NBER WORKING PAPER SERIES

INFORMATION, TECHNOLOGY ADOPTION AND PRODUCTIVITY: THE ROLE OF MOBILE PHONES IN AGRICULTURE

Apoorv Gupta
Jacopo Ponticelli
Andrea Tesei

Working Paper 27192
http://www.nber.org/papers/w27192

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2020

We received valuable comments from Siwan Anderson, Shawn Cole, Nicola Gennaioli, Sabrina Howell, Dean Karlan, Marco Manacorda, Pepita Miquel, Imran Rasul, Gabriella Santangelo, Chris Udry and seminar participants at Columbia, Universidad de Los Andes, Universidad Javeriana, University of Zurich, University of Bonn, Lancaster University, NYU Stern, Bocconi University, Queen Mary University, University of Maryland, Northwestern University, Berkeley Haas, UBC, CUNY, BGSE Summer Forum, European Economic Association 2019 and the IPA 2019 Researcher Gathering. Ponticelli gratefully acknowledges financial support received for this project from the Cohen and Keenoy Faculty Research Fund and the Fama-Miller Center at the University of Chicago Booth School of Business. The project was developed while Andrea Tesei was visiting the Ford Motor Company Center for Global Citizenship at the Kellogg School of Management, Northwestern University, whose hospitality is gratefully acknowledged. Pierre Jaffard, Gursharan Bhue and Mark He provided excellent research assistance. We are grateful to the staff at GSMA for their help with the mobile phones data. The GSMA data used in this study are covered by a confidential license agreement. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Information, Technology Adoption and Productivity: The Role of Mobile Phones in Agriculture
Apoorv Gupta, Jacopo Ponticelli, and Andrea Tesei
NBER Working Paper No. 27192
May 2020
JEL No. O33,O4,Q16,Q55

ABSTRACT

We study the effect of information on technology adoption and productivity in agriculture. Our empirical strategy exploits the expansion of the mobile phone network in previously uncovered areas of rural India coupled with the availability of call centers for agricultural advice. We measure information on agricultural practices by analyzing the content of 2.5 million phone calls made by farmers to one of India's leading call centers for agricultural advice. We find that areas receiving coverage from new towers and with no language barriers between farmers and advisers answering their calls experience higher adoption of high yielding varieties of seeds and other complementary inputs, as well as higher increase in agricultural productivity. Our estimates indicate that information frictions can explain around 25 percent of the agricultural productivity gap between the most productive and the least productive areas in our sample.

Apoorv Gupta
Northwestern University
2211 Campus Drive
Room 4470
Evanston, IL 60208
apoorv-gupta@kellogg.northwestern.edu

Jacopo Ponticelli
Kellogg School of Management
Northwestern University
2211 Campus Drive
Evanston, IL 60208
and NBER
jacopo.ponticelli@kellogg.northwestern.edu

Andrea Tesei
School of Economics and Finance
Queen Mary University of London
327 Mile End Road
London E1 4NS
and CEPR, CEP (LSE) & CESifo
a.tesei@qmul.ac.uk

1 Introduction

Agricultural workers in the richest 10 percent of countries produce, on average, 50 times more output per worker than those in the poorest 10 percent (Gollin, Lagakos, and Waugh 2014). One often cited explanation for these large productivity differences is the sluggish adoption of modern agricultural technologies in developing countries. What are the constraints on the adoption of these technologies and to what extent they contribute to the observed differences in productivity, however, remains an open question (see reviews in Jack 2013, Foster and Rosenzweig 2010 and Feder, Just, and Zilberman 1985). Since agriculture employs a significant share of workers in the developing world, addressing this issue is crucial to our understanding of income inequality across countries more generally.[1]

In this paper we focus on the role of information in explaining differences in technology adoption and productivity in the agricultural sector. Several scholars have pointed to the importance of learning and information frictions in the diffusion of new technologies in agriculture (Foster and Rosenzweig, 1995; Conley and Udry, 2010). Yet, empirically quantifying the impact of information in closing productivity differences across farmers remains challenging. This is because it is usually hard to observe access to information about agricultural technologies, actual adoption of these technologies and productivity. We address this challenge by bringing together data on the expansion of the mobile phone network in previously unconnected areas of rural India, with geo-located data on phone calls made by farmers seeking agricultural advice and detailed survey data on agricultural inputs and yields covering around 19 million farmers. Our goal is to investigate whether access to the mobile phone network – coupled with the availability of call centers for agricultural advice – helps to disseminate timely and reliable information about modern agricultural practices and inputs, and whether access to this information fosters technology adoption and productivity.

We study this question in the context of India, a country where a large share of the population is employed in agriculture and where differences in agricultural productivity across regions are large. India offers a unique setting to study the role of information on technology adoption and productivity. As late as 2003, 60 percent of Indian farmers reported not having access to any source of information on modern technology to assist them in their farming practices (National Sample Survey, 2005). In the mid-2000's, however, the Indian government launched two programs that plausibly increased farmers' access to information on agricultural practices. First, the Shared Mobile Infrastructure Program (SMIP), which brought mobile phone coverage in previously unconnected areas through the construction of more than 7,000 mobile phone towers. Second, the Kisan Call Centers (KCC), which provided a free-of-charge, phone-based service of agricultural advice for Indian farmers. In our empirical analysis we exploit the heterogeneous exposure of farmers to these two programs to generate plausibly exogenous variation in their access to information. We measure farmers' access to information on farming practices based on the content and location of approximately 2.5 million phone calls made by farmers to the KCC.

The analysis proceeds in two steps. In the first step, we use an event-study design to doc-

[1] On the magnitude of labor productivity differences across countries – both in the aggregate and in the agricultural sector – see Caselli (2005) and Restuccia, Yang, and Zhu (2008). On differences in technology adoption – and the pattern of technology diffusion – across countries see Comin and Hobijn (2004) and Caselli and Coleman (2001).

ument the evolution of farmers' calls to KCC when new mobile phone towers are constructed in previously uncovered areas. Using high-frequency (monthly) variation, we show a significant increase in the number of calls following the construction of the first tower in a given area, which is consistent with a large and underserved demand for agricultural advice among Indian farmers (Cole and Fernando, 2016). We also show that the increase in the number of calls varies depending on the local language spoken in an area. The reason is that, although the agricultural advice provided by KCC is in principle available to all farmers with access to a phone, KCC agronomists answer calls only in one of the 22 official languages recognized by the Indian Constitution.[2] This effectively creates a language barrier to access the service for around 40 million individuals, mostly concentrated among the rural population, who speak one of the approximately 100 non-official languages of India. Consistently, we observe a smaller increase in calls from areas where the majority of the local population speaks a non-official language, despite these areas are comparable in terms of observable socio-economic characteristics and pre-existing trends in the outcomes of interest.

In the second step of our analysis, we turn to study the real effects of access to information. To this end, we match detailed survey data on agricultural inputs and crop yields with geo-located data on the diffusion of mobile phone coverage. We propose an identification strategy that compares – within each administrative district – locations where new SMIP mobile phone towers were proposed and eventually constructed, with locations where new towers were also proposed but eventually not constructed. In addition, we exploit variation in the spatial diffusion of non-official languages to capture the farmers' ability to access phone-based services for agricultural advice. We think of the combination of mobile phone coverage and absence of language barriers with agricultural advisers as a positive shock to information about agricultural practices for farmers. This identification strategy allows us to disentangle the effect of information about agricultural practices from other potential mechanisms linking the arrival of mobile phones to technology adoption and productivity in a given area.

We start by documenting that areas with a larger increase in potential access to information experienced larger adoption of more advanced agricultural technologies. We focus in particular on farmers' adoption of high-yielding variety (HYV) seeds, chemical fertilizers and pesticides, as well as irrigation systems. HYV seeds are commercially developed to increase crop yields and are one of the most prominent innovations in modern agriculture.[3] Chemical fertilizers and reliable irrigation systems are key complementary inputs to maximize HYV potential. We find that areas with a one standard deviation larger increase in mobile phone coverage experienced a 1.4 percentage points larger increase in area farmed with HYV seeds between 2007 and 2012.[4] This effect corresponds to a 5.3 percent increase in land cultivated with HYV seeds for the average cell in our sample.[5] This effect is concentrated in areas with no language barriers between farmers and agricultural advisers. We find positive and significant effects also on the adoption

[2] The 2011 Census identifies 121 languages spoken in India, 22 of which are part of the Eight Schedule of the Constitution, i.e. they are recognized as official languages of the Republic of India.

[3] On the impact of high-yielding varieties on agricultural productivity and economic development see, among others, Evenson and Gollin 2002, 2003.

[4] Data on farmers' use of agricultural inputs is sourced from the Agricultural Input Survey of India, which is carried out at 5 year intervals. The last two waves for which data is available are 2007 and 2012.

[5] The units of observation are cells of 0.083 $\times$ 0.083 degree resolution, approximately corresponding to areas of 10$\times$10 km at the equator.

of chemical fertilizers, pesticides and irrigation. In line with an information mechanism, we show that the same areas also experienced a larger increase in farmers' calls seeking information exactly on these technologies.

Next, we study the effect of farmers' access to information on agricultural productivity, as measured by average crop yields. Our results indicate that areas where farmers had a larger increase in potential access to information experienced a larger increase in agricultural productivity. In particular, our estimates indicate that areas with a one standard deviation larger increase in mobile phone coverage and no language barriers between farmers and agricultural advisers experienced a 1.3 percent larger increase in agricultural yields between 2007 and 2012. Within our sample of rural areas with no initial mobile phone coverage there is large variation in the baseline level of agricultural productivity. In 2007, the average yield of an area at the 75th percentile of agricultural productivity was around twice as large as the one observed at the 25th percentile. This is a yield gap similar to the one observed in rice and wheat production between the richest 10 percent and the poorest 10 percent of countries (Gollin, Lagakos, and Waugh, 2014). Our estimates indicate that differences in access to information on agricultural practices can explain around 25 percent of this productivity gap. We also show that the effect of access to information is heterogeneous across areas with different initial productivity, and it is the largest for those in the lowest productivity quartile. This indicates that access to information on agricultural practices can reduce productivity differences across farmers by increasing the productivity of the initially less productive ones.

Our results are robust to a number of checks and alternative specifications. First, we show that there are no pre-existing trends in technology adoption and crop yields between cells that received coverage from new towers and cells that did not. In addition, we show that the lower adoption of modern agricultural technologies and the smaller productivity gains in areas characterized by language barriers between farmers and KCC agricultural advisers are not driven by other factors, such as geographical isolation or income levels, that may potentially be correlated with the diffusion of non-official languages in a given area. We also discuss and test for other potential mechanisms linking mobile coverage with technology adoption and productivity. In particular, previous studies have shown that by providing detailed and timely information on prices, mobile phones can reduce price dispersion, favor a more efficient allocation of goods across markets and generate higher incomes for goods producers (Jensen, 2007). This, in turn, can help farmers pay the fixed cost of adopting new technologies. We test to what extent our results might be driven by the effect of mobile phones on price dispersion by augmenting our main specification with fixed effects for the closest agricultural market to each cell in our sample (in addition to the district fixed effects included in all specifications). This allows us to compare outcomes across farmers who plausibly face the same prices for their products and experience the same changes in local demand. We show that all our main results are robust to this augmented specification.

Related Literature

Our paper is related to different strands of literature. An influential body of work in macroeconomics and development has documented the existence of substantial differences in productivity across countries and investigated their determinants. Some studies have focused on productivity differences in the manufacturing sector and the extent to which they can be explained by the misallocation of factors of production across heterogeneously productive firms (Restuccia and Rogerson 2008, Hsieh and Klenow 2009). Other studies have focused on productivity differences in the agricultural sector (Gollin, Lagakos, and Waugh 2014). These are larger than differences in aggregate labor productivity, suggesting that the productivity gap in agriculture is particularly important for our understanding of income differences across countries (Caselli 2005, Restuccia, Yang, and Zhu 2008). Several potential explanations of this productivity gap have been proposed, including land misallocation, lack of insurance markets, or frictions in the reallocation of workers from agriculture to the non-agricultural sectors (Adamopoulos and Restuccia 2014, Lagakos and Waugh 2013, Gollin, Lagakos, and Waugh 2014, Donovan 2016). Relative to these studies, we emphasize and quantify the role played by information frictions.

Our work also contributes to a large literature studying the impact of technology on productivity growth (Aghion and Howitt, 1992; Comin and Hobijn, 2010). Empirical studies in this area have tried to understand the determinants of technology adoption in the manufacturing sector across and within countries. Cross-country studies have documented the role played by financial development and management practices (Comin and Nanda, 2019; Bloom and Van Reenen, 2007), while studies at a finer level of geographical detail have pointed to organizational barriers and misaligned incentives (Atkin, Chaudhry, Chaudry, Khandelwal, and Verhoogen, 2017). Compared to these studies, we document the role of information diffusion on technology adoption in agriculture, a sector plausibly characterized by higher information barriers and lower organizational levels than the manufacturing sector.

Our paper is also related to the micro-development literature investigating the role of modern agricultural technologies – such as high-yielding variety seeds – in the process of development. This literature has studied several potential frictions to the adoption of modern technologies by farmers in developing countries, including credit constraints (Duflo, Kremer, and Robinson 2004), missing insurance markets (Karlan, Osei, Osei-Akoto, and Udry 2014), lack of access to high-quality inputs (Bold, Kaizzi, Svensson, and Yanagizawa-Drott 2017), or lack of a reliable transportation infrastructure (Asher and Novosad 2020).[6] Among these frictions, the lack of information on new technologies or how to use them has received extensive attention (see De Janvry, Sadoulet, Manzoor, and Kyle 2016 for a recent review). This literature includes work grounded on learning models of new technologies based on farmers' own experience or the experience of others in their social network (Foster and Rosenzweig, 1995; Conley and Udry, 2010; Munshi, 2004; Hanna, Mullainathan, and Schwartzstein, 2014; Beaman, BenYishay, Magruder, and Mobarak, 2018).[7] Studies in this area have also highlighted the mixed record of traditional agricultural extension programs (Duflo, Kremer, and Robinson, 2008). In particular, researchers and policy makers have long identified the lack of timely and personalized information as obstacles to the effectiveness of the communication between farmers and extension workers (Anderson and Feder, 2004). In this respect, our findings indicate that mobile phone-based extension programs that can provide timely and far-reaching information to farmers, adapted to their individual needs at different points in the production cycle and the specific agro-climatic characteristics

[6] Suri (2011) emphasizes how low adoption rates in developing countries mask large disparities in returns from adoption across farmers.

[7] The extent to which social networks represent a reliable source of information on agricultural practices and technologies is unclear, as neighboring farmers and agricultural input dealers may be either poorly informed or misinform farmers due to misaligned incentives (Anderson and Birner, 2007).

of their area, can have real effects on agricultural practices and productivity. At the same time, our findings emphasize that language barriers can have a significant role in preventing part of the rural population employed in agriculture from benefiting from these programs.

Our results complement those in the recent experimental literature studying the impact of mobile phone-based extension programs on the diffusion of information about agricultural practices and farmers' behavior. Casaburi, Kremer, Mullainathan, and Ramrattan (2014) show that sending text messages containing agricultural advice has significant positive effect on yields of small sugarcane farmers in Kenya. Cole and Fernando (2016) randomize access to a hot line for agricultural advice to households in the Indian state of Gujarat, finding evidence that the use of this phone service has a significant impact on agricultural practices, although relatively small effects on yields. They also find that information provided through mobile phones spread within farmers' network, amplifying the effect of the agricultural extension program.[8] These studies are part of a larger literature on the impact of mobile phones in developing countries (see Aker, Ghosh, and Burrell 2016 and Nakasone, Torero, and Minten 2014 for recent reviews).[9]

The rest of the paper is organized as follows. Section 2 introduces the data used in the analysis, and provides institutional background on the diffusion of mobile phones in India and on the two government programs – the Shared Mobile Infrastructure Program and the Kisan Call Centers for agricultural advice – that are central to our empirical analysis. Section 3 presents our identification strategy and all the empirical results. Section 4 provides concluding remarks.

2 Data, Institutional Background, and Stylized Facts

In this section we describe the main datasets used in the empirical analysis, provide some institutional background for the government programs used for identification, and present a set of stylized facts that emerge from the raw data. The unit of observation in our empirical analysis are areas of $10{\times}10$ km, which we refer to as cells. We use a grid of $10{\times}10$ km cells to match information from the datasets presented below, which come at different levels of geographical aggregation, which could be an administrative division such as a village or a subdistrict, or a geo-referenced polygon in the case of mobile phone coverage data.[10]

[8] Relatedly, Fafchamps and Minten (2012) study the impact of a text message-based agricultural information system providing market and weather information to Indian farmers and find non significant effects on cultivation practices or productivity.

[9] Jensen (2007) and Aker (2010) show that mobile phone coverage can reduce price dispersion in, respectively, fisheries in Southern India and agricultural goods markets in Niger. Jack and Suri (2014) study the impact of lowering transaction costs to transfer money among individuals on risk sharing. They find that households using a mobile phone system that reduces transaction costs are better able to smooth consumption when facing negative income shocks.

[10] Overall, India can be split into 41,495 cells distributed over 524 districts. Since cell borders do not typically correspond to district administrative borders, we assign cells spanning over more than one district to the district which occupies the largest area. One challenge that we face is that Indian districts have been changing shape, or were created or dissolved during the period under study. In order to define districts consistently over time, we created minimum comparable areas (MCAs) encompassing one or more districts that cover the same geographical space between 1997 and 2012. The main source used to re-construct district changes over time is the Population Census Map, which contains a short history of how each district was created.

2.1 Data on Mobile Phone Coverage and its Diffusion in India

We use data on the diffusion of mobile phone coverage in India provided by the Global System for Mobile Communication Association (GSMA), the association representing the interests of the mobile phone industry worldwide.[11] The data is collected by GSMA directly from mobile operators and refers to the GSM network, which is the dominant standard in India with around 89 percent of the market share in 2012 (Telecom Regulatory Authority of India, 2012). The data licensed to us provide, for all years between 1998 and 2012, geo-located information on mobile phone coverage aggregated across all operators.[12] Our analysis focuses on the 2G technology, the generation of mobile phones available in India during the period under study, which allows for phone calls and text messaging.[13]

Figure 1 reports the geographical diffusion of 2G GSM mobile phone coverage in India at five-year intervals between 1997 and 2012. As shown, India had virtually no mobile phone coverage as of 1997. From then on, the mobile phone network expanded rapidly, covering 22 percent of the population in 2002, 61 percent in 2007 and 89 percent in 2012.[14] Data from the World Bank (2014) indicate that mobile phone subscriptions per 100 people in India went from 0.08 in 1997 to 68.4 in 2012.

