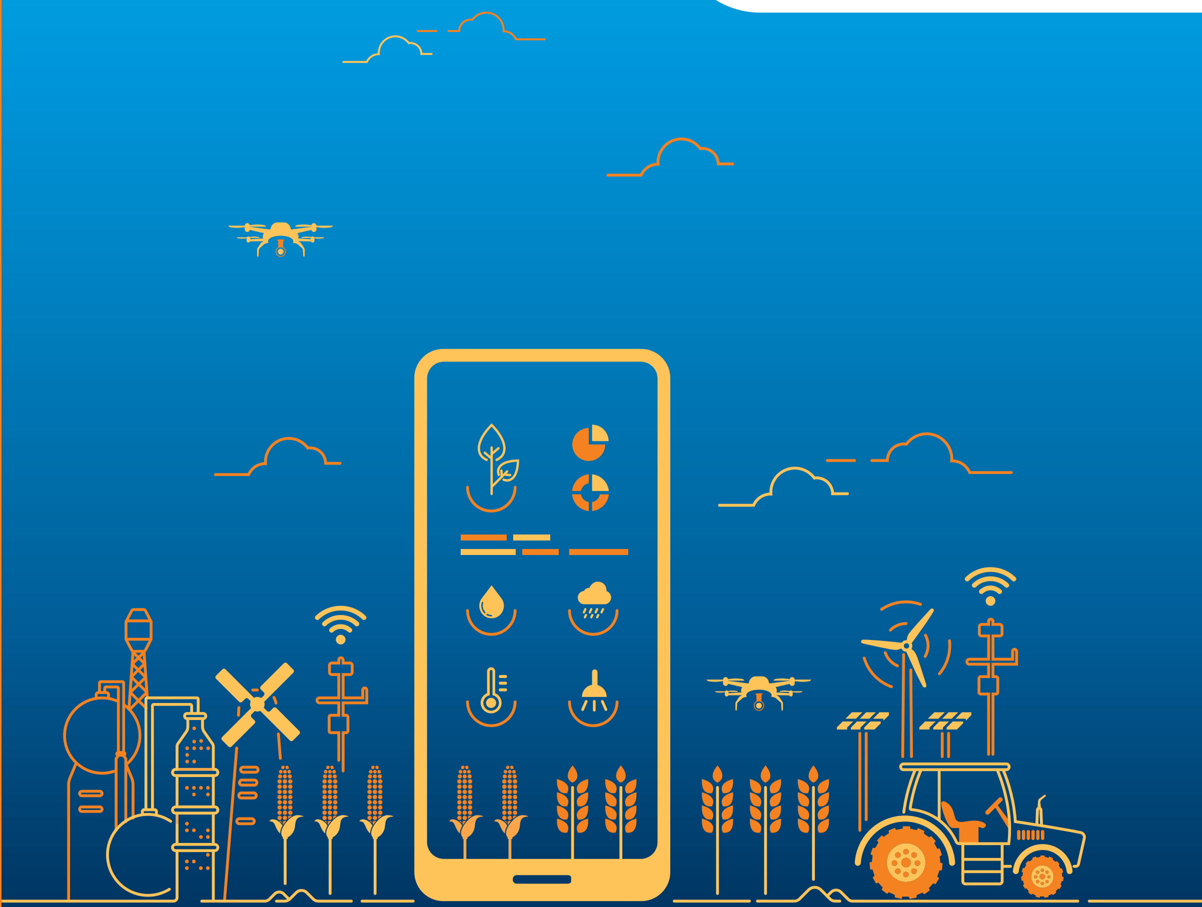


ESSAYS ON FINANCE AND DEVELOPMENT

OP No: 211

EXPORT-IMPORT BANK OF INDIA

OCCASIONAL PAPER NO. 211

ESSAYS ON FINANCE AND DEVELOPMENT

This study is based on the doctoral dissertation titled "Essays on Finance and Development," which is selected as the award-winning entry for the Export-Import Bank of India's BRICS Economic Research Annual Award (BRICS Award) 2022. The dissertation was written by Dr. Apoorv Gupta, currently Assistant Professor of Economics, Dartmouth College, USA, under the supervision of Professor Dean Karlan, Professor Dimitris Papanikolaou, Professor Efraim Benmelech, and Professor Jacopo Ponticelli (Northwestern University, USA). Dr. Gupta received his doctoral degree in 2020 from Northwestern University, USA.

June 2022

CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

This study is a collection of three self-contained essays that study various aspects of economic development, with an emphasis on the factors affecting technology adoption and productivity.

The first chapter studies how demand-side factors affect markups and propose a new methodology to correct for bias in misallocation losses generated by variable markups. Using data on Indian firms, the study first document two key correlations: marginal costs and markups are increasing in firm size. It then explores how these correlations are driven by two factors: the *assortative matching* of wealthier consumers to larger firms, and the lower demand elasticity of wealthier consumers. Results on how firms across the size distribution change their markups in response to exogenous demand shocks to poor households provide support to the demand-based markup channel: producing better quality and selling to wealthier, less demand elastic households leads larger firms to incur higher costs and charge higher markups. This demand-driven markup dispersion lowers aggregate productivity gains from reallocation because firms also adjust their markups in response to policies that may improve allocative efficiency. Firms' pass-through of changes in their costs into prices is a sufficient statistic to account for endogenous markup adjustment when estimating productivity gains from reallocation. Gains from reallocation are 50 percent lower due to demand-driven variable markups.

The second chapter studies how large but temporary aggregate shocks trigger long-run adoption in financial technologies such as electronic payment systems. The network-based nature of many fintech products implies that their adoption is subject to coordination frictions. The study provides evidence on the quantitative importance of these frictions by studying data from the largest provider of fintech payments in India in the aftermath of the demonetization of 2016. Consistent with a dynamic technology adoption model with externalities, in response to the temporary cash contraction, both the size and adoption rate of the platform increased persistently. Estimates of the model suggest that 45% of the six-month response was driven by externalities. With externalities, temporary interventions can thus lead to persistent shifts in adoption, though the study highlights an important limitation of this logic: because externalities create state-dependence, temporary interventions can also exacerbate initial differences in long-run adoption.

The third chapter studies the effect of information on technology adoption and productivity in agriculture. The 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. Information on agricultural practices is measured by analyzing the content of 2.5 million phone calls made by farmers to one of India's leading call centers for agricultural advice. The study finds 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. The 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.

1. DEMAND FOR QUALITY, VARIABLE MARKUPS AND MISALLOCATION: EVIDENCE FROM INDIA

An important line of research, starting with Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), argues that the existence of cross-sectional dispersion in marginal products across firms can be attributed to market distortions that prevent efficient allocation of resources in an economy. Recent work suggests that markups are an important source of this dispersion. Much of this literature, however, routinely treats markups as exogenous wedges and abstracts from underlying sources driving markup variation. If dispersion in markups stems largely from exogenous distortions, then eliminating those distortions will increase aggregate productivity. However, if driven by firms choosing their optimal markups based on heterogeneous demand elasticities, potential gain from reallocation will be limited. A large literature has taken markup dispersion as evidence of allocative inefficiency as given, with little consideration on how gains from reallocation are influenced by demand factors.

The chapter studies how demand-side features shape the distribution of markups and develop a new method to estimate productivity gains from reallocation under variable markups. It shows that segmentation in output product market coupled with differences in demand elasticities across consumers with different income levels can allow large systematic dispersion in markups to persist in equilibrium. An implication is that gains from reallocating resources across firms will be limited because larger firms charge higher markups as they face low demand elasticities. Specifically, firms that face lower demand elasticities will respond to policies that reallocate resources (to lower their markup in the first place) by increasing their markup. This endogenous markup adjustment by firms lowers potential gains from reallocation. The estimates of the extent to which firms pass through changes to their costs into their prices is a sufficient statistic for correcting the bias in reallocation gains under endogenous markups.

The analysis proceeds in four steps. The first step (Section 1.2) documents a systematic relation between firm size, and its marginal costs and markups across manufacturing firms in India. Detailed data from Indian Annual Survey of Industries on firms' input usage and final output is used to estimate firm-product level markups and marginal costs, by building on the work of De Loecker et al. (2016). It is shown that, first, marginal costs within a product group are increasing in firm size. This finding is consistent with the literature on product quality (Kugler and Verhoogen 2011; Atkin et al. 2015), and in line with the findings in these papers, it is documented that input material prices, wages and capital intensity are higher for larger firms. Second, markups are also increasing in firm size. These relationships are more pronounced in sectors with greater scope for quality differentiation, as proxied by the Rauch (1999) classification of non-homogeneous goods.

The second step posits that larger firms have higher markups and marginal costs because of *assortative matching* between firms producing higher quality goods and wealthier consumers, and develops new testable prediction for this mechanism. The approach is motivated by two theoretical ideas. First, following Linder (1961), consumers are asymmetric in income and their willingness to pay for product quality; and firms producing higher quality varieties cater to the demand of wealthier households. Second, firm productivity and input quality are complements in determining output quality, and in equilibrium higher quality is produced by more productive and larger firms (Kugler and Verhoogen 2011). Together, this implies that wealthier households source larger share of their consumption from goods produced by larger firms. This matching on product quality implies that markups charged by firms inversely depend on their sales-weighted average demand elasticity, where the weights are share of firms' sales across income groups. Because wealthier households are less price sensitive, larger firms charge higher markups, consistent with the correlation evidence in Section 1.2. However, the model delivers an additional testable prediction: an increase in the demand from the poorer income group makes the marginal consumer more demand elastic for firms that sell to both rich and poor households. These firms respond by lowering their markups. This is termed as the *demand composition* channel.

The third step (Section 1.3) provides evidence on the demand composition channel using weather-driven exogenous variation in consumer demand. Specifically, to test the claim that increases in demand from the poor lead to higher demand elasticity for firms and a fall in markups, an empirical strategy that uses quasi-random income shocks to poor households as a source of fluctuation in their demand is proposed. The majority of the poor in India are employed in agricultural sector and face substantial productivity risk — even today, less than one-third of the agricultural land is irrigated, making agricultural yields significantly driven by local rainfall variation. These rainfall-driven shocks to agricultural productivity have substantial impact on local income of the poor and, due to their preference to consume lower-quality products, on the demand faced by smaller firms.

The estimates from the identification strategy show that in response to increase in rural income, driven by positive rainfall shocks, firms lower their markups by 0.5 percent. The demand composition channel above posits that rain shocks should disproportionately affect the demand from lower income groups and change the weighted demand elasticity, and hence markups, only for firms selling to both rich and poor households. These firms are proxied in the data by firms in the middle of the size distribution. This channel is tested by analyzing how rain shocks affect demand and markups for firms across the size distribution. First, the effects of positive rainfall shocks on quantity sold are monotonically decreasing with firm size. Positive rain shocks increases the quantities sold for firms in the lowest and middle of the size distribution by 1-2 percent, with no effects for firms in the upper range. Second, the same shock has a non-monotonic effect on markups across the firm-size distribution. Specifically, mid-sized firms lower their markup by 1 percent in response to positive rainfall shocks. In contrast, markups of firms in the lower and upper ranges of the distribution remain unchanged. These responses are only present in sectors with larger scope for quality-differentiation, and are stronger in non-tradable industries and in regions with larger share of rural population. This non-monotonic response of markups to demand shocks to the poor is unique to the demand composition channel, and empirical evidence inconsistent with alternative explanations for lower markups in periods of increased demand is provided.

The last step (Section 1.4) concludes by quantifying the effects of demand-driven variable markups for misallocation losses. It is shown that estimate of pass-through of changes to firm's costs into its prices is a sufficient statistic of how firms will adjust their markup in response to taxes/subsidies that may improve allocative efficiency. These pass-through rates are estimated for all firms in the data and it is documented that larger firms have lower pass-through rate, and the relationship is stronger in quality-differentiated sectors. Because firms that charge higher markups also have lower pass-through rates, aggregate productivity gains from reallocation is biased upwards without this correction. These pass-through rates can be readily estimated using standard techniques (for e.g., see Haltiwanger et al. (2018)) and just requires production data to contain separate information on prices and quantities.

The modeling choices for pass-through, however, have important implications for the quantification of pass-through rates (Goldberg and Hellerstein 2008). This is particularly relevant in oligopolistic settings, as under imperfect competition both the curvature of demand and its elasticity affect firms' pass-through (Weyl and Fabinger 2013). A methodology to separately identify underlying determinants of the pass-through rate is proposed. The framework relies on minimal assumptions on the nature of demand and market structure faced by firms, and leverages the results from previous sections — i.e., firms in homogeneous goods sector face the same slope of demand curve, while the slope of demand curve can vary across firms in quality differentiated sector because of differences in composition of their demand. This observation implies that pass-through across firms in homogeneous sector varies only due to differences in competitive index faced by firms, while any additional variation in pass-through in the quality differentiated sector occurs due to differences in demand elasticities faced by firms.

It is estimated that the aggregate productivity gains form reallocation are substantially large (about 30 percent) when the pass-through is assumed to be complete. However, once accounting for incomplete pass-through across firms, the aggregate productivity gains from reallocation are only 6 percent. This substantial decrease in productivity gains arises because firms endogenously adjust markups in response to policies enacted to lower their markups and improve allocative efficiency in the first place. Part of this endogenous markup adjustment arises because high markup firms cater to consumers with low demand elasticities. The contribution of heterogeneous consumer demand is shown to be large — holding competitiveness index faced by firms fixed, aggregate productivity gains from reallocation are 12 percent under the case when all firms face the same consumer demand elasticity. This implies that demand based markup channel lowers productivity gains from reallocation by 50 percent.[1]

The remainder of the paper is organized as follows. Section 1.1 provides a brief analytical framework for misallocation and derives an expression that allows for endogenous markup adjustment to reallocation policies. Section 1.2 provides an overview of the data, describes the framework to estimate markups and documents motivating facts. Section 1.3 describes the reduced-form empirical strategy to test predictions of a theoretical framework on demand-driven markup variation, and reports the results. Section 1.4 provides a methodology to estimate pass-through rates across firms, and report estimates on productivity gains before and after correcting for bias arising from endogenous markup responses by firms. Section 1.5 concludes.

[1]Differences in demand factors only matter to the extent there is imperfect competition in the economy. If the markets are perfectly competitive to begin with, then firms have no market power, charge no markup and demand factors do not play any role in firms' price setting conditions.

1.1 Analytical framework

1.1.1 Misallocation and Aggregate Productivity

Production misallocation occurs when more production is allocated to less productive firms, and less production to more productive firms in an economy due to distortions faced by firms. These distortions are modeled as input/output taxes imposed on firms that creates an inefficient allocation of resources (Hsieh and Klenow 2009). Consider a firm i that use a single input factor x. The price of the input is p_x and firms face an input tax $(1 + \tau_i^x)$ on the price of that input. The firm's profit function is given by:

$$\pi_i = p_i f_i(x_i) - (1 + \tau_i^x) p_x x_i$$

where $f_i(.)$ is firm's production function which exhibits diminishing marginal returns in x. A cost-minimizing firm selects x_i to the point it equates it's marginal revenue product of input to its cost:

$$\text{MRPX}_i = \mu_i (1 + \tau_i^x) p_x \tag{1.1}$$

where $\mu_i \equiv \mu(mc_i, D_i, \phi_i)$ is firm's markup, and is a function of firm's marginal cost, the demand conditions D_i, and the competitive environment ϕ_i faced by the firm (described later). The marginal revenue product for the input x is directly proportional to both markups μ_i (output wedge) and input wedge τ_i^x. Firms with higher markups μ_i and input wedges τ_i^x will have higher marginal revenue product of input and demand lower x_i than their efficient input demand.

Providing subsidies to high MRPX firms could eliminate wedges, reallocate production factors across firms and increase aggregate productivity. To formalize this argument, aggregate productivity growth decomposition following Petrin and Levinsohn (2012) is described. An advantage of this approach is that it is independent of any assumptions on underlying demand and market structure. The change in aggregate productivity is the difference between changes in output and input costs.

Setting aside firms' entry and exit, and with some algebra, the aggregate productivity growth can be decomposed into a within-firm productivity improvement ("technical efficiency") term and an across-firm allocation ("reallocation") term. Using the relation between firm markups, output elasticity and expenditure share (discussed in Section 1.2), the aggregate productivity gains from reallocation are:

$$\text{APG (reallocation)} = \sum_i \lambda_i \theta_{x,i} \left(1 - \frac{1}{\mu_i}\right) d\log x_i \tag{1.2}$$

where i subscripts firm, μ_i is firm-level markups, $\theta_{x,i}$ is the output elasticity respect to the input, $\alpha_{x,i}$ is the input expenditure as share of firm's revenue, and λ_i are firm's (Domar 1961) weight.

1.1.2 Variable markups and gains from reallocation

Equation 1.1 shows that firms input demand x_i is a function of markups and input wedges ($x_i \equiv x_i(\mu_i, \tau_i)$). Without loss of generality, and a slight abuse of notation, change in input demand to a tax/subsidy τ_i can be written as :

$$d \log x_i = \left[1 + \frac{\partial \log \mu_i}{\partial \log mc_i}\right] \frac{\partial \log x_i}{\partial \log \tau_i} d \log \tau_i \equiv \Gamma_i \frac{\partial \log x_i}{\partial \log \tau_i} d \log \tau_i \tag{1.3}$$

where $\Gamma_i = \left[1 + \frac{\partial \log \mu_i}{\partial \log mc_i}\right]$ is firm's pass-through rate, defined as the effect of firm's marginal costs on its price while holding the market structure fixed. While equation 1.3 defines Γ_i for a fixed ϕ_i, it is general and allows for mc_i to affect ϕ_i. Following Weyl and Fabinger (2013) and Atkin and Donaldson (2015), the following general expression for Γ_i is used:

$$\Gamma_i = \left[1 + \frac{1 + \epsilon_i}{\phi_i}\right]^{-1} \equiv \left[1 + \frac{\delta_i}{\phi_i}\right]^{-1} \tag{1.4}$$

where ϵ_i is the elasticity of the slope of inverse demand curve. It is shown that level of pass-through depends only on (i) elasticity of slope of demand, (ii) demand elasticity σ_i and (iii) competitive structure of industry. Using relationship in 1.3, the expression for aggregate productivity gains from reallocation under endogenous markups becomes:

$$\text{APG (reallocation)} = \sum_i \lambda_i \theta_{x,i} \left(1 - \frac{1}{\mu_i}\right) \Gamma_i \frac{\partial \log x_i}{\partial \log \tau_i} d \log \tau_i \tag{1.5}$$

Equation 1.5 formulates the main argument of this paper — under variable markups, potential gains from reallocation from taxes and subsidies will be affected by the pass-through parameter Γ_i. Under well-known case of monopolistic competition and CES demand, the super-elasticity is zero and $\Gamma_i = 1$, and therefore, optimal reallocation to remove wedges will increase aggregate productivity as expected. Variable markups lower these productivity gains. This is because under incomplete pass-through ($\Gamma_i < 1$), firms also change their markups in response to policy changes, lowering gains from such reallocation.

