Bridging the Information Gap: Sowing the Seeds of Productivity with High-Speed 4G Internet*

Sumit Agarwal Mohit Desai Pulak Ghosh Nishant Vats†

Abstract

Combining granular data on 4G rollout and Rights of Way policies with remote-sensing data on agricultural productivity, we find that 4G increases productivity, fertilizer use, credit uptake, and reduces price dispersion. Six years post-introduction, agricultural household income increases by 14.5%. Exploiting spatial variations in accuracy and value of information, we argue that 4G improves productivity by expanding information access. Farmer internet-browsing data shows that 4G indeed improves access to agri-information. Moreover, internet-based information access dominates traditional call or text-based information access, as internet's decentralized nature circumvents trust barriers associated with state-controlled traditional channels. However, internet infrastructure alone is insufficient; effective dissemination requires reliable, high-quality information, making internet access a complement to information generation.

*We thank John Barrios, Raissa Fabregas, Todd Gormley, Brett Green, Sean Higgins, Xiumin Martin, Kristina Manysheva, Patrick Sun, and Margarita Tsoutsoura as well as the seminar participants at the American Finance Association and Federal Communications Commission (FCC). We take full responsibility for any errors. We have no conflicts of interest to disclose.

†Agarwal: National University of Singapore (ushakri@yahoo.com); Desai: Kenan-Flagler Business School, UNC Chapel Hill (mohit_desai@kenan–flagler.unc.edu); Ghosh: Indian Institute of Management Bangalore (pulak.ghosh@iimb.ac.in); Vats: Olin Business School, Washington University at St Louis (vats@wustl.edu)

1 Introduction

The widespread availability of high-speed internet has fundamentally changed how we access and consume information. With its seamless data flow, high-speed internet helps remove barriers to information access, making it easier to share and disseminate knowledge. This is crucial because decision-making under imperfect knowledge can lead to sub-optimal outcomes. Although there is significant discussion about the potential of high-speed internet to improve access to timely and reliable information – thereby enabling more informed and efficient economic decisions – empirical evidence remains limited. Most existing research has focused on the impact of digital infrastructure on labor markets, political economy outcomes, financial inclusion, and credit markets. However, there is scarce empirical evidence on whether high-speed internet can improve information access and, consequently, boost productivity – particularly agricultural productivity (see Hjort and Tian (2021) for a review).

In this paper, we aim to fill this gap by studying the impact of the introduction of 4G based high-speed internet on agricultural productivity and the underlying mechanism. We investigate this question in the context of agriculture, which is well suited for a number of reasons. First, lack of information access can be detrimental in economic settings plagued by uncertainty. Agriculture in developing countries is often marked by weather-induced uncertainty and lack of stable insurance markets making access to information an important determinant of the output and input choice (Deaton, 1990; Townsend, 1994; Rosenzweig and Udry, 2013; Cole and Xiong, 2017). Second, existing research indicates minimal impact on agricultural productivity through traditional approaches (Fabregas, Kremer, and Schilbach, 2019).[1] These traditional approaches suffer from several frictions such as illiteracy, lack of trust, and timeliness (Cole and Sharma, 2017; Anderson and Feder, 2004). High-speed internet through 4G has the potential to overcome these barriers by enhancing real-time information accessibility for farmers that is user-friendly and engaging (Fabregas, Kremer, and Schilbach, 2019). Third, numerous governments are striving to enhance farmers' information access through the internet (Mehrabi et al., 2021). Therefore, from a policy perspective, it is crucial to comprehend the specific conditions under which such investments are likely to yield significant returns.

We attempt to answer this question in the context of a developing economy where information

[1] In the absence of the internet, information is disbursed via traditional means including television, radio, phone calls, texts, extension workers or other informal means including word of mouth within communities. The relative advantage of 4G is that it improves access to high-speed internet, whereas 2G and 3G primarily improve access to texts, calls and slow internet.

frictions are acute and a majority of households rely on agriculture for their livelihood. Particularly, we focus on India where more than 40% of the population is engaged in agriculture and there is an unmet need for information. As per the 2005 National Sample Survey, 60% of Indian farmers indicated a lack of access to any information source to support their farming practices. Moreover, the Indian context provides an ideal natural experiment that allows us to disentangle the effect of the 4G based high-speed internet on agricultural productivity from other confounding variables.

We begin by exploiting the staggered introduction of 4G towers in a difference-in-differences (DID) framework to estimate the effect of 4G on agricultural productivity. However, the exercise poses a key challenge as geographic coverage of a 4G tower is limited to a relatively small area around the tower whereas measures of agricultural productivity are typically available at a more aggregated level such as districts. Therefore, using administrative units as a unit of analysis violates the assumption of homogeneous treatment within treated units. We address this concern by employing a two-step approach. First, we divide the map of India into hexagons, each approximating the coverage area of a 4G tower. Subsequently, we use remote-sensing satellite data to create hexagon-level measures of agricultural productivity, specifically aggregating 8-day Enhanced Vegetation Index (EVI) composites to this level. Additionally, we integrate various economic, demographic, and agricultural variables into this dataset. Therefore, a key contribution of this paper is creating a granular database that can be employed in future research; this database links telecom coverage with several economic, agricultural, and demographic variables.

Our heterogeneity-robust Callaway and Sant'Anna (2021) DID estimates exploiting the staggered introduction of 4G towers imply that six years after 4G introduction, EVI implied yield increases by 0.035 units. This increase is equivalent to 12.5% of the mean value of EVI implied yield. Economically the effect on EVI implied yield is equivalent to – (1) 1.06% increase in agricultural yields of cereal crops, (2) a 4.28% increase in the production value of cereal crops, and (3) a 3.21% increase in the production value of all crops. We estimate that six years after the introduction of 4G internet, the annual income of agricultural households grew by 14.5%. To put these numbers in perspective, as per the Situation Assessment Survey of India the income of agricultural households grew by 59.01% over the period from 2014 until 2021. Our back of the envelope calculations suggest that 24.56% of the income growth of agricultural households over this period can be explained by the introduction of high-speed 4G internet.

Our empirical strategy exploits the variation across treated and not-yet-treated hexagonal units. Specifically, we compare hexagonal units that have already experienced the introduction

of 4G towers with units that are slated for introduction at a later stage. A potential concern may be that early treated units are different from units that are treated later. However, as long as these differences are time-invariant, this selection is fully accounted for by unit fixed effects. Thus, our specification does not require the introduction of 4G to be randomly allocated, nor does it require that units must have the same pre-treatment characteristics. Rather our estimate would be unbiased if the treated units would have evolved the same way as the not-yet-treated units in the absence of the treatment. We substantiate this assumption by presenting the pre-trend analysis in all our empirical assessments.

There may still be concerns associated with the endogeneity in the spatial introduction of 4G towers. For instance, the actual placement of 4G infrastructure may be driven by unobserved time-varying factors such as profit motives of telecom providers or regional agricultural productivity, which could bias our estimates. Prior literature has employed lightning as an instrument for the introduction of 2G/3G towers. However, lightning plays an important role in the nitrogen cycle. Therefore, such an instrument is likely to violate the exclusion restriction as the intensity of lightning can directly affect agricultural output and other economic variables in an agrarian economy.

We circumvent this issue by exploiting a plausibly exogenous variation in the roll-out of 4G technology due to the staggered introduction of Rights of Way (RoW) policies at the state-level. RoW policies were adopted at the state level with the intention to promote the growth of telecom infrastructure by streamlining rules and reducing the regulatory burden on telecom infrastructure companies. In addition to addressing the endogeneity concerns, this natural experiment also allows us to examine the economic effects of 4G introduction on variables that are recorded at the district- and state-level, such as fertilizer use, credit access, and prices.

We first establish that the enactment of the RoW rules had a positive effect on the construction of 4G towers in those states. Specifically, the number of towers within a district increased by 600 within six quarters of the adoption of the policy. Next, we document an increase in agricultural productivity following the adoption of RoW policy. The two results indicate that the RoW policies lead to an increase in the number of 4G towers as well as agricultural productivity. Furthermore, the introduction of 4G due to RoW policies is associated with increased fertilizer consumption by 22% and a 20% increase in credit uptake. Lastly, we document that the introduction of 4G increases market efficiency by reducing price dispersion of the same commodity across different markets.

Our results are robust across various measures of agricultural productivity, employing dynamic two-way fixed effects estimation, alternative estimators suggested by Borusyak, Jaravel, and

Spiess (2022) and de Chaisemartin and D'Haultfœuille (2020), and a falsification test employing a sub-sample of non-cropland hexagons. Moreover, while hexagonal tessellation facilitates a nearly uniform treatment across units, it does not guarantee strict uniformity. We leverage variation in tower height and the distance of each tower from the centroid of the hexagon to show that our estimates are robust to potential violations of the uniformity assumption.

Next, we investigate the improvement in access to information as a mechanism driving the increase in agricultural productivity following the introduction of 4G technology. While the introduction of 4G improves access to a wide variety of information, we focus on weather information for two reasons. First, agricultural output and returns heavily depend on weather shocks, particularly rainfall shocks (Townsend, 1994; Jayachandran, 2006; Kaur, 2019; Rosenzweig and Udry, 2013). Second, the availability of data on weather information and its spatial distribution allows us to more credibly identify the underlying mechanism.

We start by assessing the importance of the accuracy and reliability of weather information by using the distance of a hexagon to the nearest automatic weather station (AWS) as a proxy for the accuracy of forecasts. Almost all weather information and forecasts in India are sourced from the instruments located at AWS. However, weather shocks, especially rainfall, exhibit low geospatial correlation (Cole and Xiong, 2017). Therefore, the distance of a hexagon to the nearest AWS provides a proxy for the accuracy of weather information and forecasts. Specifically, as the distance to the nearest AWS increases, the accuracy of the forecast decreases, thereby decreasing the accuracy or reliability of the information. We find that hexagons closest to AWS exhibit the highest treatment effect of increased agricultural productivity after the introduction of high-speed 4G internet, compared to the statistically insignificant and economically negligible effects for hexagonal units far from AWS.

Furthermore, we refine our understanding of the mechanism by examining the heterogeneity in the treatment effects due to the potential value of information. We construct a hexagon-level measure of sensitivity to droughts based on the historical relationship between agricultural productivity and droughts as a proxy for the value of information. The intuition behind this test is that farmers operating in regions where extreme weather shocks explain a greater share of agricultural productivity stand to gain the most from rainfall forecasts and, thus, value weather information the most. Our results indicate that hexagons with ex-ante greater sensitivity to droughts are more likely to experience gains in agricultural productivity due to the introduction of 4G towers.

These results have two key takeaways. First, it sheds light on the mechanism, i.e., the im-

provement in agricultural yields due to the introduction of 4G towers is related to improved access to information. Second, the introduction of 4G infrastructure alone may not be enough; it acts as a complement to existing information-generating technologies. Specifically, the positive impact of 4G on enhancing access to information depends critically on whether the information being shared is reliable and valuable.

We provide additional support for our mechanism by presenting evidence of technology adoption by farmers. Most of our analysis is focused on the *availability* of 4G technology through the introduction of 4G towers. However, the uptake of 4G technology by farmers – once it is accessible through 4G towers – is of first-order importance for information-related mechanisms to drive agricultural productivity growth. Moreover, the kind of information accessed by farmers may play a significant role in shaping productivity growth. We address these questions using a proprietary dataset of geocoded and time-stamped mobile application ("app") installations from Krishify – an Indian firm which runs one of the largest social networks and commerce platforms for more than ten million farmers. Our DID estimates indicate that app installations increase after the introduction of 4G. This result makes explicit what has been implicit in our analysis till now: availability of 4G is associated with 4G adoption, as proxied through the installation of the Krishify app.

We also examine the type of information farmers access on the internet using detailed browsing and activity data of farmers on the Krishify app to provide more direct evidence showing that farmers are indeed using the internet to acquire agri-related information. The foremost three topics collectively constitute about 70% of farmer interactions on the platform. Notably, weather information and agricultural news rank highest, contributing to 30.9% of all interactions, followed by government programs at 21.3%, and information about farm machinery (e.g., tractors) at 16.1%. Farmers also utilize the platform to seek advice specific to animal husbandry (9.1%) and crop management (8.18%). Interactions related to inputs like fertilizers, seeds, finance, and other miscellaneous inputs collectively make up 12.8% of all farmer engagements on the platform. The detailed activity data of farmers suggests that farmers may be using internet-based platforms extensively for gathering information on a wide variety of agri-related activities.

Next, we compare the effects of 4G technology relative to earlier versions of telecommunications technology, such as 2G and 3G. High-speed internet in the form of 4G presents several key advantages over traditional modes of information dissemination including television, radio, phone calls, texts, and slow-speed internet. For instance, Fabregas, Kremer, and Schilbach (2019) argue in their review article that internet-based platforms may be an effective tool for disseminating in-

formation in a user-friendly and engaging manner by employing videos and graphics. Moreover, 4G significantly improves real-time information accessibility for farmers which is crucial for the impact of information dissemination on agricultural productivity (Anderson and Feder, 2004).

We examine this hypothesis by comparing the effects of the introduction of 2G, 3G and 4G on agricultural productivity. We do not find any economically or statistically significant improvement in agricultural productivity following the introduction of 3G, a generation that provides access to relatively low-speed internet. Similarly, we do not find evidence of a positive effect of the introduction of 2G on agricultural productivity. Overall, these results suggest that 4G technology may possess distinct advantages in improving information access for farmers. This analysis leverages a large, regionally, and temporally extensive dataset across India, employing granular 16-day composites to reconstruct EVI-implied yields dating back to 2000. Consequently, our findings are less susceptible to critiques that null results in similar studies – particularly those based on randomized control trials – are due to limited sample sizes and low statistical power.

We investigate a particular channel that potentially makes information access through high-speed 4G internet special – role of trust. Information dissemination through traditional modes such as text messages, call centres, hotlines, etc. is often directly controlled by the state and its agents. Moreover, local governing bodies known as village panchayats often act as the central node for information dissemination by the government. Therefore, a lack of trust in local and state authorities can hinder effective communication and engagement through these traditional methods (Cole and Sharma, 2017; Fabregas, Kremer, and Schilbach, 2019). The internet, on the other hand, being a more decentralized platform, may not encounter the same issues related to trust. For instance, a user of a weather application may not realize that the weather information on the app is sourced from government weather stations. Consequently, the introduction of 4G internet may have greater effects in regions where trust in state institutions is diminished.

Using responses from the India Human Development Survey (IHDS) to questions on the level of trust in local village panchayats and state governments, we compare changes in agricultural yields after 4G introduction in areas with ex-ante high and low levels of trust. Our findings suggest that regions characterized by low trust in village panchayats or the state exhibit an increase in agricultural production following the introduction of 4G. In contrast, the impact of 4G introduction in areas with a high level of trust in local village panchayats is economically small and statistically insignificant. These results imply that the advancement of high-speed internet through 4G infrastructure can mitigate the effect of human frictions, like mistrust in local institutions and the state.

Related Literature: The key contribution of this work is examining the effect of modern communication infrastructure, specifically high-speed internet via 4G, on agricultural productivity. While existing research has predominantly examined the impact of traditional information dissemination methods – such as face-to-face interactions, SMS, and voice calls – on agricultural yields, we focus on the effects of high-speed internet deployment (Fabregas, Kremer, and Schilbach, 2019). We contribute to this literature in two ways. First, we show that 4G can have a positive effect on agricultural yield as well as usage of agricultural inputs such as fertilizers and bank credit. Second, our results suggest that 4G infrastructure plays a crucial role in improving information access and consequently productivity. Our results stand in contrast to most prior studies which generally find negligible long-term effects of traditional communications interventions on agricultural yields.[2] We complement the extant literature by documenting that earlier generations of mobile communication technology – 2G and 3G – which improve information dissemination through calls, texts, and low-speed internet has economically negligible effect on yields, highlighting the unique role of high-speed internet. Moreover, unlike previous small-sample studies which leverage experimental variation; our analysis combines a large, regionally, and temporally extensive dataset across India with granular remote sensing data to estimate agricultural yields from 2000 and straggered roll-out of 4G technology. Consequently, our methodology is better equipped to address concerns that previous null findings – often based on randomized controlled trials – may have been driven by limited sample sizes and insufficient statistical power.

The second contribution of this paper is documenting that the dissemination of information through high-speed internet may dominate conventional approaches to improving information access. Prior literature has argued that traditional approaches to information dissemination may suffer from several frictions such as lack of trust, less user-friendly content, and lack of timeliness (Anderson and Feder, 2004; Duflo, Kremer, and Robinson, 2011; Cole and Sharma, 2017). We contribute to this literature by showing that internet-based information dissemination via 4G can help reduce a key human friction: lack of trust in government.

Our paper also contributes to the literature examining the effect of infrastructure development on agricultural yields. Prior work has documented limited effects of infrastructure development on

[2] Cole and Fernando (2021) find no systematic impact on yields on the treated farmers in the Indian state of Gujarat that randomly received access to a hotline for agricultural advice. Additionally, experiments by Fafchamps and Minten (2012) that sent Indian farmers price and weather information via text messages found no evidence of value addition. document similar results among Colombian farmers. Other studies documenting null results of traditional information dissemination mechanism on yields include Camacho and Conover (2010), Van Campenhout, Spielman, and Lecoutere (2021), Udry et al. (2019), among others. A notable exception is Casaburi et al. (2014). They document a positive effects on yields of a text message based agricultural advice experiment among small sugarcane farmers in Kenya. However, a follow-up trial of the same intervention has no significant impact on yields.

agricultural yields or production. These studies include examining the effects of the construction of rural roads in India (Asher and Novosad, 2020), bridges in Nicaragua (Brooks and Donovan, 2020), and kisan call centers in India (Gupta, Ponticelli, and Tesei, 2024). Our contribution to the literature is two-fold. First, we document the positive effect of 4G infrastructure on agricultural yields. Second, our results suggest that the introduction of 4G infrastructure in isolation might not be adequate, i.e., it acts as a complement to existing infrastructure. Specifically, we show that when 4G introduction is not coupled with the reliable and valuable information-generating infrastructure, its potential to boost output may be constrained.

More broadly, we are related to the literature examining the economic effects of digital infrastructure.[3] We add to this literature in three ways. First, as discussed by Hjort and Tian (2021) in their review article, we are the first to provide evidence linking firm-level productivity to internet connectivity. We establish this relationship in the context of farms and contribute to prior work of Akerman, Gaarder, and Mogstad (2015), Hjort and Poulsen (2019) and Tian (2021) who document the effect of digital infrastructure on labor productivity.Second, we document that high-speed internet reduces price dispersion by lowering information frictions, thereby improving market efficiency à la Allen (2014). This finding complements prior research documenting similar effects from the introduction of telegraph, mobile phones, and text messaging (Jensen, 2007; Parker, Ramdas, and Savva, 2016; Jensen and Miller, 2018; Steinwender, 2018). Third, we find that 4G increases agricultural productivity, particularly in weather-sensitive areas. We argue that high-speed internet acts as a climate adaptation tool by providing farmers with timely, localized weather data. These results highlight the need to incorporate internet access into climate resilience strategies, a conclusion supported by recent experimental evidence presented in Burlig et al. (2025).

This paper proceeds as follows. Section 2 describes the data. Section 3 delineates the empirical strategy. Section 4 presents the baseline effect of the introduction of 4G on agricultural productivity. Section 6 documents the underlying mechanisms. Section 7 examines the importance of the introduction of 4G relative to earlier versions of telecommunications technology, such as 2G and 3G. Section 8 concludes.

[3]This literature has examined the impact of digital infrastructure on labor market outcomes (Akerman, Gaarder, and Mogstad, 2015; Hjort and Poulsen, 2019; Zuo, 2021), political mobilization (Manacorda and Tesei, 2020), corruption (Andersen et al., 2011; Gonzalez, 2021), confidence in governments (Guriev, Melnikov, and Zhuravskaya, 2021), financial inclusion (Aker and Mbiti, 2010; Suri and Jack, 2016; Suri, 2017; Batista and Vicente, 2023), and credit markets (Brunnermeier, Limodio, and Spadavecchia, 2023; D'Andrea and Limodio, 2023; D'Andrea et al., 2024; Ghosh et al., 2025). This by no means is an exhaustive list of the literature and we direct readers to Hjort and Tian (2021) for a systematic review of the literature on the economic impact of internet connectivity in developing countries. Furthermore, Goldfarb and Tucker (2019) present a review of the literature on digital economics in developed countries and Zhuravskaya, Petrova, and Enikolopov (2020) survey the literature on the political impact of internet connectivity and social media.

2 Data

This section discusses the various datasets employed in the analysis. Our primary datasets include Base Transceiver Stations (BTS) geolocation data and remote sensing data on agricultural yields and croplands. We use the BTS data to identify the timing of the treatment and the remote sensing data to construct our key dependent variables. We also map several other economic, demographic, and agricultural variables to this dataset. Therefore, a key contribution of this paper is creating a granular database that links telecom coverage with several economic, agricultural, and demographic variables that can be employed in future research.

2.1 BTS Data

We collect a novel dataset from the Ministry of Telecom, which provides detailed information on the geolocations of all Base Transceiver Stations (BTS) in India. A BTS serves as a communication link between user devices, such as mobile phones or computers, and the network. The dataset includes an exhaustive list of all BTSs in India, deployed at various points in time from 1995 to 2022.

The data contains information on precise geographic coordinates of each BTS, the associated technology (e.g., 2G/3G/4G), and the date of start of operation. In total, our dataset comprises information on 2.35 million unique BTS across 717,764 towers. We restrict our sample to the subset of 1.2 million BTS that were the first-operational-BTS-by-technology at that geolocation. Figure 1 presents the evolution of 4G, 3G, and 2G coverage over space and time.

2.1.1 Hexagonal Tessellation

A key challenge in working with the BTS data is deciding on the appropriate unit of analysis. Particularly, it is challenging to use conventional administrative units, such as a district or block, as the unit of analysis due to two reasons. First, the geographic coverage of BTS is typically limited around the tower. Second, BTS are activated on multiple towers within close vicinity over time. As a result, the usage of large administrative units as the unit of analysis violates the assumption of homogeneous treatment within treated units, as multiple locations within a unit receive unequal treatment over time.

We attempt to address this concern, by dividing the map of India into a grid of hexagons or hexagonal tessellation, with each hexagon representing a unit of analysis. We choose hexagonal tessellation because it is the most efficient arrangement of equal circles that fills a plane without any overlaps or gaps (Birch, Oom, and Beecham, 2007). Specifically, we draw hexagons with a distance

of 0.1-degree latitude between their opposite ends, which is roughly equivalent to a distance of 11 km. These hexagons can be thought of as circles with a radius of 5.5 km, with approximate area of 95 sq. km. This size is reasonable considering that the range of a 4G cell tower is typically between 3-6.5 km (2-4 miles).[4]

By employing hexagonal tessellation, we aim to ensure that each location within our analysis unit receives approximately equal treatment. However, two key factors may prevent the treatment from being perfectly uniform. First, achieving exact uniformity requires that the BTS location precisely coincide with the centroid of each hexagon. Second, BTS in towers with different heights may have different coverage areas. Given that our analysis focuses on India – a large and geographically diverse country – it is not feasible to guarantee perfect treatment homogeneity within each hexagon. Instead, we can only argue that treatment is *almost* uniform. We will examine the robustness of our findings to these concerns in Section 4.2.

