SOCIAL NETWORKS, REPUTATION, AND COMMITMENT: EVIDENCE FROM A SAVINGS MONITORS EXPERIMENT

EMILY BREZA† AND ARUN G. CHANDRASEKHAR‡

Abstract. We study whether individuals save more when information about their savings is shared with another village member (a "monitor"). We focus on whether the monitor's effectiveness depends on her network position. Central monitors may be better able to disseminate information, and more proximate monitors may pass information to individuals who interact with the saver frequently. In 30 villages, we randomly assign monitors. Average monitors increase savings by 35%. A one-standard deviation more central monitor increases savings by 14%; increasing proximity from social distance three to two increases savings by 16%. The increased savings persist over a year after the intervention's end, and monitored savers better respond to shocks. Information flows. 63% of monitors tell others about the saver's progress. 15 months later, others know more about the saver's progress and believe she is responsible if the saver was assigned a more central monitor. To benchmark the results, in 30 other villages, savers choose their monitors. Monitored savers save similar amounts and non-monitored savers increase their savings relative to their random-assignment village counterparts.

JEL Classification Codes: D14, D83, L14, O16, Z13

Keywords: Commitment, Reputation, Savings, Social Networks

1. Introduction

Peer effects have been found in a range of settings from schooling to exercise to savings. The literature has traditionally focused on cleanly identifying the reduced form effect, asking how an individual's savings or academic performance depends on the savings or academic performance of her peers. Individuals may be affected by the actions of their peers through a variety of channels. Using the example of group savings, the literature has shown that peer effects operate through channels such as (a) learning how to use financial products; (b)

Date: August 2017.

We thank Monisha Mason, Lisa Nestor, Shobha Dundi, and Gowri Nagraj for outstanding assistance in the field. We thank Robert Townsend, Esther Duflo, Abhijit Banerjee, and Matt Jackson for their tremendous guidance. We also thank Ran Abramitzky, Sumit Agarwal, Attila Ambrus, John Beshears, Gabriel Carroll, Pascaline Dupas, Itay Fainmesser, Erica Field, Andy Foster, Ben Golub, Rema Hanna, Andrew Hertzberg, Stephan Meier, Markus Mobius, Kaivan Munshi, Rohini Pande, Luigi Pistaferri, Dina Pomeranz, Jonathan Robinson, Duncan Thomas, Chris Udry, and Xiao Yu Wang for helpful discussions, as well as numerous seminar and conference participants. We are grateful to CFSP, NSF SES-1156182, the Chazen Institute at Columbia Business School, CIBER and the Earth Clinic at Columbia University, the Russell Sage Behavioral Economics Small Grant, and SEED for generous support. Finally, we both thank the NSF GRFP. A previous version of this paper was titled "Savings Monitors."

†Harvard University and NBER. Email: ebreza@fas.harvard.edu.

‡Stanford University and NBER. Email: arungc@stanford.edu.

reminders; (c) posting a bond; or (d) reference-dependent preferences ("keeping up with the Joneses") (see Jack and Suri (2014), Cai et al. (2013), Bryan et al. (2012), Kast et al. (2012), Beaman et al. (2014), Beshears et al. (2015), Munshi (2014), Karlan et al. (2010), Bursztyn et al. (2013), and Banerjee et al. (2013)). However, less has been written on whether peer effects may arise from individuals wanting to impress others through their actions.

This paper focuses on this last channel that is likely present in many applications - that when actions are observable to others, they may come with reputational benefits. Further, those benefits may depend on the network position of the observer given that building reputation may be more valuable with some members of the community than others.[1]

The potential for reputation-based peer effects is particularly widespread in development economics. For example, in theorizing about repayment incentives in joint liability micro-credit, Besley and Coate (1995) describe a social punishment that depends on reporting poor behavior and admonishment by others, writing:

> "the contributing member may admonish his partner for causing him or her discomfort and material loss. He might also report this behavior to others in the village, thus augmenting the admonishment felt. Such behavior is typical of the close-knit communities in some LDCs."

Peer-driven financial institutions, such as rotating savings and credit associations (RoSCAs), self-help groups (SHGs), and village savings and loan associations (VSLAs), are ubiquitous in the developing world and are thought to, in part, work on this principle. However, it is hard to get traction on how these institutions work, let alone isolate the reputational channel: they are complicated objects of anywhere from five to 30 individuals, with endogenous group formation and forces beyond simple reputation effects contributing to good behavior.[2]

We design and implement a savings field experiment to focus on the reputational force highlighted by Besley and Coate (1995) and begin to unpack the black box of peer effects in informal financial institutions. To make the problem tractable, we construct a simplified "institution" of one saver who desires to save matched to one observer and induce random group formation. Importantly, our setting is naturalistic, mimicking the business correspondent (BC) model, which is commonly used by banks in India to service rural customers.

Specifically, we conduct an experiment across 60 villages in rural Karnataka, India, where we have complete network data for almost all households in every village. We assist 1,300

[1]Our experiment was inspired by the earlier lab-in-the-field study Breza et al. (2015), where we explore the efficiency of transfers in non-anonymous sender-receiver investment games with a third-party observer. Villagers are assigned to one of three treatments: (1) sender-receiver game, (2) sender-receiver game with a third party who observes the interaction, but takes no action of her own, or (3) sender-receiver game with a third party who observes the interaction and can levy a fine against the receiver. The interaction is fully non-anonymous. We are interested in how the network position of the third party influences the efficiency of the transaction. When a more central third-party observes the transaction, efficiency increases significantly (as seen from comparing (2) to (1) for more versus less central third parties). Further, the beneficial effect of centrality is greater when the third party is also given an observable punishment technology.

[2]Reputation effects may help to explain, for example, why researchers have documented peer effects in microfinance groups even in the absence of contractual joint liability (Breza 2014, Feigenberg et al. 2013).

individuals to review their finances, set a six-month savings goal, and open a formal account at a bank or post office. A random group of savers is selected from each village, and each saver receives a (different) partner for the duration of the experiment, whom we call a monitor. In 30 randomly-selected villages, we randomly assign individuals from a pre-specified pool to serve as monitors, and in the remaining 30 villages, using random serial dictatorship, we allow savers to select their respective monitors from the pre-specified monitor pool. In all cases, the monitor receives bi-weekly information about the saver's target account savings. As monitors are drawn from a random pool of villagers, they vary in their position in the village network: some are more central (i.e., more connected directly or indirectly) than others, and some have closer relationships (i.e., proximity through the network) with the saver. Using the 30 villages in which we randomly assign saver-monitor pairs, we study how the network position of randomly-assigned monitors influences savings behavior. Further, we use the 30 villages in which savers choose their monitors to benchmark how much agents save under endogenous group formation.

Why might the monitor's position in the network be important? Each monitor learns about the saver's progress. The monitor may, in turn, pass that information or any opinion she has made on to others. Thus, the monitor's position within the village network may determine how far and to whom her opinion may spread. For example, more central agents – i.e., better connected (directly or indirectly) to a larger set of people – are well-suited to broadcast information. In turn, they may make more effective monitors, ceteris paribus, as the saver has more to gain by impressing them. Similarly, a socially proximate monitor may be more likely to speak to others with whom the saver is likely to interact in the future. Therefore, by telling individuals who are more relevant to the saver's future interactions, proximate monitors may also be more valuable.

To help clarify these issues and identify those aspects of the otherwise-complex network on which to focus our empirical analysis, we develop a simple signaling model. In this model, we assume that savers gain utility from interacting in the future with individuals who have heard about their successes.[3] Here, the network plays two roles; information is disseminated from the monitor through the network, and future interactions between the saver and other villagers (including the monitor) occur through the network. We show that a saver is incentivized to save more when randomly assigned to a more central monitor or to one that is more proximate to her.

Equipped with this framework, we pair our experiment with extremely detailed network data collected in part by the authors in previous work (Banerjee et al. (2014)). This household-level network data comprises 12 dimensions of interactions across *all* potential

[3]There are many microfoundations for such an effect. Successful savers may gain an improved reputation for being responsible, for example. Alternately, agents may feel embarrassment or shame when interacting people who have learned of their shortcomings.

pairs of households in each of the 60 study villages.[4] As described above, two moments of the network data emerge from our model and we focus on both: monitor (eigenvector) centrality, which captures how much information emanating from a monitor should spread in the network, and the social proximity between the saver-monitor pair, which is the inverse of the shortest path length through the network.[5] The framework also generates a model-specific network statistic that drives savings incentives, which we can also take to the data. This model-specific statistic is increasing in both monitor centrality and saver-monitor proximity. Of course we do not claim that our experiment shows that this mechanism is the sole force driving why monitor centrality and proximity should affect savings, but we do provide evidence consistent with such a story playing a role.

Savings is an ideal application for our experiment for several reasons. First, we require a setting where reputation is important. Anecdotal and survey evidence from the study villages suggests that a large fraction of villagers indeed want to save more, and that showing one can save more is a sign of responsibility. Second, we want to be able to accurately measure the outcome variable. Certainly savings in a bank account is easy to observe and we can verify the data through passbooks. Third, we desire a context that is naturalistic. Savings is an obvious application in which to study public commitments as many of the informal financial products commonly observed in developing countries (and in the study villages) and discussed above incorporate groups of individuals from the same social network and rely on mechanisms that are likely to include mutual monitoring/observation (Besley and Coate (1995), Beaman et al. (2014), Besley et al. (1993), Karlan (2007), Giné and Karlan (2006), Bryan et al. (2012), and Breza (2014)).[6,7]

Finally, chronic under-saving is an important issue in developing and developed countries alike. The desire to save is widespread, but many are unable due to lack of access, lack

[4]The network data we use here is essentially complete. With data surveyed from 90% of households in each village, the probability of not surveying either member of a pair is 0.1^2. So there is data on $1 - 0.1^2 \approx 0.99$ share of possible links in the network. We use what is called the OR network, drawing a link between two households if either named the other.

[5]See Katz and Lazarsfeld (1970), DeMarzo et al. (2003), Ballester et al. (2006), Banerjee et al. (2013), and Golub and Jackson (2010).

[6]Our paper is closely related to Kast et al. (2012), who conduct an experiment layering a peer savings scheme on top of an existing microfinance borrowing group. Group members were motivated to save by making public commitments in front of one another along with public contributions. The authors find large positive effects on savings. However, in a second experiment, SMS-based reminders to save led to similar savings impacts. An interesting distinction with our setting is that in their's, the monitors were both co-borrowers with the savers and were savers themselves by construction.

[7]The paper is also related to a larger dialogue about the role of social capital and observability in informal financial institutions (Platteau 2000). Jakiela and Ozier (2016) provide lab-in-the-field evidence that peer observability can have a dark side, causing some women to inefficiently distort their behavior to avoid paying a so called "kin tax." Brune et al. (2016) and Dupas and Robinson (2013a) argue that social pressure may be one mechanism underlying the results of their respective savings field experiments as well. One natural feature of opting into any group-driven financial product, including our experimental sample, is that by definition some component of savings is observable. Therefore, it is likely that individuals who worry the most about unwanted demands from others would be the least likely to participate.

of commitment, or lack of attention (Ashraf et al. (2006), Brune et al. (2013), Karlan et al. (2010), Thaler and Benartzi (2004), and Beshears et al. (2011), for example). Our intervention can be interpreted as a special kind of commitment savings device where the characteristics of the monitor determine how well it performs. Research has also shown that increased savings has numerous benefits including increased investment, working capital, income and even labor supply and can improve the ability for households to overcome shocks (Dupas and Robinson (2013b), Dupas and Robinson (2013a), Prina (2013), Schaner (2014), and Kast and Pomeranz (2014)). We can explore these issues in the short run (6 months) and medium run (21 months) through the lens of our study.

Our empirical analysis has four main components. First, using the data from the 30 villages in which we randomly assign monitors to savers, we establish that receiving an arbitrary monitor increases total savings across formal and informal vehicles by 35%. As predicted by our model, the largest increases are generated by more central monitors as well as more proximate monitors. Increases of one standard deviation in monitor centrality and proximity, respectively, correspond to increases in savings of 14% and 16%. Similarly, a one standard deviation increase in the model-specific network statistic corresponds to an increase in savings of 33.5%.

Second, we make use of novel supplemental data to support the reputational story. We show that monitors indeed speak to others about the saver, and 40% of savers even hear gossip about themselves through back-channels. Moreover, 15 months after the conclusion of the intervention, the opinions of randomly-selected households about a saver's performance and ability to follow through on self-set goals are related to the centrality of that saver's randomly-assigned monitor. To our knowledge, this is the first time such an exercise has been conducted in the literature.

Third, we provide evidence that our intervention caused lasting and positive average impacts on participant households. We show that the increases in savings caused by our intervention come from increases in labor supply and decreases in unnecessary expenditures. Fifteen months after the end of the intervention, we show that subjects randomly assigned to monitors report declines in the incidence of unmitigated shocks, which we measure following Dupas and Robinson (2013a).[8] Moreover, the increases in savings persist 15 months after the intervention. Taken together, these results suggest that monitors, especially central and proximate ones, help savers to direct financial slack towards savings, rather than wasteful expenditures or leisure, result in improved risk-coping, and yield persistent increases in savings, likely held as buffer stocks.[9]

[8]We denote a shock to be an event such as a personal health shock, bovine health shock, or other unexpected household expenditure where the household did not have enough cash on hand to cover the cost. We show that the incidence of reporting an above-median number of shocks drops for individuals assigned to a monitor.

[9]The most common savings goal purposes listed at baseline were unforeseen expenditures and emergencies.

Fourth, in the 30 villages in which savers could choose their monitors, we find that monitored savers perform approximately as well as their random-assignment village counterparts. Further, we find that non-monitored savers in the endogenous assignment villages save substantially more than the non-monitored savers in the random assignment villages. Non-monitored savers in endogenous-selection villages completely "catch up" to the saving levels of the monitored community members. This suggests that monitoring can affect the saver's propensity to save, but may also spill over to friends of the saver. Indeed, we find suggestive evidence that more unplanned conversations about savings take place in endogenous selection villages, many of which were likely overheard by non-monitored savers.

In short, our study points to the idea that reputations matter, and they matter heterogeneously within the broader village network. The experiment provides a context in which agents could respond to our monitor treatment using an important economic vehicle – formal savings – that itself stood to generate real benefits to our subjects. That the increased savings allowed our subjects to better respond to health and household shocks indicates that the monitor treatment effect was strong enough not only to change savings behavior directly but to also yield measurable and meaningful economic consequences over the next year. Again we caution that reputation is of course not the only force that could matter for our findings, but our analysis points to it being a relevant channel.

The remainder of the paper is organized as follows. Section 2 contains a description of the experimental design, setting and data. In section 3 we provide a parsimonious model that shows why it is natural to focus on centrality and proximity. Section 4 presents the results for the villages where monitors were randomly assigned to savers, and Section 5 presents a discussion of threats to validity. We discuss savings balances in the villages with endogenous monitor assignment in Section 6, and Section 7 is a conclusion.

2. Data and Experimental Design

The requirements for our study are threefold: 1) detailed social network data; 2) the presence of financial institutions where study participants can open accounts and save; 3) experimental variation in the nature of the saver-monitor relationship. Our final sample includes approximately 3,000 participants from 60 villages in rural Karnataka, India.

2.1. **Network and Demographic Data.**

We chose to set our experiment in villages that coincide with the social network and demographic data set previously collected, in part by the authors (and also described in Banerjee et al. (2014)). In our field experiment, we match participants to this unique data set.

Banerjee et al. (2014) collected network data from 89.14% of the 16,476 households living in those villages. The data concerns "12 types of interactions for a given survey respondent: (1) whose houses he or she visits, (2) who visit his or her house, (3) his or her relatives in the village, (4) non-relatives who socialize with him or her, (5) who gives him or her medical advice, (6) from whom he or she borrows money, (7) to whom he or she lends money, (8) from whom he or she borrows material goods (e.g., kerosene, rice), (9) to whom he or she lends material goods, (10) from whom he or she gets advice, (11) to whom he or she gives advice, (12) with whom he or she goes to pray (e.g., at a temple, church or mosque)." This provides a rich description of the interactions across households. Using this data we construct one network for each village at the household level and indicate that a link exists between households if any member of a household is linked to any other member of another household in at least one of the 12 ways.[10] Network-level summary statistics are displayed in Appendix Table D.1.

As such, we have extremely detailed data on social linkages, not only between our experimental participants but also about the embedding of the individuals in the social fabric at large. Every village is associated with a social network. Following the extensive work on this data, we assume that this is an undirected, unweighted network (see, e.g., Banerjee et al. (2013), Jackson et al. (2010), Chandrasekhar et al. (2013) for discussion).

Moreover, the survey data includes information about caste, elite status, house amenities and the GPS coordinates of respondent's homes. In the local context, a local leader or elite is someone who is a *gram panchayat* member, self-help group official, *anganwadi* teacher, doctor, school headmaster, or the owner of the main village shop. All our analyses include measures of network effects conditional on these numerous observables.

Given the richness and uniqueness of the data collected by Banerjee et al. (2014) and used to analyze the diffusion of microfinance, other projects have also been conducted in subsets of the 75 network villages, namely three half-day lab-in-the-field experiments (Chandrasekhar et al. 2013, Breza et al. 2015, and Chandrasekhar et al. 2012). Online Appendix I describes these other studies in detail, discusses treatment balance with respect to the other studies, and also shows robustness to saver-level controls for prior participation. The main results do not change with the addition of such controls, which is to be expected as these were several hour lab-in-the-field experiments that should not interact with a financial inclusion program several years later.

2.2. **Bank and Post Office Accounts.** In addition to the social network data, a key requirement of our study is convenient access to bank and post office branches for all participants. In each village, we identify one bank branch and the local post office branch to offer as choices to the savers. We select bank branches that satisfy several criteria: are located within 5km of the village, offer "no-frills" savings accounts,[11] and agree to expedite

[10]The main idea here is that individuals can communicate if they interact in any of the 12 ways. This is the network of potential communications and a good description of which individuals are likely to interact (in one of several ways in a day to day sense) with others.

[11]"No-frills" accounts generally have no minimum balance, charge no user fees, and require a minimal initial deposit, which is generally around Rs. 100 ($2).

our savings applications and process them in bulk.[12] Out of the 75 villages surveyed by Banerjee et al. (2014), 60 villages satisfy these criteria and constitute our final sample.

Each village in India is well-served by the postal branch network, and branches are generally within a 3km walk of each village in our sample. 35% of our study villages have a post office branch within the village boundary. We offer the post office choice because women often feel uncomfortable traveling to bank branches but feel much more comfortable transacting with the local post master. On the other hand, some individuals greatly prefer bank accounts because those accounts make it easier to obtain bank credit in the future.

Both the bank and postal savings accounts have very similar product characteristics. In each type of account, the minimum balance is typically Rs. 100 (~$2), and there are no account maintenance or withdrawal fees. The interest rates on the bank accounts range from 3% to 4.5%, and the interest rate on the postal accounts is 4%.[13] Users of both types of accounts are given a passbook, which is an official document containing the account information, the name, address and photo of the account holder, and the record of all account activity, both deposits and withdrawals. Entries in the passbooks are stamped with an official seal by branch personnel and cannot be forged. “No frills” bank and post office accounts have no formal commitment features. Participants in the study are allowed to withdraw freely from the accounts at any time. We stress that in our study, the only source of commitment comes from the presence of a monitor, described below. Our monitor treatments introduce an “informal commitment” in the parlance of Bryan et al. (2010).

2.3. **Experimental Design.** Figures 1 and 2 represent our experimental design and Figure 3 presents a time line. Study participants are randomly selected from an existing village census database, collected in conjunction with the network survey, and then randomly assigned to be part of our saver group, monitor group, or pure control (Figure 1.B).

All potential treatment savers and monitors who are interested in participating (Figure 1.C) are administered a short baseline survey, which includes basic questions on account ownership, income sources and desire to save. Only a quarter of households had a bank or post office account at baseline. We aim to test whether monitors can increase savings balances and also increase the use of already-accessible interest-bearing bank savings accounts.

Next, potential savers establish a six-month savings plan. Importantly, this plan is established before the saver knows whether she is assigned to the non-monitored treatment or one of the monitored treatments. Moreover, the saver does not know whether the village is assigned to endogenous monitor selection or random monitor selection.

The process of setting a savings goal includes listing all expected income sources and expenses month by month for six months. Savers are prompted to make their savings

[12]The 5km distance restriction meant that we were not able to work with only one bank, and instead opened accounts at branches of six different banks.

[13]Inflation rates around these periods are as follows: 2010: 9.47%, 2011: 6.49%, 2012: 11.17%, 2013: 9.13% (Inflation.eu, 2017)

goals concrete, and we record the desired uses of the savings at the end of the six-month period. Individuals are then invited to a village-level meeting in which study participation is finalized and treatment assignments are made. Potential monitors are also invited to attend the village meeting and are told that if selected, they can earn a small participation fee and incentive payment for participating.

Our sample frame for randomization is the 57% of savers who self-select into attending the village meeting (see Figure 2). We use two different data sources in Appendix Tables C.1 and C.2 to explore correlates with participation. In Appendix Table C.1, we use our responses to the short baseline survey to compare the participants with non-participants.[14] In Appendix Table C.2, we compare the participating households with the full set of non-participants using the village census data collected by Banerjee et al. (2014). Both tables show that participants disproportionately come from poorer households with a desire to save. Landless laborers are more likely, while salaried government workers are less likely to select into the sample. Moreover, the stated saving goals of participants are 8% smaller in size. This is consistent with poorer individuals having a harder time meeting their savings goals on their own. We also observe that households that actively save at a regular frequency and in which at least one member has a formal account are more likely to participate. Finally, individuals with exposure to RoSCAs and SHGs are almost 10 percentage points more likely to participate. This is a nice feature because these are the types of people who are prone to participate in social finance.[15]

From the pool of consenting participants and attendees of the village meeting, we randomly assign savers to one of three treatments (see Figure 1.E)[16]:

- T1 : Non-monitored treatment (Randomization at the individual level)

  We sometimes also refer to this treatment as the Business Correspondent treatment (BC). This is because the individual level treatment resembles a financial institution already in use in India called *business correspondents*. In this institution agents of the bank travel to villages to provide direct in-home customer service. This includes account opening procedures and deposit-taking. However, we were not legally able to collect deposits ourselves as researchers.
- T2 : Peer monitoring with random matching (Randomization to receive a monitor at the individual level, randomization of monitor assignment at the village level)
- T3 : Peer monitoring with endogenous matching (Randomization to receive a monitor at the individual level, randomization of monitor assignment at the village level)

[14]We note that, if during the initial visit, the potential savers inform the enumerators that are not interested in savings, then the baseline survey is not completed and they are not invited to the village meeting. However, they are included as potential savers in Figure 2.

[15]If there are villagers concerned about rotten-kin type of forces (Jakiela and Ozier 2016), they would be likely to both self-select out of our study as well as SHGs, which also render one's savings visible to a group.

[16]Let T0 denote the pure control treatment.

All individuals who attend the village meeting are assisted in account opening by our survey team. Savers are allowed to choose to open a bank or a post office account or to use an existing account, if applicable. We help savers to assemble all of the necessary paper work and "know your client" (KYC) identification documents for account opening and submit the applications in bulk. The savings period begins when all of the savers have received their new bank or post office account passbooks.

All savers in the individual treatment (T1, T2, and T3) are visited on a fortnightly basis. Our surveyors check the post office or bank passbooks and record balances and any transactions made in the previous 14 days and also remind savers of their goals.[17] This process gives us a reliable measure of savings in the target account on a regular basis. These home visits also serve as strong reminders to save.[18] We should note that in no treatment do our surveyors collect deposits on behalf of the savers.[19]

In our peer treatment with random matching (T2), we randomize the assignment of monitors to savers. In each village, a surplus of monitors turned up to the village meeting, so there were more than enough monitors for each T2 (or T3) saver. Every two weeks, after surveyors visit the T2 savers, they then visit the homes of the monitors. During these visits, the monitors are shown the savings balances and transaction records of their savers and are also reminded of each saver's goal. Thus, our intervention intermediates information between the saver and monitor. At the end of the savings period, monitors receive incentives based on the success of their savers. Monitors are paid Rs. 50 if the saver reaches at least half the goal, and an additional Rs. 150 if the monitor reaches the full goal.[20]

The peer treatment with endogenous matching (T3) is identical to T2, except for the method of assigning monitors. In this treatment, individuals are allowed to choose their monitor from the pool of all potential monitors attending the meeting. We only allowed one saver per monitor, so we randomized the order in which savers could choose. There was excess supply of monitors, so even the last saver in line had many choices. It is important to note here that the pool of potential monitors is recruited in an identical fashion in both sub-treatment groups (2) and (3). Table 1 presents summary statistics for the sample that attended the village meeting and also shows baseline differences between T1, T2, and T3.

For the most part, our sample is balanced across receiving a monitor or not and being in a random choice or exogenous choice village. Unfortunately, by chance, there is a small imbalance in saver centrality, which is driven by outliers in the random villages but persistent

[17]We were not able to obtain administrative data from the banks and post offices due to the large number of different institutions (Post office + branches of six different banks).

[18]Some participants even report that these visits are very motivating.

[19]This is one important difference between our product and the typical business correspondent model.

[20]Monitors also receive a participation fee at the time of the village meeting. Monitors that are ultimately selected receive Rs. 50 and those who are not ultimately selected receive Rs. 20. We had initially wanted to vary experimentally the size of the monitor incentives, but the required sample size was not feasible given our budget and the number of villages with both network data and a nearby bank branch willing to expedite our account opening. We investigate whether these incentives could be driving our results in Section 5.

in the endogenous villages.[21] Note that the vast majority of paper is concerned with the random choice villages, so this is less of a concern, and further, everywhere where controls are included, we control for saver centrality.