The pass-through rate, however, is affected by the slope of demand curve, its curvature and market structure faced by firms. Therefore, variable markups can generate bias in estimates of productivity gains from three sources. For example, if incomplete pass-through is driven by differences in firms' slope of demand, misallocation occurs even in a competitive market but gains from reallocation will be limited. This is because when firms charge higher markups due to lower consumer demand elasticity they face, any subsidies received are passed into increasing their markups as these firms have little to gain from lowering their prices. To inform on the magnitude of bias due to differences in slope of demand curve across firms, one would need to separately observe δ_i and ϕ_i. However, both δ_i and ϕ_i are not observable to researchers — indeed, if they were observed then they could have been used that directly to compute pass-through rates.

The remainder of the paper provides a novel strategy to separate out demand forces from competitive forces in the data. To do so as parsimoniously as possible, it makes an additional assumption: the elasticity of slope of demand (the super-elasticity) is constant at all prices. Under this assumption, the study shows that markup differences in homogeneous and differentiated goods sector can be used to inform on the slope of firms' inverse demand. Combining this with estimates of pass-through rates, it obtains estimates of δ_i and ϕ_i for all the firms. This allows to estimate productivity gains under CES versus variable demand, and under different underlying market structure faced by firms.

1.2 Data, Estimation and Motivating evidence

1.2.1 Data

1. Producer-level data. The primary data used in this analysis is Indian plant panel-data, the Annual Survey of Industries (ASI) maintained by the Ministry of Statistics. The basic unit of observation in the ASI is an establishment. The data from 1998 to 2009 that contain both consistent product level information and establishment location information during these years is used.[2] The sample frame for the survey is all manufacturing establishments in India that employ more than 10 workers. Establishments with more than 100 workers ("census" establishments) are surveyed every year, while smaller establishments are randomly sampled each year. The data contains establishment-level identifiers across years for both census and non-census establishments, allowing to construct panel data for both types of establishments. The establishment-level panel data is matched to a separate ASI cross-sectional data previously maintained by the Ministry, allowing to obtain the location of all the plant at the district level.[3] The ASI allows owners who have more than one establishment in the same state and industry to provide a joint return, but very few (less than 5 percent of the sample) do so, and the analysis is conducted at the level of the establishment. Each establishment is treated as a separate firm

[2]The ASI uses the accounting year, which runs from April 1 to March 31. Each accounting year is referred based on the start of the period; for example, the year we call "2000" runs from April 1, 2000 to March 31, 2001.
[3]A district is an administrative unit in India, with an average of 17 districts per state. A district is comparable to US county in size. On average, a district has approximately 2 million total residents.

but the results of the paper (discussed later) hold when explicitly allowing for only single-establishment firms.[4] The analysis is limited to domestic firms by excluding the firms that report non-zero share of their sales exported.

A key advantage of the ASI data is that it provides information on factory-gate wholesale prices for the reporting plants. ASI data enables to track firm's product mix over time because Indian firms are required by the 1956 Companies Act to disclose product-level information on capacities, production, and sales in their annual reports. Product-level information is available for 80 percent manufacturing firms, which collectively account for more than 90 percent of labor force for the ASI manufacturing firms.

Firms report products in the ASI survey using ASI Commodity Classification (ASICC) codes which is the most refined level of product available in the data. There are approximately 2000 unique products in the data. Firms in ASI report not only report total sales, but also report sales and quantity sold broken down by product. As the "product" definition is available at highly disaggregated level, unit values are interpreted as prices. This information is used to define per-unit price as ((Total Sales Value)/(Total Quantity Sold)).

2. Household consumption data. Consumption data for households are from Indian National Sample Survey (NSS) conducted between years 1998 and 2009. The survey records total household expenditure and quantity bought by households across 256 product categories, which is used to construct the per-unit prices at the household-level. The survey is a nationally representative repeated cross-sectional sample of about 500,000 households with sampling weights provided at the district-level.[5]

3. Other data. Rainfall data collected by the University of Delaware is used to construct a time series of rainfall received across Indian districts since the year 1960. These data are gridded by longitude and latitude lines. In order to match these to districts, the closest point on the grid to the center of the district is used and assigned that level of rainfall to the district for each year.

The agricultural data on district-level cropping patterns, crop prices and crop yields comes from the ICRISAT Village Dynamics in South Asia (VDSA) Macro-Meso Database.

[4]Therefore, going forward, the terms firms and establishment are used interchangeably. They always refer to the establishment.

[5]The 256 product categories categories asked in the survey can be broadly classified into food, clothing and footwear, fuel and light, educational expenses, personal care items and durable goods.

1.2.2 Estimating markups and cost from producer behavior

The procedure to estimate markups at the firm-product-level is now described. Markup are estimated using cost-minimization approach developed by Jan De Loecker in his various contributions (De Loecker and Warzynski 2012; De Loecker et al. 2016), technical details for which can be found in those papers. The main benefit of the cost-minimization approach is that it allows to measure firm's markups without having to take a stand on many aspects of the theory. This flexibility in this approach is particularly appealing in the setting as it allows to infer full distribution of markups across firms and products without imposing any parametric assumptions on consumer demand; or the underlying nature of competition; or assumptions on the returns to scale. Moreover, the method allows to estimate markups across different manufacturing sectors and over the complete time period. The estimation relies on cost minimization that provides the following expression for markups:

$$\mu_{ijt} = \theta^{v}_{ijt} \left(\alpha^{v}_{ijt}\right)^{-1}$$

where μ_{ijt} are the markups for firm i producing product j in year t. θ^{v}_{ijt} is the output elasticity for the product with respect to a variable input and α^{v}_{ijt} is the expenditure on that variable input as share of firm's revenue. Material inputs are used as the variable input in production.

Because more than half of the plants in the data produce more than one product, De Loecker et al. (2016) is followed to estimate markups at firm-product level. This estimation procedure has few advantages relative to methods that rely on productivity and markup estimation based on firms revenues (instead of physical quantities). Specifically, the availability of micro-data allows to overcome two biases in markup estimates when compared to the existing work. First, due to data limitations procedures that uses revenue-based measure of productivity estimation typically rely on industry-level price deflators. This leads to measurement error when firms produce differentiated products, price differentiate or have market power. The use of physical output allows to overcome this issue. Second, unobserved differences in input quality across firms and over time could generate bias in productivity estimation. This unobserved variation in input quality is controlled for by using information on output prices, with the intuition that output prices contain information on both output and input quality (Kugler and Verhoogen 2011).

1.2.3 Motivating evidence

The section documents four facts consistent with *assortative matching* — the tendency of wealthier consumers to source their consumption from goods produced by larger firms: (1) product-level marginal costs are increasing in firm size (2) larger firms charge higher markups

(3) the positive relation between firm size, cost and markups is stronger in quality differentiated sectors, and (4) richer households consume higher priced products.

1. Firm-level facts. Panel (a) of Figure 1.1 shows that larger firms incur higher marginal costs. The figure plots the residual value of log product marginal costs (y-axis) and log number of employees (x-axis) after controlling for district-product-year fixed effects. Therefore, within the same product-group and located in the same district, smaller firms incur lower marginal costs than larger firms.

Panel (b) of Figure 1.1 documents one of the central findings of the paper: that larger firms charge higher markups for their products. The residual values along both axes are after controlling for district-product-year fixed effects. Hence, within the same product-group and located in the same district, larger firms charge higher markups than smaller firms. Table 1.1 summarizes these correlations. Firms with 10 percent larger labor-force have 0.96 percent higher sales prices, 0.41 percent higher marginal costs and 0.56 percent higher markups. Columns 3-5 shows that higher marginal costs are associated with higher priced inputs, wages and capital intensity.

Panel (c) and (d) of Figure 1.1 shows that the positive correlation of marginal costs and markups with firm size is stronger in sectors with greater scope of quality differentiation, proxied using Rauch (1999) classification of product differentiation. Table 1.2 reports these correlations. Column 1 shows that the positive relation between marginal costs and firm size is entirely driven by more differentiated sectors. Column 2 shows that for relationship between firm size and markups is about 1.7 times higher in sectors with greater scope of quality differentiation. Columns 3-5 show that the relation also holds for input prices, wages and capital intensity.[6]

2. Household-level facts. Figure 1.2 documents the relationship between log per-unit price for a manufactured good consumed by households and their income. The numbers depict the residuals of a regression of log unit price on region-by-product fixed effects, where region is either a town or village and is finer geographical unit than a district, and household controls. Therefore, within the same region-by-product group, wealthier households pay higher average unit-price for the products they consume.

[6]As differences in markups across firms could be driven by differences in marginal costs, the study also verifies that the positive relationship between firm size and markups hold, with the relationship stronger in quality differentiated sectors, when controlling for marginal costs. This further suggests on the potential role of demand factors in driving markup dispersion across firms.

1.2.4 Discussion

The evidence presented in this section strongly rejects constant markups across firms. While the stronger positive relationship between costs and markups with firm size in quality differentiated sector is suggestive of the role of demand factors in generating variable markups, it is silent on the nature of demand factors and does not rule out supply forces. For example, demand models such as Melitz and Ottaviano (2008); Zhelobodko et al. (2012) where consumers have non-CES preferences and firms compete in monopolistic competitive environment generate positive correlation between markups and firm-size. In supply-side models of variable markups such as Atkeson and Burstein (2008); Edmond et al. (2015) where consumers have CES preferences and firms compete in imperfectly competitive environment, markups are higher for larger firms.

The study provides a new demand-side framework of variable markups by linking differences in expenditure across the consumer-income distribution to firm-size distribution through product quality. The main predictions of the model are summarize below, and more details on the framework can be found in Gupta (2021). The model delivers two comparative statics results. First, it generates predictions on relation between firm-size, markups and costs that are consistent with the empirical correlations documented above. Specifically, the model predicts that markup dispersion in quality differentiated sector is generated due to assortative matching between firms and households. Second, and more importantly, the model generates testable prediction for how firms across the size distribution should change their markups in response to demand shocks across the income distribution.

Specifically, the model predicts that firms lower their markups in response to an increase in demand from poor households. Moreover, this markup response to an increase in demand from poor households is non-monotonic across the firm-size distribution. Figure 1.3 documents this prediction for a simple example: Consider only two consumer groups in the economy - the poor and the rich households, and let there be a positive income shock to the poor population. The figure shows the markup responses across the firm size distribution in response to positive income shock to the poor population. Two findings emerge. First, the elasticity of markups is zero in absence of any heterogeneity in demand elasticities (i.e. under CES preference structure), and in absence of assortative matching on quality (i.e in homogeneous goods sector). Second, when poorer households have higher demand elasticities, markup elasticity is strictly convex with respect to share of sales made to the poor. The elasticity is highest for firms catering to both rich and poor households, while it approaches zero for firms making most of their sales to the poor households, and for firms making most of their sales to the rich households. The curvature of the function is also increasing in the demand elasticities gap between the two income groups.

Intuitively, positive demand shocks to poor have a stronger positive effect on sales of firms that cater to a heterogeneous consumer base. This makes these firms pay more attention to the demand elasticity of its more price elastic consumer base, lowering their markups. Section 1.3 formally tests prediction in the data.

1.3 Empirical Strategy and Results

This section proposes an identification strategy to test the model's predictions. The objective is to understand how firms adjust their markups in response to changes in demand across the income distribution. However, any correlation between price changes and quantities will not identify the causal effect of demand because of (i) reverse causality: high quality products could observe an increase in their demand, that is causality might run from prices to quantities; (ii) omitted variable bias: changes along the demand curve i.e. changes to marginal costs of production could change firms' prices and therefore the demand for their products; and (iii) measurement error: estimates could be mechanically negative as prices are calculated as product revenue divided by quantity sold for that product.

To address these identification issues, changes in consumer demand driven by changes to household income due to local rainfall fluctuations are used. The idea is the following: quantity demanded by a consumer group over time is increasing in the income for that group. Thus, one can obtain variation in firm-level demand from changes in income for households over time. These income changes affect the demand for firms depending on the share of firms' sales made to consumer group. Local rainfall fluctuations, by significantly affecting the income for poor households, are ideal instruments for changes in their demand and serve as quasi-random demand shifter for firms that cater to these consumers.[7] To estimate how rain shocks affect firm outcomes, the following specification is used:

$$\log y_{ijt} = \beta \, \text{Shock}_{dt} + \alpha_{ij} + \alpha_{jt} + \gamma \tilde{X}_{ijt} + \epsilon_{ijt} \tag{1.6}$$

where y_{ijt} is the year t outcome of interest (demand, quantity sold, costs, and markups) for product j produced by firm i located in district d. Shock_{dt} are local rain shocks in district d and year t as defined below. As products produced by different firms could differ across various characteristics, firm-product fixed effects α_{ij} are included which absorbs any time-invariant firm-product unobservables (for example, any constant quality differences). The presence of product-year fixed effects α_{jt} controls for product-specific inflation rates and any macro-economic shock at the product level. The firm-product and time dummies therefore capture

[7] Using weather-induced income also has an additional advantage over other measures of local income changes (for e.g., industry level wage growth) as the latter could be driven by changes in price levels in the local economy.

permanent differences in price levels among different products and common time-trends in prices. $\tilde{X}_{ijt}$ are set of firm, product and market level controls described as they are used in section 1.3.3. The reduced form coefficient β in the specification is straightforward to interpret as the elasticity of the response of firm-product level variables to rain shocks in district d across various years.

1.3.1 Rainfall shocks in India

Agricultural households in India face extremely high income volatility across years. 70 percent of farmed area in India is rain-fed; and thus the agricultural production is considerably dependent on rainfall. Rainfall exhibits significant variation across districts and over years, and are an important driver of agricultural productivity and rural income.[8] The local rainfall fluctuations generate income changes for poor households, and affects the demand for firms that cater to the demand of these consumers.

A positive shock is defined as the annual rainfall measure above the 70^{th} percentile and negative shock as rainfall measure below the 30^{th} percentile within the district. The "positive" and "negative" shocks should not be taken in an absolute sense it does not compare districts that are prone to higher rainfall to those that are prone to lower rainfall. These are simply high or low-rainfall years for each district during 1960-2009. For the analysis, rain shock is defined as equal to +1 for positive shock, -1 for negative shock, and 0 otherwise. This definition is similar to the one employed in Jayachandran (2006).

1.3.2 Identification Assumptions

Consistent estimation of β in specification 1.6 requires two conditions to be satisfied: relevance of rain shocks, that is, rain shocks and quantities should be correlated; and exclusion restriction, that is, rain shocks should be uncorrelated with ϵ_{ijt}. These conditions are discussed in detail next.

Relevance. Relevance (i.e. the first stage) can be directly tested in the data — local rainfall deviation should be strongly correlated with the local income and the quantity demanded for poor households. Two results lend strong support to the hypothesis that rain shocks change the *relative* demand of the poor households, and affect the demand disproportionately across the firm-size distribution.

[8]Similar to many other developing countries, majority of the poor in India are employed in the agricultural sector. About 66 percent of males and 82 percent of females in rural India report agriculture (as either farmers or laborers) as their principal economic activity (Mahajan and Gupta 2011).

1. Consumer-level evidence. First, rain shocks only affect the wages of agricultural population and does not affect the wages of the population employed outside agricultural sector during the rainfall months. Column 1 and 2 of Table 1.3 shows the effect of rain shocks on agricultural productivity and revenue: positive rain shocks increase average agricultural yields in the district by 5 percent. Column 3 of Table 1.3 shows the effect of rain shocks on incomes of the poor: daily wages of farmers and agricultural workers increase by 2.7 percent. Rain shocks do not affect wages for households employed outside agricultural sector in rural areas or for non-rural labor force (Column 4 and 5).

2. Firm-level evidence. Second, rainfall shocks are compared with idiosyncratic demand shocks that are estimated using firms' production data. Specifically, ASI firm-panel data is used to obtain firm's idiosyncratic demand by isolating total quantity movement from quantity movement due to a shift in firm's supply curve. Firm-product level demand-shifters are estimated following Foster et al. (2008). Changes to marginal cost are used as instrumental variables (IV) for supply-side price shifters. Marginal costs incorporate firms' idiosyncratic cost-shifters embodied in their input prices and firm's technology. Thus, it has explanatory power over firms' prices which are unlikely correlated with short-run changes to demand.[9]

The residuals from the demand function estimation are exploited to provide evidence on the relevance of rain shocks for firm-level demand. Table 1.4 reports the correlation of demand shocks η_{ijt} with rainfall shocks (using specification 1.6). Column 1 shows that firms' estimated idiosyncratic demand increases by 1.2 percent during years of positive rain shocks. This results hold if quantity sold by firms is instead used as a direct measure of firms' demand (Column 2). Figure 1.4 Panel (a) and (b) show that these correlations are strongest for smallest firms and firms in the middle of the size distribution, and gradually decreases to zero for the largest firms.

The combined evidence on the effect of rainfall shocks on income of the poor households, and the demand across the firm-size distribution provides strong empirical support for the relevance of instrument.

Exclusion Restriction. The second identification condition that rain shocks should satisfy is exclusion restriction. That is, rain shocks should affect markups only through changes to demand curve faced by firms. While this assumption cannot be directly tested, richness of the

[9]Relative to Foster et al. (2008), marginal costs instead of TFPQ are used as an instrument for prices. Unlike TFPQ which is estimated at firm-level, using marginal costs estimates have the advantage of being available at the firm-product level, and therefore, provides greater variation. Results are qualitatively similar if instead TFPQ is used as supply-side instruments for prices.

production data allows to test whether rain shocks might affect firm supply curve. As mentioned before, observing prices at firm-product level allows to back out marginal costs using estimated markups, allowing to test whether (and how) rain shocks affect marginal costs across firms. Columns 3-6 of Table 1.4 report the correlations of rain shocks on firms' marginal costs and its underlying components. No supporting evidence is found that rainfall shocks affect marginal costs on average, or firms' physical productivity (TFPQ), wages and prices of material inputs. Figure 1.4 Panel (c) shows that rain shocks do not affect marginal costs across the firm-size distribution.

However, it could still be that while rain shocks do not affect marginal costs on average, they might still have non-zero effect on costs for some firms. For example, an increase in demand could affect costs through changes in X-inefficiencies for few firms and not others. These changes in firms' costs can have an independent supply-side affect on markups (De Loecker and Goldberg 2014). Testing for effects of rain shocks on marginal costs across all firms separately is an infeasible task. Therefore, marginal costs if controlled for in specification 1.6 in order to isolate markup responses due to changes in demand from rain shocks. This addresses any omitted variable bias by absorbing any component in the error term that might be correlated with both markup changes and quantity produced.

1.3.3 Results

This section presents the empirical results from the proposed identification strategy. It starts by estimating the demand elasticities across the income distribution and document that price elasticity of demand is decreasing in household income levels. Then the effects of rain shock induced demand shocks on producer markups are analyzed and evidence on the demand composition channel by analyzing the differential effects of rain shocks on markups across the firm-size distribution is provided.

Estimates of demand elasticity across income groups. This section start by estimating price-elasticity across different income groups by only considering goods with limited scope of quality differentiation.[10] The OLS estimate of demand elasticity will be potentially biased due to unobserved taste shocks in the error term could be correlated with price changes.This issue is addressed by instrumenting local prices with state-level leave-out mean price changes. The instrument identifies the local average treatment effects where the complier group of the IV will

[10]This information is sourced from NSS Consumption Expenditure Survey and includes products recorded under "cereals" (i.e., grain) category. The category includes quantities and prices for rice, wheat, jowar, bajra, maize, barley, and small millets.

be local and regional sellers for the products (Faber and Fally 2020). Figure 1.5 shows that price-elasticity of demand is decreasing in income levels.