Following a standard pattern of diffusion (Buys, Dasgupta, Thomas, and Wheeler, 2009; Aker and Mbiti, 2010), the spatial roll-out of mobile phone coverage in India started in urban areas and only later reached rural ones. We document this pattern in Figure C.1, which reports - at 5-year intervals between 1997 and 2012 - the average share of land covered by mobile phones across cells with different initial levels of urbanization. As a proxy for urbanization we use night light intensity in 1996. As shown, in 1997, there was virtually no mobile phone coverage in either urban or rural areas. By 2002, areas in the highest decile of night light intensity had, on average, 40 percent of their area covered by the mobile phone network, more than 80 percent in 2007, and close to full coverage by 2012. On the other hand, mobile phone coverage in the lowest decile was, on average, still almost non-existent in 2002, around 20 percent by 2007 and around 40 percent by 2012.

2.2 Construction of mobile phone towers under the SMIP Government Program

The Indian government played an important role in the expansion of the mobile phone network in rural areas, where market demand did not justify infrastructural investment by private telecommunication companies. In 2007, the government launched the Shared Mobile Infrastruc-

[11] The data collection effort is in partnership with Collins Bartholomew, a digital mapping provider.

[12] The extent of geographical precision of the original data submissions ranges between 1 km^2 on the ground for high-quality submissions based on GIS vector format, and 15-23 km^2 for submissions based on the location of antennas and their corresponding radius of coverage (GSMA, 2012; Sauter, 2006).The data have been used by Manacorda and Tesei (2020) to study the effects of mobile phone coverage expansion on political mobilization in Africa.

[13] The 3G spectrum was allocated to private operators only at the end of 2010 and the roll-out of commercial operations was very slow. By 2015, 3G penetration was just 20 percent in urban areas and much lower in rural areas (Ericsson, 2015).

[14] We use data from the Gridded Population of the World, Version 4. We assume that population is uniformly distributed within each 10×10 km cell and we use information on the share of each cell's area that is covered by mobile phone technology to compute the fraction of individuals reached by the mobile phone signal in each cell/year. We then aggregate across cells to obtain the share of population covered by mobile phone signal in the country in a given year.

ture Program (SMIP), aimed at providing subsidies to telecom operators for the construction and maintenance of mobile towers in identified rural areas without existing mobile coverage. Under Phase-I of the program, a total of 7,871 sites across 500 districts were initially identified as potential locations for new towers. Villages or cluster of villages not covered by the mobile phone network and with a population of at least 2,000 were prioritized. Telecom operators receiving government subsidies were responsible for installing and maintaining the towers between 2007 and 2013.[15] Of the 7,871 proposed towers under Phase-I, 7,353 were eventually constructed.

We obtained data on the towers constructed under SMIP from the Center for Development of Telematics (C-DoT) - the consulting arm of the Department of Telecommunications of India. The C-DoT provided us with the geographical coordinates of the location of the 7,871 initially proposed towers, the geographical coordinates of the location of the 7,353 effectively constructed towers, and the operational date of each tower. The latter is the date in which the construction of the tower is completed and the tower becomes operational. For simplicity, in the remainder of the paper we refer to this date as the date of construction. From the 7,353 towers constructed under Phase I of the SMIP program we remove 350 towers for which the construction date is missing. This leaves us with 7,003 mobile towers used in our empirical analysis. Figure 2 shows a timeline of construction of these towers by month. As shown, the construction of towers effectively started in January of 2008 and ended in May of 2010, with most towers being introduced between the second half of 2008 and the first half of 2009. To estimate the potential coverage of each tower, we assume a 5-km radius of coverage around the towers' location, based on information reported in tender documents obtained from the C-DoT officials responsible for the Phase I implementation (tender document No. 30-148/2006-USF).

2.3 Data on farmers' calls to Kisan Call Centers

To investigate the role of information on agricultural practices we use data on farmers' calls to Kisan Call Centers (KCC), which we obtained from the Department of Agriculture, Cooperation and Farmers Welfare. Calls are geo-located at the subdistrict (or block) level and we assign them proportionally to all cells whose centroid is contained in the subdistrict.[16]

KCC were introduced in January 2004 by the Indian Ministry of Agriculture and were the first providers of general agricultural advice to farmers via mobile phone in India.[17] KCC are

[15] A second Phase of the scheme was also planned to be launched shortly after Phase-I to cover even more sparsely populated areas, but was never implemented.

[16] On average, there are 27 cells per subdistrict. Whenever information on the subdistrict from which the call is originated is missing, we use information on the district of the call and the crop for which the caller is seeking information to assign calls to a given cell. Our probabilistic assignment rule is described in the following equation:

$$Calls_{idt} = \sum_{c \in O_i} (Calls)_{cdt} \times \left(\frac{Area_{idc,t=2000}}{Area_{dc,t=2000}} \right)$$

The first element of the product captures the number of calls about a given crop c that are originated from district d, while the second element of the product captures the share of crop c that is farmed in cell i over the total area farmed with the same crop in district d (sourced from the FAO-GAEZ data). Thus, this assignment rule implies that if 10 percent of the area farmed with rice in district d is farmed in cell i, 10 percent of the calls about rice received from farmers located in district d will be assigned to cell i.

[17] Figure C.2 shows the timing of introduction of the largest Indian providers of agricultural advice via mobile phones. Other early development extensions, like aAQUA and NanoGanesh, established in 2003 and 2004 respectively, focused on SMS-based advice on agricultural practices and irrigation techniques, respectively.

available in all Indian states and allow farmers to call a toll-free number to get answers to their questions. In total, during the 2006-2012 period, farmers made around 2.5 million calls to KCC. The number of calls increased substantially starting in 2009, reaching over half a million per year between 2009 and 2011, and over eight hundred thousands in 2012.[18]

For every call received in one of the 25 call centers that are part of the KCC network, the agronomist collects basic information on the farmer (name, location and contact information), date and time of the call, a brief description of the question, the crop for which the query is made, and the response provided.[19] The calls are answered by trained KCC agricultural graduates, who address the query based on their knowledge and on a database of previous answers to similar queries. Approximately 98 percent of the calls are answered using this database. In case the agronomist is unable to answer the question, the call is forwarded to a senior expert.[20]

Around 50 percent of the calls to KCC are about pests and how to deal with them. In the responses, farmers receive detailed advice on which pesticide (if any) they should use, as well as information on dosage and number of applications. The second most represented category is calls about how to improve yields or – more specifically – which varieties of seeds to use in order to obtain higher yields (13 percent of calls). In these cases, farmers often receive suggestions on which HYV seeds to use based on crop, location, and irrigation system available. Other topics farmers consistently ask about are: fertilizers (10.5 percent of calls), weather conditions (5.7 percent), advice for field preparation (4.6 percent), market price information (3.6 percent), credit information (2.3 percent), and irrigation (1 percent).[21]

In Figure 3 we report the breakdown by month and topic of the call for the two largest crops by cultivated area in India, rice – panel (a) – and wheat – panel (b). A number of patterns emerge. First, the distribution of calls reflects the different farming season of the two crops. Rice is mainly grown during the $kharif$ season, where crops are grown between June and September and harvested between October and February. On the other hand, wheat is mainly grown in the *rabi* season, where crops are grown between October and November and harvested between December and the Spring months. Second, the composition of the calls is consistent with the agricultural calendar just described. For example, rice farmers mostly ask questions about which seeds to use in May and June – at the beginning of the growing season. Instead, when crops are fully grown, most of the calls are about how to defend the plants from pests. Similar patterns can be observed for wheat.

Finally, in Figure C.3 we report the overall distribution of calls to KCC by month, by time of the day and by crop. The figure shows that most calls are received during Summer months

Until 2010, no other provider of general agricultural advice entered the market.

[18] The availability of this service has been largely advertised by the Indian government. The advertising campaign mostly took the form of TV ads. Ads were broadcasted in both public and private TV channels, and at times matching farmer's preferences in different states.

[19] The version of the data provided to us by the Department of Agriculture, Cooperation and Farmers Welfare does not contain farmers' names or contact information. Thus, we cannot identify farmers that call multiple times.

[20] According to an external evaluation of the KCC program, 84% of farmers expressed satisfaction with the advice received, 99% said they would call again if there was a problem, and 96% were willing to recommend the service to their friends.

[21] In Appendix A we provide a detailed description of the keywords that we use to categorize calls to KCC by topic. We classify calls by categories based on the description provided by the operator. Based on these descriptions, we are able to classify 93 percent of the calls to KCC between 2006 and 2012.

(panel a), that the peak number of calls is around late morning hours (panel b) and that most questions are about rice and wheat (panel c).

2.4 Data on Technology Adoption and Agricultural Productivity

Our measures of technology adoption come from the Agricultural Input Survey (AIS), conducted at five-year intervals by the Ministry of Agriculture in coincidence with the Agricultural Census to collect information on input use by Indian farmers. In the survey, all operational holdings from a randomly selected 7 percent sample of all villages in a sub-district are interviewed about their input use.[22] Data from the AIS is aggregated and made available by the Ministry of Agriculture at the district-crop level. Our main empirical analysis focuses on the last two waves of the AIS, 2007 and 2012, while we use earlier survey waves to document pre-existing trends.[23] The main objective of the survey is to collect information on agricultural inputs. In particular, the survey covers the following inputs: seeds – distinguished between traditional and high-yielding varieties – chemical fertilizers, organic manures and pesticides, agricultural machinery and agricultural credit.

Our main measure of technology adoption in agriculture is the share of land farmed with high-yielding varieties (HYV) of seeds. These are hybrid seeds developed via cross-breeding in order to increase crop yields. They combine desirable characteristics of different breeds, including improved responsiveness to fertilizers, dwarfness, and early maturation in the growing season.[24] HYV seeds have been available in India since the Green Revolution (the IR8 rice, flagship of the Green Revolution, was introduced in 1966), but new varieties are constantly developed and introduced in the market. In the period between 2002 and 2013, 47 new varieties of different oilseeds, cereals and vegetables including rice, groundnut, wheat, millet, soy and cotton were introduced to the Indian market. Despite their early introduction and rapid adoption in many areas of the country, a large share of the Indian agricultural land is still not farmed using HYV seeds. The average share of HYV area across cells in our sample in 2007 was 26 percent.

The data on agricultural productivity (yield) also come from the Ministry of Agriculture. The data provide yearly information on covered area and production for all crops at the district level. The yield for each crop is defined as production (in tons) per unit of area farmed (hectares).

3 Empirics

Our empirical analysis proceeds in two steps. First, we use an event-study design to document the evolution of farmers' calls to KCC when new mobile phone towers are introduced in areas without previous coverage. This evidence relies on monthly-level variation in the number of farmers' calls originated from a given location, around the month of construction of the first tower in the area. The event-study also allows us to document the role of language barriers in the diffusion of information. In particular, we show that geographical differences in the diffusion

[22] The AIS was not conducted in the states of Bihar and Maharastra before 2012. Thus, we exclude these states from our analysis.

[23] The Agricultural Input Survey which runs from 1^{st} July, to June 30^{th} of the following year. In the paper, we use the terminology 2007 when referring to the survey carried out between July of 2006 and June of 2007.

[24] Dwarfness makes the plant consume less nutrients for growth and instead use those nutrients to increase production of grains.

of non-official languages among the rural population affect the spatial availability of agricultural advice provided by KCC. We present these results in section 3.1.

Next, we study the real effects of access to information on technology adoption and agricultural productivity. Since technology adoption and productivity are not observed at the same high frequency as farmers' calls, we cannot use the event-study design just described for these outcomes.[25] Instead, we propose an identification strategy that compares locations where new mobile phone towers were proposed and constructed under the SMIP program with similar locations where new towers were proposed but eventually not constructed. We exploit variation in tower construction along with variation in local languages spoken by farmers to capture their ability to access phone-based services for agricultural advice. We focus on the change in technology adoption and productivity between 2007 and 2012, with 2007 being the last wave in the AIS *before* the SMIP program, and 2012 the first wave *after* the SMIP program. We discuss the identification strategy in section 3.2 and present the results in sections 3.3 to 3.5. Finally, in section 3.6 we present a set of additional robustness tests.

3.1 Event-Study Evidence on Farmers' Access to Information

We estimate the evolution of farmers' calls to KCC around the introduction of new mobile phone towers using the following specification:

$$Ln\,(1 + Calls)_{it} = \alpha_i + \alpha_t + \sum_{k=-12}^{+36} \beta_k D_{it}^k + \varepsilon_{it} \tag{1}$$

The outcome variable in equation (1) is the natural logarithm of the total number of calls originated from cell i in month t. D_{it}^k is a dummy equal to 1 if month $t = k$ for cell i, and captures the time relative to the month of introduction of the first tower covering cell i, which we set at $k = 0$. We include the 12 months prior to the introduction of the first tower and the 36 months after. The specification has calendar time and cell fixed effects, denoted by α_t and α_i, respectively. Standard errors are clustered at the district level.

The objective of this exercise is to exploit the different timing of construction of mobile phone towers in different cells to document their impact on farmers' calls. Notice that we focus on cells that will eventually receive a mobile phone tower under the SMIP program described in section 2. Notice also that in this first analysis we focus on the number of calls, while the analysis of their content is discussed in detail in section 3.3.

Panel (a) of Figure 4 reports the estimated coefficients β_k along with their 95 percent confidence intervals. Several findings emerge. First, the coefficients are precisely estimated zeros in the months preceding the introduction of the first tower in a cell. This indicates that the timing of tower introduction is not correlated with pre-existing trends in calls.[26] Second, within 4 months of the construction of the first tower we observe a significant increase in calls for agricultural advice. The magnitude of the estimated coefficients indicates, on average, a 5 to 10 percent increase in the number of calls to KCC in the first year post tower construction. Third,

[25] Data on adoption of agricultural technologies is observed at 5-year intervals in the Agricultural Input Survey, while agricultural yields are observed at yearly level.

[26] Note that farmers can call KCC before the introduction of mobile phone towers using landlines, when available.

this differential continues to grow over the next 18 months, reaching a 40 percent increase in calls three years after the construction of the first tower in a cell.

As discussed in section 2, KCC agricultural advice can in principle be accessed by any farmer with either a landline or a mobile phone connection. KCC agronomists, however, answer farmers' calls only in one of the 22 official languages recognized in the Indian Constitution.[27] This effectively creates a barrier to the service for around 40 million individuals, whose mother tongue is one of the about 100 additional non-official languages spoken in India. Thus, even among areas that receive mobile phone coverage via new SMIP towers, the ability of farmers to access dedicated information on agricultural practices might vary by local language. In panel (b) of Figure 4 we explore this possibility, by estimating equation (1) separately for cells where the majority of the local population speaks one of the 22 official languages and cells where the majority speaks one of the non-official languages.[28] The figure shows that, after the construction of the first mobile phone tower, calls to KCC increase in both groups. However, the increase is much more pronounced in areas where the majority of the local population speaks the same languages as KCC agronomists. Within 3 years from the construction of the first tower, calls in these cells increase by around 30 percentage points more than in those where the majority of the local population speaks a non-official language.

Taken together, the evidence in Figure 4 suggests that the expansion of mobile phone coverage represents a large information shock to farmers, which is larger in areas where farmers may benefit from agricultural advice on best practices and inputs. In the next section, we study how this information shock translates into both technology adoption and agricultural productivity.

3.2 The Real Effects of Access to Information - Identification Strategy

In this section, we present our identification strategy to study the effect of farmers' access to information on real outcomes, namely agricultural technology adoption and productivity. Our identification strategy relies on the two sources of cross-sectional variation that emerge as important determinants of farmers' calls in the event-study setting: availability of mobile phone coverage and share of local population speaking non-official languages. We think of the combination of mobile phone coverage and absence of language barriers with agricultural advisers as a positive shock to information about agricultural practices for farmers. As we discuss in more detail below, the use of both sources of variation allows us to disentangle the effect of information about agricultural practices from other potential mechanisms linking the arrival of mobile phones with technology adoption and productivity.

Our identification strategy exploits variation in the construction of mobile phone towers under the Shared Mobile Infrastructure Program, or SMIP, described in section 2. In the initial phase of this program, the Department of Telecommunications identified 7,871 potential locations for the

[27] See `https://mkisan.gov.in/aboutkcc.aspx`. Agronomists answering in each KCC location answer calls in one (or more) of the official languages.

[28] Data on the share of local population speaking non-scheduled languages is sourced from the 2011 Indian Census and available at the subdistrict level. We assign to all cells whose centroid falls within a given subdistrict the share of local population speaking non-official languages in that subdistrict. The 22 languages recognized by the Indian Constitution as official are: Hindi, Bengali, Marathi, Telugu, Tamil, Gujarati, Urdu, Kannada, Odia, Malayalam, Punjabi, Assamese, Maithili, Santali, Kashmiri, Nepali, Sindhi, Dogri, Konkani, Manipuri, Bodo, and Sanskrit.

construction of mobile phone towers. All the locations in this initial list responded to certain specific criteria, including lack of existing mobile phone coverage and number of individuals potentially covered by the new tower. For identification purposes, we exploit the fact that not all the locations in the initial list eventually received a tower. In some cases, towers were either relocated or not constructed. Thus, we compare cells where towers were initially proposed and eventually constructed with cells in the same administrative district where towers were initially proposed but eventually not constructed. Figure C.4 provides a visual example of how we classify cells into treatment and control group based on proposed and actual tower location.[29] Our final sample consists of 6,320 cells, of which 4,569 in the treatment group and 1,751 in the control group. The summary statistics for the main variables of interest are reported in Table 1. Figure 5 presents the geographical distribution of treatment (in red) and control (in blue) cells across India, while Figure 6 zooms onto Rajasthan – the largest Indian state by area – superimposing the lattice of 10 × 10 km cells to show the level of geographical detail allowed by our data. On average, our sample includes 27 cells per district – 20 treated and 7 control. We further combine this variation with data on the share of local population speaking non-official languages. We report the geographical distribution of the share of local population speaking non-official languages in Figure 7.

The identification relies on the assumption that locations where a tower was proposed but eventually not constructed are a good control group for those that eventually received a tower.[30] The main challenge to our identification is that, although all proposed locations had to meet specific criteria, the decision to relocate or cancel a tower is not random. For example, based on conversations with the C-DoT officials responsible for the implementation of the program, towers were sometimes relocated (or canceled) when, upon visiting the actual site, technicians realized that a relocation would increase the total population covered, or when they discovered logistical issues related to terrain characteristics or lack of an available connection to the electricity grid to power the tower. In what follows, we formally test for differences in the probability of receiving coverage from new SMIP towers based on cell observable characteristics and on pre-existing trends in technology adoption and productivity. We also perform this balance test across cells with different shares of the local population speaking a non-official language, conditional on receiving coverage from new SMIP towers.

