to being exchangeable for new bills over the counter).

Our experiment covered over 200 villages in India where we varied the way we provided information about the demonetization rules. One dimension of variation was in how many people we informed: Seed vs Broadcast. A second dimension of variation, motivated by the observation that the signaling concern is most relevant when there is common knowledge about who was informed, was to vary whether to provide meta-knowledge i.e. whether who was informed was common knowledge or not. In the experiment, we compare four possible dissemination strategies: (1) (Broadcast, Common Knowledge): information is broadcast widely to all households in a village, and this fact about the information policy is itself evident to all (as in any standard broadcast method); (2) (Seed, Common Knowledge): information is delivered to a small set of (five) "seed" individuals, and this is again made evident to the community (as in extension services that publicize the identities of model farmers, etc.); (3) (Seed, No Common Knowledge): information is again seeded with a small set of (five) individuals, but no information about this is publicized (as in viral marketing); and (4) (Broadcast, No Common Knowledge): information is dispersed widely but, *unlike* with standard broadcasts, this is done in a way that does not generate public awareness of the delivery strategy. Of these, the first three are realistic options that are used in practice.[3] The last one is more artificial but serves to illuminate the role of common knowledge, which is a key aspect of the theory.

[3]Common knowledge is observed in many different seeding contexts. In the case of agricultural extension programs, for example, announcements about new initiatives, including the identity of model farmers, can be posted in a public place such as a local agricultural cooperative building. Organizations using seeding strategies could alternatively reveal this meta-information in SMSs, village meetings, or quick door-to-door visits.

We present a simple reduced-form framework to aid in organizing the effects of these arms and interpreting our experiment. An individual decides to seek information based on the net expected benefit. Seeking yields valuable information if other individuals in the community are informed. An agent has a subjective probability that others are informed based on the announcement of the policymaker. Seeking also incurs a physical cost. Finally, a friction term captures other considerations, such as equilibrium social costs of seeking or treatment-dependent misperceptions of the value of seeking. We construct testable hypotheses for the benchmark frictionless model, in which the friction term is absent. In that benchmark, the conventional wisdom holds: Namely, holding the number of seeds constant, demand for information (and also knowledge of the policy) is always higher under common knowledge. Another key prediction of the frictionless benchmark is that (Broadcast, No Common Knowledge) should *not* do worse than (Broadcast, Common Knowledge). On the other hand, as we have noted above, distortions in the seeking decision that depend on who is (perceived to be) informed could make the former policy *less* effective. Broadcasting has an obvious advantage over seeding in that information immediately reaches more people, but it also has a potential disadvantage: it may harm the social aspect of learning if fewer people ask when it is common knowledge that everyone got the same information. This can happen for a variety of reasons, including the concern about revealing inability to use one's own information that we set out above. We use the reduced-form model to flesh out some key contrasts between the frictionless benchmark and relevant alternatives.

Our experiment was conducted in the ten days (starting on December 21, 2016) leading up to the deadline when the old Rs. 500 and Rs. 1000 bills stopped being accepted by banks. We randomized how we provided information to villages and varied (1) whether information was provided to all households or to just five seed households; (2) whether it was made common knowledge who was informed within the village; and (3) the number of facts provided, which could be either 2 or 24. The information we provided always consisted of a list of facts in a short printed pamphlet, and 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 circular (released on December 19th, 2016), and thus contained the information that the policymakers themselves chose to communicate to the public. We 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 policy. 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.

We measure three outcomes: 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.

From a policy perspective, we care which of the three core strategies leads to the greatest social learning. We find that (Seed, Common Knowledge) dominates in this setting. In terms of studying the mechanisms, we are interested in whether the reversals predicted above hold. We find strong evidence for each type of reversal.

First, we look at endogenous participation in social learning.[4] 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).[5] Furthermore, (Broadcast, CK) leads to 61% *fewer* 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 are unlikely to obtain in models where there is no strategic motive behind participation in social learning.

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

[5]We often abbreviate Common Knowledge by CK.

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) increases the knowledge index by 5.6% ($p = 0.0142$). On the other hand, going from (Seed, CK) to (Broadcast, CK) leads to a 3.1% *decrease* in the knowledge index ($p = 0.062$). This shows that even though all households, rather than just five, are given signals, 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 4.9% increase in the knowledge index ($p = 0.053$). Within broadcast, (Broadcast, CK) has a 3.8% lower knowledge index than (Broadcast, No CK), though the effect is not statistically significant ($p = 0.17$).

Third, we look at the incentivized decision – 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 an 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 ($p = 0.104$). In 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, we find that in a policy-relevant context, perhaps counter-intuitively (Seed, CK) is the best of the available policy-relevant strategies. Moreover, removing common knowledge under broadcasting – which, admittedly, is generally not an available option for policymakers – leads to increased learning.

The results indicate that the assumptions of the benchmark frictionless model do not hold. Indeed, the comparisons between the experimental treatments imply a friction that affects seeking *differentially* across treatments – in particular, one that is more severe when it is common knowledge who was informed.

Next we present a richer signaling model, built around the motivating hypotheses of our experiment. This signaling model explicitly incorporates (i) potentially asymmetric information about who was informed and (ii) about how able individuals are to understand information they were given. We can then more fully model the dependence of seeking behavior on the environment. Public announcements that information is present increase the perceived value of seeking, but they also create room for seeking to signal ability through channel (ii) if they make it common knowledge that everyone was informed. This theory can explain, in a more detailed way than the reduced-form one, why (Seed, CK) is so effective: it accomplishes the maximum common knowledge that information is out there, while minimizing signaling concerns for the average person. It also helps us understand what conditions allow for (Broadcast, No CK) to be so effective – a subtlety that we discuss in this section.

We also discuss whether several alternative models can explain our experimental findings. First, we consider conventional models of social learning including "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); we also look at models of Bayesian information aggregation. We explain why, in these models, forces that make for markedly worse collective learning when information endowments are uniformly improved are largely absent. By emphasizing the role of asking questions and the strategic choices involved in doing so, this paper highlights the importance of a force relevant for the large and growing literature on social learning, but not typically studied in social learning models. Finally, we consider in light of the data several alternative behavioral models that could be devised to explain our findings, including ones based on variations in curiosity or treatment-dependent misperception of the value of seeking information. We argue that, on the whole, a mechanism based on image concerns has substantial advantages in explaining the data parsimoniously, though we do not aim for a full decomposition of effects.

The remainder of the paper is organized as follows. Section 2 describes the context and setting. Section 3 describes the experimental design, motivated by a basic theory, and its implementation. We present our empirical results in Section 4. Section 5 presents the details of a theoretical framework where agents endogenously choose to participate in social learning. We also compare the predictions to those of models in which signaling-based seeking frictions are not present and argue that, on the whole, signaling offers a more parsimonious explanation. Section 6 provides a discussion.

2. Context and setting

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. Given the hamlet structure of the study area, all of our treatments and outcomes were focused on 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 Panel A of 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;[6] 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 small fraction 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).

Panel B of Table 3 shows the incidence of the respondent reporting to us that they "don't know" the answer to the question.[7] While almost all respondents believed they knew which notes were being demonetized, more than 30% of respondents reported that they did not know about the withdrawal limits or how to deposit the demonetized notes anywhere besides a bank branch. This suggests that a large fraction of individuals were willing to acknowledge to us (and thus, to themselves) that they were uninformed about important aspects of the policy.

One might ask whether it was important for relatively poor households with limited formal savings to understand various details of the policy. One major 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 (and not just knowing a little about them) would have been useful for a variety of decisions: e.g., whether to accept an IOU from an employer or customer, or how much inventory to order for a small business.

---

[6]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).

[7]If a respondent answered "don't know" to any of the questions, they were then asked to make their best guess. These guesses are included in our measures of errors in Panel A.

3. Experiment: Design and implementation

3.1. **Motivation for the experiment.** Our experimental design was motivated by our prior work in (Chandrasekhar, Golub, and Yang, 2018). As part of that study, we conducted field surveys in rural villages in India on several topics, asking about information seeking behavior by villagers in domains of finance, health, and agriculture. Individuals report significant worries about being perceived as ignorant in these domains and therefore limit how much they engage in social learning with others. That is, they face some image-based limits on seeking advice. Further, Chandrasekhar et al. (2018) demonstrates experimentally that ability-signaling plays a substantial role in driving this reluctance: it is precisely when needing help signals low ability that seeking is most suppressed. We have since supplemented this evidence by conducting surveys to document that individuals felt that (a) the demonetization was very confusing, (b) those in the know would judge others to be "ignorant" or "dumb" if they sought out information, given that information about the policy was widely broadcast, and (c) owing to this, they curbed their own seeking behavior. We thus focus on the demand for information; the motivating evidence is discussed further in Section 5.1.[8]

Our experiment also directly answers a policy question of interest. It tells us for example that a policy of simply blanketing information everywhere, which corresponds to (Broadcast, CK) in our experiment may not always be the optimal policy.

3.2. **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 villager in the study area.[9] Our experimental protocol involved giving a randomly-selected set of facts to each village – below we describe exactly how the selection was done. All individuals receiving lists of facts in a village received the same list.

Our core design is a 2 × 2 that varies how many people got information as well as whether there was common knowledge. Because another important dimension for

[8]We discuss supply considerations in Section 4.1.1 below.

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

information policies is the volume of information given, we added an arm varying whether villages 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:[10]

(1) Information dissemination:
  - *Broadcast:* information was provided to all households in the hamlet.[11]
  - *Seed:* information was provided to 5 seed households in the hamlet, chosen via the gossip survey.[12]

(2) Common knowledge:
  - *No Common Knowledge (No CK):* we did not tell any subject that we were providing information to anyone else in the community.
  - *Common Knowledge (CK):* 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.

(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.[13]

[10]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.

[11]Pamphlets were dropped off at every household.

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

[13]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.

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.[14]

3.3. **A simple model.** We present a simple, reduced-form framework to organize our analysis of how the treatments affect demand for information. This model allows us to outline what we would expect to happen without treatment-dependent information frictions such as the signaling concerns that motivated our study, and also equips us to detect frictions induced by our information policies. A discussion of our preferred richer version of this model as well as alternatives can be found in Section 5.