Effect of rain shocks on producer markups. The identification strategy described in section 1.3 is now used to estimate the effect of rainfall-induced demand shocks on the markups charged by firms. Table 1.5 presents the results. Column 1 shows that firms lower their markups by 0.5 percent in years of positive rain shocks. In Column 2-8 show that the results are robust of inclusion to various controls. Column 2 restricts the analysis to only single plant establishment to address the concern that multi-plant establishment might not be responsive to local shocks. Column 3 includes controls for firms' age to allow for markup changes with firm growth (Peters 2020). Column 4 includes controls for firms' size quartile and its interaction with year fixed effects to allow for aggregate shocks across size groups. Column 5 includes controls for past two-years of rain shocks in the district to allow for any effects from lagged changes in demand. Column 6 controls for market access measure constructed from Allen and Atkin (2016), which is a weighted average rainfall deviation for every other district connected to district, where the weights are proportional to the distance between the two districts. Column 7 controls separately for an in-state and an out-state market access measure to allow for separate impact based on whether other districts are in the same state as original district or outside the state (Rotemberg 2019). Finally, Column 8 allows for combined effect of controls from Column 2-7. As can be seen, addition of these controls has no significant effect on the estimate of average effect of rain shocks on markups.

These findings suggests that firms lower their markups in response to positive rain shocks. This is consistent with the hypothesis that rain shocks, by increasing the relative demand from more price-sensitive households, increase the demand elasticity of firms. The next set of results provide evidence on this mechanism.

Mechanism: Demand Composition Channel. The next set of results provides evidence supporting the role of consumer heterogeneity in driving markup variation. It is shown that obtaining separate data on quantity and prices (and hence markups and marginal costs) can allow to qualitatively distinguish between sources behind markup variation. An important feature of the *assortative matching* is that demand shocks to the poor can be used to separate out the level channel from the composition channel because small and large firms cater to different segments of the income distribution. To see this, notice that the markup elasticity to firm idiosyncratic demand shocks is driven by markup change due to changes in slope, and due to markup changes to changes in level of demand.

A researcher does not observe the second term on right hand side — if she did, then it can be used directly. However, by observing both the left hand term and last term on the right hand side allows to draw inferences on markup responses due to a change in slope of demand. More specifically, markups do not change in response to idiosyncratic demand unless the slope of demand curve changes. A parallel shift in the demand curve does not change markups. Markups only change when a shift in firms' idiosyncratic demand is accompanied by shift in slope of firms' demand curve (i.e. firms' demand elasticity). Markups do not change if there is no shift to firms' demand (this is more mechanical). In fact, if supply factors dominate, then markup changes should be highest among producers for which idiosyncratic demand changes the most. Only under demand composition, we can expect non-monotonic responses. The demand composition channel posits that an increase in demand from the poor households increases the demand elasticity only for firms that sell to both rich and poor households, forcing them to lower their markups. Under assortative matching, these firms are proxied in the data by firms in the middle of the size distribution. Therefore, demand composition, and hence markups, should change for firms in the middle of the size distribution. On the other hand, smallest and largest firms cater only to poor and rich consumer base, respectively, and thus rain shocks should not effect their demand composition. To test this prediction, the effect of rain shock on each quartile of firm-size distribution are estimated.[11]

Estimation results are presented in Table 1.6. Rain shocks only affect markups in the middle of the size distribution. The estimated coefficient of -0.7 to -0.9 percent and -0.5 to -0.8 percent in the second and third quartile, respectively, of the size distribution is more than two to three times larger than the lowest quartile (which are insignificant across all specifications). Firms in the largest size quartile do not change their markups as well. The estimates remain stable after inclusion of various controls from Table 1.5 (Columns 2-8). This non-monotonic relationship is clear from Figure 1.6, and is consistent with the prediction and simulated relation (Figure 1.3) from the theoretical framework.

The finding that firms in the lowest quartile of size distribution do not lower their markups is consistent with the hypothesis that rain shocks do not change the demand composition for these firms significantly. Similarly, firms in the top quartile of the size distribution do not change their markups as rain shocks have no effect on their demand. With no changes to the level of their demand, the composition of demand also remains unaffected for the largest firms.

[11]The firm size (using firm's first occurrence in the panel) based on its labor force relative to two-digit industry average is used. Using 2-digit industries instead of products increases the number of observations within each quartile and reduce the noise associated with misclassfication.

The key take-away from this section is that differences in demand composition across firms — arising from assortative matching on quality between firms and consumers — are necessary to rationalize the patterns of markup dispersion observed in the data. However, these results leave two related questions open. First, they do not indicate how large are the misallocation losses due to variable markups. That is, in terms of the framework of Section 1.1, how large are aggregate productivity gains from reallocation when firms can respond by changing their markups. Second, from these results, no conclusion can be drawn on the contribution of demand-factors over and beyond supply-side factors (say, due to different degrees of competition). The next section addresses both questions by providing an approach that relies on firms' pass-through of changes in cost into its prices, and proposes a methodology to separately identify firms' slope of inverse demand and competitiveness index from estimated pass-through rates.

1.4 Implications and aggregate effects

Markups act as wedges across firms that create misallocation (Equation 1.1). To remove this misallocation, a central planner would impose taxes and subsidies across firms to the point that markups are equalized across firms and there are no further gains from reallocation. Equation 1.2 implies that providing subsidies/imposing taxes of $d\log x_i = [\ln \mu_i - \sum_i w_i^x . \ln \mu_i]$, where $w_i^x = \frac{x_i}{\sum_i x_i}$, across firms will equate markups.[12] The aggregate productivity gains from reallocation are:

$$\text{(1.7)} \quad \text{APG-R} \equiv \text{APG (reallocation)} = \sum_i \lambda_i \theta_{x,i} \Gamma_i(\phi_i, \delta_i) \left(1 - \frac{1}{\mu_i}\right) \left[\ln \mu_i - \sum w_i^x . \ln \mu_i\right]$$

In order to estimate productivity gains the following is used: (i) materials as the only input in production, and (ii) materials, labor and capital as inputs in production. The motivation on the use of materials comes from markup estimation that relies on materials as the variable input.

Identification of Γ, δ and ϕ. All the variables apart from pass-through rates in equation 1.7 are available either directly in the data or can be obtained through the markup estimation procedure described in section 1.2.2. Below the methodology to estimate pass-through rate Γ_i for all firms is described. Then estimated Γ_i is used to analyze how demand-driven variable markups affect aggregate productivity gains from reallocation. To make progress on this question, estimates of δ_i from pass-through rate are required. Unfortunately, without knowledge on either δ_i and ϕ_i, there is no unique mapping between Γ_i and the other variable. However, the results from last sections provide with a credible restriction that allows to separately estimate δ_i and ϕ_i from

[12]To arrive at the expression for resources to be reallocated, the following two conditions are considered: (i) the taxes or subsidies that would be required to equate markups across firms, and (ii) total supply of resources in the economy is fixed.

Γ_i. The results in previous section show that firms in homogeneous sector face same slope of demand curve, and therefore, $\delta_i = \delta^{\text{non-diff}}$ across firms in homogeneous sector. Any differences in pass-through rates across firms in homogeneous sector are thus driven by differences in competitiveness. The framework of Atkeson and Burstein (2008) is followed to model the relation the between firm's market share (proxied by its employment) and its competitiveness. For firms in quality differentiated sector, variation in pass-through is driven by differences in slope of demand as well as competitive index. It is assumed that the variation in pass-through rate coming from competitiveness index is similar across the two sectors. This methodology is described formally below in three steps.

(1) Step 1: Estimate firm-level pass-through rates Γ_i: Firm-level pass-through can be estimated using the information on prices, marginal costs. The pass-through rates are estimated both by using marginal costs directly and by instrumenting it with estimated physical productivity (TFPQ). An advantage of using marginal costs is that the estimates are available at firm-product level, while TFPQ varies only at the firm level. A drawback of using marginal costs is that it is calculated using prices and markups, and therefore any measurement error could generate upward bias in the estimates of Γ_i. Instrumenting the marginal costs with TFPQ addresses this issue. However, as is shown later the OLS and IV estimates of pass-through are not significantly different from one another, suggesting that the bias is not a primary concern.

(2) Step 2: Recover estimates of competitive index for homogeneous goods sector:

Suppose, one obtains an unbiased estimate of pass-through rate from Step 1. The finding that all firms in the homogeneous good sector face the same slope of demand is used, which implies that any variation in pass-through arises only due to imperfect competition (embedded in ϕ_i). Next, following Atkeson and Burstein (2008) and Edmond et al. (2015), it is assumed that is one to one mapping between firms' competitiveness and relative size of the firm within its industry. This step provides that relationship between competitive index and firm size.

(3) Step 3: Recover estimates of slope of curvature of demand for firms in differentiated sector: Finally, it is assumed that the relationship between firm-size and competitiveness index in Step 2 is similar across both homogeneous and differentiated sector. This assumption is based on the results in previous sections which argued that markup dispersion due to differences in slope of demand curve faced by firms should arise only in quality differentiated sector. The estimates for competitive index obtained in Step 2 is then used to estimate slope of demand for firms in quality-differentiated sector.

Results. Table 1.7 shows the results on pass-through estimates from the above strategy. The average pass-through rate is 55 percent (OLS estimates, Column 1), and 70 percent (IV estimates, Column 4). Columns 2 and 5 show that larger firms pass-through less of changes in costs into their prices. Column 3 and 6 show that the negative relation between pass-through and firm size is stronger in quality-differentiated sectors.

Figure 1.7 shows the estimates of pass-through rate, slope of demand curvature and competitiveness index across homogeneous and differentiated firms across firm-size distribution. Panel (a) shows that pass-through rates are decreasing in firm size, and this relationship is stronger in quality differentiated sectors. Panel (b) shows the negative relationship between pass-through rate and firm size also reflects lower competition faced by larger firms. Panel (c) shows that larger firms in quality-differentiated sector face higher slope of inverse demand (i.e., less elastic demand curve) relative to smaller firms.

Next, these parameters are used to estimate aggregate productivity under various scenarios, and assess the role of demand factors in reducing aggregate productivity gains from reallocation. Table 1.8 presents the results. First step calculate gains from reallocation without any endogenous markup adjustment by firms. This is the case when markups are assumed to be exogenous wedges that do not react to underlying environment or policy changes. $\Gamma_i = 1$ is plugged along with the estimates of firm-level markups, weighted-average markups and Domar weights in equation 1.7.

The first row shows that ignoring endogenous markup responses, aggregate productivity growth due to reallocation is about 30 percent. The second row of Table 1.8 allows for firms to adjust their markup in response to taxes/subsidies aimed at removing markup dispersion and to reallocate resources. As described in Section 1.1, this markup adjustment is captured by firm-level pass-through rates. The estimated coefficients of estimated pass-through rate is plugged from the data to equation 1.7. As reported in the second row, while the estimated productivity gains from reallocation are still positive (they are approximately 6 percent), they are an order of magnitude lower than the case when pass-through is assumed to be complete. This is because high markup firms are also the firms that have the lowest pass-through of subsidies into their prices.

Finally, the third row of Table 1.8 analyzes the role of demand composition channel by applying the restriction that all firms face the same slope of demand curve as faced by firms in homogeneous good sector, while holding fixed their estimated competitiveness. Specifically, estimated slope of demand for non-differentiated firms from the Step 2 are used described above and plugged into in equation 1.7. The third row shows that if all firms faced the same slope of demand curvature, aggregate productivity gains are 12 percent. While this gains are still less than 60 percent of the gains when pass-through is assumed to be completed, they are still double than the observed pass-through in the data (that allows slope of demand to vary across firms). This implies that, holding the level of competitiveness fixed, demand-driven markup dispersion lowers the aggregate productivity gains from reallocation by 50 percent.[13]

1.5 Conclusion

This study provides evidence on how demand-side characteristics affect the equilibrium distribution of markups across firms and assesses the implications of this demand-driven markup dispersion for understanding misallocation losses. The results provide strong support for macro models that feature heterogeneous demand elasticities across firms, which are able to generate variable markups. However, the study goes a step further in documenting the interaction of consumer and firm heterogeneity in driving these variable markups. It shows that heterogeneity in consumer preferences — i.e., differences in their demand elasticities and preferences over quality — across income distribution translates into heterogeneity in markups charged by firms: lower demand elasticity of wealthier households allows larger firms to charge higher markups. While this demand-driven variable markups generates misallocation across firms, the losses from this misallocation are small. This is because firms that face low demand elasticities increase their markups in response to subsidies aimed at lowering their markups in the first place.

The study proposes a sufficient statistic, firms' pass-through rate, to correct bias in aggregate productivity gains from reallocation under endogenous markup adjustments by firms. Pass-through rates are documented to be decreasing in firms size, with the relationship stronger in quality-differentiated sector. These differences in pass-through rates are driven by both differences in slope of demand curve and market structure faced by firms. A methodology — motivated and backed by empirical evidence — is proposed that uses differences in pass-through rates across homogeneous and quality-differentiated sector to separately identify how differences in the slope of demand curve across firms affect their pass-through rates. The main finding is that gains from reallocation are lower by about 50 percent under demand-driven variable markups than when pass-through is assumed to be complete.

[13]Counterfactual scenarios are also created using the estimates of slope of demand and allowing competitiveness index to vary from least competitive environment (proxied by minimum competition within 2-digit industries in data) to most competitive environment (proxied by maximum competition within 2-digit industries in data). The aggregate productivity gains range from 8.5 percent under least competitive environment (fourth row in Table 1.8) to 22.5 percent in most competitive environment (fifth row in Table 1.8).

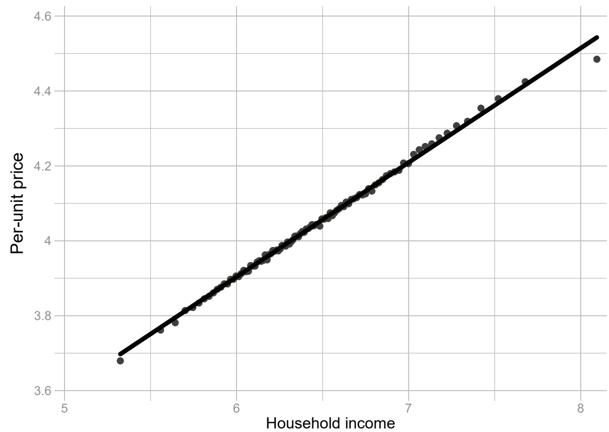
Like much of data available in developing countries, the study not directly observe the characteristics of consumers that buy from firms. Yet the paper shows that inferences on how consumer demand affects firms' prices — and its underlying components — can still be made by combining available production data for firms with natural experiments. With separate data on prices and quantities (rather than revenues), — and despite imposing minimal assumptions on demand or market structure faced by firms — differences in markups and marginal costs across firms and sectors, and how firms change their prices in response to changes in their costs can inform us to a great extent about sources behind firms' market power and how that affects aggregate productivity.

Figures and Tables

Figure 1.1. Firm's markups, marginal costs and size

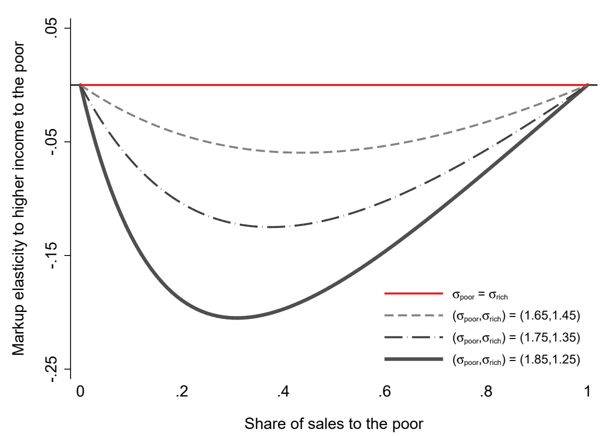
Notes: The figure shows the relation between firm's per-unit markups, marginal costs and labor force. The top panels shows the average relationship by firm-size, and the bottom panel shows the relationship by quality-differentiation using the definition in Rauch (1999). All variables are measured in logs. Both axes depict the residuals of a regression of dependent variable on district-by-product-by-year fixed effects. Each dot represents 1% of observations. Source: ASI

Figure 1.2. Household income and per-unit product prices



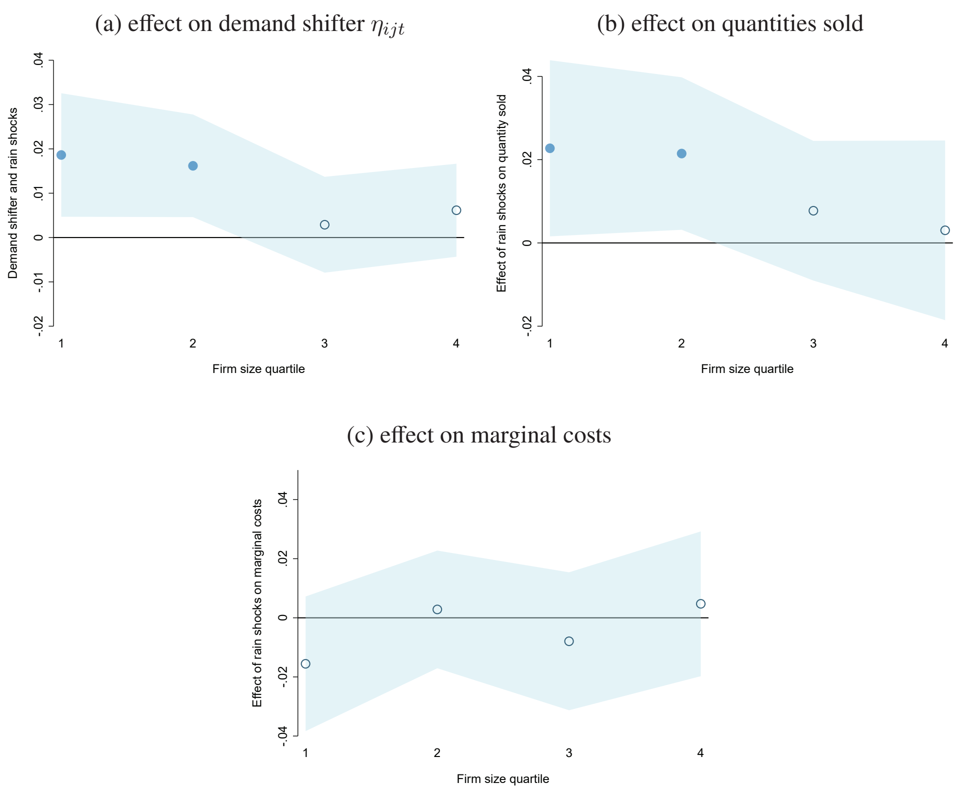
Notes: The figure shows the relation between log per-unit prices for manufactured goods paid by households and log household income. The y-axis depicts the residuals of a regression of log unit-level price on product-by-village-by-year fixed effects. The x-axis depicts the residuals of a regression of log household consumption (as proxy for income) on product-by-village-by-year fixed effects and household controls (including industry of occupation, type of occupation, religion and social group). Each dot represents 1% of observations. The figure uses NSS data for the year 1993.