The results of the balance tests are reported in Table 2. The outcome in columns (1) to (4) is an indicator variable – $\mathbb{1}$ (Tower) – which is equal to 1 for cells where a new SMIP tower was proposed and eventually constructed, and 0 for cells where a new SMIP tower was proposed but eventually not constructed. Column (1) shows that, in line with the C-DoT officials' account, the conditional probability of eventually receiving a new tower is higher for cells with higher initial population and with flatter terrain, while it does not depend on the availability of a connection to the power grid. Next, in column (2) we study whether pre-trends in agricultural technology adoption or productivity affect the probability of eventually receiving a SMIP tower. As shown, we find no significant differences in technology or productivity growth across treated

[29] We compute coverage for each new tower based on its technical specifications, which corresponds to a 5 km coverage radius around its centroid. As discussed in Section 3.6, our analysis is robust to using the share of land covered by SMIP towers instead of an indicator variable.

[30] This is similar in spirit to, e.g., Greenstone, Hornbeck, and Moretti (2010).

and control cells in the 5 years preceding the tower construction program. In column (3) we then explore the correlation with a number of cell characteristics sourced from the Village Survey of the Population Census of India.[31] Treatment and control cells appear to be comparable along a large set of observable characteristics including: agricultural employment share, share of irrigated land, presence of a school, hospital or bank branch, availability of landline phone connections, night lights intensity, income and expense per capita. The only exception is average distance to the nearest town, which is shorter for the treatment group, although very small in terms of magnitude. In column (4) we consider all previous variables together. The main takeaway is that population and terrain ruggedness remain strong predictors of tower construction, while the other variables are by and large statistically insignificant. In the empirical analysis we add these controls to our specification and show that all our estimates are stable when including the observable cell characteristics reported in Table 2. Finally, in column (5) we condition on cells eventually receiving coverage from new SMIP towers, and explore the correlation between all observable cell characteristics and an indicator variable equal to one for cells where the majority of the population speaks a non-official language, and zero otherwise. As shown, among the treated cells in our sample, the distribution of non-official language speakers is uncorrelated with observable characteristics and pre-trends in technology adoption and productivity.

3.2.1 First Stage

Our first-stage regression is as follows:

$$\Delta Cov_{id} = \alpha_d + \gamma \, \mathbb{1} \, (\text{Tower})_{id} + \delta X_{id} + u_{id} \tag{2}$$

The outcome variable is the change in the share of land covered by the mobile phone network between 2007 and 2012 in cell i, district d. It is important to underline that this variable is constructed using actual mobile coverage data as reported by Indian telecommunication companies to GSMA, i.e. it is not the predicted increase in coverage constructed using SMIP tower location.[32] The coefficient of interest is γ, which captures the effect of tower construction under the SMIP program on the change in coverage in a given cell. X_{id} is a vector of initial cell-level controls, which includes all the cell characteristics reported in Table 2. We include in our specification district fixed effects (α_d) and we cluster standard errors at the district level. Finally, in all specifications we weight each cell by its population at baseline (2001).

Table 3 reports the first-stage results. The estimated coefficient in column (1) indicates that cells covered by new SMIP towers experienced a 11 percentage points larger increase in the share of land covered by mobile phones between 2007 and 2012 relative to the control group. In column (2) we include the three main determinants of tower relocation according to C-DoT officials: population, availability of power supply and terrain ruggedness.[33] The magnitude of

[31] We assign villages to 10 × 10 km cells based on the geographical coordinates for the centroid of the village. The coordinates are obtained from `http://india.csis.u-tokyo.ac.jp`. Village-level information is then aggregated to obtain cell-level characteristics.

[32] The tower construction program we use for identification is not the only driver of changes in mobile phone coverage in these areas. During the same period, private companies also built mobile phone towers across India to extend their services and expand their market shares. Thus, we do not expect tower construction under SMIP to be the sole source of variation in change in GSMA coverage, even in rural regions.

[33] We construct a measure of cell-level terrain ruggedness using the Terrain Ruggedness Index obtained from

the estimated coefficient decreases from 0.11 to 0.073, and remains highly significant. Finally, in column (3) we add all the observable socio-economic cell characteristics. Consistent with the results presented in Table 2, the size of the point estimate is unaffected by including these additional controls. According to the specification in column (3), cells covered by new SMIP towers have, on average, 7.4 percentage points larger share of land covered by mobile phones in 2012 relative to the control group (recall that all these cells have no coverage at baseline). Below the regressions we report the Kleibergen and Paap (2006) first stage F-statistics for the validity of the instrument. We can safely reject that the first stage is weak.

3.2.2 Second Stage: Empirical Specifications

We start by modelling the overall effect of mobile phone coverage on the outcomes of interest – such as the number of calls for agricultural advice, the adoption of agricultural technologies or productivity. If we denote a generic cell by i, with $i \in d$, where d denotes a district, our regression model is:

$$\Delta y_{id} = \alpha_d + \beta \widehat{\Delta Cov_{id}} + \delta X_{id} + u_{id} \tag{3}$$

where Δy_{id} denotes the change in a given outcome between 2007 and 2012 and $\widehat{\Delta Cov_{id}}$ represents the change in the share of land covered by the mobile phone network over the same period, instrumented with the variable $\mathbb{1}\,(\text{Tower})$ from equation (2). X_{id} is the vector of cell characteristics discussed in Table 2 and α_d are district fixed effects.

The main coefficient of interest is β, which will be positive if mobile phones have a positive impact on technology adoption and agricultural productivity. Clearly, this coefficient subsumes different mechanisms linking mobile phone coverage with technology adoption and productivity. For example, the arrival of mobile phone coverage might promote local economic opportunities more generally, increasing local income and thus demand for agricultural products. Farmers might adopt new technologies to serve this increased demand.[34]

To make progress in the direction of isolating the role of information, we expand equation (3) to account for the share of population in the cell speaking a non-official language, hence with limited access to information about inputs and best agricultural practices provided by the KCC. We estimate the following augmented specification:

$$\Delta y_{id} = \alpha_d + \beta_1 \widehat{\Delta Cov_{id}} + \beta_2 \widehat{\Delta Cov_{id} \times NOLang_{id}} + \beta_3 NOLang_{id} + \delta X_{id} + u_{id} \tag{4}$$

where, compared to equation (3), we also include the share of population speaking a non-official language ($NOLang_{id}$) and its interaction with the change in mobile phone coverage.[35] The

Nunn and Puga (2012).

[34] Previous studies have also shown that, by reducing transaction costs on money transfers, mobile phones can facilitate risk sharing among farmers (Jack and Suri 2014; Blumenstock et al. 2016). This might, in turn, incentivize them to experiment with newer but riskier technologies. See Feder, Just, and Zilberman (1985) for a discussion of the role of farmers' risk-aversion in adoption models. This mechanism is unlikely to be at play in our setting given the lack of mobile-based money transfer technologies in rural India during the period under study.

[35] We instrument this latter term by interacting the share of population speaking a non-official language with the indicator variable for tower construction from equation (2). In section 3.6 we show that our results are

coefficient β_1 captures the effect of mobile coverage when the entire local population speaks an official language ($NOLang_{id} = 0$) and hence has full access to information about agricultural practices and inputs. The coefficient β_2 instead captures the differential impact of mobile phone coverage in cells with different shares of the population speaking a non-official language. Clearly, the sum of the two coefficients β_1 and β_2 identifies the effect of mobile coverage on outcomes in the absence of access to a phone-based service for agricultural advice ($NOLang_{id} = 1$).

3.3 The Effect of Access to Information on Farmers' Calls: By Topic of the Call

We start by documenting the effect of mobile phone coverage on farmers' calls for agricultural advice. In particular, we use the identification strategy described in section 3.2 to study farmers' access to information about specific technologies. Crucially for our purpose, the call-level data from KCC report the exact question asked by the farmer – as well as the answer provided by the agronomist. This allows us to distinguish between calls in which farmers seek advice regarding specific agricultural technologies such as new varieties of seeds, fertilizers, irrigation, or pesticides. Appendix A reports a detailed description of the keywords used to classify calls in different categories, as well as several examples.[36] Documenting the type of information acquired by farmers is important in order to trace a link between access to information and actual adoption of agricultural technologies, which we study in the next section.

In column (1) of Table 4 we estimate the effect of mobile phone coverage on the change in total number of calls to KCC between 2007 and 2012, as described by equation (3). The magnitude of the estimated coefficient indicates that cells with one standard deviation larger increase in mobile phone coverage experienced a 23 percent larger increase in total calls by farmers. Next, in column (2), we report the results of estimating equation (4) for the same outcome. We interpret the estimated coefficient β_1 as the combined effect of coverage and access to a phone-based service for agricultural advice on farmers' calls. Its magnitude suggests that a one standard deviation increase in coverage in cells where all farmers speak an official language increases the number of calls by 26.3 percent. The coefficient β_2, on the other hand, indicates that this effect is smaller the larger is the share of population speaking a non-official language. Indeed, in cells where the entire local population speaks non-official languages, the sum of the estimated coefficients β_1 and β_2 (0.828 - 0.716 = 0.112) implies that the increase in calls for a one standard deviation increase in mobile coverage is only 3.6 percent and not statistically different from zero.

Next, we study the effect of mobile phone coverage on farmers' calls regarding different technologies. We focus on the main technologies covered in the AIS, namely: seed varieties, fertilizers, irrigation, and pesticides. The results are reported in column (3) to (10). Odd columns refer to the average effect of mobile phone coverage, while even columns allow for the heterogeneous response across cells characterized by different shares of the population speaking a non-official language. The results indicate that, across all technologies, areas with higher increase in mobile

robust to include additional interaction terms of $\widehat{\Delta Cov_{id}}$ with other cell characteristics including measures of agricultural intensity, geographical isolation and local income.

[36] For example, we classify as calls about new seed varieties those where farmers ask advice on which seeds to use to improve yields for a given crop; those in which they ask information on how to use HYV seeds; and those in which they ask general advice on how to improve yields and the agronomist suggests to try specific HYV seed varieties.

phone coverage and where a large share of the population speaks the same language as KCC agricultural advisers experience larger increases in calls about agricultural technologies. On the other hand, the negative and statistically significant coefficients on the interaction terms indicate that access to mobile phone coverage has a limited effect on farmers' calls in areas where a significant share of local population speaks non-official languages.

Overall, the results reported in Table 4 are consistent with the existence of an underserved demand for information on farming techniques by Indian farmers. To the extent that the information provided by call centers for agricultural advice is accurate, we can think of farmers acquiring mobile phone coverage and having access to a phone-based service for agricultural advice as receiving a positive shock to their information set on farming techniques. This allows to study the effect of such shock on the actual adoption of the technologies farmers ask about, as well as on local agricultural productivity. We focus on these two outcomes in the following sections.

3.4 The Effect of Access to Information on Technology Adoption

In this section we study the effect of farmers' access to information via call centers for agricultural advice on technology adoption. We focus in particular on those technologies farmers ask about in their phone calls to KCC, namely seed varieties, fertilizers, irrigation and pesticides.

To study the effect of mobile phone coverage on adoption of a given technology we estimate equations (3) and (4) using as outcome variable $\Delta\left(\frac{Area^k}{Area}\right)_{id}$, which is the change in the share of land farmed with a given technology k (e.g. HYV seeds) in cell i located in district d. Changes in outcomes are calculated using the last 2 waves of the AIS, which were run in 2007 and 2012. Before presenting the results, let us discuss our measure of technology adoption. The data from the AIS reports information on land farmed with a given technology at the district-crop level. Thus, we compute the share of land farmed with a given agricultural technology k in a given cell i using the following neutral assignment rule:

$$\left(\frac{Area^k}{Area}\right)_{idt} = \sum_{c \in O_i} \left[\left(\frac{Area^k}{Area}\right)_{dct} \times \left(\frac{Area_{idc,t=2000}}{Area_{id,t=2000}}\right) \right] \tag{5}$$

The first element in the summation is the share of land farmed with technology k in district d among the land farmed with crop c. This variable captures the rate of technology adoption for a given crop in a given district and varies over time. The second element in the summation is the share of land farmed with crop c in cell i, which is observed at cell level in the FAO-GAEZ dataset and captures the initial allocation of land across crops in a given cell in the baseline year 2000.[37] Thus, the product of first and second element gives us an estimate of the share of land in cell i that is farmed under technology k and crop c. Summing across the set of crops farmed in cell i (O_i), we obtain an estimate of the share of land farmed with a given technology in a given cell.[38] In Appendix B we validate this measure by showing that it captures well technology

[37] The GAEZ dataset reports information on the amount of land – expressed in hectares – farmed with a specific crop in a given cell. The data refers to the baseline year 2000. We focus on the 10 major crops by area harvested in India, namely: rice, wheat, maize, soybean, cotton, groundnut, rape, millet, sugar and sorghum. According to FAOSTAT, the area harvested with these 10 crops amounts to 135.5 million hectares and accounts for 76 percent of the total area harvested in India in 2000.

[38] As an example, suppose that in district d, 20 percent of land farmed with rice and 50 percent of land farmed

adoption at the cell level using a small sample of cells for which we observe actual adoption of HYV seeds at the village level from publicly available surveys. Although our technology adoption outcomes are measured with error at the cell level, note that classical measurement error in the dependent variable does not generate bias in the estimated coefficients.

Column (1) of Table 5 reports the results of estimating equation (3) when the outcome variable is the change in the share of land farmed with HYV seeds – as opposed to traditional seeds – in a given cell. The coefficient is positive and precisely estimated. Its magnitude indicates that cells with a one standard deviation larger increase in mobile phone coverage experienced a 1.4 percentage points larger increase in the share of area farmed with HYV seeds. Among the cells in our sample, the average area farmed with HYV seeds in the baseline year 2007 was 26 percent. Thus, the 1.4 percentage point increase mentioned above corresponds to a 5.3 percent increase in land cultivated with HYV seeds for the average cell in our sample.

Column (2) reports the results of estimating equation (4), where we allow for an heterogeneous response to mobile phone coverage depending on the share of local population speaking non-official languages in an area. The estimated coefficient β_1 captures the combined effect of coverage and access to a phone-based service for agricultural advice. Its magnitude indicates that areas with full coverage and where all farmers speak official languages experienced a 4.7 percentage points larger increase in share of land farmed with HYV seeds between 2007 and 2012, compared to areas with no coverage (corresponding to 28 percent of the share at baseline). The negative and statistically significant coefficient on the interaction term β_2 indicates that limited access to information about agricultural practices reduces the impact of mobile phones on technology adoption. A one standard deviation difference in the share of local population speaking a non-official language leads to a reduction of 0.8 percentage points (-0.041×0.212) in the technology adoption differential between areas with full and no coverage. This corresponds to a reduction of 17 percent of the differential observed when the entire population have access to the KCC agricultural advice.

In columns (3) and (4) we focus on the share of land under chemical fertilizers as an additional measure of technology adoption. One important characteristic of HYV seeds is that they are highly respondent to fertilizers (Dalrymple, 1974). Thus, we expect adoption of HYV seeds by farmers to increase their demand for these complementary inputs of production. Column (4) shows that cells with larger increase in mobile phone coverage and no language barriers experienced an increase in area farmed with chemical fertilizers of similar magnitude as the increase documented for HYV seeds. The negative coefficient on the interaction term, although less precisely estimated compared to column (2), suggests that language barriers with agricultural advisers limit the impact of mobile phone coverage on adoption of fertilizers.

Next, we test for the effect of access to information on adoption of artificial irrigation. Farming with HYV seeds does not necessarily require more water than farming with traditional seeds. However, in order for HYV seeds to attain their full potential, they do require a reliable source of irrigation (Dalrymple, 1974). Thus, we expect adoption of HYV seeds by farmers to also increase their demand for irrigation. We study the effect on irrigated area in columns (5) and

with wheat are farmed using high-yielding variety seeds. Suppose also that 40 percent of land in cell i that is part of district d is farmed with rice, while the remaining 60 percent is farmed with wheat. Under our neutral assignment rule, we assign 38 percent of land in cell i to high-yielding varieties: $(0.2 \times 0.4) + (0.5 \times 0.6) = 0.38$.

(6), and find results that are similar, although smaller in magnitude, to the ones documented for chemical fertilizers.[39] Finally, columns (7) and (8) show a positive and significant effect of mobile coverage combined with access to a phone-based service for agricultural advice on the share of land under chemical pesticides.

Overall, the results presented in Tables 4 and 5 are consistent with a positive and significant effect of mobile phone coverage, coupled with access to a service for agricultural advice, on technology adoption via the diffusion of information about new technologies. We can use the estimates to calculate the implied elasticity of technology adoption to access to information about a given technology. To compute this elasticity we divide the estimated percentage increase in area farmed with a given technology by the estimated percentage increase in farmers' calls regarding that same technology for a given information shock. For HYV seeds, the obtained elasticity indicates that a 1 percent increase in mobile phone calls about this technology translates into a 0.78 percent increase in its actual adoption. Similarly, we find elasticities of 0.64 for chemical fertilizers, 1.1 for chemical pesticides and 3 for irrigation.

3.5 The Effect of Access to Information on Productivity

Finally, we study the effect of farmers' access to information via call centers for agricultural advice on agricultural productivity. Our measure of agricultural productivity is crop yield, which is defined as the quantity of crop produced (in metric tons) in a given area divided by the land farmed with that crop (in hectares) in the same area. We construct our measure of crop yield similarly to Jayachandran (2006) and Kaur (2019), who use a weighted average of normalized yields of the major crops farmed in India to generate a district-level measure of agricultural productivity. We compute yields using data from the Indian Ministry of Agriculture. This database collects information on the quantity produced and the area farmed with a given crop in a given district. Agricultural productivity at the cell level is then computed with a neutral assignment rule similar to the one reported in equation (5) as follows:

$$\log yield_{idt} = \sum_{c \in O_i} \left[\log \left(\frac{\text{quantity produced}}{\text{area farmed}} \right)_{dct} \times \left(\frac{Area_{idc,t=2000}}{Area_{id,t=2000}} \right) \right] \tag{6}$$

Equation (6) defines yield in cell i as the weighted average of log crop yields for the ten major crops by area farmed, where the weights are the share of area farmed with a given crop in a cell at baseline.[40]

The main results on the effects of access to information on productivity are reported in Table 6. In column (1) we estimate equation (3) and find a positive but not precisely estimated effect of mobile phone coverage on agricultural productivity. Column (2) reports the results of estimating equation (4), where we allow for an heterogeneous response to mobile phone coverage based on the share of local population speaking non-official languages. As shown, the combined effect

[39] The Agricultural Input Survey reports the use of fertilizers and irrigation by land farmed with HYV vs traditional seeds. In Table C.6 we estimate our main specifications splitting fertilizers and irrigation use in land farmed with HYV seeds and with traditional seeds. As shown, the effects of mobile coverage coupled with the availability of services for agricultural advice on fertilizers and irrigation are concentrated in areas farmed with HYV seeds. This is consistent with the complementarity between these inputs described above.

