Social (Media) Learning: Experimental Evidence from Indian Farmers

Vanisha Sharma, Cornell University

DRAFT NOT FOR CIRCULATION

July 2, 2023

Abstract

The recent internet expansion in India has led to an increase in the rural uptake of smartphones and social media. In this paper, I study whether farmers are willing to exchange agricultural information with their peers online, or if they need an in-person interaction to learn. I measure the effects of this information exchange on their farming expenses and revenues. Self-selection into social media networks makes it challenging to tease out effects of online peer interactions. To address this, I exogenously assign farmers to multi-village Whatsapp groups in an experiment spanning 108 villages in rural south India. In another treatment, I invite the farmers to interact in-person with their online peers. I find that treated farmers invest more in their farms, by an average of Rs. 10,000 ($\approx$ \$120), but I do not find significant increases in their farm revenues. The increase in farm investment is driven by increases in online information exchange. The added in-person interactions do significantly add to the effects of online information exchange. While farmers are able to exchange useful agricultural information online at low costs, this online information exchange does not lead to significant revenue returns in the short-run.

1 Introduction

Information constraints present significant barriers for decision-making in developing countries. With recent internet expansion in India, the costs of information exchange have drastically reduced. Since 2014, demand for social media messaging services such as Whatsapp has been steadily increasing in the Indian market (discussed in more detail in section 2.2). So much so that today, the company has its biggest market in India with over 300 millions users, with its penetration steadily expanding in rural areas as well [Iqbal, 2020]. Sharing new information with peers across space at negligible costs has the potential to economically impact multiple domains including politics, health, education, and even agricultural productivity.

There is consistent evidence in prior literature that farmers' decision making process and technology adoption is influenced by their peers [Conley and Udry, 2001] [Foster and Rosenzweig, 1995] [Bandiera and Rasul, 2006] [Beaman et al., 2018]. These studies have measured peer-effects for individuals in geographically proximate networks. A growing body of literature has also focused on the usage of mobile phones in agriculture, especially in developing countries.[1] For instance, there is some evidence that shows that communication through mobile phones can affect market price dispersion, and farmer self-efficacy beliefs [Jensen, 2007] [Aker, 2010] [Lasdun, 2022]. A more recent study revealed that farmers that had better network coverage were more likely to adopt high yielding variety seeds, and witnessed greater productivity as a result [Gupta et al., 2020].

Today, despite increasing rates of rural social media usage for communication, the prevalence of using social media for agriculture remains relatively low. This is probably why the effects of social media on farmer learning remains a relatively understudied area in economics. Are in-person interactions crucial to social learning in farmers, or can social media interactions substitute the traditional channels of farmer peer effects? These questions come with substantial identification challenges, because social media participation is endogenous in that it is correlated with other factors that affect farmer learning and technology adoption.

[1] For a detailed review of mobile phone usage and economic development in sub-Saharan Africa, refer to [Aker and Mbiti, 2010].

In this paper, I address these identification challenges through a multi-faceted randomized control trial in India. The experiment has two treatment arms. In the first arm, farmers from non-neighboring villages were connected on multi-village, moderated Whatsapp groups (one group per 3 villages) to facilitate online information sharing of farming practices. In the second arm, along with the Whatsapp groups, farmers were invited to meet in-person monthly to discuss several farming related topics over lunch, and attend a poster-presentation. Both arms facilitated online or in-person discussions about various topics ranging from pesticide use, pest/weather shocks in the region and how to prevent crop loss to such shocks, preferred locations of input dealers and so on. Using a sample of 108 villages and 1,083 farmers in the southern Indian state of Andhra Pradesh, I test the effects of online versus online+in-person interactions on adoption of Whatsapp as an input for informing farming decisions, farmer information exchange behavior, as well as farming outcomes.

From the current literature we know that agriculture technology adoption drives, and it primarily driven by diffusion of information, peer-effects and social learning. This study contributes to each of these aspects in the following ways. First, there is a vast body of literature on peer-effects and social learning. [Besley and Case, 1993] ,[Bandiera and Rasul, 2006] and [Comola and Prina, 2015] provide some useful insight on how peer effects are dynamic. [Nourani, 2016] demonstrates how strong ties affect the sophisticated farmer learning behavior and adoption in a network. The linear in-means model by [Manski, 1993] explains the typical methodology to identify peer effects, and [Goldsmith-Pinkham and Imbens, 2013] expand on this model by capturing the influence of indirect links.

A small yet important caveat here is that peer-effects and social learning are not synonymous; while peer-effects influence actions, learning from peers should reflect in greater knowledge of productivity [Foster and Rosenzweig, 1995]. [Kondylis et al., 2017] and [Cai et al., 2015] tackle this challenge in different ways, by measuring knowledge scores of individuals conditional on being in treated networks to extract the effect of learning versus imitation or scaling effects. [Magnan et al., 2015] show that mirroring adoption from peers is contingent on knowledge of adopters and does not hold for adopters without knowledge about the new technology. But in addition to the regular challenges of social network data collection [Maertens and Barrett, 2013] and informational inefficiencies [Jack, 2013], the deeper issue still remains that the information being shared between peers is not observed, or is self-reported. [Beaman and Dillon, 2018] deal with this issue by asking participants to exchange crop-cycle calendars with their friends, to get tangible evidence of knowledge exchange between peers, which requires more time and effort. One of the major contributions of this work is that by using Whatsapp groups as the primary platform for information exchange, I collect timely, high-frequency data not only on the frequency of peer-interactions, but also on the content of those interactions. This will aid in making robust causal inferences of the intervention on adoption specifically through learning.

Third, most literature on agriculture technology adoption measures adoption as input use or knowledge of input application [Foster and Rosenzweig, 1995] [Magnan et al., 2015] [Conley and Udry, 2001] [Munshi, 2004] [Bandiera and Rasul, 2006] [McNiven and Gilligan, 2012] [Kondylis et al., 2017] [Beaman et al., 2018] [Maertens et al., 2020]. Only recently, there was a study conducted in Tanzania to measure self-efficacy and learning through creating a new SMS chat platform [Lasdun, 2022]. Subsequently, to my knowledge, this remains the first paper to discuss the causal effects of online peers, versus when these peers are invited to meet in-person, on farming outcomes and information sharing. Lastly, while some literature has aimed at measuring welfare effects of adoption (for example, [Jensen, 2007], [Harou et al., 2017]), this area departs from the focus of this study.

Lastly, in the past decade, a plethora of studies has aimed to increase effective diffusion of information, through various methods- by identifying optimal injection points using the various centrality measures, or social identity of nodes [Banerjee et al., 2013], [Beaman and Dillon, 2018] [BenYishay and Mobarak, 2019] [Kondylis et al., 2017] [Cai et al., 2015], by varying threshold for contagion [Beaman et al., 2018] , or by exploiting strength of ties [Nourani, 2016] [Murphy et al., 2019]. In his seminal contribution, Granovetter [Granovetter, 1973] emphasizes the importance of weak ties in bridging disconnected information islands consisting of clusters of strong ties. [2] In this study, I exploit a pre-existing, essentially free platform of communication that is already being used by farmers who have access to smartphones, but is not being widely used for discussions around farm

[2] The "strength" of an interpersonal tie is a linear combination of the amount of time, the emotional intensity, the intimacy (or mutual confiding), and the reciprocal services which characterize each tie. See [Granovetter, 1973] for more details, and Fig. 5 for a pictorial representation. Since the endogenously formed networks can create 'echo-chambers', it is the weak ties that are responsible for bridging the information divide.

productivity.[3]

The rest of the paper is organized as follows: section 2 explains the setting of the experiment, while section 3 explains the research design. Section 4 introduces a conceptual framework to guide empirical hypotheses, Section 5 explains elaborates on the data collected, and section 6 goes over the empirical strategy. Section 7 describes the results and section 8 discusses the findings.

2 Context in rural Andhra Pradesh

Andhra Pradesh is a state in south India, which was bifurcated into itself and Telangana in 2014. According to the last census data from 2011, 38.9% of the total population of the area is urban. Over 85% of farm holdings are small and marginal, and farmers mostly rely on rainfall for irrigation [D. Raji, 2017]. Madanapalle, the nearest market to the study's location is the biggest tomato market in India, and a large part of the farmers in the sample grew tomatoes in at least one of the three seasons in the duration of the study (see Fig. 8 in Section 7).

2.1 Traditional Agriculture Extension Services

Other than the involvement of its state government, Andhra Pradesh consists of multiple agriculture extension institutes, such as the Agriculture Research Station, Institute of Crop Research in Semi-Arid Tropics (ICRISAT) a few miles outside its border, and a few others. The state also consists of several research universities that focus on providing training in extension services, such as Acharya N. G. Ranga Agricultural University. Therefore, most farmers, especially in villages neighboring cities are able to rely on agricultural advisory from extension officers. Extension services are also provided on TV and radio, and more recently, through Kisan Call Centers which allow for tele-advisory over phone-calls to experts on a toll-free number.

Tuta Absoluta: A Pest Concern

Recently, an invasive pest species migrated to India from Latin America, locally known as Uzi, which particularly affects tomato farmers in the region [Buragohaina et al., 2021]. Government and research institutes are trying to reach as many farmers to adopt an integrated pest management

[3] This was revealed to the author in pilot farmer interviews before the beginning of the experiment.

method to deal with this pest. In the current landscape, farmers tend to over-invest in pesticides, and not utilize other non-chemical methods to prevent the pest, which is what the research institutes are trying to address.

2.2 Social Media Usage

The recent internet expansion in India started in 2016, with a public-private partnership 'Jio', which laid out at least 250,000km of fiber optic cables all throughout India. This led to a significant drop in the price of 4G internet, close to merely $2 per month for unlimited mobile data. By 2020, around 450 million people were using mobile phone internet. This in turn led to a significant increase in engagement with phones and applications, especially social media platforms- even more so during the Covid-19 pandemic. As we can see in the following Google Trends graph in figure 1, the interest in Whatsapp spiked in 2016, and then peaked in 2020. [4]

A more interesting pattern was formation of Whatsapp groups with social connections that surpassed geographical boundaries, allowing for several channels of virtual information exchange to form at very low costs at the intensive margin. While it may be intuitive that these groups existed to reinforce pre-existing ties in a social context, baseline data from this study revealed that farmers mostly used their smartphones for communicating with friends and family, and not for agriculture (Fig. 2). This indicates that there exists a possibility to convey agricultural information on an already existing and popular platform to facilitate information flow between farmers across space.

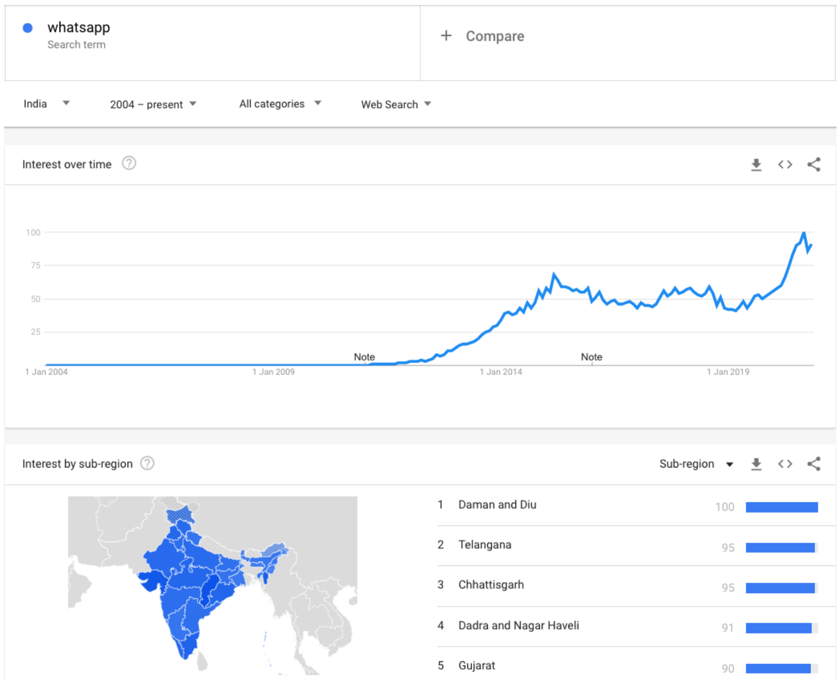
3 Research Design

The sample was selected from a 2-stage randomization process. First, 108 villages (clusters) were selected within a radius of 20km from the town of Madanapalle, Andhra Pradesh, which also consists of the biggest vegetable markets (daily and weekly) in the region. Then, within each village, on average upto 10 smartphone and regular phone owning farmers were enrolled in the study, provided they satisfied the following 3 criteria:

1. They farm in both the monsoon and winter crop cycles.

[4] The interest in other social media platforms such as Facebook did not follow a similar pattern, and in fact declined in the last 5 years [see Fig. C.1 in the Appendix], which is why Whatsapp remains the focus of this study.

Figure 1: Growth in searches for Whatsapp



Note: This figure depicts a google trends graph showing the increasing interest over time in Whatsapp.

2. They are the primary agricultural decision maker.

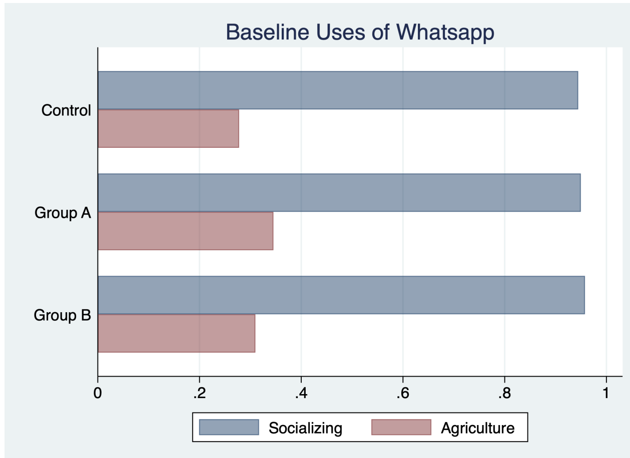
3. Their primary source of income is agriculture.

The reason for the first criterion is to allow panel data collection throughout the calendar year without losing farmers who only farm in one season. The second criterion ensures that the treated participants will be the ones making all farming-related decisions. The third criterion is listed to exclude farmers whose livelihoods do not rely solely on agriculture, and can be thought of as a proxy for intrinsic motivation to improve agricultural practices and outcomes. Out of all the enrolled farmers, 90% from each village were randomly selected into the sample. [5] Farmers were contacted through socially distanced, in-person conversations in each village.

After the baseline survey, the sample villages were stratified into quartiles based on distance from the

[5] The reason for surveying smartphone non-owners is to measure possible spillover effects of treatment to farmers who do not have direct access to smartphone technology.