Figure 1.3. Elasticity of markups to positive income shocks to poor (as function of share of sales made to the poor by firm)



Notes: The figure shows simulated relationship of elasticity of markups to positive income shocks to households that are more price-sensitive. It plots the relationship as a function of share of sales made by the firm to the poor households.

Figure 1.4. Effect of rain shocks on demand and costs (across firm-size distribution)



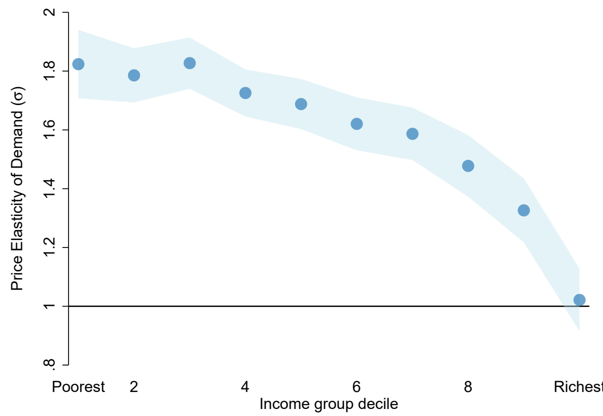
(a) effect on demand shifter η_{ijt}

(b) effect on quantities sold

(c) effect on marginal costs

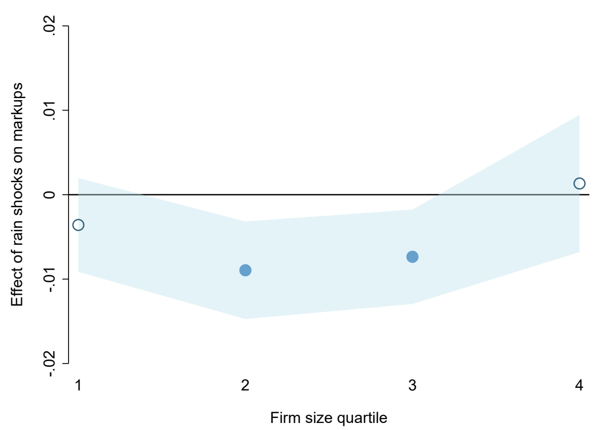
Notes: The figure shows the estimates of the effect of rain shocks on firms' demand and marginal costs across the firm-size distribution. All specifications control for firm age and size quartile-year fixed effects. 95% confidence intervals are represented by shaded blue area. Bold circles indicate results that are significant at the 10% level, and hollow circles statistically insignificant from 0 at the 10% level.

Figure 1.5. Estimates of price-elasticity of demand (σ) by income groups



Notes: The figure reports the estimates of price-elasticity of demand across income groups. The estimates are based on a IV-2SLS specification that instruments price of a good with state-level leave out mean price changes (described in Section 1.3.3). Source: NSS

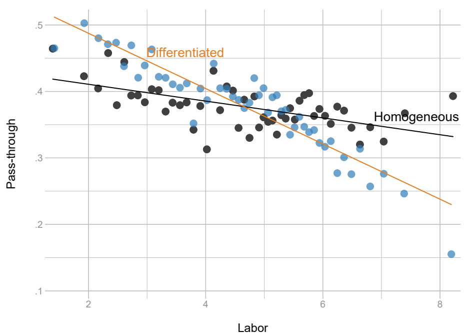
Figure 1.6. Effect of rain shocks on markups (across firm-size distribution)



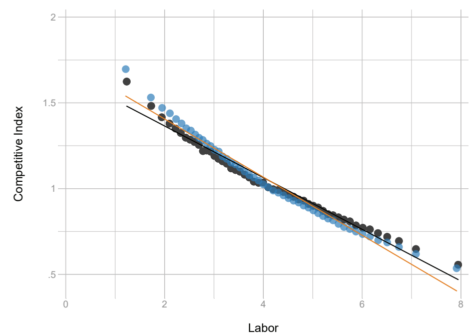
Notes: The figure shows the estimates of the effects of rain shocks on markups across the firm-size distribution. Specification controls for firm age and size quartile-year fixed effects. 95% confidence intervals are represented by vertical lines. Bold circles indicate results that are significant at the 10% level, and hollow circles statistically insignificant from 0 at the 10% level.

Figure 1.7. Pass-through rate, demand slope and competitiveness across firm size

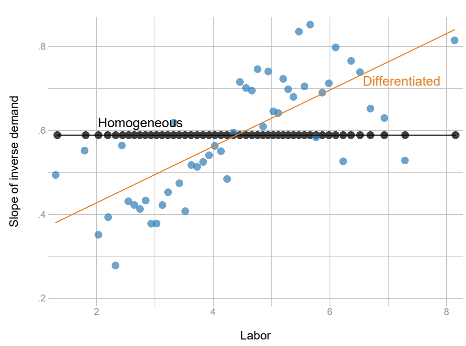
(a) Pass-through rate (ρ)



(b) Competitive index (ϕ)



(c) Slope of inverse demand (δ)



Notes: The figure shows the estimates of pass-through Γ_i (Panel (a)), competitiveness index ϕ_i (Panel (b)) , and slope of inverse demand δ_i (Panel (c)) as a function of firm-size.

**Table 1.1. Baseline Correlations:
Firm-size, markups and costs**

	Dependent variable: log of ...				
	Marg.Costs	Markups	Input Price	K/L	Wages
	(1)	(2)	(3)	(4)	(5)
(log) labor	0.041***	0.056***	0.063***	0.098***	0.189***
	[0.006]	[0.004]	[0.010]	[0.007]	[0.003]
Observ.	167,221	167,221	443,022	167,221	167,221
R-squared	0.870	0.638	0.410	0.656	0.803
Industry f.e.	✓	✓	✓	✓	✓
District-prod.-year f.e.	✓	✓	✓	✓	✓

Notes: The table reports the correlation between firm size (proxied by the labor force), and its markups, marginal costs and input factors. Column 1-2 reports the results on marginal costs and markup. Column 3-5 documents the relation between firm size and prices paid for its input factors: input material prices, capital-to-labor ratio and wages per-unit labor. All variables are measured in logs. The data is from the sample of manufacturing firms in India in the Annual Survey of Industries (ASI). Standard errors are clustered by district level are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

**Table 1.2. Baseline Correlations by quality-differentiation:
Firm-size, markups and costs**

	Dependent variable: log of ...				
	Marg. Costs	Markups	Input Prices	K/L	Wages
	(1)	(2)	(3)	(4)	(5)
(log) labor	-0.023***	0.077***	0.051***	0.073***	0.184***
	[0.008]	[0.003]	[0.007]	[0.010]	[0.004]
(log) labor ×	0.117***	0.009**	0.019**	0.046***	0.008**
different. good	[0.011]	[0.004]	[0.008]	[0.010]	[0.004]
Observations	167,221	167,221	443,022	167,221	167,221
Industry f.e.	✓	✓	✓	✓	✓
District-prod.-year f.e.	✓	✓	✓	✓	✓

Notes: The table reports the correlation between firm size (proxied by labor) and markups, marginal costs, for their products; as well as between firm size and factor prices (input prices, capital intensity and wages per-unit labor). All variables are measured in logs. The data is from the sample of manufacturing firms in India in the Annual Survey of Industries (ASI). Column 2 also includes control for firms' marginal costs. Standard errors are clustered by district level are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 1.3. Rainfall induced income shocks for poor population

	Dependent variable: log of ...				
	Agricultural output		Daily wages		
	Crop yield	Revenue per unit area	Rural agri. labor	Rural non-. -agri labor	Non-rural labor
	(1)	(2)	(3)	(4)	(5)
Shock$_{dt}$ (-1/0/+1)	0.045***	0.035***	0.027***	-0.009	0.017
	(0.005)	(0.005)	[0.008]	[0.009]	[0.011]
Observations	38,280	38,280	115,852	102,910	154,939
R-squared	0.887	0.853	0.516	0.271	0.124
District-crop f.e.	✓	✓			
Crop-year f.e.	✓	✓			
District f.e.			✓	✓	✓
Year f.e.			✓	✓	✓

Notes: The table reports the effect of rain shocks on agricultural productivity (Columns 1-2) and labor market wages (Columns 3-5). Standard errors are clustered by district level are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 1.4. Effect of rain shocks on firms' idiosyncratic demand and marginal costs

	Dependent variable:					
	Demand Shifter	log quantity	log of ... marg. cost	TFPQ	wage	input price
	(1)	(2)	(3)	(4)	(5)	(6)
Shock$_{dt}$ (-1/0/+1)	0.012**	0.014***	-0.004	-0.011	0.001	0.001
	[0.005]	[0.005]	[0.007]	[0.008]	[0.002]	[0.008]
Observations	133,094	133,094	133,094	59,965	102,541	239,100
R-squared	0.898	0.975	0.952	0.887	0.922	0.931
Firm f.e.				✓	✓	
Firm-product f.e.	✓	✓	✓			✓
Product-year f.e.	✓	✓	✓	✓	✓	✓

Notes: The table reports the effects of rain shocks on firms' demand shifters, log quantity sold, log marginal costs (and its underlying factors) based on specification1.6. Shock$_{dt}$ is defined as +1(-1) if the rainfall in the monsoon months is above(below) the 70$^{\text{th}}$(30$^{\text{th}}$) percentile of the district's usual distribution for monsoon rainfall. It takes the value of 0 if the rainfall is between 30$^{\text{th}}$-70$^{\text{th}}$ percentile of district's usual distribution. All columns include firm-product and product-year fixed effects. Standard errors are clustered by district level are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 1.5. Average effect of rain shocks on firms' markups

	Dependent variable: log markup							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Shock_{dt} (-1/0/+1)	-0.005***	-0.005***	-0.005***	-0.005***	-0.005***	-0.005**	-0.004*	-0.004*
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Observations	133,094	122,828	133,094	133,094	133,094	133,094	133,094	133,094
R-squared	0.989	0.990	0.989	0.989	0.989	0.989	0.989	0.989
Firm-product f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Product-year f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Controls	Baseline Specification	Single-plant establishment	+ Age control	+ Size-year control	Past 2-year shocks controls	National Market access control	In + out-state market access	(3)-(7) controls

Notes: The table reports the average effects of rain shocks on markups, based on specification 1.6. Shock_{dt} is defined as +1(-1) if the rainfall in the monsoon months is above(below) the 70^{th}(30^{th}) percentile of the district's usual distribution for monsoon rainfall. It takes the value of 0 if the rainfall is between 30^{th}-70^{th} percentile of district's usual distribution. All columns include firm-product, product-year fixed effects and control for log marginal costs. Standard errors are clustered by district level are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 1.6. Effects of rain shocks on markups across firm-size distribution

	Dependent variable: log markup							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Shock$_{dt}$ (-1/0/+1)								
× First size quartile	-0.003	-0.003	-0.003	-0.004	-0.004	-0.002	-0.001	-0.002
	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.004]	[0.003]
× Second size quartile	-0.009***	-0.010***	-0.009***	-0.009***	-0.010***	-0.008**	-0.007**	-0.008**
	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.004]
× Third size quartile	-0.007**	-0.005*	-0.007**	-0.007**	-0.008***	-0.006*	-0.005	-0.006*
	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]
× Fourth size quartile	0.001	0.000	0.001	0.001	0.001	0.003	0.003	0.003
	[0.003]	[0.004]	[0.003]	[0.004]	[0.003]	[0.004]	[0.004]	[0.004]
Observations	133,094	122,828	133,094	133,094	133,094	133,094	133,094	133,094
R-squared	0.989	0.990	0.989	0.989	0.989	0.989	0.989	0.989
Firm-product f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Product-year f.e.	✓	✓	✓	✓	✓	✓	✓	✓
Controls	Baseline Specification	Single-plant establishment	+ Age control	+ Size-year control	Past 2-year shocks controls	National Market access control	In + out-state market access	(3)-(7) controls

Notes: The table reports the heterogeneous effects of rain shocks on firms' markups. Shock$_{dt}$ is defined as +1(-1) if the rainfall in the monsoon months is above(below) the 70th(30th) percentile of the district's usual distribution for monsoon rainfall. It takes the value of 0 if the rainfall is between 30th-70th percentile of district's usual distribution. Q_i^r are dummy variables taking the value of 1 when firm i belongs to size quartile r within the industry k. All columns include firm-product, product-year fixed effects and control for log marginal costs. Standard errors are clustered by district level are reported in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 1.7. Estimates of pass-through rates

	Dependent variable: log price_{ijt}					
	OLS estimation			IV estimation		
	(1)	(2)	(3)	(4)	(5)	(6)
log mc_{ijt}	0.553***	0.678***	0.664***	0.700***	0.739***	0.855***
	[0.006]	[0.015]	[0.025]	[0.008]	[0.023]	[0.036]
log mc_{ijt} $\times$ log labor_{it}		-0.026***	-0.019***		-0.009*	-0.014*
		[0.003]	[0.005]		[0.005]	[0.008]
log mc_{ijt} $\times$ $1(\text{diff})_i$			0.030			-0.166***
			[0.031]			[0.046]
log mc_{ijt} $\times$ log labor_{it} $\times$ $1(\text{diff})_i$			-0.014**			0.004
			[0.007]			[0.010]
Observations	131,557	131,557	131,557	131,557	131,557	131,557
R-squared	0.408	0.411	0.411			
Kleibergen-Paap F-stat				8392.57	847.44	132.57
Firm-product f.e.	✓	✓	✓	✓	✓	✓
NIC4 - year f.e.	✓	✓	✓	✓	✓	✓

Notes: The table reports the pass-through rates. Column 1-3 report OLS estimates, and Column 4-6 report IV-2SLS estimates where firms' TFPQ is used as to instrument marginal costs. All columns include firm-product and 4 digit industry-year fixed effects. Standard errors are clustered by firm-level are reported in parentheses. Column 4-6 also report Kleibergen-Paap F statistic. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table 1.8. % change in aggregate productivity from reallocation

	Input used:	
	Material	Material, Labor, ... and Capital
	(1)	(2)
Complete pass-through ($\Gamma_i = 1$)	29.9%	31.0%
Incomplete pass-through (Γ_i estimated in data)	6.5%	5.7%
No differences in slope of demand curve ($\Gamma_i \equiv \Gamma_i(\phi_i, \delta^{\text{non-diff}})$)	12.4%	12.9%
Minimum competitiveness + no differences in demand slope ($\Gamma_i \equiv \Gamma_i(\phi^{\min}, \delta^{\text{non-diff}})$)	8.5%	8.8%
Maximum competitiveness + no differences in demand slope ($\Gamma_i \equiv \Gamma_i(\phi^{\max}, \delta^{\text{non-diff}})$)	22.2%	23.0%

Notes: The table reports gains from aggregate productivity using materials as the only input factor (Column 1), and using materials, labor and capital as the three input factors inputs (Column 2). The methodology to calculate the gains is described in Section 1.4.

2. SHOCKS AND TECHNOLOGY ADOPTION: EVIDENCE FROM ELECTRONIC PAYMENT SYSTEMS

A rich literature in economics has argued that coordination failures could be an important obstacle to the adoption of new technologies (Rosenstein-Rodan 1943; Carlton and Klamer 1983). Coordination failures arise when decisions to adopt a new technology are complements across users — that is, when the private value of adoption for each single user depends positively on adoption by other users (Katz and Shapiro 1985, 1986).[1] In these situations, expectations of low adoption can become self-fulfilling. While the possibility of coordination failures is theoretically well understood (Murphy et al. 1989; Matsuyama 1995), direct evidence of their importance is scarce. Using data on the adoption of a digital wallet technology during the 2016 Indian Demonetization, the study provides novel evidence on coordination failures in technology adoption, and studies the role that policy can play in addressing them.

There are two reasons why documenting the role of coordination failures in adoption of the digital wallet technology is useful. First, this product provides a clean test case for the general proposition that coordination failures can slow down technology adoption. Digital wallets are network goods; this makes adoption decisions complements across users, and creates scope for coordination failures (Katz and Shapiro 1994; Rysman 2007). Relative to other network goods, digital wallets are generally cheap and simple to adopt, which helps isolate the role of coordination problems. Second, digital wallets are a canonical example of financial technology ("fintech") products. The rapid diffusion of information technology over the past two decades has raised expectations about the potential for fintech to improve financial inclusion, particularly

[1] Katz and Shapiro (1985, 1986) highlight how externalities can arise both directly — when the number of users affects the quality of the product — or indirectly — in situations where the number of users affect the value of other add-on products due to compatibility (e.g. hardware/software) or post-purchase services (e.g. cars).

in developing countries, where fostering access to financial services remains a key goal for policymakers. Understanding the obstacles to their adoption is therefore also relevant to policy.

To better identify the role of coordination failures, the study analyzes adoption of the electronic wallet technology by retailers after the 2016 Indian Demonetization. This unexpected policy shock resulted in a large but temporary reduction in the availability of cash, leading to a temporary incentive to adopt the technology. The analysis is organized in three parts. First, a dynamic model of technology adoption with complementarities is developed and used to characterize the key features of the response of adoption of digital payments to a temporary shock to the availability of traditional means of payment. Second, merchant level data from the leading fintech payment system in India is used with quasi-exogenous variation in exposure to the Demonetization to test the model's predictions. Third, the contribution of complementarities to the overall adoption response is quantified by structurally estimating the model.

The main findings are the following. First, consistent with the model's predictions, the demonetization caused an adoption wave among merchants, characterized by three features: a persistent increase in the *size* of the platform, that is, the total number of merchants using it in transactions; a persistent increase in the platform's *adoption rate*, that is, the number of *new* merchants adopting the platform each month; and *state-dependence* in adoption, meaning that the long-run adoption response depends on the initial (pre-shock) strength of externalities. The latter two features are important, and it is shown that they are distinctive predictions of the model with externalities. Second, the quantitative estimation of the model shows that externalities were not only present, but played a large role: they account for approximately 60% of the long-run adoption response. Taken together, these results suggest that externalities — as opposed to pecuniary adoption costs or transaction costs — could be an important obstacle to the diffusion of fintech.

Externalities pose a challenge to policymakers interested in expanding the use of fintech, or technologies subject to adoption externalities in general, as large-scale interventions may be required to overcome the coordination problem. The bright side is that these interventions could be temporary and yet have persistent effects, as the Indian demonetization example suggests. On this point however, the analysis offers an important caveat. Because of the state-dependence mentioned above, temporary shocks — while they can boost *average* long-run adoption — can also exacerbate *differences* in long-run adoption across regions or industries. Furthermore, this effect can be offset by increasing the persistence of the shock. Thus, a policymaker that cares about inducing uniform adoption should lean toward more persistent interventions. Evidence of this mechanism is found in the data, and the study construct counterfactual experiments using the estimated model to show that, in the case of India, a more persistent but smaller intervention would have led to long-run adoption that would have been both higher on average across regions, and more evenly distributed.