The key endogenous margin we focus on is individuals' decisions of whether to seek information.[15] Individuals are trying to learn the state of the world, and the choice each makes is whether to ask questions and engage in conversations that are helpful for learning. Considering a representative individual, there are five basic moving parts affecting the decision.

- Information endowment: Different treatments vary how much people are informed, and seeking information is more attractive to those who have less information to start with. Let e denote the value of the informational endowment.
- Seeker's subjective probability of seeking being successful. This captures the probability of finding an informed source, conditional on seeking. We denote this by p.
- Perceived incremental value of information, above the endowment, conditional on seeking being successful. Let $v(e)$ denote this value, which is assumed to be decreasing in the endowment e.[16]
- A fixed cost of seeking information. Seeking information has an opportunity cost of time, as well as physical costs, etc.; we denote the total such cost by c. This fixed cost does *not* depend on anyone's beliefs about the seeker, and it also does not depend on the likelihood of finding information or its eventual value.
- A possible treatment-dependent friction, f, that increases the effective cost (or, equivalently, reduces the perceived benefit) of seeking. Some examples of what f might encode include:

[14] 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.

[15] In principle, this framework and simple adaptations of it could also model decisions to spread information. We justify our focus on seeking in Section 5.

[16] In a more realistic model, this will depend on who else participates and other social learning dimensions; we abstract from this here, but see Section 5 for more on this point.

– Image concerns: "If I received information that intelligent people can process easily, others who know this may think negatively of me if I need help making sense of it." Here f might be especially large in the (Broadcast, CK) treatment. This foundation motivated our experimental design.
– Expectations about conversation topics: "If I seek to have a conversation about something that was broadcast to everyone, it will be a boring conversation (or considered boring)."
– Perceptions (not necessarily well-calibrated) of the value of information[17]: "If the same information was given to everyone, I will not learn much by engaging in conversations."

We now consider how our treatments shed light on the sign and magnitude of f. We take v to be the same function across treatments, and p to be a number that depends only on what is announced about where information can be found. For instance, p is higher when it is common knowledge who the seeds are, because then information is easier to find. The endowment e depends on the treatment. Importantly, we focus on an individual who is not a seed, so receives a larger endowment e_h in the Broadcast treatments, and a smaller one (whatever he has outside our experiment), e_l in the Seed treatments. Finally, we take c to be an idiosyncratic draw from a distribution F.[18] The quantities p, f, and the endowment e depend on the treatment, though we often leave this dependence implicit.

The simplest model is one where the values enter in an additively separable way, so that the net return to seeking is $e + pv - c - f$ (the expected value of seeking net of costs), and the return of not seeking is e. Seeking occurs[19] if $e + pv - c - f \geq e$. Thus, aggregating across individuals, the amount of seeking we see in a particular treatment should be $F(pv - f)$. More generally, it is reasonable to assume that the amount of seeking is $\mathcal{F}(p, v, f)$, which is increasing in the first two arguments and decreasing in the last, and our key reasoning will work for such a function subject to suitable technical assumptions, though we stick with the simpler parameterization for simplicity.

3.3.1. *Frictionless model.* Our strategy is to first consider how the rate of seeking depends on the treatment in a frictionless model where $f = 0$ – i.e., where there are

[17]When nonzero, f could be a function of of other primitives of the model, such as p, as would be natural when it essentially modifies v.
[18]One could elaborate the model to have randomness in other quantities, not just c; the results would generalize under suitable conditions, but we focus on the simplest case for ease of exposition.
[19]Up to a tiebreaking rule which is inconsequential for an atomless c.d.f.

no communication frictions or other distortions. We then describe how the shifter f changes predictions relative to this frictionless model.

The incremental value of information $v(e)$ is decreasing in e, holding other parameters fixed. So a higher endowment will, all else equal, lead to less seeking. From now on, consider a *non-seed* individual – one who is not a candidate for being endowed with information in any of our interventions.[20] In terms of the experimental treatments, the endowment of this individual, e, will be low in the Seed treatments and high in the Broadcast treatments. Turning now to p, the probability of successful seeking, this depends both on the number of people informed and the information available about who they are. Under Seeding without Common Knowledge, our non-seed individual does not know there is something to be learned and does not know who is informed, so p is low. On the other hand, the Common Knowledge treatments have two effects: they inform the participants that there is something to be learned, and they also make it easier to know whom to ask; both of these raise p. We would therefore predict the following about the demand for information and its empirical analogue, the volume of conversations. We also comment on related outcomes – level of knowledge and decision quality – that depend on information acquisition.

A. Demand is higher under (Seed, CK) than under (Seed, No CK). As we have explained, in (Seed, CK) p is higher due to CK. On the other hand, subjects have a lower endowment, in both cases. Thus $F(pv)$ is higher under (Seed, CK) than (Seed, No CK). Moreover, we expect that the level of knowledge and the likelihood of making the right choice will also be higher with (Seed, CK) than with (Seed, No CK).

B. Demand is higher with (Broadcast, CK) than with (Broadcast, No CK). With CK, people know that others have potentially useful information, while with No CK they have less reason to believe this. Thus, p is higher under CK. Endowments e are the same in both cases. Thus $F(pv)$ is higher under (Broadcast, CK). The level of knowledge and the likelihood of making the right choice are also higher with (Broadcast, CK) than with (Broadcast, No CK).

The comparison of (Seed, CK) and (Broadcast, CK) is less straightforward because the endowment e is higher under Broadcast, lowering v, but at the same time it may be easier to find someone to talk to – more people have been informed, making p higher. As a result, we have:

C. Demand may be higher or lower under (Seed, CK) than with (Broadcast, CK). Moreover even if demand for information is higher under (Seed, CK), the measured level of knowledge and the likelihood of making the right choice may still be lower under

[20] In practice, one who would not be selected by our procedure to find the seeds.

(Seed, CK) than with (Broadcast, CK) since, with Broadcast, the initial information endowment is higher.[21]

D. When we consider (Seed, No CK) and (Broadcast, No CK), in both cases, our representative individual has no particular reason to believe that there is information to seek out, so p is similar across the two and low. Therefore demand is likely to be low in both cases. However, with broadcast e is higher, so the level of knowledge and the likelihood of making the right choice are lower under (Seed, No CK) than with (Broadcast, No CK).

These predictions are compatible with the conventional wisdom that more information is generally better.

3.3.2. *Model with frictions.* Within this reduced-form framework, deviations from the above predictions suggest a treatment-dependent f that is not equal to 0. Diagnosing deviations from the frictionless benchmark is already valuable, as reversals in the orderings implied by the frictionless model have direct implications for the design of real-world information dissemination strategies, regardless of their origins.

In this subsection, we briefly sketch the nature of the frictions that we expect under the alternative that seeking demand is affected by signaling concerns. We discuss a richer model providing a formal foundation for these in Section 5.2 below.

In (Broadcast, CK) it is known that everyone got the information. If evidence of incomprehension is considered compromising, the friction f would be high. On the other hand, f should be especially low under (Seed, CK) because everyone knows that the person asking did not get the information. The two other cases, (Broadcast, No CK) and (Seed, No CK), are similar: nobody knows who was informed in either case, and therefore any social friction of asking will tend to be low.[22]

The introduction of the friction makes a difference mainly for the comparison of (Broadcast, CK) with (Broadcast, No CK). Predictions A and D continue to hold.[23] However, instead of B we will get:

B': To compare (Broadcast, CK) with (Broadcast, No CK), first recall from B that p is higher under CK, while endowments e are the same in both cases. If (and only if) f is significantly higher under (Broadcast, CK) than under (Broadcast, No CK) can

[21]Recall that the expected value of information ultimately obtained is $e + pv$.

[22]However, it is possible that the (Broadcast, No CK) case could set off discussions among those who got the information and people will find out that in fact everyone was informed. That would make this case more like (Broadcast, CK).

[23]In A, if anything, the friction should be lower under (Seed, CK), which makes for a larger effect in the same direction. In D, the frictions should be low in either case.

there be lower demand under the former. In that case, there would also tend be worse knowledge and more wrong choices, as the information endowments are the same.

By the same logic we also have the following when comparing (Seed, CK) and (Broadcast, CK):

C'. If under (Seed, CK) the friction f is substantially lower than under (Broadcast, CK) then we would expect less demand under the latter. The level of knowledge and the likelihood of making the right choice could both end up lower, if the reduction in the seeking rate more than compensates for the lower endowment of the typical individual under (Seed, CK).

3.3.3. *Implications for our main comparisons.* The difference between the predictions of B and B' as we compare (Broadcast, CK) with (Broadcast, No CK) underlies our most basic rejection of the frictionless benchmark. We will also compare (Seed, CK) with (Seed, No CK) to verify our prediction A, and confirm that the perverse effect of common knowledge is only in Broadcast. Finally we will compare (Seed, CK) with (Broadcast, CK) to check if the difference in the cost of asking can overturn the natural advantage of broadcast, as suggested in prediction C'.

Jointly, the results we find will provide support for the frictions we have postulated. The model we have sketched is reduced-form, and yields comparisons only of some treatments based on simple monotonicity arguments. Our empirics actually yield a richer array of comparisons, for example allowing us to examine the sign of the interaction of the BC and CK treatments. In Section 5.2, we discuss a more detailed model that underlies our discussion of the frictions above, and also helps us to interpret the empirical results more fully.

3.4. **Sample.** We enumerated an initial list of 276 villages which were assigned to treatments. Given the short time frame required of the experiment, we were not able to scout villages before conducting the baseline. Because of this, our initial listing of villages included places where the research team had been before over the course of work on Breza et al. (2017), Breza et al. (2019), and Kaur et al. (2019).[24] We required that all villages in the study have a hamlet structure (the predominant village organization in the study area) and that each hamlet had at least 20 households. 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 included in our experiment and the other half to have their SCST hamlet included in our experiment. We randomized

[24]The presence of researchers in these villages had ended many months before the baseline survey was conducted for this study.

villages to treatments before we verified that each village met our criteria leading to only a set of 221 villages being assigned treatments. Sixteen villages were then added in a new subdistrict to increase the sample to 237.[25] 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.[26]

For each survey round, the enumerators selected households using standard circular random sampling. We asked to speak with any adult permanent resident of the household. Almost all of the survey refusals were from households in which no adult permanent resident was home at the time of the enumerator's visit.[27] For speed of execution, we drew fresh samples of households for the endline survey.