Figure 2: Reasons for using Whatsapp



Note: This graph depicts a sample subset of 407 farmers who owned smartphones and used Whatsapp in the beginning of the study. Regardless of treatment assignment, over 90% of the farmers used Whatsapp as a means to communicate with friends or family, and only around 30% used Whatsapp for agriculture.

biggest nearest market, Madanapalle (Fig. 3). Then, the villages were randomized from each strata into a pure control and 2 treatment groups. Both smartphone owners and non-owners were surveyed to measure potential spillover effects of the Whatsapp intervention. 5% of the sample reported to purchase a new smartphone during the course of the study, and these farmers were invited to the Whatsapp groups subsequently. This is not a concern since the treatment assignment is at the village level. The goal of the Whatsapp moderation was two-fold:

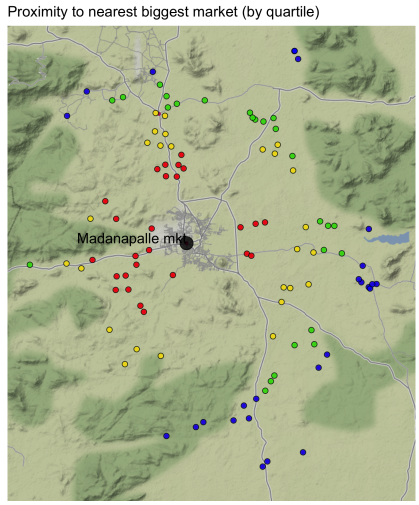
1. To solicit participation from the farmers by asking questions that were relevant in the region, for example satisfaction of input quality, or tips to prevent pests on tomato crops.

2. To share information from our end in conjunction with research institutes on issues that were brought up in the pilot surveys.

Farmers in the treatment villages who own smartphones were assigned to one of the following two treatments:

1. In the only-Whatsapp arm, randomly selected farmers from non-neighboring villages were connected on 12 Whatsapp groups, three non-adjacent villages per group with approximately each group having on average 15 farmers (see Fig. A.3 in Appendix). Random assignment of

Figure 3: Sample of 108 villages



Note: Navy dots represents the villages that are in the farthest quartile of distance from the biggest market (Madanapalle), red dots are villages in the closest quartile. Treatment assignment was stratified on distance quartiles. Fig. A.4 in the Appendix shows the distribution of these villages by distance to Madanapalle.

villages ensures exogeneity in network formation, while them being non-adjacent reduces the probability that the Whatsapp group members already know each other in-person.

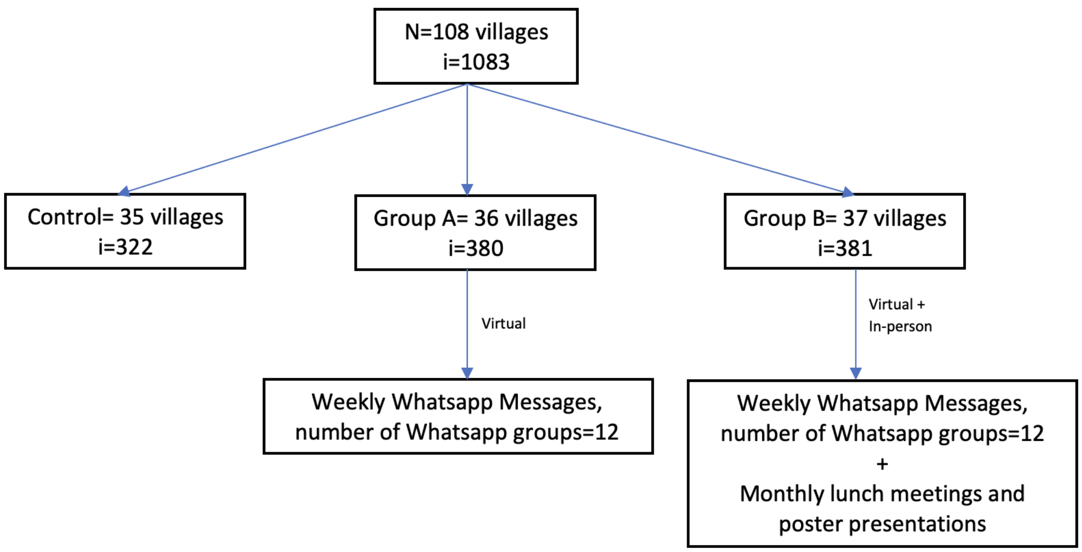
2. In the Whatsapp-plus-lunch and poster presentations arm, other than being connected on non-adjacent, 3-village Whatsapp groups, the farmers were also invited to attend monthly lunch gatherings which included poster presentations. Posters were made by our team on topics that were in-demand based on conversations with pilot farmers, and were based on the timing of the crop cycle. This is the exact same information being shared on Whatsapp groups, but in this arm, the farmers are allowed to interact with the posters and with each other in-person (see Section A in Appendix for examples). The purpose of this arm is to complement the virtual interactions on these exogenously created Whatsapp groups to extract the differential effects of adding an in-person component to virtual interactions. [6]

The Whatsapp groups were moderated by enumerators to monitor the conversations and encourage participation. To ensure homogeneity in moderation across the 24 Whatsapp groups, I created a

[6] In an ideal setting, I would also have a third pseudo-control arm where information would be shared with the farmers only in-person. Given standard sample size and cluster-size requirements, this exceeded the time and cost capacities of this study and was therefore not feasible.

schedule of messages to be sent as prompts to increase engagement, or responses to farmer questions to ensure consistency across groups (pdf in Appendix). The timeline of the project is depicted in Fig. in the Appendix. I also measure the number of days spent on the groups per farmer and use this as a proxy for exposure to treatment (see Fig. A.6 in Appendix). The following figure 4 depicts treatment assignment.

Figure 4: Treatment Assignment



Note: Treatment was assigned at the village level to avoid contamination of treatment within villages. For each village, approximately same number of both smartphone-owners and button-phone owners were interviewed. The total number of farmers enlisted was 1293, but only 1083 were retained for a balanced panel across the three survey waves.

4 Conceptual Framework

In this study, I aim to create and measure the influence of online peers as compared to in-person peers, by enabling long-distance social media links that connect farmers across space, on farmer information exchange and agriculture outcomes.

Benchmark Framework

In a benchmark framework, farmers optimize input use to maximize their profit [Weersink and Fulton, 2020]. They seek agricultural information from their geographically proximate peers, family members, input dealers, extension agents and so on. However, search for this information bears a cost. Let this cost be given by $\theta_{i,t}$ for farmer i in period t. Consider the following farmer production function:

$$y_{i,t} = f(\mathbf{x}_{i,t}, \omega(\theta)_{i,t}, \mathbf{z}_{i,t}) \tag{1}$$

where $y_{i,t}$ is the total yield for farmer i for crop cycle t. $\mathbf{x}_{i,t}$ is a vector of inputs used in that crop cycle. This vector of inputs also includes farmer adoption of Whatsapp for agriculture. The input prices are given by the vector $\omega(\theta)_{i,t}$, which is a function of the search cost of information, given by θ. $\mathbf{z}_{i,t}$ represents the individual characteristics of the farmer. Then the profit maximization problem for the farmer is:

$$\max_{\mathbf{x}_{i,t}} pf(\mathbf{x}_{i,t}, \omega(\theta)_{i,t}, \mathbf{z}_{i,t}) - c(\omega(\theta)_{i,t}, y_{i,t}) \tag{2}$$

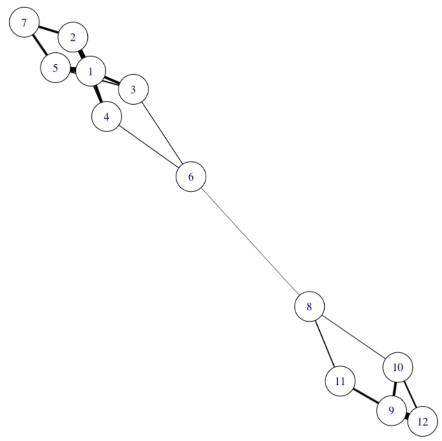
Framework with Treatment

Borrowing from the literature on strength of ties [Granovetter, 1973] [Nourani, 2016], let us assume that pre-existing farmer networks represent 'strong' ties with frequent interactions, that are more likely to influence farmer behavior. These networks can be both in-person, or online.[7] Suppose the new randomly assigned networks in the treated villages represent 'weak' ties, or acquaintances. (see Fig. 5 for reference). For tractability, let us suppose that the effects of both strong and weak peers are additively separable. The goal is to eventually test the effects of these peers on farmers' willingness to exchange information, and consequently their farming outcomes.

With increased online access, farmers' sources of information exchange increase. Direct online contact with distant farmers reduces the degree of separation, thereby reducing the search costs associated with this information and affecting farmer input prices and usage. While the search costs may increase due to misinformation or noise, rationally bounded farmers, in expectation, would

[7] While in this context, in-person farmer networks are way more prevalent, I include virtual networks to account for potentially preexisting online farmer connections.

Figure 5: Example of Weak Ties bridging Two Strong Ties Islands



Note: Here, strength of ties is proportional to thickness of the edges between two nodes. That is, nodes 6 and 8 are weakly connected, while 1-7 in the top left, and 8-12 in the bottom right are more strongly connected to each other. This is analogous to strong within-village networks being connected through a weak across-village link.

optimize their Whatsapp use and hence, their input use, excluding the noisy signals.

Intuitively, the search cost of information is a function of number of other farmers in farmer i's information neighborhood, as well as the nature of information being shared in the group. θ in equation 1 already accounts for pre-existing in-person and virtual connections for farmer i. For treated groups, let the search cost be given by $\theta - \hat{\theta}(inp_i, v_i, part_i)$, where $-\hat{\theta}(\cdot)$ is the search cost reduction for treated farmers. $\hat{\theta}(\cdot)$ is a function of the number of farmer i's new in-person and virtual connections, given by inp_i and v_i respectively, and her own participation with these links, given by $part_i$.[8] [9] Then, the production function for treated farmers will be:

$$y_{i,t} = f(\mathbf{x}_{i,t}, \omega(\theta - \hat{\theta}(\cdot))_{i,t}, \mathbf{z}_{i,t}) \tag{3}$$

The goal of the treatment is to reduce information search cost and optimize input use. Since this cost is latent, I use other outcomes such as farm revenue and expenditure to measure treatment effects. The mechanism through which treatment affects revenue and expenditure decisions is changes

[8] Participation in the Whatsapp groups measures compliance of treated farmers.

[9] Note that for the only-virtual treatment, $inp_i = 0$ such that $\frac{\partial \theta}{\partial inp} = 0$.

in farmer information sharing behavior post-treatment. While information sharing is only observed for treated groups, self-reported measures are collected for all farmers (explained further in section 5). The intuition is that greater information exchange between distant individual farmers will lessen the cost of searching information, and increase chances of optimal decision-making for input use.

Let Y_i denote the set of outcomes for farmer i including the revenue $(p \cdot f(\mathbf{x}_{i,t}, \omega(\theta - \hat{\theta}(\cdot)_{i,t}, \mathbf{z}_{i,t}))$, farm expenditure $(c(\omega(\theta - \hat{\theta}(\cdot)_{i,t}, y_{i,t}))$, and information sharing behavior (further explained in Section 5). Let $Treat_i$ denote whether farmer i was in a treated village, and $Wave_i$ denote the survey wave. A simple regression framework yields that an ITT effect is captured by the following equation:

$$Y_i = \alpha + \beta Treat_i + \gamma Wave_i + \delta(Treat_i \times Wave_i) + \epsilon_i$$

Here,

$$E[Y_{pre}^{Treat=1}] = \alpha + \beta$$

$$E[Y_{post}^{Treat=1}] = \alpha + \beta + \gamma + \delta$$

$$E[Y_{pre}^{Treat=0}] = \alpha$$

$$E[Y_{post}^{Treat=0}] = \alpha + \gamma$$

Given that the sample was randomized and balanced, we have the following:

$$\hat{\delta}_{ITT} = E[\bar{Y}_{post}^{Treat=1}] - E[\bar{Y}_{pre}^{Treat=1}] - E[\bar{Y}_{post}^{Treat=0}] - E[\bar{Y}_{pre}^{Treat=0}]$$

$$\implies \hat{\delta}_{ITT} = \alpha + \beta + \gamma + \delta - (\alpha + \beta) - (\alpha + \gamma - \alpha)$$

Which can be rearranged as:

$$\hat{\delta}_{ITT} = \underbrace{\delta}_{\text{Weak-tie effects for Treated}} + \underbrace{\alpha + \beta + \gamma - (\alpha + \beta)}_{\text{Strong-tie effects for Treated}} - \underbrace{(\alpha + \gamma + \sigma - \alpha)}_{\text{Strong-tie effects for Control}}$$

$$\implies \hat{\delta}_{ITT} = \delta$$

In expectation in a balanced panel, the strong-tie peer effects of preexisting peers should be the same regardless of treatment. This enables us to measure the additional treatment effects of new peers separately, given by δ.

Since there are two treatment arms, we can decompose δ into δ_v and δ_{inp}, where δ_v denotes the effects on the only-virtual arm and δ_{inp} measures the effect on the virtual+in-person arm. Given this premise, we get the following hypotheses:

Hypothesis 1: Given y_i denotes farm revenue (expenditure), $\delta > (<) 0$. This implies that conditional on useful information being exchanged on the Whatsapp groups, treated farmers should see positive effects on their farming outcomes.

If $\delta_v > (<) \delta_{inp}$, then the in-person component is not necessary for utilizing online farmer networks. On the other hand, if $\delta_v < (>) \delta_{inp}$, then the in-person component is necessary to optimally use the information being shared on the online networks.

Hypothesis 2: Given y_i denotes willingness to exchange agriculture-related information with unknown farmers, the treatment effect $\delta > 0$. That is, both treatments encourage farmers to exchange more information whether it is online, or in-person.

If $\delta_v > \delta_{inp}$, then the in-person component is crowding out online interaction with peers, since the only virtual arm increases the willingness to exchange information by more. On the other hand, if $\delta_v < \delta_{inp}$, then the in-person interactions complement the virtual interactions and increase the willingness to exchange information more than only virtual interactions. Conditional on the value of information being shared, this exchange is the mechanism through which treatment impacts farming outcomes.

5 Data

There are two main sources of data in this study- farmer-level survey data and observational data on Whatsapp groups for treated villages. The survey was conducted over 3 waves- baseline, midline and endline. Each wave was collected in the initial stages of the 3 seasonal crop cycles. Observational data on Whatsapp groups was manually recorded by tracking each farmer's participation in their respective group. Transcripts of the Whatsapp chats were also recorded.