The empirical setting of the paper is the Indian demonetization of 2016. On November 8th, 2016, the Indian government announced that it would void the two largest denominations of currency in circulation and replace them with new bills. At the time of the announcement, the voided bills accounted for 86.4% of the total cash in circulation. The public was not given advance warning, and the bills were voided effective immediately. A two-month deadline was announced for exchanging the old bills for new currency. In order to do so, old bills had to be deposited in the banking sector. However, withdrawal limits, combined with frictions in the creation and distribution of the new bills, meant that immediate cash withdrawal was constrained. As a result, cash in circulation fell and bank deposits spiked. Cash transactions became harder to conclude, but more funds were available for use in electronic payments. Importantly, though the shock was very large, it was also temporary, as things normalized for the most part by February 2017.

Section 2.1 starts by showing that the demonetization led to a large aggregate increase in the use of electronic payment systems. The focus is primarily on data from the largest Indian provider of non-debit card electronic payments. This payment platform operates as a digital wallet. The digital wallet consists of a mobile app that allows customers to pay at stores using funds deposited in their bank accounts. Payment is then transferred to merchants' bank accounts via the app. The costs associated with the adoption of this technology for merchants are small; there are no usage fees, and all that is required to join the platform is to have a bank account and a mobile phone, both of which had high ownership rates in India by 2016 Agarwal et al. (2017). Aggregate activity on the platform doubled in size several times during the two months following the demonetization announcement. Additionally, this adoption response was *persistent*, though the shock was not. There was no significant mean-reversion in the aggregate number of merchants using the technology or in aggregate transaction volumes once cash withdrawal constraints were lifted.

The aggregate evidence thus suggests that the *temporary* contraction in cash led to a *persistent* increase in adoption of fintech payments. However, this finding alone is not sufficient to establish that externalities played a role. First, theoretically, economic mechanisms other than externalities may also generate persistent responses to transitory shocks. Second, empirically, long-term responses in aggregate event studies are potentially confounded by subsequent aggregate shocks.

In order to address the first issue, the study characterizes further the testable implications of externalities by studying a dynamic technology adoption model. The model builds on the framework of Burdzy et al. (2001). Firms face a choice between two technologies, one of which (the "platform") is subject to positive externalities — the flow profits from operating under this technology increases with its rate of use by other firms. Moreover, the relative benefit of adopting the platform is subject to aggregate shocks. The presence of these common shocks helps eliminate the potential equilibrium multiplicity arising from complementarities in adoption decisions.

The model predicts that following a large, *temporary* shock, the total number of firms using the platform increases *persistently*, consistent with the aggregate evidence of Section 2.1. However, it delivers two additional predictions. First, the presence of externalities — on top of increasing the *size* of the platform — also generates a persistent increase in its *adoption rate*. In other words, the number of *new* firms joining the platform every period remains high even after the shock has dissipated. The reason is that, with externalities, the initial adoption triggered by the shock, by temporarily raising the size of the platform, increases the relative future value of adoption for other firms. This "snowball" effect can generate endogenous persistence in the increase in adoption rates and is unique of a technology with externalities. Second, the model predicts that adoption responses exhibit state-dependence: the long-run adoption response depends on the pre-shock adoption rates. The intuition for this result is simple: higher initial externalities, in the form of higher initial adoption rates, make it easier to reach the "tipping point" beyond which the platform has sufficient critical mass to grow without the shock.

The study then shows that the empirical predictions of externalities highlighted above are consistent with the adoption responses observed in the data after the demonetization. In order to do this, an empirical design is provided to estimate the causal impact of the cash contraction on adoption, both in the short- and medium-run. The empirical design exploits variation across districts in the importance of chest banks — local bank branches in charge of the distribution of new currency — to identify variation in exposure to the shock. As Section 2.3 discusses, this design allows to isolate the effect of the cash contraction from other effects of the demonetization, therefore overcoming some of the limitations of the aggregate evidence of Section 2.1. The districts that were more exposed to the cash crunch also experienced a larger and more persistent increase in total adoption following the demonetization, the first prediction of the model. Crucially, higher exposure also predicts a larger increase in the number of *new* firms joining the platform, even after restrictions on cash withdrawals are lifted — the second prediction of the model. Finally, the study finds robust evidence consistent with state-dependence, the third prediction of the model. Districts where the pre-shock value of joining the platform is higher — either because pre-shock adoption was high, or because they were located close to other high adoption cities — are characterized by a higher average adoption response. The same pattern holds at a disaggregated level: a firm's choice to use the technology is positively affected by the rate of adoption of firms in the same local industry. This latter result does not simply capture variation across locations and industries, since it holds conditional on these fixed effects interacted with time.

Altogether, this reduced-form evidence shows that a model with adoption externalities can account for the *qualitative* features of the adoption response caused by the demonetization. However, it is silent about the *quantitative* contribution of externalities to the adoption response. In order to address this issue, Section 2.4 estimates the dynamic adoption model of Section 2.2 via simulated method of moments, using the data on fintech payments. The key parameter of interest is the size of adoption externalities. Following the intuition described above, it is shown that this parameter can be identified using the difference between short- and long-run adoption rates following the shock.

The estimates of the model is used to provide two main results. First, externalities are quantitatively important in understanding the total adoption response: they account for approximately 45% of the total response of adoption to the demonetization, in the sense that the medium-run adoption rate would have been 45% lower (and declining), had the technology featured no externalities in adoption. Second, the persistence of the shock is crucial to understanding its effects, both in terms of *average* adoption, and for the *variance* of adoption across regions. As discussed earlier, temporary interventions may increase overall adoption. However, because of state-dependence, they can also exacerbate initial differences in adoption. Consistent with this intuition, keeping the present value of the decline in cash constant, a cash crunch with a smaller initial magnitude (by around 40%) but a longer half-life (by a factor of 2), would have led to higher long-run adoption rates (by about 10%) and lower dispersion. Thus, an implication of the model is that policymakers with a preference for uniform adoption across regions or industries should generally favor smaller but more persistent interventions.

The rest of the paper is organized as follows. Section 2 provides some background on the demonetization and documents aggregate adoption effects. Section 3 analyzes the dynamic adoption model and derives key predictions. Section 4 tests these predictions in the electronic wallet data. Section 5 estimates the model and provides counterfactuals. Section 6 concludes.

2.1 Background

2.1.1 The demonetization

On November 8, 2016, at 08:15 pm IST, Indian Prime Minister Narendra Modi announced the demonetization of Rs.500 and Rs.1,000 notes during an unexpected live television interview. The announcement was accompanied by a press release from the Reserve Bank of India (RBI), which stipulated that the two notes would cease to be legal tender in all transactions at midnight on the same day. The voided notes were the largest denominations at the time, and together they accounted for 86.4% of the total value of currency in circulation. The RBI also specified that the two notes should be deposited with banks before December 30, 2016. Two new bank notes, of Rs.500 and Rs.2,000, were to be printed and distributed to the public through the banking system. The policy's stated goal was to identify individuals holding large amounts of "black money", and remove fake bills from circulation.[2]

However, the swap between the new and old currency was not immediate, and the public was unable to withdraw cash at the same rate as they were depositing old notes. As a result, the amount of currency in circulation dropped precipitously during the first two months of the demonetization period. This can be seen in Figure 2.1, which plots the monthly growth rate of currency in circulation. Overall, it declined by almost 50% during November and continued declining in December.

This cash crunch partly reflected limits on cash withdrawals put in place by the RBI in order to manage the transition. But it was also driven by the difficult logistics of the swap itself. In order to ensure that the policy remained undisclosed prior to its implementation, the RBI had not printed and circulated large amounts of new notes beforehand. This caused many banks to be unable to meet public demand for cash, even under the withdrawal limits.

Importantly, the demonetization did not lead to a reduction in the total money supply, defined as the sum of cash and bank deposits. The total money supply was stable over this period, as reported in Figure 2.1. In its press release, the RBI highlighted that deposits to bank accounts could be freely used through "various electronic modes of transfer." The public was thus still allowed to transact using any form of noncash payment, such as cards, checks, or any other electronic payment method; cash transactions were the only ones to be specifically impaired.

[2]In its annual report for 2017-2018, the RBI reported that 99.3% of the value of voided notes had been deposited in the banking system during the demonetization.

Despite its magnitude, the cash crunch was a temporary phenomenon. Overall, things significantly improved in January and essentially normalized in February. Cash in circulation grew significantly again in January 2017, suggesting that the public was able to withdraw cash from banks (see Figure 2.1). Furthermore, by January 30th, 2017, the Government lifted most of the remaining limitations on cash withdrawals, in particular removing any ATM withdrawal limit from current accounts.

The demonetization thus had three key features relevant to the analysis. First, it led to a significant contraction of cash in circulation. Second, it did not change the total money stock, that is, the sum of cash and deposits. As a result, the public could still access and use money electronically once the notes had been deposited. Third, it was relatively short-lived: the cash shortage was particularly acute in November and December, it improved with the new year, and generally normalized with February.

2.1.2 Fintech payment systems during the demonetization

Overall, the demonetization was associated with a large uptake in electronic payments. This fact is illustrated using data from one of the leading digital-wallet companies in the country. The company allows individuals and businesses to undertake transactions with each other using only their mobile phone. To use the service, a customer needs to download an application and link their bank account to the application. Merchants can then use a uniquely assigned QR code to accept payments directly from the customers into a mobile wallet. The contents of the mobile wallet can then be transferred to the merchant's bank account.

Importantly, adoption and usage costs of this technology are low. No investment in a point of sale terminal is required; the retailer simply needs a cellphone and a bank account, which are both very common in India Agarwal et al. (2017). Furthermore, for small and medium-sized merchants — who make up the bulk of the data —, transactions using the technology do not involve fees.

Figure 2.2 reports data for the total number and total value of transactions executed by merchants using this technology. In the months before the demonetization, the weekly growth in the usage of the wallet technology had been positive on average but relatively modest. However, after the demonetization, the shift towards this payment method was dramatic. In particular, in the first week after the shock, the number of transactions grew by more than 150%, and the value of transactions increased by almost 200%. For the first month after the shock, weekly growth rates were consistently around 100%.

	No frictions $(C=0, \kappa=0)$	Fixed costs $(C=0, \kappa>0)$	Externalities $(C>0, \kappa=0)$
P1: Persistent increase in size of user base	✗	✓	✓
P2: Persistent increase in adoption rate	✗	✗	✓
P3: Positive dependence on initial adoption	✗	✗	✓

Table 2.1. Predictions across versions of the model.

Crucially, this initial positive effect of the demonetization on adoption did not dissipate, even after the cash-availability constraints were relaxed. The data show a slow-down in aggregate growth starting in January, which is when the limits on the circulation of new cash started to be lifted. However, after a small negative adjustment in early February, the average growth rate over the next two months remained positive, indicating that users did not abandon the platform as cash became widely available again. In other words, a temporary decline in the availability of cash led to a permanent increase in the usage of the platform.

The data shared with by the electronic wallet company end in June 2017. However, it is important to point out that the increase in electronic-wallet technologies in India also persisted after this period. According to the official estimates by the Reserve Bank of India, monthly mobile-wallet transactions increased from 75 million to over 300 million between September 2016 and March 2017, which is the central period in the analysis. In March 2019, monthly transactions still totaled around 385 million, suggesting that adoption has not dissipated since the demonetization.

2.2 Theory

This section analyzes a dynamic model of technology diffusion. The model, discussed in extensive details in Crouzet et al. (2021) makes a number of specific empirical predictions, which depend on whether the technology features adoption externalities, and are summarized in Table 2.1. Section 2.3 tests these predictions using the adoption of the electronic wallet technology after the demonetization. Section 2.4 uses the model to quantify the contribution of externalities to the adoption response and to produce counterfactual responses to the demonetization shock.

2.2.1 Predictions of the model from the effects of a cash crunch

The effects of a sudden, unanticipated reduction in cash are discussed. The section starts by discussing its effects in a version of the model where there are no fixed adoption costs, and externalities are the only potential barrier to adoption. Three key testable implications are highlighted: a persistent effect of the shock on the *size* of the user base; a persistent effect on the *adoption rate*, that is, the number of new users joining the platform each period; and a dependence of long-run responses on initial adoption rates. Next, the section discusses qualitatively three key features of the dynamics that arise from the model.

Prediction 1: persistent increase in the size of the user base. The negative shock to cash demand can have persistent effects on the total number of firms using the technology. If the initial number of adopters is sufficiently high, it can be the case that, after the shock, users of electronic payments does not converge back to initial level, but instead, converges to 1.

Prediction 2: persistent increase in adoption rate. Importantly, on the blue trajectory, firms with the option of changing technology always opt for e-payments, even long after the shock has dissipated. Thus, with complementarities, the shock should lead not only to a persistent increase in size of the platform, but also in its adoption rate, that is, the flow of new users into the platform each period.

Prediction 3: positive dependence on initial adoption. Finally, these adoption effects are stronger and more persistent, the higher the initial level of adoption. Thus, the model features *positive* state-dependence with respect to initial adoption rates.

2.3 Adoption dynamics

This section uses micro data from the leading electronic wallet in India to test the three empirical predictions of the model with externalities highlighted in Section 2.2. The first two predictions are tested on the long-run increase in both the *size* of the platform and in its*adoption rate*, by using quasi-random variation in the exposure to the shock. Additionally, district-level evidence consistent with the third prediction — the positive dependence of adoption responses on baseline adoption rates — is provided.

2.3.1 Data

The main data used in the analysis are merchant-level transactions from the leading digital-wallet companies in the country. Data is observed at weekly level on the sales amount and number of transactions happening on the platform for anonymized merchants between May 2016 and June 2017. For each merchant, the location of the shop at the district level is also observed, as well as the store's detailed industry. For a random sub-sample of shops, the location is provided at the more detailed level of 6-digit pincode. There are two key features of these data. First, the information is relatively high frequency, since the data can be aggregated at weekly or monthly levels. Second, the transactions are geo-localized, therefore allowing to aggregate them up at the same level as other data sources used in this study.

Data on district-level banking information is obtained from the RBI. This includes three pieces of information at the district level: first, the number of bank branches; second, information on the number of currency chests by district and the banks operating the chests; third, quarterly bank deposits at the bank-group level in each district. Finally, this data is complemented with information from the Indian Population Census of 2011 to obtain a number of district-level characteristics, including: population, quality of banking services (share of villages with an ATM and banking facility, number of bank branches and agricultural societies per capita), socioeconomic development (sex ratio, literacy rate, growth rate, employment rate, share of rural capital), and other administrative details, including distance to the nearest urban center.

2.3.2 The effects of demonetization on adoption

Next, the first two predictions of the model are tested: the long-run increase in both the *size* of the platform and its *adoption rate*. The aggregate event study evidence discussed in Section 2.1 is qualitatively consistent with these predictions. At the same time, this aggregate event study evidence may not properly capture the long-run causal response of adoption to the shock. One particularly important confounding factor are national government policies that may have affected the subsequent adoption of electronic payments for reasons unrelated to externalities. This concern is addressed by using quasi-random variation across different districts in exposure to the cash contraction. This approach allows to recover the causal effect of the temporary cash contraction on adoption of electronic payments independently of any other aggregate shocks after the demonetization.

Exposure measure. To identify heterogeneity in the exposure to the cash contraction, the heterogeneity across districts in the relative importance of chest banks — defined as banks operating a currency chest in the district — in the local banking market is exploited. In the Indian system, currency chests are branches of commercial banks that are entrusted by the RBI with cash-management tasks in the district. Currency chests receive new currency from the central bank and are in charge of distributing it locally. While the majority of Indian districts have at least one chest bank, districts differ in the total number of the chest banks, as well as in chest banks' share of the local deposit market.

Consistent with anecdotal evidence, it is expected that districts where chest banks account for a larger share of the local banking market should experience a smaller cash crunch during the months of November and December. On some level, this relationship is mechanical. Chest banks were the first institutions to receive new notes, so in districts where chests account for a larger share of the local banking market, a larger share of the population can access the new bills faster. Furthermore, the importance of chest banks may be an even more salient determinant of access to cash if these institutions were biased toward their own customers or partners. Indeed, concerns of bias in chest-bank behavior were widespread in India during the demonetization. In any case, the study show that this connection between chest bank presence and the cash contraction is supported by data.

To measure the local importance of chest banks, data on the location of chest banks is combined with information on overall branching in India and data on bank deposits in the fall quarter of the year before demonetization (2015Q4). Ideally, one would like to measure the share of deposits in a district held by banks operating currency chests in that district. However, data on deposits are not available at the district level for each bank. Instead, the data are only available at the bank-type level.[3] Since the information on the number of branches for each bank at the district level is available, the study proxies for the share of bank deposits of each bank by scaling the total deposits of the bank type in the district by the banks' share of total branches in that bank type and district.[4] The score can then be computed as share of total banking deposits in a district that were held by branches of banks that also operated currency chests in that district. Since the instrument is better to be interpreted as a measure of the strength of the shock, the final score Exposure_d is simply the converse of the above described measure. The score is characterized by a very smooth distribution centered on a median around 0.55, with large variation at both tails. Overall, exposure appears to be evenly distributed across the country, as very high and very low exposure districts can be found in every region. Consistent with this idea, in the robustness section it is shown that results do not depend on any specific part of the country.

According to the logic of the approach, it is expected that areas where chest banks are less prominent — or have higher exposure according to the index — to have experienced a higher cash contraction during the months of November and December. While one cannot directly observe the cash contraction at the local level, deposit data can be used to proxy for it. Cash declined because old notes had to be deposited by the end of the year, but withdrawals were severely limited. Therefore, the growth in deposits during the last quarter of 2016 should proxy for the cash contraction in the local area. Using this proxy for the cash crunch, evidence that supports the intuition behind the identification strategy is provided. Figure 2.3 shows that there is a strong relationship between district-level exposure to the shock and deposit growth. The same

[3]The RBI classifies banks in six bank groups: State Bank of India (SBI) and its associates (26%), nationalized banks (25%), regional rural banks (25%), private sector banks (23%) and foreign banks (1%).

[4]A simple example may help. Assume that one is trying to figure out the local share of deposit by banks A and B, both rural banks. It is known that rural banks in aggregate represents 20% of deposits in the district, and it is known that bank A has 3 branches in the district, while bank B only has one. The method will impute bank A's share of deposits to be 15%, while bank B's will be 5%.

relationship holds when using different measures of deposit growth and including district-level controls.

Econometric model. Using this measure of exposure, for different outcome variables of interest, y, the following difference-in-difference model is estimated:

$$\log\left(y_{d,t}\right) = \alpha_t + \alpha_d + \delta\left(\text{Exposure}_d \times \mathbf{1}_{\{t \geq t_0\}}\right) + \Gamma_t' Y_d + \epsilon_{d,t}, \tag{2.1}$$

where t is time (month), d indexes the district, t_0 is the time of the shock (November 2016), and Exposure_d is the measure of the district's exposure constructed with chest-bank data, as explained above. The equation is estimated with standard errors clustered at the district level, which is the level of the treatment Bertrand et al. (2004). Lastly, the specification is based on the data between May 2016 and June 2017.[5]

Results. Table 2.2 shows that districts more exposed to the cash contraction also experienced higher adoption of electronic payments. Column 1 shows that districts that were more exposed to the shock saw a larger increase in the amount transacted on the platform in the months following the demonetization. This result is both economically and statistically significant. Districts with one standard deviation higher exposure experienced a 55% increase in the amount transacted on the platform relative to the average. Similarly, the number of firms operating on the platform — the main measure of adoption — increased by 20% more in districts with one standard deviation higher exposure to the shock (Column 2).