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.[28,29] As described below, the gossips were selected to receive the information in our seeding treatments, but we are able to identify the counterfactual seeds in the other treatment arms using this survey module.

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

[26]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.

[27]In these cases the enumerators made at least two additional attempts to conduct surveys in the day of the visit. The biggest reason for doorlocks was time of day – it was much easier to find respondents early in the morning or in the evening. Because surveyors were dispatched to villages in randomized order, we control for time of entry in the village in all of our main regression specifications.

[28]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.

[29]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.

3.5. **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 assessed knowledge of facts surrounding demonetization. We surveyed 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 probability of 1/6, 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.[30] 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 on the data we collected, the median wait time 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.[31] 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

[30]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.

[31]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 reality.

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.[32]

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 engagement in social learning, measured by the number of conversations the subject had over the prior three days about demonetization. Results are from regressions of the following form:

$$y_{ivd} = \alpha_d + \beta_1 Seed_v + \beta_2 CK_v + \beta_3 Seed_v \times CK_v + \gamma X_v + \lambda X_i + \varepsilon_{ivd} \tag{4.1}$$

where i indexes the individual respondent, v indexes village, and d indexes the sub-district, which was our unit of stratification. In each regression, (Seed, No CK) is the omitted treatment arm, and our key coefficients of interest are $(\beta_1, \beta_2, \beta_3)$. Village-level controls X_v include date and time of entry into the village, the caste category of the hamlet both treated and surveyed in the village, and distance from the village to an urban center. The respondent-level controls X_i include age, gender, literacy and potential seed status. Standard errors are clustered at the village level.

Table 4 presents regressions of the number of conversations on the various treatments.[33] 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 × 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 × CK = 0) allows us to compare (Broadcast, CK) to (Broadcast, No CK), which represents a direct test of the frictionless model. The test (Broadcast + Broadcast × CK = 0) further allows us to compare (Broadcast, CK) with (Seed, CK).

The outcome variable in column 1 is the number of conversations about the demonetization in which the respondent took part over the prior three days. Going from (Seed,

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

[33]For all of our main results, we focus on our core 2 × 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.

No CK) to (Seed, CK) increases the number of conversations by 103% (0.65 more conversations, $p = 0.04$). This result alone is consistent with the frictionless model detailed above. Here, adding information about the identity of the seeds increases engagement with the policy.

While we focus on the demand for information in interpreting these results, in principle, adding common knowledge to the seeding strategy could have affected the supply of information as well. Seeds may have had a stronger motivation to spread information under (Seed, CK). However we do not think a supply response drives our results for two reasons. First, in Online Appendix H, we show the same regression 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.[34] 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 in 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 Section 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 compare strategies that employ common knowledge. Going from (Seed, CK) to (Broadcast, CK) – which typically corresponds to a tenfold increase in the number of households informed (from 5 households to all households) – leads to a 61% *decline* in the volume of conversations (0.78 fewer conversations, $p = 0.029$). Here, we see a reversal of the "more is more" intuition – increasing the amount of information in a community decreases engagement in conversations. One simple explanation might be that because everyone is informed, there is less need for conversations. However, given how little people know even in (Broadcast, CK) villages and endline, this seems

[34]Recall 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.

unlikely – and, as we will see, difficult to reconcile with what happens when we go to (Broadcast, No CK).

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 to the previous result, 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, $p = 0.048$). 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 for a reversal relative to the frictionless model is available when we compare (Broadcast, No CK) to (Broadcast, CK). This leads to a 63% *decline* in the volume of conversations (0.84 fewer conversations, $p = 0.02$). This result is consistent with our predictions in Section 3.3, and highlights that indeed, the frictionless model cannot rationalize our experimental findings.

In sum, our results show that common knowledge affects considerably the decision to engage with others to discuss the policy. When only a few individuals are seeded, it greatly increases aggregate conversations. We have also shown evidence for two non-monotonicities: 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 4 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.[35] Consistent with that, column 3 of Appendix Table H.1 shows that the increase in the quantity of conversations when seeds are CK 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. The primary driver of the increase in conversations here is conversations among non-seeds, and we see no evidence of an effort by seeds to coordinate conversations about the topic.[36]

4.2. **Information aggregation and choice.** We present regressions in Table 5 which show how knowledge of the demonetization rules and incentivized choice behavior depend on the randomized information environment. Recall that the quality of the respondents' choices depended on their understanding of the demonetization rules. We again present estimates of equation 4.1.

In column 1, we turn to whether the changes in endogenous participation in conversations correspond to changes in knowledge. This is primarily an empirical question. 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 for non-monotonicities in learning 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, we would expect (Broadcast, CK) to out-perform (Seed, CK) in terms of knowledge, even with fewer conversations.

We find evidence for strong reversals and a departure from "more is more". The outcome variable is our knowledge metric, which is based on the answers to 34 questions about the demonetization policy asked at the endline.[37] The mean in the (Seed, no CK) group is 0.566. Going from seeding to broadcast under common knowledge leads to a 3.1% *decrease* in the knowledge index ($p = 0.062$). This shows that though 100% of

[35]However, we acknowledge that the *relative* increase in conversations is larger for the purposeful variety.

[36]The fact that individuals largely discuss demonetization via incidental conversations is consistent with the structure of our more detailed theory (described below in Section 5): information aggregation occurs when individuals access conversations occurring in public places; in the model of Section 5, our metaphor for such locations is a "Town Square".

[37]Recall that our treatment only gave information on a small subset of these 34 facts. We explore whether knowledge improvements are driven by the facts that were actually on the pamphlets in Appendix G.

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 3.8% *decrease* in knowledge, though the effect is not statistically significant ($p = 0.174$). In addition, going from (Seed, No CK) to (Seed, CK) increases the score on the knowledge index by 5.6% ($p = 0.0142$) and going from (Seed, No CK) to (Broadcast, No CK) actually makes people better informed and improves knowledge by 4.9% ($p = 0.05$). It is worth noting that we see reductions in knowledge exactly where we see conversations declining. This pattern is consistent with individuals transmitting useful information in their conversations.

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 ($p = 0.014$). These results are fully consistent with the results on conversations and knowledge. More conversations led to better knowledge, which in turn, allowed for improved decision-making.

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. 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. Mechanisms underlying the information frictions

The experimental findings discussed in the previous section are inconsistent with the frictionless model in Section 3.3; instead, they imply some level of seeking frictions. To provide structure on those frictions beyond the reduced-form discussion, this section offers a fuller exploration of the mechanism that motivated our experiment. As we will see, when we explicitly include incomplete information about who is informed as well as about ability to use information, the resulting signaling motives can be subtle. Nevertheless, under assumptions guided by the structure of the experiment, such a model can, fairly parsimoniously, explain the empirical deviations from the frictionless model and provide foundations for the assumptions we made about frictions in Section 3.3.2.

The specific framework we develop builds on recent work of Bursztyn et al. (2018) and Chandrasekhar et al. (2018) on how reputational or signaling incentives affect information-seeking. The idea of the model is very simple: individuals differ in a ability, and there is asymmetric information about people's abilities. The sort of ability in question makes people better at understanding and using information. As a result, conditional on being endowed with some information, high-ability individuals value additional conversation and clarification on that topic less than those with lower ability. Suppose that it is common knowledge that a signal was broadcast to everyone. Then seeking clarifying conversations reduces the likelihood in the public eye that a seeker has high ability, and that makes it costly for some people to ask questions. Conversely, when most people are not believed to have been endowed with information, seeking information is much less correlated with low ability, and therefore seeking will be relatively unimpeded.

After providing some qualitative evidence on the motivation behind this model, we outline it in more detail (relegating details to an appendix). We then interpret our experimental treatments through the lens of this model and derive predictions about how the treatments affect engagement and learning. In Section 5.3, we consider some alternative explanations, including ones based on alternative behavioral assumptions, endogenous supply of information, and network-based distortions. We argue that the phenomena we observe appear, on the whole, most consistent with endogenous seeking decisions.

5.1. **Motivating evidence for signaling concerns.** Our motivating hypothesis – that people's desire to seek out clarification, even when it is needed, may conflict with their desire to signal desirable attributes – came out of conversations about demonetization during the field-scoping phase of the project, and was also motivated by prior work in a similar setting (Chandrasekhar et al., 2018). That paper develops theory and also reports both an experiment and a field survey. The survey asked villagers how they seek information on several topics: farming, health, and household finance. 88% of respondents reported feeling constrained in seeking advice from others, and of these, 64% felt the reason they were constrained was that they did not want to appear "weak" or uninformed. These rates were similar across all three economic domains covered by the survey. In the field experiment, we find that when signaling concerns are switched on in a controlled experiment, there is a 55% decline in the probability of a low-ability subject seeking out information that has a high monetary return.

In 2018, in order to examine reputational considerations of this sort in the context of demonetization in rural villages, we conducted a survey of 102 randomly-selected subjects across 4 villages in rural Karnataka, India.[38] We discuss the results of the survey below, after displaying some representative quotes.

When considering the period of the demonetization, individuals recall feeling confused or knowing others were confused. They also reported that, because information was abundant, asking for clarification was potentially compromising.

> "There was confusion about where to deposit money, how much to deposit, where to withdraw from, where all money could be deposited and last date. People hesitate to ask because they may think, 'even after showing so much on TV, if I ask, what will they think of me. They will think I don't understand.' " – Respondent 1

> "People with more money hesitate to ask because they will worry what others will think about them [. . .] Others will think, 'Don't they know anything? People with money should know more. But if they are still asking, they must be of less intelligence.' " – Respondent 2

Relatedly, individuals who understood the key points of the policy reported judging others for not understanding them.

> "If someone didn't exchange money till December, they must definitely be the biggest *bewakoof* (fool) in the world." – Respondent 3

> "Not everyone knew the deadline and application process. In December if someone comes and asks even after showing on TV, I will think they are dumb. They can't understand so they must be unintelligent. Fearing that

[38] We used random circular sampling in each village to draw our respondent sample.

> others will think like [me], some people who were confused didn't ask." – Respondent 4

And this of course reinforced the hesitancy to ask in the first place. That is, people were indeed cognizant of such judgments.