5.1 Outcome variables

I measure the effect of treatment on two intermediary outcomes and three main outcome variables.

- Higher-order outcomes

    1. *Farm revenue/ Value of yield*: Due to the panel nature of the data, I am also able to test for treatment effects of information exchange through Whatsapp groups on farm revenues. This is calculated by self reported price and yield data. [10]

    2. *Input expenditure*: This includes total monetary costs of procuring farming inputs including pesticides, fertilizers, seeds, irrigation etc. Costs on pesticides are also separately collected.

    3. *Crop-loss due to pesticide*: Because of the variation in type of crops grown, I am unable to test convergence to an optimal level of input use. I do collect data on area of crop lost due to pests in each season.

- Intermediary outcomes and mechanisms

    1. *Farmer willingness to exchange agricultural information*: I collect data on how farmer beliefs are updated over time with respect to receiving and sharing information from known and unknown farmers- in-person versus over the phone. This is one of the primary mechanisms I propose behind any potential treatment effects.

    2. *Whatsapp participation and adoption for agriculture*: I also collect observational data on the treated sample's Whatsapp participation- including content and frequency of message sharing. I also include data on messages from our team being delivered, seen or responded to. For the entire sample, I collect survey data on use of mobile phones and Whatsapp to inform agricultural decisions.

[10] Not every farmer ended up selling their yields. In that case, I used the median price for that specific crop in that region. If no other farmer in that village sold the crop, then I used the next spatial unit, Mandal (sub-district), and used its median price.

5.2 Variables of interest

The baseline survey covers three broad domains. The first is typical demographic information about the farmers, along with control variables for the analysis, including information on age, education, household members, income, expenditure, assets etc. In addition, I collect some unique new demographic data on phone ownership, smartphone usage and access (since not everyone who uses a smartphone owns one, in a rural setting) and digital literacy, which includes familiarity with Whatsapp interface, similar to an approach by [Badrinathan, 2020]. The third domain covers standard agricultural questions including size of land holding, crops grown in the previous season, and intention to grow crops next season, input usage, costs and yields in previous season and so on. I also collect information on damage caused by pest and weather shocks. See table B7 for an overview of baseline characteristics.

6 Empirical Strategy

6.1 Identification

Identification comes from variation in individual exposure to and uptake of information based on treatment assignment. In other words, there is heterogeneity in information exchange and uptake based on the virtual or virtual + in-person interactions, leading to changing beliefs about using Whatsapp to inform agricultural decisions. As mentioned in section 5, the main outcome variables are farm revenues and expenditure. The main mechanisms I test are changes in farmer willingness to exchange relevant information with unknown farmers, and extent of participation on the whatsapp groups (for treated farmers), and adoption of Whatsapp for informing farming decisions. The main identification assumption in this analysis is that due to stratified, randomized treatment assignment, on average, the outcomes of farmers in control and treatment groups should be comparable if no one was treated.

Some potential threats to identification are as follows. First, repeated virtual interactions may lead to in-person connections, thereby conflating the two treatment arms. In this experiment, since virtual groups are randomly created from distant, non-neighboring villages, the possibility that virtual connections would turn into in-person connections in the first treatment arm does not pose a significant threat. In addition, there is a two-fold endogeneity with networks- first, in network formation, and second, participation in the network after it has been formed. Treatment assignment randomized and clustered at the village level takes care of the potential endogenous group formation since farmers are unable to choose which Whatsapp groups to be a part of. The latter, also known as Manski's reflection problem, posits that average outcome (app adoption in this context) in a group influences and is influenced by the outcome of individuals that comprise the group (also known as endogeneous effects), as well as average group characteristics (also known as exogenous affects) [Manski, 1993]. Since individuals are randomly assigned into Whatsapp groups, average group outcomes are not likely to affect each individual [Moffitt et al., 2001] [McNiven and Gilligan, 2012]. Also, the panel nature of the data will allow me to capture the lagged effect of peer beliefs on an individual's behavior, thereby addressing the potential simultaneity bias.

While attrition remains a concern, the time spent on Whatsapp groups for both treatments does not seem to systematically differ- indicating that similar exposure to treatment across the groups (see Fig. A.6). Lastly, common shocks to farmers located in the same geographic region can lead to another type of endogeneity that [Manski, 1993] referred to as correlated effects. This variation can be absorbed by using region-level fixed effects [McNiven and Gilligan, 2012].

6.2 Outcomes

To simply estimate the Intention-To-Treat estimate, I can run the following model:

$$y_{i,v,k,t} = \alpha + \beta Treat_{i,v,k,t} + \gamma Wave_t + \delta Treat_{i,v,k,t} \times Wave_t + \mu X_{i,v,k,t} + \rho_v + \epsilon_{i,v,k,t} \qquad (4)$$

Where $y_{i,v,t}$ is a set of agricultural outcomes including farm revenue, input costs, pesticide costs and area of crop lost to pests, for individual i, in village v in treatment k at time t. $X_{i,k,t}$ is a set of demographic controls of the individual, and ρ_h are village-fixed effects (to control for correlation effects mentioned in [Manski, 1993])

6.3 Intermediary Outcomes and Mechanisms

The proposed mechanism for the treatment effects is greater information exchange with unknown virtual peers, as well as increased participation on Whatsapp to inform agricultural decisions. I test this by running the following models:

$$exchange_{i,v,k,t} = \alpha + \beta Treat_{i,v,k,t} + \gamma Wave_t + \delta Treat_{i,v,k,t} \times Wave_t + \mu X_{i,v,k,t} + \rho_v + \epsilon_{i,v,k,t} \quad (5)$$

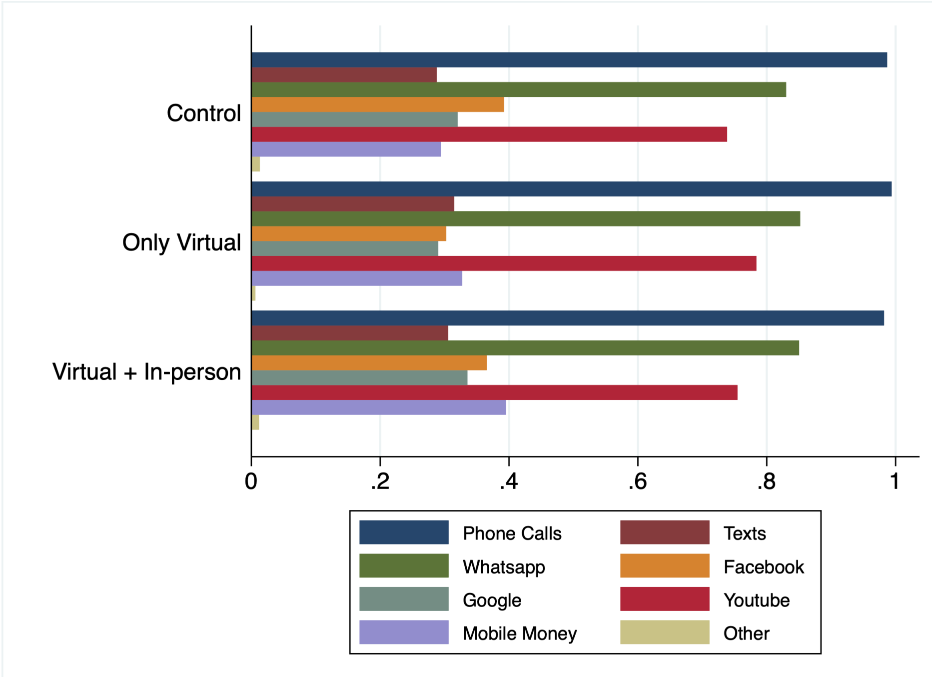
where $exchange_{i,v,k,t}$ is the willingness to exchange information with other farmers.

7 Results

7.1 Descriptives

Table B7 in the Appendix depicts the summary statistics of the sample by treatment. The sample is balanced across most demographic and agricultural variables in the baseline. Fig. 6 shows self-reported uses of smartphones at baseline, by treatment. In all three groups, most of the farmers used their smartphones for phone calls (99.6%), and Whatsapp (86.3%). There is also a surprisingly sizable proportion of farmers that use YouTube (76.2%). Other uses (7%) include other social media applications such as Instagram, as well as for children's education (e-learning).

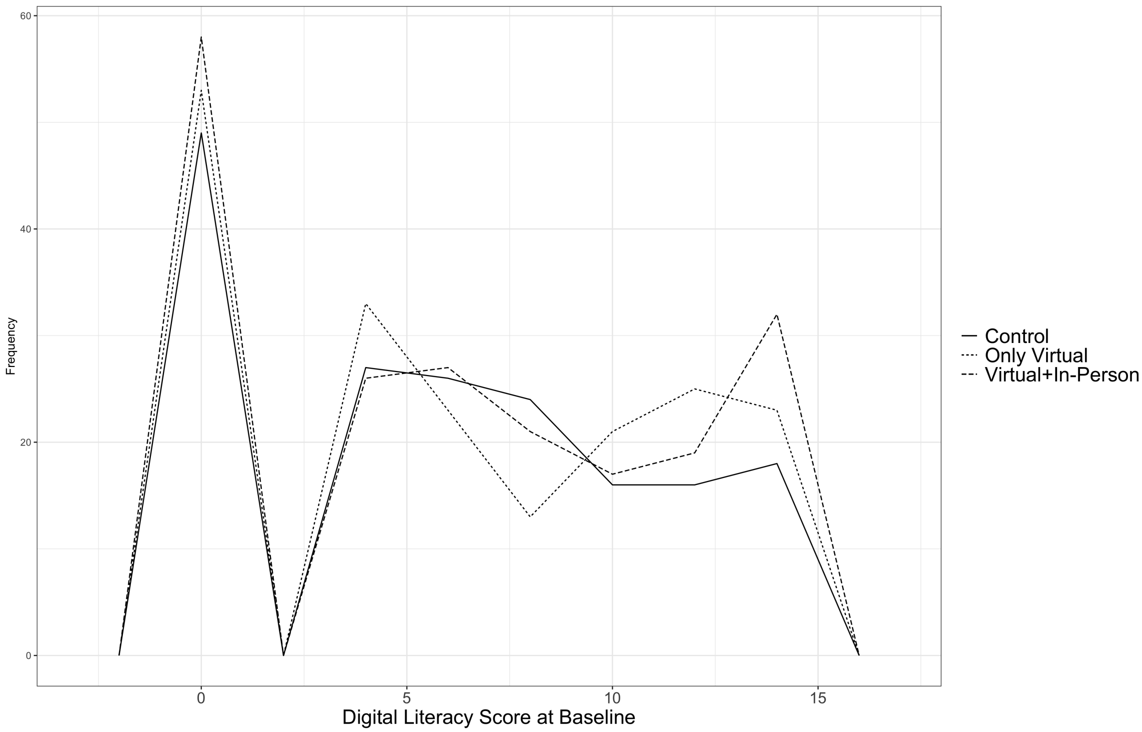
Figure 6: Range of Smartphone Uses



This graph depicts different ways in which the sample farmers used their smartphone at baseline (n=482).

Another metric for specifically measuring Whatsapp use is familiarity with the interface of phone application. I measure this through creating a score of digital literacy based on [Badrinathan, 2020]. Fig. 7 depicts baseline digital literacy score by treatment. The distribution of this score is skewed to the right, with most farmers scoring low, regardless of the treatment group. This indicates that despite high Whatsapp usage rates, farmers are using Whatsapp to view media or communicate without focusing on the technical aspects of the application.

Figure 7: Baseline Digital Literacy Score by Treatment



Note: This is a measurement of familiarity with digital messaging application such as Whatsapp. It includes an index of familiarity from 0 to 3 of Whatsapp features such as forwarding, blue tick, status, group and mute. Trends are similar across treatments.

7.2 Outcomes: Agriculture

Due to the variation in crops grown in the sample, it is not feasible to measure agricultural outcomes in terms of optimal input usage, since this would vary by crop. Instead, I measure the expenditure on inputs, especially pesticides, since a particular pest, *Uzi*, was of particular concern to the tomato farmers in the region. Table 1 shows the intent-to-treat effects on the farming investments including input expenditure and pesticide expenditure separately. All values are in Rs. 1000 and were standardized by area of land owned. Columns (1) and (2) depict ITT effects on input and pesticide expenditure without any controls. Column (3) and (4) include individual controls, and columns (5) and (6) include individual as well as mean group controls to address the exogenous effects stated in [Manski, 1993]. Controlling for the group means does not significantly change the outcomes,

Table 1: Differences in Farm Expenditure by Treatment at Endline

	(1) Input Exp.	(2) Pest. Exp.	(3) Input Exp.	(4) Pest. Exp.	(5) Input Exp.	(6) Pest. Exp.
Only Virtual	11.11^{**}	2.58^{*}	11.02^{**}	2.55^{*}	11.02^{**}	2.55^{*}
	(4.78)	(1.30)	(4.80)	(1.31)	(4.80)	(1.31)
Virtual + In-person	8.58^{*}	2.36^{**}	8.58^{*}	2.36^{**}	8.58^{*}	2.36^{**}
	(5.03)	(1.17)	(5.04)	(1.18)	(5.04)	(1.18)
Controls	×	×	✓	✓	✓	✓
Mean Group Controls	×	×	×	×	✓	✓
Baseline Control Mean	48.13	11.17	48.13	11.17	48.13	11.17
N	3249	3249	3246	3246	3246	3246

This table represents changes per acre owned in input expenditure (1) (3) and (5), pesticide expenditure (2),(4) and (6). Columns (3) and (4) include individual controls, and columns (5) and (6) include mean control values for each individual in a Whatsapp group controls, excluding that individual. All columns are in Rs. 1000 per acre. Outcome variables have been winsorized at 5-95%. Coefficients are from a regression of the dependent variable on the post-treatment indicator, as well as village and strata fixed effects. Standard error are in parentheses, and are clustered at the village level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

which is a testament to the random group assignment. Overall, there were positive and significant increases in both input expenditures (≈ Rs. 10,000), and pesticide expenditures (≈ Rs. 2,400) per acre, for both treatment groups relative to the control. The effects on the two treatment arms are not statistically significantly different from one another. This implies that the in-person component does not add any additional value to the input decision-making process of treated farmers.

I do not find any significant differences in the farm revenue of treated farmers (Table 2). This could be because of two reasons. First, the self-reported price data may be capturing noise.[11] To test this, I calculate the revenue using median prices and this does not change the results or reduce the noise. Second, benefits from greater farm investments may not reflect in revenues in the short-term. I also do not find significant effects on crop loss to pests as shown in Table .