For our endogenous matching treatment, we chose to implement random serial dictatorship (RSD). Here, savers were ordered at random and were able to then select their monitors. This was a natural choice for several reasons. First, this mechanism is easy to implement in practice and therefore policy relevant. It is easy to explain to villagers, it is rather intuitive, and owing to its randomness it seems to be equitable. There was no resistance whatsoever to implementing such a scheme. Second, this design is easier to analyze given the randomization of the choice order. Additionally, it allows us to systematically explore which network aspects are valued when an individual selects a monitor. Does an individual select a more network-central monitor? Does an individual select a socially close monitor? Third, there is an equivalence between RSD and various other matching schemes with trading which reach the core.[22]

At the end of the 6-month savings period, we administer an endline to all savers and monitors. We record the ending balance in the target accounts from the saver's passbook. We also collect complete savings information across all savings vehicles (including other formal accounts, other informal institutions, "under the mattress", etc.) to make sure that any results are not just coming from the composition of savings. Importantly, we also administer this endline survey to all available attriters or dropouts.[23] Approximately 16% of savers dropped out of our experiment at some point after the village meeting, many of which never opened a target account for the savings period. We were able to survey approximately 70% of the dropouts in our endline follow-up survey and obtain information about their ending savings balances and other key outcomes. Table 1 also shows differences in the final sampled population decomposed between T1, T2, and T3. We find no differential attrition across the sample of savers captured in our endline data.[24]

[21] As is well-known, the centrality distribution has a long right tail.

[22] Specifically, consider two allocation mechanisms in an environment of n savers and n monitors, and say each agent has strict preferences over the monitors. The first mechanism is RSD. The second is when the monitors are (for instance) randomly allocated to the various agents and then trading is allowed. In this exchange economy, there is a unique allocation in the core and it can be attained by a top trading cycle (TTC) algorithm. Results in Abdulkadiroğlu and Sönmez (2003), Carroll (2012), and Pathak and Sethuraman (2011) show that various versions of RSD and TTC are equivalent: the mechanisms give rise to the same *probability distribution over allocations* irrespective of the preferences These results both characterize optimality of RSD as well as provide a justification for real-world use.

[23] We also surveyed a random subset of those who were chosen to be savers but who were not interested in savings, and a random subset of the pure control group.

[24] Another reason for why there is no differential attrition, in addition to the high rate of endline survey participation, is the nature of the attrition itself. One common reason for dropping out is a lack of the "know your client" (KYC) legal documentation required for opening a bank account (20% of dropouts). The most frequent reason for dropping out is dis-interest in saving. Further, the composition of why savers drop out is the virtually the same (and statistically indistinguishable) for monitored and un-monitored savers.

Finally, we administer a follow-up survey 15 months after the end of the savings period to the set of savers attending the village meeting. In addition to questions on savings balances, the survey contains retrospective questions from the savings period about how the savers saved and how frequently they spoke with their monitors. It also contains questions in the style of Dupas and Robinson (2013b) about unmitigated shocks sustained in the 15 months after the savings period ended.. We ask respondents about transfers made with their monitors after the end of the savings period and also friendships made through the monitor. The follow-up survey also contains questions about each respondent's beliefs about the savings behavior and level of responsibility of 12 other arbitrarily-selected savers. Appendix Table D.2 includes control group means for all of the variables from both of our endline surveys used in the analysis.

3. Framework

In this section we present the framework, which guides our empirical analysis. The details of the formal model are presented in Appendix A.

3.1. **Motivation.** Clearly one could tell several stories about why our experiment affects information flow and savings. Underlying the framework is our hypothesis that people expect that they are more likely to be perceived as responsible if they save more in the experiment. Importantly, future benefits in the village such as access to jobs or informal loans or other leadership opportunities accompany such positive perceptions. In 2016, to provide support for our framework, we conducted a survey of 128 randomly-selected subjects across 8 villages in our study area that had never participated in this nor any prior study conducted by either of the authors. The goal was to share with them a vignette of our experiment in order to gather perceptions of villagers not exposed to our experiment. We described our savings monitors study and asked subjects about whether information about savings progress would spread, how that might depend on which monitor was assigned, whether savers would in turn save different amounts depending on who the monitor was, and whether this might lead to returns in terms of favors or hearing about job opportunities in the future. Figure 4 presents the results of our survey.

First, subjects believe that people will talk. 73% of the respondents say that the monitor would spread information about poor savings of the saver and 93% of respondents say that the monitor would spread information about successful savings by the saver.

Second, subjects perceive that savers would perform better with a more central monitor. To operationalize this, we used the technique from Banerjee et al. (2014) to elicit names of central individuals in the village from a separate sample of villagers without collecting detailed network data. We provided each respondent with the name of one randomly-chosen villager and the name of one of the central villagers. Importantly, we did not explain to the villagers how we obtained those names or comment on the centrality of the named individuals. We then asked respondents if either named individual was the monitor, which would generate more savings. 63% believe the central monitor would generate more savings, 21% the average monitor would generate more, and 16% say they couldn't decide.

Third, subjects were asked about how much information about the saver's savings would spread under each of the two monitors: the majority perceive that there would be more information flow under the central monitors. 66% of the village would come to find out if there was a central monitor but only 41% if there was an average monitor.

Fourth, survey responses suggest that subjects are cognizant that even successful savers will often fall short of stated goals. Given a goal of Rs. 1,500, under an average monitor they predict Rs. 819 in savings but Rs. 1,132 under a central monitor.

Fifth, subjects recognize that better savers would experience better rewards for their savings behavior in the future. If given the choice between a saver who saved a large versus a small amount (Rs. 1,000 versus Rs. 100, with a goal of Rs. 1,500), they would predominantly be more likely to take the more successful saver for a supervisor job, an event organizer, or to be a village funds collector. On the other hand, for a manual laborer, the choice seems more even, as one may expect. Thus, the evidence suggests that respondents are more willing to allocate jobs that require greater responsibility to those who saved more, consistent with the interpretation that respondents interpret saving more in the experiment as a signal of responsibility.[25]

Taken together, this paints a picture of a setting where savers understand that monitors would talk about their progress, that there are returns to perceived reputation, that certain monitors spread information more widely, and that recognizing this, savers would work to save more if given such a monitor.

3.2. **Sketch of the Model.** As the survey responses suggest, the reputational consequences of monitors likely constitute an important source of motivation to save in our experiment.[26]Moreover, the impacts of the monitor may be heterogeneous by network position. To explore this further we embed the standard signaling model of Spence (1973) in a network.

In our model, agents decide whether to save with a signaling motive in mind.[27] We assume that responsible individuals have a lower cost of savings, and that an agent receives utility benefits from interacting with others in the future who consider her responsible. In

[25]An alternative story is that savings behavior does not contain useful information about responsibility but instead serves purely as a coordination device. The logic can be thought of like "we do not hire people who have low savings because we do not have enough jobs to go around and we need some rule for deciding who gets the job, even though we do not think that savings is related to job performance." That the respondents tend to pick higher savers systematically for jobs that require responsibility (supervisor, event organizer, village funds collector) but not for manual labor, at least suggests that such a coordination story is unlikely.
[26]In fact, it was actually a member of a village in a different study's focus group who originally suggested the core experimental design to us, citing the idea that reputation about individuals accumulating savings when they commit to do so could be leveraged to help encourage savings behavior.
[27]As in Spence (1973), we abstract from the direct benefits savers might receive from the act of saving alone.

our setting, once monitors learn about the savings decision of the saver, they can (and do) pass that information on to others in the network, who can in turn, continue to pass on the information. In this way, the set of people who will ultimately observe the signal sent by the saver is a function of the monitor's position in the network. Thus, when the agent decides to save, she will consider both how many people will learn about her actions and the likelihood of interacting with such individuals in the future.

The novelty of the model comes from how we take an otherwise complicated object – agents interacting this way on a network – and model it naturalistically. The model demonstrates that we can focus on two aspects of the agents' network position for our analysis: monitor centrality and saver-monitor proximity. Moreover, we obtain from the model a new network measure that predicts how a monitor's network position relative to the saver impacts savings that we can also take to the data.

Types. Each individual i has a type $\theta_i \in \{H, L\}$, where θ_i is independently distributed and $\mathrm{P}(\theta_i = H) = \frac{1}{2}$. Type H denotes a responsible agent. Each individual needs to pay a type-specific cost, c_{θ_i} to save, where $c_H < c_L$. This cost maps to the effort required for an agent to overcoming (with effort) her time inconsistency, temptations, or inattention issues, and these costs are lower for responsible types. Thus, being responsible is helpful for saving but also for other types of day-to-day behaviors and interactions in the community.[28]

Timing, Actions, and Payoffs. Our model has two periods. The first summarizes the entire 6-month savings period of our experiment, during which the agent makes her savings decision, $s_i \in \{1, 0\}$. For simplicity, we take this decision to be binary; she either saves a large or a small amount over the entire period. Next, the monitor is informed of the saver's decision, s_i. The monitor can then pass this information on to others, who many in turn continue to pass it on.[29] The monitor's conversations are not strategic and are independent of s_i. Through this diffusion process, which is governed by the network structure, a subset of the community will be informed of s_i at the end of the first period.

The second period of the model captures future interactions with villagers following the intervention's end. Savers interact with members of the community, again in a process governed by network structure, and receive payoffs from each interaction. Because these future interactions may take many forms under a wide set of circumstances, we model the payoffs in a reduced form way. The saver's payoff is equal to the third party's posterior expectation of the saver's productivity. Let $y(s_i)$ denote this posterior expectation.

[28]In fact, survey data shows that a randomly chosen individual is 6pp more likely to believe that an individual who reached her goal is responsible (mean 0.46) relative to an individual who did not reach her goal. Anecdotal evidence presented in Appendix Section B suggests that this influences how people will think of the saver in a labor market situation in the future, consistent with the survey evidence described above.

[29]As we show in Section 4.3, survey evidence documents that monitors do indeed pass such information to others and further, many savers have even heard about others talking about the their progress.

The saver saves $s_i = 1$ if and only if the expected increase in payoff to saving the high amount exceeds the cost of doing so:

$$q_{ij}\left[y\left(1\right)-y\left(0\right)\right] > c_{\theta_i}, \tag{3.1}$$

where q_{ij} is the expected number of individuals that saver i with monitor j will encounter in the future who also would have heard directly or indirectly about the saver's choice of savings. Thus, Equation 3.1 captures the fact that the saver only benefits from choosing $s_i = 1$ when she interacts with informed third parties.

Network Interactions. We model network interactions in a parsimonious and natural way. Namely, we assume that information flows and meetings occur in a similar way. In the first stage, the monitor informs her friend with some probability, that friend informs another friend with the same probability, and so on. Similarly, in the second stage, the saver meets any of her friends with some probability, meets a friend's friend with a lower probability (i.e., needs to meet a friend and also be referred to the friend's friend), and so on.

So, the probability that the agent meets some third party in the future, who will in turn have heard about the saver's choice of savings s_i through the network, will depend on the network structure and the position of the saver $i$and the monitor j in the network. By modeling information flow from monitors to others in the network and the possibility of the saver running into the third party through the network in this way, we compute:

$$q_{ij} = n \cdot \text{Social Proximity of Monitor and Saver} + \frac{1}{n} \cdot \text{Monitor Centrality} \times \text{Saver Centrality}$$

in the manner as described precisely in Appendix A.

Equilibrium. Under simple parameter assumptions,[30] we show there is a Perfect Bayesian equilibrium and a cutoff $\widehat{q}$ such that

$$s_i = \begin{cases} 1 & \text{if } \theta_i = H \text{ and } q_{ij} \geq \widehat{q}, \\ 0 & \text{if } \theta_i = H \text{ and } q_{ij} < \widehat{q}, \text{ and} \\ 0 & \text{if } \theta_i = L. \end{cases}$$

With this stylized structure on interactions, the signaling model predicts that in the data, we should see more savings where we have randomly assigned a saver to a monitor with higher q_{ij}. Ceteris paribus, q_{ij} is higher when (1) saver-monitor proximity is higher or (2) monitor centrality is higher. Therefore random assignment to higher centrality monitors or monitors closer to the saver in the network should yield higher savings.[31]

[30] If parameters are such that the cost of accumulating savings exceeds any possible posterior update about one's productivity, then of course everyone has $s_i = 0$.

[31] Note, if there was no signaling effect altogether because either productivity type was unrelated to cost of savings or because costs were systematically too high or too low for both parties and therefore only pooling, the regression would have no slope since all agents pool on $s_i = 0$.

4. The value of Central and Proximate Monitors

4.1. **Random Monitors.** Our main results analyze how the centrality, proximity, and combined model-based regressor value of randomly-assigned monitors influence savings. Before turning to this, we briefly discuss the average impact of monitors relative to the baseline treatment bundle (non-monitored treatment). The main outcome of interest is the log of total savings across all accounts. Conceptually this is the key outcome, as subjects could simply move funds from other places into the target formal account.[32]

Table 2 presents the results, showing effects on the log of total savings across all formal and informal savings vehicles. In column 2, we also include village fixed effects as well as saver controls for saving goal, age, marital status, number of children, preference for bank or post office account, baseline bank or post office account ownership, caste, elite status, number of rooms in home and type of electrical connection.

Finally, in column 3, we take a strict approach and use machine learning to select what out of the long list of controls we should include, which could potentially account for why we are seeing a treatment effect. This is the new technique called double post-LASSO of Belloni et al. (2014a) (see also Belloni et al. (2014b)). The idea is straightforward –because networks are not randomly assigned, and because we have many characteristics for which we could control, we allow machine learning (specifically LASSO) to pick out which covariates to include in the final regression specification. Here, our goal is to regress an outcome y on a treatment T, observing a a large vector of Xs. We use LASSO twice: first y on X to select X^{RF} and second T on X to select X^{FS}. Taking the union of these selected regressors as $X^{\star} = X^{RF} \cup X^{FS}$, in a final step, we regress y on T and $X^{\star}$. The coefficient on T resulting from this procedure is the estimated causal treatment effect. The underlying idea is that if some component of observables either explained treatment or the outcome variable, and therefore could explain the relationship of T to y, then we allow that component to be selected. Of course, because the monitoring treatment is random, the double post-LASSO procedure for estimating the treatment effect of receiving any random monitor it only deals with covariate imbalance. However, when estimating the effect of monitor network position (q_{ij} or proximity to monitor and centrality of monitor), below, double post-LASSO allows us to look at how the relationship of monitor network position and savings are affected or explained away by the other characteristics that the double post-LASSO selects.

Columns 1-3 present qualitatively similar results. We describe the results from column 3 and find that being randomly assigned to a monitor leads to a 0.35 log point increase in

[32]We developed our survey instrument after conducting numerous conversations about savings in similar communities. We attempted to be as comphrensive as possible in enumerating both formal and informal savings, including banks, post offices, SHGs, RoSCAs, MFIs, , insurance schemes, cash at home, and so on.

the total savings across all accounts. This corresponds to a 42% increase in savings across all savings vehicles of the households.[33][34]

Given these large impacts on overall savings, we next explore whether this increase is driven by a few individuals dramatically increasing their savings or by individuals across the group of savers more broadly. In Panel A of Figure 5, we plot the cumulative distribution functions of the log of total savings normalized by the savings goal for monitored vs. non-monitored savers. The figure suggests that the average treatment effects are not simply capturing large increases experienced by a small number of savers in the tail of the distribution. Indeed the intervention shifts savers to save more across the entire distribution (a Kolmogorov-Smirnov rejects that the CDFs are the same with $p = 0.08$).

In sum, having a randomly-assigned monitor helps increase savings significantly.[35]

4.2. **Monitor Centrality and Proximity.** We now turn to our main results: how q_{ij} (the model-based regressor), monitor centrality, and saver-monitor proximity influence saving behavior. Table 3 presents the results of regressions of log total savings across all accounts on monitor centrality, saver-monitor proximity, and q_{ij}, as well as a battery of controls.[36]

Columns 1-4 look at monitor centrality, saver-monitor proximity, both, and q (the model-based regressor), and all include village fixed effects, controls for the savings goal, saver centrality, and controls for saver and monitor characteristics including age, marital status, number of children, preference for bank or post office account (saver), baseline bank or post office account ownership, caste, elite status, number of rooms in home, and type of electrical connection, along with geographic distance between the homes of the saver and monitor.

In columns 5 and 6, we repeat the same exercises of columns 3 and 4, but use double post-LASSO of Belloni et al. (2014a). These provide our preferred estimates, though the results are comparable across the board. They are preferred because double post-LASSO employs a selection of regressors such that if some combination of covariates was effectively driving the effect on savings and we attributed it to networks, then the selector would include these and would actually kill the network effect we estimate. On the other hand, if regressors that predict neither the treatment (network position of monitor relative to saver) nor the outcome are being included, then it simply adds noise.

Consistent with our model, we find that being assigned to a central monitor or a proximate monitor generates large increases in savings. Namely, a one-standard deviation increase in the centrality of the monitor corresponds to a 0.153 log point increase in the log total

[33] Given that we find an increase in savings across all accounts, we need not fear that the treatment effects are simply the cause of moving funds from one location into the target account. In Online Appendix Table F.2, we find no evidence of such behavior.

[34] We have also checked whether participating in our experiment and receiving a bank account and bi-weekly visits increases total savings. We find very small statistically insignificant effects in Online Appendix P.

[35] In Appendix Table E.1, we show similar impacts on goal attainment (6.3pp increase, corresponding to an 80% increase in the likelihood relative to non-monitored savers).

[36] We stress that all accounts includes all formal and informal savings, including cash “under the mattress.”

savings, or a 14.5% increase – a large effect. Further, in Panel B of Figure 5, we explore the distributional effects of receiving a high centrality monitor vs. a low centrality monitor. Receiving a high centrality monitor does shift most of the distribution to the right, again suggesting that increases are not only driven by a small number of highly-impacted savers (a Kolmogorov-Smirnov test rejects equality of the CDFs, $p = 0.01$).

Turning to proximity, moving from a monitor of distance three to two leads to a 18.4% increase in the total savings across all accounts – again a large effect. Finally, in column 6, we look at the effect of our model-based regressor, q. A one standard deviation increase in the model-based regressor corresponds to a 33.5% increase in savings.

In Appendix Table E.2 we present the results of monitor centrality on the incidence of the saver reaching her goal. A one standard deviation more central monitor corresponds to a 2.9pp increase in the likelihood of a goal being met, which is just under half the effect size of being assigned an average centrality monitor. Similarly going from a monitor of distance three to two results in a 2pp increase in the likelihood of a goal being met.

Thus we show that randomly assigning a better monitor in terms of the model (q_{ij}), or randomly assigning a more central and more proximate monitors encourages savings across all accounts, including both formal and informal. That these results hold controlling for numerous demographic characteristics of both savers and monitors suggests that observables that may be correlated with network position cannot explain our proximity and centrality results. The covariate controls described above include caste group fixed effects and even the geographic distance between homes of savers and monitors. Traits such as these could have been thought to be driving the network effect through omitted variables, but our results are estimated conditional on this variation and a machine learning technique actually jettisons a number of controls and improves our estimates. Furthermore, magnitudes and significance are essentially the same even when entirely removing this bevy of characteristics (available upon request), which bolsters the idea that the effects are not driven by these characteristics.

4.2.1. *Multigraphs: Investigating multiple link-types.* We next investigate whether the observed patterns are driven by a specific slice of the multigraph. It is possible that the financial component or the advice component of the network could be driving the effects. This would be true if, e.g., financial information were only passed between individuals conducting financial transactions with one another. Table 4 presents a version of our main specification, but allows centrality, proximity, and the model-based regressor to vary by relationship type. While the results get noisy, we find that only the centrality, proximity, and model-based regressor in the union of all relationships appear to matter. This is natural as individuals could pass information across link types: for instance, to a coworker, who then tells a friend, who then tells his neighbor about the information when borrowing rice.

4.3. **Effect of Central Monitors on Beliefs about Savers.** We next make use of novel supplemental data to provide evidence in support of the reputational mechanism of Section

3. One necessary condition for reputation to be at play is for the monitors and other community members to actually discuss the savings of participants. In fact, more than 60% report doing so in the last two weeks of the savings period. Further, 40% of monitored savers also report that the monitor passed information about their progress to others.[37]

Moreover, we attempt to track this information flow from monitors to other community members. Our follow-up survey, administered 15 months after the end of the intervention, asks respondents their views about a randomly-chosen set of 12 savers who participated in our experiment. Namely, we capture a measure of responsibility – whether the saver is viewed generally (in avenues beyond savings alone) as being good at meeting self-set goals. We test whether community members update their beliefs about the saver's ability to meet goals more in response to their behavior in our experiment when the monitor is central.

Table 5 presents the results of this exercise. We examine a regression of whether the interviewee updated her beliefs about the general ability of the saver to reach her goals in the direction of the saver's savings goal attainment on the centrality of the randomly assigned monitor as well as the proximity between the interviewee and the saver's monitor (Columns 1-3). We repeat this exercise, changing the outcome variable to whether the interviewee knows correctly whether the saver reached her goal (Columns 4-6). Our preferred outcome is the update in the responsibility metric: if the monitor is more central, a random interviewee in the village is more likely to have a better view of the saver's general responsibility if she succeeded or a worse view of it if the saver failed. Our regression specifications include no fixed effects, village fixed effects or interviewee fixed effects, the latter of which therefore captures variation within an interviewee but across randomly assigned saver-monitor pairs. We find that indeed if a saver is randomly assigned a more central monitor, the respondent is more likely to believe that the saver is good at meeting her goals and also is more likely to know if the saver reached her goal.

While interesting, this dynamic is not necessary for our story. Specifically, it need not be the case that the information has already or immediately spread. What is important in our framework is that when the saver impresses the monitor, there may be benefits at some point when a new opportunity arises (much like sending a letter of recommendation).

It should go without saying that this is an admittedly imperfect exercise. We use self-reported data on whether people chat about others, whether people hear gossip about themselves through back channels, and several questions about respondents perspectives on other savers' responsibility profiles and savings habits in the experiment. The usual caveats about self-reported data certainly apply here and, further, we are not making a causal claim that this shift in beliefs exactly corresponds to the shift in savings. Nonetheless, we want to emphasize that the evidence provided here is (a) largely consistent with our

[37]We asked "did your monitor tell others about your savings goal or about your progress towards trying to meet it?" This is striking because it requires enough communication such that savers hear gossip about themselves. Our own reflection suggests it is rare for people to gossip about a subject in front of him/her.

framework/story, (b) mostly self-consistent, and (c) agrees with the anecdotal evidence provided in Appendix B. Further, given the difficulties in digging into such a mechanism in a networks setting, we argue that this simple idea – simply asking whether conversations happened, asking whether people changed their views of others, etc. – which has not been used much in this literature, has tremendous value for this research program.

Consistent with the perception effects, conversations with study participants and other villagers support the idea that reputational mechanisms are at play in our experiment. In fact, our experimental design was based, in part, on a conversation with a gentleman in a rural village. In Appendix Section B, we present short excerpts of conversations with participants that we recorded. Many villagers described wanting to impress their monitor in general and paying special attention when that monitor was important. Some respondents gave us specific examples of why impressing the monitor would be helpful in the future. Finally, we remind the readers that in our follow-up survey across 128 subjects in 8 new villages, the responses were consistent with what has been documented here (Figure 4).

4.4. **Longer-Run Impacts.** Given that our treatment increased total savings across all accounts, we next ask whether we can detect any lasting benefits of the additional savings. This is a difficult question, so to address this, in our 15-month follow-up survey, we adopt the methods proposed by Dupas and Robinson (2013b). We asked subjects about their ability to cope with various shocks. Given that our intervention helped savers to increase their stock of savings, we can ask if in the subsequent 15 months, they were less likely to be in a situation where they did not have money to cope with a shock.[38]

Specifically, we posed a series of questions to the savers as to whether they faced a specific hardship for which they did not have enough savings to purchase a remedy (e.g., falling ill and being unable to purchase medicine). Table 6 presents the results. We measure effects on the total number of unmitigated shocks (columns 1-2), whether the household experienced fewer unmitigated shocks than the median (columns 3-4), incidence of unmitigated health shocks (columns 5-6) and incidence of unmitigated household consumption shocks (7-8). Specifications are shown with and without village fixed effects, and all regressions use the standard saver controls. We find that being randomly assigned a monitor leads to a decline in the rate at which individuals face a shock and are unable to purchase a remedy. For instance, there is a 0.197 decline in the total number of shocks (on a base of 1.769, column 1, p-value 0.13). Further, there is a 7.5pp decline in the probability that a household has greater than median number of instances where they were unable to cope with the shock (p-value 0.076). We find suggestive, though not statistically significant effects when we look at health and household expenditures as separate categories. We acknowledge that the types

[38]Note that this could arise for two reasons. First, and perhaps the ex ante more likely reason, agents would have more money to deal with the same distribution of shocks. Second, agents could have invested in shock mitigation. Like Dupas and Robinson (2013b), our analysis does not need to take a stand.

of shocks that the intervention helped savers to mitigate are likely of modest scale.[39] The key point is that there are, nonetheless, tangible benefits of savings for situations like these. Note that it could also be another channel: the tangible benefit of improved reputation, which may cause others to be willing to help the saver in times of need.[40]

Finally, in the last two columns of Table 6, we present the effects of the random monitor treatment on log savings balances 15 months after the intervention. Remarkably, while a bit noisier, the size of the increase in savings is as large as that reported in Table 2 (p-value 0.13). This suggests that individuals are able to maintain their savings even after the monitors are no longer actively receiving information. Appendix Figure D.1 shows that the increases in savings across the distribution are still apparent 15 months later.

We have documented increases in savings both during and considerably after our experiment. One concern the reader may have is whether there is wide-spread over-savings. We emphasize our results are for total savings across all household accounts; recall, only the target account is revealed to the monitor. So the fact that we find effects here suggest that over-savings is not likely. Additionally, there are other reasons for why we believe over-savings was not a wide-spread issue: (1) only information in the target account is provided to the monitors, (2) the main effects across all household savings accounts are persistent; (3) recall people could have opted out at any time, ending the revelation of information and chose not to; (4) as just discussed we demonstrate increased labor supply likely drives the savings increases rather than taking costly loans or reducing essential consumption; (5) we document positive effects on substantive outcomes like shock mitigation.