> I came to know a little later that I had 2 old notes with me. I didn't exchange because I didn't know when the last date was. If I ask someone, I was worried what they will say about me. What will people think? They will say, 'Were you lazy? Were you sleeping till now? Everything was shown on TV.' " – Respondent 5

Turning to quantitative summaries of the survey responses: 80% of respondents said they felt confused, and 79% felt that even at the end of the demonetization period, they did not understand the note-ban's policy relevant implications completely. 94% reported that others in the village were confused as well. At the same time, 96% of the individuals felt that people were responsible for understanding the policy. If someone in the village asked about the policy in December (after extensive public information campaigns), 80% of respondents said that the individual would seem unintelligent, while 85% said the individual would appear irresponsible. Finally, 85% said that even if they were confused, they held back from asking questions of acquaintances for fear of being judged. Figure 5 displays the results.

In short, this is a setting in which individuals felt confused; felt that confusion was associated with being unintelligent or irresponsible; worried that seeking out information would therefore look bad; and therefore reduced their information-seeking. Though a large fraction of people were somewhat confused themselves, they readily admitted they were willing to pass judgment on others who did not understand how to behave. The model we develop reflects these features and examines the equilibrium behavior they imply.

5.2. **Outline of our endogenous seeking model.** Here we outline the model, deferring details to Section D. In the model, we focus on the choice of one focal individual, a decision-maker we call "D," of whether to seek or not. For simplicity in this exposition, we make D a non-seed individual. The timing is:

- First, a policymaker chooses a policy, which determines both the breadth (Broadcast, Seed, None) of the dissemination and any public announcement about it.
- When information is present (i.e., when Broadcast or Seed was chosen by the policymaker), information is available in public places in the village.

- The individual learns an idiosyncratic draw – his value of seeking that period – and can choose whether to seek information; this decision is observed by others.

The reason that the prospect of being observed matters is that individuals differ in their ability to understand information on their own, and value being perceived as more able in this sense. Lower-ability individuals have more to gain by seeking, but also stand to lose reputation if they do so.

The expected benefit of seeking depends on what D knows about who is informed: it is greater when more people are believed to be informed. In the CK treatments, forming such a belief is straightforward, but in other cases it involves inferences by D. At the same time, the endogenous reputational cost of seeking depends on what others know about whether the Seeker is informed. We examine how the information policy affects both margins, and derive detailed predictions about the seeking rates in all four treatments. The model is richer than the reduced-form one from Section 3.3.2 but extends the predictions of that model and obtains more precise implications about the interaction of Broadcast and Common Knowledge treatments.

We present the implications. We focus on the volume of seeking (conversations), and comment on other outcomes afterward.

(1) *(Broadcast, No CK) dominates (Broadcast, CK).* In both cases, D assesses the same informational benefits. But under CK, seeking has the potential to signal low ability, whereas under No CK, under natural assumptions about priors, observers do not consider it likely that D is informed, and therefore do not infer much from his behavior.

(2) *(Seed, CK) dominates (Broadcast, CK).* In both cases, D assesses the same informational benefits of seeking, but has less information under Seed. Moreover, Broadcast turns on signaling concerns (since it is known D got information) whereas Seed makes it plain that D is uninformed, eliminating them.

(3) *(Seed, CK) dominates (Seed, No CK).* In the latter case, there is no reason to expect information to be available, whereas in the former, it is known that it can easily be found. Signaling concerns are small in either case, because others either know D is not informed or have no reason to believe that he is.

(4) *(Broadcast, No CK) dominates (Seed, No CK).* In the former case, the fact of D's being informed makes it likely, from his perspective, that information is out there (it is rare for exactly one person in the village to get a pamphlet – in

practice, pamphlets always come during at least a Seed-type intervention).[39] On the other hand, there is no reason for a non-seed (which, recall, D is) to believe there is information in (Seed, No CK). Signaling concerns are not substantial under (Broadcast, No CK) because D does not think (given the absence of common knowledge) that others see him as informed.[40]

In other words, in our model more is not always more. Under No CK, the Broadcast arm *increases* engagement by alerting people to at least the existence of information, without activating signaling concerns, as explained above in (4). But under CK it *decreases* engagement: with CK, people know the existence of information regardless, but Broadcasting makes it clear that D is informed and activates ability-signaling concerns.

So far we have focused on volume of conversation, rather than knowledge or choice outcomes. We consider these other outcomes in the appendix. These comparisons are more delicate, because in some cases the information endowment is decreased even as engagement in learning is decreased. For cases (1) and (3), the results noted above for volume of conversations extend straightforwardly, as endowments do not change in the comparisons. Nevertheless, we show that under assumptions that are reasonable in our setting, the other comparisons also extend. For changes in endowments not to reverse our effects, we need that social learning is important enough for enough of the population, relative to private processing of information. The details are in Section D.

5.3. **Some alternative models.** The predictions about the perverse effect of more knowledge contrast with many conventional models of social learning, which suggest monotone benefits. For instance, in "infection-type" models often used to study information transmission (Bass, 1969; Bailey, 1975; Jackson, 2008; Jackson and Yariv, 2011; Aral and Walker, 2012), if more individuals are seeded with information, ultimate diffusion is improved. We flesh out a version of this benchmark and the monotonicity result, in a learning setting in the section immediately below. It is also worth considering whether sufficiently strong "less is more" forces can be produced by standard

[39]This implication actually requires some assumptions on priors about information-delivery – most importantly, that broadcasts without common knowledge are uncommon, which we formalize in the appendix.

[40]There is an important subtlety here: because (Broadcast, No CK) is a fairly unnatural and uncommon policy, most people in that condition would not think that a broadcast is in fact happening, and would not think that others think this. This is what protects people from signaling concerns; these assumptions are elaborated formally in the appendix.

models of imperfect social learning (possibly in networks) or alternative behavioral frictions, and we turn to those next. In each case, we discuss the predictions of alternative models in light of the evidence found in our study.

5.3.1. *Tagged information transmission.* The first network learning model we look at adapts the models of, e.g., Acemoglu et al. (2014), Möbius et al. (2015). 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 models of tagged information aggregation, 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 is well represented by the frictionless benchmark presented in Section 3.3, where $f = 0$. Again, these models predict 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 6 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, pooling across treatments. 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.[41] More strikingly, when we provide information to ten times the number of people, we do not see the expected increase 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.

5.3.2. *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 sources. A way to study these difficulties is to use canonical *sequential social learning* models from the literature on herding or information cascades, which seem 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 tends to 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 strong "less is more" forces such as those our main model produces are unlikely to be explained by standard sequential models. We flesh out the details of the argument in Appendix E.3.

[41]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.

Consider, for simplicity, a binary decision – say, 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 about the right decision. 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 their superior information and persuading others. Remarkably, the same remains true even when agents observe only some of their predecessors, under certain 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. Thus, though we do not know that the friction arising from sequential learning forces will be zero in general, sequential social learning models will have difficulty producing large values of f when, for instance, Broadcasting occurs.

5.3.3. *Curiosity.* A potential alternative explanation is that the treatments themselves may have piqued households' interest to differing degrees. Let's consider a world with no signaling concerns, but with curious agents. Such agents, for instance, might become intrigued to learn the information possessed by others in (Seed, CK), but not under (Broadcast, CK), where they had the information, and knew this.

Our fist observation is that if villagers are curious about demonetization in general, and facts they did not receive, they should also be at least somewhat curious about the meaning of facts they received but did not understand. In (Broadcast, No CK), enough conversations occurred that knowledge was meaningfully improved, which entails that the topics had some grip on people's interest even when they *did* receive information. But in this case, we would expect to see measurably more conversation in (Broadcast, CK) compared to (Seed, No CK): In (Broadcast, CK), people know that others are in a position to clarify any facts they did not understand.[42] But we do not see that difference in the data, as shown in Table 4.

To sustain a curiosity story, one might try to augment it with an additional degree of freedom, adding an assumption of overconfidence or "unaware ignorance." Suppose that individuals in Broadcast treatments incorrectly believed that seeking clarification of things they did not understand would *not* enhance their knowledge, but the information held by others in (Seed, CK) *would* be useful or interesting. This could explain a low amount of conversation in (Broadcast, CK). This ties into a class of explanations based on mistaken beliefs. We discuss this class of explanations next.

5.3.4. *Mistaken perceptions and overconfidence.* To explain the low amount of talking in (Broadcast, CK) – in the data it is comparable to (Seed, No CK) – one could posit that participants mistakenly believed they understood the facts they were told (although in fact they did not understand them). This, however, runs counter to several different kinds of evidence we collected. First, it runs counter to the direct evidence from the knowledge surveys, in which many participants admitted ignorance even to us (Panel B of Table 3); this evidence shows that substantial scope of learning remained. Second, and more fundamentally, such a theory does not predict less seeking in (Broadcast, CK) than in (Broadcast, No CK), which is what we observe. Indeed, insofar as subjects overcome overconfidence and ask others, those others' being informed about the facts should make it more, rather than less, appealing to ask them for clarification. Yet more degrees of freedom would then be needed to explain this comparison.

We close with a few observations about curiosity, understanding, and satiation, as these relate to the context in which our study took place. Media information about the policy, outside of our intervention, was extensive. In our supplemental 2018 survey, we find that even though a great deal of attention was devoted to the policy in mass media, the majority of respondents did not understand the policy implementation even if they learned about it through broadcast media, and felt they needed more clarification; we also find that some people did feel they understood the policy. It thus seems unlikely that (Broadcast, CK), which is just a small-scale version of the public media efforts, somehow satiated villagers with information or created a situation where everybody felt that nobody could clarify things. Instead, our intervention was an increment in a large public information campaign, in which there was a fairly wide distribution of knowledge both before and after we delivered information. If individuals had even

[42] In neither case is there "intrigue," since information is either common knowledge *or* a typical non-seed does not even know that it is present.

moderately well-calibrated beliefs about their environment, they knew that clarification from others could help them – and, as we document, under (Broadcast, No CK), many in fact took advantage of this opportunity.

5.3.5. *Spam.* Finally, we consider the related issue of whether agents interpreted or anticipated the value of the information differently across treatments. This could also generate treatment-dependent frictions. Here, one specific story that could match many of our key predictions is that agents inferred that the information was thought to be of less value when it was distributed to more individuals. Thus, they might not have even looked at the pamphlets in (Broadcast, CK), throwing them away as "spam."