Figure 2.4 (first two panels) plots the dynamics of the main effect, i.e. the month-by-month estimates of how districts characterized by different levels of exposure responded to the shock. This figure highlights three main findings. First, it confirms that the main effect is not simply driven by differential trends across high- vs. low-affected areas. Second, the shift in adoption across districts happened as early as November. Third, the difference in the response also persists after the cash availability returns to normal level. In particular, the effects are still large and significant after the month of February. These findings, taken together with the aggregate-level evidence in Section 2.1, confirms that the temporary cash contraction led to a persistent increase in size of the user base of the electronic payment technology, consistent with the first prediction of the model.

[5] Sparsely populated northeastern states and union territories are excluded from the analysis due to missing information on either district-level characteristics or banking variables. Altogether these regions account for 1.5% of the Indian population.

Next, the second prediction of the model is tested, which is that the shock led to a persistent increase in the adoption rate, that is, the flow of new users in the platform. This is empirically tested by analyzing whether districts more affected by the shock witnessed a more persistent increase in new adopters. THe new adopters at time t are defined as the firms using the technology for the first time at time t. The third panel of Figure 2.4 shows that districts experiencing a larger contraction in cash saw a larger increase in new adopters joining the platform as early as on November 2016. Crucially, the relative increase in the number of new adopters continued even after January 2017, the last month during which cash withdrawal was constrained, and persisted for the whole of spring 2017. This persistent increase in new users is consistent with the second prediction of the model.

2.3.3 State-dependence in adoption

One of the key predictions of the model with complementarities is the state-dependence of adoption. In particular, the model suggests that a temporary shock may lead to a permanent shift in adoption, but that the increase in adoption will not be uniform across regions: it will crucially depend on the initial strength of complementarities in the area. This section uses the data on electronic payments to present evidence that are consistent with this prediction.

This section tests how the increase in adoption differs depending on the distance between a district and areas in which the usage of there electronic wallets was large *prior* to November (*hubs*). The mapping between the strength of complementarities and distance to the electronic payment hub is intuitive. In the model, the heterogeneity in the strength of complementarities is completely determined by the number of users in the same area. In reality, individuals move across districts and therefore the size of adoption in neighboring districts will also be important. Therefore, being located close to a large hub — a center where electronic payment use is relatively common — may significantly increase the benefits of adoption Comin et al. (2012).[6]

This is implemented by running a simple difference-in-difference model where the usage of wallet technologies is compared around the demonetization period across districts that are differentially close to a digital wallet hub. Despite the clear advantages relative to the naive evidence presented above, there still are two concerns with this approach. First, by sorting on distance might capture variation coming from areas that are located in more extreme or remote parts of the country. Second, since the electronic hubs are some of the largest and most important

[6]In particular, district is defined to be an electronic payment hub if there were more than 500 active firms pre-demonetization (September 2016). The results are essentially identical if a threshold of 1,000 firms is used to define the hub districts. The nine hubs are spread evenly across the country. In particular, these districts are: Delhi, Chandigarh and Jaipur (North), Kolkata (East); Mumbai and Pune (West); Chennai, Bangalore and Rangareddy (South). The distance to the hub is defined as the minimum of the distance between the district and all the hubs.

cities in the country, one should expect that being located close to them will have benefits that go beyond the effect of complementarities.

These limitations are dealt in three ways. First, the comparison to districts is limited that are located within the same state, adding state-by-month fixed-effects. In this way, only distance variation between areas that are already located in similar parts of the country is exploited. Second, the distance to the capital of the state, also interacted with time effects is controlled for. This control allows to isolate the effect of the distance to a major electronic payment hub from the effect of being located close to a large city. Third, as in the previous analyses, the specification is augmented with a wide set of district-level covariates interacted with the time dummies. This implies a specification of the following form:

$$X_{d,s,t} = \alpha_{st} + \alpha_d + \delta \left(D_d \times \mathbf{1}_{\{t \geq t_0\}} \right) + \gamma \left(\tilde{D}_{d,s} \times \mathbf{1}_{\{t \geq t_0\}} \right) + \Gamma_t' Y_d + \epsilon_{d,t}, \tag{2.2}$$

where t indicates time, defined at the monthly level in this analysis, d indexes the district and s identifies the state of the district. D_d is the district's distance to the nearest electronic-wallet hub and $\tilde{D}_{d,s}$ is the district's distance to the capital district of the state. As before, standard errors clustered at district level. The main coefficient of interest is δ — which provides the difference in the level of adoption pre- and post-demonetization depending on how far the district is from its closest electronic-wallet hubs.

These results are reported in Table 2.3. Across all the outcomes — the amount of transactions, number of operating firms and number of new adopters — the districts farther away from major hubs experienced a lower increase in the post-demonetization period. The most conservative of the estimates indicates that a 50km increase in distance translates into a 19% lower increase in the amount of transactions. Importantly, these effects are not driven by differential trends in adoption between areas that are closer and further from hub cities (Figure 2.5). In general, distance does not matter before November, but it predicts differential responses starting in December.

2.3.4 Discussion

The key takeaway from this section is that the demonetization caused an adoption wave with features that are *qualitatively* consistent with three predictions of the model with externalities: a persistent increase in the *size* of user base; a persistent increase in the *adoption rate*, that is, the flow of new users into the platform; and state-dependence in responses, that is, a positive relation between initial adoption rates and the initial strength of adoption externalities, broadly defined. Because these predictions are *specific* to the presence of externalities, the reduced-form results of this section thus support the notion that externalities played a role in shaping the adoption response following the shock.

However, these reduced-form results also leave open two related questions. First, they do not indicate how large these externalities are in the data. Second, from these reduced-form results, no conclusions can be drawn about the effects of alternative shocks (say, a more transitory cash contraction), or about how a policymaker seeking to promote adoption, for instance, should structure a policy intervention. In other words, the reduced-form results of this section are silent about what can be learned from the demonetization experiment for adoption of similar technologies in other contexts. These two issues are connected: model implications depend on the strength of externalities. The next section addresses them both by structurally estimating the model using the data on electronic wallet adoption, and studying the model's counterfactual and policy implications.

2.4 Quantifying the role of externalities in the adoption decision

2.4.1 Estimation

The simulated method of moments is used to estimate the key parameters of the model. Details for the methodology can be found in Crouzet et al. (2021). The section below discusses the results and the counterfactual exercises.

Results Table 2.4 reports estimates of the five structural parameters. The point estimate for the size of the shock is 21.5% (with a 90% coverage interval of $[13.2\%, 29.7\%]$). The parameter expresses the decline in profits associated with cash-based transactions, relative to their long-run mean. Next, the point estimate of C is 0.062, with a 90% coverage interval of $[0.047, 0.076]$. The findings therefore reject the null of no adoption complementarities. The point estimates imply that relative to cash, profits under the electronic technology are on average 2.6% lower if there are no other adopters, and 3.6% higher if there is full adoption.

2.4.2 Counterfactuals

Next, the estimated model is used to construct the quantitative answer to three questions about the effects of the shock, and the role played by complementarities in the adoption process.

1. How would adoption have responded, in the absence of complementarities? Figure 2.6 reports empirical and model-based paths of average adoption across districts, in the aftermath of the shock. At the point estimates reported in Table 2.4, adoption rises by approximately 4p.p. by the end of December, and 6.5p.p. by the end of May, in line with the empirical estimates. This result is not surprising, since these moments were explicitly targeted. The figure also reports a counterfactual path of adoption rates, under the assumption that there are no complementarities, that is, when $C = 0$. With respect to the data, and to the baseline estimate, the adoption path is similar during the first three months, when the cash crunch is still ongoing. After that, it diverges from the data and from the model with complementarities, declining in the medium-run. The gap is fairly substantial: the predicted increase in adoption rates without complementarities would have been 3p.p. (or approximately 45%) lower than observed. Thus, the model attributes a important share of the response of adoption rates to complementarities.

2. What if the cash swap had been completed more quickly? Figure 2.6 also reports counterfactual adoption paths which speak to the role of the size and persistence of the shock. First, adoption paths are constructed under the assumption that a 90% decay rate of the shock is one month, instead of three months; this captures an alternative world in which the cash swap would have been executed as rapidly as initially intended. Under this scenario, adoption would only have risen by approximately 1p.p., and the increase in the dispersion of adoption would have been negligible. Figure 2.6 also indicates that, if the shock had been smaller in magnitude — which could capture a situation in which only one denomination would have been replaced, for instance — the long-run response would have been smaller. With a shock half as large, the average adoption rate only rises by approximately 4.5p.p., versus 6.5p.p. in the baseline case. The model thus suggests that the persistence and size of the cash crunch might have had substantial, though unintended, positive effects on adoption overall.

3. What sort of intervention maximizes long-run adoption? Next, the model is used to ask whether a hypothetical policymaker could have achieved higher long-run changes in adoption rates by implementing the cash swap differently. Table 2.5 reports the numerical solution under different values of aversion to dispersion in adoption rates, $g = 0$. Additionally, the first column reports the model estimates of the size and persistence of shocks, and the implied long-run first and second moments of the change in adoption rates.

Results in Table 2.5 show that a policymaker who cares about dispersion has a motive to *further* increase the persistence of the intervention. Compared to $g = 0$, the "constrained optimal" plan with an aversion to dispersion of $g = 0.6$ is associated with a shock that is smaller (by about 1-10.7/14.2=25%) but more persistent (by about 1-10.7/14.2=26%). Thus, while the size and persistence of the shock had positive effects on long-run adoption — as discussed above —, the model also suggests that if the objective of the policy had been to increase long-run adoption while minimizing the dispersion in outcomes across districts, a more persistent but smaller intervention would have been preferable. That said, long-run adoption gains under these alternative policies are relatively mild, in the order of 25% to 35% of the long-run adoption increase implied by the baseline estimates.

The analysis of this section has shown that the dynamic model with adoption complementarities can account well for key moments of the data. Counterfactuals suggest that complementarities account for 45% of the medium-run response of adoption, and that a smaller, but more persistent intervention may have led to a larger increase in long-run adoption rates, along with a lower long-run dispersion in adoption across districts.

2.5 Conclusion

An increasing number of new technologies feature network externalities. When this is the case, the technology's ability to grow and scale is subject to coordination frictions. How can this coordination friction be overcome? Furthermore, how can a policy intervention help to foster adoption? This study used the Indian demonetization of 2016, and its subsequent effect on the adoption of electronic wallets, as a laboratory to study these questions.

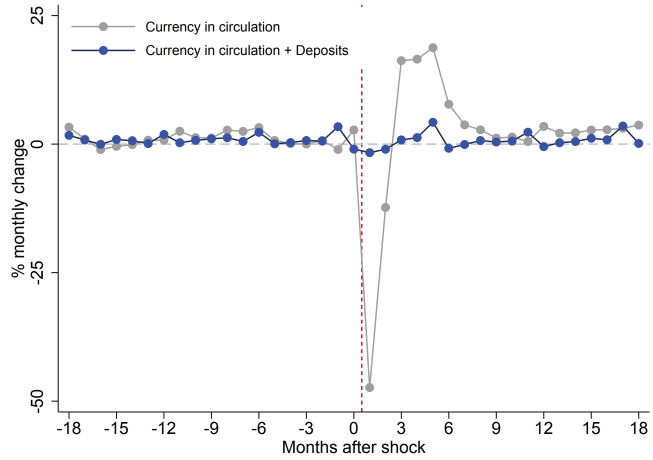
The study started by showing that the demonetization led to a large and persistent increase in the overall use of this technology, even though the demonetization shock itself was temporary. This large and persistent increase is argued to be consistent with a dynamic technology adoption model with externalities, and some additional testable predictions unique to externalities are derived. In particular, the study showed that in this model, a temporary shock can cause a persistent increase in the adoption rate of the platform (as opposed to only its size), and that the response of adoption rates depends positively on initial adoption levels.

Using micro data on electronic payments, the study showed that these additional testable predictions are supported by the data. At the the district level, a novel identification strategy is proposed based on heterogeneity in the presence of chest banks to estimate the causal impact of the cash crunch. The results show that the cash crunch caused a persistent increase in the adoption rate by firms of electronic wallets. Additionally, the adoption responses are characterized by positive state-dependence, both at the district and the firm level. Finally, a structural estimation of the dynamic model is provided. This estimation suggests that about 45% of the total adoption response is due to complementarities.

The analysis also highlighted some of the challenges faced by policymakers in environments with complementarities. In those environments, large, punctual interventions can have permanent effects on adoption because they effectively act as coordinating devices that help firms overcome coordination frictions. However, because of state-dependence, an intervention that is too brief can also exacerbate inequality in adoption rates. Policymakers may therefore face a trade-off between the length the intervention and how much it will exacerbate initial difference in adoption rates.

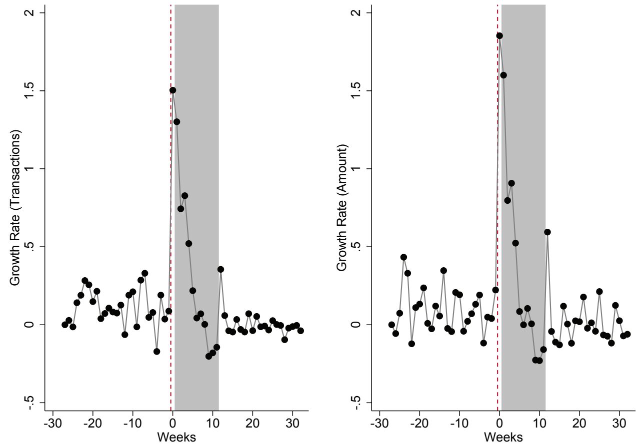
Figures and Tables

Figure 2.1. Change in nominal value of currency in circulation



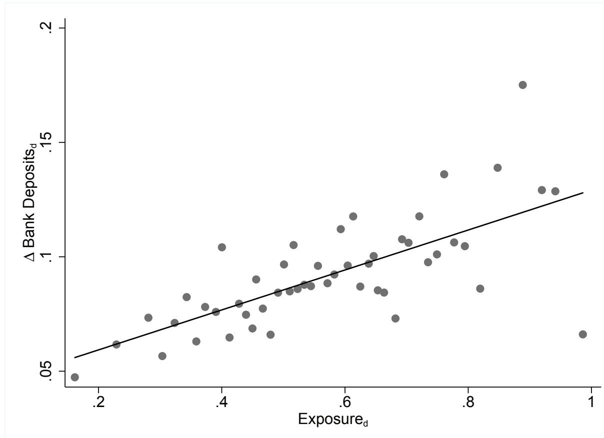
Notes: The figure shows the change in the nominal value of the stock of currency in circulation (in grey) and change in the value of the total money supply (in blue) in India. Month 0 is the month of October 2016; the figures are end-of-month estimates. Source: Reserve Bank of India.

Figure 2.2. Amount and Transaction Growth on Mobile Payment Platform



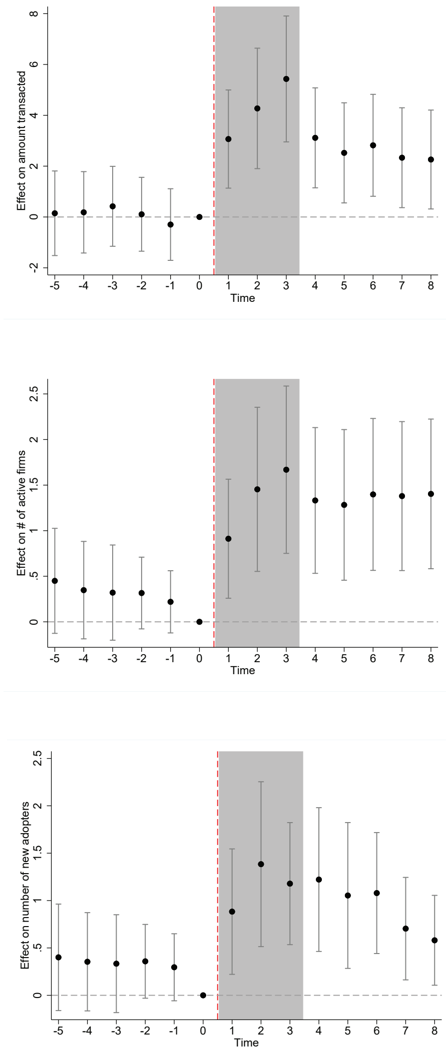
Notes: Week-over-week growth rate in the number of transactions (left panel) and total amounts (right panel) on the electronic wallet platform. The dashed red line indicates the week of November 8th, 2016.

Figure 2.3. Relation between Exposure and 2016 Q4 deposit growth



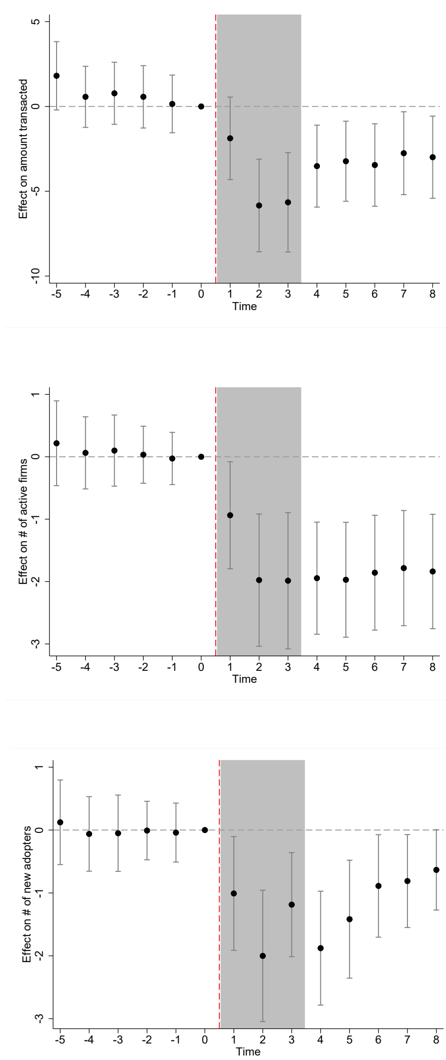
Notes: The figure shows the relation between our measure of Exposure$_d$ (as described in Section 2.3) and the change in bank deposits in the district between September 30, 2016 and December 31, 2016 $i.e.$ during the quarter of demonetization. Source: Reserve Bank of India.