We overlay the geographic coordinates of the BTS onto the tessellated map to assign each first-operational-BTS-by-technology to a unique hexagon. See Appendix Figures A.1 and A.2 for an illustration of the overlay of BTS on the tessellated map of India. Figure 2 plots the number of hexagons first treated by BTS of different technologies from 1995 until 2022. We identify the date of treatment for each hexagon, i.e., the date on which a hexagon first saw the introduction of a 4G BTS, by using the date of operation for 4G BTS.

2.2 Remote-sensing data on Agricultural Production

We use remote-sensing data to construct measures of unit-level (hexagon-level) agricultural production, as no such granular agricultural production data exists for India. This data comes from NASA's Earth Observation Satellite - Landsat 8. Specifically, we use Enhanced Vegetation Index (EVI) to construct our unit-level measures of agricultural production.

EVI is a chlorophyll-sensitive composite measure of plant matter generated by NASA's Earth Observation satellite – Landsat 8. The composites are created from all the scenes in each 8-day period beginning from the first day of the year and continuing to the 360th day of the year. Each pixel value is optimized considering cloud cover obstruction, the influence of background vegetation, image quality, and viewing geometry. We direct the readers to Huete et al. (2002) for details on the construction of an earlier version of EVI and the usage of this earlier version – 16-day EVI composites – in economics research by Asher and Novosad (2020). More recently, Ghosh and

[4] See here.

Vats (2022) employ the EVI values based on 8-day composites to examine the effect of guaranteed income to farmers on agricultural productivity.

Our sample period extends from 2013 until 2021 as the 8-day EVI composites are available from April of 2013 until January of 2022. We query EVI values over this period for our desired micro-regions, hexagonal units, by supplying the geometry of the micro-region to the Google Earth Engine's API. EVI values obtained from each pixel are spatially averaged over the hexagon to obtain a time-series of EVI values with an 8-day interval. Appendix Figure A.3a plots the evolution of average, the 25^{th} percentile, and the 75^{th} percentile EVI value for *all* hexagons in our sample.

Following the methodology outlined in Asher and Novosad (2020) and Ghosh and Vats (2022), we construct hexagon-level measures of agricultural yields for the primary growing season of kharif that begins in June and ends in October of each year. We construct two measures using the 8-day composite based EVI values. Our preferred measure – EVI implied agricultural yield – is constructed by subtracting the average value of EVI during the initial weeks of kharif season from the maximum EVI value during the season. Appendix Figure A.3a shows that the EVI value increases with the start of the kharif season and reaches its highest value around the kharif harvest time. Therefore, the difference measure effectively measures production per unit of pixel during the season. Moreover, the difference measure implicitly controls for differences in non-crop vegetation, such as perennial non-crop green cover that maybe captured by EVI composites. Our second measure – maximum EVI – uses the maximum EVI value for the kharif season. Panel A of Table 1 presents the summary statistics for the two measures. The two measures are highly correlated with a correlation coefficient of 0.53 (see Appendix Figure A.3b).

An assumption that justifies the usage of these measures is that they exhibit significant correlation with both agricultural productivity measures and real production measures. We verify this assumption in Appendix A.3 for our preferred measure of agricultural production based on 8-day composite exhibits. We find that the EVI implied yield based on 8-day EVI composites exhibit significant correlation with real production measures such as yield and production value measured in USD.

In addition to constructing hexagon-level EVI measures using 8-day composites from 2013 to 2022, we also construct hexagon-level EVI values based on 16-day composites to extend our dataset back to 2000. As shown in Appendix Figure A.5, the two yield measures – derived from 8-day and 16-day composites – are highly correlated. We construct this alternative measure to assess the impact of 2G and 3G technology rollouts on agricultural productivity, which primarily occurred

by 2015, as indicated in Figure 2.

2.3 Other datasets

This section describes a range of other datasets used in our analysis.

Cropland Data: The remote-sensing cropland extent data comes from the Global Food Security-Support Analysis Data (GFSAD) and identifies cropland at the resolution of 1 km for the year 2010.[5] We map the cropland extent data to our hexagons and remove all hexagons that do not have any cropland to focus attention on agrarian units.

Weather Station Location: Data on geo-locations of automatic weather stations (AWS) comes from the Ministry of Earth Sciences, Government of India.

Local Drought Shocks: We use the Standardized Precipitation and Evapotranspiration Index (SPEI) to construct local drought shocks. Specifically, we follow the methodology outlined in Cramer et al. (2024) to map this index coherently and uniformly to our hexagons and the methodology outlined in Jayachandran (2006) to construct drought shocks from 2000 until 2012. Appendix A.5 presents details of the precise methodology followed to construct these local drought shocks.

Agricultural Production Potential: Agricultural production potential is measured using two measures provided in Advancing Research on Nutrition and Agriculture's (ARENA) Demographic and Health Surveys (DHS)-GIS Database. Our measures of agricultural production from the ARENA-DHA database is based on exogenous characteristics such as soil characteristics and water availability in the area. Survey clusters in the DHS data are mapped to hexagons based on the minimum distance between centroids of hexagons and geo-coordinates of survey clusters. These measures are calculated as of 2014.

Trust in Village Panchayats & State Government: The data on trust in village panchayat and state government comes for the India Human Development Survey (IHDS-II) conducted in 2011, which includes a question that queries respondents about the level of trust they have in their local village panchayats and state government. The exact questions are: "How much trust do you place in the ability of your village panchayat to implement public projects?" and "How much trust do you place in the ability of the state government to take care of people?" The response options consist of three choices: a) A great deal of confidence, b) Only some confidence, and c) Hardly any confidence at all. To quantify these responses, we assign a numerical value of 1 to option (a), 0.33 to option (b), and 0 to option (c). We take the average of this measure across all households within a

[5] GFSAD is a NASA funded project providing high-resolution global cropland data. Data documentation of can be found here.

district based on their weights. This allows us to compute a continuous measure of confidence in village panchayats and state government at the district level. We use the response to this question to construct a district-level proxy for confidence in village panchayats and divide the districts into two subsets based on the median response value. Districts with confidence levels above the median are categorized as "high confidence" districts, while those below the median are classified as "low confidence" districts. We then link this measure to the hexagonal grids situated within the districts.

Fertilizer Consumption data: We collect district-level data on consumption of total fertilizers, fertilizer by nutrient type, and gross sown area from the States of India (SoIdx) database maintained by the Center for Monitoring the Indian Economy (CMIE). This data spans from 2000 until 2020. We map hexagons in our dataset to districts based on the extent of overlap, i.e., if a hexagon falls in multiple districts it is mapped to the district with the maximum area of overlap.

Agricultural Credit data: We obtain quarterly branch level data from one of the largest state owned bank in India. The data spans 2013-2021 and contains information about total agricultural credit disbursed by the branch during the period.

Mandi-level Price data: We obtain daily market-level (*mandi*) prices from Agmarkonline Portal maintained by the Government of India. We obtain this data for nine key crops– Bajra, Maize, Jowar, Wheat, Paddy, Soyabean, Potato, Tomato, and Onions – from 2011 to 2021.

Krishify app installations data: Krishify is an Indian company aimed at connecting farmers on a social network where they can discuss agriculture related issues.[6] We obtain a proprietary, geolocated and time-stamped dataset of app installations for the years 2019-2021. We superimpose the coordinates of app installations over our hexagonal grid to identify the number of app installed at the hexagonal level. Additionally, we also get a 10% random sample of all farmers in the Krishify database along with their detailed search, browsing, likes, and comment history.

Table 1 presents the summary statistics of the key variables used in the analysis. The average value of our preferred measure of agricultural productivity, EVI implied yield, is 0.27 with a standard deviation of 0.14. The hexagons, on average, are located 40km away from the nearest weather station. Rainfall exhibits significant variability, as indicated by the coefficient of variation, which ranges from 0.127 to 1.414.

[6]Link to their official website can be found here. Appendix A.6 presents more details about the company and the data employed in this paper.

3 Empirical Strategy

Our objective is to estimate the effect of high-speed internet on agricultural productivity. We estimate this effect by employing a difference-in-differences (DiD) framework that leverages the staggered rollout of Base Transceiver Station (BTS) within hexagons. Particularly, our empirical strategy examines the effect on agricultural productivity in hexagons around the time when BTSs first become operational in the *treated* hexagons relative to the agricultural productivity in *control* hexagons which are *not-yet-treated*.

Recent econometric developments have highlighted that difference-in-differences (DID) estimates derived from the traditional two-way fixed effects approach rely on the assumption of treatment effect homogeneity in staggered adoption settings. This assumption of treatment effect homogeneity imposes the condition of homogeneous treatment effects across cohorts, i.e., farmers gaining access to 4G services in 2014 experience similar productivity gains as farmers that gain 4G access in 2021. Intuitively, the assumption of treatment effect homogeneity may be violated due to a variety of reasons, such as changes in forecasting skills over time, mergers and acquisitions among telecom service providers, and variation across cohorts due to calendar time-varying macroeconomic conditions could govern the effects on outcomes across cohorts. We direct readers to Appendix B for a detailed econometric discussion of this issue.

We address the issue of violation of the treatment effect homogeneity assumption by presenting our event-time DID estimates using the Callaway and Sant'Anna (2021) estimator. We choose this estimator for two reasons discussed in Roth et al. (2023). First, this approach provides sensible estimands even under arbitrary heterogeneity of treatment effects. Second, this approach makes transparent which units are being used as a control group to infer the unobserved potential outcomes. Specifically, we use the set of not-yet-treated cohorts as the control group because the no-never-treated group is minuscule in our design. For robustness, we also present our event-time DID estimates using the estimators outlined in Borusyak, Jaravel, and Spiess (2022) and De Chaisemartin and d'Haultfoeuille (2020).

The Callaway and Sant'Anna (2021) framework is based on causal parameters referred to as *group-time average treatment effects*. These effects represent the average treatment effect for units (hexagons) that belong to a specific group g at a particular time period t, i.e., group of hexagons

treated in 2014, 2015,... etc.:

$$ATT(g,t) = \mathbb{E}[Y_t(g) - Y_t(0)|G_g = 1] \tag{1}$$

where $Y_t(g)$ is the outcome variable for treated group g at time t, $Y_t(0)$ is the untreated potential outcome for group g at time t, and G_g is a binary variable that equals one for units belonging to group g, i.e., group first treated in period g.

Most importantly, the construction of these $ATT(g,t)$'s does not rely on the assumption of treatment effect homogeneity across groups or across time. These $ATT(g,t)$'s are then aggregated across groups to arrive at event-time parameters analogous to coefficients in a dynamic TWFE specification while avoiding the pitfalls associated with such specifications.

3.0.1 Identifying Assumption

Our empirical strategy exploits the variation across treated and not-yet-treated hexagonal units (hexagons). Therefore, one natural concern is that early treated units are different from units that are treated later. However, as long as these differences are time-invariant, this selection is fully accounted for by unit fixed effects. Thus, our specification does not require that the introduction of high-speed 4G BTS was randomly allocated, nor does it require that units must have the same pre-treatment characteristics. Rather our estimate would be unbiased if the treated units would have evolved the same way as the not-yet-treated units in the absence of the treatment. Our identifying assumption implies that unit-level trends in agricultural production would have been the same in treated and not-yet-treated in the absence of the policy. While this assumption is untestable, we present parallel pre-trends for all our analysis to provide supporting evidence to our identifying assumption.

4 Baseline Results

This section presents the effect of 4G introduction on agricultural yields. We use the heterogeneity-robust, dynamic difference-in-differences estimator outlined in Callaway and Sant'Anna (2021) that exploits the staggered introduction of 4G across hexagons in our sample.

Figure 3 presents the results using the Callaway and Sant'Anna (2021) estimator.[7] The dependent variable is standardized to mean zero and standard deviation of one. Our sample consists of

[7] Appendix Figure C.1a presents the dynamic coefficients estimated using the Callaway and Sant'Anna (2021) estimator with pre-treatment effects calculated as long differences to make the visual presentation more comparable with the traditional TWFE event-study regressions.

all hexagons that received the 4G treatment at some point between 2013 and 2021, thus comparing treated units with not-yet-treated units. Our sample is a balanced panel of hexagons from 2013 until 2021. We restrict our sample to agrarian units by removing hexagons that did not contain any cropland within its boundaries as identified by the GFSAD data on cropland extent. The outcome variable is the EVI implied yield, constructed by subtracting the average value of EVI during the initial weeks of the kharif season from the maximum EVI value over the entire season. Each plotted coefficient measures the average treatment effect on the treated in event-time e constructed by aggregating over event-time coefficients for groups treated in each year, i.e., $ATT(g, g+e)$'s.

There are two key takeaways from Figure 3. First, we do not observe significant pre-trends in EVI implied yields in the event-time leading up to the treatment. Second, we observe an increase in EVI implied yield following the introduction of high-speed 4G internet. Specifically, we observe that the treatment effect increases gradually after treatment.

We also provide estimates of the overall effects of treatment. We characterize the overall treatment effect using three different estimators outlined in Callaway and Sant'Anna (2021) that aggregate $ATT(g,t)$'s into a single estimate of the aggregate average treatment effect. First, in simple aggregation, the aggregate ATT is defined as the weighted average of all group-time average treatment effects. Second, in group-specific aggregation, the aggregate ATT is defined as the average of partially aggregated group-specific ATT's across all groups. Third, in event-time aggregation, the aggregate ATT is defined as the average of partially aggregated event-time ATT's across all lengths of post-treatment exposure.

The table in Figure 3 presents the three measures of aggregate ATT. The last column of the table presents our estimates of the average effect of 4G introduction on agricultural yields. The estimates indicate that the introduction of high-speed 4G internet significantly increased agricultural yields across all three measures of aggregate ATT.

4.1 Discussion of the Magnitude of Estimate

We find that six years after the treatment EVI implied yield increases by 0.24 times the standard deviation or 0.035 unit increase in EVI. This increase is equivalent to the 12.5% of the mean value of EVI implied yield.

We translate the estimates of 4G introduction on EVI implied yield to the effect on real agricultural yield and production value. Appendix A.3 presents the detailed calculations. Our estimates imply that the introduction of high-speed 4G internet is associated with a – (1) 1.06% increase in

agricultural yields of cereal crops, (2) a 4.28% increase in the production value of cereal crops, and (3) a 3.21% increase in the production value of all crops. Overall, our calculations imply that six years after the introduction of 4G internet, the annual income of agricultural households grew by 14.5%.[8] According to the Situation Assessment Survey of India, the income of agricultural households grew by 59.01% over the period from 2014 until 2021. Our back-of-the-envelope calculations suggest that 24.56% of this income growth of agricultural households over this period can be explained by the introduction of high-speed 4G internet.

Our estimated effect on income is comparable with the estimates in the existing research on the economic impact of internet access. For instance, Keleher, Rezaee, and Troland (2020) find a 17% increase in household income following the deployment of mobile towers in the Philippines. More closely related to our context, Gunsilius and Van Dijcke (2023) document a 25-35% decline in economic activity following the internet shutdown in India.

We also compare our estimate of the increase in yield with the effects of climate change, infrastructure developments, and other interventions on agricultural yields. Guiteras (2009) estimates that projected climate change in India will reduce major crop yields by 4.5-9% over the period 2010-2039. Our estimate of increase in yield of 1.02% for major cereal crops, six years after the introduction of high-speed 4G internet, may be sufficient to offset this projected climate change impact on agricultural yield in India. We discuss the role of 4G in improving the resilience of agricultural production to climate change in detail in Section 6.1.2.

Next, we compare our estimates to the documented effects of other infrastructure development on agricultural yields. Asher and Novosad (2020) study the development of rural road infrastructure in India, and do not find any significant effect on agricultural income four years after road construction. Similarly, Brooks and Donovan (2020) do not find a statistically significant effect of construction of bridges in Nicaragua on agricultural production of maize and beans. In contrast, our estimate indicates a 1.02% increase in agricultural yield, five years after the introduction of high-speed 4G internet.

Our estimate is smaller when compared to other agricultural interventions that reduce risk or directly increase productivity. Emerick et al. (2016) conduct a randomized controlled trial in India that involves distribution of a new rice variety that reduces downside risk by providing flood tolerance. They document that this intervention has positive effects on adoption of a more labor-intensive planting methods, area under cultivation, fertilizer usage, and credit utilization translat-

[8]We direct readers to Appendix A.4 for the detailed calculations.

ing to improvement in productivity. Their most conservative estimate indicates that productivity increases by 6%. Our estimate of increase in yield is roughly one-sixth of the estimate reported in Emerick et al. (2016). Beaman et al. (2013) document that yield increases by 31% and 16.6% when farmers in Mali receive fertilizers equal to the recommended quantity and half the recommended quantity, respectively. Similarly, Gollin, Hansen, and Wingender (2021) estimate that the Green Revolution in the developing world led to a 44% increase in crop yields between 1965 and 2010, following the adoption of high-yielding crop varieties (HYVs).

4.2 Robustness

This section presents a battery of robustness tests to strengthen the validity of our baseline results.

Alternative Estimation Methodology: First, we test the robustness of our results using alternative estimators proposed in Borusyak, Jaravel, and Spiess (2022) and de Chaisemartin and D'Haultfœuille (2020). Figure 4 presents the dynamic event-time coefficients obtained from these estimators, with our baseline measure of agricultural production — EVI implied yield – as the dependent variable. For reference, we also show the dynamic TWFE estimator and our baseline estimator of Callaway and Sant'Anna (2021). The tests employing alternative estimation methodologies echo our baseline results.[9]

Falsification Test: Second, we present a falsification test using a sub-sample of hexagons that do not contain any cropland within their boundaries as of 2010. The objective of this test is to show that our empirical design is valid, the results are unlikely to be driven by spurious factors, and our results represent an actual increase in agricultural production. The falsification sub-sample contains no cropland. Therefore, we should not observe any change in agricultural production in these units. Appendix Figure C.1b presents these results. We do not find any economically or statistically significant effect on agricultural production for these units.

Alternative Measure: Third, we demonstrate the robustness of our results using an alternative measure of agricultural production. We use the maximum EVI value during the kharif season as our dependent variable. Appendix Figure C.1c presents these results and documents results similar to our baseline results.

Controlling for 3G Entry: Fourth, we demonstrate the robustness of our results to the simultane-

[9]We do not use the Sun and Abraham (2021) estimator as it uses the set of never treated units as the control group, which is very small in our setting. Moreover, the standard errors in the pre-period for the Borusyak, Jaravel, and Spiess (2022) estimator are large. This issue of the large standard errors is consistent with the discussion on standard error estimation in Borusyak, Jaravel, and Spiess (2022) and Braghieri, Levy, and Makarin (2022). Similar to Braghieri, Levy, and Makarin (2022), we also find that the standard errors increase dramatically as we increase the number of pre-periods.

ous entry of 3G. Specifically, we control for event time for 3G entry at the hexagon level. Appendix Figure C.1d presents these results and documents that our 4G entry results are unlikely to be affected by the simultaneous entry of 3G.

Assumption of Equal Treatment within Hexagons: Fifth, we examine the robustness of our estimates to potential violations of the assumption of uniform treatment within a hexagon. Specifically, we consider two concerns: BTS locations not aligning with hexagon centroids and variations in tower height. Appendix Figure C.2 shows that our baseline results remain qualitatively consistent across sub-samples adjusted for these factors. This suggests our findings are unlikely to be driven by violations of the equal-treatment assumption within a hexagon.

Heterogeneity by Agricultural Production Potential: Lastly, we present a sanity check for our results by exploiting the heterogeneity in agricultural production potential. Intuitively, if high-speed internet increases agricultural production, we should document a greater effect in areas with high potential for agricultural production. Appendix Figure C.3 presents these results for two measures of agricultural production from the ARENA-DHA database. We find that the effect of the introduction of high-speed 4G internet is higher in areas with high agricultural production potential (shown in red) relative to areas with low agricultural production potential (shown in blue).

5 Alternative Identification Strategy

So far, our estimation approach leverages the staggered deployment of 4G technology across regions. While this approach doesn't require identical pre-treatment characteristics there may still be concerns of endogeneity. For instance, the actual placement of 4G infrastructure may be driven by unobserved time-varying factors such as profit motives of telecom providers or regional agricultural productivity, which could bias our estimates. We address this concern by exploiting variation in 4G rollout timing driven by the staggered adoption of Rights of Way (RoW) regulations by different states. Additionally, tests using this variation allow us to examine the economic effects of 4G introduction on variables that are recorded at the district- and state-level, such as fertilizer use, credit access, and prices.

5.1 Details of the RoW policy

Laying of optical fibers and installing telecom towers requires significant digging and trenching work on public lands (GSMA, 2020). As such, access to public Rights of Way (RoW) forms an

essential part of deploying telecom infrastructure at a large scale. The permissions required for such RoW are the jurisdiction of states and local municipalities.

However, the telecom industry often faced strong opposition from local bodies and the public due to issues related to estate taxes and concerns over electromagnetic frequency emissions that allegedly cause cancer. For instance, the Cellular Operators Association of India (COAI), the key lobby group for the major telecom operators in the country, presented its grievances to the telecom regulator, saying, "State bodies continue to initiate actions such as disconnecting electricity, sealing the premises and even dismantling of tower sites without any prior notice, leading to coverage disruptions and network congestion."[10]

The federal government established a framework to standardize regulations for the deployment and maintenance of telecom infrastructure nationwide, aiming to streamline rules and lessen the regulatory burden on telecom providers. This effort led to the introduction of the Indian Telegraph Right of Ways Rules on 15th November 2016. The policy provides a blueprint for rules to be followed by telecom companies and local administrative bodies. Anecdotal evidence suggests that the blueprint was inspired by the RoW policy implemented in the state of Jharkhand in December 2015 following a joint pilot between the Telecom Regulatory Authority of India (TRAI) and the government of Jharkhand.[11]

These rules were designed to streamline the process by which companies apply for approvals and resolve disputes related to telecom infrastructure. Moreover, the rules required the establishment of specific timeframes. For example, the policy dictates that the state authority must respond to applications within 60 days from the date of submission, either by granting or denying the application. If the state authority fails to respond within this timeframe, permission is automatically granted, and the authority cannot reject the application without first giving the company an opportunity to be heard and providing a written reason for the rejection.

While the policy was drafted by the federal government as a blueprint, its adoption was left to the discretion of individual states. Consequently, different states implemented the RoW framework at various times between 2015 and 2019. Appendix Table D.1 presents the specific adoption dates of the RoW framework by different states. Sixteen states had adopted the framework by December 2019.

The policy was perceived as a positive development by the telecom industry. For instance,

[10] See here for more details.
[11] See here for details.

Rajan Matthews, director general of the COAI, noted, "It is a great move to assist the industry with improving the quality of service experience of customers...This will provide a great fillip to expanding cell site coverage as well as fiber implementation to support broadband services."[12] Overall, the passage of a state-wide RoW policy significantly reduced the regulatory and operational burden on telecom companies.

5.1.1 Identifying Assumption

The selection of states into the adoption of RoW policies could be endogenously linked to agricultural productivity. For instance, states with proactive leadership and bureaucracies are more likely to adopt such policies. This proactivity might also influence other agricultural policies, creating a correlation between RoW adoption and agricultural productivity at the state level.

We address this concern in two ways. First, we use the set of adopters, i.e., states that eventually adopted RoW policy by 2020. Second, we include unit and time-fixed effects in our empirical specification. As before, we use a heterogeneity-robust Callaway and Sant'Anna (2021) estimator. Therefore, our empirical strategy exploits variation across treated and not-yet-treated states. Moreover, our estimation strategy fully accounts for selection concerns as long as the differences across states that lead to the adoption of RoW policies are time-invariant. Under this identifying assumption, the date of adoption of RoW policies is likely to be random or unlikely to be driven by pre-existing differences across adopters. While this assumption is untestable, we present parallel pre-trends for all our analysis to provide supporting evidence to our identifying assumption.