We do not view this as the likely channel, for several reasons. First, agents were aware at baseline that they were broadly uninformed about much of the policy (Table 1). Second, we can show that agents in the (Broadcast,CK) treatment did in fact learn from the pamphlets that they received. In Appendix Table G.2, we consider only the households that were not potential seeds and compare outcomes across (Broadcast, CK) and (Seed, No CK). Given that we already know that there was not a detectable increase in conversations in (Broadcast, CK), we can compare a person who received the pamphlet to a person who did not receive the pamphlet, both in low-communication settings. Column 2 presents the effects of moving from (Seed, No CK) to (Broadcast, CK) on the knowledge index for such an individual. Here, we see a small insignificant increase in total knowledge in (Broadcast, CK) (p-value = 0.34). In column 3, we focus on knowledge of the facts that were listed on the pamphlets.[43] (Broadcast, CK) leads to a 13.4% increase (p-value = 0.0633) in the likelihood of correctly knowing each of the facts that were on the pamphlets.

Thus, there is evidence that villagers read the pamphlets in Broadcast, and conditional on reading them it seems unlikely that mode of delivery would still make a large difference in their perceived value.[44] This point is buttressed by the observation that, in our setting, many trusted, critical messages are broadcast in a common-knowledge way: key examples include disaster alerts and monsoon forecasts. So common knowledge of broadcasting equally could emphasize and amplify the perceived importance of an information delivery. Thus, both on empirical grounds and features of the context, we doubt that a "spam" story is the main driving force.

[43]For the (Seed, No CK) households, we consider the facts that were given to the seeds in the respondent's village.

[44]The analysis here ties back into the "curiosity" explanation and the counterpoints considered in the previous section.

5.4. **Seed effort and public goods.** 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 fairly 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. The basic idea is that the marginal provider of a public good equalizes marginal benefit and marginal costs of provision, and so if different individuals' efforts are substitutable, an essentially constant amount should be provided. We flesh out this point and note some empirical evidence in Appendix E.1.

5.5. **Taking stock.** We have presented a number of alternative frameworks in this section. In each case, we have argued that, while these frameworks can potentially explain some aspects of our observations, they require several degrees of freedom to rationalize the empirical findings. Moreover, some important hypotheses in alternative stories are in tension with the context reported by the population in question.

On the whole, we believe signaling forces provide a fairly simple and unified account of the key reversals relative to a frictionless model, and have fleshed out this argument in the detailed model of Section D.

Nevertheless, given the simplicity of our treatment, there may well be alternative behavioral mechanisms that could rationalize our findings. While we believe that signaling is an important component of what we find, our main finding is a friction in seeking that depends on meta-knowledge about the information policy. A definitive decomposition of the friction into its ultimate constituents is beyond our scope and an important question for further studies.

6. 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. (2018), 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 inconsistent with a more standard model. 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. These findings are at odds with the predictions of a simple frictionless benchmark ($f = 0$).

Our model of signaling concerns provides a mechanism that can explain both why the "more is more" logic holds when it does, and reversals that we observe in the data. The forces in the model are consistent with villagers' reports of their experiences in the context of the Indian demonetization.

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 and was implemented in our experiment as a theoretical benchmark. 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.

An important question for future work is when policymakers should anticipate reversals and when, in contrast, the "more is more" logic prevails. Our view is that when information is simple to interpret – low dimensionality, needing little further clarification – broadcast is likely to work better. However, in cases like ours where the information is complex and where recipients have much to gain from conversations with others, targeted seeding strategies with common knowledge are likely to be more effective.

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. 3.2

Acemoglu, D., K. Bimpikis, and A. Ozdaglar (2014): "Dynamics of information exchange in endogenous social networks," *Theoretical Economics*, 9, 41–97. 5.3.1

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, 5.3

Bailey, N. (1975): *The Mathematical Theory of Infectious Diseases and Its Applications.*, Griffin: London. 1, 5.3

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*. 3.4, 28

Banerjee, A., L. Iyer, and R. Somanathan (2007): "Public action for public goods," *Handbook of Development Economics*, 4, 3117–3154. 5.4, 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, 5.3

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. 3.2

Breza, E., S. Kaur, and N. Krishnaswamy (2019): "Scabs: Wage Norms and the Social Suppression of Labor Supply," *Working Paper*. 3.4

Breza, E., S. Kaur, and Y. Shamdasani (2017): "The morale effects of pay inequality," *The Quarterly Journal of Economics*, 133, 611–663. 3.4

Bursztyn, L., G. Egorov, and R. Jensen (2018): "Cool to be smart or smart to be cool? Understanding peer pressure in education," *Review of Economic Studies*, Forthcoming. 1, 5

Butera, L., R. Metcalfe, W. Morrison, and D. Taubinsky (2019): "The Deadweight Loss of Social Recognition," Tech. rep., National Bureau of Economic Research. 2

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," . 3.2, 41

Chandrasekhar, A. G., B. Golub, and H. Yang (2018): "Signaling, Shame, and Silence in Social Learning," *Working paper, Stanford University.* 1, 3.1, 5, 5.1, 6, 48, 52, 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

Gerber, A. S., D. P. Green, and C. W. Larimer (2008): "Social pressure and voter turnout: Evidence from a large-scale field experiment," *American political Science review*, 102, 33–48. 2

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, 5.3

Jackson, M. O. (2008): *Social and Economic Networks.*, Princeton University Press. 1, 5.3

Karing, A. (2018): "Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone," *Working Paper.* 2

Kaur, S., S. Mullainathan, S. Oh, and F. Schilbach (2019): "Does Financial Strain Lower Worker Productivity?" *Working Paper.* 3.4

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. 5.3.2, 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. 5.3.1

Niehaus, P. (2011): "Filtered Social Learning," *Journal of Political Economy*, 119, 686–720. 4

Perez-Truglia, R. and U. Troiano (2018): "Shaming tax delinquents," *Journal of Public Economics*, 167, 120–137. 2

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

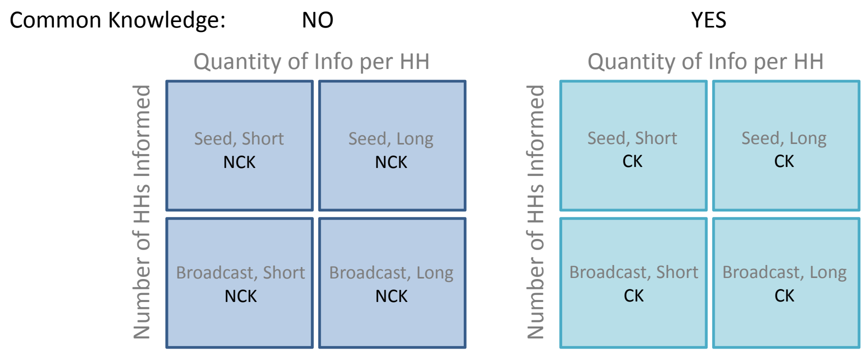


Figure 1. Experimental Design

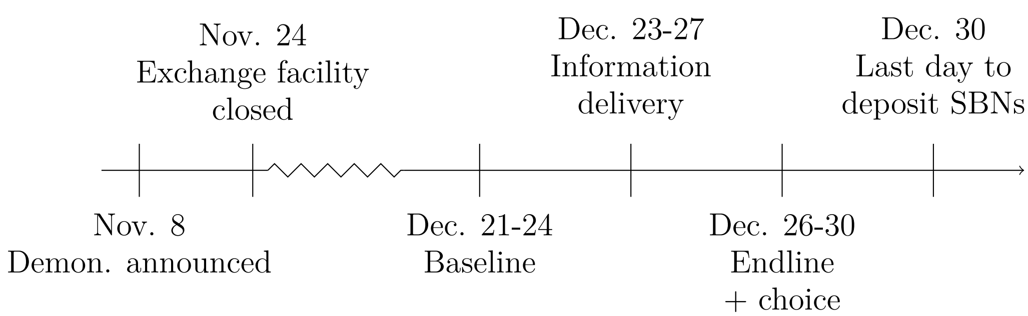


Figure 2. Intervention Timeline

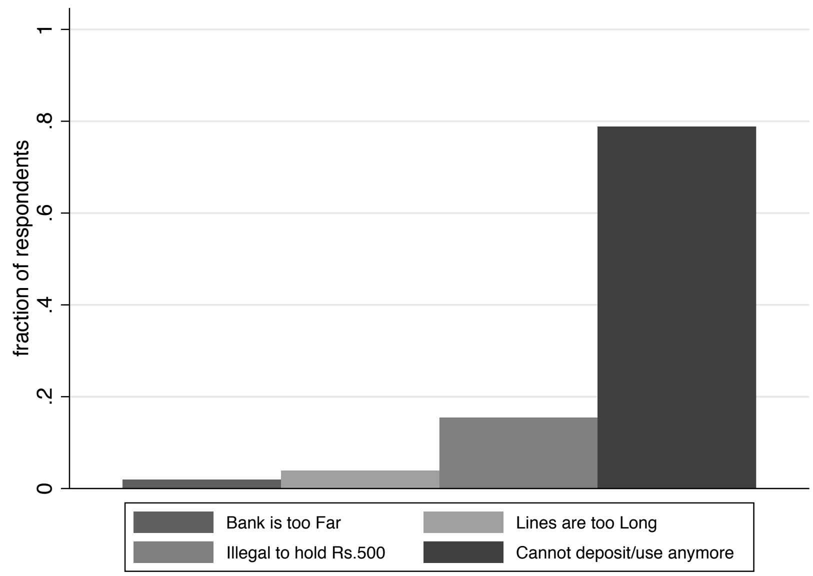
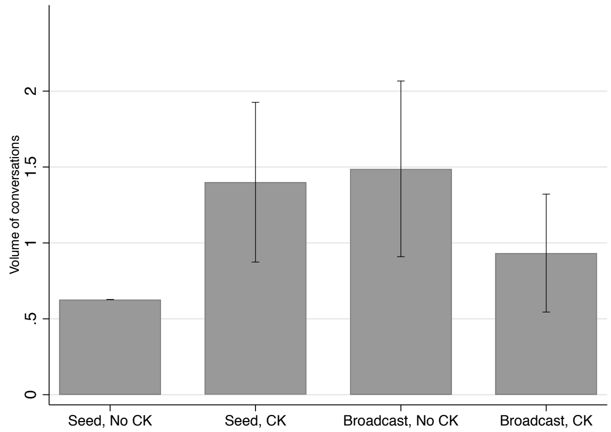
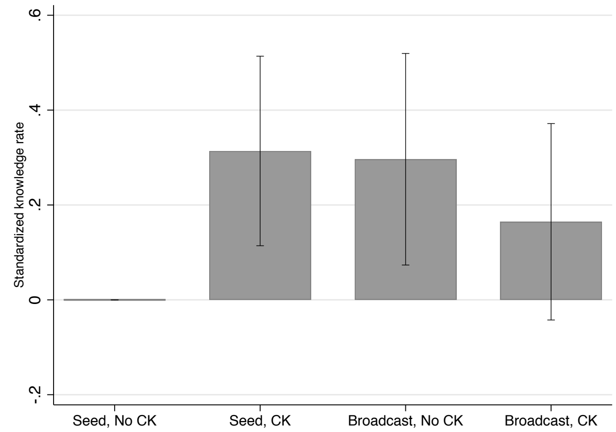