These findings serve to show that there was truly an increase in savings (since they were better able to make purchases to cope with shocks) that persisted after the intervention and moreover show that there were important, real consequences of the increased savings.

5. Threats to Validity

5.1. **Negligibility of Monitor Incentives.** There are two natural questions one may ask when it comes to monitors in this study. First, is it the case that the presence of the monitor causes individuals to unwind their savings from other accounts? Second, does the fact that the monitors received a small incentive drive the results?

We show evidence against both of these hypotheses. Conditional on reaching her goal, a saver exceeds 200% of her goal in 65% of the cases. Further, over 75% of individuals who reach their goal in the target account save in excess of 200% of their target amount across all accounts. This suggests that individuals are not likely subject to undue pressure. They save immensely, mostly do not bunch at their goal, and don't unwind across other accounts.

Turning to the monitor incentives, we do the following. Recall that the monitor incentive function has two discontinuities. In addition to the payment made at the full goal, we

[39]We are not claiming that the gains in savings had large persistent health benefits, for example.
[40]We thank an anonymous referee for this comment.

added a second discontinuity at the half goal to generate a test. In terms of personal value to the saver, the incentives above and below the half goal should be smooth. So testing for bunching above this threshold should identify how the monitor incentive may have differentially led to behavior nudging people across the threshold. Turning to the full goal amount, this is a mix of potential monitor incentives but also natural incentives to simply reach one's stated goal: they may be saving up for something specific and furthermore, after all, it is a goal. Both are natural motivations to bunch at the goal.

Table 7 presents the results. The outcome variable is a dummy for whether the saver who is in the window of the specified value (1/2 or full goal) has saved weakly more than the value. In column 1 we look at the 1/2 goal and full goal savings amounts for each saver and look within a window of the bonus (Rs. 50 or Rs. 150) of each. The first three rows constitute our test of interest as they focus on the 1/2 goal mark. We see that unmonitored savers, conditional on being in the window around the 1/2 goal, are 88.9% likely to be weakly greater than 1/2 their goal. This drops by 24.6pp (p-value 0.086) or 20.5pp (p-value 0.15) when one has a random or endogenous monitor, respectively. This suggests that if anything, the fact that the monitor may have an incentive makes it less likely for the saver to just clear the threshold. Of course we interpret this as the monitor incentive having no meaningful effect, not that it disincentivizes clearing the threshold.

In columns 2-4, we repeat the exercise scaling the window by 3/2, 2, and 7/3 (so Rs. 66/Rs. 200, Rs. 100/Rs. 300, and Rs. 116/Rs. 350 respectively). Notice that the set of observations in the window do not change across columns 1 and 2 and similarly 3 and 4 for the 1/2 goal mark. Our results remain essentially the same and we gain precision for the endogenous monitoring case. Notice that the endogenous and random monitoring case cannot be distinguished from half the savers on either side of the window.

Overall this rejects the bunching hypothesis since, first, in the monitored groups it is as good as random that people are on either side of the window but, further, if anything the unmonitored group is significantly more likely to bunch on the right of the 1/2 goal mark despite not facing any monitor incentives by definition. We believe that this is a good test of the impact of monitor incentives because 1/2 goal is not a particularly salient milestone for the saver aside from the monitor incentive.

Rows 4-6 present the same estimates but for the full goal. Note that by construction there is likely to be more bunching here (ex ante) simply because individuals set goals for themselves and they may be saving for specific goods. With the most conservative window we find that 60% of unmonitored savers fall at or within Rs. 150 above the goal whereas that fraction is 86% and 70% for the monitored savers. As the window widens, the share at or above the goal stays roughly the same under monitoring and declines to 20%-29% for unmonitored savers. This is not surprising because bunching should happen irrespective of the incentive. Also, as the window is widens, one is adding in the treatment effect.

Because we find so little evidence of gaming, we believe that many of our monitoring results would still hold even in absence of financial incentives. However, an experimental test is required to confirm this hypothesis.

5.2. **Robustness of our Results.** We now describe the results of two robustness exercises. First, we might be worried about measurement error: it is important to see that in fact savings were achieved and also that we can at least partially understand the source of the increased savings. Second, because we do have survey attrition in the sample, we show robustness of our results to corrections for that attrition.

In Appendix F, we deal with the first concern and examine how the savers saved. We tackle this in two ways: first, using a detailed expenditure survey in the sixth month of our savings period and second, using a retrospective survey in our follow-up fifteen months after the savings period ended. Table F.1 Panel A presents the results of the first exercise. We find that being assigned a random monitor leads to noisy 9% decline in total expenditures ($p = 0.118$). In levels, with a 5% winsorizing to deal with outliers, there is a Rs. 560 decline in total expenditures ($p = 0.057$). We see, consistent with anecdotes and our retrospective survey, evidence of decline in festival expenditure (by Rs. 238), decline in transportation expenditure (by Rs. 157), and an increase in tea consumption (by Rs. 35, which is a common drink to take on the job). These results are somewhat noisy but suggestive.

Panel B provides a more-powered view, albeit through a retrospective survey. Assignment to a random monitor corresponds to a claimed 7pp increase in labor supply on a base of 15%, a 2pp increase in business profits on a base of 3%, and 7.8pp reduced unnecessary expenditures on a base of 15%. The first and third of these effects are significantly different from zero in addition to being qualitatively large. Reassuringly, while it seems more work and better budgeting led to savings, there is no increase in borrowing money from one's network, no reduction in transfers to others, and no borrowing to save.

Throughout the paper, we drop observations for which we do not have total savings information from our main total savings regressions. Recall from Table 1 that attrition is balanced across monitored and non-monitored savers. Nonetheless, our main regression estimates might be conservative if monitors disproportionately caused individuals with large savings balances to attrit from the study and those with small savings balances to remain, or they might be overstated if monitors caused the better (worse) savers to remain (attrit). For that reason, we conduct an exercise using Lee bounds in Appendix G. Note that the method is constructed for binary treatment variables. Thus applying Lee bounds to the treatment effect of receiving any random monitor is straightforward. When we look for the effect of a more central and proximate monitor, we construct a binary indicator for whether the saver's monitor has a value of the model-based regressor in the top 25th percentile.

Table G.2 presents the results. Looking both at the effect of having a random monitor and, conditional on the random monitoring sample, the effect of having a monitor with a high value of the model-based regressor generates lower bounds that are only modestly smaller than our main regression estimates (for instance a lower bound of 0.31 on the value of having a high model-based regressor monitor as compared to a point estimate of 0.451). Of course, our bounds are noisy, and we have used female as a binary predictor of attrition (see Table G.1) to tighten the bound.

As mentioned above, our results all go through when we use goal attainment in the target account as our dependent variable (see Online Appendix E). In these specifications, the outcome data is generated from the passbooks of the savers. If a participant dropped out of the study before its completion we denote that individual as failing. The combination of these exercises strongly suggests that our results are likely to be robust.

6. Endogenous Monitors

6.1. **Endogenous Monitors Benchmark.** The previous results suggest that a social planner interested in maximizing savings could "optimize" the allocation of monitors to savers using the network as an input. However, such an allocation mechanism is likely infeasible for most real-world institutions.In many of the informal peer-based financial arrangements that are commonly found in developing countries, individuals endogenously sort into groups.[41] Thus, measuring savings under endogenous monitor choice is a useful policy-relevant benchmark. For obvious reasons, causally determining what drives choice is beyond the scope of the paper. That is neither our aim nor what we claim to measure. Our goal, rather, is to assess how well the community does when left to its own devices.

To measure this benchmark, we analyze the savings outcomes in the 30 villages with endogenous monitor choice. It is important to note that a priori, savings could be higher or lower in endogenous relative to random monitor allocation. On the one hand, savers might completely unwind all savings benefits of monitors. It may also be the case that some individuals feel constrained socially in their ability to choose their preferred monitor.[42] On the other hand, individuals might arrive at the "optimal" savings-maximizing allocation of savers to monitors. Thus, any outcome between "optimality" and full unwinding is feasible.

Table 8 presents the log total savings of participants in endogenous and random choice villages with and without monitors. In column 1, we include village fixed effects., so the estimated coefficients measure the effects of receiving a monitor relative to non-monitored savers in the same village. The savers who picked their own monitors save no more than the savers who were not assigned to receive a monitor (insignificant coefficient -0.0830).

However, when we remove the village fixed effects, column 2 suggests that the negative and insignificant coefficient can be explained by a large spillover effect on the control group.

[41]Examples include Stickk.com, which asks individuals to choose a "referee" to monitor their progress toward a goal. Also, MFIs, ROSCAs, and SHGs often involve endogenous group formation. We also note that a financial institution in India approached us to implement a similar program in one of their urban branches.

[42]For example, low caste individuals may feel uncomfortable choosing high caste monitors. Similarly, low income day laborers may feel that they aren't entitled to pick important people in the village.

Relative to the non-monitored savers in the villages with random monitor assignment (omitted category), non-monitored savers in the endogenous choice villages increase savings by 0.35 log points. Moreover, the total savings effect is not statistically different from that of monitored savers in either monitor treatment.[43] Thus, we find that endogenous monitors are about as good as having a randomly assigned monitor and, more interestingly, that even the unmonitored individuals in endogenous villages save similarly well.

Why the non-monitored savers save more in endogenous choice villages is an interesting question. Given that we did not expect such an outcome, we can only speculate as to the exact mechanism. We think that it is most likely that endogenous choice led to an increase in the number of conversations in the village about savings. For example, we observe that savers ran into their endogenously chosen monitors more than their randomly assigned monitors (5.1 versus 4.0 times per fortnight - difference significant at the 1% level).[44] These conversations may have motivated some of the non-monitored savers to save more. In Appendix M we conduct an exercise to test for spillovers from monitored to non-monitored savers.[45] We do find evidence that the monitors, and especially the high centrality monitors, affect the savings of the friends of their savers.[46] Better understanding these spillovers is an interesting direction for future research.

Thus the results show that the community does well implementing our informal peer-based financial product on its own. So even if it is not feasible to optimize the matching of savers to monitors, the community can still benefit from an endogenous implementation.

6.2. **Exploring Choice.** While our experiment was not designed to fully unpack monitor choice, we end by exploring one specific aspect of choice. To do this, we extend our signaling model in Appendix N.1 to develop intuitions for which individuals might pick more central and proximate monitors and where choice order may matter. The model extension also provides a framework for thinking about who might self-select into the experiment.

We consider agents of both heterogeneous quality and centrality, who first decide whether or not to opt into the experiment, knowing that if they do, they will be assigned to BC, random monitoring, or endogenous monitoring. In the endogenous treatment, agents also choose their monitors from the available pool, and agents know this. Our extended model shows the complexities in modeling choice in the endogenous treatment, even abstracting

[43]Appendix Table E.3 investigates the effects of random and endogenous monitors on goal attainment. There we see that again, endogenous and random monitors generate similar levels of goal attainment. However, we do not observe a goal attainment spillover onto the non-monitored savers in the endogenous villages.

[44]In contrast, planned meetings between savers and their monitors changed by a much smaller, insignificant amount (2.5 vs. 2.3, p-value 0.4).

[45]We also show that allowing for such spillovers does not change our main results in the random allocation villages. The logic is that having a friend who is randomly assigned a monitor, conditional on participating, is orthogonal to receiving a monitor oneself or that monitor's location in the network.

[46]Finally, it is also possible that the ability for savers to choose their own monitors increases the desirability of the program and the buy-in of the village.

away from the likely forces that may also affect choice (whether people are amicable, forgiving, encouraging, etc.). We focus on one specific subtlety – that H types have an incentive to enter our experiment and choose highly central monitors in the endogenous treatment, if they are available, whereas not only do L types want to choose low centrality monitors to avoid being revealed, but highly central L types may not even opt into the experiment.[47] Therefore, when we look at choice, the theory suggests that high centrality savers should be more likely to choose better monitors. Further, if high centrality monitors are scarce, there should be a relationship between choosing early and choosing more central monitors, but only among the highly central. This is indeed what we observe in the data in Online Appendix N.1, Figure N.1 and Table N.1, respectively.

7. Conclusion

Reputations matter. Our subjects enunciate this both in direct surveys and through their economic decisions. When it is known that information about their savings is transmitted to others in the community, participants increase their savings in meaningful enough amounts that they are better able to mitigate shocks.

But reputation in *whose* eyes also matters, and the social network provides an apt lens to examine this. Individuals benefit from impressing their monitors because, in the future, they might need to rely either on the monitor directly or on parties who have come to learn about them from the monitor. This motive to impress is undoubtedly asymmetric in communities. Certain sets of people interact more or less frequently with others, and a network perspective puts discipline on thinking about how reputational stakes may vary with the position of one's monitor in the community.

Our field experiment is carefully designed to quantify impacts on a measurable and important behavior – savings. Further, we collect evidence pertaining to how the households saved, whether the savings had follow-on benefits, and whether the savings accumulation persisted. We make a methodological contribution toward measuring reputation by tracking the information flow itself from the monitors to other members of the community.

The findings of this experiment speak to a general discussion in development economics about the nature and role of social sanctions that may support informal financial institutions. In our simplified setup, a benefit or sanction is simply getting a good or bad name as demonstrated by one's effort to accumulate savings. We show that monitors do pass on information, savers desire to be perceived as responsible, and savers make payments into the monitored accounts. These findings document empirically the forces alluded to in the literature (e.g., Besley and Coate, 1995; Munshi, 2014). Further, because the degree of information that is passed on is correlated in a convincing manner with the network position

[47]Consistent with this observation, saver centrality is correlated with total savings in our data, conditional on savings goal, though this may be a spurious correlation for a variety of other obvious reasons.

of the monitor, the identity of who in a community can leverage this reputational motive is an important factor when considering whether networks can sustain good behavior.

The forces described here are likely to operate in settings where a primary barrier to savings accumulation is a failure of responsibility (including things like time inconsistency or inattention). Further, it is essential that social reputation carries weight, or more generally network relationships with community members carry great weight. These conditions are likely to hold in exactly the types of communities that are able to sustain RoSCAS, SHGs, VSLAs, and other types of informal financial structures. However, contexts where the primary barriers include lack of access to financial institutions and urban settings where individuals often have access to markets or a number of alternatives beyond their networks, are unlikely to be ones where our mechanism would be strongly at play.

References

Abdulkadiroğlu, A. and T. Sönmez (2003): "School choice: A mechanism design approach," *American economic review*, 729–747.

Alatas, V., A. Banerjee, A. G. Chandrasekhar, R. Hanna, and B. A. Olken (2012): "Network structure and the aggregation of information: Theory and evidence from Indonesia," *NBER Working Paper 18351*.

Ashraf, N., D. Karlan, and W. Yin (2006): "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *Quarterly Journal of Economics*, 121, 635 – 72.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006): "Who's who in networks. wanted: the key player," *Econometrica*, 74, 1403–1417.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. Jackson (2013): "The Diffusion of Microfinance," *Science*, 341 (6144).

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2014): "Gossip: Identifying central individuals in a social network," *NBER Working Paper 20422*.

Beaman, L., D. Karlan, and B. Thuysbaert (2014): "Saving for a (not so) Rainy Day: A Randomized Evaluation of Savings Groups in Mali," *NBER Working Paper*.

Belloni, A., V. Chernozhukov, and C. Hansen (2014a): "High-dimensional methods and inference on structural and treatment effects," *The Journal of Economic Perspectives*, 28, 29–50.

——— (2014b): "Inference on treatment effects after selection among high-dimensional controls," *The Review of Economic Studies*, 81, 608–650.

Beshears, J., J. J. Choi, D. Laibson, B. Madrian, and J. Sakong (2011): "Self control and liquidity: How to design a commitment contract," *RAND Corporation Working Paper*.

Beshears, J. L., J. J. Choi, D. Laibson, B. C. Madrian, and K. L. Milkman (2015): "The Effect of Providing Peer Information on Retirement Savings Decisions," *Journal of Finance, Forthcoming.*

Besley, T. and S. Coate (1995): "Group lending, repayment incentives and social collateral," *Journal of development economics*, 46, 1–18.

Besley, T., S. Coate, and G. Loury (1993): "The economics of rotating savings and credit associations," *The American Economic Review*, 792–810.

Bramoulle, Y., H. Djebbari, and B. Fortin (2009): "Identification of peer effects through social networks," *Journal of Econometrics*, 150, 41–55.

Breza, E. (2014): "Peer Effects and Loan Repayment: Evidence from the Krishna Default Criris," *Working Paper.*

Breza, E., A. G. Chandrasekhar, and H. Larreguy (2015): "Network Centrality and Institutional Design: Evidence from a Lab Experiment in the Field," *NBER Working Paper 20309.*

Brune, L., X. Giné, J. Goldberg, and D. Yang (2013): "Commitments to Save: A Field Experiment in Rural Malawi," *Working Paper.*

——— (2016): "Facilitating savings for agriculture: Field experimental evidence from Malawi," *Economic Development and Cultural Change*, 64, 187–220.

Bryan, G., D. Karlan, and S. Nelson (2010): "Commitment devices," *Annu. Rev. Econ.*, 2, 671–698.

Bryan, G. T., D. Karlan, and J. Zinman (2012): "Referrals: Peer Screening and Enforcement in a Consumer Credit Field Experiment," *NBER Working Paper.*

Bursztyn, L., F. Ederer, B. Ferman, and N. Yuchtman (2013): "Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions," *Econometrica.*

Cai, J., A. de Janvry, and E. Sadoulet (2013): "Social Networks and the Decision to Insure," *Working Paper.*

Carroll, G. (2012): "When Are Local Incentive Constraints Sufficient?" *Econometrica*, 80, 661–686.

Chandrasekhar, A. G., C. Kinnan, and H. Larreguy (2011): "Information, Savings, and Informal Insurance: Evidence from a lab experiment in the field," *MIT Working Paper.*

——— (2013): "Informal Insurance, Social Networks, and Savings Access: Evidence from a lab experiment in the field," *Working Paper.*

Chandrasekhar, A. G., H. Larreguy, and J. Xandri (2012): "Testing Models of Social Learning on Networks: Evidence from a Framed Field Experiment," *Working Paper.*

DeMarzo, P. M., D. Vayanos, and J. Zwiebel (2003): "Persuasion bias, social influence, and unidimensional opinions," *The Quarterly Journal of Economics*, 909–968.

Dupas, P. and J. Robinson (2013a): "Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya," *American Economic Journal: Applied Economics*, 5, 163–92.

——— (2013b): "Why Don't the Poor Save More? Evidence from Health Savings Experiments," *The American Economic Review*, 103, 1138–71.

Feigenberg, B., E. Field, and R. Pande (2013): "The Economic Returns to Social Interaction: Experimental Evidence from Microfinance," *Review of Economic Studies*, 80, 1459–1483.

Giné, X. and D. Karlan (2006): "Group versus Individual Liability: Evidence from a Field Experiment in the Philippines," Yale University Economic Growth Center working paper 940.

Golub, B. and M. O. Jackson (2010): "Naive learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics*, 112–149.

——— (2012): "How Homophily Affects the Speed of Learning and Best-Response Dynamics," *The Quarterly Journal of Economics*, 127, 1287–1338.

Hagen, L. and A. Kahng (1992): "New spectral methods for ratio cut partitioning and clustering," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 11, 1074–1085.

Inflation.eu (2017): "Historic inflation India - CPI inflation," *http://www.inflation.eu/inflation-rates/india/historic-inflation/cpi-inflation-india.aspx*.

Jack, W. and T. Suri (2014): "Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution," *The American Economic Review*, 104, 183–223.

Jackson, M., T. Barraquer, and X. Tan (2010): "Social Capital and Social Quilts: Network Patterns of Favor Exchange," .

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

Jakiela, P. and O. Ozier (2016): "Does Africa need a rotten kin theorem? Experimental evidence from village economies," *The Review of Economic Studies*, 83, 231–268.

Karlan, D. (2007): "Social connections and group banking*," *The Economic Journal*, 117, F52–F84.

Karlan, D., M. McConnell, S. Mullainathan, and J. Zinman (2010): "Getting to the Top of Mind: How Reminders Increase Saving," *NBER Working Paper*.

Kast, F., S. Meier, and D. Pomeranz (2012): "Under-savers anonymous: Evidence on self-help groups and peer pressure as a savings commitment device," *NBER Working Paper*.

Kast, F. and D. Pomeranz (2014): "Saving More to Borrow Less: Experimental Evidence from Access to Formal Savings Accounts in Chile," *NBER Working Paper*.

Katz, E. and P. F. Lazarsfeld (1970): *Personal Influence, The part played by people in the flow of mass communications*, Transaction Publishers.

Munshi, K. (2014): "Community networks and the process of development," *The Journal of Economic Perspectives*, 49–76.

Pathak, P. A. and J. Sethuraman (2011): "Lotteries in student assignment: An equivalence result," *Theoretical Economics*, 6, 1–17.

Platteau, J.-P. (2000): *Institutions, social norms, and economic development*, Harwood Academic, chap. Egalitarian norms and economic development, 189–240.

Prina, S. (2013): "Banking the poor via savings accounts: Evidence from a field experiment," *Working Paper*.

Schaner, S. (2014): "The Persistent Power of Behavioral Change: Long-Run Impacts of Temporary Savings Subsidies for the Poor," *Working Paper*.

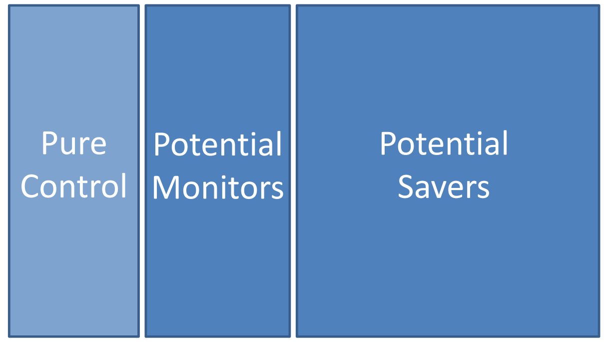
Spence, M. (1973): "Job market signaling," *The quarterly journal of Economics*, 355–374.

Thaler, R. H. and S. Benartzi (2004): "Save More Tomorrow: Using behavioral economics to increase employee saving," *Journal of political Economy*, 112, S164–S187.

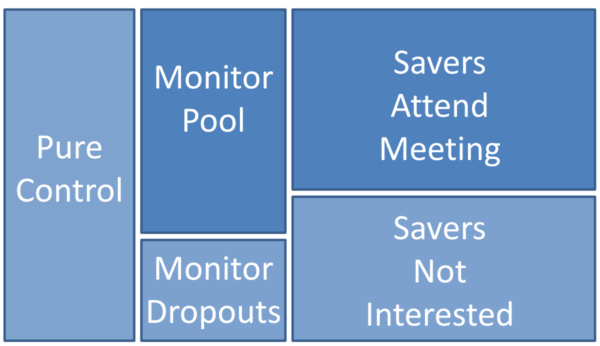
Figures

Figure 1. Experimental Design and Randomization

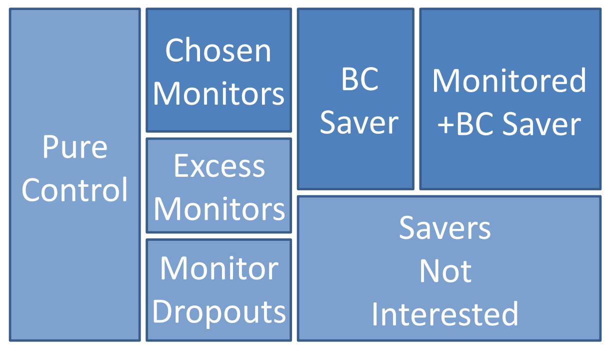
(a) Village



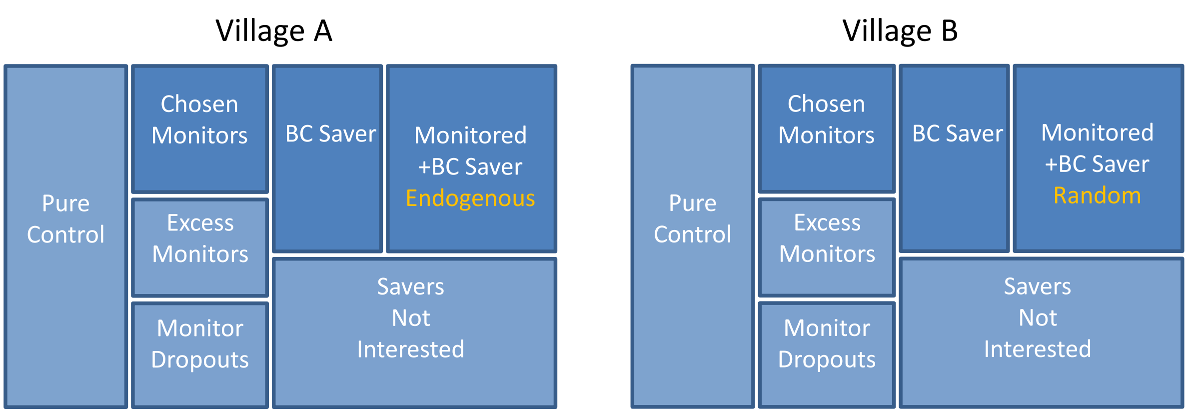
(b) Individual- level randomization to pure control, monitor pool or savers pool.



(c) Participating samples.