5.2 Effect of RoW Adoption on Installation of 4G Telecom Towers

We quantitatively establish the effect of the staggered adoption of the RoW rules across states on the installation of telecom towers. Figure 5 plots the dynamic event-time coefficients using the Callaway and Sant'Anna (2021) estimator. The unit of analysis is at the district level. The outcome variable is the number of operational 4G BTSs in districts measured at the quarterly frequency from 2013 Q1 to 2021 Q4.[13] For robustness, we report results without any controls as well as after controlling for the natural logarithm of the area of the district. We include this control because districts vary in geographic size and may differ in their capacity to accommodate new towers.

Our results indicate that the adoption of RoW policy lead to a significant increase in the

[12] See here for more details.

[13] We exploit the granular information on set up of 4G BTSs in our dataset to move our analysis a step finer, up to the quarterly frequency. The annual frequency of our analysis so far has been dictated by the annual frequency of the EVI-implied agricultural yield.

number of 4G tower installations. Specifically, the number of towers within a district increased by 600 within six quarters of the adoption of the policy, and the effect tapers off thereafter. The increase in the number of operational 4G towers after the adoption of RoW policy is consistent with the anecdotal evidence that the RoW policy significantly reduced the regulatory and operational burden on telecom companies.

5.3 Effect of RoW Adoption on Agricultural Production

Next, we examine the effect of the adoption of the RoW policy on agricultural production. Figure 6 plots the dynamic event-study coefficients using our preferred Callaway and Sant'Anna (2021) estimator. There are two key takeaways from Figure 6. First, we do not observe significant pre-trends in EVI implied yields in the event time leading to the treatment. Second, we observe an increase in EVI implied yield following the adoption of RoW policy. Specifically, we observe that agricultural production increases by 0.8 times the standard deviation or 0.11 unit increase in EVI implied yield two years after the adoption of the RoW policy. This increase is equivalent to the 39.1% of the mean value of EVI implied yield. For robustness, we also supplement Figure 6 with OLS TWFE estimator, Borusyak, Jaravel, and Spiess (2022) estimator, and de Chaisemartin and D'Haultfœuille (2020) estimator. Overall, our results using the alternative identification strategy of staggered adoption of RoW policy indicate that our baseline results are unlikely to be driven by spatial endogeneity associated with the installation of 4G towers.

5.4 Effect on Operational Inputs: Fertilizer Consumption

This section documents the relationship between the entry of 4G and the usage of agricultural inputs such as fertilizers. The intuition behind this analysis is that the dissemination of information through 4G can decrease uncertainty, ultimately encouraging higher levels of investment (Bloom, 2009). Specifically, the introduction of 4G towers, which increases access to information about upcoming weather shocks, is expected to lead to higher usage of inputs as farmers are more well-equipped to navigate these shocks (Rosenzweig and Udry, 2013).

We use the annual district-level data on fertilizer usage to test the association between 4G penetration and fertilizer consumption. Specifically, our analysis examines the effect of staggered adoption of Rights of Way (RoW) rules across states on fertilizer consumption.

Figure 7a presents the results. We document an increase in the total fertilizer consumption after the adoption of Rights of Way rules that lead to the rapid introduction of 4G towers. Moreover,

the increase in fertilizer consumption does not seem to be driven by pre-trends. Specifically, we find that the total fertilizer consumption is driven by an increase in the consumption of phosphate and potash based fertilizers. Meanwhile, the consumption of nitrogen based fertilizers does not increase to a similar extent. Specifically, we find that the total fertilizer consumption per unit of gross sown area increased by 21.88% after the adoption of Rights of Way rules. While the consumption of nitrogen based fertilizers increased by 8.6%, the consumption of phosphate and potash based fertilizers increased by 72.72% and 89.54%.

This is not surprising given that nitrogen based fertilizers are extensively used in Indian agriculture whereas the usage of potash and phosphate based fertilizers is very limited. For instance, Panel C of Table 1 indicates that on average 101,102 kg of Nitrogen based fertilizers are used per hectare of gross sown area. In contrast, only 39,230 kg of phosphate and 14,230 kg of potash-based fertilizers are used per hectare of gross sown area.

5.5 Effect on Financial Input: Agricultural Bank Credit

This section examines the impact of 4G introduction on agricultural credit using branch-level data from one of the largest public-sector banks in India. Apart from the productivity shock at the farmer level, greater awareness, better monitoring, and reduced information asymmetry fostered by the introduction of 4G can lead to greater demand and supply for agricultural credit.

We combine the branch-level quarterly credit data with the dates of adoption of RoW policies by various states to examine the effect of 4G introduction on agricultural credit. Figure 7b presents the results using the Callaway and Sant'Anna (2021) estimator. The results suggest that agricultural bank credit increases by 20% one year after the implementation of the Right of Way (RoW) laws.

5.6 Effect on Price Dispersion

This section examines the effect of 4G deployment on price dispersion in agricultural markets. The intuition is that improved internet connectivity reduces information frictions between farmers and buyers, allowing the farmers to compare prices more effectively across markets. This increased transparency limits opportunities for buyers to set divergent prices, fostering price convergence and adherence to the Law of One Price. Consequently, regional price disparities for the same agricultural products are likely to reduce.

To test this hypothesis, we collect daily mandi-level prices for nine key crops: bajra, maize, jowar, wheat, paddy, soybean, tomato, potato, and onion. We measure price dispersion at the state

level by calculating the natural logarithm of the ratio between the maximum and the minimum value of modal price of each crop across markets within a state on the same date. We use state-crop as our unit of analysis because India's domestic agricultural trade is largely regulated to restrict farmers to selling only to government-licensed intermediaries within state-regulated markets (Chand, 2012; Chatterjee, 2023). Specifically, the Essential Commodities Act (1955) prohibits hoarding and the free movement of produce across state borders. Therefore, a state-crop is an appropriate unit of analysis, as farmers can primarily sell their produce within their own state to exploit any local price differences, given the regulatory constraints on inter-state agricultural trade.

Figure 8 presents the results combining the state-crop-month level measure of price dispersion with the staggered timing of RoW adoption across states to examine the effect of 4G introduction on price dispersion. The results suggest that price dispersion reduces after the implementation of the Right of Way (RoW) laws. Specifically, we document a 20% reduction in price dispersion two years after the adoption of the RoW policy. Moreover, the effect is unlikely to be driven by pre-trends. Overall, the effect on price dispersion indicates the role of the internet in improving market efficiency.

6 Mechanism

This section investigates the mechanism driving the increase in agricultural productivity following the introduction of 4G technology.

6.1 Improvement in Information Access

We document the significance of enhanced access to timely information due to the availability of high-speed 4G internet connectivity. Moreover, we show that high-speed internet in isolation might not be adequate to improve output. The ultimate impact of improved information accessibility on agricultural output hinges on factors such as the reliability, accuracy, and value of the information. Therefore, 4G infrastructure that improves access to information is likely to act as a complement to other information-generating infrastructure.

For the majority of this section, we will focus on weather information since agricultural output and returns heavily depend on weather shocks, particularly rainfall shocks. As such, access to information about the season's impending weather shocks could change the agricultural output. Specifically, access to the forecast of the season's weather shocks can allow farmers to choose inputs more appropriately leading to significant improvement in agricultural productivity (Rosenzweig and

Udry, 2013).[14] Additionally, information about the timing and the quantity of rainfall plays a crucial role in choosing the timing of plowing, planting, cultivating, fertilizing, and harvesting crops.

6.1.1 Accuracy & Reliability of Information

We posit that the introduction of high-speed 4G internet increases agricultural productivity by improving access to weather information for farmers. We test this claim by exploiting the heterogeneity in the distance to the nearest automatic weather station (AWS). Specifically, as the distance to the nearest AWS increases, the accuracy of the forecast decreases, thereby decreasing the accuracy or reliability of the information.

The proximity of a hexagon to the nearest AWS functions as a proxy for the accuracy and reliability of information, especially rainfall forecasts. The Indian Meteorological Department (IMD) combines observational data from various AWS with multiple global, regional, and meso-scale forecast models to generate weather forecasts. However, weather shocks, particularly rainfall, exhibit low geo-spatial correlation, as extensively documented in research on demand for index rainfall insurance (see Cole and Xiong (2017) for a comprehensive review).[15] This weak spatial correlation implies that forecasts for farmers in hexagons further from AWS are less reliable, impairing their capacity to make optimal pre-season input decisions. Such decisions, as Rosenzweig and Udry (2013) demonstrate, have a direct bearing on yields and economic returns. Furthermore, using primary data from India, Cole, Giné, and Vickery (2017) documents that farmers seem to be aware of the issue of low spatial correlation in rainfall due to the distance to the nearest weather station. Thus, the spatial proximity to AWS is one of the key determinants of whether the information received by farmers – rainfall forecasts in this case – is accurate or reliable.

According to the Ministry of Earth Sciences, India had a network of 697 AWS in 2014. We map each hexagonal unit to a unique AWS that is geographically closest. On average, each AWS is connected to approximately 42 hexagonal units. Panel D of Table 1 presents the summary statistics for the distance to the nearest AWS. The average distance between hexagons and their closest AWS is approximately 35 km, with a standard deviation of 19 km. Notably, there exists significant

[14] Prior work has documented that rainfall fluctuations can have a significant impact on aggregate consumption in village economies and agricultural productivity (Townsend (1994), Jayachandran (2006), Kaur (2019), among others). Rosenzweig and Udry (2013) find if rainfall were at the mean of the observed historical rainfall distribution in India, ₹10,000 increase in planting-stage investments would lead to an increase in profits of about ₹20,000 (over a base of ₹33,000). Similarly, if rainfall were at the 75th percentile, the same increase in planting-stage investment would generate an additional profit of ₹40,000.

[15] Hill, Robles, and Ceballos (2016) document that doubling the distance to a reference weather station increases basis risk and decreases insurance demand in India by 18%. Mobarak and Rosenzweig (2013) estimate that for every kilometer increase in the (perceived) distance of the weather station for a farmer without any informal risk protection, there is a drop-off in demand for formal index insurance of 6.4 percent.

variability in the distances to the closest AWS: the 25th percentile value stands at 20 km, while the 75th percentile value reaches 46 km. We divide the hexagonal units into three distinct sub-samples based on their proximity to the nearest weather station. The first sub-sample, termed low, includes hexagons with distances less than or equal to the 25th percentile value. The second sub-sample, medium, comprises hexagons with distances greater than the 25th percentile value but less than the 75th percentile value. Lastly, the third sub-sample, high, includes hexagons with distances above or equal to the 75th percentile value.

Figure 9a presents the results, and the table underneath reports the associated ATT estimates for each sub-sample along with the difference between the high and low sub-samples. As before, we do not observe significant pre-trends in EVI implied yields in the event time leading to the treatment. The most significant takeaway from this figure is the heterogeneity in the treatment effect across the three sub-samples. Hexagons closest to AWS exhibit the highest treatment effect of increased agricultural output after the introduction of high-speed 4G internet. Whereas hexagonal units far from AWS observe a statistically insignificant and economically negligible increase in agricultural productivity. Moreover, the difference between the point estimates for the high and low sub-samples is statistically significant, indicating that the treatment effect is higher in areas closer to the AWS.

Overall, our results suggest that the gains in productivity after 4G introduction are higher in areas where weather information is more accurate or reliable, i.e., the hexagons located closer to AWS.

6.1.2 Value of Information

We refine our understanding of the underlying mechanism by examining the heterogeneity in the treatment effects due to the potential value of information. We begin by constructing the historical sensitivity of each hexagon to droughts as a proxy for the value of information. The intuition of this assertion is that farmers operating in regions where production is more sensitive to such shocks stand to gain the most from rainfall forecasts and, thus, value weather information the most.

We use Standardized Precipitation and Evapotranspiration Index (SPEI) values mapped to our hexagons to construct measures of drought shocks covering the period from 2000 to 2012. Appendix A.5 details the methodology employed to construct these local drought shocks. To assess the impact of droughts on agricultural productivity, we estimate the sensitivity of each hexagon by regressing EVI-derived yields, based on 16-day historical composites, on the drought shocks. The R^2 from each regression serves as a measure of drought sensitivity, representing each hexagon's

proportion of variability in agricultural output explained by drought conditions.

We posit that hexagons with greater model R^2 are likely to value information about weather more, as the productivity in those areas is more sensitive to weather shocks. We divide the hexagonal units into three distinct sub-samples based on their model R^2. The first sub-sample, termed low, includes hexagons with model R^2 less than or equal to the 25th percentile value. The second sub-sample, medium, comprises hexagons with R^2 greater than the 25th percentile value but less than the 75th percentile value. Lastly, the third sub-sample, high, includes hexagons with R^2 above or equal to the 75th percentile value.

Figure 9b presents the results, and the table underneath reports the associated ATT estimates for each sub-sample along with the difference between the high and low sub-samples. As before, we do not observe significant pre-trends in EVI implied yields in the event time leading to the treatment. The most significant takeaway from this figure is the heterogeneity in the treatment effect across the three sub-samples. Hexagons with the highest model R^2 exhibit the highest treatment effect of increased agricultural output after the introduction of high-speed 4G internet. Whereas hexagonal units with low model R^2 exhibit smaller, or somewhat negative, treatment effects. Moreover, the difference between the point estimates for the high and low sub-samples is statistically significant, indicating that the treatment effect is higher in areas which have a higher model R^2 or greater sensitivity to droughts. This result suggests that the gains in productivity after 4G introduction are higher in areas where weather information is more valuable.

Our findings suggest that the diffusion of 4G infrastructure significantly increases agricultural productivity, especially in regions with high sensitivity to weather shocks. This has important policy implications for the growing body of research at the intersection of climate change and economics, emphasizing the role of digital connectivity as a form of climate adaptation. In particular, high-speed internet can serve as a vital tool for mitigating the adverse effects of increasingly frequent, severe, and persistent weather shocks associated with climate change, effects that disproportionately threaten the livelihoods of farmers in emerging economies. By providing access to timely and localized weather information, digital infrastructure functions as a buffer, reducing the volatility and negative impact of climate shocks on agricultural output and household income, highlighting the importance of integrating digital connectivity into climate resilience strategies. The implications of our findings – that improved information access can improve farmers' capacity to adapt to climate change – have recently been validated by the experimental evidence presented in Burlig et al. (2025).

Overall, there are two key takeaways from the analysis in this section. First, the results suggest

that one channel through which high-speed 4G internet increases agricultural output is through increased information accessibility. Second, in cases where the internet isn't coupled with the provision of reliable and valuable information, its potential to boost output may be constrained. Specifically, high-speed internet serves as the conduit for information, but the actual yield stems from the information. The positive effect of such infrastructure development that improves access to information crucially relies on the availability of reliable and valuable information. Therefore, the introduction of infrastructure to support high-speed internet in isolation might not be adequate, i.e., it acts as a complement to existing technology.

6.2 Technology Adoption: App Installations after 4G

This section examines the effect of 4G introduction on the adoption of 4G specific agricultural technology. Our proxy for 4G adoption is installation of agricultural applications that improve information access among farmers. Specifically, we exploit a proprietary dataset of geocoded and time-stamped mobile application ("app") installations from Krishify – an Indian firm which runs one of the largest social networks and commerce platforms for farmers. Appendix section A.6 presents a detailed description of app and the dataset.

Our dataset of Krishify app installation contains geolocations of users along with the dates on which they installed the app. The earliest date of app installation is May 2019. Thus, we subset our hexagon-level dataset to keep hexagons which got treated after May 2019 i.e. received their first 4G BTS after May 2019. We superimpose the geolocations of users' app installations on our hexagons and calculate the number of monthly installations at the hexagon level.

Using dates on which the first 4G BTS became operational as the treatment date, we run a difference-in-differences regression at the hexagonal level to study the effect of 4G *availability* on Krishify app installations. Figure 10 presents the results using the Callaway and Sant'Anna (2021) estimator. We observe a positive, statistically significant, and persistent increase in app installations starting from the first month after the introduction of 4G. Specifically, the likelihood of an app installation rises by 20% two years post-4G rollout. Moreover, we observe no significant pre-trends in the months leading up to the treatment, i.e. the introduciton of 3G or 4G. This result makes explicit what has been implicit in our analysis till now: availability of 4G is associated with 4G adoption, as proxied through the installation of the Krishify app.

6.2.1 What type of information do farmers access?

This section examines the type of information farmers access on the internet using detailed browsing and activity data of farmers on the Krishify app. The objective of this section is to present direct evidence suggesting that internet-based app allows farmers to access relevant information.

We use a 10% random sample of all farmers in Krishify. This random sample includes over 400,000 farmers and provides details of the browsing, search, likes, and comment history of all farmers since they joined the application. Additionally, we also observe information on their land holdings and their risk-taking tendency. Appendix section A.6 presents a detailed description of the dataset.

We categorize all interactions from these farmers into ten distinct topics: weather updates and agricultural news; government program information; farm machinery; animal husbandry; inputs such as fertilizer and seeds; crop management; financial services; market prices; as well as recreational and political content. We consolidate weather information and agricultural news into a single category, as the majority of agricultural news concerns weather conditions.Table 2 presents the list of these topics.

The second column of Table 2 presents the percentage of interactions for each topic among our sample farmers. The foremost three topics collectively constitute about 70% of farmer interactions on the platform. Notably, weather information and agricultural news rank highest, contributing to 30.9% of all interactions, followed by government programs at 21.3%, and information about farm machinery (e.g., tractors) at 16.1%. Furthermore, farmers also utilize the platform to seek advice specific to animalhusbandry (9.1%) and crop management (8.18%). Interactions related to inputs like fertilizers, seeds, finance, and other miscellaneous inputs collectively make up 12.8% of all farmer engagements on the platform. Columns (2)-(6) and columns (7)-(8) report numbers for farmers with different land ownership and risk-taking abilities, respectively, and find that the topic-wise interaction of farmers is similar across these divisions.

The detailed activity data of farmers suggests that farmers may be using internet-based platforms extensively for gathering information on a wide variety of agri-related activities. This descriptive analysis, coupled with the findings presented in section 6.2, reinforces the confidence in our main finding that the introduction of 4G technology imporved information accessibility for farmers.

6.2.2 What type of farmers are more likely to use such services?

The objective of this section is to discuss some correlations indicating what type of farmers are more likely to use high-speed internet-powered services. Specifically, we examine the correlation between the activity on the Krishify app and farmer characteristics such as landholding and risk-taking ability. We use two measures of the level of adoption – (1) engagement score which measures the total activity of the farmer on the app, and (2) direct measures of active engagement such as likes, comments, and videos posted.

First, we use the engagement score to measure the amount of total activity of the farmer on the app (see Appendix Figure A.9). The engagement score is a composite proprietary measure created by the firm to account for the total activity of the farmer based on their likes, comments, posts, and shares. Column 1 of Table 3 presents the results from the Poisson regression using the engagement score as the dependent variable. We include cohort $\times$ district fixed effects effectively comparing farmers in the same district that installed the application during the same month. Including this fixed effect is crucial because internet engagment might be mechanically higher for farmers who have been using the platform for a longer duration.

Two key facts emerge. First, land-owning farmers have greater total engagement relative to landless or tenant farmers. However, conditional on owning land, small and medium farmers have a greater engagement relative to large farmers with land above ten acres. Second, farmers with higher risk-taking ability have greater overall engagement.

Next, columns 2, 3 and 4 of Table 3 use the number of likes, comments and videos posted on the application as dependent variable, respectively. As in column 1, all specifications include cohort $\times$ district fixed effects. Farmers with higher levels of risk-taking ability are associated with higher levels of active engagement. Farmers who own less than two acres of land show lower levels of active engagement compared to landless or tenant farmers. However, active engagement tends to increase in land-holding as land holding size surpasses two acres. This relationship between land-holding and active engagement is different from the relationship between land-holding and total engagement presented in column 1, indicating that land-holding size may be associated with the type of engagement on internet-based platforms.

7 Is High-Speed Internet Special?

This section examines the importance of the introduction of 4G relative to earlier versions of telecommunications technology, such as 2G and 3G. The intuition for comparing the different

technologies arises from the commonality that while all have the potential to improve access to information, 4G presents several key advantages over its predecessors. First, 4G internet access can dramatically improve real-time information availability for farmers, a critical factor highlighted by Anderson and Feder (2004), who argue that the impact of information dissemination on agricultural productivity depends largely on farmers' ability to access relevant data promptly. Second, internet-based platforms can provide valuable insights in a user-friendly and engaging manner by employing videos and graphics, and evidence shows that video-based interventions improve farmers' knowledge and practices (Vasilaky et al., 2015; Van Campenhout, Spielman, and Lecoutere, 2021). Third, internet-based platforms do not face limitations similar to those faced by traditional text or call-based information delivery systems. For example, farmers might encounter challenges in reading text messages due to illiteracy or experiencing delays in connecting with attendants at farmer call centers.

We evaluate this hypothesis by comparing the impact of 2G, 3G, and 4G introduction on agricultural productivity. Figures 11a and 11b present the effects of 3G and 2G introduction, respectively. Unlike the significant improvements in productivity observed after the introduction of 4G, we do not find an economically meaningful or statistically significant increase in productivity following the introduction of 3G or 2G.

The null effects following the introduction of 2G and 3G align with prior research, which consistently finds little evidence that traditional information methods improve productivity. Casaburi et al. (2014) show that sending text messages with agricultural advice to small sugarcane farmers in Kenya didn't result in long-term positive effects on yields. In an experiment in the Indian state of Gujarat, Cole and Fernando (2021) find no systematic impact on yields on the treated farmers that randomly received access to a hotline for agricultural advice. Additionally, experiments by Fafchamps and Minten (2012) that sent Indian farmers price and weather information via text messages found no evidence of value addition. Similar results were documented by Camacho and Conover (2010) among Colombian farmers. We direct readers to Fabregas, Kremer, and Schilbach (2019) for a more detailed review of the effects of extension workers, phone calls and SMS-based information programs on agricultural yields.

While Anderson and Feder (2004) and Duflo, Kremer, and Robinson (2011) attribute the poor performance of traditional face-to-face programs to the lack of timely, personalized information for farmers, Fabregas, Kremer, and Schilbach (2019) argue that their null results may stem from small sample sizes and insufficient statistical power, increasing the risk of Type II errors. In contrast,

our analysis leverages a large, regionally and temporally extensive dataset across India, utilizing granular 16-day composites to reconstruct EVI-based yields dating back to 2000. This broad scope captures the majority of variation associated with the rollout of 2G and 3G, as shown in Figure 2, making it less susceptible to the underpowered critique. Consequently, our findings not only bolster external validity but also suggest that prior null results are unlikely to be solely due to lack of power. Overall, our evidence supports the view that high-speed internet, unlike earlier telecom technologies, is special and has the potential to activate channels that can meaningfully boost agricultural productivity.

7.1 Why Might High-Speed Internet be Special?: Role of Trust

This section examines a potential channel that may make high-speed internet access, due to introduction of 4G, an effective tool for information dissemination. Specifically, we examine the heterogeneity in the treatment effect of 4G introduction by the trust in local and state authorities.