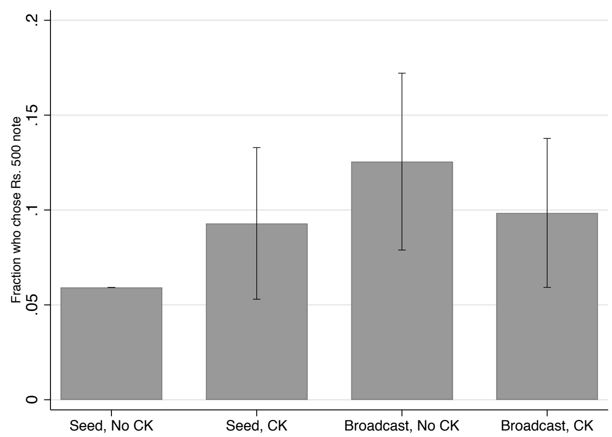
Figure 3. Why did you not choose 500?



(A) Volume of conversations



(B) Knowledge error



(C) Chose old 500

FIGURE 4. Raw Data: Core Experiment Outcomes

(1) Were you confused about any of these changes?
(2) Do you think anyone in your village were confused about these changes?
(3) Do you think you understood how much to exchange, where to exchange, how to exchange, till when you could exchange, application process etc, completely?
(4) Do you think everyone in your village completely understood how much to exchange, till when they could exchange, where to exchange from, how much to exchange etc?
(5) After the policy was introduced, did you ever hesitate to ask someone from your village about the note-ban policy because you were concerned about what they might think about you?
(6) After the policy was introduced, did you ever hesitate to ask an **acquaintance** from your village about the note-ban policy because you were concerned about what they might think about you?

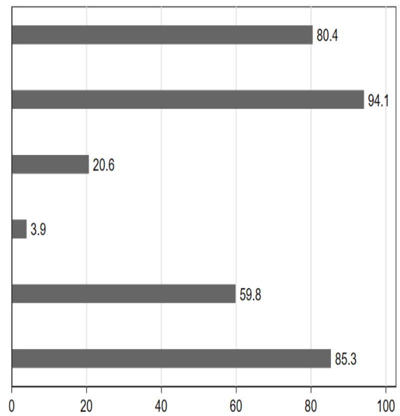


Figure 1: Percent responding yes

(7) If yes, why did you hesitate?

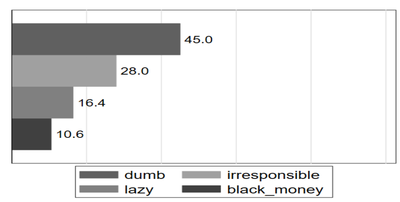
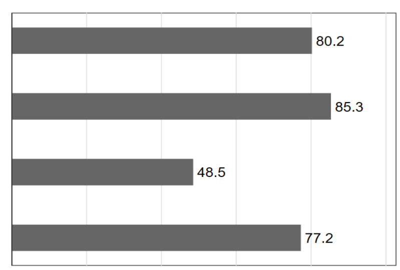


Figure 2: Percent choosing the option

(8) If someone from your village asks about the note-ban policy in December after it was heavily broadcasted on TV, do you think people would think
   (a) he is dumb for not understanding even after being broadcast?
   (b) he is irresponsible for not checking earlier?
   (c) he is dealing in black money?
   (d) he is lazy?



(9) In December, since the news about the note-ban policy was being heavily broadcasted on TV, do you think it was the responsibility of people in your village to know everything/ completely about the note-ban policy?

(10) In December after being heavily broadcasted on TV, do you think some people in your village reduced asking about the note-ban policy, even though they were confused because they were scared/ worried that they would be judged as dumb/ lazy/ irresponsible/dealing in black money?

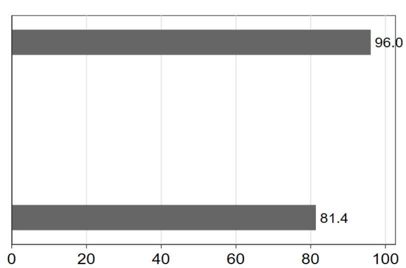


Figure 3: Percent responding yes

Note : The sample consists of 102 randomly sampled respondents across 4 villages in Karnataka.

FIGURE 5. Survey Results

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. On the last day on which SBNs were accepted, we surveyed as many banks as possible near the study villages. Our enumerators made it to 51 banks, where employees were surveyed. 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

Panel A: Error rates

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

Panel B: Incidence of "don't know" responses

	mean	sd	obs
General currency	0.01	(0.11)	966
Exchange locations other than banks	0.30	(0.46)	966
Weekly withdrawal limits from bank accounts	0.33	(0.47)	966
Withdrawal limits on Jan Dhan accounts	0.78	(0.41)	966
Daily withdrawal limits on ATMs	0.32	(0.47)	966

Notes: Panel A gives error rates on knowledge about demonetization in the baseline sample. Panel B gives the incidence of "don't know" responses for the relevant questions. All respondents giving a "don't know" response were asked to make their best guess of the response.

Table 4. 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 5. Knowledge and decision-making

	(1)	(2)
	OLS	OLS
VARIABLES	Knowledge	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	0.566	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 (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. The effects of adding more information

Panel A: Short vs. Long

VARIABLES	(1) OLS Volume	(2) OLS Knowledge	(3) OLS 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	0.583	0.0954

Panel B: Seed vs. Broadcast

VARIABLES	(1) OLS Volume	(2) OLS Knowledge	(3) OLS 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	0.582	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.

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

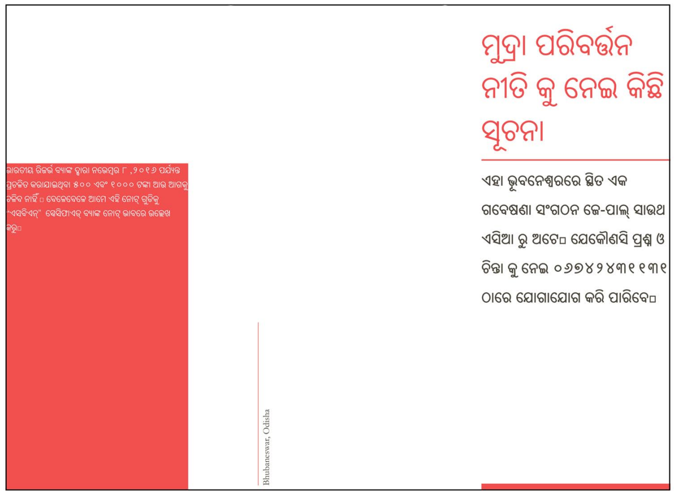
Nov-28	Relaxation in norms of withdrawal from deposit accounts of deposits made in legal tender note wef Nov-29
Nov-29	For account holders of Pradhan Mantri Jan Dhan Yojana: limit of Rs. 10,000 withdrawal per month for full KYC customers; Rs. 5000 with customers with partial KYC
Dec-02	Aadhaar-based 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

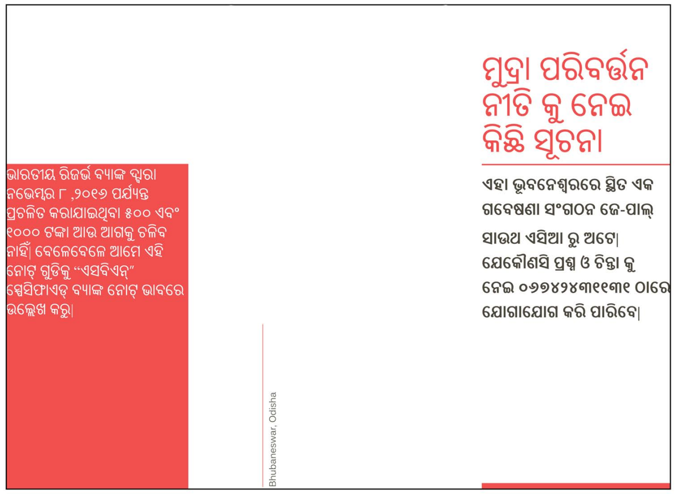
Appendix C. Example Pamphlet Excerpts



(A) Front

(B) Back

Figure C.1. Short pamphlet (2 facts)



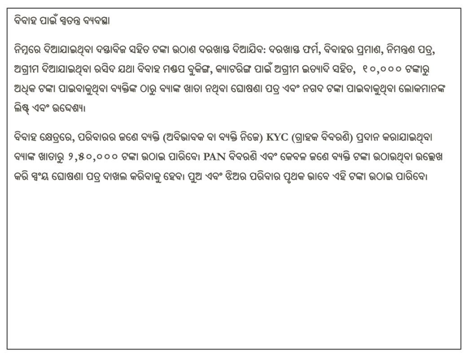
ମୁଦ୍ରା ପରିବର୍ତ୍ତନ ନୀତି କୁ ନେଇ କିଛି ସୂଚନା

ଭାରତୀୟ ରିଜର୍ଭ ବ୍ୟାଙ୍କ ଦ୍ବାରା ନଭେମ୍ବର ୮ ,୨୦୧୬ ପର୍ଯ୍ୟନ୍ତ ପ୍ରଚଳିତ କରାଯାଇଥିବା ୫୦୦ ଏବଂ ୧୦୦୦ ଟଙ୍କା ଆଉ ଆଗକୁ ଚଳିବ ନାହିଁ| ବେଳେବେଳେ ଆମେ ଏହି ନୋଟ୍ ଗୁଡିକୁ "ଏସବିଏନ୍" ସ୍ପେସିଫାଏଡ୍ ବ୍ୟାଙ୍କ ନୋଟ୍ ଭାବରେ ଉଲ୍ଲେଖ କରୁ|