[11] I use revenue instead of yield to compare outcomes for farmers that grow different crops and sell them in different units. I collected the per-kilogram price and per-unit yield and converted each unit to kilogram to calculate the revenue.

Table 2: Differences in Revenue by Treatment at Endline

	(1)	(2)	(3)
Only Virtual	-8.34	-8.37	-8.37
	(15.93)	(15.97)	(15.99)
Virtual + In-person	-19.06	-19.06	-19.06
	(18.88)	(18.90)	(18.92)
Controls	×	✓	✓
Mean Group Controls	×	×	✓
Baseline Control Mean	31.46	31.46	31.46
N	3249	3246	3246

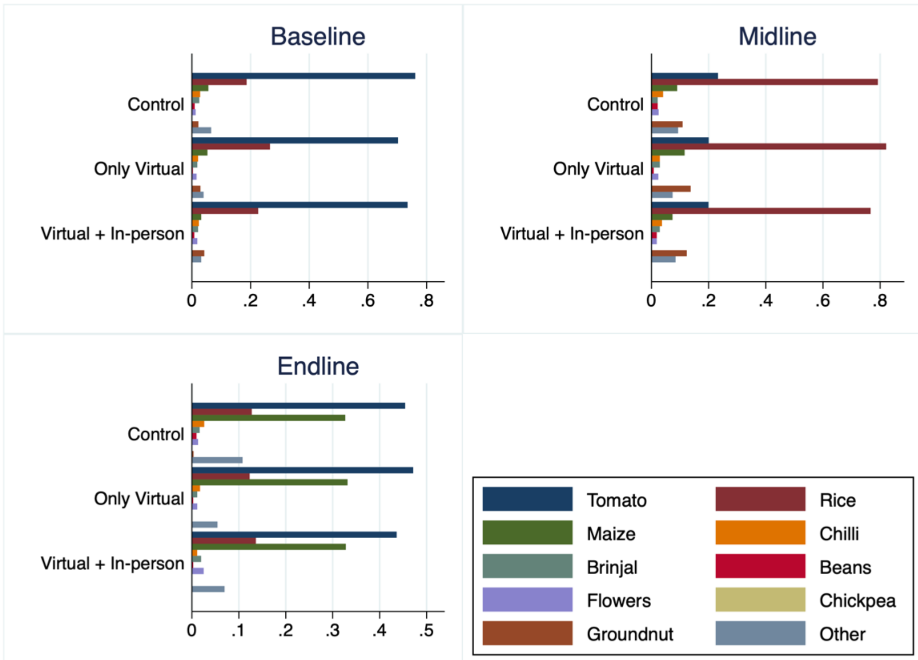
This table represents changes per acre owned in total farm revenue. Column (2) includes individual controls, and column (3) includes mean control values for each individual in a Whatsapp group controls, excluding that individual.. Both columns are in Rs. 1000 per acre. Outcome variables have been winsorized at 5-95%. Coefficients are from a regression of the dependent variable on the post-treatment indicator, as well as village and strata fixed effects. Standard error are in parentheses, and are clustered at the village level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

7.2.1 Heterogeneity in Tomato Farmers

Over the period of the study, a large number of the sample farmers grew tomatoes in at least one season (see Fig. 8). This is not surprising as Madanapalle, the study region, is known for its vast tomato market as mentioned in Section 2. Dividing the outcomes by tomato and non-tomato growing farmers yields the following results.[12]

[12] Here, I separate the sample into farmers that grew tomatoes in at least one of the three crop cycles, and those that never grew tomatoes in the duration of the study.

Figure 8: Variation in Crops by Treatment over Time



Note: This graph depicts variation in crop choice for the season preceding the mentioned survey wave. Tomato (baseline and endline) and rice (midline) were the main crops grown in the region, with over 80% farmers growing these crops before the first two waves, and 60% before the endline.

From Table 3, it is evident that farm expenses by tomato farmers are around double that of the pooled sample, indicating that the significant increases in input expenditures in Table 1 are likely being driven by tomato farmers. Within the two treatments, these increases are not significantly different. I check for differences in quantity of crop lost to pests (in acres). I do not find any significant differences in crop lost to pests for the pooled sample or tomato farmers (see Table B10). I fail to reject the hypothesis that yields and revenues for tomatoes farmers does not significantly change in the short term.

7.3 Mechanisms: Information Sharing

This section is divided into three parts: whether information is being shared, with whom it is being shared, and what kind of information is being shared (followed by a section on text analysis of the

Table 3: Differences in Farm Expenditure for Tomato Farmers by Treatment at Endline

	Pooled		Tomato Farmers	
	(1)	(2)	(3)	(4)
	Input Exp.	Pest. Exp.	Input Exp.	Pest. Exp.
Only Virtual	11.21**	2.64**	23.04***	4.41**
	(4.80)	(1.31)	(7.77)	(2.12)
Virtual + In-person	8.64*	2.40**	16.69**	3.71*
	(5.06)	(1.18)	(8.31)	(1.89)
Baseline Control Mean	25.87	4.65	44.44	8.75
N	3249	3249	1480	1480

This table represents changes per acre owned, for tomato farmers, input expenditure (1) and (3), and pesticide expenditure (2) and (4). All columns are in Rs. 1000 per acre. Coefficients are from a regression of the dependent variable on the post-treatment indicator, as well as village and strata fixed effects. Standard error are in parentheses, and are clustered at the village level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

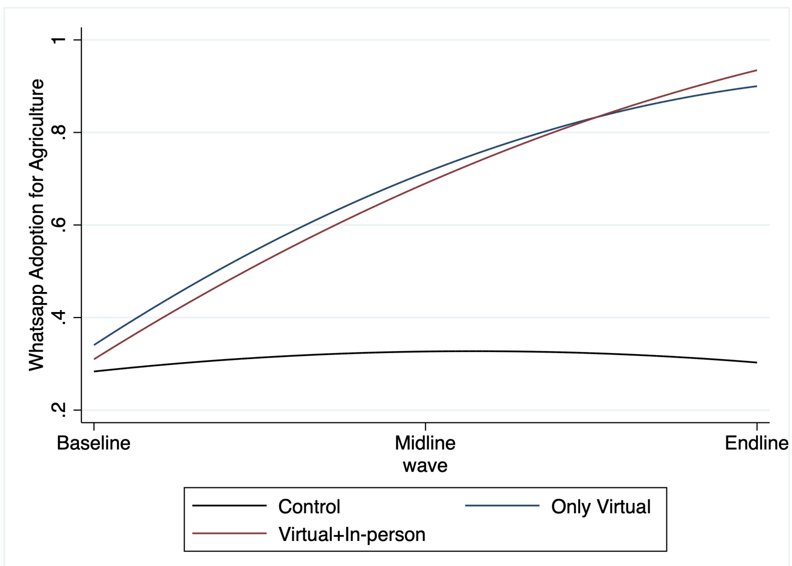
Whatsapp chats). To start off, I use an indicator variable to measure changes in farmer adoption of Whatsapp for agriculture. Table 4 shows the results by treatment. Farmers in both treatment groups were 50% significantly more likely to use Whatsapp for agriculture as compared to the control. Fig. 9 shows that in the treated groups, a greater proportion of farmers reported using Whatsapp for agriculture. But this is by design since the information intervention required treated farmers to use Whatsapp. Note that this fitted curve follows the bottom part of an S-shaped diffusion curve [Sunding and Zilberman, 2001]. However, this is self-reported data and may be affected by social desirability bias. To check for this, from the manually recorded Whatsapp data, we can track the participation of individual treated farmers over time. Fig. 10 shows the mean participation rates, including delivery, reading, responding to messages sent by the moderators in the groups, and whether the farmers initiated any agriculture related or unrelated messages on the groups. The rates are calculated by dividing the participation metric by number of total messages sent in a week to a group.

Table 4: Differences in Using Whatsapp for Agriculture at Endline

	(1)	(2)
Only Virtual	0.521***	0.528***
	(0.07)	(0.07)
Virtual + In-person	0.599***	0.601***
	(0.08)	(0.08)
Controls	×	✓
Baseline Control Mean	0.27	0.27
N	1402	1379

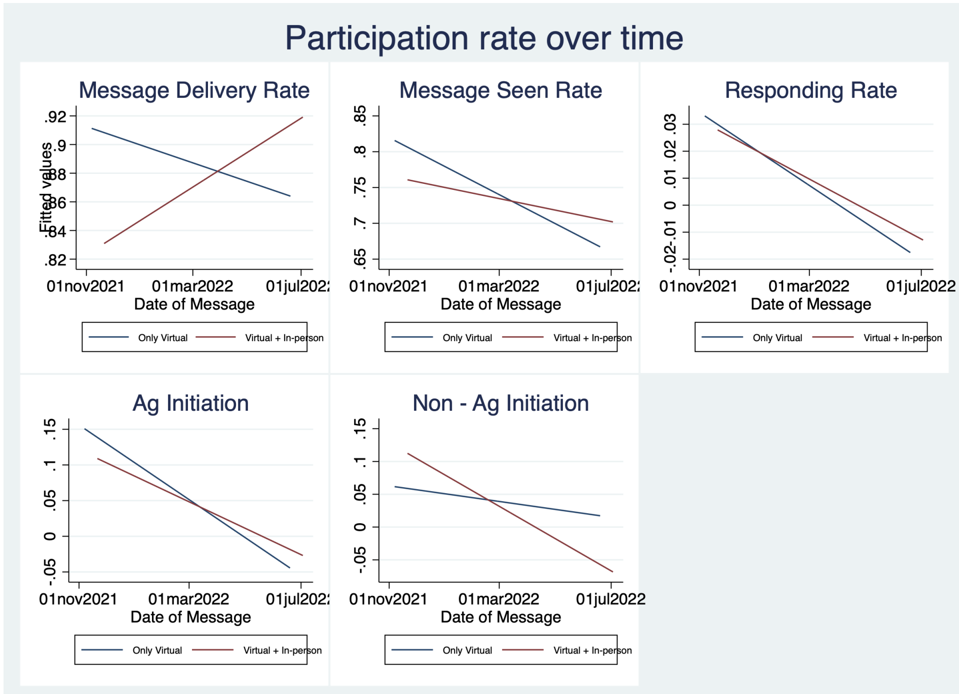
This table represents data from questions of the form ”In the last crop cycle, did you use Whatsapp for agricultural groups?”. The two columns are identical except controls are included in the second column. This data was collected in the second and third survey waves. Relative to the control, treated farmers report using Whatsapp more for participating in agricultural groups. Coefficients are from a regression of the dependent variable on the post-treatment indicator, as well as village and strata fixed effects. Standard error are in parentheses, and are clustered at the village level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 9: Proportion of Farmers using Whatsapp over Time



Note: Farmers in the treated groups are more likely to use Whatsapp for agriculture. The differences are significant between treatment and control, but not significant between the two treatments.

Figure 10: Whatsapp Participation Rates



Note: Overall, participation seems to decline over time for both treatments, and this decline is almost always less in the Virtual+In-person group.

To test with whom the treated farmers were sharing information, we gathered survey data on whether farmers were likely to discuss agriculture-related information with unknown farmers in-person or on Whatsapp. Results are shown in Table 5. While there are no significant changes for sharing information in-person or over Whatsapp for nearby farmers, significantly more farmers in the only virtual treatment reported sharing information with farmers not geographically close to them over Whatsapp. As explained in section 4, this likely entails that the in-person component in the second treatment arm seems to have crowded out willingness to exchange information online. It is worth noting that the coefficients for the two treatments are not significantly different from each other.

The above table shows that treated farmers are significantly more willing to discuss agriculture on Whatsapp with other farmers, regardless of their location. While farmers in the only virtual

Table 5: Differences in Information Sharing Behaviour by Endline

	Near		Far	
	(1)	(2)	(3)	(4)
	In-person	Whatsapp	In-person	Whatsapp
Only Virtual	0.033	0.165***	0.084*	0.204***
	(0.04)	(0.06)	(0.05)	(0.05)
Virtual + In-person	0.048	0.123**	0.069	0.110**
	(0.04)	(0.05)	(0.05)	(0.05)
Baseline Control Mean	0.81	0.22	0.31	0.04
N	3247	1244	3247	1244

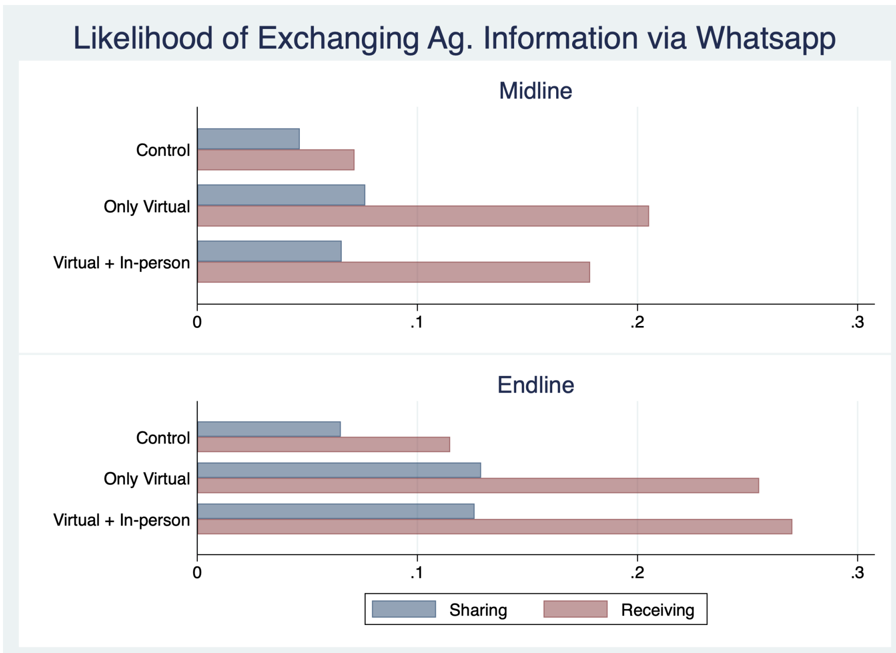
This table represents the proportion of farmers in each treatment group willing to exchange agricultural information with farmers from near and far villages. The dependent variables represent answers to yes or no questions of the form "Do you discuss agricultural issues with nearby farmers in the neighboring village or farther farmers (within the same mandal/outside the mandal) in-person or over Whatsapp?". Coefficients are from a regression of the dependent variable on the post-treatment indicator, as well as village and strata fixed effects. The results do not change significantly when controls are added (see Table B8). Standard error are in parentheses, and are clustered at the village level.

group are more willing than the farmers in the in person group to share agriculture information with farmers on Whatsapp, this difference is significant at the 10% level for distant farmers, not for nearby farmers.