[40] As in Kaur (2019) we first normalize the yield for each of the 10 major crops in India by the mean yield of that crop in each district (using the years 1998 to 2012 to construct the mean).

of coverage and access to a phone-based service for agricultural advice – as captured by β_1 – is positive and statistically significant. Its magnitude indicates that areas with a one standard deviation increase in coverage and where all farmers speak official languages experienced a 1.3 percent larger increase in agricultural yields between 2007 and 2012. Language barriers between farmers and agricultural advisers can significantly hinder this effect. The magnitude of the estimated coefficients on the interaction term β_2 and the share of non-official language speakers indicate that the effect of mobile coverage on productivity is muted in areas where 50 percent or more of the population speak non-official languages.

We conclude this section by discussing the role of access to information in explaining productivity differences across regions implied by our estimates. Even within our sample of rural areas with no initial mobile phone coverage, there is large variation in the baseline level of agricultural productivity as measured by yields. In 2007, the average yield of a cell at the 75th percentile of agricultural productivity was almost twice as large as the one observed in a cell at the 25th percentile. This gap in yield is similar to the one documented in rice and wheat production between the top decile and the bottom decile of countries in the world income distribution (Gollin, Lagakos, and Waugh 2014). The estimates presented in column (2) of Table 6 indicate that providing full mobile phone coverage to the areas in our sample, coupled with the availability of a phone-based service for agricultural advice, can close around 25 percent of this productivity gap.

Note that this quantification uses the average effect obtained in Table 6 to compute the productivity gain of farmers in the 25th percentile of the initial productivity distribution. However, the effect of access to information might be heterogeneous across farmers with different initial productivity. We test for these heterogeneous effects in Table 7, where we estimate equation (4) separately by quartile of initial productivity. As shown, the effect of access to information on productivity is largest – and precisely estimated – for farmers with the lowest initial level of productivity. The point estimate on β_1 for this group is 0.058, around 40 percent larger than the average effect reported in Table 6. The effect is positive but small for farmers in the middle of the initial productivity distribution and large but extremely noisy for farmers in the top quartile. The estimate obtained for the lowest quartile indicates that providing access to information to farmers at the 25th percentile of the productivity distribution can close up to 36 percent of the productivity gap with farmers at the 75th percentile.

3.6 Additional Robustness Tests

The goal of our empirical analysis is to test whether the arrival of mobile phone coverage, coupled with the availability of phone-based services for agricultural advice, favored farmers' adoption of modern technologies and – thus – increased agricultural productivity via an information mechanism. The idea is that farmers might lack information about the very existence of a new technology, or how to use it productively. In our data, for example, farmers' questions suggest that they often do not know which new seed varieties better meet their specific needs, or what are the best practices to use them.

Mobile phones, however, promote access to information above and beyond the information on agricultural practices provided by KCC. Jensen (2007) and Aker (2010), for example, document that mobile phone diffusion reduced price dispersion in, respectively, fishing markets in the Indian state of Kerala and grain markets in Niger. Similarly, by allowing farmers to share information on crop prices in different markets, mobile phones could have favored a more efficient allocation of goods across markets in our sample and generate higher incomes for goods producers, potentially helping them pay the fixed cost of technology adoption. Although our data does not allow us to measure prices in local agricultural markets, we do observe the precise location of such markets. Thus, to gauge the importance of higher access to price information in our setting, we augment our main specification with agricultural market fixed effects. To this end, we collect data on the latitude and longitude of 3,255 agricultural markets in rural India from the AGMARKNET service of the Ministry of Agriculture of India. We assign each cell in our sample to its closest agricultural market based on minimum geographical distance within the same state. We think of the closest geographical market as capturing the market where produces farmed in each cell are sold. This gives us 1,017 agricultural markets covering the cells in our sample, each market serving on average six cells. Including agricultural market fixed effects to our specification allows us to compare outcomes across farmers who are heterogeneously exposed to the increase in mobile phone coverage and access to services for agricultural advice, but that plausibly serve the same local market, and thus face the same prices for their products.

The results of this augmented specification are reported in Table C.7. In terms of outcomes, we focus on the adoption of the four agricultural technologies studied in section 3.4, and on agricultural productivity. As shown, all our main results are robust to including agricultural market fixed effects. In particular, the point estimates on the coefficient β_1, which captures the combined effect of mobile coverage and availability of phone-based services for agricultural advice, are similar in magnitude to those presented in Tables 5 and 6. This suggests a limited effect of access to price information on technology adoption and productivity in our sample. The negative coefficient on the interaction term β_2 is also similar in magnitude compared to our baseline estimates, although less precisely estimated.

Another potential concern with our main estimates is that the share of local population speaking non-official languages is not randomly assigned across geographical areas. Our empirical model interprets the differential impact of mobile phone coverage in areas with different diffusion of official languages as the effect of language barriers between farmers and KCC agricultural advisers. However, a potential concern with this interpretation is that areas with a greater share of the local population speaking non-official languages might also be characterized by different levels of agricultural intensity, might be more geographically isolated or simply be poorer. In this case one would load on the interaction between local languages and mobile phone coverage also variations in local economic conditions. In section 3.2 we showed that, among treated cells in our sample, the distribution of non-official language speakers is uncorrelated with observable characteristics and pre-trends in technology adoption and productivity. In this section we bring this analysis one step further and augment our model with a set of additional interaction terms. Table C.8 presents estimates of the parameters of the model where we include, in addition to the baseline interaction of mobile phone coverage with the share of population speaking non-official languages (column 1), also the interaction of coverage with measures of: agricultural intensity (share of irrigated land and of population employed in agriculture, column 2); geographical isolation (terrain ruggedness and distance from closest city, column 3); cell income (average income per capita and night lights intensity, column 4); as well as a fully saturated specification that includes all the previous dimensions interacted with mobile phone coverage (column 5). The inclusion of these additional interaction terms makes virtually no difference to our results, irrespective of the outcome considered. If anything, estimates of the parameters of interest become slightly larger compared to our baseline specification.

Finally, we show that all our results are robust to using the share of land covered by SMIP towers instead of the indicator variable used in the main analysis. Specifically, we estimate equation (4) by instrumenting the change in mobile phone coverage using the intensive margin of land covered by a SMIP tower in a given cell. Table C.9 presents the results from this analysis. Column (1) shows that mobile phone coverage is 7.5 percent higher in cells with a 50 percent higher share of land covered by a SMIP tower relative to the control group. Column (2) to (6) confirm that an increase in mobile phone coverage leads to higher number of calls to KCC. Similarly, the positive and significant coefficient β_1 in column (7) to (11) is consistent with the results from Tables 5 and 6. Finally, the negative coefficient on the interaction term β_2 confirm that the effects are mitigated in areas where the majority of the population does not have access to the agricultural advice provided by the KCC.

4 Concluding Remarks

Mobile phones have experienced a widespread and fast diffusion in developing countries over the last 20 years. The benefits – as well as the costs – of this diffusion are still to be understood, especially in previously unconnected areas, such as rural areas of developing countries. In this paper we exploit the diffusion of the mobile phone network in rural India to study the role of information in fostering technology adoption and productivity in the agricultural sector. India represents a natural setting to investigate this question. The country is characterized by a large share of the population employed in agriculture and by large differences in agricultural productivity across regions.

We argue that farmers receiving mobile phone coverage and without language barriers with the agricultural advisers answering their calls are the ones receiving a positive information shock on agricultural practices and inputs. We present evidence based on geo-located call data to support this argument. Our findings indicate that access to information has significant real effects. Farmers able to receive information from agricultural advisers via mobile phones experienced faster adoption of modern agricultural technologies, such as high-yielding varieties of seeds, and of complementary inputs of production, such as chemical fertilizers, pesticides and irrigation, as well as faster increase in agricultural productivity.

Our findings indicate that lack of access to information on agricultural practices and inputs can explain a significant fraction of the agricultural productivity gap across Indian regions. In particular, our estimates indicate that providing full mobile phone coverage to the areas in our sample, coupled with the availability of a phone-based service for agricultural advice, can close around 25 percent of the initial gap in agricultural yields between regions in the 25th and the 75th percentile. We also show that the effect of access to information is heterogeneous across farmers with different initial productivity, and it is the largest for farmers with the lowest initial level of productivity.

Although nowadays the mobile phone network covers almost the entirety of India, advancements have been made in recent years towards the expansion of 3G/4G mobile services and universal availability of broadband Internet. These ICT enhancements have been contemporaneously met with the rise of social media, online information-sharing websites and smart-phone applications. These digital platforms can further help the diffusion of information among farmers. We leave the question of how advancements in digital ICT foster technological adoption for future research.

References

Adamopoulos, T. and D. Restuccia (2014). The Size Distribution of Farms and International Productivity Differences. *American Economic Review 104*(6), 1667–97.

Aghion, P. and P. Howitt (1992). A Model of Growth through Creative Destruction. *Econometrica 60*, 323–51.

Aker, J. C. (2010). Information from Markets Near and Far: Mobile Phones and Agricultural Markets in Niger. *American Economic Journal: Applied Economics 2*(3), 46–59.

Aker, J. C., I. Ghosh, and J. Burrell (2016). The Promise (and Pitfalls) of ICT for Agriculture Initiatives. *Agricultural Economics 47*(S1), 35–48.

Aker, J. C. and I. M. Mbiti (2010). Mobile Phones and Economic Development in Africa. *Journal of Economic Perspectives 24*(3), 207–32.

Anderson, J. and R. Birner (2007). How to make agricultural extension demand-driven? : The case of India's agricultural extension policy. *IFPRI Discussion Paper 00729.*.

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

Asher, S. and P. Novosad (2020). Rural Roads and Local Economic Development. *American Economic Review 110*(3), 797–823.

Atkin, D., A. Chaudhry, S. Chaudry, A. K. Khandelwal, and E. Verhoogen (2017). Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan. *The Quarterly Journal of Economics 132*(3), 1101–1164.

Beaman, L., A. BenYishay, J. Magruder, and A. M. Mobarak (2018). Can Network Theory-based Targeting Increase Technology Adoption? Technical report, National Bureau of Economic Research.

Bloom, N. and J. Van Reenen (2007). Measuring and Explaining Management Practices Across Firms and Countries. *The Quarterly Journal of Economics 122*(4), 1351–1408.

Blumenstock, J. E., N. Eagle, and M. Fafchamps (2016). Airtime Transfers and Mobile Communications: Evidence in the Aftermath of Natural Disasters. *Journal of Development Economics 120*, 157–181.

Bold, T., K. C. Kaizzi, J. Svensson, and D. Yanagizawa-Drott (2017). Lemon Technologies and Adoption: Measurement, Theory and Evidence from Agricultural Markets in Uganda. *The Quarterly Journal of Economics 132*(3), 1055–1100.

Buys, P., S. Dasgupta, T. S. Thomas, and D. Wheeler (2009). Determinants of a Digital Divide in Sub-Saharan Africa: A Spatial Econometric Analysis of Cell Phone Coverage. *World Development 37*(9), 1494–1505.

Casaburi, L., M. Kremer, S. Mullainathan, and R. Ramrattan (2014). Harnessing ICT to Increase Agricultural Production: Evidence from Kenya. *Harvard University*.

Caselli, F. (2005). Accounting for Cross-country Income Differences. *Handbook of Economic Growth 1*, 679–741.

Caselli, F. and W. J. Coleman (2001). Cross-country Technology Diffusion: The Case of Computers. *American Economic Review 91*(2), 328–335.

Cole, S. A. and A. Fernando (2016). Mobile-izing Agricultural Advice: Technology Adoption, Diffusion and Sustainability.

Comin, D. and B. Hobijn (2004). Cross-country Technology Adoption: Making the Theories Face the Facts. *Journal of Monetary Economics 51*(1), 39–83.

Comin, D. and B. Hobijn (2010). An Exploration of Technology Diffusion. *American Economic Review 100*(5).

Comin, D. and R. Nanda (2019). Financial Development and Technology Diffusion. *IMF Economic Review*.

Conley, T. G. and C. R. Udry (2010). Learning about a New Technology: Pineapple in Ghana. *American Economic Review 100*(1), 35–69.

Dalrymple, D. G. (1974). Development and Spread of High-yielding Varieties of Wheat and Rice in the Less Developed Nations. Technical report, United States Department of Agriculture, Economic Research Service.

De Janvry, A., E. Sadoulet, D. Manzoor, and E. Kyle (2016). The Agricultural Technology Adoption Puzzle: What Can We Learn From Field Experiments? Technical report.

Donovan, K. (2016). Agricultural Risk, Intermediate Inputs, and Cross-country Productivity Differences. *Working Paper, University of Notre Dame*.

Duflo, E., M. Kremer, and J. Robinson (2004). Understanding Technology Adoption: Fertilizer in Western Kenya, Preliminary Results from Field Experiments. *Unpublished manuscript, Massachusetts Institute of Technology*.

Duflo, E., M. Kremer, and J. Robinson (2008). Why Don't Farmers Use Fertilizers? Experimental Evidence from Kenya. *Mimeo, Massachusetts Institute of Technology*.

Ericsson (2015). *The Changing Mobile Broadband*.

Evenson, R. E. and D. Gollin (2002). *Crop Variety Improvement and its Effect on Productivity*. CABI Pub.

Evenson, R. E. and D. Gollin (2003). Assessing the Impact of the Green Revolution, 1960 to 2000. *Science 300*(5620), 758–762.

Fafchamps, M. and B. Minten (2012). Impact of SMS-based Agricultural Information on Indian Farmers. *The World Bank Economic Review 26*(3), 383–414.

Feder, G., R. E. Just, and D. Zilberman (1985). Adoption of Agricultural Innovations in Developing Countries: A Survey. *Economic Development and Cultural Change 33*(2), 255–298.

Foster, A. D. and M. R. Rosenzweig (1995). Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture. *Journal of Political Economy 103*(6), 1176–1209.

Foster, A. D. and M. R. Rosenzweig (2010). Microeconomics of Technology Adoption. *Annual Review of Economics 2*(1), 395–424.

Gollin, D., D. Lagakos, and M. E. Waugh (2014). The Agricultural Productivity Gap. *The Quarterly Journal of Economics 129*(2), 939–993.

Greenstone, M., R. Hornbeck, and E. Moretti (2010). Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. *Journal of Political Economy 118*(3), 536–598.

GSMA (2012). GSMA Network Coverage Maps - Submission Guide January 2012. *London: GSMA*.

Hanna, R., S. Mullainathan, and J. Schwartzstein (2014). Learning Through Noticing: Theory and Evidence from a Field Experiment. *The Quarterly Journal of Economics 129*(3), 1311–1353.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics 124*(4), 1403–1448.

Jack, B. K. (2013). Constraints on the adoption of agricultural technologies in developing countries. *Literature review, Agricultural Technology Adoption Initiative, J-PAL (MIT) and CEGA (UC Berkeley).*

Jack, W. and T. Suri (2014). Risk Sharing and Transactions Costs: Evidence from Kenya's mobile money revolution. *American Economic Review 104*(1), 183–223.

Jayachandran, S. (2006). Selling Labor Low: Wage Responses to Productivity Shocks in Developing Countries. *Journal of Political Economy 114*(3), 538–575.

Jensen, R. (2007). The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector. *The Quarterly Journal of Economics 122*(3), 879–924.

Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry (2014). Agricultural Decisions After Relaxing Credit and Risk Constraints. *The Quarterly Journal of Economics 129*(2), 597–652.

Kaur, S. (2019). Nominal Wage Rigidity in Village Labor Markets. *American Economic Review 109*(10).

Lagakos, D. and M. E. Waugh (2013). Selection, Agriculture, and Cross-country Productivity Differences. *American Economic Review 103*(2), 948–80.

Manacorda, M. and A. Tesei (2020). Liberation Technology: Mobile Phones and Political Mobilization in Africa. *Econometrica 88*(2), 533–567.

Munshi, K. (2004). Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution. *Journal of Development Economics 73*(1), 185–213.

Nakasone, E., M. Torero, and B. Minten (2014). The Power of Information: The ICT Revolution in Agricultural Development. *Annual Review of Resource Economics 6*(1), 533–550.

National Sample Survey (2005). Report No. 499: Situation Assessment Survey of Farmers: Access to Modern Technology For Farming. Technical report, Government of India. Ministry of Statistics and Programme Implementation.

Nunn, N. and D. Puga (2012). Ruggedness: The Blessing of Bad Geography in Africa. *The Review of Economics and Statistics 94*(1), 20–36.

Restuccia, D. and R. Rogerson (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Establishments. *Review of Economic Dynamics 4*(11), 707–720.

Restuccia, D., D. T. Yang, and X. Zhu (2008). Agriculture and Aggregate Productivity: A Quantitative Cross-country Analysis. *Journal of Monetary Economics 55*(2), 234–250.