Information dissemination through traditional modes such as text messages, call centres, hotlines, etc. is often directly controlled by the state and its agents. Moreover, local governing bodies known as village panchayats often act as the central node for information dissemination by the government. Therefore, a lack of trust in local and state authorities can hinder effective communication and engagement through these traditional methods (Fabregas, Kremer, and Schilbach, 2019). For instance, using a survey of cotton-producing farmers in Gujarat, Cole and Sharma (2017) document that nearly 70% of Indian farmers distrust information provided by extension workers. On the contrary, the internet, being a more decentralized platform, may not encounter the same issues related to trust.[16] Consequently, the introduction of 4G internet may have greater effects in regions where trust in state institutions is low,

We investigate this hypothesis by combining the India Human Development Survey (IHDS) with our dataset. Specifically, we utilize data from IHDS-II, conducted in 2011, which includes a question that queries respondents about the level of trust they have in their local village panchayats.[17] We use the response to this question to construct a district-level proxy for confidence in village panchayats and divide the districts into two subsets based on the median response value. Districts

[16] For instance, while looking at the Apple Weather App many may not realize that the app ultimately sources all its data from the National Weather Service, which is part of the National Oceanic and Atmospheric Administration (NOAA), a scientific and regulatory agency within the United States Department of Commerce, a United States federal government department.

[17] The exact question is as follows: "How much trust do you place in the ability of your village panchayat to implement public projects?" The response options consist of three choices: a) A great deal of confidence, b) Only some confidence, and c) Hardly any confidence at all. To quantify these responses, we assign a numerical value of 1 to option (a), 0.33 to option (b), and 0 to option (c). We take the average of this measure across all households within a district based on their weights. This allows us to compute a continuous measure of confidence in village panchayats at the district level.

with confidence levels above the median are categorized as "high confidence" districts, while those below the median are classified as "low confidence" districts. We then link this measure to the hexagonal grids situated within the districts.

Figure 12a presents the results. The outcomes for districts with low confidence levels are displayed in red, while the results for districts with high confidence levels are shown in blue. The findings suggest that regions characterized by low trust in village panchayats exhibit a notable increase in agricultural production. In contrast, the impact of introducing 4G technology is economically marginal and lacks statistical significance in areas with a high level of trust in local village panchayats.

We also present results using an alternative measure of trust, i.e., the confidence in the state government to look after people.[18] Similar to before, in regions where trust in state government is low, the marginal benefit of the introduction of 4G internet for information-sharing is high. Figure 12b presents the results. The findings suggest that regions characterized by low trust in state government exhibit a notable increase in agricultural production. In contrast, the impact of introducing 4G technology is economically marginal and lacks statistical significance in areas with a high level of trust in state government.

Overall these results imply that the advancement of high-speed internet through 4G infrastructure can mitigate the effect of human frictions, like mistrust in local institutions. These barriers typically hinder the dissemination of information through traditional channels.

8 Conclusion

Can high-speed internet boost information access and enhance productivity? We examine this question in the context of Indian agriculture where information frictions, especially frictions related to information dissemination, are rampant. Combining granular geographic data on the introduction of 4G with remote-sensing data on agricultural productivity we show that the improvement in information dissemination due to the introduction of 4G leads to improvement in productivity.

Our analysis reveals that the introduction of high-speed 4G internet is associated with significant improvements in agricultural outcomes, including - (1) a 1.06% increase in the yields of cereal crops, (2) a 4.28% increase in the production value of cereal crops, and (3) a 3.21% increase in the production value of all crops. Six years post the introduction of 4G internet, agricultural households

[18] The two variables trust in village panchayat and trust in state government are highly correlated with each other with a correlation coefficient of 0.632. We think of these two measures as alternative ways to target a similar economic parameter, i.e., the trust in institutions that are the primary modes of the traditional information dissemination processes.

experienced a substantial 14.50% growth in annual income. Our calculations attribute 24.56% of this income growth from 2014 to 2021 to the improved accessibility of information. Moreover, we find that 4G introduction leads to increased fertilizer consumption and greater credit uptake.

Our study indicates that the effectiveness of information dissemination is contingent on the usefulness, reliability, accuracy, and value of the information. Additionally, areas with lower trust in local and state institutions, experience a stronger impact from the introduction of high-speed 4G internet.

Overall our results indicate that while high-speed internet acts as an important tool for information dissemination, our findings underscore the importance of considering the quality of information. Merely introducing internet infrastructure may not be sufficient; instead, it should be viewed as a complement to other factors. Thus, comprehensive infrastructure development that enhances information accessibility along with the availability, reliability, and accuracy of information holds the key to maximizing the effect.

References

Aker, Jenny C and Isaac M Mbiti. 2010. "Mobile phones and economic development in Africa." *Journal of economic Perspectives* 24 (3):207–232.

Akerman, Anders, Ingvil Gaarder, and Magne Mogstad. 2015. "The skill complementarity of broadband internet." *The Quarterly Journal of Economics* 130 (4):1781–1824.

Allen, Treb. 2014. "Information frictions in trade." *Econometrica* 82 (6):2041–2083.

Andersen, Thomas Barnebeck, Jeanet Bentzen, Carl-Johan Dalgaard, and Pablo Selaya. 2011. "Does the Internet reduce corruption? Evidence from US states and across countries." *The World Bank Economic Review* 25 (3):387–417.

Anderson, Jock R and Gershon Feder. 2004. "Agricultural extension: Good intentions and hard realities." *The World Bank Research Observer* 19 (1):41–60.

Asher, Sam and Paul Novosad. 2020. "Rural roads and local economic development." *American economic review* 110 (3):797–823.

Batista, Catia and Pedro C Vicente. 2023. "Is mobile money changing rural Africa? Evidence from a field experiment." *Review of Economics and Statistics* :1–29.

Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry. 2013. "Profitability of fertilizer: Experimental evidence from female rice farmers in Mali." *American Economic Review* 103 (3):381–386.

Birch, Colin PD, Sander P Oom, and Jonathan A Beecham. 2007. "Rectangular and hexagonal grids used for observation, experiment and simulation in ecology." *Ecological modelling* 206 (3-4):347–359.

Bloom, Nicholas. 2009. "The impact of uncertainty shocks." *econometrica* 77 (3):623–685.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2022. "Revisiting Event Study Designs: Robust and Efficient Estimation."

Braghieri, Luca, Ro'ee Levy, and Alexey Makarin. 2022. "Social media and mental health." *American Economic Review* 112 (11):3660–3693.

Brooks, Wyatt and Kevin Donovan. 2020. "Eliminating uncertainty in market access: The impact of new bridges in rural Nicaragua." *Econometrica* 88 (5):1965–1997.

Brunnermeier, Markus K, Nicola Limodio, and Lorenzo Spadavecchia. 2023. "Mobile money, interoperability, and financial inclusion." Tech. rep., National Bureau of Economic Research.

Burlig, Fiona, Amir Jina, Erin M Kelley, Gregory Lane, and Harshil Sahai. 2025. "The value of forecasts: Experimental evidence from India." .

Callaway, Brantly and Pedro HC Sant'Anna. 2021. "Difference-in-differences with multiple time periods." *Journal of Econometrics* 225 (2):200–230.

Camacho, Adriana and Emily Conover. 2010. "The impact of receiving price and climate information in the agricultural sector." *Documento CEDE* (2010-40).

Casaburi, Lorenzo, Michael Kremer, Sendhil Mullainathan, and Ravindra Ramrattan. 2014. "Harnessing ICT to increase agricultural production: Evidence from Kenya." *Harvard University* .

Chand, Ramesh. 2012. "Development policies and agricultural markets." *Economic and Political Weekly* :53–63.

Chatterjee, Shoumitro. 2023. "Market Power and Spatial Competition in Rural India." *The Quarterly Journal of Economics* URL https://doi.org/10.1093/qje/qjad004. Qjad004.

Cole, Shawn and A Nilesh Fernando. 2021. "'Mobile'izing agricultural advice technology adoption diffusion and sustainability." *The Economic Journal* 131 (633):192–219.

Cole, Shawn, Xavier Giné, and James Vickery. 2017. "How does risk management influence production decisions? Evidence from a field experiment." *The Review of Financial Studies* 30 (6):1935–1970.

Cole, Shawn and G Sharma. 2017. "The promise and challenges in implementing ICT for agriculture." Tech. rep., Working paper. Available at: https://precisionag. org/uploads/cole-sharma ….

Cole, Shawn and Wentao Xiong. 2017. "Agricultural insurance and economic development." *Annual review of Economics* 9:235–262.

Cramer, Kim Fe, Pulak Ghosh, Nirupama Kulkarni, and Nishant Vats. 2024. "Shadow Banks on the Rise: Evidence Across Market Segments." *Olin Business School Center for Finance & Accounting Research Paper (2024/18)* .

D'Andrea, Angelo, Patrick Hitayezu, Roland Kangni Kpodar, Nicola Limodio, and Andrea Presbitero. 2024. "Mobile Internet, Collateral, and Banking." Tech. rep., International Monetary Fund.

De Chaisemartin, Clément and Xavier d'Haultfoeuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review* 110 (9):2964–2996.

de Chaisemartin, Clément and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110 (9):2964–96. URL https://www.aeaweb.org/articles?id=10.1257/aer.20181169.

Deaton, Angus. 1990. "On Risk, Insurance and Intra-Village Consumption Smoothing." *Research Program in Development Studies. Princeton University* .

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2011. "Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya." *American economic review* 101 (6):2350–2390.

D'Andrea, Angelo and Nicola Limodio. 2023. "High-Speed Internet, Financial Technology, and Banking." *Management Science* .

Emerick, Kyle, Alain De Janvry, Elisabeth Sadoulet, and Manzoor H Dar. 2016. "Technological innovations, downside risk, and the modernization of agriculture." *American Economic Review* 106 (6):1537–1561.

Fabregas, Raissa, Michael Kremer, and Frank Schilbach. 2019. "Realizing the potential of digital development: The case of agricultural advice." *Science* 366 (6471):eaay3038.

Fafchamps, Marcel and Bart Minten. 2012. "Impact of SMS-based agricultural information on Indian farmers." *The World Bank Economic Review* 26 (3):383–414.

Ghosh, Pulak, Apoorv Gupta, Jacopo Ponticelli, and Andrea Tesei. 2025. "Information Frictions and Take-up of Government Credit Programs." .

Ghosh, Pulak and Nishant Vats. 2022. "Safety Nets, Credit, and Investment: Evidence from a Guaranteed Income Program." *University of Chicago Working Paper* .

Goldfarb, Avi and Catherine Tucker. 2019. "Digital economics." *Journal of economic literature* 57 (1):3–43.

Gollin, Douglas, Casper Worm Hansen, and Asger Mose Wingender. 2021. "Two blades of grass: The impact of the green revolution." *Journal of Political Economy* 129 (8):2344–2384.

Gonzalez, Robert M. 2021. "Cell phone access and election fraud: evidence from a spatial regression discontinuity design in Afghanistan." *American Economic Journal: Applied Economics* 13 (2):1–51.

GSMA. 2020. "Rights of Way (RoW) in India: A submission from the GSMA relating to India Mobile Sector." Tech. rep., GSMA.

Guiteras, Raymond. 2009. "The impact of climate change on Indian agriculture." *Working Paper, University of Maryland* .

Gunsilius, Florian and David Van Dijcke. 2023. "Free Discontinuity Regression: With an Application to the Economic Effects of Internet Shutdowns." *arXiv preprint arXiv:2309.14630* .

Gupta, Apoorv, Jacopo Ponticelli, and Andrea Tesei. 2024. "Language barriers, technology adoption and productivity: Evidence from agriculture in india." *Review of Economics and Statistics* :1–28.

Guriev, Sergei, Nikita Melnikov, and Ekaterina Zhuravskaya. 2021. "3g internet and confidence in government." *The Quarterly Journal of Economics* 136 (4):2533–2613.

Hill, Ruth Vargas, Miguel Robles, and Francisco Ceballos. 2016. "Demand for a simple weather insurance product in India: theory and evidence." *American Journal of Agricultural Economics* 98 (4):1250–1270.

Hjort, Jonas and Jonas Poulsen. 2019. "The arrival of fast internet and employment in Africa." *American Economic Review* 109 (3):1032–1079.

Hjort, Jonas and Lin Tian. 2021. "The economic impact of internet connectivity in developing countries." *Annual Review of Economics* 17.

Huete, Alfredo, Kamel Didan, Tomoaki Miura, E Patricia Rodriguez, Xiang Gao, and Laerte G Ferreira. 2002. "Overview of the radiometric and biophysical performance of the MODIS vegetation indices." *Remote sensing of environment* 83 (1-2):195–213.

Jayachandran, Seema. 2006. "Selling labor low: Wage responses to productivity shocks in developing countries." *Journal of political Economy* 114 (3):538–575.

Jensen, Robert. 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.

Jensen, Robert and Nolan H Miller. 2018. "Market integration, demand, and the growth of firms: Evidence from a natural experiment in India." *American Economic Review* 108 (12):3583–3625.

Kaur, Supreet. 2019. "Nominal wage rigidity in village labor markets." *American Economic Review* 109 (10):3585–3616.

Keleher, Niall, Arman Rezaee, and Erin Troland. 2020. "The impact of mobile phones: Experimental evidence from the random assignment of new cell towers." .

Manacorda, Marco and Andrea Tesei. 2020. "Liberation technology: Mobile phones and political mobilization in Africa." *Econometrica* 88 (2):533–567.

Mehrabi, Zia, Mollie J McDowell, Vincent Ricciardi, Christian Levers, Juan Diego Martinez, Natascha Mehrabi, Hannah Wittman, Navin Ramankutty, and Andy Jarvis. 2021. "The global divide in data-driven farming." *Nature Sustainability* 4 (2):154–160.

Mobarak, Ahmed Mushfiq and Mark R Rosenzweig. 2013. "Informal risk sharing, index insurance, and risk taking in developing countries." *American Economic Review* 103 (3):375–380.

Parker, Chris, Kamalini Ramdas, and Nicos Savva. 2016. "Is IT enough? Evidence from a natural experiment in India's agriculture markets." *Management Science* 62 (9):2481–2503.

Rosenzweig, Mark and Christopher R Udry. 2013. "Forecasting profitability." Tech. rep., National Bureau of Economic Research.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe. 2023. "What's trending in difference-in-differences? A synthesis of the recent econometrics literature." *Journal of Econometrics* .

Steinwender, Claudia. 2018. "Real effects of information frictions: When the states and the kingdom became united." *American Economic Review* 108 (3):657–696.

Sun, Liyang and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics* 225 (2):175–199.

Suri, Tavneet. 2017. "Mobile money." *Annual Review of Economics* 9:497–520.

Suri, Tavneet and William Jack. 2016. "The long-run poverty and gender impacts of mobile money." *Science* 354 (6317):1288–1292.

Tian, Lin. 2021. "Division of labor and productivity advantage of cities: Theory and evidence from Brazil." .

Townsend, Robert M. 1994. "Risk and insurance in village India." *Econometrica: journal of the Econometric Society* :539–591.

Udry, Christopher, Federica di Battista, Mathias Fosu, Markus Goldstein, Alev Gurbuz, Dean Karlan, and Shashidhara Kolavalli. 2019. "Information, market access and risk: Addressing constraints to agricultural transformation in northern Ghana." *Draft Report* .

Van Campenhout, Bjorn, David J Spielman, and Els Lecoutere. 2021. "Information and communication technologies to provide agricultural advice to smallholder farmers: Experimental evidence from Uganda." *American Journal of Agricultural Economics* 103 (1):317–337.

Vasilaky, Kathryn, Kentaro Toyama, Tushi Baul, Mohak Mangal, and Urmi Bhattacharya. 2015. "Learning Digitally: Evaluating the Impact of Farmer Training via Mediated Videos." .

Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov. 2020. "Political effects of the internet and social media." *Annual review of economics* 12 (1):415–438.

Zuo, George W. 2021. "Wired and hired: Employment effects of subsidized broadband Internet for low-income Americans." *American Economic Journal: Economic Policy* 13 (3):447–482.

Table 1: Summary Statistics

Panel A: Agricultural Productivity Measures based on 8-day composites (2013-2022)								
	# Obs	Min	p25	p50	p75	Max	Mean	SD
EVI Implied Yield	255,330	0.000	0.165	0.267	0.369	0.700	0.270	0.144
Maximum Yield	255,330	0.158	0.511	0.602	0.676	0.896	0.584	0.138
Panel B: Agricultural Productivity Measures based on 16-day composites (2000-2022)								
	# Obs	Min	p25	p50	p75	Max	Mean	SD
EVI Implied Yield	633,144	26.287	1336.688	2013.979	2729.066	5185.972	2064.124	998.134
Maximum Yield	633,144	182.902	3805.398	4553.492	5177.993	7039.151	4397.172	1189.108
Panel C: Standardized Precipitation Evapotranspiration Index (SPEI) (2000-2012)								
	# Obs	Min	p25	p50	p75	Max	Mean	SD
SPEI	357,851	-2.431	-0.543	-0.208	0.121	1.511	-0.215	0.480
Panel D: Unit (Hexagon) Level Measures								
	# Obs	Min	p25	p50	p75	Max	Mean	SD
Agricultural Landcover	28,370	1.000	4.000	4.000	7.000	8.000	5.064	2.251
Dist. to Automatic Weather Station (km)	28,370	3.645	20.449	32.166	46.168	92.507	34.683	18.852
Sensitivity to Droughts	27,055	0.000	0.055	2.244	9.028	78.221	6.387	9.421
Panel E: District Level Measures								
	# Obs	Min	p25	p50	p75	Max	Mean	SD
Trust (Village Panchayat)	382	0.056	0.368	0.447	0.553	0.967	0.463	0.141
Trust (State Government)	382	0.070	0.374	0.451	0.567	0.906	0.468	0.148
LN(Gross Sown Area, GSA in hectares)	5,649	1.723	4.873	5.301	5.808	7.400	5.277	0.780
Total Fertilizer (in '000 kg)/GSA	5,649	0.189	63.227	124.625	221.007	789.677	155.147	121.657
Nitrogen Fertilizer (in '000 kg)/GSA	5,649	0.126	37.432	75.876	144.535	515.728	101.102	85.767
Phosphate Fertilizer (in '000 kg)/GSA	5,649	0.031	17.523	32.564	55.007	232.372	39.230	29.664
Potash Fertilizer (in '000 kg)/GSA	5,649	0.005	2.126	7.440	17.732	139.473	14.230	18.905
Panel F: Branch-Level Agricultural Credit								
	# Obs	Min	p25	p50	p75	Max	Mean	SD
LN(Agricultural Credit)	133,056	12.2326	16.9629	18.0691	19.0249	20.5999	17.8341	1.5743

This table reports the descriptive statistics for the key variables used in the analysis. Panel A reports the summary statistics for agricultural productivity variables based on 8-day EVI composites from 2013 to 2022. Panel B reports the summary statistics for agricultural productivity variables based on 16-day EVI composites from 2000 to 2022. Panel C reports the summary statistics of the SPEI index at the hexagon-year level from 2000 to 2012. Panel D reports the summary statistics for unit (hexagon) level variables. Panel E reports variables measured at the district level. Panel F reports the summary statistics for branch-level agricultural credit. All variables are winsorized at the 1% level.

Table 2: Topic-wise Interaction of Farmers on Krishify

Topic	Overall	By Land Ownership					By Risk-Taking Ability		
		No Land	Below 2 Acres	2-5 Acres	5-10 Acres	Above 10 Acres	Low	Medium	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Weather Info & Ag News	30.89%	31.96%	29.78%	29.55%	29.84%	31.59%	34.51%	26.29%	27.60%
Government Programs	21.29%	21.24%	21.98%	20.70%	21.13%	20.52%	21.46%	22.71%	19.60%
Farm Machinery	16.09%	15.92%	15.05%	16.62%	18.09%	17.85%	18.12%	13.73%	14.03%
Animal Husbandry	9.07%	8.39%	11.78%	8.55%	7.38%	6.84%	5.30%	14.10%	12.27%
Fertilizers, Seeds & Oth Inputs	8.48%	9.01%	7.21%	9.02%	8.32%	8.24%	8.22%	7.74%	9.71%
Crop Management	8.18%	7.41%	8.23%	9.74%	9.08%	9.16%	7.01%	8.69%	10.15%
Finance	2.96%	2.67%	3.43%	3.09%	3.08%	2.90%	2.20%	3.99%	3.60%
Price Information	1.39%	1.53%	1.08%	1.28%	1.59%	1.46%	1.45%	1.27%	1.36%
Recreation	0.89%	0.92%	0.84%	0.91%	0.92%	0.73%	0.82%	0.95%	0.97%
Politics	0.77%	0.96%	0.63%	0.54%	0.57%	0.70%	0.91%	0.54%	0.71%

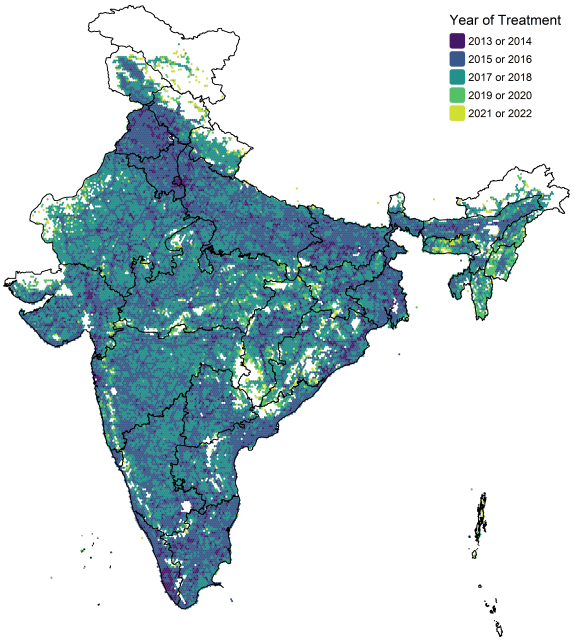
This table presents the percentage of interactions for each topic among our sample farmers. We classify all interactions of these farmers into 13 distinct topics which include – agricultural news, information about government programs, farm machinery, crop management, weather, animal husbandry, fertilizer, seeds, prices, and other inputs, as well as content that is recreational and political. Column (1) reports the numbers of all farmers. Columns (2)-(6) present the numbers by different land-ownership buckets. Land ownership is divided into five buckets no land ownership, land below 2 acres, land between 2 to 5 acres, land between 5 to 10 acres and land above 10 acres. Columns (7)-(9) reports the numbers for farmers in different buckets of risk-taking ability. Risk-taking ability is divided into three categories of low, medium, and high.