ଏହା ଭୁବନେଶ୍ୱରରେ ସ୍ଥିତ ଏକ ଗବେଷଣା ସଂଗଠନ ଜେ-ପାଲ୍ ସାଉଥ ଏସିଆ ରୁ ଅଟେ| ଯେକୌଣସି ପ୍ରଶ୍ନ ଓ ଚିନ୍ତା କୁ ନେଇ ୦୭୭୪୨୪୩୧୧୩୧ ଠାରେ ଯୋଗାଯୋଗ କରି ପାରିବେ|

Bhubaneswar, Odisha

(A) Front

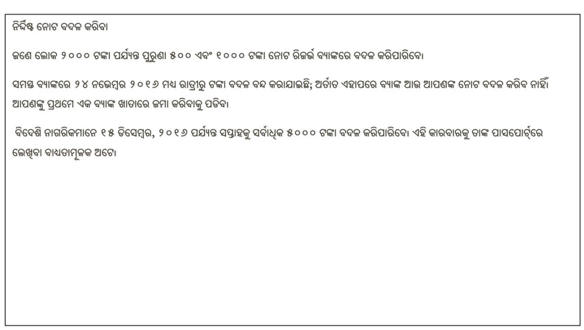


ବିବାହ ପାଇଁ ସ୍ୱତନ୍ତ୍ର ବ୍ୟବସ୍ଥା

ନିମ୍ନରେ ଦିଆଯାଇଥିବା ଦସ୍ତାବିଜ ସହିତ ଟଙ୍କା ଉଠାଣ ଦରଖାସ୍ତ ଦିଆଯିବ: ଦରଖାସ୍ତ ଫର୍ମ, ବିବାହର ପ୍ରମାଣ, ନିମନ୍ତ୍ରଣ ପତ୍ର, ଅଗ୍ରୀମ ଦିଆଯାଇଥିବା ରସିଦ ଯଥା ବିବାହ ମଣ୍ଡପ ବୁକିଙ୍ଗ, କ୍ୟାଟରିଙ୍ଗ ପାଇଁ ଅଗ୍ରୀମ ଇତ୍ୟାଦି ସହିତ, ୧୦,୦୦୦ ଟଙ୍କାରୁ ଅଧିକ ଟଙ୍କା ପାଇବାକୁଥିବା ବ୍ୟକ୍ତିଙ୍କ ଠାରୁ ବ୍ୟାଙ୍କ ଖାତା ନଥିବା ଘୋଷଣା ପତ୍ର ଏବଂ ନଗଦ ଟଙ୍କା ପାଇବାକୁଥିବା ଲୋକମାନଙ୍କ ଲିଷ୍ଟ୍ ଏବଂ ଉଦ୍ଦେଶ୍ୟ।

ବିବାହ କ୍ଷେତ୍ରରେ, ପରିବାରର ଜଣେ ବ୍ୟକ୍ତି (ଅବିଭାବକ ବା ବ୍ୟକ୍ତି ନିଜେ) KYC (ଗ୍ରାହକ ବିବରଣି) ପ୍ରଦାନ କରାଯାଇଥିବା ବ୍ୟାଙ୍କ ଖାତାରୁ ୨,୫୦,୦୦୦ ଟଙ୍କା ଉଠାଇ ପାରିବେ। PAN ବିବରଣି ଏବଂ କେବଳ ଜଣେ ବ୍ୟକ୍ତି ଟଙ୍କା ଉଠାଉଥିବା ଉଲ୍ଲେଖ କରି ସ୍ଵଂୟ ଘୋଷଣା ପତ୍ର ଦାଖଲ କରିବାକୁ ହେବ। ପୁଅ ଏବଂ ଝିଅର ପରିବାର ପୃଥକ ଭାବେ ଏହି ଟଙ୍କା ଉଠାଇ ପାରିବେ।

(B) Page 1/8



ନିର୍ଦ୍ଦିଷ୍ଟ ନୋଟ ବଦଳ କରିବା

ଜଣେ ଲୋକ ୨୦୦୦ ଟଙ୍କା ପର୍ଯ୍ୟନ୍ତ ପୁରୁଣା ୫୦୦ ଏବଂ ୧୦୦୦ ଟଙ୍କା ନୋଟ ରିଜର୍ଭ ବ୍ୟାଙ୍କରେ ବଦଳ କରିପାରିବେ।

ସମସ୍ତ ବ୍ୟାଙ୍କରେ ୨୪ ନଭେମ୍ବର ୨୦୧୬ ମଧ୍ୟ ରାତ୍ରୀରୁ ଟଙ୍କା ବଦଳ ବନ୍ଦ କରାଯାଇଛି; ଅର୍ତାତ ଏହାପରେ ବ୍ୟାଙ୍କ ଆଉ ଆପଣଙ୍କ ନୋଟ ବଦଳ କରିବ ନାହିଁ। ଆପଣଙ୍କୁ ପ୍ରଥମେ ଏକ ବ୍ୟାଙ୍କ ଖାତାରେ ଜମା କରିବାକୁ ପଡିବ।

ବିଦେଶି ନାଗରିକମାନେ ୧୫ ଡିସେମ୍ବର, ୨୦୧୬ ପର୍ଯ୍ୟନ୍ତ ସପ୍ତାହକୁ ସର୍ବାଧିକ ୫୦୦୦ ଟଙ୍କା ବଦଳ କରିପାରିବେ। ଏହି କାରବାରକୁ ତାଙ୍କ ପାସପୋର୍ଟ୍‌ରେ ଲେଖିବା ବାଧ୍ୟତାମୂଳକ ଅଟେ।

(C) Page 2/8

FIGURE C.2. Long pamphlet (24 facts)

Appendix D. Detailed signaling model

D.1. **The model.** 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.

D.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.[45]

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

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

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 5.3, we consider some models with richer social learning.

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.[46] We will assume these are generic.[47] Let $I_{\mathrm{D}} \in \{0,1\}$ 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:

$$\Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})} := V^{(I_{\mathrm{D}},I_{\mathrm{T}})}(\mathrm{S}) - V^{(I_{\mathrm{D}})}(\mathrm{NS}) \sim F_a^{(I_{\mathrm{D}},I_{\mathrm{T}})}. \tag{D.1}$$

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.

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

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

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.[48]

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

$$u^{(I_{\mathrm{D}},I_{\mathrm{T}})}(d) = V^{(I_{\mathrm{D}},I_{\mathrm{T}})}(d) + \lambda \mathbb{P}(a = \mathrm{H} \mid d, \mathfrak{p}, I_{\mathrm{O}}). \tag{D.2}$$

It will be useful to write the difference

$$u^{(I_{\mathrm{D}},I_{\mathrm{T}})}(\mathrm{S}) - u^{(I_{\mathrm{D}},I_{\mathrm{T}})}(\mathrm{NS}) = \Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})} - \lambda \Pi \tag{D.3}$$

where $\Delta^{(I_{\mathrm{D}},I_{\mathrm{T}})}$ is defined in (D.1) and

$$\Pi = \mathbb{P}(a = \mathrm{H} \mid d = \mathrm{NS}, \mathfrak{p}, I_{\mathrm{O}}) - \mathbb{P}(a = \mathrm{H} \mid d = \mathrm{S}, \mathfrak{p}, I_{\mathrm{O}}). \tag{D.4}$$

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.

D.1.2. *Equilibrium: Definition and basic observations.* We study a Bayesian equilibrium of this game. A strategy of D determines beliefs of the Observer – i.e. $\mathbb{P}(a = \mathrm{H} \mid$

[48]Foundations for this assumption are discussed in Chandrasekhar et al. (2018).

$d, \mathfrak{p}, I_{\mathrm{O}})$ – for both values $d = \mathrm{S}, \mathrm{NS}$.[49] 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.[50]

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 (D.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: 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.[51]

D.1.3. *Assumptions.*

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

[49] 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.

[50] 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, (D.3), does not depend on his private ability type. Because his ability type is unobservable, the reputational payoff cannot depend on it, either.

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

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 $v \geq 0$.

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.

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.

D.2. **Analysis and results.**

D.2.1. *Dependence of seeking rates on treatment.* In general the model may have multiple equilibria.[52] However, under our assumptions (the key one being P3) the game has some nice structure. In particular, as the cutoffs[53] $\underline{v}(\mathfrak{p}, I_{\mathrm{D}})$ 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*.[54]

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 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 Assumptions P1–P3:

(a) s(Broadcast, No CK) $>$ s(Broadcast, CK);
(b) s(Seed, CK) $>$ s(Broadcast, CK).

The proof of this and all other propositions appears in Section D.3 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 Assumptions P1–P3 and B1–B2, and assuming $\overline{k}/n$ is small enough, it holds that s(Seed, CK) $>$ s(Seed, No CK).

Finally, the prediction that requires the most structure is:

Proposition **3.** Take Assumptions P1–P3 and B1–B2, and, fixing all other parameters, suppose the following three quantities are small enough: (i) $\overline{k}/n$; (ii) β_{Seed}; and (iii) $\frac{1-\chi_{\mathrm{Broadcast}}}{(\overline{k}/n)^2}$. Then s(Broadcast, No CK) $>$ s(Seed, No CK).

[52]For more on this multiplicity, see Chandrasekhar et al. (2018).
[53]Introduced in Section D.1.2 above.
[54]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.

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 $I_{\mathrm{D}} = 1$, D knows that $I_{\mathrm{T}} = 1$. 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 $d = \mathrm{S}$ (given either value of $\mathfrak{p}$) as a convex combination over values of I_{D}. The term corresponding to $I_{\mathrm{D}} = 1$ 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 $I_{\mathrm{D}} = 0$ involves a weakly greater posterior that $a = \mathrm{H}$ 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.[55]

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

LEMMA **1.** Under the assumptions of Section D.1.3, suppose that $(1-\chi_{\text{Broadcast}})$ is small enough relative to $(\overline{k}/n)^2$. 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 negligibly small.