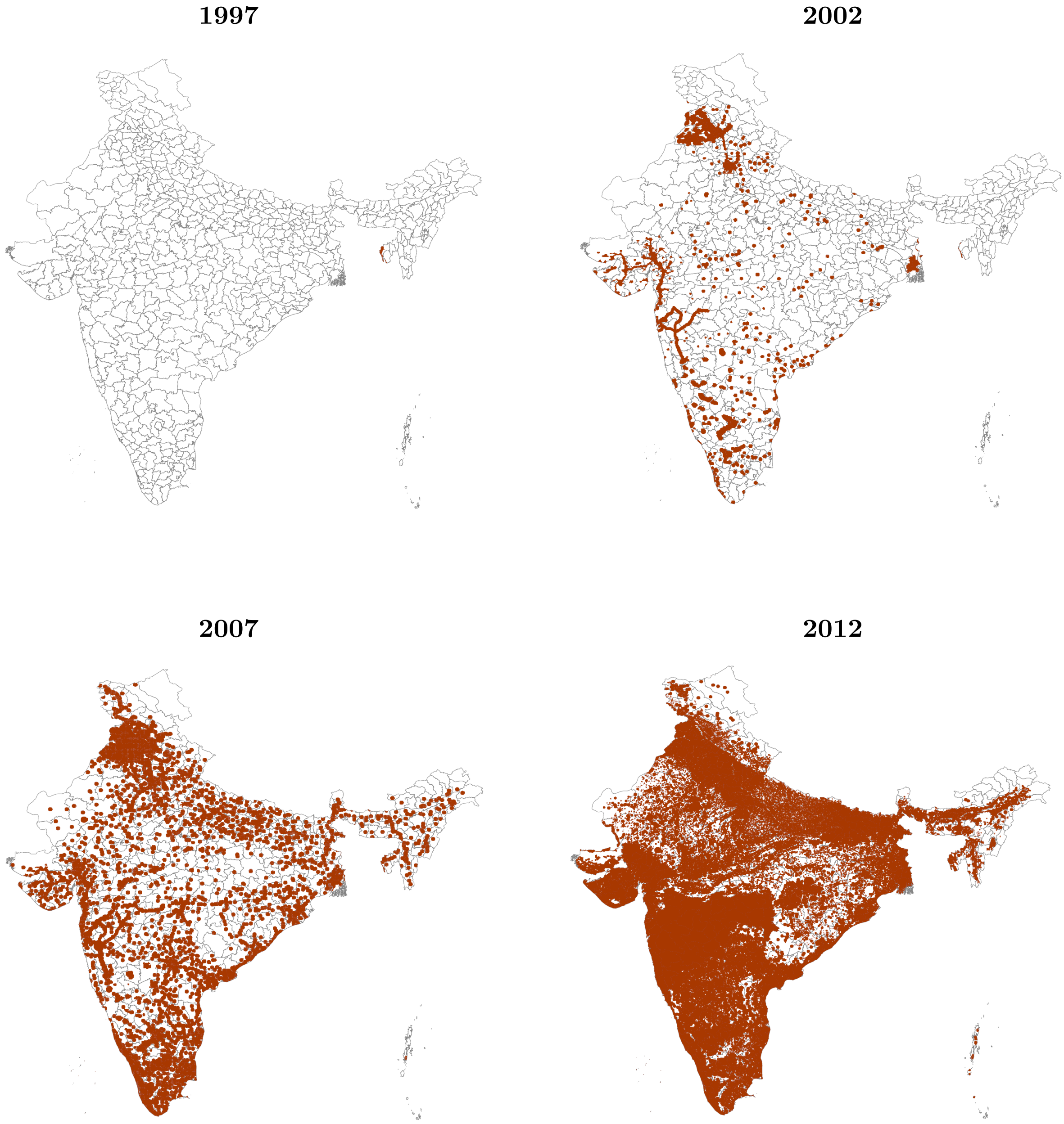
Sauter, M. (2006). Communication Systems for the Mobile Information Society. *John Wiley & Sons*.

Suri, T. (2011). Selection and Comparative Advantage in Technology Adoption. *Econometrica 79*(1), 159–209.

World Bank (2014). *World Development Indicators 2014*. World Bank Publications.

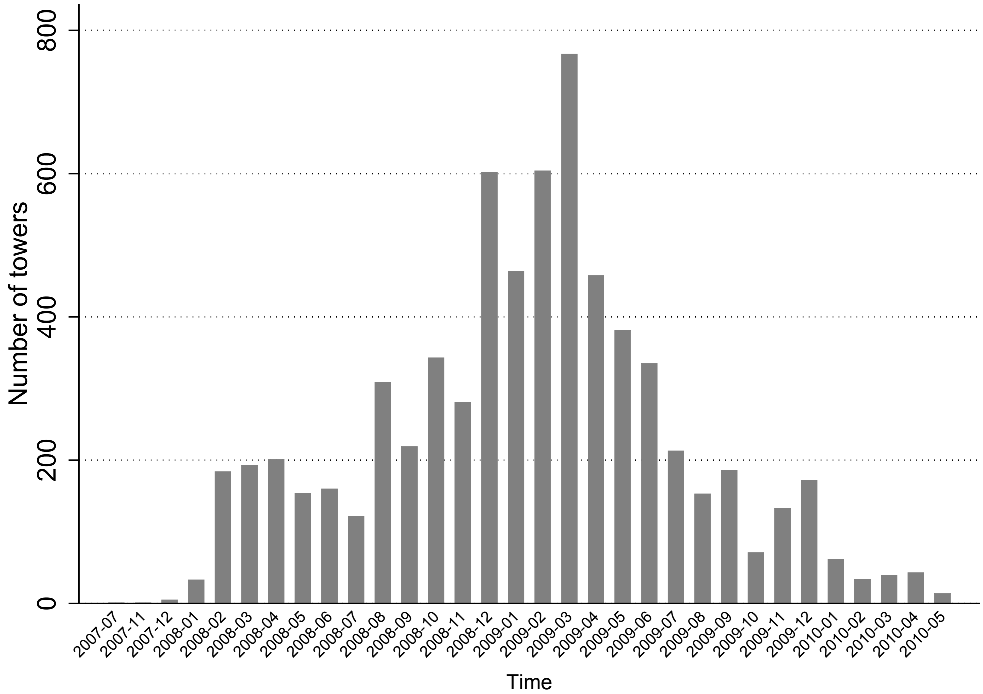
Figures and Tables

Figure 1: Mobile Phone Coverage Evolution, India 1997-2012



Notes: The figure reports geo-referenced data on mobile phone coverage for all of India at five-year intervals between 1997 and 2012. Source: GSMA.

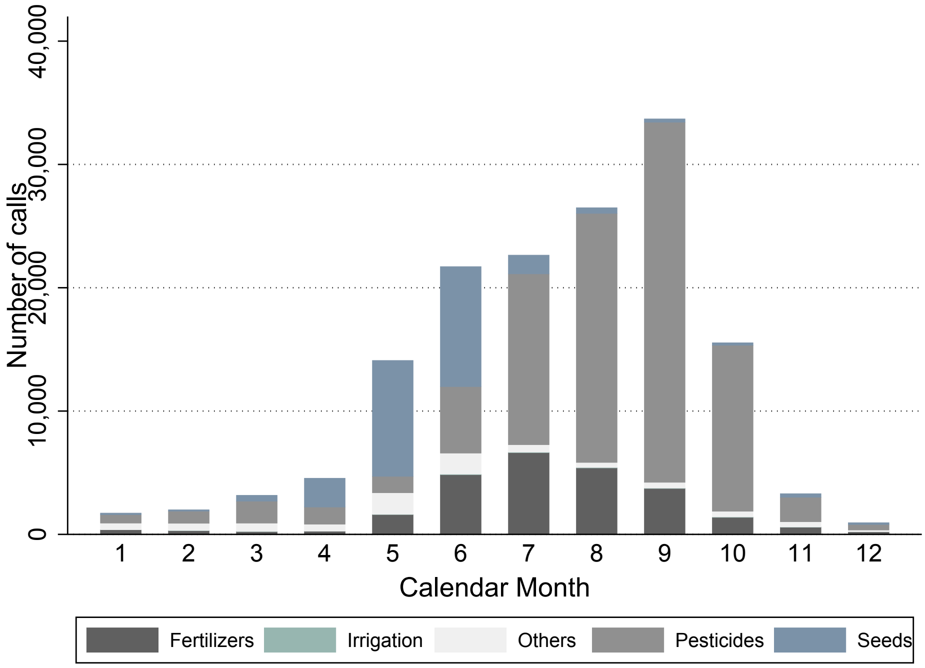
Figure 2: Timeline of tower construction under SMIP Phase I



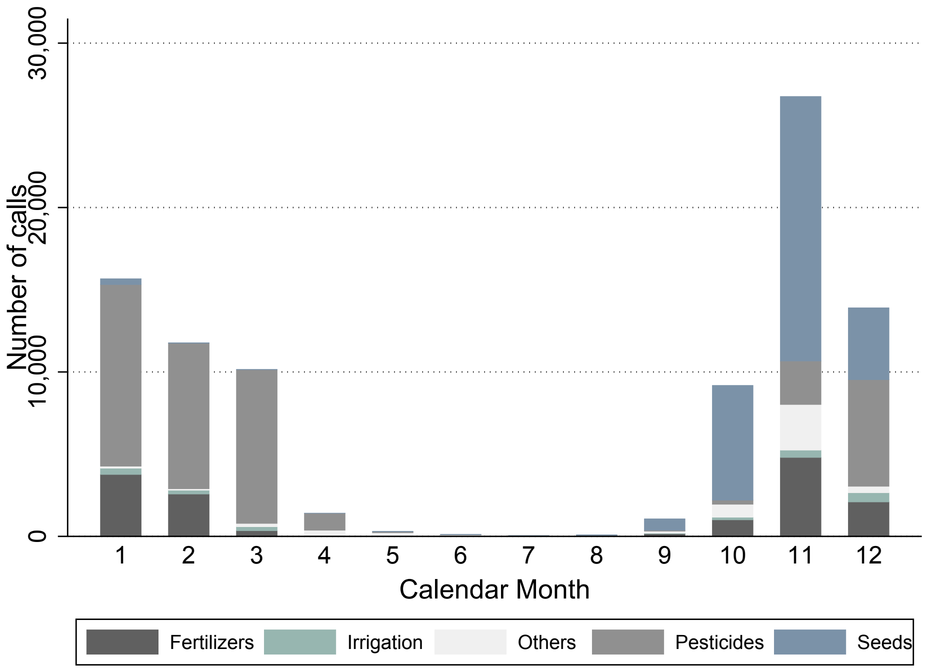
Notes: Source: Department of Telecommunications, India. Month captures the time at which the construction of the tower is completed and the tower becomes operational.

Figure 3: Distribution of calls on rice and wheat across agricultural cycle

(a) Rice



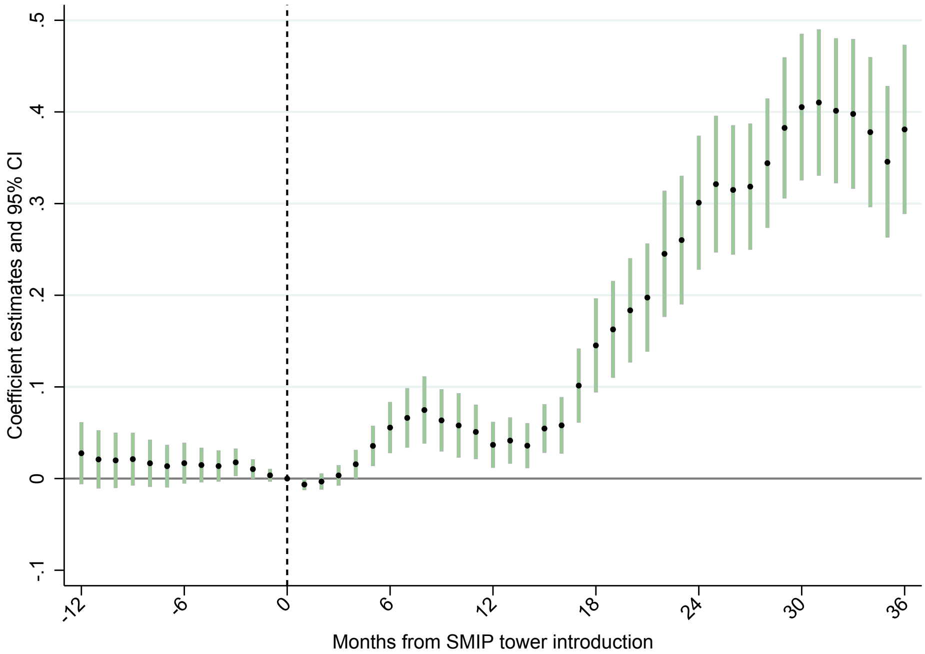
(b) Wheat



Notes: Source: Kisan Call Center, Ministry of Agriculture

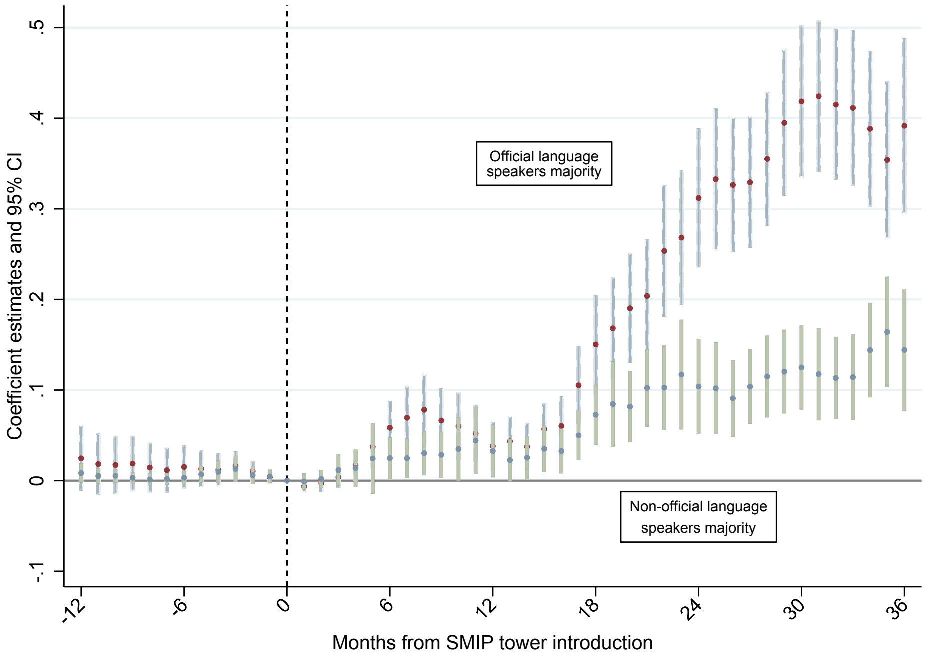
Figure 4: Farmers' Calls to KCC relative to Tower Construction - Event Study

(a) Average effects



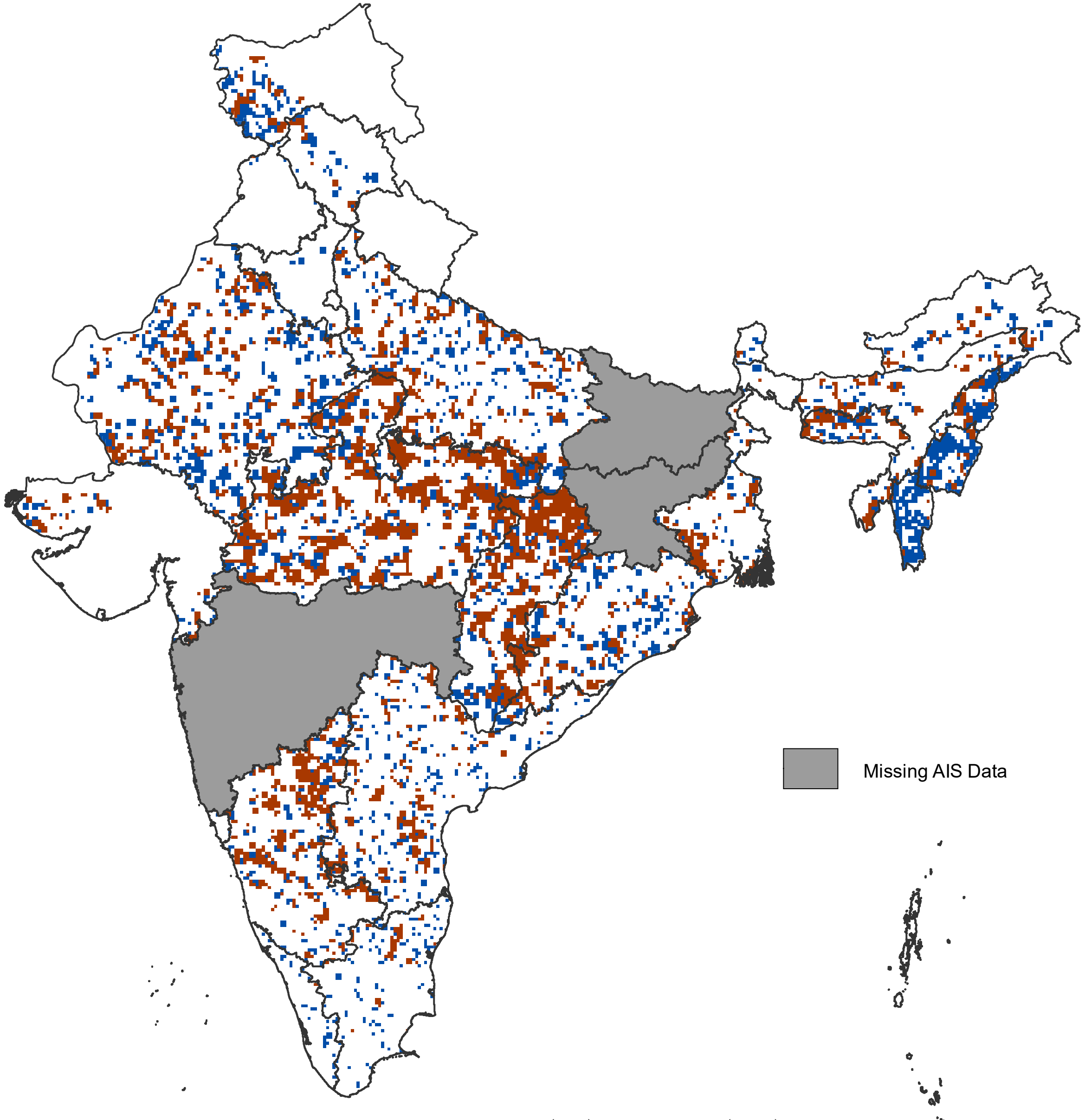
Notes: The figure plots the coefficients obtained with the following specification $Ln\,(1 + \text{calls})_{it} = \alpha_i + \alpha_t + \sum_{k=-12}^{+36} \beta_k D_{it}^k + \varepsilon_{it}$. Where i cell, t month, D_{it}^k dummy equal to 1 if month $t = k$ for cell i

(b) Heterogeneous Effects by Language



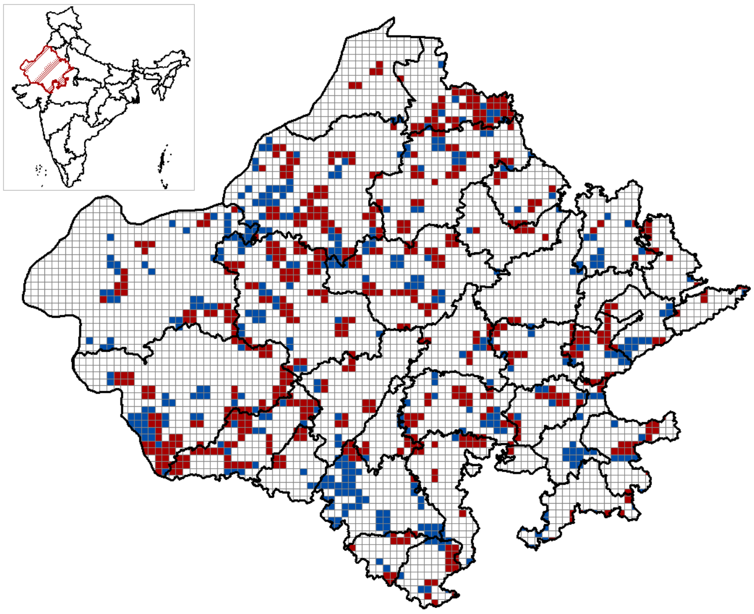
Notes: The figure plots the coefficients obtained with the following specification $Ln\,(1 + \text{calls})_{it} = \alpha_i + \alpha_t + \sum_{k=-12}^{+36} \beta_k D_{it}^k + \varepsilon_{it}$. Where i cell, t month, D_{it}^k dummy equal to 1 if month $t = k$ for cell i. We estimate this specification separately for two groups of cells based on the share of population speaking non-official languages.