Table 3: Level of Active Engagement & Farmer Characteristics

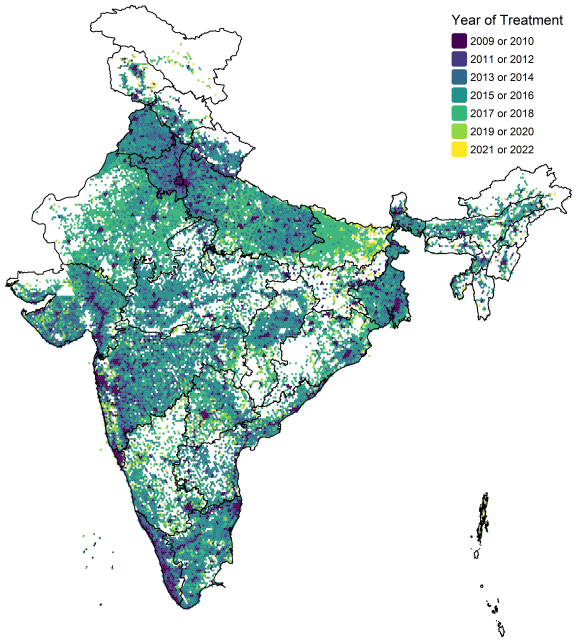
	(1)	(2)	(3)	(4)
	Engagement Score	# Likes	# Comments	# Videos Posted
Land, Below 2 Acres	0.0631***	-0.0415***	-0.0373***	-0.0396
	(0.0037)	(0.0130)	(0.0140)	(0.0429)
Land, 2-5 Acres	0.0776***	0.0905***	0.1006***	0.0795*
	(0.0048)	(0.0152)	(0.0147)	(0.0478)
Land, 5-10 Acres	0.0677***	0.1159***	0.1225***	0.2170***
	(0.0055)	(0.0173)	(0.0181)	(0.0578)
Land, Above 10 Acres	0.0175***	0.1466***	0.0474**	0.2648***
	(0.0060)	(0.0188)	(0.0187)	(0.0587)
Risk-Taking Ability, Medium	0.2160***	0.7682***	1.1795***	1.3651***
	(0.0038)	(0.0137)	(0.0131)	(0.0377)
Risk-Taking Ability, High	0.2846***	0.9317***	1.2855***	1.4531***
	(0.0044)	(0.0142)	(0.0148)	(0.0351)
District X Cohort FE	Yes	Yes	Yes	Yes
# Obs	405,717	402,828	396,728	276,631
Pseudo R2	0.071	0.0976	0.1203	0.0904
Dep Var Sample	1882.2690	1.6178	0.3676	0.0193
Dep Var Sample SD	1640.7710	4.5000	1.0400	0.1377

This table presents the relationship between farmer-level active engagement on the Krishify application and their characteristics. The dependent variable is engagement score in column (1), the total number of likes in column (2), total number of comments in column (3), and total number of videos posted in column (4). Engagement score is provided by Krishify and is a proprietary score that measures the total amount of engagement of the farmer on the platform. Land ownership is divided into five buckets no land ownership, land below 2 acres, land between 2 to 5 acres, land between 5 to 10 acres and land above 10 acres. For land ownership, farmers with no land ownership are the omitted category. Risk-taking ability is divided into three categories of low, medium, and high. Low risk-taking ability is the omitted category. We use Poisson regressions to estimate the relationship. Cohort refers to the cohort of farmers who install the app during the same month. District refers to the district of farmer's location. All variables are winsorized at the 1% level. Standard errors in parentheses are estimated by clustering at the district level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

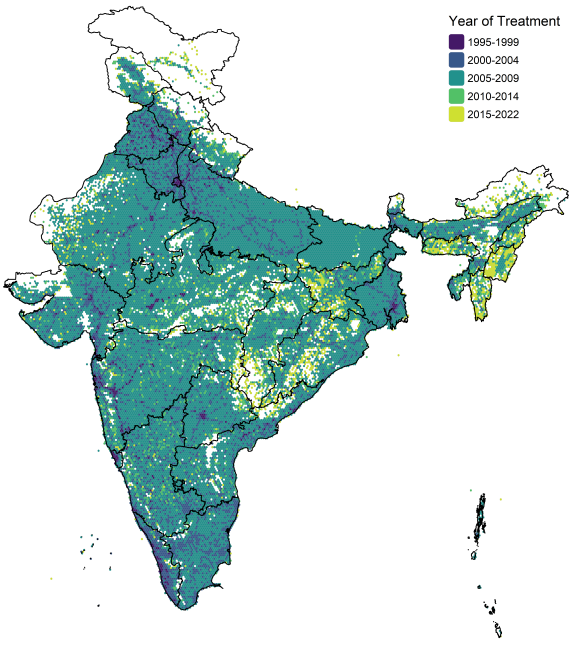
Figure 1: Evolution of 4G, 3G, and 2G Treatment Across Hexagons



(a) Evolution of 4G coverage



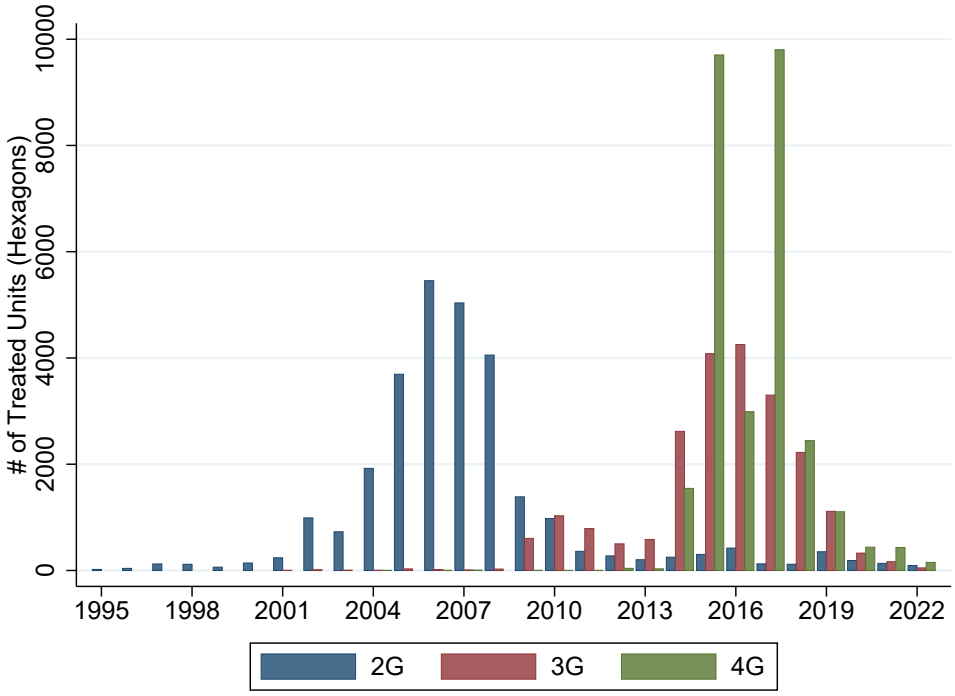
(b) Evolution of 3G coverage



(c) Evolution of 2G coverage

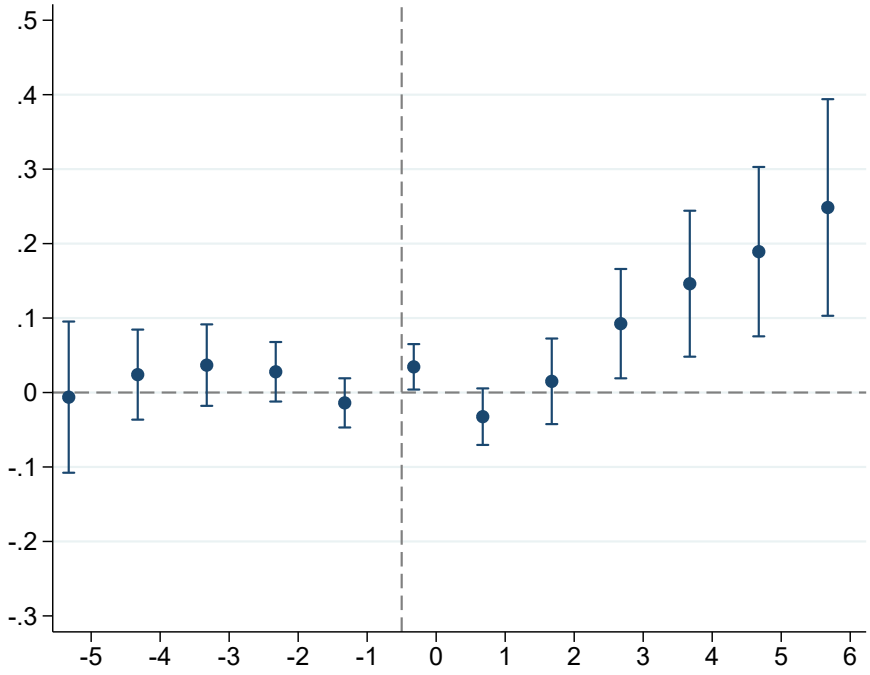
This figure presents a tessellated heat map of India. Each color denotes the year in which the 4G, 3G, and 2G technologies were introduced in the hexagons for the first time. Panels 1a, 1b, and 1c plot the treatment years for 4G, 3G, and 2G, respectively.

Figure 2: Staggered introduction of BTS by Technology



This figure presents the staggered introduction of BTs by technology type from 1995 until 2022. The X-axis denotes the calendar year. The Y-axis denotes the number of units (hexagons) treated by technology type for each year. The three technology types include 2G, 3G, and 4G.

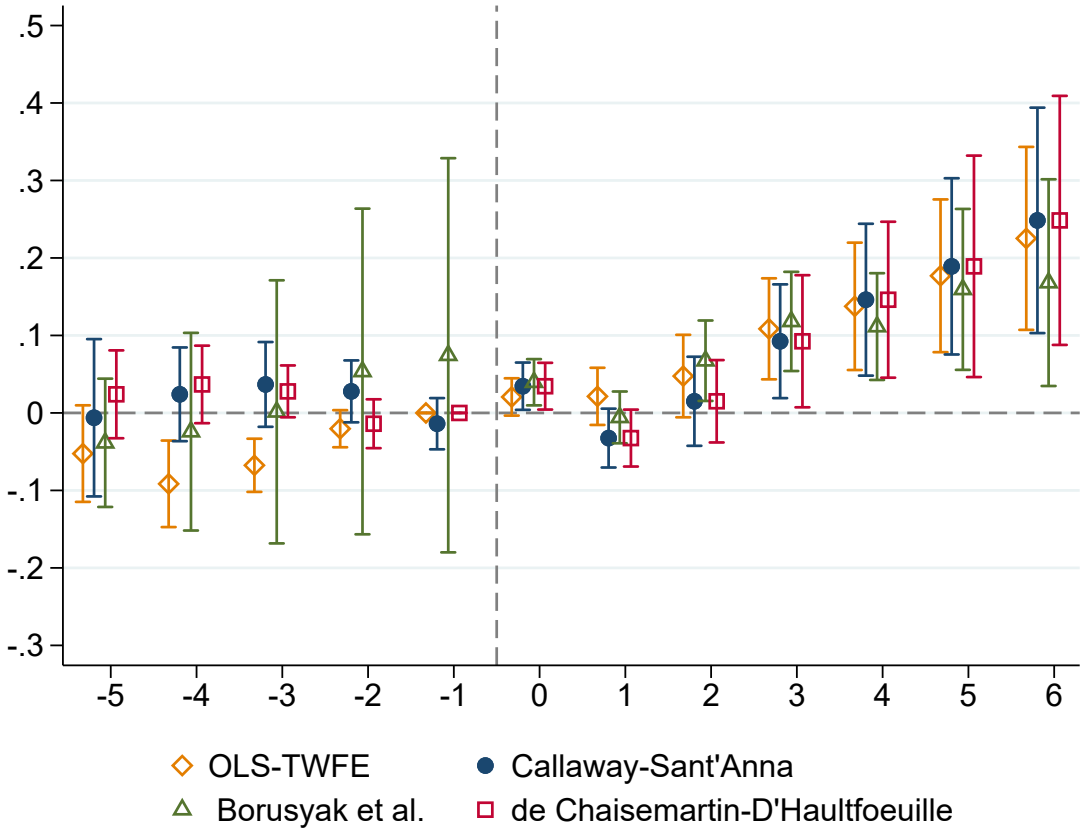
Figure 3: Effect of 4G introduction on EVI implied agricultural yield



Specification	Partially Aggregated							ATT
Simple								0.053**
								(0.022)
Group-specific	g=2014	g=2015	g=2016	g=2017	g=2018	g=2019	g=2020	
	0.050	0.053*	0.082**	0.078**	-0.123**	0.057	-0.105	0.047**
	(0.035)	(0.032)	(0.039)	(0.039)	(0.054)	(0.055)	(0.075)	(0.022)
Event study	e = 0	e = 1	e = 2	e = 3	e = 4	e = 5	e = 6	
	0.034**	-0.033*	0.014	0.090**	0.145***	0.186***	0.245***	0.097***
	(0.015)	(0.019)	(0.028)	(0.037)	(0.048)	(0.057)	(0.073)	(0.028)

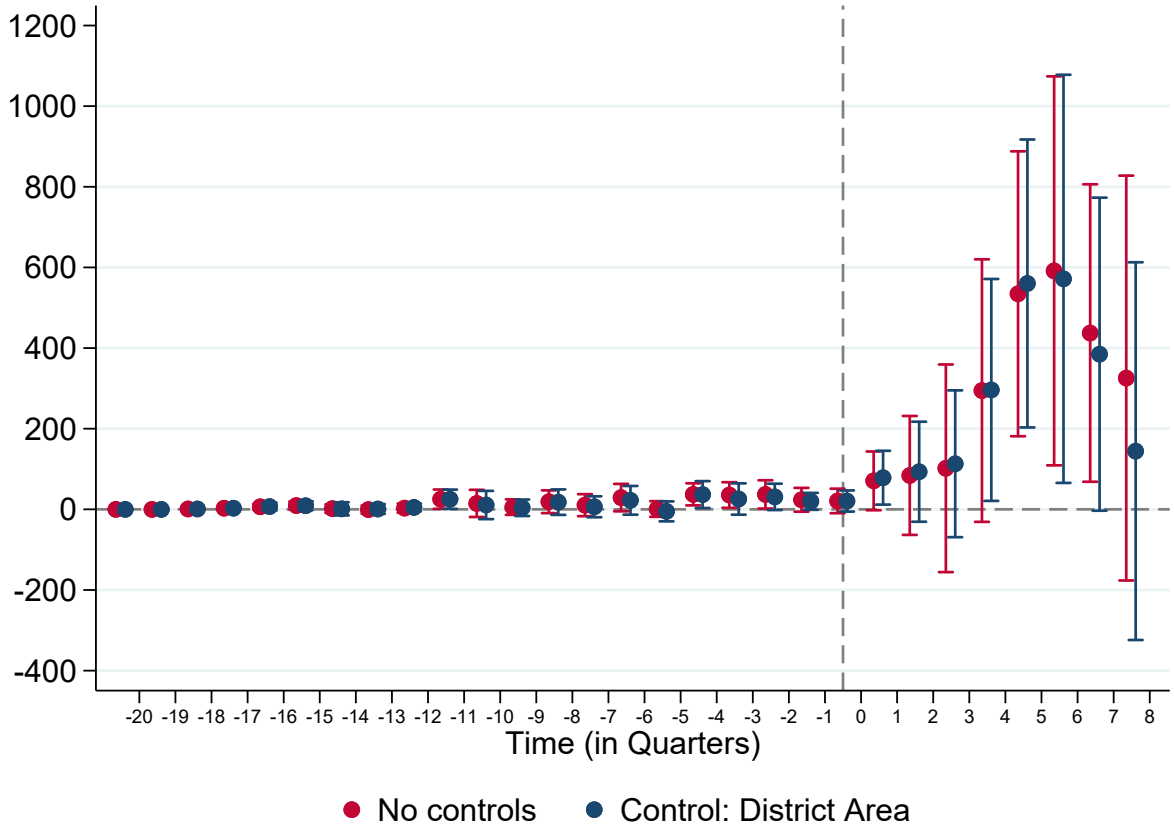
DiD estimates for our baseline measure – EVI implied agricultural yield: This figure plots the event-study dynamic coefficients estimated using the methodology outlined in Callaway and Sant'Anna (2021). The outcome variable is EVI implied agricultural yield constructed by subtracting the average value of EVI during the initial weeks of kharif season from the maximum value of EVI during the kharif season. The dependent variable is standardized to mean zero and standard deviation of one. The sample consists of all unique hexagons that saw the introduction of 4G BTS between 2013 and 2021. All measures are winsorized at 1% level. Standard errors for all estimators are clustered at the district level. The bars represent 95 percent confidence intervals.

Figure 4: Robustness using alternative DID estimators



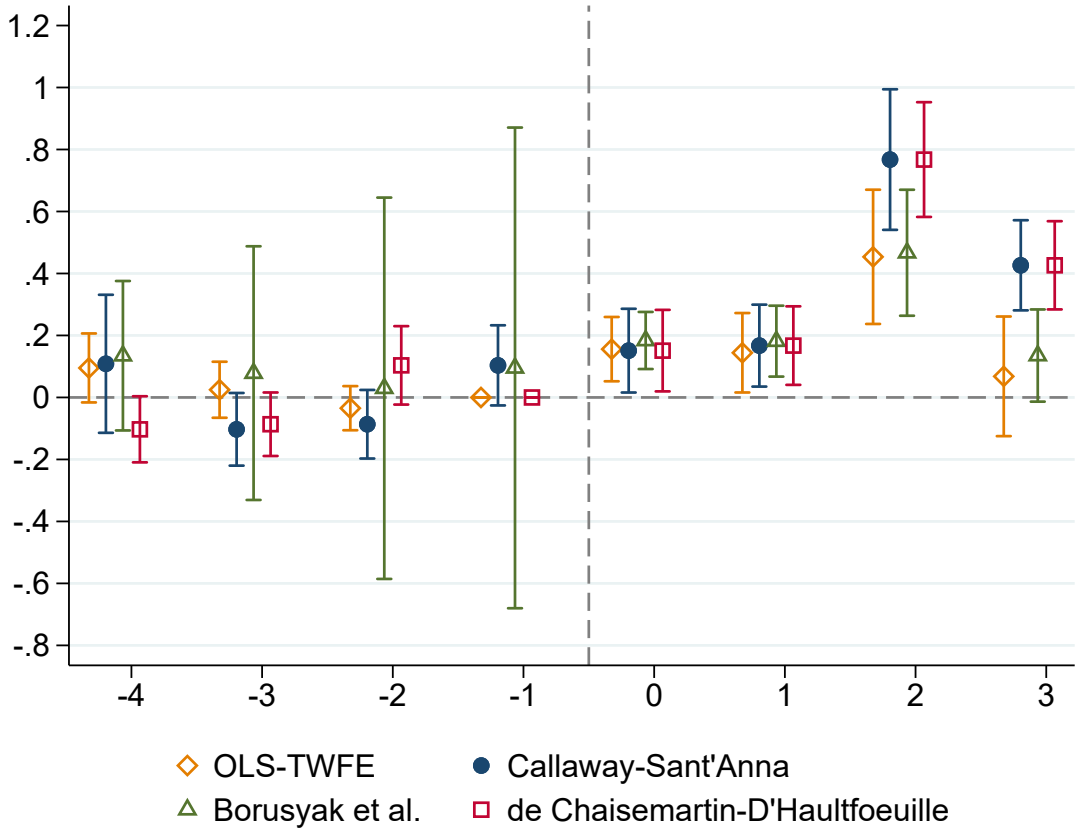
Robustness with multiple DiD estimators for our baseline measure – EVI implied agricultural yield This figure plots the event-study dynamic coefficients using four different estimators: (1) dynamic TWFE estimator (in orange with hollow diamond markers), (2) Callaway and Sant'Anna (2021) estimator (in blue with solid circle markers), (3) Borusyak, Jaravel, and Spiess (2022) estimator (in green with hollow triangle markers), and (4) De Chaisemartin and d'Haultfoeuille (2020) estimator (in pink with hollow square markers). The outcome variable is EVI implied agricultural yield constructed by subtracting the average value of EVI during the initial weeks of Kharif season from the maximum value of EVI during the Kharif season. The outcome variable is standardized to mean zero and standard deviation of one. The sample consists of all unique hexagons that saw the introduction of 4G BTS between 2013 and 2021. All measures are winsorized at 1% level. Standard errors for all estimators are clustered at the district level. The bars represent 95 percent confidence intervals. The table below the figure reports the aggregate treatment effect parameters with standard errors clustered at the district level shown in parenthesis. Rows "Group-specific" and "Event study" list the partially aggregated group-specific and event-time specific average treatment effects where g indexes the year of treatment of the group and e indexes event-time. The last column of each row summarizes the aggregate average treatment effect on the treated (ATT) by taking a weighted average of partially aggregated group-specific and event-time specific average treatment effects following Callaway and Sant'Anna (2021). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Figure 5: Effect of RoW Adoption on Installation of 4G Telecom Towers



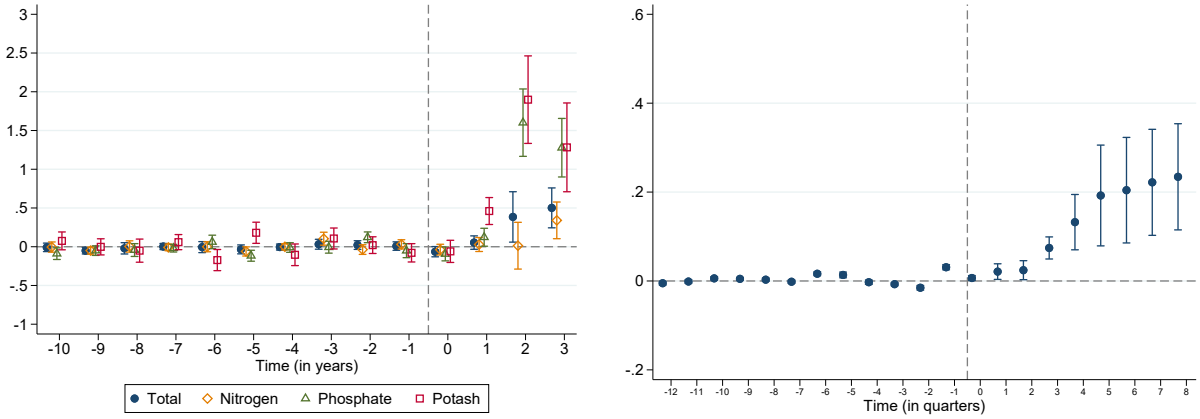
DiD estimates for the impact of staggered adoption of Rights of Way (RoW) rules across states on number of BTSs (first-stage results for RoW-policy based identification strategy): This figure plots the event-study dynamic coefficients estimated using the methodology outlined in Callaway and Sant'Anna (2021). The outcome variable is the quarterly stock of the number of 4G BTSs operational at the district-level. Consequently, event-time is measured in quarters. Treatment timing is constructed using the dates of introduction of RoW rules across different States/Union Territories as listed in D.1. The sample consists of the districts that fall within States/Union Territories which passed RoW rules from 2015 to 2019. All measures are winsorized at 1% level. Standard errors are clustered at the state level. The bars represent 95 percent confidence intervals.

Figure 6: Effect of RoW Adoption on Agricultural Production



DiD estimates for the impact of (staggered) passage of Rights of Way (RoW) rules across States on EVI implied agricultural yields (second-stage results for RoW-policy based identification strategy): This figure plots the event-study coefficients estimated using four different estimators: (1) dynamic TWFE estimator (in orange with hollow diamond markers), (2) Callaway and Sant'Anna (2021) estimator (in blue with solid circle markers), (3) Borusyak, Jaravel, and Spiess (2022) estimator (in green with hollow triangle markers), and (4) De Chaisemartin and d'Haultfoeuille (2020) estimator (in pink with hollow square markers). The outcome variable is EVI implied agricultural yield constructed by subtracting the average value of EVI during the initial weeks of Kharif season from the maximum value of EVI during the Kharif season. The outcome variable is standardized to mean zero and standard deviation of one. Treatment timing is constructed using the dates of introduction of RoW rules across different States/Union Territories as listed in D.1. The sample consists of the subset of hexagons that fall within States/Union Territories which passed RoW rules from 2015 to 2019, and saw the introduction of 4G BTS between 2013 and 2021. All measures are winsorized at 1% level. Standard errors for all estimators are clustered at the district level. The bars represent 95 percent confidence intervals.