Figure 2.4. District adoption dynamics in electronic payments data based on exposure to shock



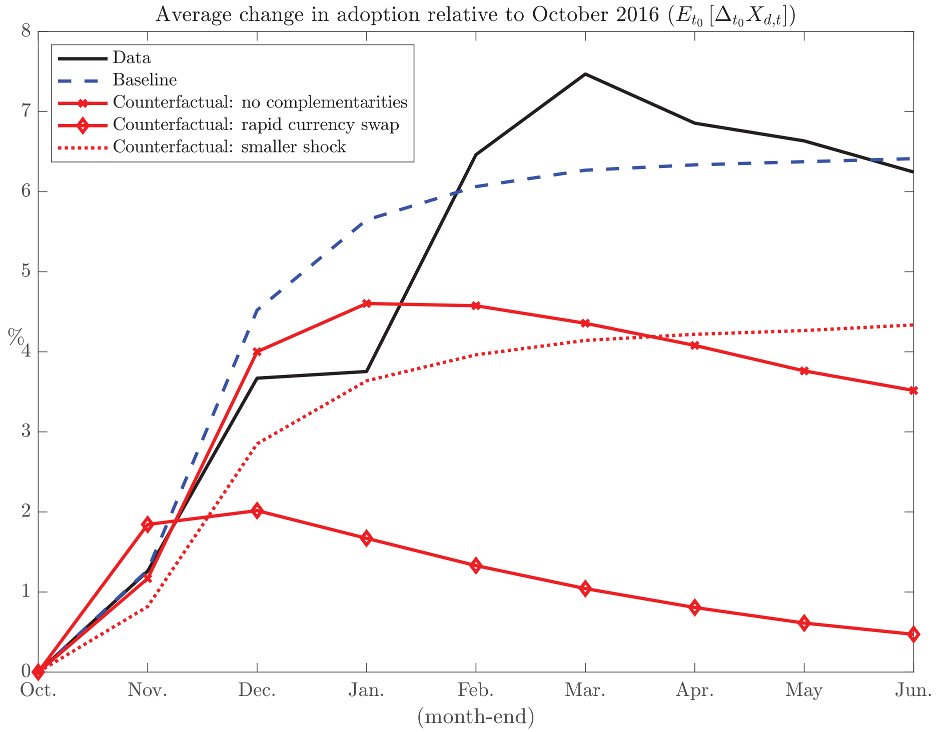
Notes: The figure plots the dynamic treatment effects of the demonetization shock on technology adoption of electronic payment systems. The graphs report the coefficients δ_t from specification 2.1; the top panel reports the effects for the total amount of transactions (in logs), the middle panel reports the effects for the total number of active firms on the platform (in logs), and the bottom panel reports the effect for the total number of new firms on the platform (in logs). The x-axis represents the month, where October 2016 is normalized to be zero. 95% confidence intervals are represented with the vertical lines; standard errors are clustered at the district level.

Figure 2.5. District adoption dynamics in electronic payments data based on distance to electronic hub



Notes: The figure plots the dynamic effects of adoption across districts based on a district's initial adoption rates as proxied by the distance of that district to the closest district with more than 500 active firms before demonetization. The top panel reports the effects for the total amount of transactions (in logs), the middle panel looks at the total number of firms, while the bottom panel reports the effects for the total number of new firms transacting on the platform (in logs). The x-axis represents month, where October 2016 is normalized to be zero. 95% confidence intervals are represented with the vertical lines; standard errors are clustered at the district level.

Figure 2.6. Counterfactual paths of average adoption rates across districts



Notes: The black solid line reports the empirical change in average adoption rates across districts. The other lines report average changes in adoption rates constructed using $S = 100$ simulations from the model, each of a dataset of the same size as the actual data. The dashed blue line is the change in adoption rate obtained from the model evaluated at the point estimates reported in table 2.4. The solid crossed red line is the average change in adoption rate in the absence of complementarities, assuming that the switching frontier (which is flat without externalities) has the same level as the switching frontier with externalities when adoption is 0. The solid diamond red line is the change in adoption rate when $\theta = 4.6$, corresponding to a 90% decay time of 15 days. The dotted red line is the change in adoption rate when the shock has half the initial size as estimated in table 2.4.

Table 2.2. Exposure$_d$ and adoption of digital wallet

	log(amount)	log(# users)	log(# switchers)
	(1)	(2)	(3)
(Exposure)$_d$ $\times$ $\mathbf{1}\,(t \geq t_0)$	3.134***	1.078**	0.720**
	[0.884]	[0.424]	[0.315]
Observations	6,846	6,846	6,846
R-squared	0.849	0.868	0.818
District f.e.	✓	✓	✓
Month f.e.	✓	✓	✓
District Controls $\times$ Month f.e.	✓	✓	✓

Notes: Difference-in-differences estimates of the effect of the shock on the adoption of digital wallet. District controls include (log) pre-shock banking deposits, share of villages with ATM facilities, share of villages with banking facility, share of rural population and level of population in the district. Standard errors clustered at the district level are reported in parentheses. $*** : p < 0.01$, $** : p < 0.05$, $* : p < 0.1$.

Table 2.3. District adoption rate of digital wallet based on distance to the hubs

	log(amount)		log(# users)		log(# switchers)	
	(1)	(2)	(3)	(4)	(5)	(6)
(Distance to hub)$_d$ $\times$ $\mathbf{1}\,(t \geq t_0)$	-5.098***	-3.958***	-2.233***	-1.724***	-1.613***	-1.100***
	[0.936]	[1.190]	[0.468]	[0.497]	[0.361]	[0.387]
Observations	6,846	6,846	6,846	6,846	6,846	6,846
R-squared	0.852	0.886	0.871	0.912	0.821	0.871
District Controls $\times$ Month f.e.	✓	✓	✓	✓	✓	✓
State $\times$ Month f.e.		✓		✓		✓

Notes: Difference-in-differences estimate of the effect of initial conditions, using the distance to the nearest hub (defined as districts with more than 500 retailers in September 2016) as a proxy for the initial share of adopters. District-level controls include (log) pre-shock banking deposits, share of villages with ATM facilities, share of villages with banking facility, share of rural population, level of population and distance to state capital. Standard errors clustered at district level are reported in parentheses. $*** : p < 0.01$, $** : p < 0.05$, $* : p < 0.1$.

Table 2.4. Point estimates and standard deviations for parameters

Parameter		**Estimate**	**Standard error**
S	Size of aggregate shock	0.215	(0.042)
C	Adoption complementarities	0.062	(0.009)
k	Speed of technology adjustment	0.166	(0.035)
σ	Volatility of idiosyncratic innovations	0.042	(0.018)
M^e	Returns to electronic payments when $X_{d,t} = 0$	0.974	(0.001)

Notes: The parameters are estimated on a balanced panel with 512 districts and 8 months. The estimation procedure uses the simulated method of moments and is described in section 2.4. Standard errors computed using bootstrap are reported in parenthesis.

Table 2.5. Alternative Interventions

	Baseline	Alternative interventions			
		$g = 0$	$g = 0.2$	$g = 0.4$	$g = 0.6$
Shock size (p.p.)	21.5	14.2	13.1	11.5	10.7
Shock half-life (months)	0.9	1.4	1.5	1.8	1.9
$_{t_0}\left[\Delta_{t_0} X_{d,t_0+T}\right]$ (p.p.)	6.4	8.8	8.8	8.2	7.8
$sd_{t_0}\left[\Delta_{t_0} X_{d,t_0+T}\right]$ (p.p.)	28.6	23.6	21.9	20.5	19.9

Notes: The column marked "Baseline" report the estimated shock size, the shock half-life, and the mean and standard deviation of long-run changes in average adoption rates; $T = 3$ years and $s = 100$ simulations is used to compute these moments. The other columns report these moments under alternative scenarios.

3. ACCESS TO INFORMATION, TECHNOLOGY ADOPTION AND PRODUCTIVITY: LARGE-SCALE EVIDENCE FROM AGRICULTURE IN INDIA

Nearly 80 percent of the world's extreme poor live in rural areas, with most relying on agriculture for their livelihoods (World Bank 2019). These individuals are often trapped in a vicious circle of low yields, due to limited adoption of modern agricultural technologies capable of improving their productivity and raising their incomes. One of the main barriers to adoption is farmers' imperfect knowledge of these technologies and of the best practices associated with their use.[1]

Over the past two decades, the rapid diffusion of mobile phones and telecommunication services in rural areas of developing countries has raised expectations about their ability to reduce informational frictions, promote technology adoption and increase farmers' productivity. A number of recent randomized controlled trials have shown that access to mobile-based agricultural advice services may indeed affect agricultural practices (Cole and Fernando 2016; Casaburi et al. 2014; Fabregas et al. 2019). Yet, several important questions remain unanswered. First, there is little empirical evidence on the distributional consequences of greater access to information. Does its expansion amplify or reduce productivity differences across farmers? Second, there is limited empirical evidence on its long-run consequences. Is the effect of access to information on agricultural practices and productivity temporary or long-lived? A key challenge to tackle these questions is that one needs to observe, for a large sample of farmers and over a long period of time, data on access to information about modern agricultural technologies, the actual

[1] On the constraints to the adoption of new technologies in agriculture in developing countries, see reviews in Jack (2013), Foster and Rosenzweig (2010) and Feder et al. (1985). On the role of information frictions see, among others, Foster and Rosenzweig (1995) and Conley and Udry (2010).

adoption of these technologies, and the evolution of agricultural productivity. Furthermore, one needs to be able to separate the role of information from additional dimensions through which mobile phones can influence farmers' decisions to modernize their technologies.

The study addresses these challenges using large-scale data from India. First, variation in the rollout of mobile phone coverage generated by the Shared Mobile Infrastructure Scheme (SMIS), a large government program launched in 2007 that financed the construction of about 7,000 mobile phone towers in previously unconnected areas of India is exploited. Second, the geographical coverage brought by new SMIS towers is matched with data on the location and content of 2.5 million toll-free phone calls made by farmers to one of India's leading agricultural advice services, the Kisan Call Centers (KCC). This data allows to observe farmers' questions about specific agricultural technologies and the answers they receive from agronomists. A unique feature of the KCC service is exploited – that agricultural advice is offered in a limited number of languages, effectively excluding farmers who do not speak any of these – to isolate the effect of access to information. Finally, data on mobile phone coverage and phone calls is matched with detailed district-level survey data on crop yields and adoption of agricultural inputs – including seed varieties, pesticides and herbicides – in an area covering around 19 million farmers.

The combination of these datasets allows to map farmers' calls about specific agricultural technologies with their actual adoption. The data on agricultural inputs used by farmers is observed 5 years after the introduction of the SMIS and annual agricultural yields for 10 years after the introduction of the SMIS, which allows to study the long-run effects of access to information.

The empirical analysis proceeds in two steps. The first step uses an event-study design to document the evolution of farmers' calls to seek agricultural advice when new mobile phone towers are constructed in previously uncovered areas. Using high-frequency (monthly) variation, is it documented that the construction of the first mobile phone tower in a given area is followed by a significant increase in the number of farmers' calls. This is consistent with a large and underserved demand for agricultural advice in rural India.[2]

The event study also documents that linguistic differences can generate unequal gains in access to information. Although the government-sponsored agricultural advice is in principle available

[2]As of 2003, 60 percent of Indian farmers in a nationally representative survey reported not having access to any source of information on modern technology to assist them in their farming practices (National Sample Survey 2005).

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.[3] This effectively creates a language barrier for the over 40 million individuals whose main language belongs to the 100 non-official ones recorded in the Indian Census. The calls for agricultural advice from areas where the majority of the local population speaks a non-official language only increase by 20 to 30 percent of the increase observed in areas where the majority speaks an official language. This is despite the fact that, within the sample, these areas are comparable in terms of initial socio-economic characteristics and pre-existing trends in agricultural performance.

The second step of the analysis studies the real effects of access to information on technology adoption and productivity. To account for the potentially endogenous location of SMIS mobile phone towers, an identification strategy is proposed that compares – within each administrative district – locations where new SMIS towers were proposed and eventually constructed, with locations where they were also proposed but eventually not constructed. It is shown that these two types of location are balanced on initial observable characteristics once determinants of tower relocation such as terrain ruggedness and population covered are controlled for, and that they experienced similar pre-existing trends in both technology adoption and agricultural yields in the 5 years preceding the introduction of new towers. In addition, variation in the spatial diffusion of non-official languages is exploited to capture the heterogeneous ability of farmers to access phone-based services for agricultural advice. The combination of mobile phone coverage and absence of language barriers with agricultural advisors is considered as a positive shock to information about agricultural practices for farmers.

The measures of technology adoption include farmers' adoption of high-yielding variety (HYV) seeds, chemical fertilizers and pesticides, as well as artificial irrigation systems. HYV seeds are commercially developed to increase crop yields and are one of the most prominent innovations in modern agriculture.[4] Chemical fertilizers and reliable irrigation systems are key complementary inputs to maximize HYV potential. Data on the adoption of these technologies is sourced from the Agricultural Input Survey of India, which is carried out at 5-year intervals and whose last two waves were in 2007 and 2012.

[3]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. The 22 officially-recognized languages are: Hindi, Bengali, Marathi, Telugu, Tamil, Gujarati, Urdu, Kannada, Odia, Malayalam, Punjabi, Assamese, Maithili, Santali, Kashmiri, Nepali, Sindhi, Dogri, Konkani, Manipuri, Bodo, and Sanskrit.
[4] On the impact of high-yielding varieties on agricultural productivity and economic development see, among others, Evenson and Gollin (2002, 2003).

The estimates indicate that in areas where the entire population speaks an official language and can therefore access agricultural advice, a 1 s.d. larger increase in mobile phone coverage is associated to a 1.4 percentage points larger increase in area farmed with HYV seeds between 2007 and 2012. This effect corresponds to a 5.3 percent increase in land cultivated with HYV seeds for the average cell in the sample.[5] Positive and significant effects are also found on the adoption of chemical fertilizers, pesticides and irrigation. Consistent with an information mechanism, it is shown that these areas also experienced a larger increase in farmers' calls seeking information on the adopted technologies. On the other hand, in areas where the population cannot access agricultural advice due to language barriers with KCC advisors, the impact of mobile phones on the modernization of agricultural technologies is significantly more limited. The estimates indicate that, for any level of mobile phone coverage increase, a 1 s.d. increase in the share of non-official language speakers reduces the adoption of new agricultural technologies by 18 percent. This effectively represents the share of technology adoption attributable to access to information about agricultural practices.

Next, the effect of farmers' access to information on agricultural productivity, measured by average crop yields, is studied. The estimates indicate a significant relative increase in yields in the years following the construction of a new SMIS tower in previously uncovered areas. The yearly frequency of the data to document the timing of this effect is exploited. While there are no pre-existing trends in agricultural yields in the 5 years prior to the launch of the program, the effect of new mobile phone towers materializes about one year after their construction and increases in magnitude for the first three years. It is also shown that this effect is persistent in the long run: areas that received a new SMIS tower and did not face language barriers to access agricultural advice still displayed higher agricultural yields in 2017, about a decade after the start of the program. As with technology adoption, however, the positive effects of mobile phones on agricultural productivity are strongly mitigated by the inability of farmers to access agricultural advice. The estimates indicate that in areas where more than 50 percent of the population speaks a non-official language, the effect of mobile phone coverage on productivity is completely muted. These results, like the previous on technology adoption, are robust to controlling for the interaction between mobile coverage expansion and 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.

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

Finally, the returns to mobile phone coverage and access to information are shown to be highly heterogeneous, depending on farmers' initial productivity. Within the 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 the productivity distribution was around twice as large as the one observed at the 25th percentile. This is a yield gap similar to that observed in rice and wheat production between the richest 10 percent and the poorest 10 percent of countries (Gollin, Lagakos, and Waugh 2014). The results show that the effect of access to information is the largest for areas in the lowest productivity quartile. The estimates suggest that providing agricultural advice on mobile phones can close about 36 percent of the productivity gap between farmers in the 25th percentile of the productivity distribution and those in the 75th percentile.

The lower returns to mobile phone coverage in areas where farmers face language barriers with KCC advisors strongly suggest that access to agricultural advice plays a key role in the modernization of agriculture. However, alternative mechanisms potentially linking mobile phone coverage with technology adoption and productivity are also discussed and tested. For one, previous evidence suggests 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, could help farmers pay the fixed cost of adopting new technologies. To account for this possibility, estimates of the model where a full set of fixed effects for the closest agricultural market to each cell in the sample were included are presented. This allows to compare outcomes across farmers who plausibly face the same prices for their products and experience the same changes in local demand. All main results are robust to this augmented specification.

A second alternative mechanism through which mobile phones could also affect technology adoption and productivity is social learning. Models of social learning suggest that individuals adopt new technologies once they have gathered enough evidence from previous adopters that the new technology is actually worthy of uptake. In the context of the study, the expansion of the mobile phone network could facilitate the diffusion of such information across farmers and encourage the modernization of agriculture, regardless of the availability of call centers for agricultural advice. A prediction of this interpretation is that social learning is more likely to happen in areas where individuals tend to speak the same language. It is attempted to capture this potential channel in the main specification by controlling for the interaction between mobile phone diffusion and local linguistic fragmentation. The main estimates on access to information are not significantly affected by the introduction of this additional interaction term.

The rest of the study is organized as follows. Section 3.1 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 Scheme and the Kisan Call Centers for agricultural advice – that are central to the empirical analysis. Section 3.2 presents the identification strategy and all the empirical results. Section 3.3 provides concluding remarks.

3.1 Data, Institutional Background, and Stylized Facts

This section describes 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 the empirical analysis are areas of 10×10 km, which is referred to as cells. A grid of 10×10 km cells is used 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.[6]

3.1.1 Data on Mobile Phone Coverage and its Diffusion in India

Data on the diffusion of mobile phone coverage in India comes from the Global System for Mobile Communication Association (GSMA), the association representing the interests of the mobile phone industry worldwide. 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 provide, for all years between 1998 and 2012, geo-located information on mobile phone coverage aggregated across all operators. The 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.[7]

[6]Overall, India can be split into 41,495 cells distributed over 524 districts. Since cell borders do not typically correspond to district administrative borders, cells spanning over more than one district as assigned to the district which occupies the largest area. One challenge faced 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, minimum comparable areas (MCAs) encompassing one or more districts that cover the same geographical space between 1997 and 2012 were created. 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.

[7]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).

While the country had virtually no mobile phone coverage until 1997, the mobile phone network began to expand rapidly shortly afterwards, covering 22 percent of the population in 2002, 61 percent in 2007 and 89 percent in 2012. 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 started in urban areas and only later reached rural ones. We document this pattern in Figure 3.1, 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.

3.1.2 Construction of mobile phone towers under the SMIS 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 Infrastructure Scheme (SMIS), 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.[8] Of the 7,871 proposed towers under Phase-I, 7,353 were eventually constructed.

Data on the towers constructed under SMIS was obtained from the Center for Development of Telematics (C-DoT) - the consulting arm of the Department of Telecommunications of India. The C-DoT provided 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, this date is referred to as the date of construction. From the 7,353 towers constructed under Phase I of the SMIS program

[8]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.

350 towers for which the construction date is missing were removed. This leaves with 7,003 mobile towers used in the empirical analysis. Figure 3.2 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, a 5-km radius of coverage around the towers' location is assumed 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).

3.1.3 Data on farmers' calls to Kisan Call Centers

To investigate the role of information on agricultural practices data on farmers' calls to Kisan Call Centers (KCC) obtained from the Department of Agriculture, Cooperation and Farmers Welfare is used. Calls are geo-located at the subdistrict (or block) level and they are proportionally assigned to all cells whose centroid is contained in the subdistrict.[9]

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. KCC are 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.[10]

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.[11] 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.[12]

[9]On average, there are 27 cells per subdistrict. Whenever information on the subdistrict from which the call is originated is missing, information on the district of the call and the crop for which the caller is seeking information to assign calls to a given cell is used.