Figure 5: Treatment and Control cells under SMIP



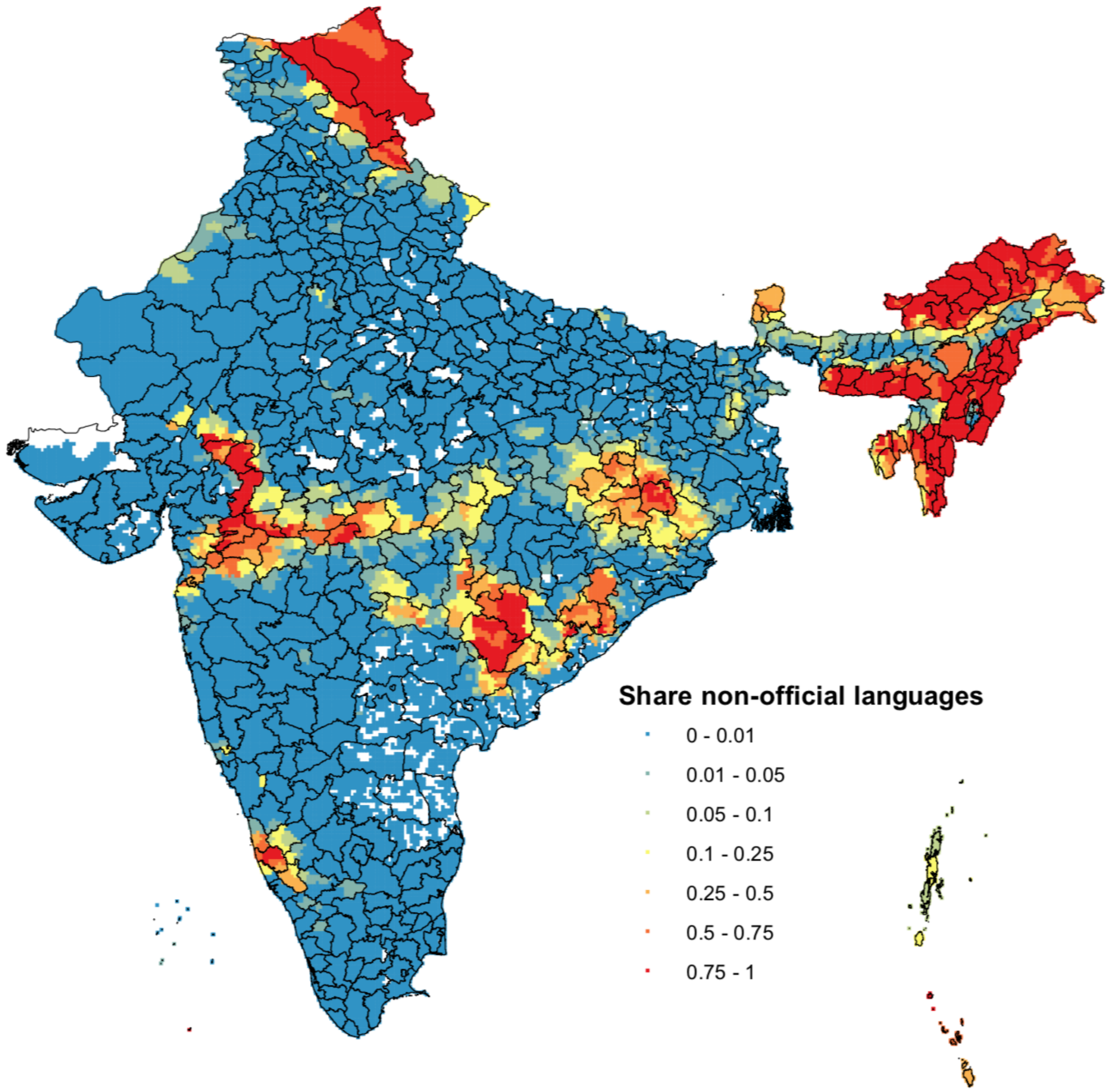
Notes: The figure shows the 6,320 10×10 km identification cells distributed across treatment (red) and control (blue) cells for all of India. State borders are marked in black. Treatment cells are those that are both proposed *and* covered by mobile tower under SMIP Phase I. Control cells are those that are proposed *and not* covered by mobile towers under SMIP Phase I. Grey areas represent states with missing AIS information.

Figure 6: Treatment and Control Cells
(Rajasthan State)



Notes: 10×10 Km treatment (red) and control (blue) cells for the state of Rajasthan. District boundaries are labeled in black. Treatment cells are those that are both proposed *and* covered by mobile tower under SMIP Phase I. Control cells are those that are proposed *and not* covered by mobile tower under SMIP Phase I.

Figure 7: Share of non-official languages in India



Notes: Share of non-official languages is the share of population speaking non-official languages in a given sub-district. Source: Population Census of India (2011).

Table 1: Summary Statistics

	Mean	**Median**	**Std. Deviation**	**N**
Δ Coverage	0.756	0.926	0.321	6320
Non-official Languages (%)	0.075	0	0.212	6320
Δ HYV Share	0.034	0.018	0.068	6320
Δ Fertilizer Share	0.022	0.023	0.081	6310
Δ Irrigation Share	0.017	0.013	0.043	6320
Δ Pesticides Share	0.025	0.018	0.108	6142
Δ log(yield)	0.058	0.055	0.069	5033
Δ log $(1+\text{calls}_{\text{All}})$	1.294	1.167	0.918	6320
Δ log $(1+\text{calls}_{\text{Yield}})$	0.461	0.222	0.517	6320
Δ log $(1+\text{calls}_{\text{Fertilizers}})$	0.374	0.193	0.436	6320
Δ log $(1+\text{calls}_{\text{Irrigation}})$	0.093	0.034	0.13	6320
Δ log $(1+\text{calls}_{\text{Pesticides}})$	0.948	0.763	0.777	6320

Notes: Changes in variables are calculated over the 5-year interval 2007-2012. The unit of observation is a 10×10 km cell and the sample includes all cells used for identification. Only cells with non-missing Δ HYV values are considered.

Table 2: SMIP coverage ($\mathbb{1}$ (Tower)) and cell characteristics (Balance Test)

Dependent variable:	1(**Tower**)				1(non-off. lang.| **Tower**)
	(1)	(2)	(3)	(4)	(5)
Determinants of Tower Relocation					
log(Population)	0.097***			0.097***	0.014
	(0.021)			(0.026)	(0.024)
Power Supply	0.019			0.010	-0.059
	(0.038)			(0.049)	(0.052)
Ruggedness	-0.080***			-0.093***	0.030
	(0.018)			(0.023)	(0.024)
Pre-trends technology/productivity					
Δ log(yield) (2002-2007)		-0.034		0.090	0.199
		(0.456)		(0.435)	(0.200)
Δ HYV Share (2002-2007)		0.091		-0.143	-0.355
		(0.455)		(0.438)	(0.300)
Socio-economic characteristics					
Agri. Workers/Working Pop.			0.078	0.109	-0.034
			(0.076)	(0.087)	(0.040)
Percent Irrigated			0.060	0.047	-0.031*
			(0.043)	(0.046)	(0.018)
Education Facility			0.071	-0.053	0.020
			(0.057)	(0.059)	(0.022)
Medical Facility			0.026	0.025	-0.003
			(0.032)	(0.038)	(0.016)
Banking Facility			-0.032	-0.068	-0.013
			(0.061)	(0.062)	(0.016)
# Phone conn. per 1000 people			0.002	0.003*	-0.001
			(0.001)	(0.002)	(0.001)
Dist. to nearest town(kms)			-0.001***	-0.001	0.000
			(0.000)	(0.001)	(0.000)
Night Lights (2006)			-0.003	-0.012	-0.000
			(0.006)	(0.007)	(0.002)
Income per capita			0.000	0.000	0.000
			(0.000)	(0.000)	(0.000)
Expense per capita			-0.000	-0.000	-0.000
			(0.000)	(0.000)	(0.000)
District f.e.	✓	✓	✓	✓	✓
Observations	6,320	5,019	6,320	5,019	3,570
R-squared	0.193	0.174	0.182	0.192	0.706

Notes: The table reports the correlation of cell-characteristics across treatment and control cells (columns 1-4) and across cells with and without a majority of non-official language speakers, conditional on treatment (column 5). The treatment variable $\mathbb{1}$ (Tower) in columns (1)-(4) is a dummy variable that takes the value of 1 if a cell is both proposed *and* covered by a tower (Treatment) under SMIP Phase I and takes the value of 0 if a cell is proposed *and not* covered (Control). The dependent variable in column (5) that takes the value of 1 if the share of non-official language speakers is greater than 50% of the total population in the cell, and 0 otherwise. Column (1) focuses on the main determinants of tower relocation, *i.e.* cell's population, the availability of power supply and average ruggedness; column (2) on pre-trends in technology/productivity; column (3) on socio-economic characteristics; columns (4) and (5) consider simultaneously all observable cell characteristics. All specifications include district fixed effects. The sample includes all cells with zero cell phone coverage in 2006. Standard errors clustered at district level are reported in brackets (number of clusters = 285). Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table 3: First Stage

Outcome:	Δ Coverage		
	(1)	(2)	(3)
$\mathbb{1}$ (Tower)	0.110***	0.073***	0.074***
	[0.015]	[0.012]	[0.012]
log(Population)		0.118***	0.074***
		[0.014]	[0.013]
Power Supply		0.254***	0.164***
		[0.028]	[0.029]
Ruggedness		-0.168***	-0.139***
		[0.019]	[0.018]
Observations	6,320	6,320	6,320
F-stat	56.54	34.24	36.72
District f.e.	✓	✓	✓
Other Controls			✓

Notes: The table reports first-stage regression of Δ Coverage on treatment variable $\mathbb{1}$ (Tower). The unit of observation is a 10×10 *km* cell. Δ Coverage is the change in the share of cell area under mobile coverage from 2007 to 2012, based on the data provided by telecom companies to GSMA. $\mathbb{1}$ (Tower) is a dummy variable that takes the value of 1 if a cell is both proposed *and* covered by a tower under SMIP Phase I and takes the value of 0 if a cell is proposed *and not* covered. All specifications control for district fixed effect. Column (1) reports estimates of regression of Δ Coverage on treatment variable. Column (2) includes baseline controls of cell's (log) population, the availability of power supply and average ruggedness. Column (3) includes other controls for the cell including share of labor force employed in agricultural sector, share of agricultural land that is irrigated, access to an educational facility, access to a medical facility, access to a banking facility, number of landline phone connections per 1000 people, distance to nearest town (in *kms.*), night lights intensity, income per capita (in rupees), and expense per capita (in rupees). The sample includes all cells with zero cell phone coverage in 2006. All regressions are weighted by the cell's population. The value of the first stage Kleibergen-Paap Wald F-statistics for the validity of the instruments is also reported in all columns. Standard errors clustered at district level are reported in brackets (number of clusters = 285). Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4: Mobile Coverage and Farmers' Calls

Outcome:	Δ log (1+ number of calls)									
Topic of the calls:	All		Seeds		Fertilizer		Irrigation		Pesticides	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Δ Coverage	0.742***	0.828***	0.322***	0.357***	0.269***	0.304***	0.059**	0.071**	0.656***	0.731***
	[0.199]	[0.206]	[0.113]	[0.119]	[0.099]	[0.104]	[0.028]	[0.030]	[0.170]	[0.175]
Δ Coverage × Non-official Languages (%)		-0.716**		-0.300***		-0.296***		-0.099***		-0.619**
		[0.316]		[0.107]		[0.103]		[0.032]		[0.261]
Non-official Languages (%)		-0.185*		-0.061**		-0.047		-0.025*		-0.169**
		[0.096]		[0.030]		[0.030]		[0.013]		[0.084]
Observations	6,320	6,320	6,320	6,320	6,320	6,320	6,320	6,320	6,320	6,320
R-squared	0.901	0.901	0.923	0.922	0.916	0.916	0.891	0.890	0.907	0.907
District f.e.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Other Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The table reports IV-2SLS estimates of the effect of mobile phone coverage on change in (log) calls received at Kisan Call Centers (KCC). The dependent variable in Columns (1)-(2) is change in all calls received at KCC; Columns (3)-(4) is change in calls about seeds; Columns (5)-(6) is change in calls about fertilizers; Columns (7)-(8) is change in calls about irrigation; Columns (9)-(10) is change in calls about pesticides. All changes are calculated between 2007-2012. The unit of observation is a 10×10 *km* cell. Δ Coverage is the change in the share of cell area covered under GSM mobile coverage between 2007-2012 instrumented using $\mathbb{1}$(Tower). $\mathbb{1}$(Tower) is a dummy variable that takes the value of 1 if a cell is both proposed *and* covered by a tower under SMIP Phase I and takes the value of 0 if a cell is proposed *and not* covered. Odd columns reports the average effect, even columns report the heterogenous effects depending on share of cell's population speaking non-official languages. All columns include district-fixed effects, baseline controls as well as other controls. Baseline controls include cell's (log) population, the availability of power supply and average ruggedness. Other controls for the cell include share of labor force employed in agricultural sector, share of agricultural land that is irrigated, access to an educational facility, access to a medical facility, access to a banking facility, number of landline phone connections per 1000 people, distance to nearest town (in *kms.*), night lights intensity, income per capita (in rupees), and expense per capita (in rupees). The sample includes all cells with zero cell phone coverage in 2006. All regressions are weighted by the cell's population. Standard errors clustered at district level are reported in brackets (number of clusters = 285). Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5: Mobile Coverage and technology adoption

Outcome:	Δ Technology Adoption							
Technology:	HYV Seeds		Fertilizers		Irrigation		Pesticides	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Coverage	0.043**	0.047**	0.037	0.040*	0.023*	0.027*	0.062**	0.068**
	[0.018]	[0.019]	[0.023]	[0.023]	[0.014]	[0.015]	[0.029]	[0.029]
Δ Coverage × Non-official Languages (%)		-0.041**		-0.022		-0.027		-0.048
		[0.019]		[0.031]		[0.017]		[0.037]
Non-official Languages (%)		-0.002		-0.013		-0.006		-0.013
		[0.009]		[0.017]		[0.007]		[0.013]
Observations	6,320	6,320	6,310	6,310	6,320	6,320	6,142	6,142
R-squared	0.856	0.856	0.885	0.885	0.809	0.808	0.883	0.883
District f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓	✓	✓	✓
Other Controls	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The table reports IV-2SLS estimates of the effect of mobile phone coverage on changes in technology adoption between 2007-2012. The dependent variable in Columns (1)-(2) is change in share of area cultivated under HYV; Columns (3)-(4) is change in share of area cultivated under fertilizers; Columns (5)-(6) is change in share of area cultivated under irrigation; Columns (7)-(8) is change in share of area cultivated under pesticides. All changes are calculated between 2007-2012. The unit of observation is a 10×10 *km* cell. Δ Coverage is the change in the share of cell area covered under GSM mobile coverage between 2007-2012 instrumented using $\mathbb{1}$ (Tower). $\mathbb{1}$ (Tower) is a dummy variable that takes the value of 1 if a cell is both proposed *and* covered by a tower under SMIP Phase I and takes the value of 0 if a cell is proposed *and not* covered. Odd columns reports the average effect, even columns report the heterogenous effects depending on share of cell's population speaking non-official languages. All columns include district-fixed effects, baseline controls as well as other controls. Baseline controls include cell's (log) population, the availability of power supply and average ruggedness. Other controls for the cell include share of labor force employed in agricultural sector, share of agricultural land that is irrigated, access to an educational facility, access to a medical facility, access to a banking facility, number of landline phone connections per 1000 people, distance to nearest town (in *kms.*), night lights intensity, income per capita (in rupees), and expense per capita (in rupees). The sample includes all cells with zero cell phone coverage in 2006. All regressions are weighted by the cell's population. Standard errors clustered at district level are reported in brackets (number of clusters = 285). Significance level: *** p<0.01, ** p<0.05, * p<0.1 .

Table 6: Mobile Coverage and agricultural productivity

Outcome:	Δ log(yield)	
	(1)	(2)
Δ Coverage	0.029	0.041**
	[0.020]	[0.020]
Δ Coverage $\times$ Non-official Languages (%)		-0.093***
		[0.033]
Non-official Languages (%)		-0.014
		[0.012]
Observations	5,033	5,033
R-squared	0.904	0.901
District f.e.	✓	✓
Baseline Controls	✓	✓
Other Controls	✓	✓

Notes: The table reports IV-2SLS estimates of the effect of mobile phone coverage on changes in (log) agricultural productivity between 2007-2012. The unit of observation is a 10×10 *km* cell. Δ Coverage is the change in the share of cell area covered under GSM mobile coverage between 2007-2012 instrumented using $\mathbb{1}$ (Tower). $\mathbb{1}$ (Tower) is a dummy variable that takes the value of 1 if a cell is both proposed *and* covered by a tower under SMIP Phase I and takes the value of 0 if a cell is proposed *and not* covered. Column (1) reports the average effect and Column (2) reports the heterogenous effects depending on share of cell's population speaking non-official languages. All columns include district-fixed effects, baseline controls as well as other controls. Baseline controls include cell's (log) population, the availability of power supply and average ruggedness. Other controls for the cell include share of labor force employed in agricultural sector, share of agricultural land that is irrigated, access to an educational facility, access to a medical facility, access to a banking facility, number of landline phone connections per 1000 people, distance to nearest town (in *kms.*), night lights intensity, income per capita (in rupees), and expense per capita (in rupees). The sample includes all cells with zero cell phone coverage in 2006. All regressions are weighted by the cell's population. Standard errors clustered at district level are reported in brackets (number of clusters = 285). Significance level: *** p<0.01, ** p<0.05, * p<0.1 .