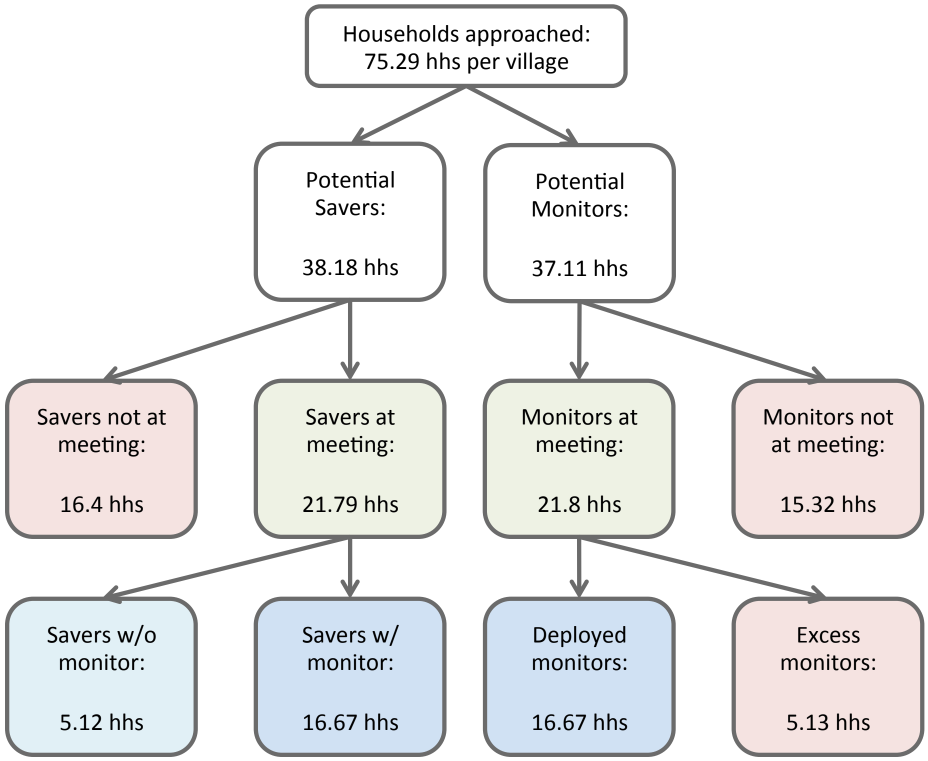
(d) Individual-level randomization of participating savers to treatments. Monitors selected.



(e) Village-level randomization. Village A is randomly assigned to endogenous monitoring treatment. Village B is randomly assigned to exogenous monitoring treatment.

Notes: "BC Saver" refers to our non-monitored treatment (T1) described in section 2.

FIGURE 2. Sample Description



Notes: This figure shows the average number of households per village in our sample in each cell. There are 60 villages in the sample, with an average population of 222.12 households. Every village has an average of 146.83 pure control households that were not approached at all before or during the savings intervention.

FIGURE 3. Time Line of Experiment

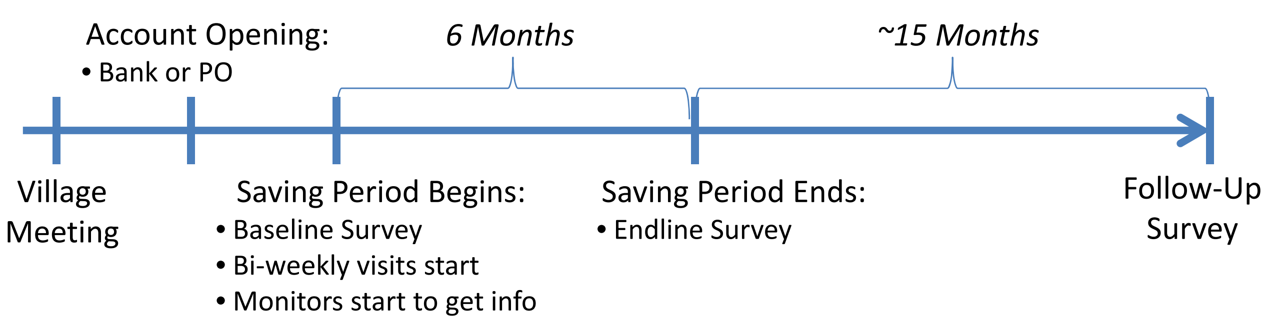
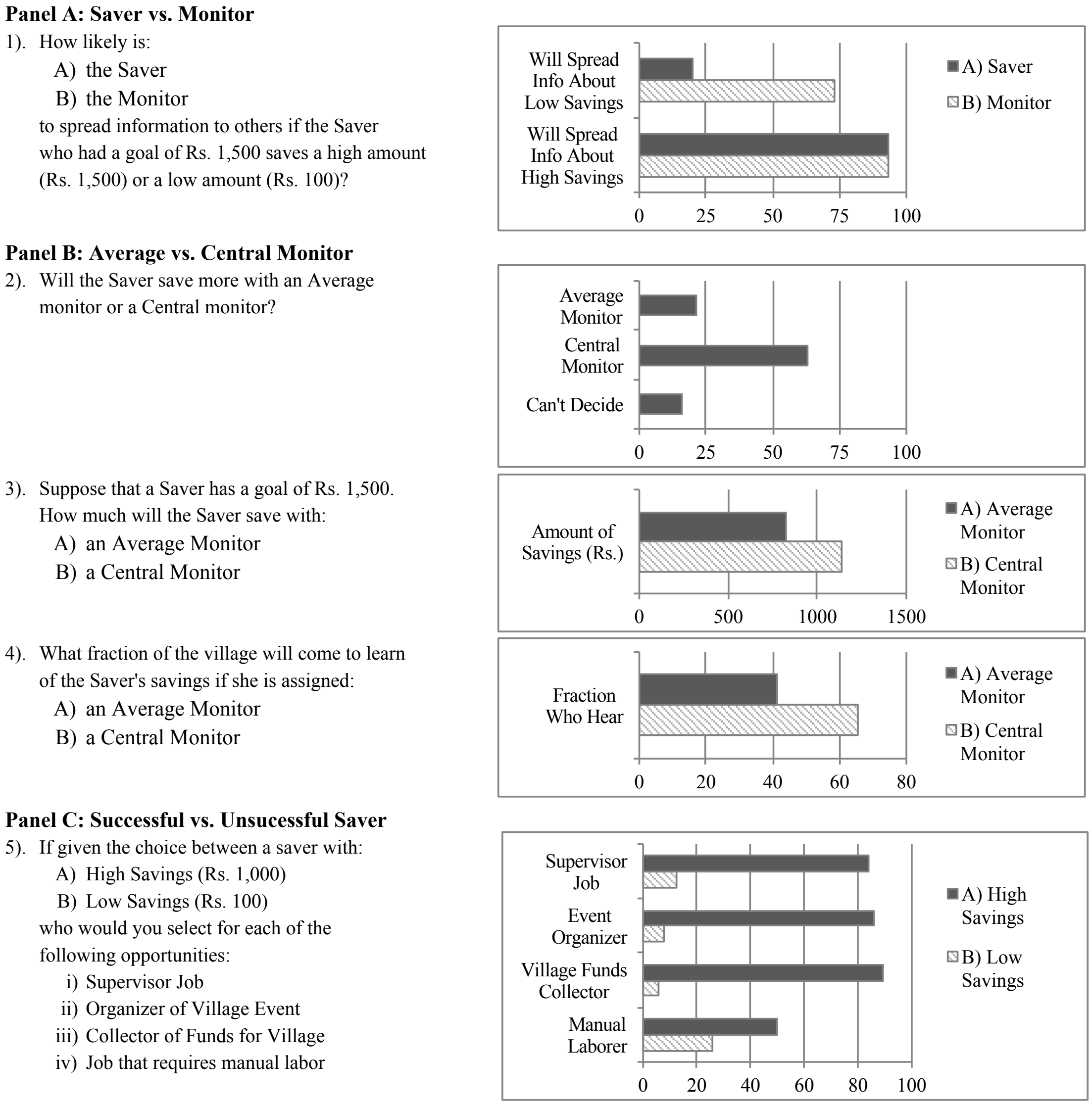


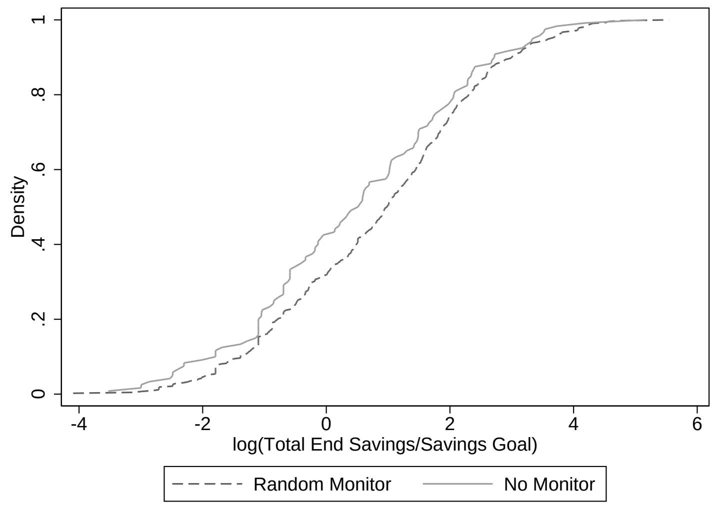
Figure 4. Supplemental Survey Evidence



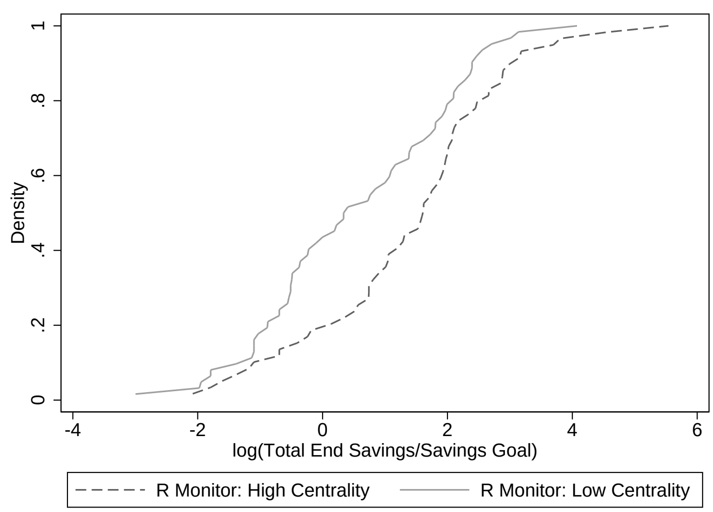
Notes: Surveys conducted with 128 individuals across 8 villages. The villages were all in the study districts and were selected to be comparable to the study villages. Before the surveys were asked, four randomly selected households were selected to conduct the gossip questionnaire from Banerjee et al. (2014). In the questions presented in Panel B, actual names of villagers were given for the Average Monitor and the Central Monitor. The Average Monitor name was selected by visiting houses according to the right-hand rule. The name of the Central monitor was obtained from the gossip questionnaires.

FIGURE 5. Distributions (CDF) of log(Total Savings/Savings goal) by Treatment

Panel A: Non-Monitored Savers vs. Savers with Random Monitors



Panel B: Savers with High Centrality vs. Low Centrality Monitors



Notes: The panels plot the CDFs of $\log(\frac{TotalSavings}{SavingsGoal})$ for different experimental subsamples. In Panel A, we plot the CDFs for the non-monitored savers and the monitored savers, both in random assignment villages. $p = 0.081$ from a Kolmogorov-Smirnov test for the difference in distributions. In Panel B, we plot the CDFs for the monitored savers in random assignment villages with high versus low centrality. Here high centrality is defined as top 15% of monitor centrality and low as bottom 15%. $p = 0.01$ for a Kolmogorov-Smirnov test for the difference in distributions.

Tables

Table 1. Summary Statistics, Treatment Assignment, and Attrition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Treatment (Village Meeting Sample)					Treatment (Endline Sample)				
Dependent Variable	Mean of Non-Monitored Savers	Diff. Random vs. No Monitor	Diff. Endogenous vs. No Monitor	Diff. Endogenous vs. Random	Obs.	Mean of Non-Monitored Savers	Diff. Random vs. No Monitor	Diff. Endogenous vs. No Monitor	Diff. Endogenous vs. Random	Obs.
Age	33.09	-0.147	0.158	0.306	1,307	33.45	0.0414	0.0254	-0.0160	1,146
	(0.385)	(0.458)	(0.528)	(0.458)		(0.387)	(0.454)	(0.551)	(0.468)	
Female	0.756	-0.0411	-0.0253	0.0158	1,307	0.78	-0.0412	-0.0248	0.0164	1,146
	(0.0243)	(0.0316)	(0.0343)	(0.0308)		(0.0261)	(0.0304)	(0.0339)	(0.0292)	
Married	0.857	-0.0287	-0.0244	0.00429	1,307	0.875	-0.0334	-0.0409	-0.00757	1,146
	(0.0192)	(0.0208)	(0.0272)	(0.0253)		(0.0204)	(0.0218)	(0.0268)	(0.0249)	
Widowed	0.0358	0.00954	0.0151	0.00559	1,307	0.033	0.0175	0.0246	0.00712	1,146
	(0.00984)	(0.0126)	(0.0161)	(0.0157)		(0.0101)	(0.0137)	(0.0174)	(0.0174)	
Positive Savings 6 Mos Prior to Baseline	0.717	0.0244	0.0116	-0.0128	1,307	0.725	0.0181	0.0157	-0.00241	1,146
	(0.0319)	(0.0346)	(0.0358)	(0.0420)		(0.0333)	(0.0371)	(0.0370)	-0.0434	
Has Post Office or Bank Acct. at Baseline	0.378	-0.0111	0.0404	0.0515	1,307	0.396	-0.00964	0.0241	0.0337	1,146
	(0.0316)	(0.0362)	(0.0340)	(0.0400)		(0.0329)	(0.0367)	(0.0354)	(0.0407)	
Has BPL Card	0.84	0.0197	-0.00175	-0.0215	1,307	0.842	0.0150	0.0112	-0.00374	1,146
	(0.0211)	(0.0251)	(0.0269)	(0.0288)		(0.0235)	(0.0296)	(0.0280)	(0.0295)	
Savings Goal	1838	-239.1**	131.1	370.1**	1,307	1751	-207.7	185.8	393.5**	1,146
	(117.1)	(117.4)	(165.2)	(178.2)		(126.6)	(117.8)	(166.2)	(180.2)	
Savings Goal (1% outliers trimmed)	1650	-106.5	-55.07	51.41	1,286	1538	-35.75	34.54	70.28	1,127
	(76.04)	(78.99)	(101.0)	(94.23)		(75.69)	(81.40)	(103.4)	(96.58)	
Log Savings Goal	7.253	-0.0631	0.00868	0.0718	1,307	7.209	-0.0350	0.0467	0.0817	1,146
	(0.0398)	(0.0415)	(0.0464)	(0.0533)		(0.0421)	(0.0408)	(0.0476)	(0.0543)	
Saver Centrality	1.700	0.151*	0.255***	0.104	1,307	1.709	0.168**	0.237***	0.0693	1,146
	(0.0578)	(0.0756)	(0.0722)	(0.0940)		(0.0552)	(0.0812)	(0.0686)	(0.0989)	
Saver Centrality (1% outliers trimmed)	1.690	0.113	0.229***	0.116	1,294	1.709	0.112	0.214***	0.102	1,136
	(0.0562)	(0.0715)	(0.0710)	(0.0881)		(0.0552)	(0.0775)	(0.0656)	(0.0930)	
Endline Survey Administered (Non-Attriters)						0.889	-0.0272	-0.00390	0.0233	1,307
No Fixed Effects						(0.0179)	(0.0252)	(0.0219)	(0.0279)	
Endline Survey Administered (Non-Attriters)						0.887	-0.00413	-0.0248	-0.0207	1,307
Village Fixed Effects						(0.0150)	(0.0295)	(0.0254)	(0.0389)	
15-Month Follow-Up Survey Administered (Non-Attriters)						0.893	-0.000259	0.0353	0.0356	1,307
No Fixed Effects						(0.0175)	(0.0243)	(0.0212)	(0.0254)	
15-Month Follow-Up Survey Administered (Non-Attriters)						0.896	0.0139	0.0101	-0.00384	1,307
Village Fixed Effects						(0.0143)	(0.0287)	(0.0232)	(0.0370)	

Notes: Table displays summary statistics and treatment balance of baseline characteristics. Columns 1-5 consider the full sample of savers who opted into the village meeting, while columns 6-10 consider the sample for which we have endline survey responses. Columns 1 and 6 present means (and standard deviations) for the non-monitored savers. Columns 2 and 7 present the differences between random monitor treatment savers with the non-monitored savers. Columns 3 and 8 compare the endogenous monitor treatment with non-monitored savers. Columns 4 and 9 compare monitored savers under random versus endogenous assignment. Standard errors from regressions of baseline characteristic on treatment are in parentheses. Standard errors are clustered at the village level.

Table 2. Effect of Random Monitors on Savings

Dependent Variable	(1) Log Total Savings	(2) Log Total Savings	(3) Log Total Savings
Monitor Treatment: Random Assignment	0.370**	0.284*	0.353**
	(0.146)	(0.162)	(0.138)
Observations	544	544	544
R-squared	0.008	0.125	0.086
Dependent Variable Mean (Omitted Group)	7.647	7.647	7.647
Fixed Effects	None	Village	
Controls	None	Saver	Double-Post LASSO

Notes: Table shows the effects of receiving a randomly allocated monitor on log total savings in the 30 random assignment villages. Total savings is the amount saved across all savings vehicles – the target account and any other account, both formal and informal including money held "under the mattress." Sample constrained to individuals who answered our questionnaire and whose goals were not in the top 1%. Saver controls include the following saver characteristics: savings goal, saver centrality, age, marital status, number of children, preference for bank or post office account, whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. The double-post LASSO specification in column 3 considers all saver controls and individual village fixed effects in the possible control set. Standard errors clustered at the village level.

Table 3. Total Savings by Network Position of Random Monitor

Dependent Variable	(1) Log Total Savings	(2) Log Total Savings	(3) Log Total Savings	(4) Log Total Savings	(5) Log Total Savings	(6) Log Total Savings
Monitor Centrality	0.178**		0.134*		0.153**	
	(0.0736)		(0.0729)		(0.0675)	
Saver-Monitor Proximity		1.032***	0.865**		1.108***	
		(0.352)	(0.334)		(0.294)	
Model-Based Regressor				0.217*		0.289**
				(0.118)		(0.106)
Observations	424	424	424	424	424	424
R-squared	0.150	0.155	0.161	0.147	0.101	0.081
Fixed Effects	Village	Village	Village	Village		
Controls	Saver, Monitor	Saver, Monitor	Saver, Monitor	Saver, Monitor	Double-Post LASSO	Double-Post LASSO

Notes: Table shows impacts on log total savings by monitor network position. Total savings is the amount saved across all savings vehicles – the target account and any other account, both formal and informal including money held "under the mattress." Sample constrained to savers who received a monitor in the 30 random-assignment villages, who answered our questionnaire, and whose goals were not in the top 1%. The variable "Model-Based Regressor" is defined as q_{ij} in the framework. Saver and Monitor controls include savings goal and saver centrality, along with the following variables for each monitor and saver: age, marital status, number of children, preference for bank or post office account (saver only), whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Saver and Monitor controls additionally include the geographical distance between their homes. The double-post LASSO specifications in columns 5-6 consider all saver and monitor controls and individual village fixed effects in the possible control set. Standard errors clustered at the village level.

Table 4. Total Savings by Network Position of Random Monitor: Multigraph Analysis

	(1)	(2)	(3)	(4)
Dependent Variable	Log Total Savings	Log Total Savings	Log Total Savings	Log Total Savings
Monitor Centrality: Full Network	0.180*		0.139	
	(0.100)		(0.0960)	
Monitor Centrality: Financial Network	0.00246		-0.0111	
	(0.154)		(0.146)	
Monitor Centrality: Advice Network	-0.00589		-0.00228	
	(0.121)		(0.112)	
Saver-Monitor Proximity: Full Network		0.739	0.642	
		(0.546)	(0.515)	
Saver-Monitor Proximity: Financial Network		0.273	0.258	
		(0.924)	(0.928)	
Saver-Monitor Proximity: Advice Network		0.236	0.147	
		(0.744)	(0.770)	
Model-Based Regressor: Full Network				0.232
				(0.175)
Model-Based Regressor: Financial Network				-0.00373
				(0.189)
Model-Based Regressor: Advice Network				-0.0132
				(0.162)
Observations	424	424	424	424
R-squared	0.152	0.158	0.164	0.149
Fixed Effects	Village	Village	Village	Village
Controls	Saver, Monitor	Saver, Monitor	Saver, Monitor	Saver, Monitor

Notes: Table shows impacts on log total savings by monitor network position, using different definitions of link-types. Total savings is the amount saved across all savings vehicles – the target account and any other account, both formal and informal including money held "under the mattress." Sample constrained to savers who received a monitor in the 30 random-assignment villages, who answered our questionnaire, and whose goals were not in the top 1%. The variable "Model-Based Regressor" is defined as q_{ij} in the framework. Saver and Monitor controls include savings goal and saver centrality, along with the following variables for each monitor and saver: age, marital status, number of children, preference for bank or post office account (saver only), whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Saver and Monitor controls additionally include the geographical distance between their homes. Standard errors clustered at the village level.

Table 5. Beliefs About Savers and Monitor Centrality

Dependent Variable: Beliefs about Saver	(1) Good at Meeting Goals	(2) Good at Meeting Goals	(3) Good at Meeting Goals	(4) Reached Goal	(5) Reached Goal	(6) Reached Goal
Monitor Centrality	0.0426***	0.0398***	0.0373**	0.0268**	0.0206**	0.0206**
	(0.0137)	(0.0137)	(0.0146)	(0.00976)	(0.00782)	(0.00820)
Respondent-Monitor Proximity	0.0455	0.0142	0.0311	0.00503	-0.00331	-0.00257
	(0.0432)	(0.0378)	(0.0354)	(0.0188)	(0.0196)	(0.0237)
Observations	4,743	4,743	4,743	4,743	4,743	4,743
Dependent Variable Mean	0.240	0.240	0.240	0.061	0.061	0.061
Fixed Effects	No	Village	Respondent	No	Village	Respondent
Controls	Saver, Monitor	Saver, Monitor	Saver, Monitor	Saver, Monitor	Saver, Monitor	Saver, Monitor

Notes: Table explores beliefs of 615 respondents across the 30 random villages, each of whom was asked in the 15-month follow-up survey to rate approximately 8 randomly selected savers who had a monitor from their village. "Good at Meeting Goals" is constructed as 1(Saver reached goal)*1(Respondent indicates saver is good or very good at meeting goals) + (1-1(Saver reached goal))*1(Respondent indicates saver is mediocre, bad or very bad at meeting goals). "Reached Goal" measures whether the saver reached her goal and the respondent correctly believed this to be true. Saver and Monitor controls include savings goal and saver centrality, along with the following variables for each monitor and saver: age, marital status, number of children, preference for bank or post office account (saver only), whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Saver and Monitor controls additionally include the geographical distance between their homes. All specifications include controls for saver and respondent centrality and for the proximity between the respondent and monitor. Columns 2 and 5 include village fixed effects. Columns 3 and 6 include respondent fixed effects. Standard errors clustered at the village level.

Table 6. Shock Mitigation for Monitored Savers in Random Villages

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Shocks								*Final Endline Savings*	
Dependent Variable:	Total Number	Total Number	Above Median	Above Median	Health	Health	HH Expenditure	HH Expenditure	log(Tot. Sav.) 15 mos.	log(Tot. Sav.) 15 mos.
Monitor Treatment: Random Assignment	-0.197	-0.244*	-0.0751*	-0.0924**	-0.0753	-0.102	-0.0468	-0.0641	0.308	0.278
	(0.129)	(0.131)	(0.0417)	(0.0438)	(0.0625)	(0.0674)	(0.0390)	(0.0424)	(0.199)	(0.194)
Observations	1,153	1,153	1,153	1,153	1,153	1,153	1,153	1,153	1,152	1,152
Mean of Dep. Var (Non-Monitored)	1.769	1.769	0.577	0.577	0.862	0.862	0.500	0.500	7.659	7.659
Fixed Effects	Village	No	Village	No	Village	No	Village	No	Village	No
Controls	Saver	Saver	Saver	Saver	Saver	Saver	Saver	Saver	Saver	Saver

Notes: Table presents effects of a monitor on shock mitigation in the 30 random-allocation villages. All outcome variables are measured in the 15 month follow-up survey. The outcome variables in columns 1-8 are all measures of unmitigated shocks experienced by the savers between the end of the six month savings period and the survey. Columns 9-10 report log total savings across all accounts at the time of the survey. The total number of shocks measures the number of unmitigated shocks experienced, including deaths, family illnesses, health shocks, livestock shocks, and unexpected HH expenditures. Sample constrained to all savers in the sample who answered our questionnaire. If a response was missing for a category, the observation is missing in the regression. Controls include the following saver characteristics: log savings goal, saver centrality, age, marital status, number of children, preference for bank or post office account, whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home, type of electrical connection, a dummy for endogenous village (when no village fixed effects) and a dummy for endogenous monitor. Columns 1, 3, 5, 7, and 9 include village fixed effects. Standard errors clustered at the village level.

Table 7. No Evidence of Bunching or Gaming

	(1) Exceeded Payment Threshold	(2) Exceeded Payment Threshold	(3) Exceeded Payment Threshold	(4) Exceeded Payment Threshold
Monitor R x In window of Half Goal	-0.246*	-0.246*	-0.315**	-0.292**
	(0.141)	(0.140)	(0.146)	(0.143)
Monitor E x In window of Half Goal	-0.205	-0.205	-0.250*	-0.238*
	(0.142)	(0.141)	(0.146)	(0.142)
In window of Half Goal	0.889***	0.889***	0.750***	0.750***
	(0.0980)	(0.0972)	(0.109)	(0.109)
Monitor R x In window of Full Goal	0.257	0.514***	0.470***	0.492***
	(0.237)	(0.136)	(0.117)	(0.112)
Monitor E x In window of Full Goal	0.1000	0.286*	0.304**	0.288**
	(0.269)	(0.158)	(0.146)	(0.138)
In window of Full Goal	0.600**	0.286**	0.222**	0.200**
	(0.228)	(0.119)	(0.0980)	(0.0905)
Observations	88	114	174	183
R-squared	0.751	0.693	0.575	0.568
Size of window around Half/Full	± 50/150	± 66/200	± 100/300	± 116.66/350

Notes: Table explores different windows of the half goal and the full goal thresholds by treatment. Column 1 considers widows equal to the size of the monitor's incentive: Rs. 50 around the half goal and Rs. 150 of the full goal. In column 2, these windows are Rs. 66 and Rs. 200, respectively. In column 3, the windows are Rs. 100 and Rs. 300 around the half and full goal thresholds. Finally, in column 4, these windows are Rs. 116.66 and Rs. 350.