These results are consistent with self-reported Whatsapp usage for agriculture in Table 4. In any case, these results indicate a change in perspective, or increased familiarity, for using Whatsapp as an input in the farmer production function as in equation 3. As per the hypothesis in section 4, this indicates that the in-person component may have crowded out online participation in the second treatment arm, for instance, if farmers believe they can share the information in-person and need not put the effort to share it online.

This does not tell us much about the content of information being shared. To gain a better understanding of the content of information shared, I collected survey data on whether farmers shared 'tips' to prevent crop loss with other farmers. I find that over time, treated farmers were more open to sharing and receiving agricultural information (Fig. 11). This figure depicts the differences in sharing behavior across treatment, these differences are significant (Table B9). It should be noted that by design, the treated groups enabled farmers to receive (though not necessarily share) more information than the control groups. While the slightly higher rates of information exchange in the only-virtual group hint towards potential substitution effects of the in-person component, the differences between the two treatment arms is not significant.

Figure 11: Sharing/Receiving Agri. Information Over Time



Note: This graph depicts varying information exchange behavior across treatment groups over both midline and endline waves. This data is based on questions of the form "In the last crop cycle, did you 1) share and 2) receive tips to prevent crop damage or other agriculture-related information with other farmers on Whatsapp, other than our study Whatsapp group?". Relative to the control, treated farmers report using Whatsapp more for both receiving and sharing agricultural information.

To get an even better grip on details of text exchange, I conducted a rigorous text analysis from the farmer Whatsapp chats as explained below.

7.4 Text Analysis

7.4.1 Cleaning

I first exported their group chat histories of farmers who complied with the treatment, then translated them through google translate. I also converted the voice notes to text in the translation process. Then, I converted each of the 24 translated documents to text files (with standard encoding UTF-8) for readability for text mining. Despite the UTF-8 coding, there were some texts that were not readable due to spacing issues in the date, which I manually fixed. I removed all non-author messages including Whatsapp notes on encrypted data or when a group member was added. Using R, I was able to create a data frame of the text files, with separate columns for date, sender and text. This allowed me to measure frequency of total messages per group, per sender, content of the text messages send including frequency of certain words. For the total frequency graph, I appended the chat histories (12 each) within each treatment arm and removed messages sent by enumerators to observe only-farmer participation in the group. To eyeball each Whatsapp conversation, I created word clouds for the most and least active groups as shown in Fig. 12.

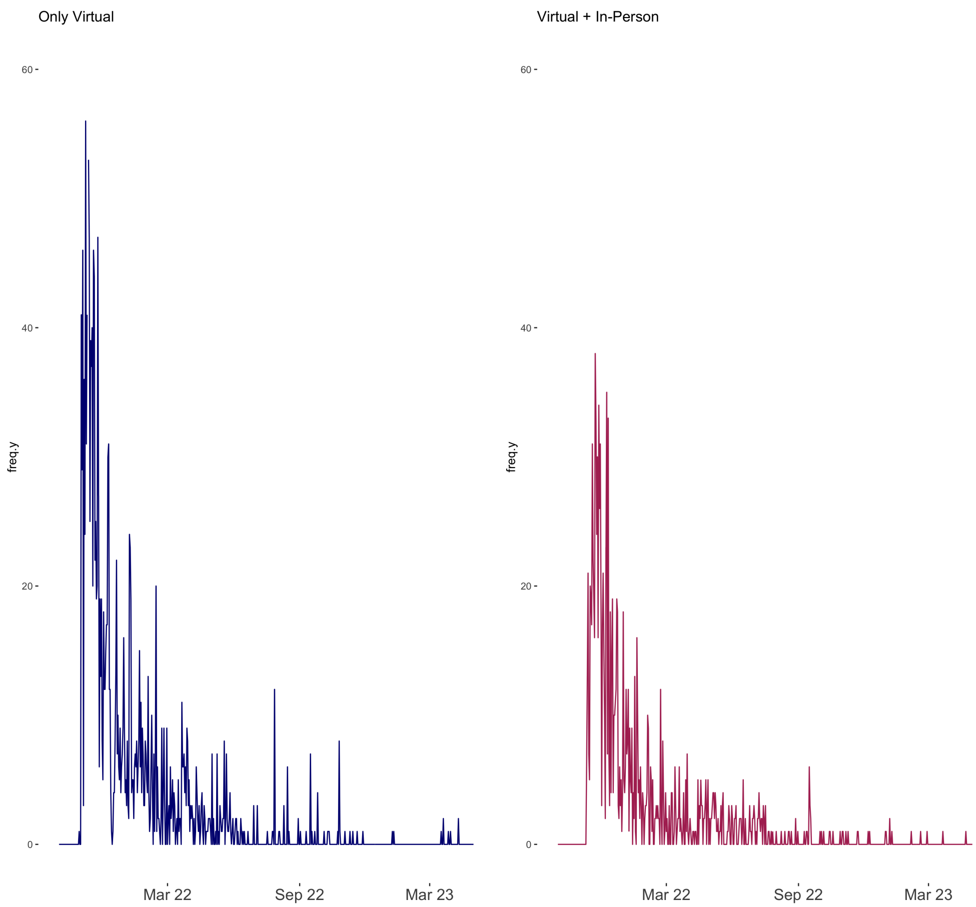
Figure 12: Word cloud of words used most in Whatsapp Group A1

Note: The size of the word is proportional to the frequency. For this image, the transcripts were translated to English and then coded.

7.4.2 Frequency

Figure 13 shows daily total frequency of messages exchanged in both groups. In the Only Virtual treatment, initial participation was higher.

Figure 13: Daily Frequency of Texts



Note: This graph depicts daily text messages sent on each set of the 12 Whatsapp groups in each treatment. Participation for both groups was higher at the beginning of the project.

When we split the graph by groups, we can see that the overall participation declined over time for all groups. Meeting in-person did not seem to systematically increase participation on the Virtual+In-Person Groups as shown in Fig. 15.

Figure 14: Frequency of Daily Messages per Whatsapp Group (Only Virtual)

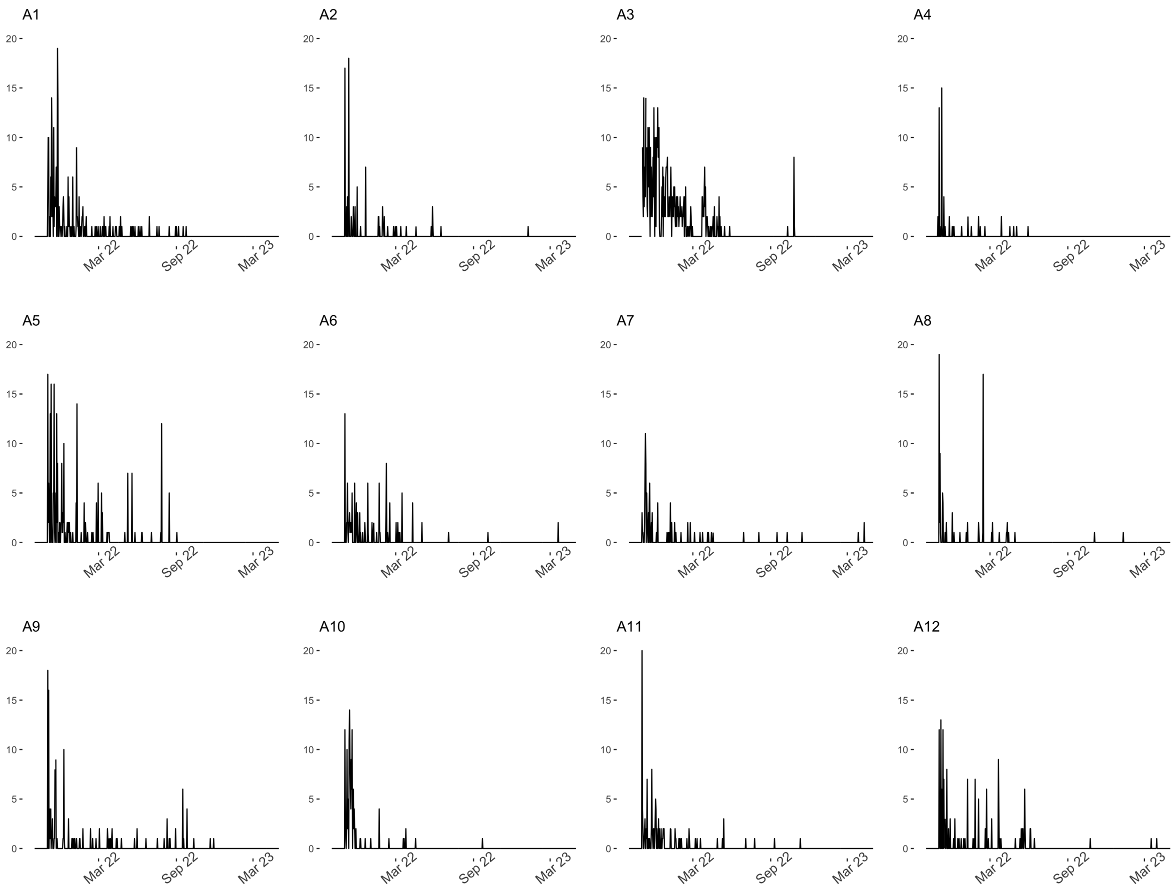
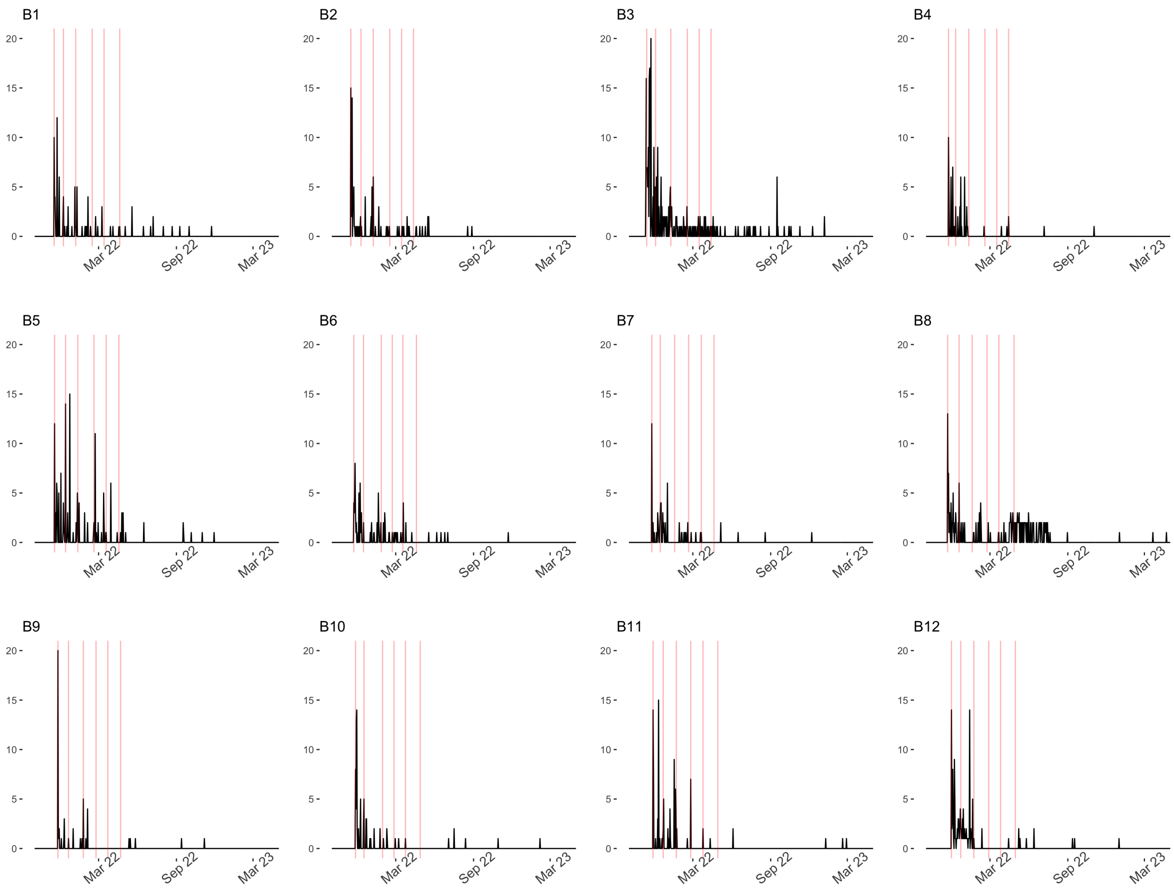


Figure 15: Frequency of Daily Messages per Whatsapp Group (Virtual+In-person)



Note: The red lines indicate 6 monthly meetings held on different days for each group.

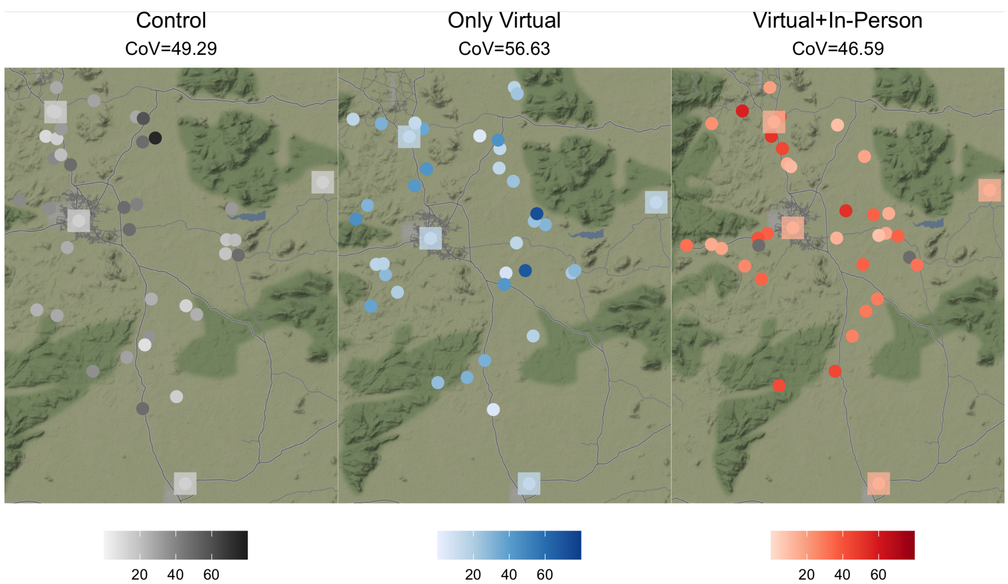
7.4.3 Content

I classified the group conversations into 7 main categories, namely- Agriculture, Market, Pests, Inputs, Crop Damage, Greetings and Compliments. This does not include images or other media shared. The following table shows how the content of information varied across both treatments:

Table 6: Frequency of Topics Discussed

Category	Freq (Only Virtual)	Freq (Virtual+In-Person)
agriculture	128	127
compliments	87	41
input	86	86
crop damage	82	48
pests	36	49
market	30	222
greetings	21	9

Figure 16: Spatial Variation in Tomato Price/Kg (in Rs.)