Figure 7: Rights of Way Adoption & Agricultural Input Usage

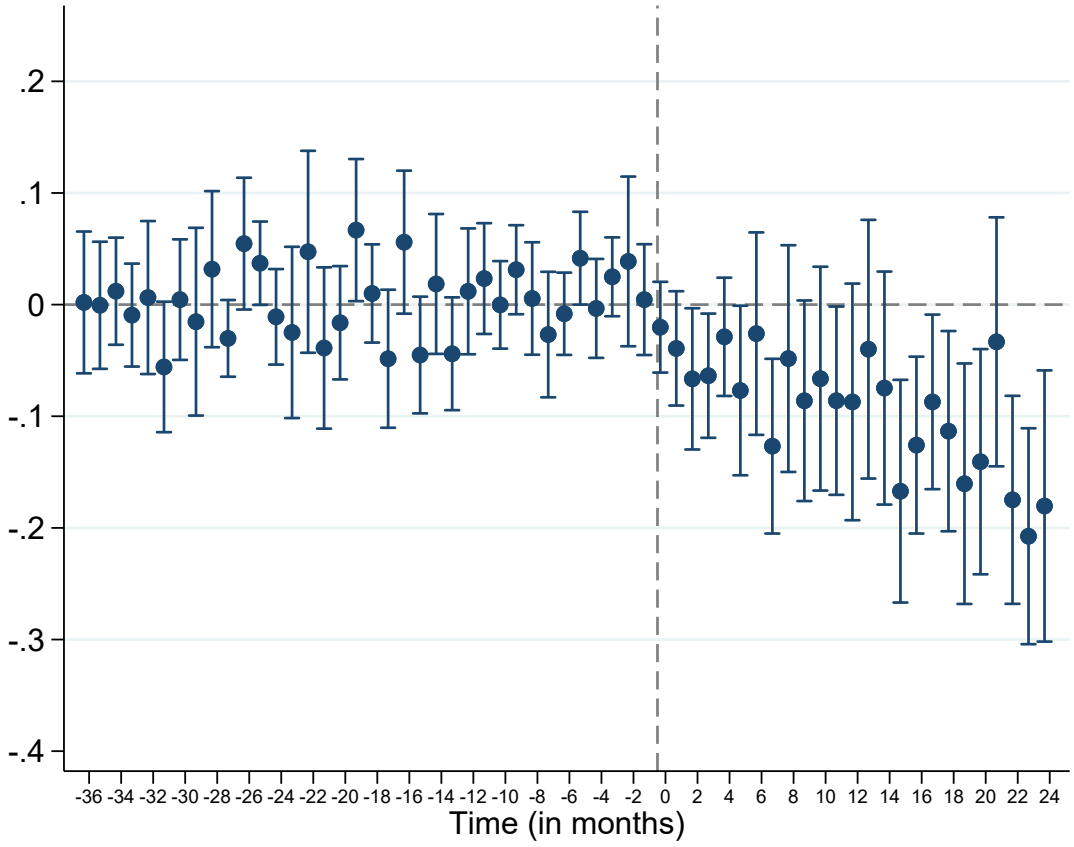


(a) Fertilizer Consumption

(b) Agricultural Bank Credit

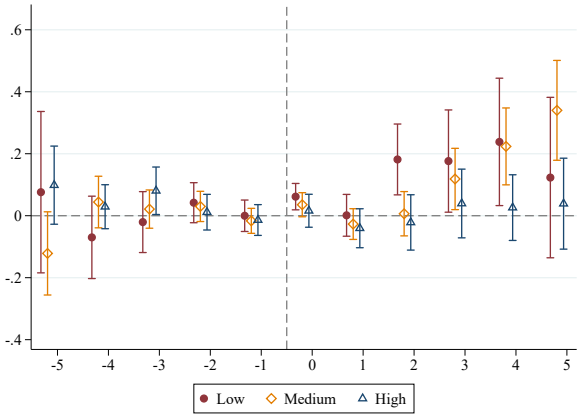
DiD estimates for impact of (staggered) passage of Rights of Way (R0W) rules across states on the usage of agricultural inputs: Figure 7a plots the event-study dynamic coefficients for the effect of RoW on fertilizer consumption estimated using the methodology outlined in Callaway and Sant'Anna (2021). The outcome variables are measured as the natural logarithm of the amount of consumption of total fertilizers (NPK, in blue with solid circle markers), nitrogen (N, in orange with hollow diamond markers), phosphate (P, in green with hollow triangle markers), and potash (K, in pink with hollow square markers) per unit of gross sown area (GSA). All variables are measured at the district-year level from 1998 until 2020. Standard errors for all estimators are clustered at the district level. Figure 7b plots the event-study dynamic coefficients for the effect of RoW on agricultural bank credit estimated using the methodology outlined in Callaway and Sant'Anna (2021). Event-time (X-axis) is measured in quarters. The outcome variable is the natural logarithm of quarterly credit disbursed at the branch level. The sample consists of all the zipcodes that saw the introduction of a 4G BTS between 2014 to 2022. Standard errors for all estimators are clustered at the zipcode level. For both panels, treatment timing is constructed using the dates of introduction of RoW rules across different States/Union Territories as listed in D.1. All measures are winsorized at 1% level. Standard errors are clustered at the zipcode level. The bars represent 95 percent confidence intervals. Standard errors for all estimators are clustered at the district level.

Figure 8: Rights of Way Adoption & Price Dispersion of Agricultural Commodities

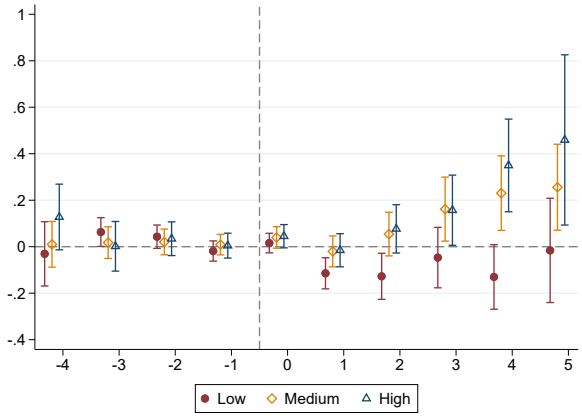


DiD estimates for impact of (staggered) passage of Rights of Way (R0W) rules across states on price dispersion: This figure plots the event-study dynamic coefficients estimated using the methodology outlined in Callaway and Sant'Anna (2021). The outcome variable is the monthly average of the natural logarithm of the ratio of the maximum and minimum price of the same crop across markets within the same state on the same day. We use daily data on nine key crops – Bajra, Maize, Jowar, Wheat, Paddy, Soyabean, Potato, Tomato, and Onions – across all markets in India from 2011 to 2021. Treatment timing is constructed using the dates of introduction of RoW rules across different States/Union Territories as listed in D.1. The sample consists of the subset of markets that fall within States/Union Territories which passed RoW rules from 2015 to 2019. All measures are winsorized at 1% level. Standard errors for all estimators are clustered at the market level. The bars represent 95 percent confidence intervals.

Figure 9: Mechanism: Heterogeneous Treatment Effect by Accuracy & Value of Information



(a) Accuracy of Information: Distance to AWS

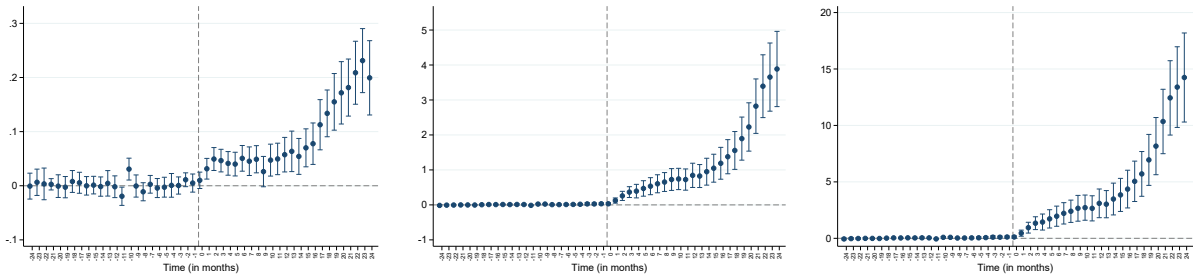


(b) Value of Information: Sensitivity to Drought

	Low	Medium	High	Difference (Low-High)
Distance to Automatic Weather Station (AWS) (Figure 9a)	0.1209** (0.0513)	0.0786*** (0.0286)	0.0043 (0.0328)	0.1166*** [8.62]
Sensitivity to Drought (Figure 9b)	-0.0705* (0.0395)	0.0962*** (0.0373)	0.1358*** (0.0460)	-0.2062*** [19.47]

DID estimates for heterogeneous treatment effect by accuracy or reliability of information (Figure 9a) and value of information (Figure 9b): This figure plots the event-study dynamic coefficients estimated using the methodology outlined in Callaway and Sant'Anna (2021). The outcome variable is EVI implied agricultural yield, constructed by subtracting the average value of EVI during the initial weeks of the Kharif season from the maximum value of EVI during the Kharif season. The outcome variable is standardized to mean zero and a standard deviation of one. The sample consists of all unique hexagons that saw the introduction of 4G BTS between 2013 and 2021. Figure 9a plots the DID event study plots for three sub-samples with varying levels of accuracy or reliability of information. The accuracy of weather information is measured using the distance of the centroid of the hexagon to the nearest weather station. **Hexagons located closer to the AWS have lower noise in information or a greater accuracy or reliability of information.** We split the hexagons into three sub-samples based on the values of the distance to the nearest automatic weather station (AWS) – (1) low, is a sub-sample of hexagons with distance lower than or equal to the 25th percentile value, (2) medium, is a sub-sample of hexagons with distance greater than the 25th percentile value and less than the 75th percentile value, and (3) high, is a sub-sample of hexagons with distance greater than or equal to the 75th percentile value. Figure 9b plots the DID event study plots for three sub-samples with varying levels of value of information. The value of weather information is measured using the pre-period sensitivity of agricultural yield in each hexagon to droughts. **Hexagons with higher sensitivity to droughts have a greater value of information.** We split the hexagons into three sub-samples based on the values of the sensitivity to droughts – (1) low, is a sub-sample of hexagons with distance lower than or equal to the 25th percentile value, (2) medium, is a sub-sample of hexagons with distance greater than the 25th percentile value and less than the 75th percentile value, and (3) high, is a sub-sample of hexagons with distance greater than or equal to the 75th percentile value. All measures are winsorized at 1% level. The bars represent 95 percent confidence intervals. The table below the figure plots the Callaway and Sant'Anna (2021) ATT point estimates for each subsample: low, medium, and high. The numbers in the parentheses present the standard errors. The first row labelled *Distance to AWS* corresponds to the event-study plot presented in Figure 9a. The second row labelled *Sensitivity to Drought* corresponds to the event-study plot presented in Figure 9b. The last column presents the difference between the ATT estimates for the low and high sub-samples, along with the associated χ^2 statistic for the difference in square brackets. Standard errors for all estimators are clustered at the district level. Statistical significance is indicated by *, **, and ***, corresponding to the 10%, 5%, and 1% levels, respectively.

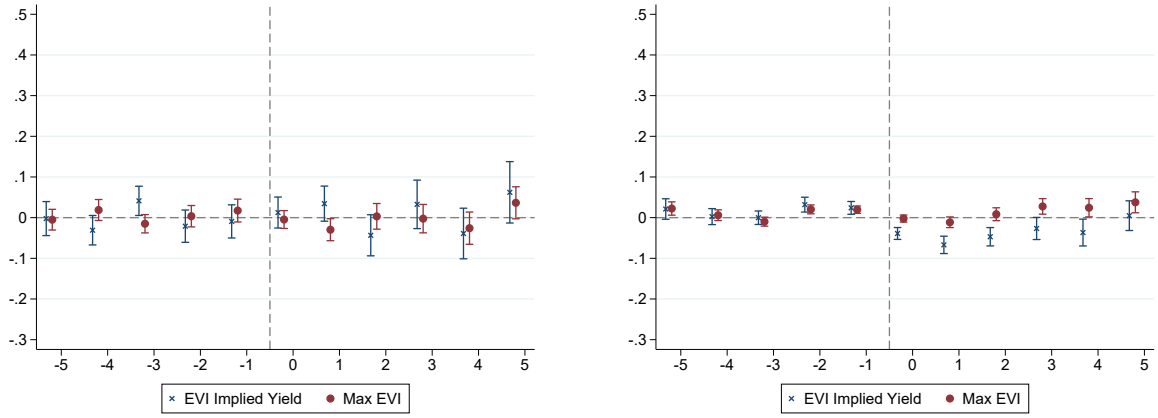
Figure 10: Effect of 4G Introduction on Krishify App Installations



(a) Any installation (b) # Installations (c) $\frac{\#Installations}{\text{Pre-Period Average}}$

DID estimates for impact of 4G introduction on monthly installations of Krishify mobile application at the hexagon level: This figure plots the event-study dynamic coefficients estimated using the methodology outlined in Callaway and Sant'Anna (2021). Figure 10a uses a binary variable which takes the value of one if total monthly app installations at the hexagon level are greater than zero, representing any installation of the app in the month. Figure 10b uses the number of monthly installations at the hexagon level as the dependent variable. Figure 10c uses the number of monthly installations scaled by pre-period average number of monthly installations at the hexagon level as the dependent variable. Geolocations and timestamps of Krishify app installations by users are superimposed on the hexagons in our sample to calculate monthly downloads at the hexagon level. Treatment timing in both the figures is the month of introduction of the first 4G BTS inside the boundary of a hexagon. Standard errors are clustered at the district level. The bars represent 95 percent confidence intervals.

Figure 11: Effect of 3G and 2G Introduction on Agricultural Yield (16-day EVI Composites)

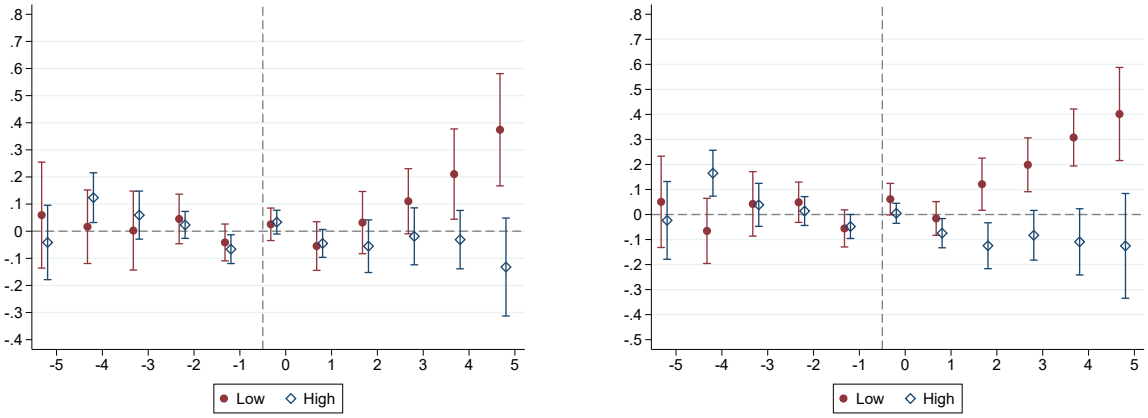


(a) 3G Introduction

(b) 2G Introduction

DiD estimates for our baseline measures – EVI implied agricultural yield and maximum EVI using 16-day EVI Composites: This figure plots the event-study dynamic coefficients estimated using the methodology outlined in Callaway and Sant'Anna (2021). Panels 11a and 11b plot the treatment effects for the sample of hexagons that saw the introduction of 3G BTS and 2G BTS, respectively. Furthermore, in Panel 11a we limit our sample to hexagons that neither experienced the introduction of a 4G BTS prior to the deployment of their first 3G BTS nor saw a 4G BTS introduced within two years of the first 3G BTS. Similarly, in Panel 11b we limit our sample to hexagons that neither experienced the introduction of a 3G or 4G BTS prior to the deployment of their first 2G BTS nor saw a 3G or 4G BTS introduced within two years of the first 2G BTS. These restrictions help us to isolate the causal impact of 3G and 2G deployment on agricultural productivity by minimizing confounding effects from early or overlapping 4G or 3G infrastructure. For both the panels, the outcome variable is EVI implied agricultural yield (in blue) constructed by subtracting the average value of EVI during the initial weeks of kharif season from the maximum value of EVI during the kharif season. The second outcome variable is Max EVI (in maroon) constructed by using the maximum value of EVI during the kharif season. The dependent variables are standardized to mean zero and standard deviation of one. Additionally, because the standard 8-day EVI data begins only in 2013, we construct the measures used in this analysis using 16-day EVI to facilitate a longer historical analysis. This extension was necessary to assess the impact of 3G and 2G deployment, requiring data coverage extending back to 2000. All measures are winsorized at 1% level. Standard errors for all estimators are clustered at the district level. The bars represent 95 percent confidence intervals.

Figure 12: Heterogeneous Treatment Effect by Trust



(a) Trust in Local Village Panchayat

(b) Trust in State Government

	Low	High	Difference (Low-High)
Trust in Local Village Panchayat (Figure 12a)	0.0819** (0.0393)	-0.0319 (0.0349)	0.1138*** [13.92]
Trust in State Government (Figure 12b)	0.1431*** (0.0327)	-0.0762** (0.0335)	0.2193*** [53.37]

DID estimates for heterogeneous treatment effect by trust in village panchayat: This figure plots the event-study dynamic coefficients estimated using the methodology outlined in Callaway and Sant'Anna (2021). The outcome variable is EVI implied agricultural yield constructed by subtracting the average value of EVI during the initial weeks of Kharif season from the maximum value of EVI during the Kharif season. The outcome variable is standardized to mean zero and standard deviation of one. The sample consists of all unique hexagons that saw the introduction of 4G BTS between 2013 and 2021. In Figure 12a, we utilize data from India Human Development Survey (IHDS-II), conducted in 2011, which includes a question that queries respondents about the level of trust they have in their local village panchayats. The exact question is as follows: "How much trust do you place in the ability of your village panchayat to implement public projects?" The response options consist of three choices: a) A great deal of confidence, b) Only some confidence, and c) Hardly any confidence at all. To quantify these responses, we assign a numerical value of 1 to option (a), 0.33 to option (b), and 0 to option (c). We take the average of this measure across all households within a district based on their weights. This allows us to compute a continuous measure of confidence in village panchayats at the district level. We use the response to this question to construct a district-level proxy for confidence in village panchayats and divide the districts into two subsets based on the median response value. Districts with confidence levels above the median are categorized as "high confidence" districts, while those below the median are classified as "low confidence" districts. We then link this measure to the hexagonal grids situated within the districts. In Figure 12b, we utilize data from India Human Development Survey (IHDS-II), conducted in 2011, which includes a question that queries respondents about the level of trust they have in their local village panchayats. The exact question is as follows: "How much trust do you place in the ability of the state government to take care of people?" The response options consist of three choices: a) A great deal of confidence, b) Only some confidence, and c) Hardly any confidence at all. To quantify these responses, we assign a numerical value of 1 to option (a), 0.33 to option (b), and 0 to option (c). We take the average of this measure across all households within a district based on their weights. This allows us to compute a continuous measure of confidence in state government at the district level. We use the response to this question to construct a district-level proxy for confidence in state government and divide the districts into two subsets based on the median response value. Districts with confidence levels above the median are categorized as "high confidence" districts, while those below the median are classified as "low confidence" districts. We then link this measure to the hexagonal grids situated within the districts. All measures are winsorized at 1% level. Standard errors for all estimators are clustered at the district level. The bars represent 95 percent confidence intervals. The table below the figure plots the Callaway and Sant'Anna (2021) ATT point estimates for each subsample: low and high. The numbers in the parentheses present the standard errors. The first row labelled *Trust in Local Village Panchayat* corresponds to the event-study plot presented in Figure 12a. The second row labelled *Trust in State Government* corresponds to the event-study plot presented in Figure 12b. The last column presents the difference between the ATT estimates for the low and high sub-samples, along with the associated χ^2 statistic for the difference in square brackets. Standard errors for all estimators are clustered at the district level. Statistical significance is indicated by *, **, and ***, corresponding to the 10%, 5%, and 1% levels, respectively.

Internet Appendix for:

"Bridging the Information Gap: Sowing the Seeds of Productivity with High-Speed 4G Internet"

Contents

A Data **A2**
- A.1 Hexagonal Tesselation of the Indian Map A2
- A.2 Properties of EVI A3
- A.3 EVI Implied Yields & Agricultural Production A4
- A.4 Quantifying the Effect of 4G Introduction A4
- A.5 Construction of Local Weather Shocks A5
- A.6 Description of Krishify Data A7

B Discussion on the Estimator **A10**
- B.1 TWFE Dynamic Specification A10
- B.2 Treatment effect homogeneity A10

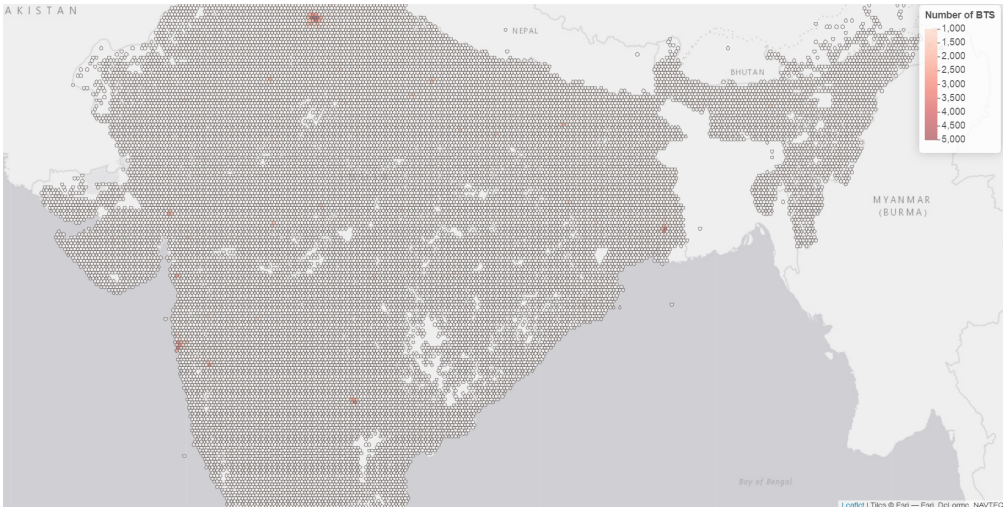
C Robustness **A13**
- C.1 Uniform Treatment Effect Assumption A14
- C.2 Heterogeneous Treatment Effect by Agricultural Production Potential A15

D Identification Using Rights of Way **A16**

Appendix A Data

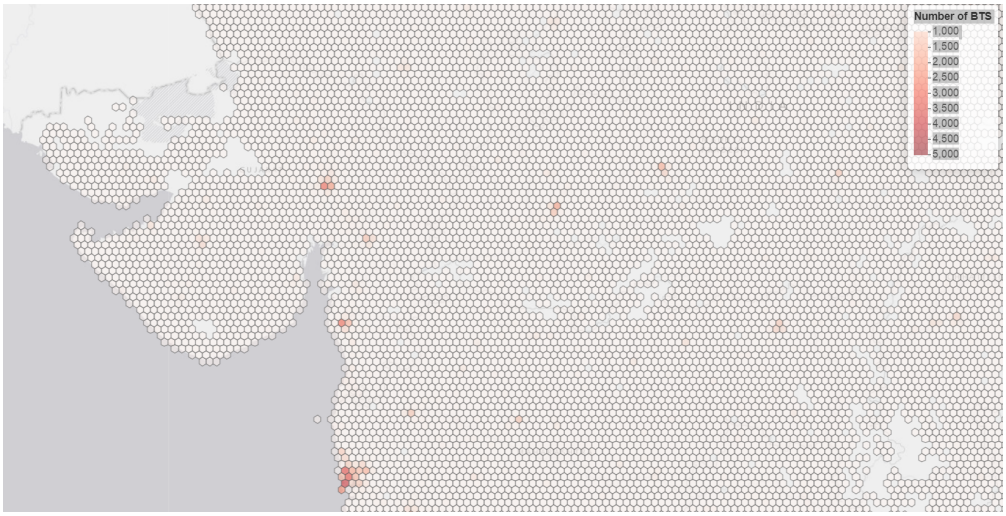
A.1 Hexagonal Tesselation of the Indian Map

Figure A.1: Hexagonal Tesselation of the Indian Map



This figure shows the hexagonal units of equal area (≈ 95 sqkm.) drawn on the Indian map with at least one BTS.