Table 8. Random vs. Endogenous Monitors

	(1)	(2)
Dependent Variable	Log Total Savings	Log Total Savings
Monitor Treatment: Random Assignment Village	0.281*	0.288*
	(0.151)	(0.147)
Monitor Treatment: Endogenous Assignment Village	-0.0961	-0.0822
	(0.160)	(0.146)
Non-Monitored Treatment: Endogenous Assignment Village		0.356*
		(0.204)
Observations	1,042	1,042
R-squared	0.127	0.146
Fixed Effects	Village	No
Controls	Saver	Saver

Notes: Table reports effects of receiving a monitor in random versus endogenous allocation villages. Total savings is the amount saved across all savings vehicles – the target account and any other account – by the saver. Sample includes savers who responded to our questionnaire and whose goals were not in the top 1%. Controls include the following saver characteristics: savings goal, saver centrality, age, marital status, number of children, preference for bank or post office account, whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Column 1 includes village fixed effects, while column 2 does not. Standard errors clustered at the village level.

Appendix A. Appendix: Formal Model

A.1. **Description of the signaling environment.** There are n agents and each has privately known type $\theta_i \in \{H, L\}$. We assume each agent's type is drawn, i.i.d., to be $\theta_i = H$ with probability 1/2 and $\theta_i = L$ with probability 1/2. This data generating process is commonly known. A subset $m \ll n$ of agents participate in a signaling game and are called "savers". Another subset of m agents are matched bijectively to the set of savers, and we call them "monitors". The remaining $n - 2m$ agents have no designated role. We assume that whether a given individual participates in the signaling game is private at the outset, so the fact that i is a saver is known only to her monitor j, and vice versa.

The game proceeds as follows. Each saver i decides whether to take a potentially costly action ($s_i = 1$) at cost c_{θ_i} or not ($s_i = 0$) at no cost. We assume high types find the action less costly, $c_H < c_L$. Each agent also has a productivity A_{θ_i} with $A_H > 0 > A_L$. That is, high types are productive whereas low types are not. For simplicity, we assume $c_L > A_H - A_L > c_H$.

When selecting s_i, every saver i knows that her choice will be observed by a unique monitor j. After i chooses s_i, every agent k in the network is informed, independently, about i's decision with probability p_{jk} and receives no information with probability $1 - p_{jk}$. Note that this probability can depend on the identity of the monitor j. Let r_{jk} denote an indicator for whether k received this information.

Finally, each agent i then meets each agent k, independently, with probability p_{ik}. If the meeting happens, agent k then offers agent i a payoff which corresponds to her belief about i's productivity: $\mathrm{E}_k [A_{\theta_i} | s_i, r_k]$. Note that this will depend on whether k as received information about whether i participates in the game as well as the choice if she does participate. We assume that the structure of probabilities $\mathbf{P} = \{p_{kl} : \ k, l \in \{1, ..., n\}\}$ is commonly known by all agents.

The expected payoff of agent i for choice s_i is given by

$$U(s_i) := \sum_k p_{ik} p_{jk} \mathrm{E}_k [A_{\theta_i} | s_i, r_k = 1] + \sum_k (1 - p_{jk}) p_{ik} \mathrm{E}_k [A_{\theta_i} | s_i, r_k = 0] - s_i c_{\theta_i}.$$

It will be useful to define $q_{ij} := \sum_k p_{ik} p_{jk}$. As a normalization, assume the probabilities are such that $q_{ij} < \bar{q} < 1$ for any i, j.

A.2. **Relationship to our experiment.** The mapping to our experiment is as follows. In the first phase, a potential saver decides whether to save a low ($s_i = 0$) or a high ($s_i = 1$) amount. This decision sends a signal to the monitor as to whether the saver is responsible or not. The type θ_i represents responsibility. The idea is that it is relatively costlier for irresponsible individuals to overcome their time inconsistency, temptations or inattention and accrue high savings. Single crossing means more responsible people are better able to overcome their time-inconsistency, given by $c_H < c_L$.

In the second phase, the saver has a future interaction with a fellow community member from the village network. The saver again meets a community member through the graph. The returns to this interaction can depend on whether this community member knows about the saver's "type" via the signaling process in period 1. This is because responsible people are more productive: so $A_H > A_L$. If the member of the community believes individual is irresponsible, she offers the saver a lower return and the saver has less to gain in the second period since she receives the low wage. Otherwise, if the member believes that she is responsible, she offers the saver the high wage. However, it is possible that the community member simply has not heard any rumor about the individual's type whatsoever, in which case the saver receives a pooled wage, which we normalize to $\frac{A_H+A_L}{2}$. In a typical signaling model, the costly signal is always transmitted to the market. Here the signal is more likely to be transmitted to someone who will give the saver a payoff if the saver is more likely to meet this person and this person is more likely to have heard directly/indirectly about the information via the monitor, captured by the $p_{ik}p_{jk}$ term. Below, we describe a model based on the social network that provides a concrete specification for p_{ik} and p_{jk}.

A.3. **The Network Environment and Interpretation of q_{ij}.**

A.3.1. *The network environment.* In order to model the network environment, it is necessary to first define what we mean by a network interaction. Our perspective is informed by our data. A link between households in our data captures whether respondents indicate in a survey a strong social or financial relationship. Surely in village communities, any two arbitrary households interact on occasion, even in absence of a direct link in our data. For instance, one may gossip with someone who is merely an acquaintance at the local tea shop, one may learn of a job opportunity indirectly through a friend's relative, etc. Therefore, we interpret the network as a medium through which we can parametrize interactions; an individual is more likely to pass information to or meet with direct contacts, is less likely to pass information to or meet friends of friends, and is even less likely to interact with friends of friends of friends, and so on.

In our model, agents interact in an undirected, unweighted graph with associated adjacency matrix $\mathbf{A}$. Each element $\mathbf{A}_{ij}$ is a binary indicator for a strong social or financial relationship between nodes i and j. The network mediates two types of interactions. First, an agent can pass information to another agent. We suppose that this happens stochastically, with information traveling from node i to j (or from j to i) when $\mathbf{A}_{ij} = 1$ with some fixed probability θ. Second, agents may meet others. We assume that every node i travels to a neighboring node j where $\mathbf{A}_{ij} = 1$ with probability θ, to a neighbor's neighbor (a node k with $\mathbf{A}^2_{ij} = 1$) with probability θ^2 (if there is only one such path there), and so on. The model is parsimonious, with both types of interactions essentially depending on the single parameter θ.

We use $p_{ij}(\mathbf{A}, \theta)$ to denote the probability that nodes i and j interact in a particular stage of the game. We further subdivide each stage of the model into T sub-periods, which allows for information to diffuse through the network. All information passing (and meetings) along the network occur in the following manner. Given $\mathbf{A}$, there is some probability θ, that a given piece of information crosses any given link ij each sub-period, t.[48] Let us define

$$p_{ij}(\mathbf{A}, \theta) \propto \left[\sum_{t=1}^{T} (\theta \mathbf{A})^t\right]_{ij}.$$

Observe that the right-hand side counts the expected number of times a piece of information starting from node i hits node j and takes into account the potentially numerous paths information may take between i and j. The constant of proportionality ensures that this term is a probability bounded by 1, but is not otherwise relevant for the model. Let $\mathbf{P}$ denote the full matrix with entries p_{ij}. This formulation equips us with expressions for the key probabilities in the signaling model: p_{jk}, the probability that monitor j transmits his observation of i's savings to third party k, and p_{ik}, the probability that saver i encounters third party k for a payoff.

Given this framework for interactions on a network, observe that certain households will be more central than others (reaching directly or indirectly more individuals). As will become clear, this has nothing to do with the strategic interactions themselves but rather only with the assumed physical interactions on the network.

It is useful to formally define

$$DC(\mathbf{A}, \theta) = \sum_{t=1}^{T} (\theta \mathbf{A})^t \cdot 1$$

as the *diffusion centrality* with T sub-periods of communication. This object is a vector where $DC_i(\mathbf{A}, \theta)$ gives the expected number of times information starting from a given node i hits all other nodes in the graph. Note that this also – equivalently – gives the expected number of times that i interacts in total with all other nodes over T periods. This is the notion of centrality that emerges from our simple model of interaction on a network. Note that $DC(\mathbf{A}, \theta)$ is related to the commonly-studied eigenvector centrality in the following way. Let λ_1 be the first (maximal) eigenvalue corresponding to the matrix $\mathbf{A}$ and let $e(\mathbf{A})$ be the corresponding eigenvector. Taking the limit as $T \to \infty$ with $\theta \geq \frac{1}{\lambda_1}$ leads to a vector $\lim_{T\to\infty} \sum_{t=1}^{T} (\theta \mathbf{A})^t \cdot 1 \propto e(\mathbf{A})$, the eigenvector centrality.[49]

We also consider measures of social proximity between nodes i and j in the graph. Note that if two agents are closer in the graph, the rows of $\mathbf{P}$ corresponding to those agents must be more correlated. This is because if i and j are neighbors, any path to a given k of length

[48] Assume that $\theta \geq \frac{1}{\lambda_1(\mathbf{A})}$ where $\lambda_1(\mathbf{A})$ is the maximal eigenvalue of $\mathbf{A}$.

[49] This is the same modeling structure used in Banerjee et al. (2013). For a more general discussion about eigenvector centrality in network economic models, see Jackson (2008). See also DeMarzo et al. (2003),Golub and Jackson (2012), Golub and Jackson (2010), and Hagen and Kahng (1992).

ℓ from i to k must be either of length $\ell+1$, ℓ or $\ell-1$ from j to k. This notion of proximity, $\operatorname{cov}(p_{i\cdot}, p_{j\cdot})$, is key for our proofs, below. Note additionally, that this measure is correlated with the more standard inverse-distance measure used in the network literature, $1/d(i,j)$, where $d(i,j)$ is the shortest path between i and j.

This certainly is not the only sensible way to model interactions, and different models would generate predictions for slightly different notions of centrality. However, the core idea would be the same. The key point is that once equipped with a simple framework describing how agents in the society interact, it sheds light on why we may be prone to see differences across treatments based on the network position of the parties.

A.3.2. *Computing q_{ij}.* We now show that we can write q_{ij} in terms of the centralities and proximity, as defined above.[50]

$$\begin{aligned} q_{ij} = \sum_k p_{ik}(\mathbf{A},\theta)\cdot p_{jk}(\mathbf{A},\theta) &= n\cdot \operatorname{cov}(p_{i\cdot}, p_{j\cdot}) + \frac{1}{n}\sum_k p_{jk}\sum_k p_{ik} \\ &= n\cdot \operatorname{cov}(p_{i\cdot}, p_{j\cdot}) + \frac{1}{n} DC_j(\mathbf{A},\theta)\cdot DC_i(\mathbf{A},\theta). \end{aligned}$$

A.4. **Analysis.** Let us define $\widehat{q} := \frac{c_H}{A_H - A_L}$. By assumption, note that $\widehat{q} < 1$. Let r_k denote a dummy variable for whether a third party k in the network hears a report about i (that i was a saver, the amount s_i, and the identity of monitor j).

Lemma A.1. *Under the maintained assumptions, there is a Perfect Bayesian equilibrium such that*

(1) *for H types,*
$$s_i = \begin{cases} 1 & \textit{if } q_{ij} \geq \widehat{q} \\ 0 & \textit{otherwise.} \end{cases}$$
(2) *for L types, $s_i = 0$ irrespective of q_{ij}.*

This PBE is supported by beliefs by each third party k,

- $\mathrm{P}(\theta_i = H | r_k = 0) = \frac{1}{2}$,
- $\mathrm{P}(\theta_i = H | s_i = 1, r_k = 1, q_{ij} \geq \widehat{q}) = 1$ *and* $\mathrm{P}(\theta_i = H | s_i = 0, r_k = 1, q_{ij} \geq \widehat{q}) = 0$,
- $\mathrm{P}(\theta_i = H | s_i = 1, r_k = 1, q_{ij} < \widehat{q}) = x$ *and* $\mathrm{P}(\theta_i = H | s_i = 0, r_k = 1, q_{ij} < \widehat{q}) = \frac{1}{2}$, *for any $x \in [0,1]$.*

Proof. Consider a saver of type θ_i. She chooses $s_i = 1$ if and only if

$$\begin{aligned} \sum_k p_{ik}p_{jk}\mathrm{E}_k[A_{\theta_i}|s_i=1,r_k=1] + \sum_k (1-p_{jk})\,p_{ik}\mathrm{E}_k[A_{\theta_i}|s_i=1,r_k=0] > \sum_k p_{ik}p_{jk}\mathrm{E}_k[A_{\theta_i}|s_i=0,r_k=1] \\ + \sum_k (1-p_{jk})\,p_{ik}\mathrm{E}_k[A_{\theta_i}|s_i=0,r_k=0]. \end{aligned}$$

[50] In this derivation we ignore the constant of proportionality (or assume that it is 1) for parsimony. This has no consequence for the result.

Since $\mathrm{E}_k\left[A_{\theta_i}|s_i=1,r_k=0\right]=\mathrm{E}_k\left[A_{\theta_i}|s_i=0,r_k=0\right]$, she chooses $s_i=1$ if and only if

$$\sum_k p_{ik}p_{jk}\mathrm{E}_k\left[A_{\theta_i}|s_i=1,r_k=1\right]-c_{\theta_i}>\sum_k p_{ik}p_{jk}\mathrm{E}_k\left[A_{\theta_i}|s_i=0,r_k=1\right].$$

Let $\mu_k\left(s_i\right):=\mathrm{P}\left(\theta_i=H|s_i,r_k=1\right)$ be defined as the posterior that k has about i's type given she has received a report and observes s_i. Notice that because network structure is common knowledge and all third parties k hold the same prior about θ_i, we can write $\mathrm{E}_k\left[A_{\theta_i}|s_i,r_k\right]=\mathrm{E}\left[A_{\theta_i}|s_i,r_k\right]$ and $\mu_k\left(s\right)=\mu\left(s\right)$ for any k with $r_k=1$.

We can then write

$$\mathrm{E}\left[A_{\theta_i}|s_i,r_k=1\right]=\mu\left(s_i\right)\left(A_H-A_L\right)+A_L \text{ and } \mathrm{E}\left[A_{\theta_i}|s_i,r_k=0\right]=\frac{A_H+A_L}{2}.$$

From the above, $s_i=1$ if and only if

$$\left(\mu\left(1\right)-\mu\left(0\right)\right)\left(A_H-A_L\right)\sum_k p_{ik}p_{jk}\geq c_{\theta_i},$$

which can be written, since $q_{ij}=\sum_k p_{ik}p_{jk}$, as $q_{ij}\geq\frac{c_{\theta_i}}{\Delta\mu\Delta A_\theta}$.

In this equilibrium, if $q_{ij}>\widehat{q}$, then $\mu\left(s_i=1\right)=1$ and $\mu\left(s_i=0\right)=0$. (Recall that network structure is common knowledge, so q_{ij} is known to k when making this calculation.) Observe that

$$\frac{c_L}{\left(A_H-A_L\right)\left(1-0\right)}>\bar{q}>q_{ij}\geq\widehat{q}=\frac{c_H}{\left(A_H-A_L\right)\left(1-0\right)}.$$

Therefore, no L-type finds it worthwhile to try to secure a reputation gain by investing in $s_i=1$, because c_L is too high.

Meanwhile, if $q_{ij}<\widehat{q}$, then

$$q_{ij}<\widehat{q}=\frac{c_H}{\left(A_H-A_L\right)\left(1-0\right)}<\frac{c_L}{A_H-A_L}$$

and therefore neither type finds it worthwhile, even for the maximal reputation gain, $(1-0)$, to have $s_i=1$. In this case, $\mu\left(s_i=0\right)=\frac{1}{2}$ and $\mu\left(s_i=1\right)$ can take any value since even the maximal increase in reputation (1-0) would not make it worthwhile. □

This result immediately implies the following.

Proposition A.2. *Under the maintained assumptions* $\mathrm{P}\left(s_i=1|q_{ij}\right)$ *is a (weakly) monotonically increasing function in* q_{ij}*. Consequently,* $\mathrm{P}\left(s_i=1|q_{ij}\right)$ *must be (weakly) monotonically increasing in both social proximity,* $\mathrm{cov}\left(p_{i\cdot},p_{j\cdot}\right)$*, and monitor centrality,* DC_j *.*

Proof. By random assignment of monitors j to savers i, and by orthogonality of private type θ_i to network position, it follows that half of those with a sufficiently high q_{ij} must be of low type, so $\mathrm{P}\left(s_i=1|q_{ij}\right)=\frac{1}{2}$ if $q_{ij}\geq\widehat{q}$. Meanwhile, if $q_{ij}<\widehat{q}$, $\mathrm{P}\left(s_i|q_{ij}\right)=0$. As shown above, since

$$q_{ij}=n\cdot\mathrm{cov}\left(p_{i\cdot},p_{j\cdot}\right)+\frac{1}{n}DC_j\left(\mathbf{A},\theta\right)\cdot DC_i\left(\mathbf{A},\theta\right),$$

ceteris paribus an increase in either $\mathrm{cov}\left(p_{i\cdot},p_{j\cdot}\right)$ or $DC_j\left(\mathbf{A},\theta\right)$ increases q_{ij}. □

A.5. **Interpretation of Results.** Our framework suggests that we should focus our empirical analysis on two network features: centrality, in particular eigenvector centrality which follows directly from the model, and proximity. We have the following predictions: (1) as q_{ij} increases, a greater proportion of savers should be saving high amounts; (2) as monitor centrality increases, a greater proportion of savers should be saving high amounts; (3) as saver-monitor proximity increases, a greater proportion of savers should be saving high amounts. These directly motivate regressions of savings on network position, as conducted in the paper. Figure A.1 presents an example where $\operatorname{cov}(p_{i\cdot}, p_{j\cdot})$ is varied between saver i and monitor j but DC_j is held fixed. This is to give an idea of how to envision holding distance fixed as we vary centrality, or vice versa.

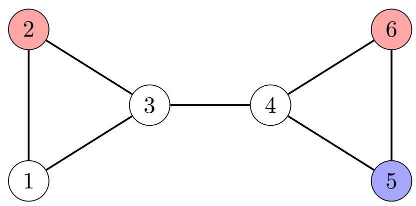


FIGURE A.1. Let node 5 be the saver and let nodes 2 and 6 be potential monitors. This presents a situation where $DC_2 = DC_6$, by symmetry, but clearly $\operatorname{cov}(p_{5\cdot}, p_{2\cdot}) \neq \operatorname{cov}(p_{5\cdot}, p_{6\cdot})$.

A reasonable question to raise is whether individuals already know each others' types, especially those who are socially close. We think that there is significant scope for learning about even a close individual's type for several reasons. The first piece of evidence comes from our data. 15 months after our intervention, individuals were asked to rate 12 random subjects about whether the subjects reached their goals as well as answer several questions concerning their level of responsibility. The respondents were no more likely to rate their unmonitored friends (who reached their goal throughout the experiment) as responsible as more distant individuals despite there being a positive correlation on average between responsibility and goal reaching. If anything, they were slightly worse at rating their friends. Second, the work of Alatas et al. (2012) examines how well individuals are able to rank others' wealth in their communities. While individuals are slightly better at ranking those to whom they are socially closer, the error rates are still very high indicating highly imperfect local information. Third, we have anecdotal evidence from our subjects that indicate that there is scope to build reputation among even their friends, neighbors or important individuals in their communities. Thus, while it is entirely possible ex ante for the scope for reputation building to be lower among the socially proximate (due to heterogeneous priors), our own prior is that this is unlikely to be the case.

ONLINE APPENDIX: NOT FOR PUBLICATION

Appendix B. Supplemental Appendix: Quotes

"For those who want to save in a bank or post office account but do not have the habit of doing so, having a monitor may help... Having a more important person as a monitor may help in comparison to a person who is not well known by people in the village. A person may save more if it is an important person knowing they might get more benefits from this person later on." – Subject 1

"If the monitor was a very important person in the village, and the saver did not meet a goal that she set, the monitor would lose trust in the saver. The monitor will feel that if in the future he or his friends gives her some job or tasks or responabilities, the saver may not fulfill them." – Subject 2

"When paired with an important person, they will save more to build the monitor's confidence in them. That way the person builds trust with me [sic]... If the person does not fulfill savings, the monitor will be disappointed and think 'I used to place trust in that person but now I can't'. They would speak less to the saver and feel 'cheated to trust' [sic]. They may tell others... But if someone is too irresponsible then monitor or no monitor, the saver will not save." – Subject 3

"People will only reach their goals if their monitors are family, friends, neighbors, or important people." - Subject 4

"I would like to choose the important person except if there are close friends. Then I may hesitate if I do not know him well." – Subject 5

Appendix C. Selection into Saver Sample

Table C.1. Determinants of Participation in Savings Program: Potential Savers

Outcome: Participates in Village Meeting	(1) Uni-Variate Regressions	(2) Multi-Variate Regression
Age	0.000701	-0.00188
	(0.00145)	(0.00165)
Female	0.157***	0.124***
	(0.0235)	(0.0272)
Married	0.0613**	-0.0215
	(0.0271)	(0.0352)
Widowed	0.0261	-0.0330
	(0.0490)	(0.0640)
Number of Children	0.0293***	0.00723
	(0.0106)	(0.0134)
Eigenvector Centrality	0.240	0.317
	(0.293)	(0.289)
Saving Goal	-1.42e-05***	-8.73e-06***
	(3.26e-06)	(2.72e-06)
Log Saving Goal	-0.0883***	
	(0.0182)	
Had Non-Zero Savings in Prior 6 Months	0.0660**	0.0551*
	(0.0266)	(0.0303)
Saves at Bimonthly Frequency or Higher	0.114***	0.0434
	(0.0217)	(0.0272)
Already Has a Bank Account	-0.0353	-0.0265
	(0.0253)	(0.0238)
Prefers a Bank to a Post Office Account	0.00203	0.00987
	(0.0253)	(0.0243)
Daily Wage Laborer	0.0694***	0.0535**
	(0.0235)	(0.0229)
Saving Purpose: Children	0.0154	0.0222
	(0.0266)	(0.0345)
Saving Purpose: Household Expenses	0.0206	0.0398
	(0.0244)	(0.0365)
Saving Purpose: Emergency Fund	-0.00673	0.0175
	(0.0268)	(0.0358)
Overall Fraction Participating in Village Meeting	57.10%	57.10%
Observations	2,288	2,288

Notes: Table presents differences in characteristics of individuals who participated in the village meeting, thus becoming savers in the experiment, with individuals who were given the opportunity to attend, but who did not attend. Variables in the table come from the baseline survey administered with all potential savers. Each row in the table corresponds to a different uni-variate regression. Standard errors clustered at the village level.

Table C.2. Determinants of Participation in Savings Program: Full Village

	Selection into Saver Sample	
Dependent Variable	Mean of Non-Participant HHs	Diff. Non-Participants vs. Savers
HH Size	1.902	0.0822**
	(0.0536)	(0.0384)
Max Education in HH	7.784	0.158
	(0.253)	(0.17)
Any HH Member Speaks English	0.0968	-0.0207**
	(0.00776)	(0.00846)
HH has BPL Card	0.777	0.0764***
	(0.0151)	(0.0141)
HH has TV	0.825	0.0391**
	(0.013)	(0.0184
HH Participates in SHG or RoSCA	0.392	0.0955***
	(0.0246)	(0.0196)
HH has Any Formal Account	0.739	0.0445***
	(0.0156)	(0.015)
Primary Occupation of at least one HH Member		
Land Owner	0.298	-0.0173
	(0.0189)	(0.015)
Agricultural Laborer	0.317	0.0624***
	(0.0131)	(0.0171)
Dairy and Animal Husbandry	0.0876	0.00621
	(0.00766)	(0.00908)
Non-Agricultural Laborer	0.101	0.000522
	(0.0138)	(0.0111)
Small Business Owner	0.098	-0.00259
	(0.00821)	(0.0108)
Government Worker	0.0281	-0.00886*
	(0.00256)	(0.00445)

Notes: Table presents differences in characteristics of households who participated in the village meeting, thus becoming savers in the experiment, with the full set of non-participant households in the village (who did not attend). Variables in the table come from the census survey conducted alongside the network elicitation by Banerjee et al. (2013). Each row in the table corresponds to a different uni-variate regression. Standard errors clustered at the village level. N=11,531.

Appendix D. Final Endline Supplemental Tables and Figures

Table D.1. Network Summary Statistics: Village-Level

Summary Statistics: Sample Villages	Obs.	Mean	Std. Dev.
Number of Households	60	222.12	65.85
Average Degree	60	17.57	3.96
Average Clustering	60	0.30	0.05
Average Path Length	60	2.34	0.19

Notes: Table presents selected village-level network characteristics for the study area.