Table 7: Heterogeneous Effects: Effect on agricultural productivity by baseline productivity in 2007

Outcome:	Δ log(yield)			
Baseline Productivity (2007):	First Quartile (1)	Second Quartile (2)	Third Quartile (3)	Fourth Quartile (4)
Δ Coverage	0.052*	0.012	-0.004	0.052
	[0.030]	[0.029]	[0.030]	[0.603]
Δ Coverage × Non-official Languages (%)	-0.046*	-0.024	-0.025	-1.028
	[0.026]	[0.021]	[0.177]	[8.199]
Non-official Languages (%)	-0.010	-0.000	-0.001	-0.235
	[0.010]	[0.009]	[0.065]	[1.913]
Observations	1,254	1,174	1,181	1,254
R-squared	0.921	0.971	0.975	0.334
District f.e.	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓
Other Controls	✓	✓	✓	✓

Notes: The table reports IV-2SLS estimates of the effect of mobile phone coverage on changes in agricultural productivity between 2007-2012, depending on the baseline productivity levels in 2007. Column (1) considers cells in the lowest quartile of baseline productivity and Column (4) cells in the highest quartile. The unit of observation is a 10×10 *km* cell. Δ Coverage is the change in the share of cell area covered under GSM mobile coverage between 2007-2012 instrumented using $\mathbb{1}$ (Tower). $\mathbb{1}$ (Tower) is a dummy variable that takes the value of 1 if a cell is both proposed *and* covered by a tower under SMIP Phase I and takes the value of 0 if a cell is proposed *and not* covered. All columns include district-fixed effects, baseline controls as well as other controls. Baseline controls include cell's (log) population, the availability of power supply and average ruggedness. Other controls for the cell include share of labor force employed in agricultural sector, share of agricultural land that is irrigated, access to an educational facility, access to a medical facility, access to a banking facility, number of landline phone connections per 1000 people, distance to nearest town (in *kms.*), night lights intensity, income per capita (in rupees), and expense per capita (in rupees). The sample includes all cells with zero cell phone coverage in 2006. All regressions are weighted by the cell's population. Standard errors clustered at district level are reported in brackets (number of clusters = 285). Significance level: *** p<0.01, ** p<0.05, * p<0.1 .

Appendix

A Calls to Kisan Call Center

In this section we describe the methodology followed to extract crop information and type of query made by farmers in all calls to Kisan Call Centers (KCC). KCC agronomists record the correct information on crop and the category of the query in less than 10% of the calls. In the remaining cases, we use the details contained in two text fields available in the KCC data, i.e. farmer's query and agronomist's answer, to obtain the information. To illustrate the procedure, consider the following calls received by the KCC:

Sno	Date	State	District	Crop	QueryType	QueryText	Answer
1	07/22/2009	Uttar Pradesh	Ambedkar Nagar	-	-	Fertilizer Dose in Paddy	Give NPK 120kg 60kg 60kg/hac
2	09/07/2009	Madhya Pradesh	Sagar	-	-	How to control temite in soyabean?	Spray Chlorpyrifos @ 30ml/pump

In Call 1, the farmer calls KCC to get information on the fertilizer dose in Paddy (Rice). The information on crop in the KCC data is missing under the "Crop" field but it is clearly available in the text of the query (variable "QueryText"). Similarly in Call 2, the farmer inquires how to control termites (which is incorrectly recorded as "temites" in QueryText) for Soyabean crop. Similar to the previous call, both the crop information and category of call are missing in the recorded data. We use the information in "QueryText" to deduce what is the crop the farmer is enquiring about (*Soyabean*). We also use the information in the "Answer" field which recommends using Chlorpyrifos to assign the "QueryType" of the call as *Pesticides*.

A.1. Categorizing Crops

We extract crop information based on methodology described above – using information within the text of the query or the answer of the KCC agronomist to the query. In many cases, crops names are recorded in Hindi. For example, Rice is commonly known as *Dhan* in Hindi. Similarly, Wheat is recorded as *Gehun*; Maize is recorded as *Makka*. We detect all these instances and convert the corresponding crop names to English.

A.2. Categorizing Query Categories

We classify calls into 17 broad categories.[41] Here we describe in detail the assignment of the main query categories used in the paper – calls on seeds, fertilizers, irrigation and pesticides.

Calls on Seeds: We classify as farmers' calls on seeds those calls made to obtain information on hybrid seed varieties or calls made to inquire about seed varieties. We use information in either the text of the query or in the answer of the KCC agronomist. In particular, we classify as calls on seeds: (i) calls directly asking about the hybrid varieties related to a crop (ii) inquires or answers about specific high-yielding varieties seeds. For example, farmers ask about the following high-yielding varieties of wheat: DHM-1, WH-542, UP-2338, HUW-468, PVM-502 or

[41] These categories include Pesticides, Yields, Fertilizers, Weather, Field Preparation, Market Information, Credit, Cultivation, Irrigation, Contact Information, Soil Testing, Mechanization, Government Schemes, Seed Availability, Crop Insurance, General Information and Others. The first seven categories are associated with 90% of the calls. We collapse all categories with less than 1% of calls into a combined category of "Others" which in total makes up about 10% of the calls.

about the following high-yielding varieties of cotton: RCH-134, RCH-208, RCH-317, MRC-6301, MRC-6304. Table A.1 below provides an illustrative example for this:

Table A.1: Sample calls on Seeds

Sno	Date	State	District	Crop	QueryType	QueryText	Answer
1	10/17/2010	Haryana	Mahendra- -garh	Wheat	Seeds	Improved varieties of wheat	PBW-343,WH-711, WH-542,DBW-1
2	03/28/2009	Andhra Pradesh	East Godavari	Maize	Seeds	Asked about Varieties	Recommended DHM-107 or 109

Calls on Fertilizers: We classify as farmers' calls on fertilizers: (i) calls seeking general information on fertilizer dosage (ii) calls directly asking remedies for nutrient deficiencies in crops (iii) queries or replies based on required dosage of specific fertilizers, *e.g.* N-P-K or Urea (iv) calls seeking information on plant growth regulators, seed treatment or solution to leaf drop. For example, in many calls farmers asks about the dosage of specific fertilizers, *e.g.* D.A.P(Diammonium phosphate). In few other calls, the agronomist prescribes specific amounts to be used for different chemicals of the fertilizer N-P-K. Table A.2 below provides an illustrative example from our exercise.

Table A.2: Sample calls on Fertilizers

Sno	Date	State	District	Crop	QueryType	QueryText	Answer
1	02/17/2011	Punjab	Amritsar	Wheat	Fertilizers	Sulphur deficiency in wheat	Apply 100 kg gympsum per acre before sowing
2	07/03/2009	Uttar Pradesh	Firozabad	Rice	Fertilizers	Fertilizer dosage in rice	N-120kg, P-60kg K-120kg, ZN-20kg/hec.
3	07/20/2011	Punjab	F.G.Sahib	Rice	Fertilizers	D.A.P dose in paddy	27 kg per acre
4	12/06/2010	West Bengal	Midnapore (East)	Rape	Fertilizers	Flower dropping in mustard	Apply Zinc Sulfate 2 gram/liter water
5	08/09/2011	Mahara- -shtra	Parbhani	Cotton	Fertilizers	Stunted growth of cotton	Spray Urea 100 grams in 10 litre water

Calls on Irrigation: In order to classify calls on irrigation, we use farmers' queries seeking information: (i) directly about irrigation practices (ii) or about water management in the field. Table A.3 below provides an illustrative example: in the first two calls farmers ask about the suitable time for particular stages of irrigation. In the last case, a farmer is seeking information on the quantity of water for irrigating the field.

Table A.3: Sample calls on Irrigation

Sno	Date	State	District	Crop	QueryType	QueryText	Answer
1	01/15/2011	Madhya Pradesh	Sehore	Wheat	Irrigation	Suitable time for 2^{nd} irrigation in wheat	At tillering stage *i.e.* 40-45 days
2	03/11/2010	Bihar	Palamu	Wheat	Irrigation	Minimum irrigation schedule for wheat	20-25,40-45,70-75,90 -95,105 days after sowing
3	06/10/2011	Bihar	Rohtas	Rice	Irrigation	Water management in rice	5-6 *cm* water given in rice field

Calls on pesticides: We classify as farmers' calls on pesticides: (i) calls seeking information specifically about pesticides (ii) agronomist suggesting the use of certain pesticides like Quinalphos, Carbofuran and Chlorpyrifos [42] (ii) calls asking about solutions for pest infection (iii) calls related to plant protection (iv) inquiries about weed control. Table A.4 below provides few examples of calls on pesticides after applying our methodology described above.

Table A.4: Sample calls on Pesticides

Sno	Date	State	District	Crop	QueryType	QueryText	Answer
1	08/29/2010	Andhra Pradesh	Anant- -hapur	Groundnut	Pesticides	Asked about spodoptera damage in groundnut	Spray Quinalphos 2ml/1 liter water
2	07/26/2011	Punjab	Mansa	Rice	Pesticides	Info. regarding control of termite in rice	Apply dilute 1 litre Chlorpyrifos 20ec in 2 litres
3	11/29/2009	Rajasthan	Alwar	Wheat	Pesticides	Prevention of Nematod problem in Wheat	Use Carbofuran 3G 20KG. per hectare soil treatment
4	09/04/2010	Uttar Pradesh	Bareilly	Rice	Pesticides	Insect Control in rice	Apply Endosulphon 35EC at 1.5 ml/lit of water
5	03/09/2011	Gujarat	Surat	Sugarcane	Pesticides	Ask for weed control	Suggested hand weeding
6	08/09/2011	Bihar	Deoghar	Rice	Pesticides	Plant protection in paddy	Given details about plant protection

[42] Quinalphos is an pesticide widely used in India for wheat, rice, coffee, sugarcane, and cotton. Carbofuran is a pesticide used to control insects in a wide variety of field crops, including potatoes, corn and soybeans. Chlorpyrifos is a pesticide used to kill a number of pests, including insects and worms.

B Data Validation: HYV Adoption

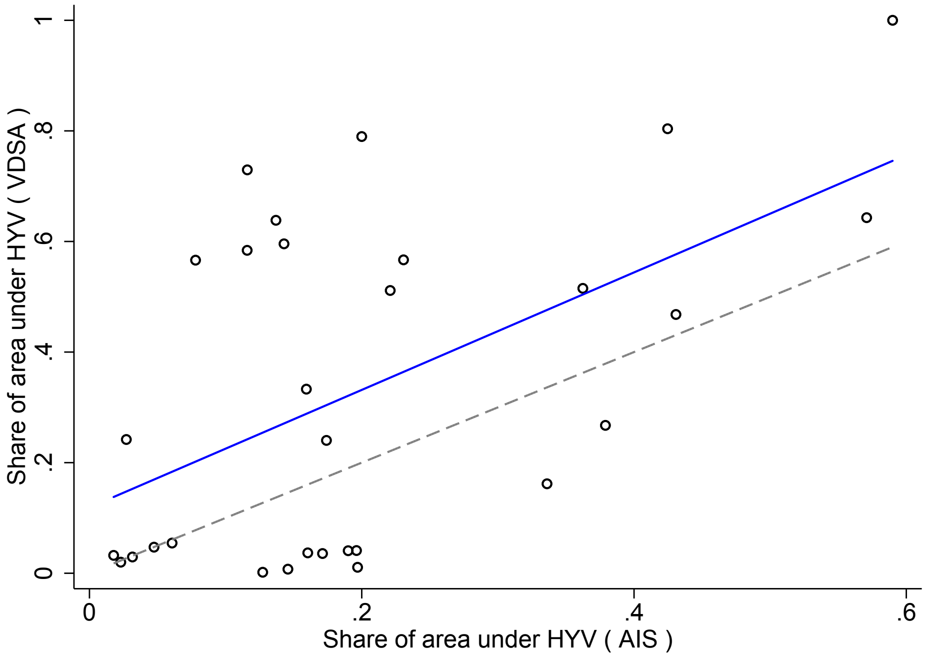
In this section, we validate our main measure of technology adoption — i.e. share of cell area farmed under HYV seeds (described in Section 3.4) — using alternative datasets that are publicly available to researchers. Information on the use of HYV seeds at village level is seldom available. Two publicly available survey data sets that report such information are the ICRISAT Village Dynamics in South Asia (VDSA) and the Tamil Nadu Socioeconomic Mobility Survey (TNSMS) conducted by the Economic Growth Center at Yale University. Both data sets are based on household surveys that collect information on cultivation practices. Both data sets record the crops farmed by each household, the total area farmed under each crop and how much of the farmed area is cultivated with improved or HYV seed variety.

The two data sets differ in their finest identifiable geographic unit of observation. The finest geographical unit of observation in the VDSA data is a village. The survey covers 17 villages in 2012 with non-missing information on HYV seeds.[43] While the TNSMS covers more villages than VDSA, it does not provide village identifiers like VDSA. The finest geographical unit of observation available in TNSMS is much larger than our 10×10 km cell and therefore it is not well suited to validate our measure. Moreover, while TNSMS only covers villages within the state of Tamil Nadu, VDSA spans the five states of Andhra Pradesh, Gujarat, Karnataka, Madhya Pradesh and Maharashtra. Therefore, we use information available in the VDSA data to cross-validate our measure of HYV adoption, as described next.

We compare our measure of share of area farmed with HYV seeds against the one reported in the VDSA data. To do so, we use information in the VDSA data to calculate the total area farmed in each village under a given crop as well as how much of that area is cultivated using HYV seeds. Similarly, we use the share of area farmed with a given crop in a given cell using the data from the Agricultural Input Survey and the methodology described in section 3.4. We then map each 10×10 km cell to VDSA villages based on village centroids. This provides us with 30 observations at the cell-crop level for which we observe HYV adoption in both the VDSA and with our measure. Appendix Figure B.1 shows that our measure is positively correlated with the VDSA measure: the slope of the line is 1.06 and statistically significant ($t = 4.33$).

[43] VDSA only covers six villages consistently between 2002-2012. Four of these villages are in the state of Maharashtra. This limits our ability to compare our measure of *changes* in share of area under HYV seeds as AIS does not cover Maharashtra until 2012. We therefore only compare the *levels* of share of area under HYV seeds in 2012.

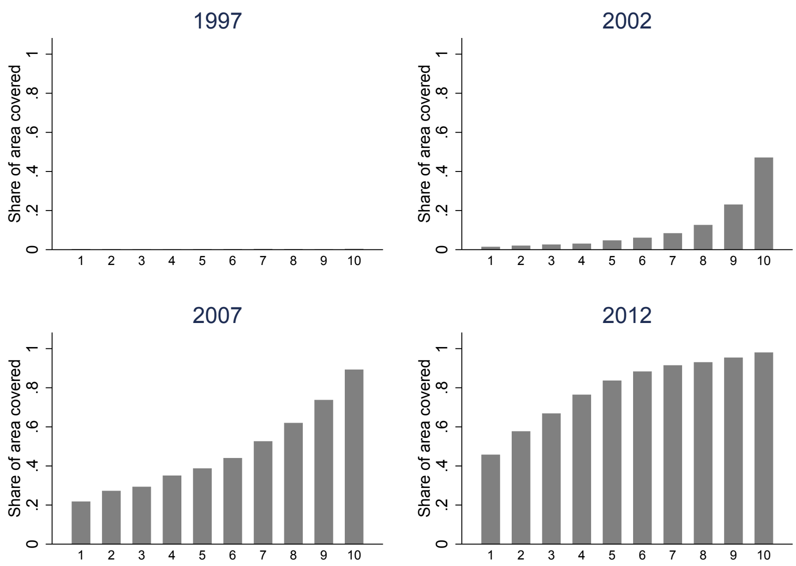
Figure B.1: Data Validation: HYV Adoption



Notes: The graph reports the share of crop area under HYV as calculated from ICRISAT VDSA (Village Dynamics in South Asia) micro data against the share of crop area under HYV seeds as calculated from AIS (Agricultural Input Survey). Each dot represents a cell-crop observation for the two measures of share of area under HYV seeds in 2012. The figure has 30 observations and the slope of the line is 1.06 (t = 4.33). The dashed gray line is the 45 degree line.

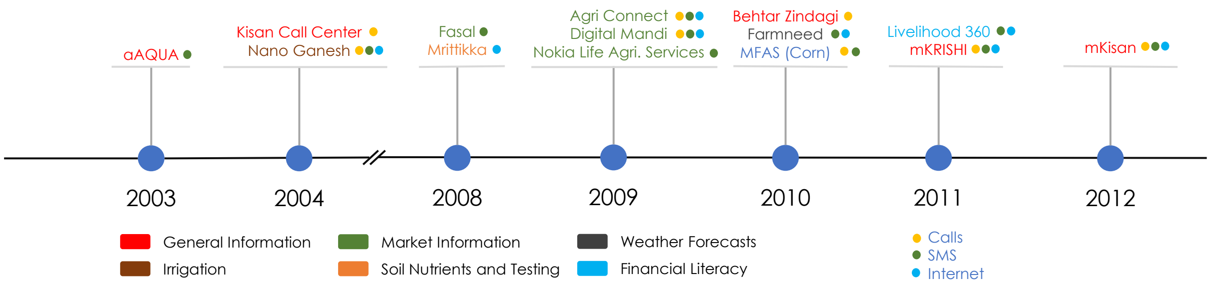
C Empirics: Additional Results

Figure C.1: Mobile Phone Coverage by night lights intensity



Notes: The average share of land with mobile phone coverage in each decile is calculated for the 4 years in which the Agricultural Input Survey was conducted: 1997, 2002, 2007 and 2012. night lights intensity data refers to 1996.