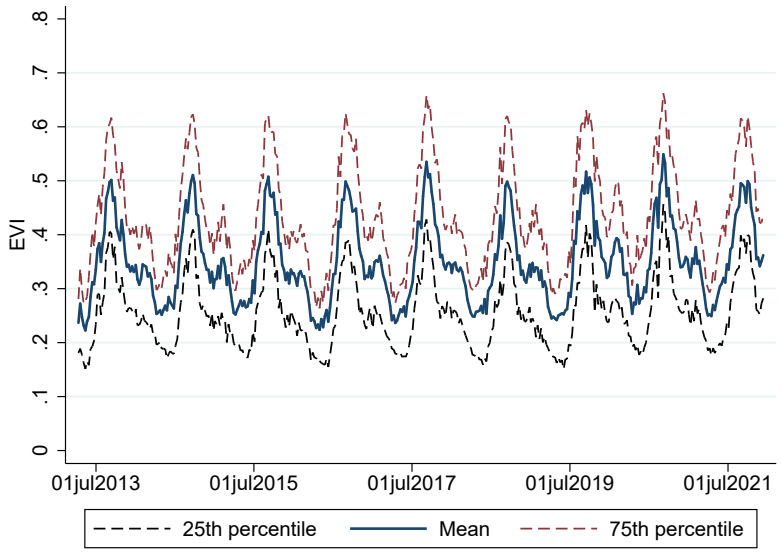
Figure A.2: Hexagonal Tesselation of the Indian Map (zoomed-in)



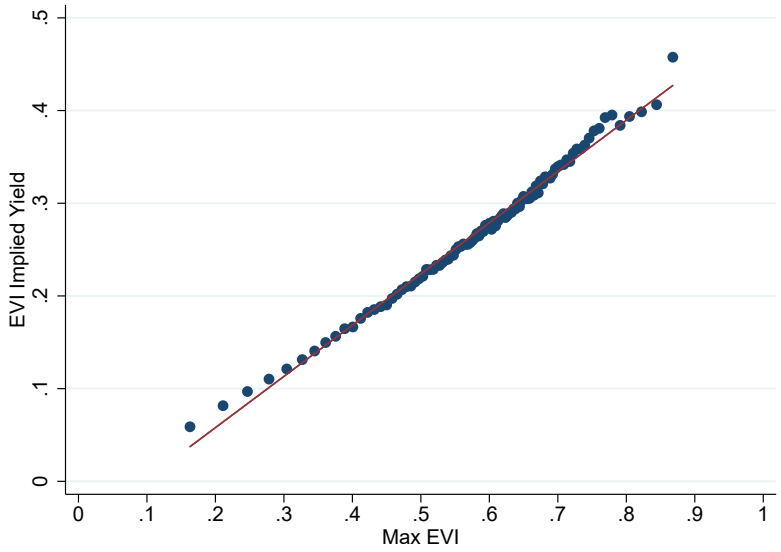
This figure shows the zoomed-in image of hexagonal units presented in Appendix Figure A.1.

A.2 Properties of EVI

Figure A.3: Properties of Enhanced Vegetation Index (EVI)



(a) Summary Statistics of EVI



(b) Relationship between EVi Implied Yield and Maximum EVI

Appendix Figure A.3a plots the evolution of average, the 25th percentile, and the 75th percentile EVI value for *all* hexagons in our sample. Appendix Figure A.3b presents the binscatter plot of the EVI implied yield measure and the maximum EVI measure. EVI implied yield is constructed by subtracting the average value of EVI during the initial weeks of kharif season from the maximum value of EVI during the kharif season. All measures are winsorized at 1% level.

A.3 EVI Implied Yields & Agricultural Production

This section discusses the relationship between EVI implied yield and measures of agricultural production. Administrative data on actual crops yields and production values is not available beyond the district level. Therefore, we construct hexagon-level measures of actual crop yields and production values from Advancing Research on Nutrition and Agriculture's (ARENA) Demographic and Health Surveys (DHS)-GIS Database. We map survey clusters in the DHS data to hexagons based on the minimum distance between the centroids of hexagons and geo-coordinates of survey clusters. Specifically, we use three measures of agricultural production available in the ARENA-DHS database. These measures include – (1) Yield of cereal crops, (2) the total value of produce of all cereal crops in USD, and (3) the total value of produce of all crops in USD. Cereal crops include a variety of different plants including wheat, rice, oats, barley, maize, rye, millet, corn, and sorghum. These measures are calculated as of the year 2014. We take a natural logarithm of these three measures.

We start by regressing these three measures against EVI-implied yields at the hexagonal level for the year 2014. Table A.1 and Figure A.4 report the regression results. Column 1 uses the natural logarithm of yield for cereal crops as the dependent variable. Yield is measured as kg of cereal crops per hectare of land. Column 2 uses the natural logarithm of total USD value of cereal crops as the dependent variable. Column 3 uses the natural logarithm of the total USD value of all crops as the dependent variable. The estimate of interest is the coefficient associated with EVI implied yield. Consistent with the findings of Asher and Novosad (2020), we observe the coefficient of interest is positive and statistically significant across all measures of agricultural production.

The size of the estimate indicates that a 0.5 unit increase in EVI implied yield is associated with a 15.1% increase in cereal yield, a 61.2% increase in USD value of cereal production, and a 45.9% increase in USD value of all crop production.

A.4 Quantifying the Effect of 4G Introduction

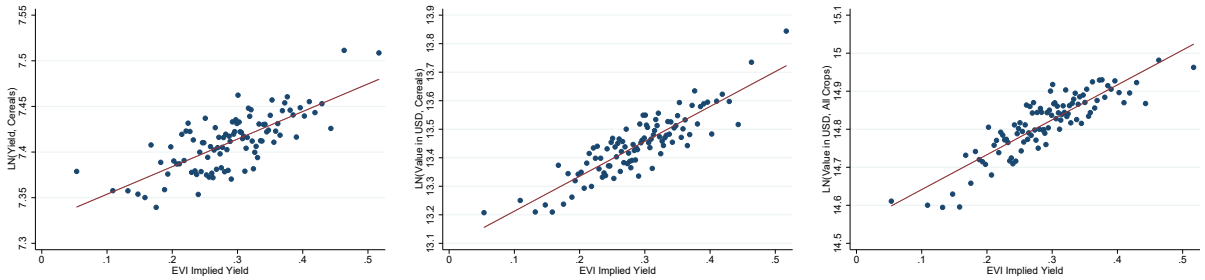
Next, we use these estimates to quantify the average impact of 4G introduction on actual yields and production values. Section 4 documents a 0.035 unit increase in EVI implied yield six years after the introduction of the 4G network. We use this change in EVI implied yield to impute the effect on real yields and production value. Specifically, the estimates in Table A.1 imply that the introduction of high-speed 4G internet is associated with a – (1) 1.06% increase in agricultural yields of cereal crops, (2) a 4.28% in USD production value of cereal crops, and (3) a 3.21% increase in the USD production value of all crops.

We further break down the hexagonal-level, dollar increase in production values from all crops into an estimate of impact on household-level income. Using the 2011 census data on village-level count of households, we aggregate total number of households at the hexagon-level. We do this by mapping village shapefiles to our hexagons. On average, 698.86 households reside in each hexagon. We multiply the average annual income of agricultural households in 2014 with the average number of households in each hexagon. The average monthly income for an agricultural household comes from the 2014 Situation Assessment Survey and is equal to ₹6,426. The average annual income for an agricultural household is equal to ₹77,112 (₹6,426 X 12). The average annual income for all households in a hexagon is ₹53,890,793.06 (₹77,112 X 698.86) or $ 883,464.31.

Comparing the average increase presented in columns 3 of Table A.1 with the average annual

income for all households in the hexagon, we find that the introduction of 4G internet led to a 14.5% increase in household income.

Figure A.4: EVI Implied Yield & Agricultural Production



(a) Yield for Cereals (kg per hectare)

(b) Value of Cereals Produced in USD

(c) Value of All Crops Produced in USD

This figure plots the relationship between EVI implied yield and three measures of agricultural production. Panel A.4a plots the relationship with the natural logarithm of yield on cereal crops measured in kg per hectare. Panel A.4b plots the relationship with the natural logarithm of the total USD value of cereal crops. Cereal crops include a variety of different plants including wheat, rice, oats, barley, maize, rye, millet, corn, and sorghum. Panel A.4c plots the relationship with the natural logarithm of total USD value of all crops. All variables are winsorized at 1%. All plots include district fixed effects.

A.5 Construction of Local Weather Shocks

Here, we provide a brief description of the Standardized Precipitation Evapotranspiration Index (SPEI) and the construction of drought shocks from this index.

The foundation of the SPEI is a measure of monthly water balance for a 0.5×0.5-degree area, with approximately four thousand squares in India. This water balance is calculated as the difference between precipitation and potential evapotranspiration. The latter describes the loss of water from the soil both by evaporation from the soil surface and by transpiration from the leaves of the plants growing on it. Potential evapotranspiration is also a function of temperature.

This monthly water balance measure closely follows a log-logistic distribution. An individual distribution is fitted for each month and geographic area. The parameters of these distributions are estimated using historical data for that specific month in a specific area, starting in 1901. For instance, to fit the distribution of March in a 0.5 x 0.5-degree area in Delhi, the historical water balance measures starting March 1901 from that location are utilized.

The next step is to make the measure comparable across months and geographical areas. For this purpose, the water balance measures are standardized, utilizing characteristics of their individual distributions. This results in the SPEI, which has an average value of zero and a standard deviation of one. A SPEI value of zero indicates no change in water balance relative to observed historical values for that month in that given area. An SPEI value greater than zero indicates a water surplus and, in extreme cases, a flood. An SPEI lower than zero indicates a water deficit and, in extreme cases, a drought.

To integrate the SPEI with the other data variables measured at the hexagon-level, we need to translate the 0.5×0.5-degree rectangles from the SPEI data to the hexagons. We calculate a weighted average of SPEI for each hexagon, where the weights are the proportion of the area covered by each 0.5×0.5-degree rectangle within the hexagon. We use the hexagon-level SPEI to

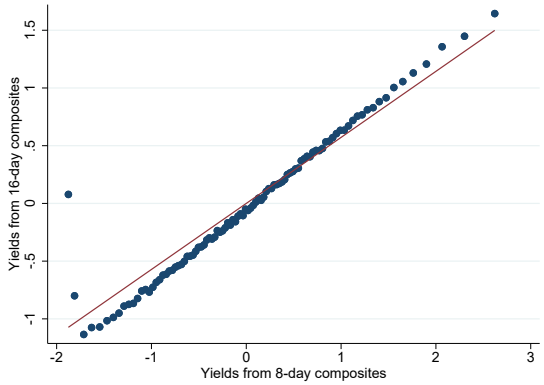
construct our local drought shock variable. Our local weather shock is a binary variable that takes a value of one if the SPEI observation (at the year-month-hexagon level) is below the 20th percentile of its historical distribution in that hexagon from January 1950 to December 2012. Our approach – including the choice of percentiles – follows a long-standing literature that investigates the impact of weather shocks in similar contexts since Jayachandran (2006).

Table A.1: EVI Implied Yield & Agricultural Production

	(1)	(2)	(3)
	LN(Yield, Cereals)	LN(Value, Cereals)	LN(Value, All Crops)
EVI Implied Yield	0.3021***	1.2232***	0.9172***
	(0.0715)	(0.1854)	(0.1357)
District FE	Yes	Yes	Yes
# Obs	28,118	28,118	28,118
R^2	0.7827	0.7168	0.7104
Coef	0.3021	1.2232	0.9172
ΔX	0.035	0.035	0.035
$\frac{\Delta Y}{Y}$	1.06%	4.28%	3.21%
Mean (Y)	2,018.54 kg/hectare	$ 1,551,215.00	$ 3,989,287.00
Increase (Y)	21.34 kg/hectare	$ 66,410.62	$ 128,064.09

This table reports the relationship between EVI implied yield and three measures of agricultural production. EVI implied agricultural yield is constructed by subtracting the average value of EVI during the initial weeks of kharif season from the maximum value of EVI during the kharif season. Column 1 uses the natural logarithm of yield for cereal crops as the dependent variable. Yield is measured as kg of cereal crops per hectare of land. Column 2 uses the natural logarithm of total USD value of cereal crops as the dependent variable. Column 3 uses the natural logarithm of the total USD value of all crops as the dependent variable. Cereal crops include a variety of different plants including wheat, rice, oats, barley, maize, rye, millet, corn, and sorghum. All variables are measured at the hexagon level for the year 2014. The sample consists of all unique hexagons that saw the introduction of 4G BTS between 2013 and 2021. All variables are winsorized at 1%. Figure A.4 reports the same regressions in graphical format. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Figure A.5: Correlation between EVI Implied Yields from 16-day and 8-day Composites



This figure plots a bin-scatter plot of standardized EVI implied yields from 16-day composites against 8-day composites.

A.6 Description of Krishify Data

Krishify is an Indian company aimed at connecting farmers on a social network where they can discuss agriculture-related issues. Link to their official website can be found here.

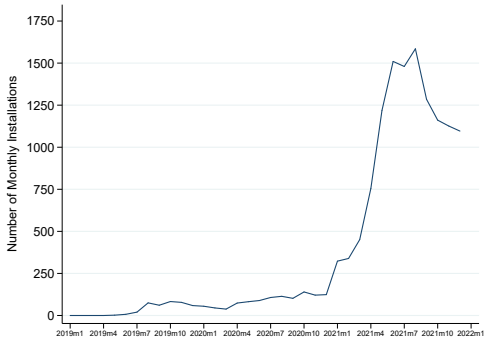
Krishify is a mobile application that can be downloaded through the Google play store with a "vision to democratize the flow of information and scale up economic opportunities for every farmer in this country". It provides a social network for farmers to connect with each other and a commerce platform for agri-businesses to connect with a community of more than 10 million farmers. They help businesses target farmers on their network mainly through engaging and informative videos. Krishify offers subscription plans to businesses to develop content for them as well as incentives to farmers and other content creators to upload videos by allowing them to monetize their content. They have more than 3.1 million daily video views, 140,000 daily active users, and 500,000 daily business interactions between farmers and businesses.

We obtain two datasets from Krishify. First, we obtain a proprietary, geolocated and time-stamped dataset of their app installations from 2019 until 2021. Second, we also obtain a 10% random sample of farmers in the Krishify database along with their detailed search, browsing, like, and comment history.

We superimpose the coordinates of app installations over our hexagonal grid to identify the number of app installations at the hexagonal level. We use this dataset to conduct a differences-in-differences analysis that examines the effect of 4G introduction on internet adoption. Therefore, we only keep hexagons where 4G was introduced after January 2019. Appendix Figure A.6 presents the evolution of monthly app installations in our sample. Appendix Figure A.7 presents the geographic distribution of total Krishify app installations by December 2021.

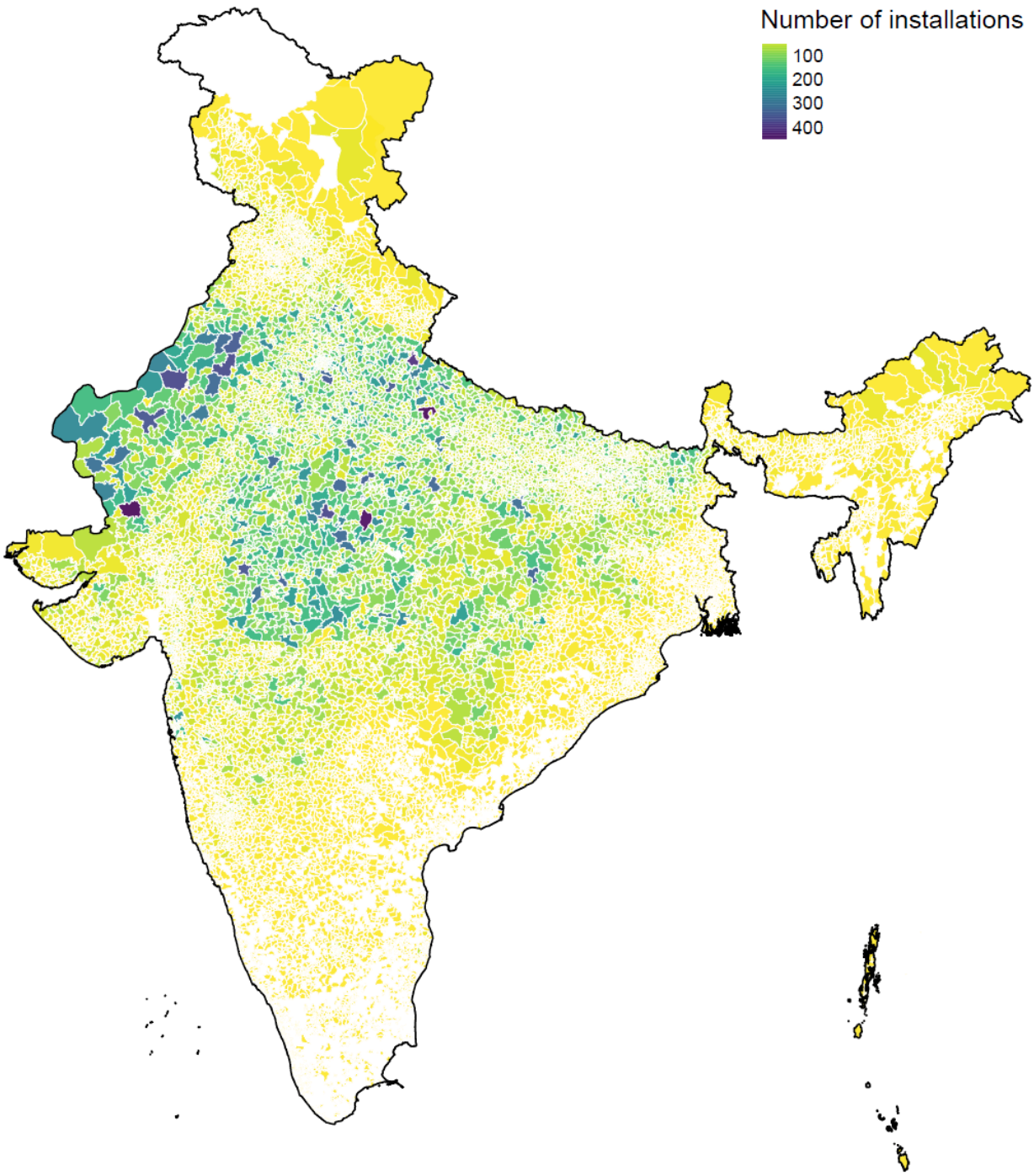
Second, we also obtain a 10% random sample of farmers in the Krishify database along with their detailed search, browsing, like, and comment history. This section discusses the aspects of Krishify database. the random sample also provides other information such as the buckets of land-holdings of the farmers as well as their risk-taking ability. Appendix Figure A.8 presents the distribution of farmers in our sample by their land-holding buckets and their risk-taking ability. Appendix Figure **??** presents the kernel density of the engagement score that measures the total engagement of the farmer on the platform.

Figure A.6: Monthly Installations of Krishify App



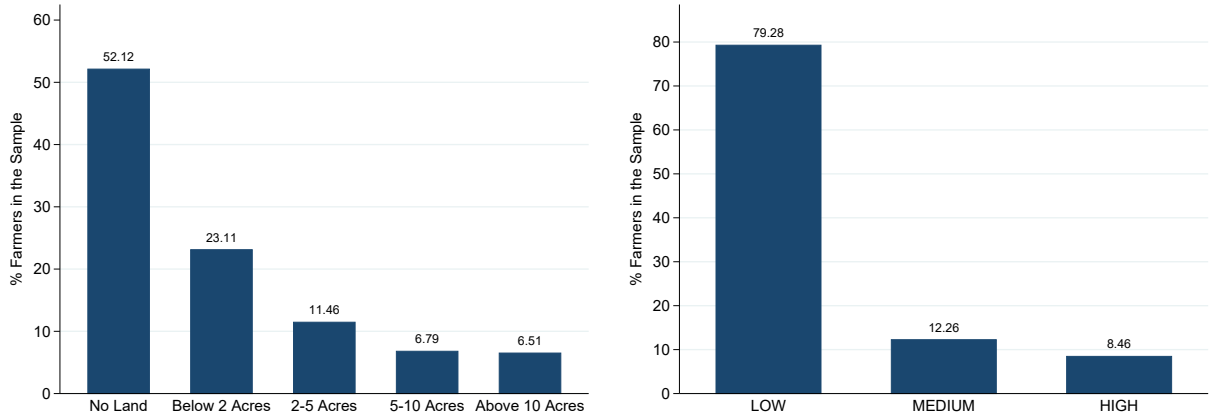
This figure presents the total number of monthly installations of the Krishify app for our sample.

Figure A.7: Geographic distribution of Krishify app installations



This figure plots a heatmap of the number of total Krishify app installations at the pincode level.

Figure A.8: Description of Farmers in the Krishify Database

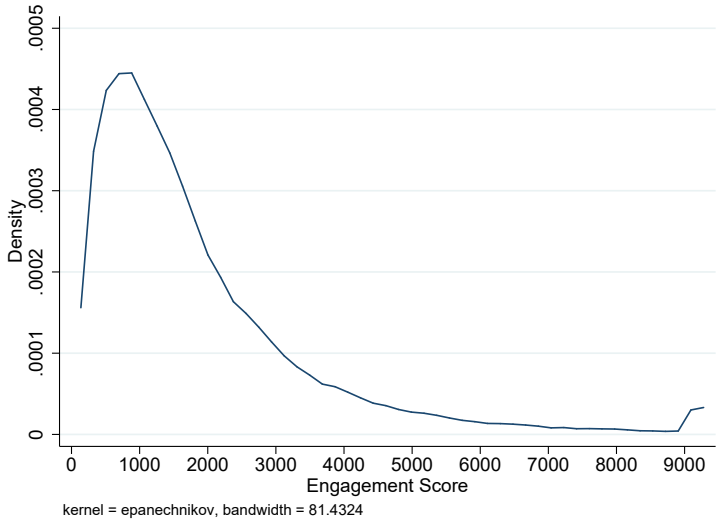


(a) Land Holding

(b) Risk Taking Ability

This figure presents the descriptives of the characteristics of farmers in the Krishify database. Figure A.8a presents the distribution of farmers by land-ownership buckets. Figure A.8b presents the distribution of farmers by their risk-taking ability. The sample comprises of a 10% random sample of farmers in the Krishify database.

Figure A.9: Kernel Density of Engagement Score



This figure presents the kernel density of farmer-level engagement score on the Krishify app. The engagement score is a proprietary score created by Krishify that measures the total level of farmer engagement on the platform. The sample comprises of a 10% random sample of farmers in the Krishify database.