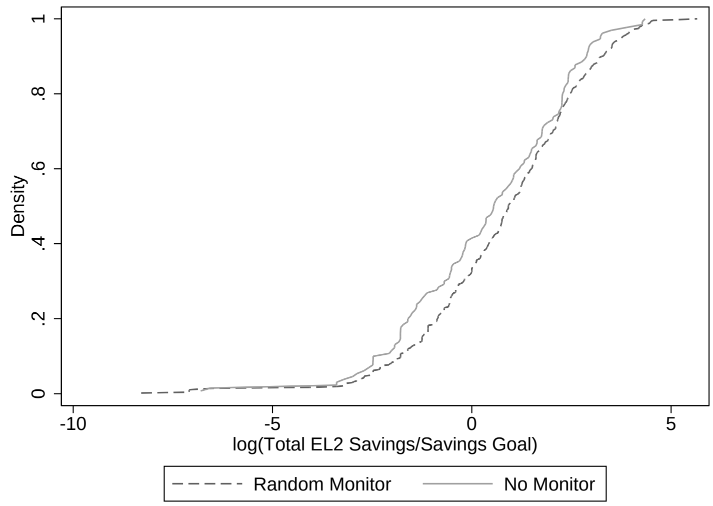
Table D.2. Endline Survey Summary Statistics: Non-Monitored Savers

Summary Statistics: Non-Monitored Savers, R Villages	Obs.	Mean	Std. Dev.
Endline Survey: Conclusion of Intervention			
Total Savings	123	8,890	17,616
Log Total Savings	123	7.67	1.83
Total Expenditures (past month), 1% winsorized	123	7,561	8,683
Total Expenditures (past month), 5% winsorized	123	6,517	3,943
Log Total Expenditures (past month)	133	8.66	0.78
Expenditure Categories (past month):			
Festivals	133	825	1,335
Pan	133	196	220
Tea	133	275	228
Meals Away	133	255	476
Eggs and Meat	133	598	781
Other Food	133	1,527	1,348
Transport	133	636	1,059
Entertainment and Phone	133	245	214
Final Endline Survey: 15 Months Following Conclusion			
Total Savings	133	9,263	16,125
Log Total Savings	133	7.65	2.08
How the Savers Saved:			
Increased Labor Supply	117	0.15	0.36
Business Profits	117	0.03	0.18
Cut Unnecessary Expenditures	117	0.15	0.35
Money from Spouse, Family, and Friends	117	0.19	0.39
Reduced Transfers to Others	117	0.01	0.09
Took a Loan	117	0.04	0.20
Shocks			
Total Number of Shocks	133	1.77	1.43
Greater than Median Number of Shocks	133	0.58	0.50
Health Shock Indicator	133	0.86	0.66
HH Expenditure Shock Indicator	133	0.50	0.50
Beliefs about Non-Monitored Savers in R Villages			
Reached Goal	2141	0.03	0.18
Good at Meeting Goals	2141	0.21	0.41

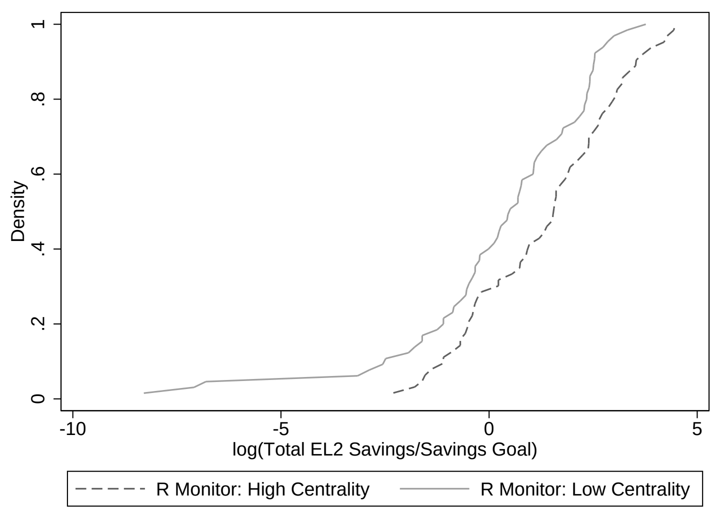
Notes: Table presents summary statistics for variables measured in the second endline for the non-monitored savers in random assignment villages.

FIGURE D.1. Distributions (CDF) of log(Total Savings/Savings goal) by Treatment

Panel A: Non-Monitored Savers vs. Savers with Random Monitors



Panel B: Savers with High Centrality vs. Low Centrality Monitors



Notes: The panels plot the CDFs of $\log(\frac{TotalSavings}{SavingsGoal})$ for different experimental subsamples 15 months after the end of the experiment. In Panel A, we plot the CDFs for the non-monitored savers and the monitored savers, both in random assignment villages. $p = 0.168$ from a Kolmogorov-Smirnov test for the difference in distributions. In Panel B, we plot the CDFs for the monitored savers in random assignment villages with high versus low centrality. Here high centrality is defined as top 15% of centrality and low as bottom 15%. $p = 0.061$ from a Kolmogorov-Smirnov test for the difference in distributions.

Appendix E. Reached Goal Outcomes

Table E.1. Effect of Random Monitors on Goal Attainment

Dependent Variable	(1) Reached Goal	(2) Reached Goal
Monitor Treatment: Random Assignment	0.0662*	0.0630*
	-0.0325	-0.0316
Observations	673	673
R-squared	0.007	0.021
Dependent Variable Mean (Omitted Group)	0.072	0.072
Fixed Effects	None	Village
Controls	None	Saver

Notes: Table shows the effects of receiving a randomly allocated monitor on goal attainment in the 30 random assignment villages. Reached Goal is a dummy for whether the saver (weakly) exceeded her savings goal. Sample constrained to individuals whose goals were not in the top 1%. Saver controls include the following saver characteristics: savings goal, saver centrality, age, marital status, number of children, preference for bank or post office account, whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Standard errors clustered at the village level.

Table E.2. Goal Attainment Network Position of Random Monitor

Dependent Variable	(1) Reached Goal	(2) Reached Goal	(3) Reached Goal	(4) Reached Goal
Monitor Centrality	0.0339**		0.0288*	
	(0.0156)		(0.0162)	
Saver-Monitor Proximity		0.147**	0.118	
		(0.0698)	(0.0718)	
Model-Based Regressor				0.0518**
				(0.0204)
Observations	523	523	523	523
R-squared	0.048	0.046	0.053	0.050
Fixed Effects	Village	Village	Village	Village
Controls	Saver, Monitor	Saver, Monitor	Saver, Monitor	Saver, Monitor

Notes: Table shows impacts on savings goal attainment by monitor network position. Reached Goal is a dummy for whether the saver (weakly) exceeded her savings goal. Sample constrained to savers who received a monitor in the 30 random-assignment villages and whose goals were not in the top 1%. The variable "Model-Based Regressor" is defined as q_{ij} in the framework. Saver and Monitor controls include savings goal and saver centrality, along with the following variables for each monitor and saver: age, marital status, number of children, preference for bank or post office account (saver only), whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Saver and Monitor controls additionally include the geographical distance between their homes. Standard errors clustered at the village level.

Table E.3. Random vs. Endogenous Monitors

	(1)	(2)
Dependent Variable	Reached Goal	Reached Goal
Monitor Treatment: Random Assignment Village	0.0607*	0.0626*
	(0.0311)	(0.0321)
Monitor Treatment: Endogenous Assignment Village	0.0597**	0.0640***
	(0.0229)	(0.0218)
Non-Monitored Treatment: Endogenous Assignment Village		-0.00417
		(0.0335)
Observations	1,277	1,277
R-squared	0.022	0.024
Fixed Effects	Village	No
Controls	Saver	Saver

Notes: Table reports effects of receiving a monitor in random versus endogenous allocation villages. Reached Goal is a dummy for whether the saver (weakly) exceeded her savings goal. Sample includes savers whose savings goals were not in the top 1%. Controls include the following saver characteristics: savings goal, saver centrality, age, marital status, number of children, preference for bank or post office account, whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Column 1 includes village fixed effects, while column 2 does not. Standard errors clustered at the village level.

Appendix F. How did the Savers Save?

Table F.1. How Did the Savers Save?

Panel A: Expenditures During Month 6 of Savings Period

Dependent Variable	(1) Log Expenditure	(2) Expenditure (1% Win.)	(3) Expenditure (5% Win.)	(4) Festivals	(5) Pan	(6) Tea	(7) Meals Away	(8) Eggs and Meat	(9) Other Food	(10) Transport
Monitor Treatment: Random Assignment	-0.0978	-924.6	-560.1*	-237.8*	16.58	35.55*	16.85	-73.90	-186.1	-156.9*
	(0.0627)	(812.1)	(294.6)	(133.2)	(27.21)	(17.97)	(38.65)	(57.61)	(131.0)	(82.46)
Observations	1,119	1,119	1,119	1,119	1,119	1,119	1,119	1,119	1,119	1,119
Number of village	0.041	0.022	0.047	0.029	0.023	0.006	0.041	0.037	0.072	0.049
Mean of Dep. Var (Non-Monitored)	8.62	7561	6517	825	198	277	259	607	1,527	641
Fixed Effects	Village	Village	Village	Village	Village	Village	Village	Village	Village	Village
Controls	Saver	Saver	Saver	Saver	Saver	Saver	Saver	Saver	Saver	Saver

Panel B: Retrospective Assessment from Follow-Up Survey

Dependent Variable	(1) Increased Labor Supply	(2) Business Profits	(3) Cut Unnecessary Expenditures	(4) Money from Spouse, Family, Friends	(5) Reduced Transfers to Others	(6) Took a Loan
Monitor Treatment: Random Assignment	0.0704**	0.0204	0.0777*	-0.0253	0.0152	-0.0224
	(0.0334)	(0.0160)	(0.0422)	(0.0351)	(0.0120)	(0.0188)
Observations	1,026	1,026	1,026	1,026	1,026	1,026
R-squared	0.055	0.026	0.020	0.057	0.016	0.014
Mean of Dep. Var (Non-Monitored)	0.15	0.03	0.15	0.19	0.01	0.04
Fixed Effects	Village	Village	Village	Village	Village	Village
Controls	Saver	Saver	Saver	Saver	Saver	Saver

Notes: Panel A measures the effect of receiving a randomly assigned monitor on selected measures of expenditures in the sixth month of the savings period measured at the end of the monitoring intervention. Panel B reports survey responses from the 15 month follow-survey. Sample constrained to all savers in the sample who answered our questionnaire and whose savings goals were not in the top 1%. Controls include the following saver characteristics: savings goal, saver centrality, age, marital status, number of children, preference for bank or post office account, whether the individual has a bank or post office account at baseline, caste, elite status, forecasted expenditure at baseline, number of rooms in the home, type of electrical connection and an indicator for endogenous monitor. All regressions include village fixed effects. Standard errors clustered at the village level.

Table F.2. Treatment Effects in Target and Non-Target Savings Vehicles

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Level				Log			
Dependent Variable: Savings	Target Account	Target Account	Non-Target Account	Non-Target Account	Target Account	Target Account	Non-Target Account	Non-Target Account
Monitor Treatment: Random Assignment	328.7*	314.4*	2,637	2,523	0.885**	0.899**	0.250	0.229
	(167.9)	(169.8)	(2,188)	(2,180)	(0.344)	(0.341)	(0.162)	(0.155)
Observations	540	540	540	540	540	540	540	540
R-squared	0.003	0.022	0.001	0.003	0.013	0.120	0.003	0.089
Dependent Variable Mean (Omitted Group)	369.7	369.7	8,279	8,279	1.776	1.776	7.559	7.559
Fixed Effects	None		None		None		None	
Controls	None	Double-Post LASSO	None	Double-Post LASSO	None	Double-Post LASSO	None	Double-Post LASSO

Notes: Table shows effects of receiving a random monitor on savings both in the target account and in non-target accounts. Sample restricted to only those individuals responding to the savings questions in the endline survey and whose savings goals were not in the top 1%. Odd columns include no controls, while even columns perform double-post LASSO using village fixed effects and the full set of saver controls. In columns 5 and 6 the outcome is log(target account savings + 1). Non target account savings includes both formal and informal sources. Standard errors are clustered at the village level.

Appendix G. Lee Bounds

Table G.1. Predictors of Attrition

	(1) Endline Participation
Dependent Variable	
Female	0.112***
	(0.0316)
Constant	0.784***
	(0.0265)
Observations	682
R-squared	0.022

Notes: Table shows that females are more likely to participate in the endline survey than males. Sample restricted to random assignment villages. Standard errors are clustered at the village level.

Table G.2. Main Analysis with Lee Bounds

Panel A: Savers in villages with random monitor assignment

Dependent Variable: Log Total Savings	(1) Raw Regression	(2) Lower and Upper Bounds
Treatment: Monitor with Random Assignment		
Estimate	0.370	[0.237, 0.496]
Confidence Interval: [5%, 95%]	[0.073, .0668]	[-0.036, 0.774]
Confidence Interval: [10%, 90%]	[0.123, 0.618]	[0.022, 0.774]

Panel B: Savers with randomly assigned monitor

Dependent Variable: Log Total Savings	(1) Raw Regression	(2) Lower and Upper Bounds
Treatment Variable: High Model-Based Regressor (25th percentile)		
Estimate	0.451	[0.306, 0.782]
Confidence Interval: [5%, 95%]	[0.009, 0.894]	[-0.086, 1.107]
Confidence Interval: [10%, 90%]	[0.083, 0.819]	[0.015, 1.050]

Notes: Panels A and B show the main results (Tables 2 and 3, respectively) with Lee Bounds for endline survey attrition. The sample is restricted to the 30 random assignment villages. Because gender predicts attrition, a dummy for female is used to tighten the bounds. The Lee bounds methodology requires a binary treatment variable. Thus, in Panel B, we consider receiving a monitor in top 25% of realizations of the model-based regressor as the treatment indicator. Confidence intervals for the bounds calculated using 500 bootstrap iterations on the upper and lower bounds, with clustering at the village level.

Appendix H. Supplemental Exercises: Beliefs

Table H.1. Beliefs About Savers and Monitor Centrality: Alternate Controls

Dependent Variable: Beliefs about Saver	(1) Good at Meeting Goals	(2) Good at Meeting Goals	(3) Good at Meeting Goals	(4) Reached Goal	(5) Reached Goal	(6) Reached Goal
Monitor Centrality	0.0389**	0.0374**	0.0353**	0.0206**	0.0157*	0.0157*
	(0.0144)	(0.0140)	(0.0148)	(0.00937)	(0.00804)	(0.00854)
Respondent-Monitor Proximity	0.0476	0.0181	0.0360	0.00357	-0.00252	-0.00160
	(0.0422)	(0.0366)	(0.0342)	(0.0194)	(0.0196)	(0.0239)
Observations	4,743	4,743	4,743	4,743	4,743	4,743
Dependent Variable Mean	0.240	0.240	0.240	0.061	0.061	0.061
Fixed Effects	No	Village	Respondent	No	Village	Respondent
Controls	Saver	Saver	Saver	Saver	Saver	Saver

Notes: Table explores beliefs of 615 respondents across the 30 random villages, each of whom was asked in the 15-month follow-up survey to rate approximately 8 randomly selected savers who had a monitor from their village. "Good at Meeting Goals" is constructed as 1(Saver reached goal)*1(Respondent indicates saver is good or very good at meeting goals) + (1-1(Saver reached goal))*1(Respondent indicates saver is mediocre, bad or very bad at meeting goals). "Reached Goal" measures whether the saver reached her goal and the respondent correctly believed this to be true. Controls include the following saver characteristics: savings goal, saver centrality, age, marital status, number of children, preference for bank or post office account, whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Columns 2 and 5 include village fixed effects. Columns 3 and 6 include respondent fixed effects. Standard errors clustered at the village level.

Table H.2. Beliefs About Savers and Random Monitor Assignment

Dependent Variable: Beliefs about Saver	(1) Good at Meeting Goals	(2) Good at Meeting Goals	(3) Good at Meeting Goals	(4) Reached Goal	(5) Reached Goal	(6) Reached Goal
Monitor Treatment: Random Assignment	0.0241	0.0315	0.0308	0.0252	0.0277	0.0279
	(0.0240)	(0.0250)	(0.0266)	(0.0200)	(0.0225)	(0.0240)
Observations	6,894	6,894	6,894	6,894	6,894	6,894
Dependent Variable Mean (Omitted Group)	0.213	0.213	0.213	0.035	0.035	0.035
Fixed Effects	No	Village	Respondent	No	Village	Respondent
Controls	Saver	Saver	Saver	Saver	Saver	Saver

Notes: Table explores beliefs of 615 respondents across the 30 random villages, each of whom was asked in the 15-month follow-up survey to rate approximately 11 savers enrolled in the study. "Good at Meeting Goals" is constructed as 1(Saver reached goal)*1(Respondent indicates saver is good or very good at meeting goals) + (1-1(Saver reached goal))*1(Respondent indicates saver is mediocre, bad or very bad at meeting goals). "Reached Goal" measures whether the saver reached her goal and the respondent correctly believed this to be true. Controls include the following saver characteristics: savings goal, saver centrality, age, marital status, number of children, preference for bank or post office account, whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Columns 2 and 5 include village fixed effects. Columns 3 and 6 include respondent fixed effects. Standard errors clustered at the village level.

Appendix I. Relation to Prior Work

It is true that this is not the first project to take place in a subset of the villages where detailed network data was collected by Banerjee et al. (2013). Here we discuss the other experiments that took place in these villages prior to our intervention and whether we should worry that those projects somehow bias or alter the interpretation of our results.

The first project was that of Banerjee et al. (2013) itself, which traced the diffusion of microfinance through the networks of 43 of the 75 study villages. We should first note that for all of our analysis, we use the second wave of network data, completed in 2012 after the introduction of microfinance in those 43 villages, through 2010. Thus any rewiring of the network that was a result of the intervention is already captured in our baseline network measures. The 60 villages in this study overlap with 36 of the villages that were "treated" with microfinance. Row 1 of Panel A of Table I.1 shows that this treatment is roughly balanced across Endogenous (17 out of 30 villages) and Random (19 out of 30 villages) selection. We are not very worried that the act of bringing microfinance to these villages alters the interpretation of the savings study in any way. Microfinance is a very common borrowing vehicle across rural India, and the partner microfinance institution is only one of many operating in the region. Moreover, the marketing and spread of microfinance was conducted entirely by the microfinance institution with no aid from the research team, and thus, there is no reason that households should draw any connection between microfinance and the enumerators.

The network villages were also the home to a series of laboratory-in-the-field projects that occurred after the rollout of microfinance. Each of these projects had a similar implementation format. Several days before the experimental lab session, members of the survey team would enter the village to seek permission from representatives of local government and would typically inform a set of randomly selected households of the upcoming session. The fact that participants could earn up to a full day's wage from participating was made salient at the time. On the day of the experimental session, the research team would arrive in the village an hour before the session was meant to begin. Some participants would typically already be gathered at the appointed location. The research team then walked through the village to gather the required number of participants for each session. Once the set of participants was assembled, the lab exercises then began and typically lasted approximately three hours. Average compensation was approximately half a day's wage.

Households could participate in up to three of these sessions between 2009 and 2010. The first set of laboratory sessions was conducted during the summer of 2009 and included the experimental protocols of Chandrasekhar et al. (2013) and pilots for Chandrasekhar et al. (2012). The second set of sessions included the experimental protocols of Chandrasekhar et al. (2011) and pilots for Breza et al. (2015) and was conducted during the summer of 2010. Finally, the main experiment for Breza et al. (2015) was rolled out during the summer and fall of 2010. Given the contained nature of each of these experiments, it is unlikely that they should interact in any way with the much more involved savings intervention. Nevertheless, we check for treatment balance and whether controls for past participation alter our results in any way.

Table I.1 also shows treatment balance by prior participation in the laboratory sessions. 29 of the study villages had participated in at least one of these three lab experimental sessions, with Endogenous selection villages exhibiting lower rates of prior participation in Panel A. Panel B further explores treatment balance at the saver level. 23% of households that enrolled in the savings study as savers had a household member who participated in at least one of the prior lab experiments. However, savers receiving a monitor in Random assignment villages are no more or less likely to have participated in the past. We do, however, detect treatment imbalance in receiving a monitor in Endogenous assignment villages.

We next check the robustness of our main results by including controls for prior household participation. Tables I.2, I.3, and I.4 present versions of Tables 2, 3, and 8, respectively, adding indicators for past participation by a household member in one, two, or three laboratory sessions. All of the key coefficients are very close in magnitude and of similar statistical significance to the main specifications in the body of paper. This gives us confidence that the prior laboratory experiments did not interact in a meaningful way with the savings intervention.

Table I.1. Treatments and Prior Experiments: Balance

Panel A: Village-Level Balance

	Treatment		
Dependent Variable	Mean of Random Selection Vills	Diff. Endogenous vs. Random	Obs.
Microfinance Entered Village	0.633	-0.0667	60
	(0.0895)	(0.128)	
Any Prior Lab Experiment	0.500	-0.200	60
	(0.0928)	(0.126)	
Number Prior Lab Experiment	1.167	-0.367	60
	(0.204)	(0.286)	

Panel B: Household-Level Balance

	Treatment			
Dependent Variable	Mean of Non-Monitored Savers	Diff. Random vs. No Monitor	Diff. Endogenous vs. No Monitor	Obs.
Any Prior Lab Experiment Participation	0.270	-0.0133	-0.0580*	1,307
No Fixed Effects	(0.0349)	(0.0365)	(0.0328)	
Any Prior Lab Experiment Participation	0.268	0.00607	-0.0726**	1,307
Village Fixed Effects	(0.0208)	(0.0424)	(0.0325)	
Number Prior Lab Experiment Participation	0.313	-0.00836	0.0147	1,307
No Fixed Effects	(0.0435)	(0.0447)	(0.0492)	
Number Prior Lab Experiment Participation	0.313	0.0147	-0.0898*	1,307
Village Fixed Effects	(0.0259)	(0.0492)	(0.0459)	

Notes: Table shows baseline balance by exposure to prior experiments. In each panel, the first column shows means and standard deviations in parentheses for the subgroup. Treatment differences are taken from regressions of each dependent variable on treatment indicators, with standard errors in parentheses. Panel A shows village-level characteristics, while Panel B shows saver-level characteristics.

Table I.2. Treatments and Prior Experiments: Receipt of a Random Monitor

Dependent Variable	(1) Log Total Savings	(2) Log Total Savings	(3) Log Total Savings	(4) Log Total Savings
Monitor Treatment: Random Assignment	0.361**	0.281*	0.359**	0.278
	(0.151)	(0.163)	(0.152)	(0.166)
Observations	544	544	544	544
R-squared	0.015	0.127	0.015	0.129
Dependent Variable Mean (Omitted Group)	8.647	9.647	7.647	7.647
Prior Experience Controls	Linear Controls	Linear Controls	Fixed Effects	Fixed Effects
Other Fixed Effects	None	Village	None	Village
Other Controls	None	Saver	None	Saver

Notes: Table shows the effects of receiving a randomly allocated monitor on log total savings in the 30 random assignment villages. Total savings is the amount saved across all savings vehicles – the target account and any other account, both formal and informal including money held "under the mattress." Sample constrained to individuals who answered our questionnaire and whose goals were not in the top 1%. Saver controls include the following saver characteristics: savings goal, saver centrality, age, marital status, number of children, preference for bank or post office account, whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Table also controls for prior participation in previous laboratory sessions. Columns 1 and 2 use linear controls in the number of prior sessions, while columns 3 and 4 include a saturated set of prior participation indicators. Standard errors clustered at the village level.

Table I.3. Treatments and Prior Experiments: Network Position of Random Monitor

	(1)	(2)	(3)	(4)
Dependent Variable	Log Total Savings	Log Total Savings	Log Total Savings	Log Total Savings
Monitor Centrality	0.173**		0.130*	
	(0.0737)		(0.0731)	
Saver-Monitor Proximity		1.021***	0.858**	
		(0.359)	(0.342)	
Model-Based Regressor				0.205*
				(0.120)
Observations	424	424	424	424
R-squared	0.154	0.159	0.165	0.151
Prior Experience Controls	Fixed Effects	Fixed Effects	Fixed Effects	Fixed Effects
Other Fixed Effects	Village	Village	Village	Village
Other Controls	Saver, Monitor	Saver, Monitor	Saver, Monitor	Saver, Monitor

Notes: Table shows impacts on log total savings by monitor network position. Total savings is the amount saved across all savings vehicles – the target account and any other account, both formal and informal including money held "under the mattress." Sample constrained to savers who received a monitor in the 30 random-assignment villages, who answered our questionnaire, and whose goals were not in the top 1%. The variable "Model-Based Regressor" is defined as q_{ij} in the framework. Saver and Monitor controls include savings goal and saver centrality, along with the following variables for each monitor and saver: age, marital status, number of children, preference for bank or post office account (saver only), whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Saver and Monitor controls additionally include the geographical distance between their homes. Table also controls for prior participation in previous laboratory sessions. All columns include a saturated set of prior participation indicators. Standard errors clustered at the village level.

Table I.4. Treatments and Prior Experiments: Effects of Random and Endogenous Monitors

	(1)	(2)
Dependent Variable	Log Total Savings	Log Total Savings
Monitor Treatment: Random Assignment Village	0.277*	0.288*
	(0.154)	(0.150)
Monitor Treatment: Endogenous Assignment Village	-0.0815	-0.0657
	(0.162)	(0.148)
Non-Monitored Treatment: Endogenous Assignment Village		0.352*
		(0.207)
Observations	1,042	1,042
Prior Experience Controls	Fixed Effects	Fixed Effects
Other Fixed Effects	Village	No
Other Controls	Saver	Saver

Notes: Table reports effects of receiving a monitor in random versus endogenous allocation villages. Total savings is the amount saved across all savings vehicles – the target account and any other account – by the saver. Sample includes savers who responded to our questionnaire and whose goals were not in the top 1%. Controls include the following saver characteristics: savings goal, saver centrality, age, marital status, number of children, preference for bank or post office account, whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Column 1 includes village fixed effects, while column 2 does not. Standard errors clustered at the village level.