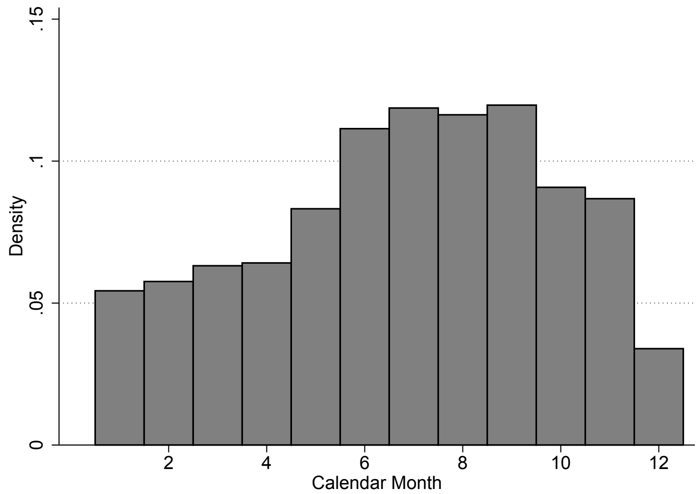
Figure C.2: Indian Providers of Agricultural Advice Services: A Timeline



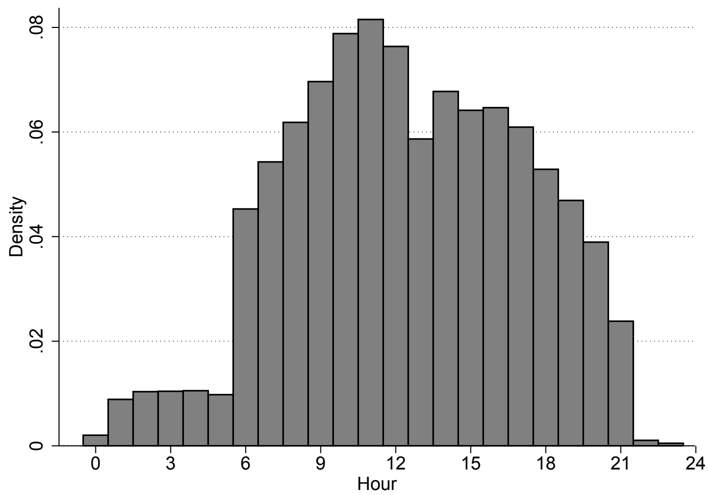
Notes: Source: GSMA mAgri Deployment Tracker

Figure C.3: Distribution of calls made to kisan call center

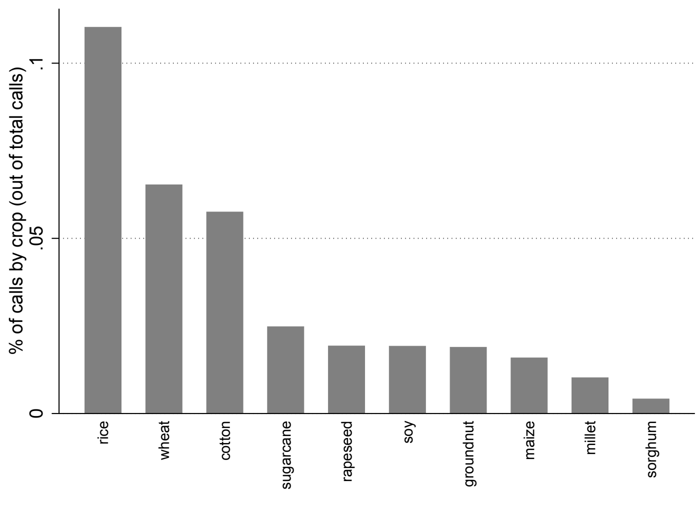
(a) Calls by Calendar Month



(b) Calls by Time of Day

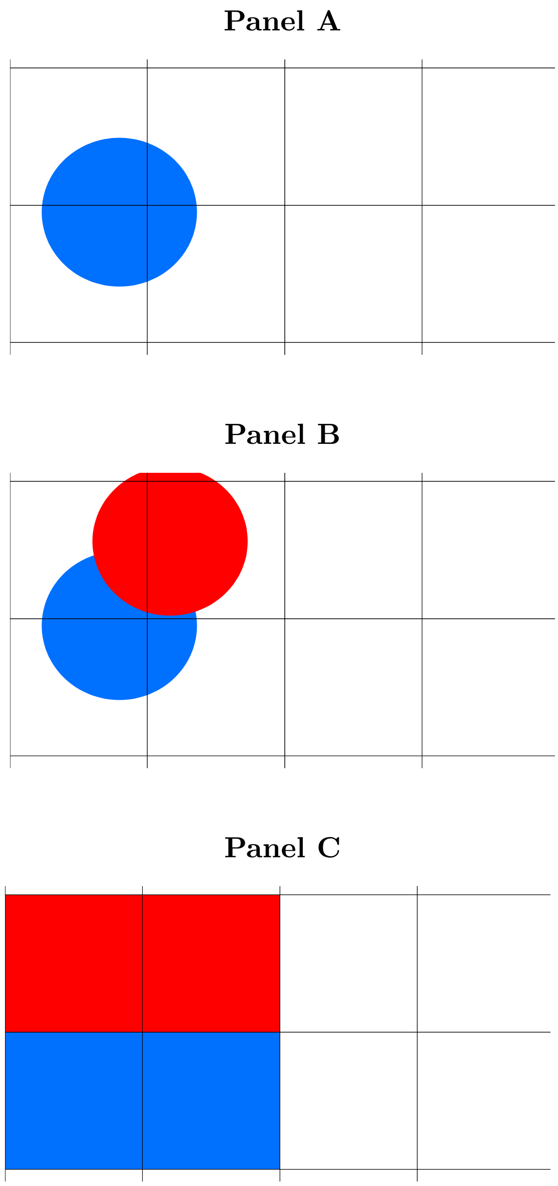


(c) Calls by Crop



Notes: Source: Kisan Call Center, Ministry of Agriculture

Figure C.4: An example of classification of cells into treatment and control groups



Notes: The figure provides an illustration of classification of cells into treatment (red) and control (blue) group. Panel A shows area covered by a *proposed* tower under SMIP. Panel B shows the area covered by an *actual* tower eventually constructed. Panel C shows the assignment of cells into treatment and control groups.

Table C.5: Summary Statistics for Cell Characteristics

	Mean	**Median**	**Standard Deviation**	**N**
log (Population)	10.06	9.99	0.76	6320
Power Supply	0.78	0.92	0.29	6320
Ruggedness	0.47	0.20	0.89	6320
Agri. Workers/Working Pop.	0.57	0.57	0.14	6320
Agri. Land/Cultivable Area	0.45	0.47	0.22	6320
Percent Irrigated	0.36	0.27	0.32	6320
Δ HYV Share (2002-2007)	0.01	0.01	0.06	5019
Δ HYV Share (1997-2002)	0.05	0.04	0.11	4986
Literacy Rate	0.43	0.44	0.12	6320
Education Facility	0.85	0.91	0.17	6320
Medical Facility	0.35	0.29	0.26	6320
Banking Facility	0.06	0.03	0.10	6320
# Phone conn. per 1000 people	1.22	0.30	3.33	6320
Dist. to nearest town(kms)	26.40	20.00	22.31	6320
Night Lights (2006)	1.43	0.72	1.84	6320
Income per capita	75.46	16.76	351.36	6320
Expense per capita	66.44	16.15	268.09	6320

Notes: The unit of observation is a 10×10 km cell. The variables reported are (log) population, fraction of villages in the cell with access to power supply, ruggedness of the cell, share of agricultural workers, share of cultivable land under agriculture, percentage of irrigated land, changes in share of land under HYV, literacy rate, education facility, medical facility, banking facility, number of telephone connections per 1000 people, night lights, distance to nearest town, (monthly) income per capita, and (monthly) expense per capita.

Table C.6: Robustness: Mobile Coverage and Technology Adoption

Outcome:	Δ Technology Adoption							
Technology:	Fertilizers in areas under HYV		Fertilizers in areas not under HYV		Irrigation in areas under HYV		Irrigation in areas not under HYV	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ Coverage	0.047**	0.051**	-0.005	-0.005	0.035**	0.042**	-0.012	-0.015*
	[0.022]	[0.023]	[0.013]	[0.014]	[0.018]	[0.019]	[0.008]	[0.009]
Δ Coverage × Non-official Languages (%)		-0.030		0.007		-0.054***		0.027
		[0.024]		[0.017]		[0.020]		[0.017]
Non-official Languages (%)		0.001		-0.013**		-0.019*		0.013*
		[0.015]		[0.005]		[0.010]		[0.007]
Observations	6,310	6,310	6,310	6,310	6,320	6,320	6,320	6,320
R-squared	0.891	0.891	0.886	0.887	0.822	0.821	0.830	0.829
District f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓	✓	✓	✓
Other Controls	✓	✓	✓	✓	✓	✓	✓	✓

Notes: The table reports IV-2SLS estimates of the effect of mobile coverage on the share of area under fertilizers (Columns 1-4) and the share of area irrigated (Columns 5-8) between 2007-2012. The unit of observation is a 10×10 *km* cell. Δ Coverage is the change in the share of cell area covered under GSM mobile coverage between 2007-2012 instrumented using $\mathbb{1}$ (Tower). $\mathbb{1}$ (Tower) is a dummy variable that takes the value of 1 if a cell is both proposed *and* covered by a tower under SMIP Phase I and takes the value of 0 if a cell is proposed *and not* covered. Odd columns reports the average effect, even columns report the heterogenous effects depending on share of cell's population speaking non-official languages. Columns (1)-(2) and (5)-(6) report the estimates for area cultivated with HYV seeds and Columns (3)-(4) and (7)-(8) report the estimates for area not cultivated with HYV seeds. All columns include district-fixed effects, baseline controls as well as other controls. Baseline controls include cell's (log) population, the availability of power supply and average ruggedness. Other controls for the cell include share of labor force employed in agricultural sector, share of agricultural land that is irrigated, access to an educational facility, access to a medical facility, access to a banking facility, number of landline phone connections per 1000 people, distance to nearest town (in *kms.*), night lights intensity, income per capita (in rupees), and expense per capita (in rupees). The sample includes all cells with zero cell phone coverage in 2006. All regressions are weighted by the cell's population. Standard errors clustered at district level are reported in brackets (number of clusters = 285). Significance level: *** p<0.01, ** p<0.05, * p<0.1 .

Table C.7: Robustness: market fixed effects

Outcome:	Δ Technology Adoption				Δ log(yield)
Technology:	HYV Seeds	Fertilizers	Irrigation	Pesticides	
	(1)	(2)	(3)	(4)	(5)
Δ Coverage	0.054***	0.050**	0.027	0.053**	0.038**
	[0.017]	[0.023]	[0.017]	[0.023]	[0.018]
Δ Coverage × Non-official Languages (%)	-0.033	-0.026	0.001	-0.053	-0.161
	[0.028]	[0.044]	[0.025]	[0.060]	[0.108]
Non-official Languages (%)	-0.007	-0.020*	-0.003	-0.020	-0.028
	[0.008]	[0.011]	[0.007]	[0.014]	[0.031]
Observations	6,092	6,081	6,092	5,914	4,840
R-squared	0.906	0.922	0.876	0.938	0.923
District f.e.	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓
Other Controls	✓	✓	✓	✓	✓
Market f.e.	✓	✓	✓	✓	✓

Notes: The table tests the robustness of our baseline IV-2SLS estimates to the inclusion of agricultural market fixed-effects. The unit of observation is a 10×10 *km* cell. ΔCoverage is the change in the share of cell area covered under GSM mobile coverage between 2007-2012 instrumented using $\mathbb{1}$ (Tower). $\mathbb{1}$ (Tower) is a dummy variable that takes the value of 1 if a cell is both proposed *and* covered by a tower under SMIP Phase I and takes the value of 0 if a cell is proposed *and not* covered. The dependent variable in Column (1) is change in share of area cultivated under HYV; Column (2) is change in share of area cultivated under fertilizers; Column (3) is change in share of area cultivated under irrigation; Column (4) is change in share of area cultivated under pesticides; Column (5) is change in (log) agricultural productivity. All changes are calculated between 2007-2012. All columns include market-fixed effects in addition to district-fixed effects, baseline controls as well as other controls. Baseline controls include cell's (log) population, the availability of power supply and average ruggedness. Other controls for the cell include share of labor force employed in agricultural sector, share of agricultural land that is irrigated, access to an educational facility, access to a medical facility, access to a banking facility, number of landline phone connections per 1000 people, distance to nearest town (in *kms.*), night lights intensity, income per capita (in rupees), and expense per capita (in rupees). The sample includes all cells with zero cell phone coverage in 2006. All regressions are weighted by the cell's population. Standard errors clustered at district level are reported in brackets (number of clusters = 285). Significance level: *** p<0.01, ** p<0.05, * p<0.1 .

Table C.8: Robustness: additional interaction terms

	Baseline (1)	+ Δ Coverage × Agriculture (2)	+ Δ Coverage × Isolation (3)	+ Δ Coverage × Income (4)	+ Δ Coverage × Controls (5)
			Panel A: Δ log(1+ number of calls)		
Δ Coverage	0.828***	1.149**	0.680***	0.809***	0.965*
	[0.206]	[0.508]	[0.181]	[0.202]	[0.578]
Δ Coverage × Non-official Languages (%)	-0.716**	-0.690**	-0.878*	-0.713**	-0.916
	[0.316]	[0.331]	[0.481]	[0.307]	[0.729]
Non-official Languages (%)	-0.185*	-0.203*	-0.234**	-0.188*	-0.273
	[0.096]	[0.108]	[0.117]	[0.098]	[0.202]
Observations	6,320	6,320	6,320	6,320	6,320
			Panel B: Δ Technology Adoption (HYV seeds)		
Δ Coverage	0.047**	0.095**	0.044**	0.049***	0.093*
	[0.019]	[0.044]	[0.018]	[0.018]	[0.049]
Δ Coverage × Non-official Languages (%)	-0.041**	-0.050***	-0.049	-0.041**	-0.059
	[0.019]	[0.019]	[0.046]	[0.019]	[0.056]
Non-official Languages (%)	-0.002	-0.006	-0.004	-0.002	-0.008
	[0.009]	[0.010]	[0.016]	[0.009]	[0.019]
Observations	6,320	6,320	6,320	6,320	6,320
			Panel C: Δ log(yield)		
Δ Coverage	0.041**	0.091*	0.037*	0.046**	0.091**
	[0.020]	[0.051]	[0.021]	[0.020]	[0.045]
Δ Coverage × Non-official Languages (%)	-0.093***	-0.095***	-0.101**	-0.095***	-0.107**
	[0.033]	[0.034]	[0.039]	[0.032]	[0.042]
Non-official Languages (%)	-0.014	-0.016	-0.016	-0.014	-0.018
	[0.012]	[0.011]	[0.013]	[0.011]	[0.013]
Observations	5,033	5,033	5,033	5,033	5,033
District f.e.	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓
Other Controls	✓	✓	✓	✓	✓

Notes: The table tests the robustness of our baseline IV-2SLS estimates to the inclusion on an additional set of interaction terms. The dependent variable in Panel A is the change in (log) calls received at KCC; in Panel B is the change in share of area cultivated under HYV; in Panel C is the change in (log) agricultural productivity between 2007-2012. Column (1) reports baseline estimates of equation (4). Column (2) includes additionally the interactions of share of labor force employed in agricultural sector and share of agricultural land that is irrigated × Δ Coverage. Column (3) includes the interactions of distance to nearest town and average ruggedness × Δ Coverage. Column (4) includes the interactions of night lights intensity and income per capita × Δ Coverage. Column (5) includes simultaneously all the interactions in the previous columns. The unit of observation is a 10×10 *km* cell. Δ Coverage is the change in the share of cell area covered under GSM mobile coverage between 2007-2012 instrumented using $\mathbb{1}$ (Tower). $\mathbb{1}$ (Tower) is a dummy variable that takes the value of 1 if a cell is both proposed *and* covered by a tower under SMIP Phase I and takes the value of 0 if a cell is proposed *and not* covered. All columns include district-fixed effects, baseline controls as well as other controls. Baseline controls include cell's (log) population, the availability of power supply and average ruggedness. Other controls for the cell include share of labor force employed in agricultural sector, share of agricultural land that is irrigated, access to an educational facility, access to a medical facility, access to a banking facility, number of landline phone connections per 1000 people, distance to nearest town (in *kms.*), night lights intensity, income per capita (in rupees), and expense per capita (in rupees). The sample includes all cells with zero cell phone coverage in 2006. All regressions are weighted by the cell's population. Standard errors clustered at district level are reported in brackets (number of clusters = 285). Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table C.9: Robustness: continuous measure (2SLS) (2007-2012)

Outcome:	Δ Coverage	Δ log (1+ number of calls)				
Topic of the calls:		All	Seeds	Fertilizer	Irrigated	Pesticides
	(1)	(2)	(3)	(4)	(5)	(6)
% covered by SMIP	0.149***					
	[0.017]					
Δ Coverage		0.546***	0.189**	0.161***	0.042***	0.489***
		[0.132]	[0.074]	[0.062]	[0.015]	[0.121]
Δ Coverage × Non-official Languages (%)		-0.560**	-0.227***	-0.219***	-0.062**	-0.469**
		[0.221]	[0.085]	[0.075]	[0.025]	[0.200]
Non-official Languages (%)		-0.170**	-0.058**	-0.040	-0.017*	-0.152**
		[0.071]	[0.026]	[0.025]	[0.010]	[0.066]
Observations	6,320	6,320	6,320	6,320	6,320	6,320
F-stat	79.60					
R-squared		0.915	0.933	0.927	0.895	0.922

Outcome:		Δ Technology Adoption				Δ log(yield)
Technology:		HYV Seeds	Fertilizers	Irrigation	Pesticides	
		(7)	(8)	(9)	(10)	(11)
Δ Coverage		0.030***	0.025*	0.017	0.037*	0.026**
		[0.011]	[0.014]	[0.011]	[0.019]	[0.011]
Δ Coverage × Non-official Languages (%)		-0.027*	-0.011	-0.020	-0.044	-0.065***
		[0.014]	[0.024]	[0.016]	[0.031]	[0.022]
Non-official Languages (%)		0.000	-0.011	-0.005	-0.015	-0.008
		[0.008]	[0.013]	[0.005]	[0.011]	[0.008]
Observations		6,320	6,310	6,320	6,142	5,033
R-squared		0.864	0.890	0.815	0.892	0.905
District f.e.	✓	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓	✓
Other Controls	✓	✓	✓	✓	✓	✓

Notes: The table reports the robustness of our baseline IV-2SLS estimates to using as the treatment variable the share of cell covered by SMIP towers instead of an indicator variable. The unit of observation is a 10×10 *km* cell. Column (1) reports the first-stage regression of Δ Coverage on cell area covered by a SMIP tower (% covered by SMIP tower). In Columns (2)-(11), Δ Coverage is the change in the share of cell area under GSM mobile coverage from 2007 to 2012, instrumented using % of cell covered by SMIP. Columns (2)-(6) estimate the effect of change in mobile coverage on change in number of (log) calls to the KCC. Column (2) estimates the effect on total calls, Column (3) on calls about seeds, Column (4) on calls about fertilizers, Column (5) on calls about irrigation, and Column (6) on calls about pesticides. Columns (7)-(10) estimate the effect of change in mobile coverage on change in technology adoption. Column (7) focuses on share of land under HYV seeds, Column (8) on share of land under fertilizers, Column (9) on share of irrigated land, Column (10) on share of land under pesticides. Columns (11) estimates the effect of change in mobile coverage on change in agricultural productivity. All columns include district-fixed effects, baseline controls as well as other controls. Baseline controls include cell's (log) population, the availability of power supply and average ruggedness. Other controls for the cell include share of labor force employed in agricultural sector, share of agricultural land that is irrigated, access to an educational facility, access to a medical facility, access to a banking facility, number of landline phone connections per 1000 people, distance to nearest town (in *kms.*), night lights intensity, income per capita (in rupees), and expense per capita (in rupees). The sample includes all cells with zero cell phone coverage in 2006. All regressions are weighted by the cell's population. Standard errors clustered at district level are reported in brackets (number of clusters = 285). Significance level: *** p<0.01, ** p<0.05, * p<0.1.