Note: This graph depicts spatial variation in tomato prices at the endline, by treatment. Coefficients of variation are listed at the top. Darker colors depict higher prices. The treatment with both Whatsapp and in-person intervention has the least price dispersion at endline.

Decomposing the content further by group, we can see that regardless of groups, markets, pests, inputs and crop damage are topics that were consistently discussed across all groups. Additionally, while not significantly different statistically, the higher communication about prices in the second treatment in Table 6 reflect in a lower coefficient of variation in tomato prices among farmers in the second treatment group. Comparing this to self-reported Whatsapp usage, it seems that farmers in the virtual group reported to exchange more agricultural information online,whereas text analysis reveals that there are more instances of social greeting exchanges in this arm, and more market-relevant information in the in-person arm.

This discrepancy between self-reported data and text analysis highlights a difference in farmer perceptions of what constitutes agricultural information exchange - sharing only agriculture-related information with other farmers or any information as long as it is being shared in a group of farmers.

Figure 17: Frequency of Category of Words (Only Virtual)

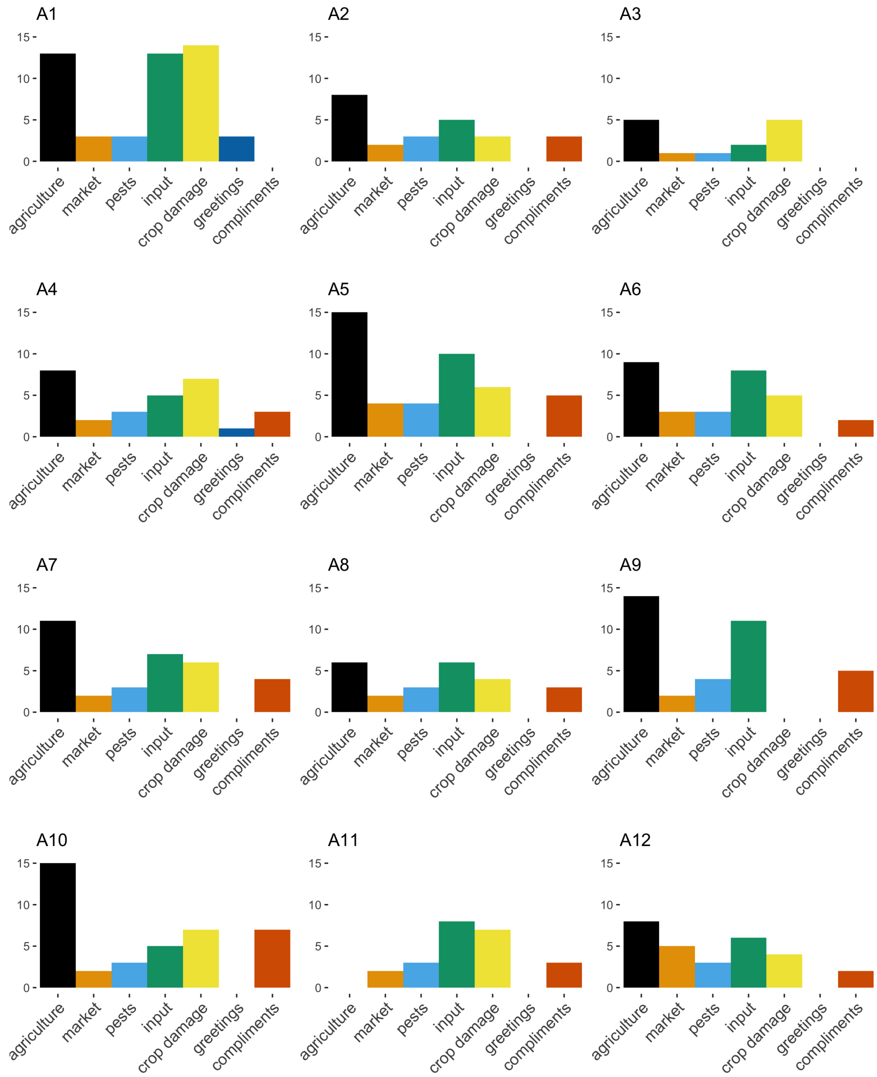
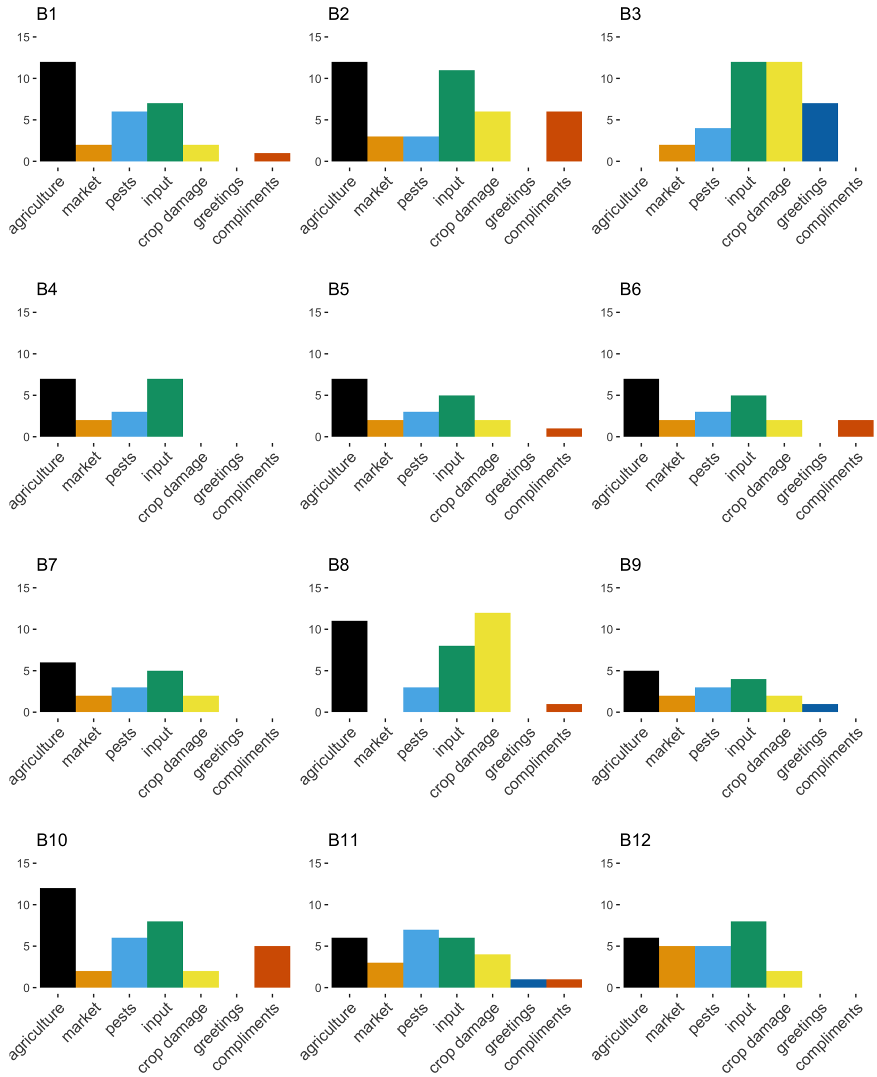


Figure 18: Frequency of Category of Words (Virtual+In-person)



8 Discussion

This study aims to measure the effects of online information exchange in exogenously created farmer Whatsapp groups in rural India. Since most of our online networks are self-selected, the online groups in this experiment provide useful and unique insights on how information travels in new networks when costs of sharing are negligible. I find that farmers in treated villages, when connected online but not induced to meet each other in-person, are more likely to share information with unknown farmers over Whatsapp. The content of information shared ranges from tips to prevent crop loss due to weather or pest shocks, to prices, inputs and other information. While treated farmers receive information being shared on Whatsapp, very few are willing to initiate conversations from their end. This can be due a gamut of factors, including literacy, digital literacy, social status, network centrality in pre-existing networks and so on.

I also find that farmers in treated villages are more likely to invest more in their farms, by an average of Rs. 10,000. This effect is primarily driven by tomato farmers in the region, who are currently dealing with a new pest, Uzi, in the region. The greater farm expenditures do not seem to translate to higher farm revenues in the short run. This effects are driven by changes in beliefs about farmer information sharing on Whatsapp. Treated farmers are significantly more likely to exchange agricultural information on Whatsapp. Instances of these exchanges are verified through text analysis.

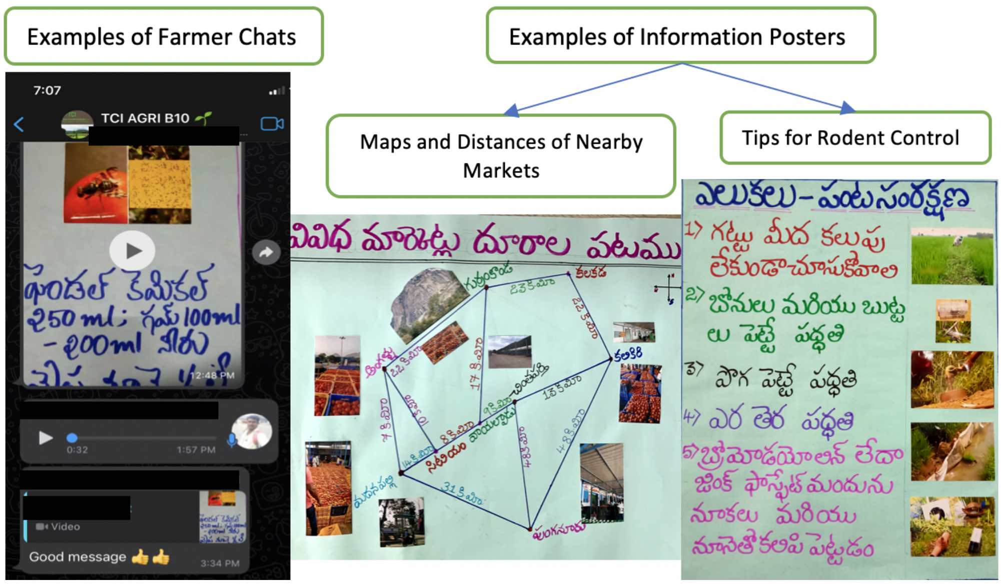
From a policy perspective, this study highlights the use of a seemingly cost effective method to deal with information constraints in the developing world. Online networks provide opportunities to disseminate information and facilitate learning in not just agriculture communities, but any other realm that faces the challenges of information constraints. However, it is not just the quantity of information shared, but also the quality that affects livelihoods. Additionally, the effects of information sharing can take months before reflecting in behavioral change.

Further work in this study involves utilizing the Whatsapp data to get treatment effects for compliers, as well as measuring spillover effects to farmers that do not own smartphones, and conducting more heterogeneity analysis by farmers' age and wealth quintile, and additional robustness checks.

Appendices

A Treatment

Figure A.1: Treatment intervention



Note: Farmer chat groups were created for both treatment arms. Information posters were presented in person to Group B, and were shared online for both groups.