Appendix J. Altonji-Type Tests

Table J.1. Total Savings by Network Position of Random Monitor: No Controls

	(1)	(2)	(3)	(4)
Dependent Variable	Log Total Savings	Log Total Savings	Log Total Savings	Log Total Savings
Monitor Centrality	0.183***		0.127*	
	(0.0665)		(0.0680)	
Saver-Monitor Proximity		1.244***	1.067***	
		(0.334)	(0.330)	
Model-Based Regressor				0.254***
				(0.0901)
Observations	424	424	424	424
R-squared	0.014	0.025	0.031	0.023
Fixed Effects	None	None	None	None
Controls	None	None	None	None

Notes: Table shows impacts on log total savings by monitor network position. Total savings is the amount saved across all savings vehicles. Sample constrained to savers who received a monitor in the 30 random-assignment villages, who answered our questionnaire, and whose goals were not in the top 1%. The variable "Model-Based Regressor" is defined as q_{ij} in the framework. Standard errors clustered at the village level.

Table J.2. Total Savings by Network Position of Random Monitor: Saver and Monitor Controls, No Village Fixed Effects

	(1)	(2)	(3)	(4)
Dependent Variable	Log Total Savings	Log Total Savings	Log Total Savings	Log Total Savings
Monitor Centrality	0.195**		0.149**	
	-0.0715		-0.0716	
Saver-Monitor Proximity		1.111***	0.920***	
		-0.335	-0.31	
Model-Based Regressor				0.248**
				-0.109
Observations	424	424	424	424
R-squared	0.167	0.171	0.179	0.166
Fixed Effects	None	None	None	None
Controls	Saver, Monitor	Saver, Monitor	Saver, Monitor	Saver, Monitor

Notes: Table shows impacts on log total savings by monitor network position. Total savings is the amount saved across all savings vehicles. Sample constrained to savers who received a monitor in the 30 random-assignment villages, who answered our questionnaire, and whose goals were not in the top 1%. The variable "Model-Based Regressor" is defined as q_{ij} in the framework. Saver and Monitor controls include savings goal and saver centrality, along with the following variables for each monitor and saver: age, marital status, number of children, preference for bank or post office account (saver only), whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Saver and Monitor controls additionally include the geographical distance between their homes. Standard errors clustered at the village level.

Table J.3. Total Savings by Network Position of Random Monitor: No Geography Controls

	(1)	(2)	(3)	(4)
Dependent Variable	Log Total Savings	Log Total Savings	Log Total Savings	Log Total Savings
Monitor Centrality	0.180**		0.139*	
	(0.0732)		(0.0718)	
Saver-Monitor Proximity		0.991***	0.821**	
		(0.356)	(0.336)	
Model-Based Regressor				0.212*
				(0.121)
Observations	424	424	424	424
R-squared	0.148	0.150	0.158	0.144
Fixed Effects	Village	Village	Village	Village
Controls	Saver, Monitor	Saver, Monitor	Saver, Monitor	Saver, Monitor

Notes: Table shows impacts on log total savings by monitor network position. Total savings is the amount saved across all savings vehicles. Sample constrained to savers who received a monitor in the 30 random-assignment villages, who answered our questionnaire, and whose goals were not in the top 1%. The variable "Model-Based Regressor" is defined as q_{ij} in the framework. Saver and Monitor controls include savings goal and saver centrality, along with the following variables for each monitor and saver: age, marital status, number of children, preference for bank or post office account (saver only), whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Standard errors clustered at the village level.

Table J.4. Total Savings by Network Position of Random Monitor: Multigraph Analysis, No Controls

	(1)	(2)	(3)	(4)
Dependent Variable	Log Total Savings	Log Total Savings	Log Total Savings	Log Total Savings
Monitor Centrality: Full Network	0.175**		0.127	
	(0.0839)		(0.0879)	
Monitor Centrality: Financial Network	0.108		0.0953	
	(0.141)		(0.126)	
Monitor Centrality: Advice Network	-0.112		-0.108	
	(0.137)		(0.128)	
Saver-Monitor Proximity: Full Network		1.150*	1.057*	
		(0.585)	(0.575)	
Saver-Monitor Proximity: Financial Network		0.0991	0.0129	
		(0.852)	(0.838)	
Saver-Monitor Proximity: Advice Network		0.0513	-0.00636	
		(0.656)	(0.667)	
Model-Based Regressor: Full Network				0.248*
				(0.145)
Model-Based Regressor: Financial Network				0.00757
				(0.177)
Model-Based Regressor: Advice Network				0.00167
				(0.153)
Observations	424	424	424	424
R-squared	0.016	0.025	0.032	0.023
Fixed Effects	None	None	None	None
Controls	None	None	None	None

Notes: Table shows impacts on log total savings by monitor network position, using different definitions of link-types. Total savings is the amount saved across all savings vehicles. Sample constrained to savers who received a monitor in the 30 random-assignment villages, who answered our questionnaire, and whose goals were not in the top 1%. The variable "Model-Based Regressor" is defined as q_{ij} in the framework. Standard errors clustered at the village level.

Appendix K. Main Results, No Goal Trimming

Table K.1. Effect of Random Monitors on Savings: No Goal Trimming

	(1)	(2)	(3)
Dependent Variable	Log Total Savings	Log Total Savings	Log Total Savings
Monitor Treatment: Random Assignment	0.358**	0.279*	0.329**
	(0.150)	(0.161)	(0.142)
Observations	549	549	549
R-squared	0.008	0.131	0.086
Dependent Variable Mean (Omitted Group)	7.647	7.647	7.647
Fixed Effects	None	Village	
Controls	None	Saver	Double-Post LASSO

Notes: Table shows the effects of receiving a randomly allocated monitor on log total savings in the 30 random assignment villages. Total savings is the amount saved across all savings vehicles – the target account and any other account, both formal and informal including money held “under the mattress.” Sample constrained to individuals who answered our questionnaire. Saver controls include the following saver characteristics: savings goal, saver centrality, age, marital status, number of children, preference for bank or post office account, whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Standard errors clustered at the village level.

TABLE K.2. Total Savings by Network Position of Random Monitor: No Goal Trimming

Dependent Variable	(1) Log Total Savings	(2) Log Total Savings	(3) Log Total Savings	(4) Log Total Savings
Monitor Centrality	0.180**		0.136*	
	(0.0732)		(0.0725)	
Saver-Monitor Proximity		1.049***	0.879**	
		(0.349)	(0.330)	
Model-Based Regressor				0.219*
				(0.118)
Observations	426	426	426	426
R-squared	0.154	0.158	0.165	0.151
Fixed Effects	Village	Village	Village	Village
Controls	Saver, Monitor	Saver, Monitor	Saver, Monitor	Saver, Monitor

Notes: Table shows impacts on log total savings by monitor network position. Total savings is the amount saved across all savings vehicles. Sample constrained to savers who received a monitor in the 30 random-assignment villages and who answered our questionnaire. The variable "Model-Based Regressor" is defined as q_{ij} in the framework. Saver and Monitor controls include savings goal and saver centrality, along with the following variables for each monitor and saver: age, marital status, number of children, preference for bank or post office account (saver only), whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Saver and Monitor controls additionally include the geographical distance between their homes. Standard errors clustered at the village level.

Appendix L. Supplemental Multigraph Analysis

Table L.1. Reached Goal by Network Position of Random Monitor: Multi-graph Analysis

	(1)	(2)	(3)	(4)
Dependent Variable	Reached Goal	Reached Goal	Reached Goal	Reached Goal
Monitor Centrality: Full Network	0.0698**		0.0656**	
	(0.0276)		(0.0265)	
Monitor Centrality: Financial Network	-0.0501		-0.0523	
	(0.0367)		(0.0370)	
Monitor Centrality: Advice Network	0.000760		0.00115	
	(0.0291)		(0.0286)	
Saver-Monitor Proximity: Full Network		0.193	0.172	
		(0.115)	(0.116)	
Saver-Monitor Proximity: Financial Network		-0.0868	-0.0300	
		(0.117)	(0.130)	
Saver-Monitor Proximity: Advice Network		0.0274	-0.0166	
		(0.179)	(0.183)	
Model-Based Regressor: Full Network				0.0682**
				(0.0301)
Model-Based Regressor: Financial Network				-0.00250
				(0.0324)
Model-Based Regressor: Advice Network				-0.0216
				(0.0340)
Observations	523	523	523	523
R-squared	0.071	0.059	0.079	0.062
Fixed Effects	Village	Village	Village	Village
Controls	Saver, Monitor	Saver, Monitor	Saver, Monitor	Saver, Monitor

Notes: Table shows impacts of goal attainment by monitor network position, using different definitions of link-types. Reached Goal is a dummy for whether the saver (weakly) exceeded her savings goal. Sample constrained to savers who received a monitor in the 30 random-assignment villages and whose goals were not in the top 1%. The variable "Model-Based Regressor" is defined as q_{ij} in the framework. Saver and Monitor controls include savings goal and saver centrality, along with the following variables for each monitor and saver: age, marital status, number of children, preference for bank or post office account (saver only), whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. Saver and Monitor controls additionally include the geographical distance between their homes. Standard errors clustered at the village level.

Table L.2. Random Monitor Analysis: Direct Financial Relationships

	(1)	(2)
Dependent Variable	Log Total Savings	Log Total Savings
Saver and Monitor Direct Friends: Any Relationship	0.510**	0.472*
	(0.218)	(0.262)
Saver and Monitor Direct Friends: Borrowing or Lending Relationship		0.136
		(0.447)
Observations	424	424
R-squared	0.149	0.150
Fixed Effects	Village	Village
Controls	Saver, Monitor	Saver, Monitor

Notes: Total savings is the amount saved across all savings vehicles – the target account and any other account, both formal and informal including money held "under the mattress" – by the saver. We define a link as having a financial component if the nodes report borrowing or lending small amounts of money or material goods to one another. In our sample, 27% of direct links have a financial component. Sample constrained to individuals who answered our questionnaire and who set a savings goal at baseline of less than Rs. 10,000. Controls include savings goal, saver centrality, and the following variables for each monitor and saver: age, marital status, number of children, preference for bank or post office account (saver only), whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. We also control for the geographical distance between the homes of the saver and monitor. All regressions include village fixed effects. Standard errors clustered at the village level.

Appendix M. Monitor Spillovers

M.1. **Measuring Monitor Spillovers.** Here, we use our experimental variation in monitor assignment in the random villages to look for spillovers from monitored to non-monitored savers. Non-monitored savers in both random- and endogenous-selection villages may be affected if their friends receive monitors and may experience larger spillovers if those monitors are especially effective.[51] The random variation in both the assignment of savers to treatment groups and of monitors to savers in the random selection villages allows us to measure such causal spillover effects.

We use the following regression specification to explore spillovers onto non-monitored savers in the villages with random monitor selection:

$$y_{ir} = \alpha_r + \beta_1 \sum_j A_{ij,r} SM_j + \gamma \sum_j A_{ij,r} AttSaver_j + \delta' X_{ir} + \epsilon_{ir}. \tag{M.1}$$

This estimating equation allows the savings of non-monitored individuals to depend on having more friends randomly assigned to receive a monitor (SM), and β_1 is the coefficient of interest. All of this is conditional on the number of friends participating as savers and monitors in the experiment and a third degree polynomial for the number of each individual's friends. The standard set of controls is included in X.[52]

We can also augment Equation M.1 to analyze any impacts of the centrality of friends' monitors on savings:

$$\begin{aligned} y_{ir} = {} & \alpha_r + \beta_1 \sum_j A_{ij,r} SM_j + \beta_2 \sum_j A_{ij,r} SM_j MC_j + \gamma \sum_j A_{ij,r} AttSaver_j \\ & + \delta' X_{ir} + \epsilon_{ir}. \end{aligned} \tag{M.2}$$

where SM*MC measures the sum of the centralities of the monitor's of individual i's friends.

[51] This could happen for a variety of reasons. For instance, "keeping up with the Joneses", increased motivation to save, receiving reminders from the friend's monitor, an overhearing/participating in more conversations about savings in general, etc. Our model in Section A abstracts from this to focus on the signaling story, just like it abstracts from the direct value of savings itself.

[52] Recall that the standard controls include savings goal, gender, age, marital status, widow status, caste, elite status, material measures of wealth, whether the saver had a pre-existing bank or PO account, preference for bank or PO account during the savings period, and village fixed effects. We also control for the eigenvector centrality of the saver.

Table M.1. Spillovers from Monitored Savers: Non-Monitored Sample

	(1)	(2)
Dependent Variable	Log Total Savings	Log Total Savings
Number of Friends Assigned a Monitor	0.342*	0.187
	(0.189)	(0.216)
Sum of Normalized Centralities of Friends' Monitors		0.123
		(0.127)
Observations	123	123
R-squared	0.310	0.318
Fixed Effects	Village	Village

Notes: Table looks for spillovers onto non-monitored savers through friends being assigned a monitor. Total savings is the amount saved across all savings vehicles – the target account and any other account – by the saver. Sample is restricted to the non-monitored savers in the 30 villages who responded to our questionnaire. Controls include the following saver characteristics: saver centrality, log savings goal, age, marital status, number of children, preference for bank or post office account, whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. All regressions include village fixed effects. Regressions also include the following 1% winsorized network controls: third order polynomials in saver degree, number of saver friends, and number of potential and attending monitor friends. Standard errors clustered at the village level.

Table M.1 presents the results estimating Equations M.1 and M.2. We find in Column 1 that individuals save 0.34 log points more (p-value 0.08) when a friend is assigned a monitor. In Column 2, it appears that this effect is stronger when those monitors are more central, but the effects are not statistically significant. We take the results in Table M.1 as suggestive evidence of spillovers from the monitoring treatments. These spillovers likely bundle many different channels of influence including (but not limited to) "keeping up with the Joneses", increased motivation to save, receiving reminders from the friend's monitor, and overhearing more conversations about savings.[53] Anecdotal evidence suggests that conversations between savers and monitors tend to take place in public and are likely to be overheard by the saver's friends. We do not attempt to decompose these multiple possible effects. We should also note that the average complier for this spillover analysis will be different than the full sample.

M.2. **Treatment Effect Robustness in the Presence of Spillovers.** More generally, all agents – un-monitored and monitored – may face spillover effects from the monitoring treatments. While a saver's own treatment assignment is orthogonal to the treatment assignments of her friends, one might still worry that the peer effects could contaminate our main results presented in the body of the paper. We present a robustness exercise in Appendix Table M.2 where we run our main specifications from Tables 2 and 3, but

[53]This multitude of channels is also the reason why we do not try to estimate and instrument a more structured model of spillovers in the spirit of Bramoulle et al. (2009).

include the treatment status of the saver's friends along with the centralities of the friends' monitors.

Table M.2. Monitor Treatment Effects: Robustness to Inclusion of Peer Spillovers

Panel A: Effects of Random Monitors

Dependent Variable	(1) Log Total Savings	(2) Log Total Savings
Monitor Treatment: Random Assignment	0.329**	0.280*
	(0.148)	(0.165)
Observations	544	544
R-squared	0.044	0.129
Dependent Variable Mean (Omitted Group)	7.647	7.647
Fixed Effects	None	Village
Controls (non-network)	None	Saver

Panel B: Effects of Monitors by Network Heterogeneity

Dependent Variable	(1) Log Total Savings	(2) Log Total Savings	(3) Log Total Savings	(4) Log Total Savings
Monitor Centrality	0.167**		0.133*	
	(0.0721)		(0.0707)	
Saver-Monitor Proximity		0.889**	0.723**	
		(0.362)	(0.340)	
Model-Based Regressor				0.191*
				(0.111)
Observations	424	424	424	424
R-squared	0.162	0.163	0.169	0.158

Notes: Tables show robustness of main results to allowing for spillovers through monitored friends. Total savings is the amount saved across all savings vehicles – the target account and any other account, both formal and informal including money held "under the mattress" – by the saver. Panel A sample constrained to all savers in random monitor villages. Panel B sample constrained to savers who received a monitor in those villages. All included savers answered our questionnaire. Controls include log savings goal, saver centrality, and the following 1% winsorized network controls: third order polynomials in saver degree, number of saver friends, and number of potential and attending monitor friends. The Panel A saver controls include age, marital status, number of children, preference for bank or post office account, whether the individual has a bank or post office account at baseline, caste, elite status, number of rooms in the home and type of electrical connection. All specifications in Panel B contain the saver controls and controls for the same variables but for the monitors (where applicable). All regressions include village fixed effects. Standard errors clustered at the village level.

In Panel A, we show that we can closely replicate the treatment effects of receiving a randomly-assigned monitor (2). Table even when we allow for peer effects. Similarly, the results in Panel B considering the network heterogeneity of the monitor look similar to those in Table 3. Again, given the orthogonality of treatment across individuals, this result is not surprising.

Appendix N. Endogenous Monitors and Choice

N.1. **Model Extension: Selection and Heterogeneity.** The core model presented in Section 3 and Appendix A was developed to study the random monitor assignment treatment and develop a vocabulary for how we should think about network position affecting the signaling game. Here we extend the model to incorporate both the choice of the monitor in the endogenous treatment and entry into our experiment. We simplify algebra by modeling both savers and monitors as only having high or low centrality (which is an aesthetic, but not substantive choice) and there being just one third party k. In addition to illustrating the complexity of thinking about choice in our setting – that certain savers may prefer central monitors while others will not – the goal of the model is to help us think through which types of savers will pick which monitors, who might benefit most from the choice, and more generally, for which patterns to look in the data. We are setting aside a number of real-world issues that may affect the monitor choices of savers in the experiment: for example, there could be other unobserved dimensions of heterogeneity (how nice or forgiving a person is) that may make some potential monitors more attractive than others. We do not claim nor is it our aspiration to fully explain choice in our study.

The model works as follows. Potential savers are either H or L types, where the cost of saving $s_i = 1$ is c_{θ_i} and as before $c_H < c_L$. Potential savers also vary in their centrality, they can be of high or low centrality. In this way, a potential saver decides to join our experiment, knowing that she may be randomly assigned to have no monitor, have a random monitor (in a random treatment village), or have the opportunity to select a monitor via random serial dictatorship (in an endogenous treatment village). The potential subjects realize that having a more central monitor means information can spread more, reaping rewards or costs, depending on their actions.

In our equilibrium, H types always choose to participate: if they receive a high centrality monitor in the random treatment, they save the high amount ($s_i = 1$), and the low amount otherwise ($s_i = 0$). In the endogenous treatment, H types have an incentive to choose a high centrality monitor if one is available to maximize the dissemination of the signal.

On the other hand, L types face a more delicate decision and one that depends on whether the person is of high or low centrality herself. Participating in the experiment has benefits because in the BC treatment, subjects receive in-kind services and a bank account. However, there is the potential cost of receiving monitors and signaling that they are low types. In this case, a high centrality L-type opts not to enter; because of her centrality, in the monitored treatments, she is likely to run into a third party in the future who has heard about her low savings amount (which she would be incentivized to do), averaged across random assignment treatment and endogenous treatment where she would pick a low centrality monitor. And therefore, it won't be worthwhile to participate. On the other hand, the low centrality L faces a similar cost, but one that is lower because of her lower centrality. Therefore, she is willing to participate, saves s_L in the random monitor treatment regardless of the monitor, and picks a low centrality monitor and saves in the endogenous treatment, effectively minimizing the degree to which information about her is every spread.

Equilibrium beliefs calculated by Bayes' rule support this equilibrium, and because our sample is small relative to population, it is easy to see that if a third party never receives a report about a given person, then their posterior remains the prior (1/2) that the person is of high type.

N.1.1. *Population.* There are four types of potential savers, denoted by $\eta_i = (\theta_i, p_{ik})$. Let $\theta_i \in \{L, H\}$ denote the quality of the savers. As in the body of the paper L-types face higher costs ($c_L > c_H$) of saving $s_i = 1$ (i.e., overcoming their time inconsistency, devoting attention to saving). The type θ_i determines a productivity A_{θ_i}, which is the output that this person will produce if hired for a task/project in the future. Let $p_{ik} \in \left\{\bar{p}_{ik}, \underline{p}_{ik}\right\}$ denote the centrality of the savers. For simplicity we assume this to be just binary. We assume these features are independent and uniform in the population, so (θ, p_{ik}) has a population share of $\frac{1}{4}$ for every type combination.

There are two types of potential monitors, denoted by $p_{jk} \in \left\{\bar{p}_{jk}, \underline{p}_{jk}\right\}$ for high or low centrality monitors. We assume again that $\frac{1}{2}$ the population of monitors are $\bar{p}_{jk}$.

N.1.2. *Timing.* In every village:

- Phase 1: the savings experiment
  - Each village has M people, of whom $N \ll M$ are given the opportunity to participate.
  - N potential savers decide whether or not to participate in the experiment resulting in $n \leq N$ savers participating. Let $x \in \{0, 1\}$ denote the participation decision.
  - Those who enter are randomly assigned to treatments: BC, Random monitor, or Endogenous monitor, where the latter two are village-assignments.
  - Monitor assignments are realized.
    - In random villages, $m = \alpha n$, for $\alpha \in (0, 1)$, savers are randomly assigned one-to-one to m monitors.
    - In endogenous villages, m savers pick their monitors via random serial dictatorship.
    - In both types of villages $(1 - \alpha) n$ savers are assigned to the BC treatment.
  - Savers decide how much to save $s_i \in \{0, 1\}$.
    - It costs an agent of type η_i, $c(\theta_i, p_{ik}) = c_{\theta_i}$ to save $s_i = 1$ with $c_H < c_L$.
- Phase 2: future interactions in the village
  - The saver interacts with an individual k(with a probability p_{ik}). This individual has either heard or not heard (denote hearing by $r \in \{1, 0\}$) of the saver's choice

of savings and this happens with a probability that depends on the position of the monitor and this third party individual. The probability that the saver meets this third party who has heard of her savings is given by

$$f\left(p_{ik}, p_{jk}\right) = p_{ik} p_{jk}$$

which depends on both saver centrality and monitor centrality.[54] We assume $\bar{p}_{jk}/\underline{p}_{jk} > \bar{p}_{ik}/\underline{p}_{ik}$, which means that having a more central monitor affects the spread more than being more central, which makes sense because words move faster than meetings. Note that in our base model we simplified this by having them be equal, but that was only because we chose the same parameter to model information flow and meetings.

– This individual offers to pay the saver for a task where the output is the saver's productivity, but of course the saver's type θ is unobserved by this individual. This individual may have heard of the saver's choice of s, if the saver chose to participate and information was transmitted from the monitor to this person, and can make inferences accordingly.

To sum up, relative to the main model, this model adds an entry decision and a monitor choice decision. Again, for algebraic transparency, we allow for only two levels of centrality.

N.1.3. *Payoffs and Participation decision.* Payoffs are as follows:

- By not entering the experiment, the agent has some autarky payoff $v_{aut} < 0$. The negative value captures the absence of the basic account opening services, reminders and small payment made in the account offered in our BC treatment, which will be normalized to 0.
- Individuals encountered in the future can offer agents projects with payoffs which depend on productivities and beliefs about type given what the individual observes.
- The BC treatment generates payoff $\pi^{BC} = 0$. This is just a normalization and note that by entering the experiment all treatments provide this payoff plus or minus the potential wage earnings in Phase 2.

  Note that the payoff to an agent from interacting with an uninformed individual is equivalent to the payoff from not receiving a monitor, $\pi^{BC} = 0$. This comes from the fact that we assume that individuals did not discuss the participation choices of invited individuals, but only the savings progress of those who did participate. This is consistent with equilibrium beliefs provided our assumption above that $M \gg N$. It is easy to check that $\mathrm{P}\left(\theta = H | r = 0, x = 1, s\right) = \frac{1}{2} + O\left(\frac{N}{M}\right)$, which can be made

[54]This is for simplicity. In the body of the paper note $f\left(p_{ik}, p_{jk}\right)$ is

$$\sum_k p_{ik} p_{jk}.$$

arbitrarily close to $\frac{1}{2}$. We also should note that in practice, invitations to participate were made privately.

- A saver in the random treatment receives equilibrium expected payoff $\pi^R(\eta)$.
- A saver in the endogenous treatment receives equilibrium expected payoff $\pi^E(\eta)$.

An agent of type η chooses to enter if and only if

$$\frac{\alpha}{2}\pi^R(\theta, x) + \frac{\alpha}{2}\pi^E(\theta, x) > v_{aut}.$$

N.1.4. *An SPE.* It is useful to define

$$\psi^R := \frac{2\bar{p}_{ik}\underline{p}_{jk} + 2\underline{p}_{ik}\underline{p}_{jk}}{\bar{p}_{ik}\bar{p}_{jk} + \underline{p}_{ik}\bar{p}_{jk} + 3\bar{p}_{ik}\underline{p}_{jk} + 3\underline{p}_{ik}\underline{p}_{jk}}$$

and

$$\psi^E := \frac{\bar{p}_{ik}\underline{p}_{jk} + \underline{p}_{ik}\underline{p}_{jk}}{\bar{p}_{ik}\bar{p}_{jk} + 5\underline{p}_{ik}\underline{p}_{jk}}.$$

These terms, which depend only on the probabilities of someone in the future meeting a saver of low or high centrality, will reflect equilibrium beliefs about a saver being a high type when the individual observes s_L savings in a village of treatment R or E.

We make the following high-level assumptions on parameters to obtain our equilibrium. Feasible parameters satisfy these conditions. We discuss interpretation of these conditions below as they arise.

Assumptions:

(1) $A_H\psi + A_L(1-\psi) < 0$ for $\psi \in \left\{\psi^R, \psi^E\right\}$.

(2) $\frac{c_H}{(1-\psi)\underline{p}_{ik}\bar{p}_{jk}} < A_H - A_L < \min\left\{\frac{c_L}{(1-\psi)\bar{p}_{ik}\bar{p}_{jk}}, \frac{c_H}{(1-\psi)\bar{p}_{ik}\underline{p}_{jk}}\right\}$ for $\psi \in \left\{\psi^R, \psi^E\right\}$..

(3) $\frac{\underline{p}_{ik}\bar{p}_{jk}+\underline{p}_{ik}\underline{p}_{jk}}{2}\left(A_H\psi^R + A_L\left(1-\psi^R\right)\right)+\underline{p}_{ik}\underline{p}_{jk}\left(A_H\psi^E + A_L\left(1-\psi^E\right)\right) > \frac{2v_{aut}}{\alpha} > \frac{\bar{p}_{ik}\bar{p}_{jk}+\bar{p}_{ik}\underline{p}_{jk}}{2}\left(A_H\psi^R + A_L\left(1-\right.\right.$ $\bar{p}_{ik}\underline{p}_{jk}\left(A_H\psi^E + A_L\left(1-\psi^E\right)\right)$.