Now we can establish the proposition. Concerning the direct benefit: in (Seed, No CK), when D receives no information ($I_{\text{D}} = 0$), the fact that β_{Seed} is small means that his expectations approximate those when $I_{\text{T}} = 0$. In contrast, in (Broadcast, No CK), given that $I_{\text{D}} = 1$, 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 ($I_{\text{T}} = 1$). 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 $\mathfrak{b} \neq \text{Broadcast}$. 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.

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.

D.2.2. *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(Broadcast, No CK) $>$ p(Broadcast, CK)
(b) p(Seed, CK) $>$ p(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(Seed, CK) $>$ p(Broadcast, CK)
- p(Broadcast, No CK) $>$ p(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.

D.3. **Proofs.**

D.3.1. *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 (D.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 (D.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.[56]

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

[56] 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.

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.

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.3.2. *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

(D.5)
$$\mathbf{E}^\mathrm{D}\left[\Pi(\boldsymbol{c}) \mid \mathcal{I}_\mathrm{D}\right] = \xi \mathbb{P}_{\boldsymbol{c}}(a = \mathrm{H} \mid d = 1, \mathfrak{p}, I_\mathrm{D} = 1) + (1-\xi)\mathbb{P}_{\boldsymbol{c}}(a = \mathrm{H} \mid d = 1, \mathfrak{p}, I_\mathrm{D} = 0).$$

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.3.3. *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.5) 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.3.4. *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 $\mathfrak{b} \neq$ 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.5) 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 (2018) 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 4 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 5, 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.[57] 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 $\widetilde{I}_{i,t}$.[58]

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.

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

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

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

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 had

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.

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

	(1)	(2)
	OLS	OLS
VARIABLES	Knowledge	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	0.564	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 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 in the endline survey.

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) is distinguishable 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 increases knowledge by 15.5% (column 1, $p = 0.014$). Under no Common Knowledge, Broadcast increases knowledge by 13.6% (column (1), $p = 0.0345$) relative to Seed. Similarly, in column 2 we see that holding useful facts fixed, (Seed, CK) increases knowledge by 6.8% ($p = 0.008$) and (Broadcast, No CK) increases knowledge by 6.1% ($p = 0.0345$), 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.3 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 and the second is Rs. 500 relative to dal. We see that going to (Seed, CK) from (Seed, No CK) leads

Table G.1. Heterogeneity in knowledge

VARIABLES	(1) OLS Knowledge (Told)	(2) OLS Knowledge (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	0.569	0.457
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 knowledge if the fact being asked about was told during information delivery. Column (2) displays effects on knowledge 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.

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),

Table G.2. Did the Broadcast, Common Knowledge Group Learn Anything?

VARIABLES	(1) OLS Volume	(2) OLS Knowledge Index	(3) OLS Knowledge Panel (Told)	(4) OLS Chose 500
Broadcast x Common Knowledge	-0.119	0.0129	0.0654	-0.00805
	(0.242)	(0.0134)	(0.0350)	(0.0227)
	[0.623]	[0.337]	[0.0633]	[0.723]
Observations	1,078	1,082	36,788	1,067
Mean: Seed x No CK, Non-seed HH	0.868	0.580	0.489	0.0677

Notes: Regressions compare outcomes for the Broadcast, Common Knowledge treatment relative to the Seed, No Common Knowledge treatment. The regression coefficient only includes households that were not potential seeds. 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. Columns (1), (2), and (4) use the same specifications as Table 5. Column (3) considers a respondent x question panel and focuses only on knowledge of the facts that were told in the respondent's village. Standard errors (clustered at the village-level) are reported in parentheses and p-values are reported in brackets.

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.

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

Table G.3. 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.

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 $\times$ CK	-1.364	-1.058	-0.306
	(0.507)	(0.429)	(0.175)
	[0.00778]	[0.0146]	[0.0821]
Seed HH $\times$ CK	1.305	1.251	0.0540
	(1.499)	(1.156)	(0.619)
	[0.385]	[0.280]	[0.931]
Seed HH $\times$ BC	-0.505	-0.694	0.189
	(1.161)	(0.616)	(0.816)
	[0.664]	[0.261]	[0.817]
Seed HH $\times$ BC $\times$ 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 $\times$ CK = 0 p-val	0.0168	0.0168	0.397
BC + BC $\times$ 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	(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	0.564	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 (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. 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 × 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 × CK = 0 p-val	0.0224	0.0332	0.248
BC + BC × 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	(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	0.562	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 (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.

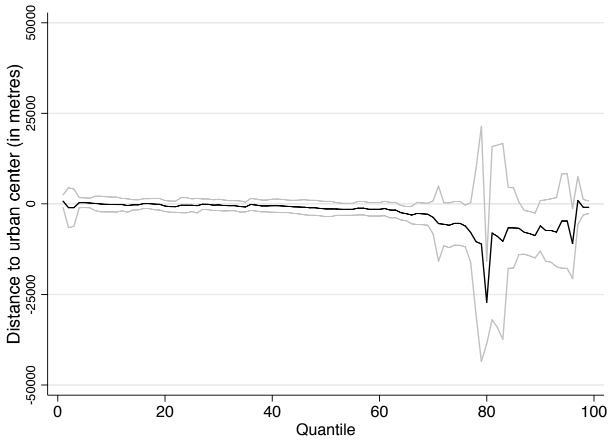
Appendix L. Status Quo Appendix

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.

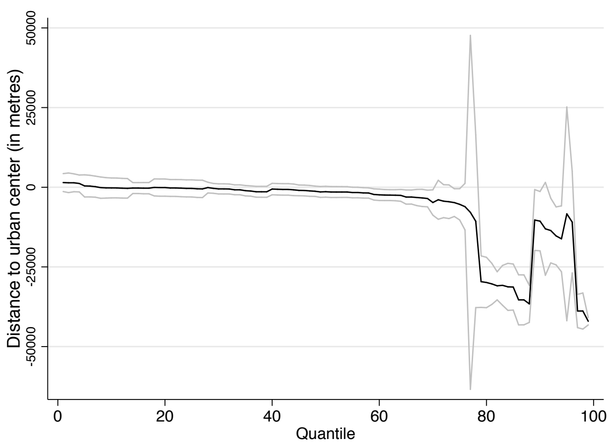
We can include "status quo" in a regression analysis to compare it to our other treatments, but we need to keep in mind that this is observational, and relies on controlling for the distribution of distance from cities, survey timing, etc. That means when we compare to "status quo" we should interpret 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.

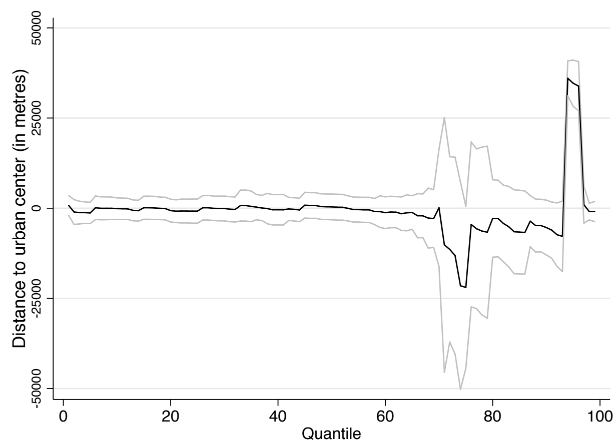
We begin by looking at the distance distributions for the "status quo" and intervention villages. Figure L.1, Panels A, B, and C present coefficients from a quantile regression of distance from urban center against "status quo", conditional on caste of the hamlet. Panel A conditions on caste, and Panels 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	(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	0.588	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

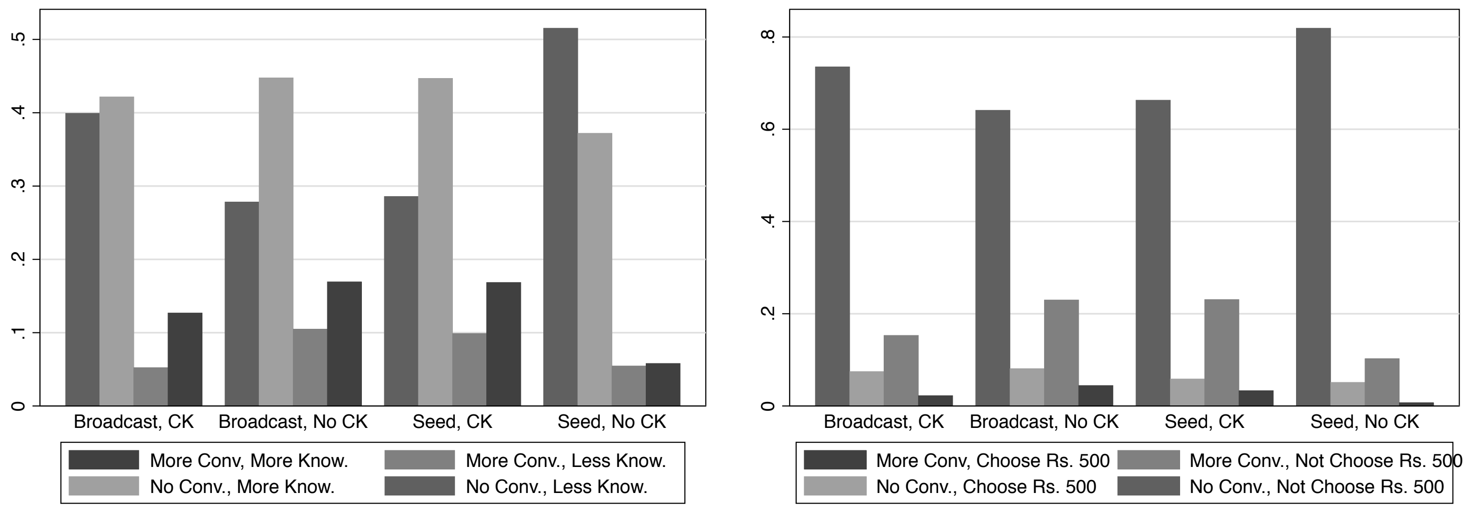
Here we look at how the joint distribution of conversations and information quality move. Table N.1 presents multinomial logistic regressions. In column 1, the outcome variable takes on values of "Conversations and High Knowledge", "Conversations and Low Knowledge," "No Conversations and High Knowledge," 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.



(A) Knowledge

(B) Choice of Rs. 500

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