Figure A.2: In-person intervention

Figure A.3: Number of Members in each of the 24 Whatsapp groups

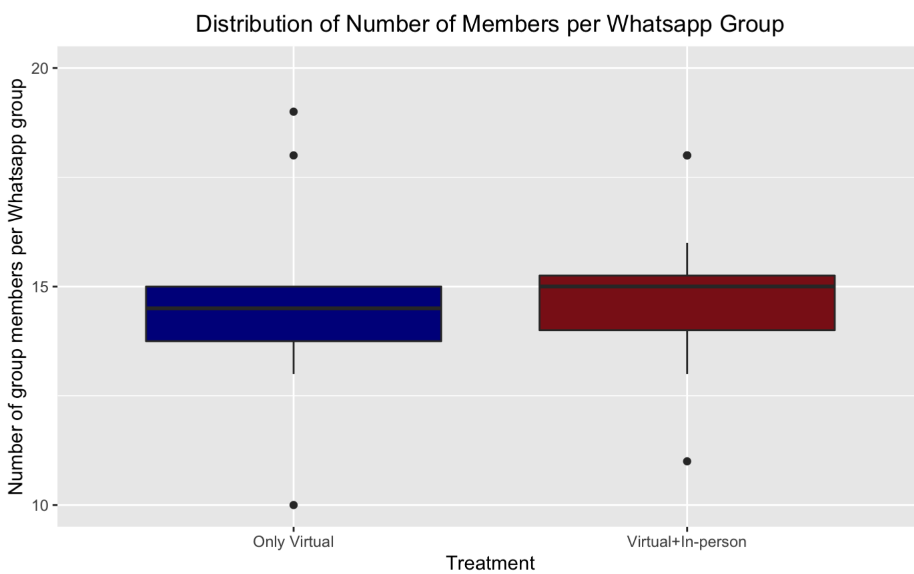


Figure A.4: Distance of Villages to Madanapalle

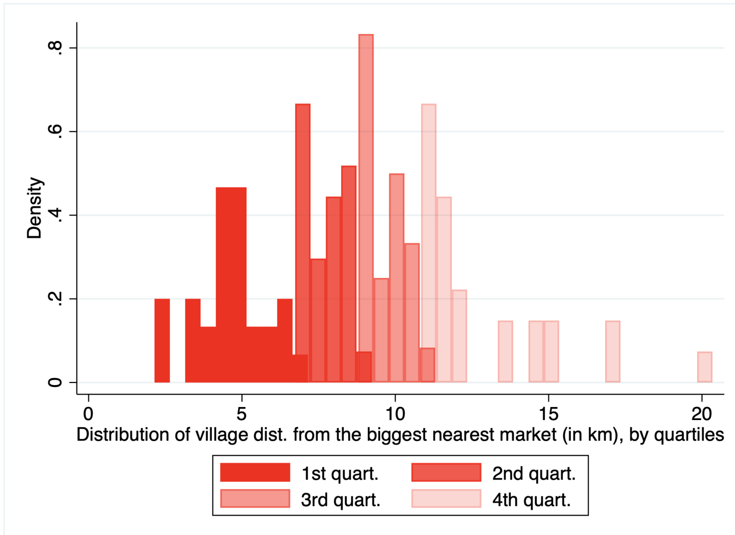


Figure A.5: Timeline of study

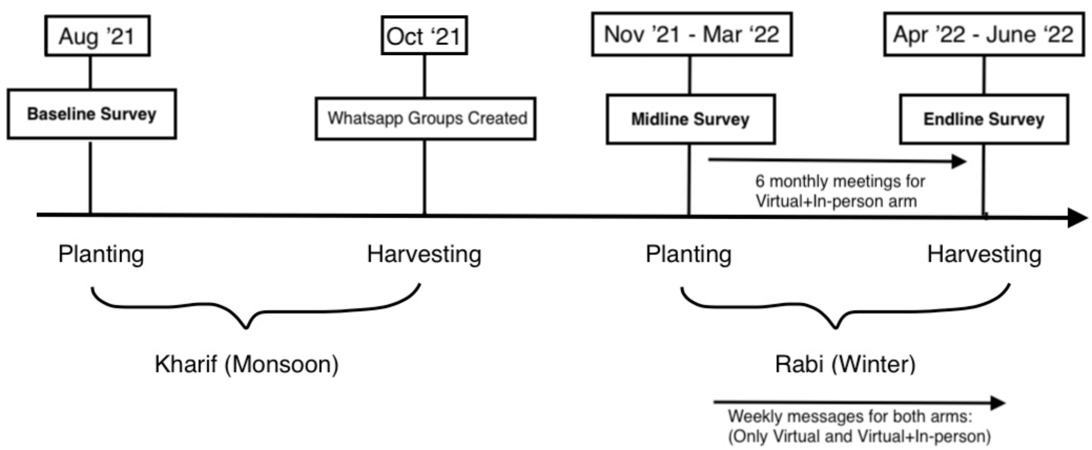
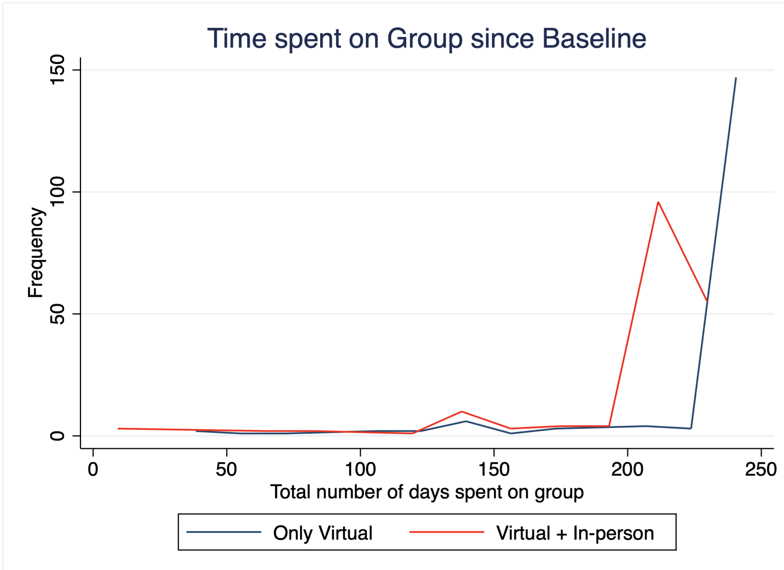


Figure A.6: Exposure to Treatment



Note: Majority of farmers on both treatments stayed on the groups for more than 200 days. This data does not include farmers who withdrew from the survey in the midline or endline.

Table B7: Baseline Summary Statistics

	Control		Treated		Diff
	mean	sd	mean	sd	p
Demographics					
Age	47.80	11.67	47.79	11.37	0.99
Gender	0.97	0.17	0.96	0.20	0.38
Education					
No educ.	0.20	0.40	0.20	0.40	0.88
Elementary	0.18	0.38	0.20	0.40	0.44
Middle School	0.47	0.50	0.46	0.50	0.65
High School and above	0.15	0.36	0.14	0.35	0.70
No. of Household members	4.72	2.15	5.06	2.21	0.02
Annual HH expenditure (in Rs. 1000)	97.96	94.73	85.33	63.27	0.01***
Annual HH income (in Rs. 1000)	79.80	75.43	76.27	77.13	0.53
Nearest phone tower (km.)	2.45	2.28	2.33	2.29	0.45
Phone Access					
Phone ownership	0.97	0.18	0.94	0.24	0.06
No. of phones in HH	2.59	5.55	2.30	1.00	0.18
No. of smartphones in HH	1.39	0.94	1.36	1.00	0.63
Digital literacy score	6.41	4.95	6.72	5.29	0.50
Hours spent on smartphone					
Less than 5	0.82	0.39	0.84	0.37	0.51
5-10	0.12	0.33	0.12	0.32	0.83
More than 10	0.06	0.24	0.04	0.20	0.40
Agriculture					
Land owned (in Acres)	3.50	3.51	3.38	4.68	0.67
Last season's input cost (in Rs. 1000)	123.05	121.80	100.23	111.08	0.00***
Last season's pesticide cost (in Rs. 1000)	27.70	32.87	23.45	31.54	0.05
Last season's crop lost to pests (in Acres)	0.27	0.58	0.21	0.48	0.046
N	322		761		1083

This table shows the differences in means across control and treated farmers. For continuous variables, a two sample t-test is used and for categorical variables, a Pearson's chi-squared test is used. Digital literacy score is measured on a 0-15 scale of familiarity with the interface of digital messaging apps such as Whatsapp.

B Results

Table B8: Differences in Information Sharing Behaviour by Endline

	Near		Far	
	(1)	(2)	(3)	(4)
	In-person	Whatsapp	In-person	Whatsapp
Only Virtual	0.017	0.064	0.060	0.126**
	(0.03)	(0.09)	(0.06)	(0.06)
Virtual + In-person	-0.001	0.019	0.069	0.031
	(0.03)	(0.08)	(0.06)	(0.06)
Baseline Control Mean	0.81	0.22	0.31	0.04
N	3247	1244	3247	1244

This table represents the proportion of farmers in each treatment group willing to exchange agricultural information with farmers from another mandal. The dependent variables represent answers to yes or no questions of the form "Do you discuss agricultural issues with nearby farmers (in the same or neighboring village) or farther farmers (within the same mandal/outside the mandal) in-person or over Whatsapp?". Coefficients are from a regression of the dependent variable on the post-treatment indicator, demographic controls, as well as village and strata fixed effects. Standard error are in parentheses, and are clustered at the village level.

Table B9: Differences in Sharing/Receiving Agricultural Information by Treatment at Endline

	(1) Sharing	(2) Receiving
Only Virtual	0.064***	0.140***
	0.022	0.027
Virtual + In-person	0.061***	0.155***
	0.020	0.029
Midline Control Mean	0.065	0.114
N	3249	3249

This table represents data from questions of the form "In the last crop cycle, did you 1) share and 2) receive tips to prevent crop damage or other agriculture-related information with other farmers on Whatsapp, other than our study Whatsapp group?". This data was collected in the second and third survey waves. Relative to the control, treated farmers report using Whatsapp more for both receiving and sharing agricultural information. Coefficients are from a regression of the dependent variable on the post-treatment indicator, as well as village and strata fixed effects. Standard error are in parentheses, and are clustered at the village level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

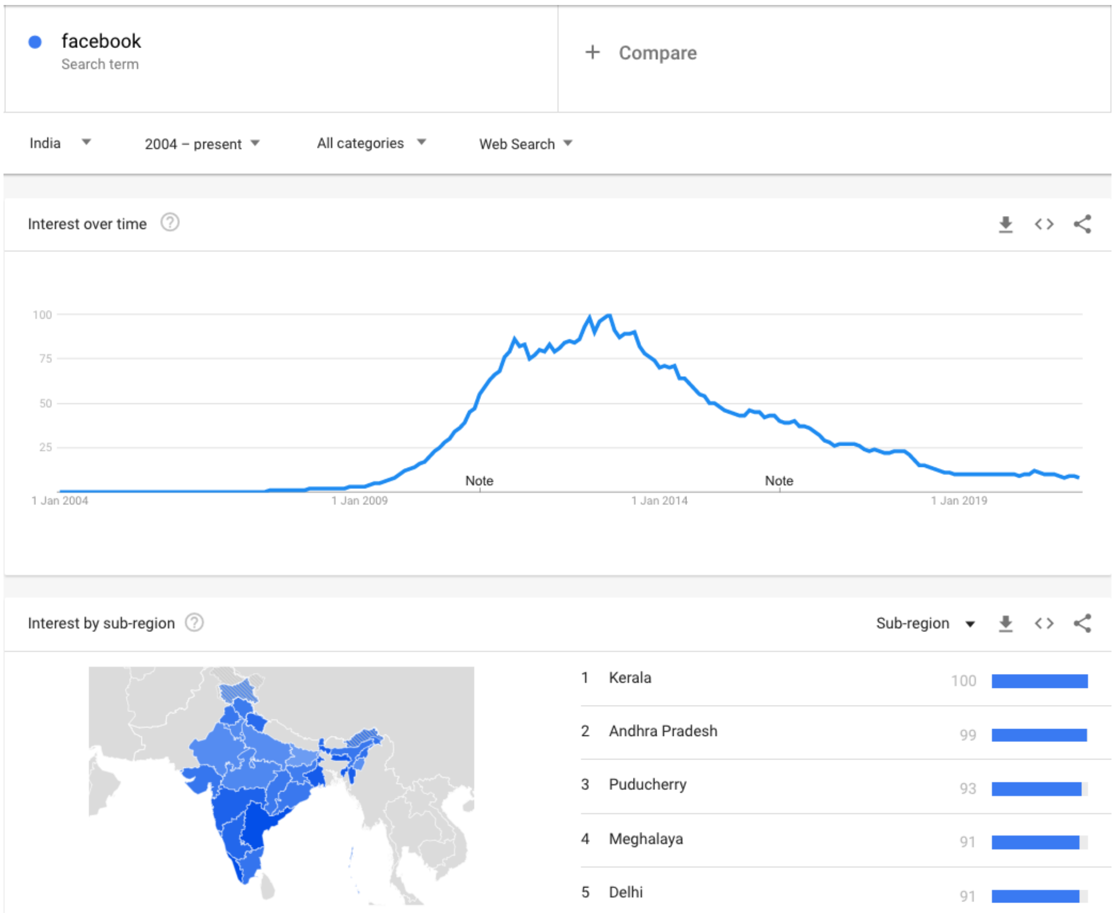
Table B10: Differences in Pest Loss by Treatment at Endline

	Pooled (1)	Tomato Farmers (2)
Only Virtual	0.02	0.02
	(0.02)	(0.03)
Virtual + In-person	0.00	-0.02
	(0.03)	(0.05)
Baseline Control Mean	0.11	0.12
N	3237	1474

This table represents changes in crop lost to pests, standardized by land ownership. Column (2) is identical to (1) but for tomato farmers. Both columns are in proportions of acres lost. Coefficients are from a regression of the dependent variable on the post-treatment indicator, as well as village and strata fixed effects. Standard error are in parentheses, and are clustered at the village level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C Miscallaneous

Figure C.1: Growth in searches for Facebook



Note: Interest in Facebook faded over time, in contrast to Whatsapp that has steadily increased with time.

This script is for the enumerators to follow-up with farmers based on the extent of Whatsapp participation. Please follow the script so that we can maintain consistency across groups.

STEP 1: Call farmers and read from the script. In case farmers do not respond by 6pm today, follow step 2.

STEP 2: Send voice notes with the same script.

Purpose	English	Telugu
No Participation	Namaste Anna, I am ____ speaking from TCI. We are looking forward to getting to know farmers like you and are hoping you could respond to the introduction message in the group? If you have any questions, please let me know.	నమస్తే అన్నా, నేను TCI నుండి __________ మాట్లాడుతున్నాను. మీలాంటి రైతుల గురించి తెలుసుకోవాలని మేము ఎదురుచూస్తున్నాము మరియు గ్రూప్‌లో మీ గురించి పరిచయం చేసుకోవడానికి ఒక మెసేజ్ పెట్టాము దానికి దయచేయి రిప్లై ఇస్తారని ఆశిస్తున్నాము. మీకు ఏవైనా ప్రశ్నలు ఉంటే, దయచేసి నాకు తెలియజేయండి
Initiation/ response (Agriculture relevant)	Namaste Anna, thank you so much for your participation in the group! Please feel free to share more pictures or information of your farming practices so other farmers can know and learn more. We will also be sharing more information with you soon!	నమస్తే అన్నా, మీరు గ్రూప్‌లో పాల్గొన్నందుకు చాలా ధన్యవాదాలు! దయచేసి మీ వ్యవసాయ పద్ధతులకు సంబంధించిన మరిన్ని ఫోటోలు లేదా సమాచారాన్ని పంచుకోండి, తద్వారా ఇతర రైతులు కూడా మరింత తెలుసుకోవచ్చు మరియు మరింత నేర్చుకోవచ్చు . మేము త్వరలో మీతో మరింత సమాచారాన్ని కూడా అందిస్తాము.
Thank you for participation (Agri related)	Namaste Anna, thank you so much for your participation in the group!	నమస్తే అన్నా, మీరు గ్రూప్‌లో పాల్గొన్నందుకు చాలా ధన్యవాదాలు!
Initiation/ response (Not relevant to Agriculture)	Namaste Anna, thank you so much for your participation in the group! Please feel free to share more pictures or information of your farming practices so other farmers can know and learn more. We will also be sharing more information with you soon! We are using this group only for agriculture related topics. Thank you again for your	నమస్తే అన్నా, మీరు గ్రూప్‌లో పాల్గొన్నందుకు చాలా ధన్యవాదాలు! దయచేసి మీ వ్యవసాయ పద్ధతులకు సంబంధించిన మరిన్ని ఫోటోలు లేదా సమాచారాన్ని పంచుకోండి, తద్వారా ఇతర రైతులు కూడా మరింత తెలుసుకోవచ్చు మరియు మరింత నేర్చుకోవచ్చు . మేము త్వరలో మీతో మరింత సమాచారాన్ని కూడా అందిస్తాము.మేము ఈ గ్రూప్ ను వ్యవసాయ సంబంధిత అంశాలకు మాత్రమే ఉపయోగిస్తున్నాము. మీ