(4) $5p_{ik}\bar{p}_{jk}A_H + 3p_{ik}\underline{p}_{jk}\left(A_H\psi^R + A_L\left(1-\psi^R\right)\right) > \frac{8}{\alpha}v_{aut} + 5c_H$ for any p_{ik}.

Proposition N.1. *Under the above assumptions there is an SPE in which*

(1) $(H, \bar{p}_{ik})$ *and* $\left(H, \underline{p}_{ik}\right)$ *always enter and*
- *in random villages,*
  - *save* $s_i = 1$ *with* $\bar{p}_{jk}$ *centrality monitors*
  - *save* $s_i = 0$ *with* $\underline{p}_{jk}$ *centrality monitors*
- *and in endogenous choice villages,*
  - *pick* $\bar{p}_{jk}$*-monitor if available and save* s_H
  - *pick* $\underline{p}_{jk}$*-monitor when an* $\bar{p}_{jk}$ *centrality monitor is not available and save* s_L.

(2) (L, h) *never enter.*

(3) (L, l) *always enter and*

- *in random villages save* $s_i = 0$ *with any monitor*
- *in endogenous choice villages pick* $\underline{p}_{jk}$*-monitors and save* $s_i = 0$.

(4) *Any type who enters and is assigned to the BC treatment saves* $s_i = 0$.

Below we compute the beliefs that support this equilibrium and check that it is indeed an SPE. This setup has the following predictions that are consistent with the data:

- Savings should be higher with monitoring in random villages because those in the BC treatment choose $s_i = 0$.
- Savings should be higher with more central monitors in random villages.
- In endogenous villages, having an earlier choice should matter for $\bar{p}_{ik}$ savers but less so for $\underline{p}_{ik}$ savers:
  - For $\bar{p}_{ik}$ savers, if available $\bar{p}_{jk}$ monitors are selected.
  - For $\underline{p}_{ik}$ savers, because the distribution includes $\left(L, \underline{p}_{ik}\right)$ types, there will be $\underline{p}_{jk}$ monitor choices both early and late.

N.1.5. *Random Assignment of Monitors.* We want to compute the belief that the third party has that the saver is of type θ given they have received a report and therefore the saver has participated and has saved an amount s:

$$\mathrm{P}\left(\theta|s, r=1, x=1\right) = \frac{\mathrm{P}\left(s, r=1|\theta, x=1\right)\mathrm{P}\left(\theta|x=1\right)}{\mathrm{P}\left(s, r=1|H, x=1\right)\mathrm{P}\left(H|x=1\right) + \mathrm{P}\left(s, r=1|L, x=1\right)\mathrm{P}\left(L|x=1\right)}.$$

In our equilibrium observe that the following hold:

- Conditional on $\theta = H$:
  - $\mathrm{P}\left(s_i=1, r=1|H, x=1\right) = \frac{\alpha}{4}\left[\bar{p}_{ik}\bar{p}_{jk} + \underline{p}_{ik}\bar{p}_{jk}\right]$
  - $\mathrm{P}\left(s_i=0, r=1|H, x=1\right) = \frac{\alpha}{4}\left[\bar{p}_{ik}\underline{p}_{jk} + \underline{p}_{ik}\underline{p}_{jk}\right]$
- Conditional on $\theta = L$:
  - $\mathrm{P}\left(s_i=1, r=1|L, x=1\right) = 0$
  - $\mathrm{P}\left(s_i=0, r=1|L, x=1\right) = \frac{\alpha}{4}\left[\bar{p}_{ik}\bar{p}_{jk} + \underline{p}_{ik}\bar{p}_{jk} + \bar{p}_{ik}\underline{p}_{jk} + \underline{p}_{ik}\underline{p}_{jk}\right]$
- Type composition given participation:
  - $\mathrm{P}\left(H|x=1\right) = \frac{2}{3}$
  - $\mathrm{P}\left(L|x=1\right) = \frac{1}{3}$

In this case we can compute

- $\mathrm{P}\left(\theta = H|s_i=1, r=1, x=1\right) = \frac{\frac{\alpha}{4}\left[\bar{p}_{ik}\bar{p}_{jk}+\underline{p}_{ik}\bar{p}_{jk}\right]\times\frac{2}{3}}{\frac{\alpha}{4}\left[\bar{p}_{ik}\bar{p}_{jk}+\underline{p}_{ik}\bar{p}_{jk}\right]\times\frac{2}{3}} = 1,$

- $\mathrm{P}(\theta = H | s_i = 0, r = 1, x = 1) = \frac{2\bar{p}_{ik}\underline{p}_{jk} + 2\underline{p}_{ik}\underline{p}_{jk}}{\bar{p}_{ik}\bar{p}_{jk} + \underline{p}_{ik}\bar{p}_{jk} + 3\bar{p}_{ik}\underline{p}_{jk} + 3\underline{p}_{ik}\underline{p}_{jk}}$, as

$$\mathrm{P}(\theta = H | s_i = 0, r = 1, x = 1) = \frac{\frac{\alpha}{4}\left[\bar{p}_{ik}\underline{p}_{jk} + \underline{p}_{ik}\underline{p}_{jk}\right] \times \frac{2}{3}}{\frac{\alpha}{4}\left[\bar{p}_{ik}\underline{p}_{jk} + \underline{p}_{ik}\underline{p}_{jk}\right] \times \frac{2}{3} + \frac{\alpha}{4}\left[\bar{p}_{ik}\bar{p}_{jk} + \underline{p}_{ik}\bar{p}_{jk} + \bar{p}_{ik}\underline{p}_{jk} + \underline{p}_{ik}\underline{p}_{jk}\right] \times \frac{1}{3}}$$
$$= \frac{2\bar{p}_{ik}\underline{p}_{jk} + 2\underline{p}_{ik}\underline{p}_{jk}}{\bar{p}_{ik}\bar{p}_{jk} + \underline{p}_{ik}\bar{p}_{jk} + 3\bar{p}_{ik}\underline{p}_{jk} + 3\underline{p}_{ik}\underline{p}_{jk}}.$$

The wages are

$$y^R(1) = A_H$$

and

$$y^R(0) = A_H\psi^R + A_L\left(1 - \psi^R\right).$$

To check the incentive constraint

$$p_{ik}\bar{p}_{jk}y^R(1) - c_H > p_{ik}\bar{p}_{jk}y^R(0) > p_{ik}\bar{p}_{jk}y^R(1) - c_L \text{ for } p_{ik} \in \left\{\bar{p}_{ik}, \underline{p}_{ik}\right\}$$

or equivalently

$$y^R(1) - \frac{c_H}{p_{ik}\bar{p}_{jk}} > y^R(0) > y^R(1) - \frac{c_L}{p_{ik}\bar{p}_{jk}} \text{ for } p_{ik} \in \left\{\bar{p}_{ik}, \underline{p}_{ik}\right\}$$

and this must be true for the worst case on either side of the bound

$$y^R(1) - \frac{c_H}{\underline{p}_{ik}\bar{p}_{jk}} > y^R(0) > y^R(1) - \frac{c_L}{\bar{p}_{ik}\bar{p}_{jk}}.$$

This bound holds by Assumption (2). In this case both low and high centrality of H quality will save $s_i = 1$ with a high centrality monitor, irrespective of the saver centrality, and save $s_i = 0$ with a low centrality monitor, irrespective of saver centrality.

N.1.6. *Endogenous Assignment of Monitors.* Endogenous choice of monitor happens through random serial dictatorship. m participating agents are randomly ordered and then select a monitor in sequence, and the chosen monitor is removed from the pool.

Again, we want to compute the belief that the third party has that the saver is of type θ given they have received a report and therefore the saver has participated and has saved an amount s:

$$\mathrm{P}(\theta | s, r = 1, x = 1) = \frac{\mathrm{P}(s, r = 1 | \theta, x = 1)\,\mathrm{P}(\theta | x = 1)}{\mathrm{P}(s, r = 1 | H, x = 1)\,\mathrm{P}(H | x = 1) + \mathrm{P}(s, r = 1 | L, x = 1)\,\mathrm{P}(L | x = 1)}.$$

In our equilibrium observe that the following hold:

- Conditional on $\theta = H$:
  - $\mathrm{P}(s_i = 1, r = 1 | H, x = 1) = \alpha\frac{3}{8}\left[\bar{p}_{ik}\bar{p}_{jk} + \underline{p}_{ik}\bar{p}_{jk}\right]$
  - $\mathrm{P}(s_i = 0, r = 1 | H, x = 1) = \frac{\alpha}{8}\left[\bar{p}_{ik}\underline{p}_{jk} + \underline{p}_{ik}\underline{p}_{jk}\right]$
- Conditional on $\theta = L$:
  - $\mathrm{P}(s_i = 1, r = 1 | L, x = 1) = 0$

  - $\mathrm{P}(s_i = 0, r = 1 | L, x = 1) = \alpha \underline{p}_{ik} \underline{p}_{jk}$
- Type composition given participation:
  - $\mathrm{P}(H | x = 1) = \frac{2}{3}$
  - $\mathrm{P}(L | x = 1) = \frac{1}{3}$

In this case we can compute

- $\mathrm{P}(\theta = H | s_i = 1, r = 1, x = 1) = 1$,
- $\mathrm{P}(\theta = H | s_i = 0, r = 1, x = 1) = \frac{\bar{p}_{ik} \underline{p}_{jk} + \underline{p}_{ik} \underline{p}_{jk}}{\bar{p}_{ik} \bar{p}_{jk} + 5 \underline{p}_{ik} \underline{p}_{jk}}$, as

$$\begin{aligned}
\mathrm{P}(\theta = H | s_L, r = 1, p = 1) &= \frac{\frac{1}{8} \left[\bar{p}_{ik} \underline{p}_{jk} + \underline{p}_{ik} \underline{p}_{jk} \right] \times \frac{2}{3}}{\frac{1}{8} \left[\bar{p}_{ik} \underline{p}_{jk} + \underline{p}_{ik} \underline{p}_{jk} \right] \times \frac{2}{3} + \underline{p}_{ik} \underline{p}_{jk} \times \frac{1}{3}} \\
&= \frac{\bar{p}_{ik} \underline{p}_{jk} + \underline{p}_{ik} \underline{p}_{jk}}{\bar{p}_{ik} \bar{p}_{jk} + \underline{p}_{ik} \underline{p}_{jk} + 4 \underline{p}_{ik} \underline{p}_{jk}} \\
&= \frac{\bar{p}_{ik} \underline{p}_{jk} + \underline{p}_{ik} \underline{p}_{jk}}{\bar{p}_{ik} \bar{p}_{jk} + 5 \underline{p}_{ik} \underline{p}_{jk}}.
\end{aligned}$$

The wages are

$$y^E(1) = A_H$$

and

$$y^E(0) = A_H \psi^E + A_L \left(1 - \psi^E\right).$$

Consider an L quality agent. Note that as long as $y^E(0) < 0$, which happens in equilibrium by assumption (1), if the agent is planning to save the low amount, then it is trivially better to do so under a low centrality monitor since $0 > p_{ik} \underline{p}_{jk} y^E(0) > p_{ik} \bar{p}_{jk} y^E(0)$. If the agent is planning to save the high amount, then so long as $y^E(1) > 0$, it is trivially better to do so with a high centrality monitor.

The L type will prefer the low monitor and save s_L provided it exceeds the maximal possible benefit under a high monitor/high savings combination

$$0 > p_{ik} \underline{p}_{jk} y^E(0) > p_{ik} \bar{p}_{jk} y^E(1) - c_L$$

which is implied by

$$\frac{c_L}{p_{ik} \bar{p}_{jk}} > y^E(1) - y^E(0),$$

which in turn is implied by Assumption (2).

Similarly one can check that by Assumption (1), the incentive constraint is met for the H-type as well.

N.1.7. *Entry Decision.* Let us compute the expected payoff to entering:

$$\frac{\alpha}{2} \pi^R(\eta) + \frac{\alpha}{2} \pi^E(\eta)$$

and consider the case of low quality agents. Under the maintained assumptions, even when L quality agents enter, they will not signal by investing s_H, since it is too costly. In our equilibrium entry, L quality agents will always be able to choose their preferred monitor type, because L-types comprise $\frac{1}{3}$ of the saver pool but $\underline{p}_{jk}$-monitors are $\frac{1}{2}$ the pool and H types prefer $\bar{p}_{jk}$-monitors . Under this and the assumption (3), L-quality agents do not enter if they are of $\bar{p}_{ik}$-centrality whereas $\underline{p}_{ik}$-centrality agents do enter.

So now consider m agents which are comprised of only $\left\{\left(H,\underline{p}_{ik}\right),\left(H,\bar{p}_{ik}\right),\left(L,\underline{p}_{ik}\right)\right\}$ agents, each with equal proportions. There are m monitors which are $\frac{1}{2}$ $\bar{p}_{jk}$-centrality and $\frac{1}{2}$ $\underline{p}_{jk}$-centrality. Under random serial dictatorship, an H-quality agent who goes in the first $\frac{3}{4}$ of the H-order will have the payoff

$$\pi^E\left(H,p_{ik}\right)=p_{ik}\bar{p}_{jk}y^E\left(1\right)-c_H>0$$

whereas the H-quality agent allocated in the last $\frac{1}{4}$ of the H-order gets

$$\pi^E\left(H,p_{ik}\right)=p_{ik}\underline{p}_{jk}y^E\left(0\right)<0.$$

Then the expected utility of entering (scaled by $\frac{2}{\alpha}$)

$$\begin{aligned}
& \frac{p_{ik}\bar{p}_{jk}+p_{ik}\underline{p}_{jk}}{2}y^R\left(0\right)+\frac{1}{2}\left(p_{ik}\bar{p}_{jk}\left(y^R\left(1\right)-y^R\left(0\right)\right)-c_H\right) \\
& +\frac{3}{4}\left[p_{ik}\bar{p}_{jk}y^E\left(1\right)-c_H\right]+\frac{1}{4}p_{ik}\underline{p}_{jk}y^E\left(0\right) \\
= \quad & \frac{5}{4}p_{ik}\bar{p}_{jk}A_H+p_{ik}\underline{p}_{jk}\left(A_H\frac{2\psi^R+\psi^E}{4}+A_L\frac{3-2\psi^R-\psi^E}{4}\right)-\frac{5}{4}c_H
\end{aligned}$$

and therefore entry occurs as long as

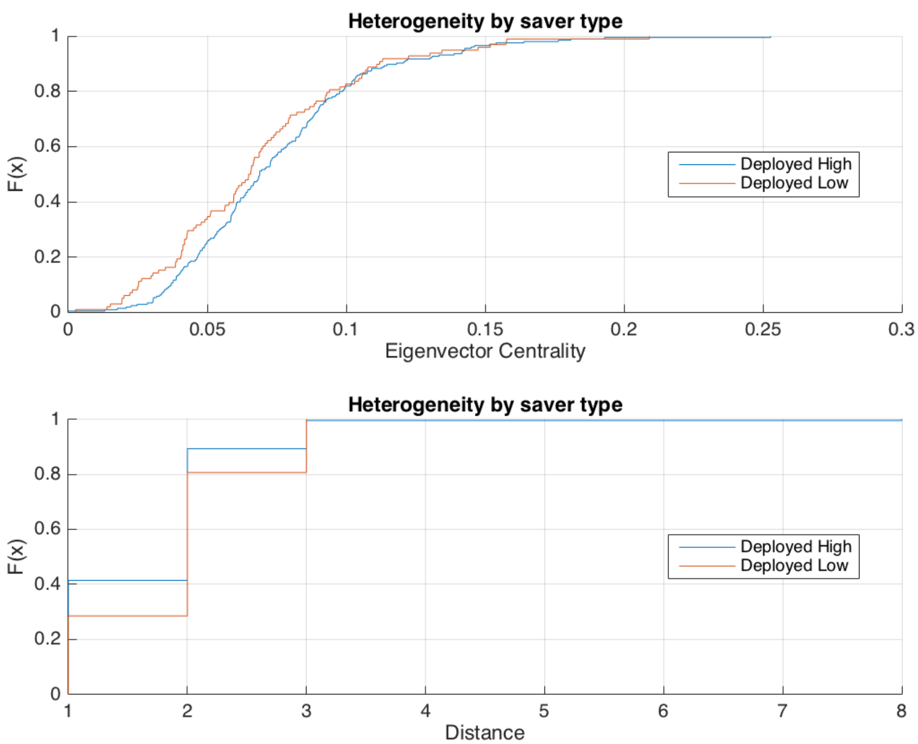
$$\frac{5}{4}p_{ik}\bar{p}_{jk}A_H+p_{ik}\underline{p}_{jk}\left(A_H\frac{2\psi^R+\psi^E}{4}+A_L\frac{3-2\psi^R-\psi^E}{4}\right)>\frac{2}{\alpha}v_{aut}+\frac{5}{4}c_H$$

and a sufficient condition is just assumption (4),

$$5p_{ik}\bar{p}_{jk}A_H+3p_{ik}\underline{p}_{jk}\left(A_H\psi^R+A_L\left(1-\psi^R\right)\right)>\frac{8}{\alpha}v_{aut}+5c_H.$$

N.2. **Endogenous Monitor Choice: Empirical Evidence.** We now turn to the data. Figure N.1 shows the CDFs of chosen monitors in endogenous choice villages, broken by whether the saver is of high or low centrality, where high means above the median. As anticipated above, over the distribution high centrality savers pick more central monitors and more proximate monitors. A Kolmogorov-Smirnov test shows that the centrality CDFs are statistically distinguishable ($p = 0.06$) whereas the proximity distributions are not ($p = 0.28$).

FIGURE N.1. Centrality Distribution of Chosen Monitors



Notes: Figures plot cdfs of the chosen monitor characteristics in endogenous choice villages, separately by high and low centrality savers. We define high centrality as above the median value. The top figure plots monitor centrality, while the bottom figure plots saver-monitor proximity.

Next, in Table N.1 we look at how the choice order affects the centrality of the monitor. We find that picking earlier leads to a choice of more central monitors that we can detect if the saver is of high centrality, but there is no such relationship when we look at low centrality savers. This too is consistent with our stylized model that explores choice.

Table N.1. Monitor Choice Order in Endogenous Allocation Villages

	High Centrality Savers		Low Centrality Savers	
	(1)	(2)	(3)	(4)
Dependent Variable	Log Total Savings	Monitor Centrality	Log Total Savings	Monitor Centrality
Choice Order: 6-10	-0.584**	-0.325	0.542*	0.0499
	(0.257)	(0.230)	(0.299)	(0.234)
Choice Order: 11-15	-0.847**	-0.395*	0.0818	0.0443
	(0.315)	(0.215)	(0.335)	(0.223)
Choice Order: >15	-0.813***	-0.279	-0.145	-0.281
	(0.273)	(0.262)	(0.333)	(0.358)
Observations	202	202	168	168
R-squared	0.138	0.033	0.027	0.035

Notes: Table shows impacts of monitor selection choice order on log total savings and monitor centrality by saver centrality. We define high centrality as above the median value. Total savings is the amount saved across all savings vehicles. Sample constrained to savers who received a monitor in the 30 endogenous random-assignment villages and who answered our questionnaire. Regressions control for log savings goal and saver centrality and also include village fixed effects. Standard errors clustered at the village level.

Appendix O. Discussion of Implementation Costs

One relevant policy consideration is the cost of implementing and scaling a peer monitoring product. Our specific treatments were implemented with research goals in mind and were never meant to be profitable or scalable. However, we do think that there are many opportunities for financial institutions to reduce the costs of product delivery. One of our largest costs was personnel. In order for the research team to have more control over the implementation, we chose to send individuals to each village on a bi-weekly basis to meet the savers, physically verify the passbooks, and pass the relevant information on to the monitors. Many financial institutions in India already use the Business Correspondent (BC) model, in which agents of the bank travel to villages to provide direct in-home customer service. This includes account opening procedures and deposit-taking. One could easily imagine a small tweak to this model, where the BC could intermediate information to others in the village after his pre-specified appointments. Further, banks could use technologies such as SMS to implement a peer monitoring scheme.

The other main cost associated with our intervention was the incentive given to monitors. First, as discussed previously, we think that the incentives had negligible effects on savings outcomes. Second, we certainly did not attempt to "optimize" the size of these incentives.[55] Nevertheless in the endogenous monitor case, the aggregate monitor incentives paid to participants correspond to a 6% semi-annual interest rate on all additional savings that were caused by our interventions, which – while not cheap – is not outlandish.[56] Experimenting with the size of the incentives would likely yield significant cost reductions.

[55] In fact, we believe that the optimal incentive would be close to, if not equal to, zero.

[56] To reach this 6% value, we first calculate the aggregate payments that we made to monitors in the Endogenous villages. We then calculate the excess savings across all savings vehicles that were caused by our treatments. We include both the direct effects of receiving a monitor on savings and also the spillovers onto non-monitored savers.

Appendix P. The Effects of the "BC" Bundle on Savings

Here we measure whether the basic bundle of services given to all participating saver by itself affected savings. Recall that all interested savers received the following services: goal elicitation, account opening facilitation, initial required deposit (Rs. 100) into the account, biweekly visits from an enumerator. While all of our previous analysis restricts the sample to savers who selected into the program, we now draw upon additional data for this exercise. Namely, as shown in Figure 2, we randomized households into participation as potential savers. Moreover. we surveyed a random subsample of the pure control group – households that were never approached for participation – along with a random subsample of non-takers – households that were approached to participate in the program but who did not attend the village meeting. Using our sampling rates, we construct an intent to treat estimate of the benefits of the "BC" bundle on log savings. We omit all monitored savers.

Table P.1 presents the results of this exercise. In column 1 we pool across Endogenous and Random assignment villages and find a modest positive, but statistically insignificant effect of the treatment on log total savings. In column 2, we allow the treatment effect to vary by the village-allocation method. We find in Column 2 that the effect is larger in Endogenous villages, but again neither coefficient is statistically significant.

Table P.1. Effects of the "Business Correspondent" Bundle

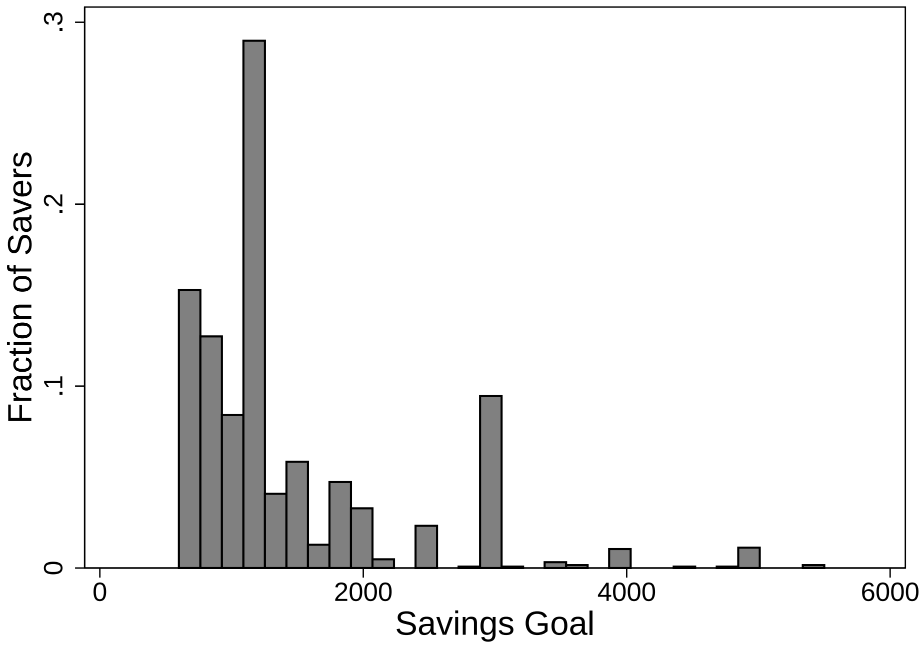
	(1)	(2)
Dependent Variable	Log Total Savings	Log Total Savings
Business Correspondent (BC) Bundle, No Monitor	0.0731	0.0548
	(0.0903)	(0.128)
BC Bundle: Endogenous Assignment Village		0.0387
		(0.180)
Observations	1,835	1,835
R-squared	0.059	0.059
Dependent Variable Mean (Omitted Group)	7.475	7.475
Fixed Effects	Village	Village

Notes: Table measures the effects of receiving the baseline bundle of services for non-monitored potential savers compared to the pure control. To do this, we use the random sample of dropouts who we surveyed and reweight the observations to reconstitute the full potential savers. The control group is comprised of individuals with whom we never interacted during the intervention and whom we randomly sampled at endline. Total savings is the amount saved across all savings vehicles – the target account and any other account, both formal and informal including money held "under the mattress" – by the saver. Given that we do not have any baseline data for the control group, we include no controls other than the village fixed effects. Standard errors clustered at the village level.

These results are consistent with the findings of Table 8, but reflect a severe lack of power. In Table 8, recall that we are only focusing on savers opting into the village meeting. Here, in Table P.1, this effect is diluted by the mass of individuals that did not opt in and were never treated with the "BC" bundle.

Appendix Q. Baseline Savings Goals

Figure Q.1. Histogram of Baseline Savings Goals



Notes: The figure shows the distribution of the baseline savings goals. We clip the top 5% tail of the distribution to make the figure more readable.

Figure Q.1 presents the histogram of savings goals, censoring the top 5%.[57] There are a few large outliers (maximum goal Rs. 26,000), so the mean of Rs. 1838 shrinks to Rs. 1650 when we trim 1% outliers. In all specifications of our key results we drop the top 1% of savings goal observations.

[57]Note that the minimum goal is Rs. 600, the lower bound of allowed goals for participants.