Appendix B Discussion on the Estimator

We start by estimating treatment effects using the conventional two-way fixed effects (TWFE) methodology. Next, we critically examine the limitations of these difference-in-differences (DID) estimates. In particular, TWFE-based DID relies on the assumption of treatment effect homogeneity across units and over time in staggered adoption contexts. We argue that this assumption is unlikely to hold in our setting, which may lead to biased estimates.

B.1 TWFE Dynamic Specification

We begin by estimating a dynamic specification B.1 in relative time, which allows for non-parametric changes in treatment effects over time. Borusyak, Jaravel, and Spiess (2022) argue that we need to exclude at least two relative period indicators when there are no never treated units with a panel balanced in calendar time.[19] Therefore, we exclude two relative time periods corresponding to $t \in \{-8, -1\}$ to address potential multi-collinearity concerns.

$$Y_{i,t}^{EVI} = \alpha_i + \lambda_t + \sum_{\substack{y=-8, \\ y\neq -1,-8}}^{y=8} \beta_y \mathbb{1}(t - t_i^* = y) + \varepsilon_{i,t} \tag{B.1}$$

where $Y_{i,t}$ denotes the EVI-implied agricultural yield for the Kharif season in year (t) and hexagon (i). Indicator variables $\mathbb{1}(t - t_i^* = y)$ measure time relative to the year hexagon i first got treated (t_i^*). α_i and λ_t denote hexagon and time fixed effects, respectively. Standard errors are clustered at the district level.

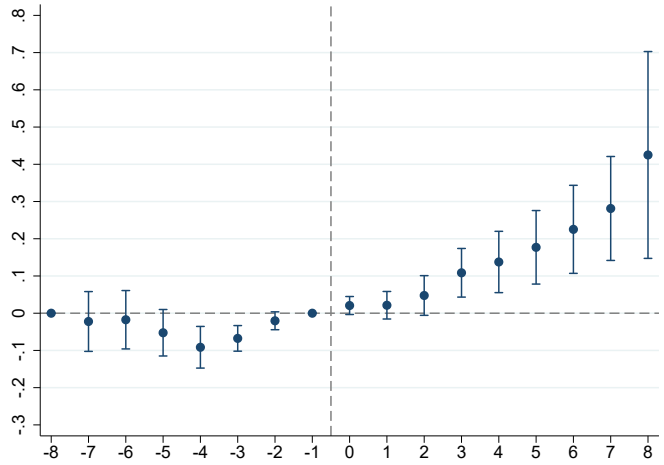
Appendix Figure B.1 presents the sequence of dynamic coefficients $\{\beta_y\}$ estimated using equation B.1. These estimates can be interpreted as causal under the assumptions of no-anticipation of the treatment, treatment effect homogeneity, and parallel trends. Negative and statistically significant coefficients for relative time periods $t = -3$ and $t = -4$ suggest meaningful pre-trends and thus, a possible violation of parallel trends assumption. However, Sun and Abraham (2021) show that testing for parallel trends using pre-treatment leads from a dynamic TWFE specification can be problematic as pre-treatment leads can be contaminated with treatment effects from other relative time periods, including post-treatment lagged effects. Therefore, a rejection of parallel trends based on significant pre-treatment leads necessarily assumes treatment effect homogeneity.

B.2 Treatment effect homogeneity

We begin by evaluating the plausibility of the assumption of treatment effect homogeneity in our context. Treatment effect homogeneity imposes the condition of homogeneous treatment effects across cohorts, i.e., farmers gaining access to 4G services in 2014 experience similar productivity gains as farmers that gain 4G access in 2021. Intuitively, the assumption of treatment effect homogeneity may be violated due to a variety of reasons, such as changes in forecasting skills over time, mergers and acquisitions among telecom service providers, and variation across cohorts due to calendar time-varying macroeconomic conditions could govern the effects on outcomes across cohorts.

[19] There are two sources of multi-collinearity. First, relative period indicators sum to one for every unit. Second, the presence of a linear relationship between two-way fixed effects and the relative period indicators.

Figure B.1: Dynamic TWFE specification for baseline EVI measure



This figure plots the event-study coefficients estimated using the dynamic TWFE specification in Equation B.1. The outcome variable is EVI implied agricultural yield constructed by subtracting the average value of EVI during the initial weeks of Kharif season from the maximum value of EVI during the Kharif season. The dependent variable is standardized. The sample consists of all unique hexagons that saw the introduction of 4G BTS between 2013 and 2021. Excluded time periods are $t = -8, -1$. All measures are winsorized at 1% level. Standard errors for all estimators are clustered at the district level. The bars represent 95 percent confidence intervals.

We formally diagnose the source and extent of contamination of the pre-treatment leads in our dynamic TWFE specification using the methodology outlined in Sun and Abraham (2021). The estimate for β_{-3} can be decomposed as follows:

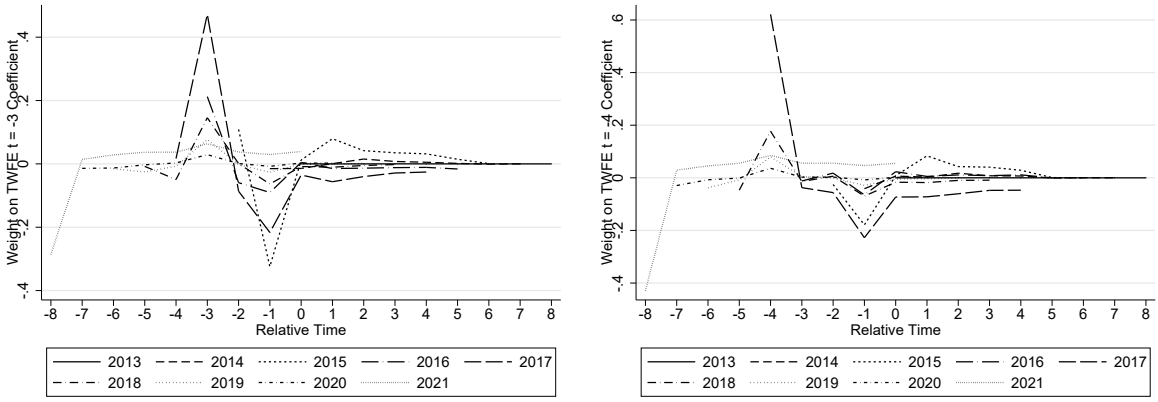
$$\beta_{-3} = \sum_{e=2013}^{2021} \omega_{e,-3}^{-3} CATT_{e,-3} + \sum_{l \in \{-7,-6,\ldots,0,\ldots,6,7\}} \sum_{e=2013}^{2021} \omega_{e,l}^{-3} CATT_{e,l} + \sum_{l' \in \{-8,-1\}} \sum_{e=2013}^{2021} \omega_{e,l'}^{-3} CATT_{e,l'} \tag{B.2}$$

where $CATT_{e,l}$ refers to average treatment effect for cohort e in relative time period l. For instance, $CATT_{2015,-2}$ is the average treatment effect for the relative time period $t = -2$ for the cohort that got treated in the year 2015. The first term captures the true parameter of interest associated with $t = -3$. The third term reduces to zero under the no anticipation assumption. The second term reduces to zero under the assumption of treatment effect homogeneity. Sun and Abraham (2021) show that if treatment effects are heterogeneous, dynamic effects of 4G introduction across different cohorts will affect the TWFE estimates of β_{-3} making it a function of post-treatment $CATT_{e,l \geq 0}$. Therefore, a test of pre-trends using the statistical significance of pre-treatment coefficients becomes invalid even in the absence of pre-trends.

Appendix Figures B.2a and B.2b respectively illustrate the decomposition of weights, as described in equation B.2, for the pre-treatment leads at $t = -3$ and $t = -4$ in specification B.1. Weights associated with $CATT_{2015,1}$ and $CATT_{2017,1}$ receive the highest absolute weights of magnitudes of 0.079 and -0.056, respectively. If treatment effects for relative time period $t = 1$ are different across these cohorts, the interaction of non-zero weights and heterogeneous treatment effects for these cohorts will contaminate the estimates of β_{-3}. Thus, under treatment effect heterogeneity, β_{-3} becomes a function of post-treatment $CATT_{e,l \geq 0}$. making any test for pre-trends using

the statistical significance of pre-treatment coefficients invalid.

Figure B.2: Weight underlying pre-treatment coefficients β_{-3} and β_{-4}



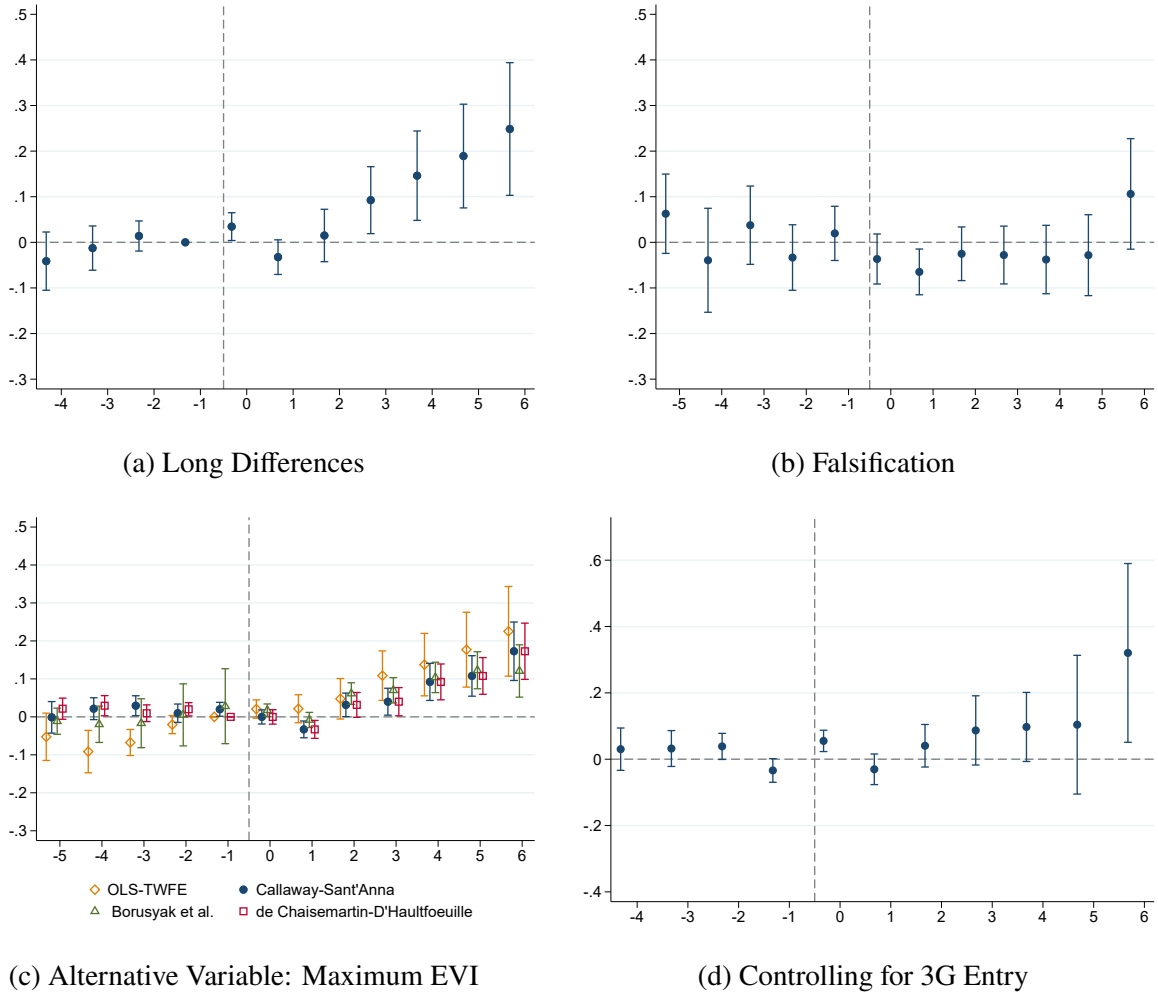
(a) Weights underlying β_{-3}

(b) Weights underlying β_{-4}

This figure plots the magnitude of weights associated with each $CATT_{e,l}$ in the weights decomposition of β_{-3} and β_{-4} as specified in Equation B.2. $CATT_{e,l}$ refers to average treatment effect for cohort e in relative time period l. For instance, $CATT_{2015,-2}$ is the average treatment effect for the relative time period $t = -2$ for the cohort that got treated in the year 2015. The weights are estimated using the methodology outlined in Sun and Abraham (2021). Figures B.2a and B.2b plot the weights underlying β_{-3} and β_{-4}, respectively.

Appendix C Robustness

Figure C.1: Robustness: Effect of 4G on EVI implied agricultural yield



(a) Long Differences

(b) Falsification

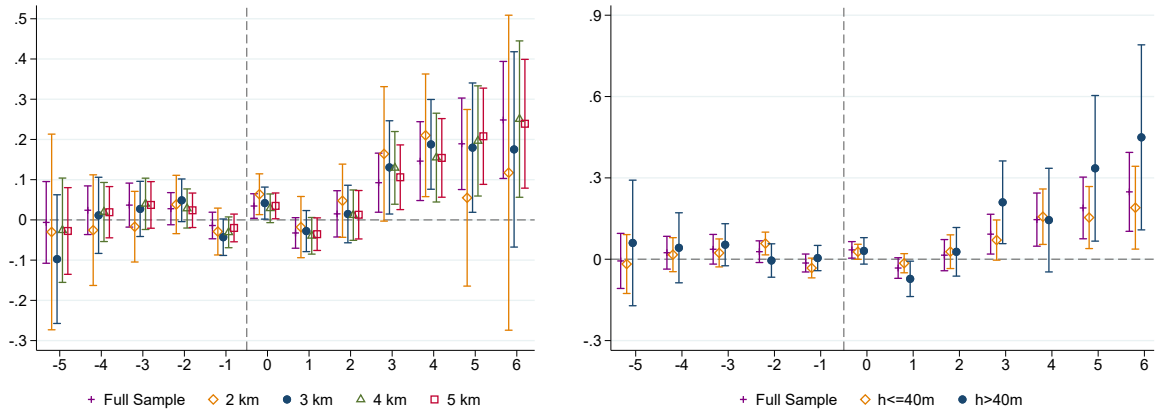
(c) Alternative Variable: Maximum EVI

(d) Controlling for 3G Entry

This figure plots the event-study dynamic coefficients estimated using the methodology outlined in Callaway and Sant'Anna (2021). Figure C.1a presents the Callaway and Sant'Anna (2021) estimator for the baseline presented in Figure 3 with pre-treatment effects calculated as long differences to make these estimators comparable to conventional TWFE plots. The outcome variable is EVI implied agricultural yield constructed by subtracting the average value of EVI during the initial weeks of Kharif season from the maximum value of EVI during the Kharif season. Figure C.1b presents a falsification test with a sample of all unique hexagons that saw the introduction of 4G BTS between 2013 and 2021 and have zero cropland as of 2010 according to the GFSAD database. The outcome variable is the maximum EVI of the kharif season. The dependent variable is standardized to mean zero and standard deviation of one. We use the maximum EVI measure here instead of EVI implied yield because the latter is a difference measure and is mostly centered around zero for non-croplands. This is because, unlike the cropped areas, the measure does not exhibit significant variation across the kharif season. Figure C.1c replicates the results in baseline Figure 3 while using the maximum value of EVI during the kharif season as a measure for agricultural yield and the four estimators: (1) dynamic TWFE estimator (in orange with hollow diamond markers), (2) Callaway and Sant'Anna (2021) estimator (in blue with solid circle markers), (3) Borusyak, Jaravel, and Spiess (2022) estimator (in green with hollow triangle markers), and (4) De Chaisemartin and d'Haultfoeuille (2020) estimator (in pink with hollow square markers). Figure C.1d replicates the baseline Figure 3 while controlling for the event time of 3G entry in the hexagon. All measures are winsorized at 1% level. Standard errors for all estimators are clustered at the district level. The bars represent 95 percent confidence intervals.

C.1 Uniform Treatment Effect Assumption

Figure C.2: Robustness (Uniform Treatment Effect Assumption): Effect of 4G on EVI implied agricultural yield



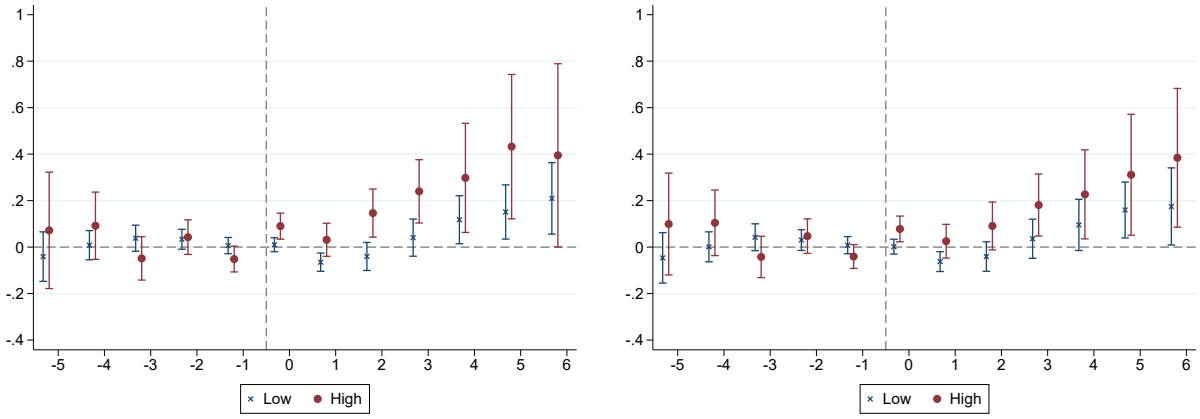
(a) Distance from hexagon centroid

(b) Height of the 4G tower

This figure plots the event-study dynamic coefficients estimated using the methodology outlined in Callaway and Sant'Anna (2021). Figure C.2a tests for violations of the assumption of uniform treatment within a hexagon by constructing sub-samples of hexagons based on the distance of the first 4G BTS from the centroid of the hexagon. Specifically, we create these sub-samples by keeping hexagons where the first 4G BTS was introduced within a radius of – (i) 5km (in cranberry with hollow square markers), (ii) 4km (in green with hollow triangle markers), (iii) 3km (in blue with solid circle markers), (iv) 2km (in dark orange with hollow diamond markers) – from the centroid of the hexagon. The baseline results for the full sample are plotted in purple with a plus sign for comparison. Figure C.2b presents another test for violations of uniform treatment assumption within hexagons by constructing sub-samples of hexagons based on the height of the 4G telecom tower. Based on the distribution of the height of the 4G towers given in Panel D of Table 1, we split the hexagons into two sub-samples – (i) $h <= 40m$ (in dark orange with hollow diamond markers), is a sub-sample of hexagons where the height of the first 4G BTS tower was less than or equal to 40m, (ii) $h > 40m$ (in blue with solid circle markers), is a sub-sample of hexagons where the height of the first 4G BTS tower was greater than 40m. The baseline results for the full sample are plotted in purple with a plus sign for comparison. The outcome variable is EVI implied agricultural yield constructed by subtracting the average value of EVI during the initial weeks of Kharif season from the maximum value of EVI during the Kharif season. The dependent variable is standardized to mean zero and standard deviation of one. All measures are winsorized at 1% level. Standard errors for all estimators are clustered at the district level. The bars represent 95 percent confidence intervals.

C.2 Heterogeneous Treatment Effect by Agricultural Production Potential

Figure C.3: Heterogeneous Treatment Effect by Agricultural Production Potential



(a) Available Land-based measure

(b) Global Land-based measure

This figure plots the event-study dynamic coefficients estimated using the methodology outlined in Callaway and Sant'Anna (2021). The outcome variable is EVI implied agricultural yield constructed by subtracting the average value of EVI during the initial weeks of kharif season from the maximum value of EVI during the kharif season. The dependent variable is standardized to mean zero and standard deviation of one. The sample consists of all unique hexagons that saw the introduction of 4G BTS between 2013 and 2021. Agricultural production potential is measured using two measures provided in Advancing Research on Nutrition and Agriculture's (ARENA) Demographic and Health Surveys (DHS)-GIS Database. Measure 1 provides the combined suitability of currently available land for pasture and rainfed crops. Measure 2 provides the combined suitability of the global land area for pasture and rainfed crops. We split the hexagons into two sub-samples based on the values of agricultural production potential. A hexagon is defined as an area with high agricultural production potential if it is marked as land well-suited or prime land for rainfed crops, i.e., the CSI (Crop Suitability Index) is greater than equal to 50; otherwise, the hexagon is classified as low potential area. Panel C.3a and C.3b report result for measures 1 and 2 based on total available land and total global land, respectively. All measures are winsorized at 1% level. Standard errors for all estimators are clustered at the district level. The bars represent 95 percent confidence intervals.

Appendix D Identification Using Rights of Way

Table D.1: Rights of Way (RoW) rules status across states

	State	RoW policy notified/ approved by Cabinet	Approval Date	Draft Policy released	Existing policy under advance discussion	No uniform policy
1	Jharkhand	Yes	2015-12-04	-	-	-
2	Rajasthan	Yes	2017-02-06	-	-	-
3	Tripura	Yes	2017-09-06	-	-	-
4	Odisha	Yes	2017-09-14	-	-	-
5	Haryana	Yes	2017-10-06	-	-	-
6	Assam	Yes	2018-02-16	-	-	-
7	Maharashtra	Yes	2018-02-17	-	-	-
8	Tamil Nadu	Yes	2018-02-18	-	-	-
9	Arunachal Pradesh	Yes	2018-05-10	-	-	-
10	Uttar Pradesh	Yes	2018-06-15	-	-	-
11	Uttarakhand	Yes	2018-09-13	-	-	-
12	Meghalaya	Yes	2018-12-20	-	-	-
13	Madhya Pradesh	Yes	2019-03-08	-	-	-
14	Karnataka	Yes	2019-05-29	-	-	-
15	Manipur	Yes	2019-11-28	-	-	-
16	Nagaland	Yes	2019-12-02	-	-	-
17	Gujarat	No	-	-	Yes	-
18	Daman and Diu	No	-	-	-	Yes
19	Lakshadweep	No	-	-	-	Yes
20	Dadra and Nagar Haveli	No	-	-	-	Yes
21	Goa	No	-	-	Yes	-
22	Jammu and Kashmir	No	-	Yes	-	-
23	Punjab	No	-	Yes	-	-
24	Kerala	No	-	Yes	-	-
25	Puducherry	No	-	-	Yes	-
26	Himachal Pradesh	No	-	-	-	-
27	Andhra Pradesh	No	-	-	Yes	-
28	Chandigarh	No	-	-	Yes	-
29	NCT of Delhi	No	-	Yes	-	-
30	Telangana	No	-	-	Yes	-
31	Chhattisgarh	No	-	-	Yes	-
32	Bihar	No	-	-	Yes	-
33	West Bengal	No	-	-	-	Yes
34	Sikkim	No	-	Yes	-	-
35	Mizoram	No	-	Yes	-	-
36	Andaman and Nicobar	No	-	-	-	Yes

This table lists the status of RoW policies (as of September 2020) across States/Union territories in India as per the GSMA (2020) report. The report highlights the importance of States' adoption of Rights of Way rules in line with the Indian Telegraph RoW Rules, 2016, issued by the Ministry of Telecom, Government of India. These policies were rolled out to promote the expansion of telecom infrastructure. According to the report, as of September 2020, 16 states had adopted RoW rules. Additionally, 15 states and union territories were deliberating on the passage of these rules in the state legislature.