	time!	యొక్క సమయం కొరకు మళ్ళీ ధన్యవాదాలు!
Initiation with selfies pictures	Thank you ___ anna for sharing your picture- it is good to see you!	మీ చిత్రాన్ని పంచుకున్నందుకు ధన్యవాదాలు ___ అన్నా- మిమ్మల్ని చూడటం ఆనందంగా ఉంది!
For answering questions	Thank you anna for asking the question.. Can any other group member answer this question?	ప్రశ్న అడిగినందుకు ధన్యవాదాలు అన్నా.. ఈ ప్రశ్నకు ఇతర గ్రూప్ సభ్యులెవరైనా సమాధానం చెప్పగలరా?
Input procurement - seeds and saplings	Thank you! Could you please share with us and other farmers where you get your seeds or saplings from? And what do you think of the quality?	అందరికి ధన్యవాదాలు! మీరు మీ విత్తనాలు లేదా మొక్కలు ఎక్కడ నుండి తెస్తారో దయచేసి మాతో మరియు ఇతర రైతులతో పంచుకోగలరా? మరియు వాటి నాణ్యత గురించి మీరు ఏమనుకుంటునారు?
Who created the group	TCI Agriculture,	TCI వ్యవసాయం
Rain crop damage	Hello, how are you all? Hope you and your families are safe during the rain storm. Hope your crops are also not damaged. In case you saw crop damage, do you have any tips to share with the group?	హలో అన్న, మీరందరూ ఎలా ఉన్నారు? వర్షపు తుఫాను సమయంలో మీరు మరియు మీ కుటుంబాలు సురక్షితంగా ఉన్నారని ఆశిస్తున్నాము. మీ పంటలు కూడా నష్టపోకూడదని ఆశిస్తున్నాము. మీరు పంట నష్టాన్ని చూసినట్లయితే, వాటిని తగ్గించటానికి ఏవైనా చిట్కాలు ఉంటె గ్రూప్ వారితో పంచుకోమని కోరుచున్నాము
Valuable info sharing	Thank you so much for sharing that useful information Anna, it is really helpful for our study.	ఆ ఉపయోగకరమైన సమాచారాన్ని పంచుకున్నందుకు చాలా ధన్యవాదాలు అన్నా, ఇది మా అధ్యయనానికి నిజంగా ఉపయోగకరంగా ఉంది
Clarifying no benefits	Anna we just want to clarify that the purpose of these groups is to share valuable information between farmers to improve yield and minimize crop damage. Unfortunately we are not able to provide monetary assistance as this study is for education purpose.	అన్నా, దిగుబడిని మెరుగుపరచడానికి మరియు పంట నష్టాన్ని తగ్గించడానికి రైతుల మధ్య విలువైన సమాచారాన్ని పంచుకోవడం ఈ గ్రూప్ యొక్క ఉద్దేశమని మేము స్పష్టం చేయాలనుకుంటున్నాము. మేము ఈ అధ్యయనం విద్య ప్రయోజనం కోసం చేసినందున ఎటువంటి ఆర్థిక సహాయం అందించలేకపోతున్నాము.

Farmers who have used Plantix	That's great anna! Could you tell us more about where you learned how to use it? If there are any useful links or resources, please feel free to share on this group for other farmers.	చాలా మంచింది అన్నా! మీరు దీన్ని ఎలా ఉపయోగించాలో ఎక్కడ నేర్చుకునారో మాకు మరింత చెప్పగలరా? ఏదైనా ఉపయోగకరమైన లింక్‌లు లేదా వనరులు ఉంటే, దయచేసి ఇతర రైతుల కోసం ఈ గ్రూప్‌లో పంచుకొండి
Introduce yourself in group B	Hello! For those farmers who were not able to join us for the lunch this month, could you please introduce yourself, the crop you are growing, and which village you're from?	హలో! ఈ నెలలో మధ్యాహ్న భోజనానికి మాతో చేరలేకపోయిన రైతులు, దయచేసి మీ గురించి, మీరు పండిస్తున్న పంట మరియు మీరు ఏ గ్రామానికి చెందినవారో తెలియజేయగలరా?
Crop damage due to unnatural calamity	Really sorry for your crop loss Anna. This is the first time in so many years that rain has affected this area so badly.	మీ పంట నష్టానికి చాలా బాధగా ఉంది అన్నా. ఇన్ని సంవత్సరాలలో మొదటి సారి ఈ ప్రాంతంలో వర్షాలు ఇంత తీవ్రంగా ప్రభావితం చూపింది.
If farmer thanks for creating group	Anna it is our pleasure to get to know you. We hope we can share valuable information with you and learn from you during this project. Thank you 🙏	నమస్తే అన్నా మీ గురించి తెలుసుకోవడం మాకు ఆనందంగా ఉంది. ఈ ప్రాజెక్ట్ సమయంలో మేము మీతో విలువైన సమాచారాన్ని పంచుకోగలమని మరియు మీ నుండి నేర్చుకోగలమని మేము ఆశిస్తున్నాము. ధన్యవాదాలు🙏
To ask for youtube link sharing to prevent crop damage due to rain	Namaste anna, do any of you follow any youtube channels to prevent crop damage due to rain? If you have any useful videos please share here. Thank you!	నమస్తే అన్నా, వర్షం వల్ల పంట నష్టం జరగకుండా మీలో ఎవరైనా యూట్యూబ్ ఛానెల్స్ చూస్తారా? మీకు ఉపయోగకరమైన వీడియోలు ఏవైనా ఉంటే దయచేసి ఇక్కడ షేర్ చేయండి. ధన్యవాదాలు!
Input procurement - fertilizers/pesticides	Namaste anna, could you please share with us and other farmers where you get your fertilizers and pesticides from in this area? We can compare prices and qualities across villages	నమస్తే అన్నా, మీరు ఎరువులు మరియు పురుగుమందులు ఎక్కడ నుండి తీసుకుంటారో దయచేసి ఈ గ్రూప్‌లో మాతో మరియు ఇతర రైతులతో పంచుకోగలరా? మనం దుకాణాల్లోని ధరలు మరియు నాణ్యతలను సరిపోల్చవచ్చు.
Acknowledging responses	Thank you for your response ___, ___ and ___ Anna!	మీ ప్రతిస్పందనకు ధన్యవాదాలు ___, ___ మరియు ___ అన్నా!

Question on which mandal	Namaste anna, we hope this information is useful for you. Can you tell us which mandal you are from?	నమస్తే అన్నా, ఈ సమాచారం మీకు ఉపయోగకరంగా ఉంటుందని మేము ఆశిస్తున్నాము. మీరు ఏ మండలానికి చెందిన వారో అందరికీ చెప్పగలరా?
Tips for Oozi 1	Namaste anna- do you have any tips to share for prevention of Oozi insect?	నమస్తే అన్న - ఊజీ ఈగ నివారణకు ఏవైనా చిట్కాలు ఉన్నాయా? ఉంటె ఈ గ్రూప్ లో ఉండే ఇతర రైతులతో పంచుకుంటారా?
Tips for Uzi 2	Namaste anna... we heard from some of you how bad the effects of Oozi pest are. Does anyone know the scientific name of the fly? Or any tips to control it	నమస్తే అన్నా... ఊజీ ఈగ తెగుళ్ళు ఎంత దారుణంగా ఉంటుందో మీలో కొందరి ద్వారా విన్నాం. ఊజీ ఈగ శాస్త్రీయ నామం/పేరు ఎవరికైనా తెలుసా? దీనిని నియంత్రించడానికి ఏవైనా చిట్కాలు చెప్తారా?
Tips for food loss	Namaste anna, unfortunately many farmers lose a lot of tomato yield due to price fluctuations. Do you have any tips on how to prevent this food loss?	నమస్తే అన్నా, దురదృష్టవశాత్తు చాలా మంది రైతులు ధరల హెచ్చుతగ్గుల కారణంగా చాలా టమోటా దిగుబడిని కోల్పోతున్నారు. ఈ ఆహార నష్టాన్ని అరికట్టడానికి మీకు ఏవైనా చిట్కాలు తెలుసా?
Tips for harvesting	Namaste anna, do you have any suggestions to share with fellow farmers on how to harvest tomato crop?	నమస్తే అన్నా, ఈ నెల మీటింగ్ కి , మీరు ప్రత్యేకంగా టమోటాలు కోయడం మరియు పంట కోయడం గురించి ఏదైనా సమాచారాన్ని తెలుసుకోవాలనుకుంటున్నారా?
Tips for heat prevention	Namaste anna, in this hot weather, do you have any tips on how to prevent crops from heat damage?	నమస్తే అన్నా, ఈ వేడి వాతావరణంలో, వేడిగాలుల నుండి పంటలను ఎలా నిరోధించాలో మీకు ఏమైనా చిట్కాలు తెలిస్తే పంచుకోండి
Reply for farmer asking for pest diagnosis	Namaste anna, please send this picture to the Plantix Plant doctor Whatsapp number at : 7876171002. They should answer your question within a few seconds. If you have any problem accessing it please let us know.	

References

[Aker, 2010] Aker, J. C. (2010). Information from markets near and far: Mobile phones and agricultural markets in niger. *American Economic Journal: Applied Economics*, 2(3):46–59.

[Aker and Mbiti, 2010] Aker, J. C. and Mbiti, I. M. (2010). Mobile phones and economic development in africa. *Journal of economic Perspectives*, 24(3):207–32.

[Badrinathan, 2020] Badrinathan, S. (2020). Educative interventions to combat misinformation: Evidence from a field experiment in india.

[Bandiera and Rasul, 2006] Bandiera, O. and Rasul, I. (2006). Social networks and technology adoption in northern mozambique. *The economic journal*, 116(514):869–902.

[Banerjee et al., 2013] Banerjee, A., Chandrasekhar, A. G., Duflo, E., and Jackson, M. O. (2013). The diffusion of microfinance. *Science*, 341(6144).

[Beaman et al., 2018] Beaman, L., BenYishay, A., Magruder, J., and Mobarak, A. M. (2018). Can network theory-based targeting increase technology adoption? Technical report, National Bureau of Economic Research.

[Beaman and Dillon, 2018] Beaman, L. and Dillon, A. (2018). Diffusion of agricultural information within social networks: Evidence on gender inequalities from mali. *Journal of Development Economics*, 133:147–161.

[BenYishay and Mobarak, 2019] BenYishay, A. and Mobarak, A. M. (2019). Social learning and incentives for experimentation and communication. *The Review of Economic Studies*, 86(3):976–1009.

[Besley and Case, 1993] Besley, T. and Case, A. (1993). Modeling technology adoption in developing countries. *The American economic review*, 83(2):396–402.

[Buragohaina et al., 2021] Buragohaina, P., Saikiaa, D. K., Sotelo-Cardonac, P., and Srinivasanc, R. (2021). Development and validation of an integrated pest management strategy against the invasive south american tomato leaf miner, tuta absoluta in south india. *Crop Protection*.

[Cai et al., 2015] Cai, J., De Janvry, A., and Sadoulet, E. (2015). Social networks and the decision to insure. *American Economic Journal: Applied Economics*, 7(2):81–108.

[Comola and Prina, 2015] Comola, M. and Prina, S. (2015). Treatment effect accounting for network changes: Evidence from a randomized intervention. *Available at SSRN 2250748*.

[Conley and Udry, 2001] Conley, T. and Udry, C. (2001). Social learning through networks: The adoption of new agricultural technologies in ghana. *American Journal of Agricultural Economics*, 83(3):668–673.

[D. Raji, 2017] D. Raji, R. (2017). Agriculture in telangana. *Professor Jayashankar Telangana State Agriculture University, Niti Aayog*.

[Foster and Rosenzweig, 1995] Foster, A. D. and Rosenzweig, M. R. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of political Economy*, 103(6):1176–1209.

[Goldsmith-Pinkham and Imbens, 2013] Goldsmith-Pinkham, P. and Imbens, G. W. (2013). Social networks and the identification of peer effects. *Journal of Business & Economic Statistics*, 31(3):253–264.

[Granovetter, 1973] Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, 78(6):1360–1380.

[Gupta et al., 2020] Gupta, A., Ponticelli, J., and Tesei, A. (2020). Information, technology adoption and productivity: The role of mobile phones in agriculture. Technical report, National Bureau of Economic Research.

[Harou et al., 2017] Harou, A. P., Walker, T. F., and Barrett, C. B. (2017). Is late really better than never? the farmer welfare effects of pineapple adoption in ghana. *Agricultural economics*, 48(2):153–164.

[Iqbal, 2020] Iqbal, M. (2020). Whatsapp revenue and usage statistics. *Business of Apps*.

[Jack, 2013] Jack, B. K. (2013). Market inefficiencies and the adoption of agricultural technologies in developing countries.

[Jensen, 2007] Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the south indian fisheries sector. *The quarterly journal of economics*, 122(3):879–924.

[Kondylis et al., 2017] Kondylis, F., Mueller, V., and Zhu, J. (2017). Seeing is believing? evidence from an extension network experiment. *Journal of Development Economics*, 125:1–20.

[Lasdun, 2022] Lasdun, V. (2022). Peer learning in a digital farmer-to-farmer network: Effects on technology adoption and self-efficacy beliefs.

[Maertens and Barrett, 2013] Maertens, A. and Barrett, C. B. (2013). Measuring social networks' effects on agricultural technology adoption. *American Journal of Agricultural Economics*, 95(2):353–359.

[Maertens et al., 2020] Maertens, A., Michelson, H., and Nourani, V. (2020). How do farmers learn from extension services? evidence from malawi. *American Journal of Agricultural Economics*.

[Magnan et al., 2015] Magnan, N., Spielman, D. J., Lybbert, T. J., and Gulati, K. (2015). Leveling with friends: Social networks and indian farmers' demand for a technology with heterogeneous benefits. *Journal of Development Economics*, 116:223–251.

[Manski, 1993] Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3):531–542.

[McNiven and Gilligan, 2012] McNiven, S. and Gilligan, D. O. (2012). Networks and constraints on the diffusion of a biofortified agricultural technology: Evidence from a partial population experiment. *Unpublished, University of California, Davis*.

[Moffitt et al., 2001] Moffitt, R. A. et al. (2001). Policy interventions, low-level equilibria, and social interactions. *Social dynamics*, 4(45-82):6–17.

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

[Murphy et al., 2019] Murphy, D. M., Nourani, V., and Lee, D. R. (2019). Chatting at church: Information diffusion through religious networks. *Review of Economics and Statistics*, pages 1–46.

[Nourani, 2016] Nourani, V. (2016). Social network effects of technology adoption: Investing with family, learning from friends & reacting to acquaintances. *Cornell University*, 54.

[Sunding and Zilberman, 2001] Sunding, D. and Zilberman, D. (2001). The agricultural innovation process: research and technology adoption in a changing agricultural sector. *Handbook of agricultural economics*, 1:207–261.

[Weersink and Fulton, 2020] Weersink, A. and Fulton, M. (2020). Limits to profit maximization as a guide to behavior change. *Applied Economic Perspectives and Policy*, 42(1):67–79.