[10]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.

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

[12]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.

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 on how to improve yields or – more specifically – on which seed varieties to use 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).

Figure 3.3 reports 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.

3.1.4 Data on Technology Adoption and Agricultural Productivity

The 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. The main empirical analysis focuses on the last two waves of the AIS, 2007 and 2012, earlier survey waves are used to document pre-existing trends.[13] 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.[14] The AIS reports information on land farmed with these input technologies at the district-crop level. The share of land farmed with a given agricultural technology k in a given cell i using the following neutral assignment rule is computed as:

[13] The Agricultural Input Survey runs from 1^{st} July to June 30^{th} of the following year. The study used the terminology 2007 when referring to the survey carried out between July of 2006 and June of 2007.
[14] The AIS was not conducted in the states of Bihar and Maharastra before 2012. Thus, these states were excluded from the analysis.

$$\left(\frac{Area^k}{Area}\right)_{idt} = \sum_{c \in O_i} \left[\left(\frac{Area^k}{Area}\right)_{dct} \times \left(\frac{Area_{idc,t=2000}}{Area_{id,t=2000}}\right) \right] \tag{3.1}$$

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.[15] Thus, the product of first and second element gives 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), an estimate of the share of land farmed with a given technology in a given cell is obtained.[16]

Effectively, the within-district variation generated by the assignment rule is driven by the baseline crop composition of each cell coupled with district-crop level variation in technology adoption. One potential concern with this assignment rule is that it may generate non-classical measurement error. This would happen if, for example, new SMIS towers are systematically constructed in cells (within a district) where farmers grow crops characterized by fast technology adoption. This concern is addressed by showing that the treatment and control cells are balanced in terms of initial shares of area farmed with crops that experienced faster increase in HYV adoption at district level. Section 3.2.2 shows that treatment and control cells have similar trends in technology adoption in the five years before the introduction of the SMIS program, which rules out the concern that baseline crop composition captures long-term trends in adoption.

The AIS covers the following agricultural input technologies: seeds – distinguished between traditional and high-yielding varieties – chemical fertilizers, organic manures and pesticides, agricultural machinery and agricultural credit. The preferred 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

[15]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. The focus is 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.

[16]As an example, suppose that in district d, 20 percent of land farmed with rice and 50 percent of land farmed 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 the neutral assignment rule, 38 percent of land in cell i is assigned to high-yielding varieties: $(0.2 \times 0.4) + (0.5 \times 0.6) = 0.38$.

desirable characteristics of different breeds, including improved responsiveness to fertilizers, dwarfness, and early maturation in the growing season. 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 the 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 each crop at the district level. The 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. The measure of crop yield is constructed similarly to Jayachandran (2006), who use a weighted average of normalized yields of the major crops farmed in India to generate a district-level measure of agricultural productivity. Agricultural productivity at the cell level is then computed with a neutral assignment rule similar to the one reported in equation (3.1).

3.2 Empirics

The empirical analysis proceeds in two steps. First, an event-study design is used to document the evolution of farmers' calls to KCC when new SMIS 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 to document the role of language barriers in the diffusion of information. In particular, it shows that that geographical differences in the diffusion of non-official languages among the rural population affect the spatial availability of agricultural advice provided by the KCC. These results are presented in section 3.2.1.

Next, the real effects of access to information on technology adoption and agricultural productivity are studies. Since technology adoption and productivity are not observed at the same high frequency as farmers' calls, the event-study design just described for these outcomes cannot be used.[17] Instead, an identification strategy that compares locations where new mobile phone towers were proposed and constructed under the SMIS program with similar locations where

[17]Data on adoption of agricultural technologies is observed at 5-year intervals in the Agricultural Input Survey. Agricultural yields are instead observed at yearly level, which allows to document the timing of the effect around the construction of new towers.

new towers were proposed but eventually not constructed is proposed. The 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 is exploited. The focus is on the change in technology adoption and productivity between 2007 and 2012, with 2007 being the last wave in the AIS *before* the SMIS program, and 2012 the first wave *after* the SMIS program. The identification strategy is discussed in section 3.2.2 and the results presented in sections 3.2.3 to 3.2.5.

3.2.1 Event-Study Evidence on Farmers' Access to Information

The evolution of farmers' calls to KCC around the introduction of new mobile phone towers is estimated using the following specification:

$$Ln\,(1 + Calls)_{it} = \alpha_i + \alpha_t + \sum_{k=-12}^{+36} \beta_k D_{it}^k + \varepsilon_{it} \tag{3.2}$$

The outcome variable in equation (3.2) 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 is set at $k = 0$. The 12 months prior to the introduction of the first tower and the 36 months after are included. 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 the focus is on cells that will eventually receive a mobile phone tower under the SMIS program described in section 3.1. Notice also that in this first analysis the focus is on the number of calls, while the analysis of their content is discussed in detail in section 3.2.3.

Panel (a) of Figure 3.4 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.[18] Second, within 4 months of the construction of the first tower there is 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, this differential

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

continues to grow over the next 18 months, reaching a 40 to 50 percent increase in calls three years after the construction of the first tower in a cell.

As discussed in section 3.1, 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.[19] 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 similar mobile phone coverage via new SMIS towers, the ability of farmers to access dedicated information on agricultural practices might vary by local language. Panel (b) of Figure 3.4 estimates equation (3.2) 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.[20] 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 3.4 suggests that the expansion of mobile phone coverage represents a large information shock to farmers, and that this shock has been largely heterogeneous depending on linguistic differences between farmers and government advisors working at KCC. The next section studies how differences in this shock to access to information map into technology adoption and agricultural yields among farmers.

3.2.2 The Real Effects of Access to Information - Identification Strategy

This section presents the identification strategy to study the effect of farmers' access to information on real outcomes, namely agricultural technology adoption and productivity. The 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. The combination of mobile phone coverage and absence of language barriers with agricultural advisors is considered as a positive shock to information about agricultural practices for farmers.

[19] See `https://mkisan.gov.in/aboutkcc.aspx`. Agronomists answering in each KCC location answer calls in one (or more) of the official languages.
[20] Data on the share of local population speaking non-official languages is sourced from the 2011 Indian Census and available at the subdistrict level. To each cell whose centroid falls within a given subdistrict the share of local population speaking non-official languages is assigned in that subdistrict.

The identification strategy exploits variation in the construction of mobile phone towers under the Shared Mobile Infrastructure Scheme, or SMIS, described in section 3.1. In the initial phase of this program, the Department of Telecommunications identified 7,871 potential locations for the 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, the fact that not all the locations in the initial list eventually received a tower is exploited. In some cases, towers were either relocated or not constructed. Thus, cells where towers were initially proposed and eventually constructed are compared with cells in the same administrative district where towers were initially proposed but eventually not constructed.[21] The final sample consists of 6,320 cells, of which 4,569 in the treatment group and 1,751 in the control group. Figure 3.5 presents the geographical distribution of treatment (in red) and control (in blue) cells across Rajasthan – the largest Indian state by area – superimposing the lattice of 10×10 km cells to show the level of geographical detail allowed by the data. On average, the sample includes 27 cells per district – 20 treated and 7 control. This variation is further combined with data on the share of local population speaking non-official languages.

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. The main challenge to the 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. The test for differences in the probability of receiving coverage from new SMIS towers based on cell observable characteristics and on pre-existing trends in technology adoption and productivity as performed. This balance test is also performed across cells with different shares of the local population speaking a non-official language, conditional on receiving coverage from new SMIS towers.

The results of the balance tests are reported in Table 3.1. The outcome in columns (1) to (4) is an indicator variable – $\mathbb{1}$ (Tower) – which is equal to 1 for cells where a new SMIS tower was proposed and eventually constructed, and 0 for cells where a new SMIS tower was proposed but

[21] The coverage for each new tower is based on its technical specifications, which corresponds to a 5 km coverage radius around its centroid. The analysis is robust to using the share of land covered by SMIS towers instead of an indicator variable.

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 appear to depend on the availability of a connection to the power grid. Next, Column (2) studies whether pre-trends in agricultural technology adoption or productivity affect the probability of eventually receiving a SMIS tower. As shown, there are no significant differences in technology or productivity growth across treated and control cells in the 5 years preceding the tower construction program. Column (3) then explores the correlation with a number of cell characteristics sourced from the Village Survey of the Population Census of India. 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. Column (4) considers 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. The empirical analysis adds these controls to the specification and show that all the estimates are stable when including the observable cell characteristics reported in Table 3.1.

Finally, in Column (5) conditions on cells eventually receiving coverage from new SMIS 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 the sample, the distribution of non-official language speakers is uncorrelated with observable characteristics and pre-trends in technology adoption and productivity.

First Stage. The first-stage regression is as follows:

$$\Delta Cover_{id} = \alpha_d + \gamma \mathbb{1}\left(\text{Tower}\right)_{id} + \delta X_{id} + u_{id} \tag{3.3}$$

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 SMIS tower location.[22] The coefficient of interest is γ, which captures the effect of tower construction

[22]The tower construction program used 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

under the SMIS 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 3.1. The specification includes district fixed effects (α_d) and standard errors are clustered at the district level. Finally, in all specifications each cell is weighted by its population at baseline (2001).

Table 3.2 reports the first-stage results. The estimated coefficient in column (1) indicates that cells covered by new SMIS 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. Column (2) includes the three main determinants of tower relocation according to C-DoT officials: population, availability of power supply and terrain ruggedness. The magnitude of the estimated coefficient decreases from 0.11 to 0.073, and remains highly statistically significant. Finally, Column (3) adds all the observable socio-economic cell characteristics. Consistent with the results presented in Table 3.1, the size of the point estimate is unaffected by including these additional controls. According to the specification in column (3), cells covered by new SMIS 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 the Kleibergen and Paap (2006) first stage F-statistics for the validity of the instrument are reported. The hypothesis that the first stage is weak can be safely rejected.

Second Stage: Empirical Specifications. The section starts 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 one denotes a generic cell by i, with $i \in d$, where d denotes a district, the regression model is:

$$\Delta y_{id} = \alpha_d + \beta \Delta \widehat{Cover}_{id} + \delta X_{id} + u_{id} \tag{3.4}$$

where Δy_{id} denotes the change in a given outcome between 2007 and 2012 and $\Delta \widehat{Cover}_{id}$ represents the change in the share of land covered by the mobile phone network over the same period, instrumented with the variable $\mathbb{1}$ (Tower) from equation (3.3). X_{id} is the vector of cell characteristics discussed in Table 3.1 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. 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

their services and expand their market shares. Thus, tower construction under SMIS is not expected to be the sole source of variation in change in GSMA coverage, even in rural regions.

more generally, increasing local income and thus demand for agricultural products. Farmers might adopt new technologies to serve this increased demand.

To make progress in the direction of isolating the role of information, equation (3.4) is expanded 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. The following augmented specification is estimated:

$$(3.5)\quad \Delta y_{id} = \alpha_d + \beta_1 \widehat{\Delta Cover}_{id} + \beta_2 \widehat{\Delta Cover_{id} \times NOLang_{id}} + \beta_3 NOLang_{id} + \delta X_{id} + u_{id}$$

where, compared to equation (3.4), the share of population speaking a non-official language ($NOLang_{id}$) and its interaction with the change in mobile phone coverage is also included.[23] The 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. 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 ($NOLang_{id} = 1$).

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

The section starts by documenting the effect of mobile phone coverage on farmers' calls for agricultural advice. In particular, the identification strategy described in section 3.2.2 is used to study farmers' access to information about specific technologies. Crucially for the 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 to distinguish between calls in which farmers seek advice regarding specific agricultural technologies such as new varieties of seeds, fertilizers, irrigation, or pesticides. 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 is studied in the next section.

Column (1) of Table 3.3 estimates the effect of mobile phone coverage on the change in total number of calls to KCC between 2007 and 2012, as described by equation (3.4). The estimated coefficient suggests that cells with 1 s.d. larger increase in mobile phone coverage experienced a 23 percent larger increase in total calls by farmers. Next, Column (2) reports the results from the unrestricted model in equation (3.5), where the effect of coverage is allowed to vary across areas

[23]The latter is instrumented by the interaction of the share of population speaking a non-official language with the indicator variable for tower construction from equation (3.3).

facing different language barriers with KCC advisors. The estimated coefficient β_1 is interpreted as the combined effect of coverage and access to a phone-based service for agricultural advice on farmers' calls. Its magnitude suggests that a 1 s.d. increase in coverage in cells where all farmers speak an official language increases the number of calls by 26.6 percent. The coefficient β_2, on the other hand, indicates that calls for agricultural advice are less responsive to changes in coverage when the local population does not speak an official language. The sum of the estimated coefficients β_1 and β_2 (0.828 - 0.716 = 0.112) implies that, when the entire population does not speak an official language, the effect of a 1 s.d. increase in mobile phone coverage on calls is only 3.6 percent and not statistically different from zero.

Next, the analysis focuses on farmers' calls about specific agricultural technologies: seed varieties, fertilizers, irrigation, and pesticides. The results are shown in column (3) to (10). Odd columns refer to the average effect of mobile phone coverage, while even columns allow for the heterogeneous response to coverage depending on the share of non-official language speakers. The results are in line with those on the total number of calls and very similar for all agricultural technologies. An increase in mobile phone coverage is associated with more calls for agricultural advice on specific technologies, but the effect is limited by the existence of language barriers between the local population and the KCC advisors, as shown by the negative and statistically significant coefficients on the interaction terms in all specifications.

Overall, the results reported in Table 3.3 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, one 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. The following sections focus on these two outcomes.

3.2.4 The Effect of Access to Information on Technology Adoption

This section studies the effect of farmers' access to information via call centers for agricultural advice on technology adoption. The focus in particular is 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 equations (3.4) and (3.5) are estimated 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.

Column (1) of Table 3.4 reports the results of estimating equation (3.4) 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 1 s.d. 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 the 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 the sample.

Column (2) reports the results of estimating equation (3.5), which allows for the heterogeneous response to mobile phone coverage depending on the share of local population speaking non-official languages. 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. For any level of mobile phone coverage increase, a 1 s.d. increase in the share of non-official language speakers reduces the adoption of new agricultural technologies by 18 percent. This differential captures the portion of mobile phone impact that is attributed to access to information on agricultural practices.

Columns (3) and (4) focuses 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 et al. 1974). Thus, adoption of HYV seeds by farmers is expected 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 advisors limit the impact of mobile phone coverage on adoption of fertilizers.

Next, the study 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 et al. 1974). Thus, adoption of HYV seeds by farmers is expected to also increase their demand for irrigation. The effect on irrigated area is studied in columns (5) and (6). Similar results, although smaller in magnitude, to the ones documented for chemical fertilizers are found. 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. As in the previous columns, the point estimate on the interaction term suggests that the impact of mobile phones on technology adoption is limited by the presence of language barriers to obtain information about agricultural practices.

Overall, the results presented in Tables 3.3 and 3.4 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. One can use the estimates to calculate the implied elasticity of technology adoption to access to information about a given technology. To compute this elasticity the estimated percentage increase in area farmed is divided 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, elasticities of 0.64 for chemical fertilizers, 1.1 for chemical pesticides and 3 for irrigation are obtained.

3.2.5 The Effect of Access to Information on Productivity

This section studies the effect of farmers' access to information via new mobile phone towers on agricultural productivity.

The section starts by studying the effects of access to information on productivity using the same specification used to study its effects on technology adoption. The results are reported in Table 3.5. Column (1) shows a positive but insignificant effect of mobile phone coverage on the change in agricultural productivity between 2007 and 2012. Column (2) shows that the impact of mobile phones varies significantly across areas, depending on farmers' access to agricultural advice. In areas where the entire population speaks an official language, a 1 s.d. increase in mobile phone coverage leads to a 1.3 percent larger increase in productivity, an effect 40 percent larger than the average. On the contrary, in areas where 50 percent or more of the population speaks a non-official language the positive effect of mobile phone coverage on productivity is completely offset, as implied by the magnitude of the negative and significant coefficient on the interaction term β_2.

The remaining columns of Table 3.5 investigate the differential returns to mobile phone coverage and access to information across areas with different initial levels of agricultural productivity. In the 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 a cell at the 75th percentile 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 test for heterogeneous effects across farmers with different initial productivity are provided in columns (3) to (6), which estimate equation (3.5) separately for each quartile of initial productivity. The results indicate that the effect of access to information is largest – and most precisely estimated – for farmers with the lowest initial level of productivity. The point estimate on β_1 for this group is 0.052, around 30 percent larger than the average effect reported in column (2). 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.

The nature of the agricultural yield data, which are available at the yearly level for the period until 2017, allows to further characterize the relationship between mobile phone coverage, access to information and productivity. First, the timing of the impact is explored. To this end, the staggered introduction of SMIS towers among the treatment cells is exploited, estimating an event-study equation similar to (3.2). Notice that in this specification the focus is exclusively on cells initially selected for the SMIS program and that eventually received a tower at some point between 2007 and 2010. The results are reported in Figure 3.6, which plots the estimated coefficients β_k on years relative to tower construction along with their 95 percent confidence intervals. The study finds no pre-existing trends in agricultural yields in the 4 years before the construction of the first tower in a cell. The effect of new mobile phone towers on productivity materializes about one year after their construction and increases in magnitude for the first three years. The magnitude of the estimated coefficients indicates, on average, a 2 to 4 percent increase in agricultural yields in the first five year post tower construction. The timing of the effect is consistent with a permanent and long-lived shift in productivity.

Next, the analysis of long-term effects is expanded by investigating the impact of access to information 10 years after the introduction of the SMIS program. To this end, the results presented in Table 3.5 are replicated using as outcome variable the decadal change in agricultural yields between 2007 and 2017 (the latest year in the productivity data). The results are reported in Table 3.6. Comparing the magnitudes of the estimated coefficients with those in Table 3.5 suggests that the differential in productivity between areas covered or not by mobile phones, and between areas with and without access to agricultural advice, increases over time. Column 1 shows that the average effect of mobile phones on productivity over the ten-year horizon is positive and precisely estimated: a 1. s.d. larger increase in mobile phone coverage is associated with a 1.7 percent larger increase in productivity. Column (2) shows that, over ten years, a 1 s.d. increase in mobile phone coverage in areas with full access to agricultural advice leads to a 2.2 percent larger increase in productivity. This represents an additional 65 percent increase respect to the differential observed over the period 2007-2012. The negative and significant coefficient on the interaction term in column (2) also confirms the absence of effect of mobile phones on productivity in areas where the share of population speaking a non-official language exceeds 60 percent. Finally, the results in columns (3) to (6) display a pattern similar to what observed in Table 3.5 with regard to the effect at different levels of initial productivity. The estimates suggest that the returns to access to mobile phones and agricultural advice are the largest for farmers in the initially least productive areas.

Overall, the results in this subsection show that increased access to information can set rural areas on a different path of agricultural development, encouraging the adoption of modern technologies that generate higher yields. This effect manifests rapidly and persists over time. At the same time, the results indicate that language barriers between agricultural advisors and local communities represent an obstacle to widespread access to information, potentially increasing disparities between areas and significantly hampering the returns of telecommunications infrastructure programmes designed to include rural areas in the mobile phone network.

3.3 Concluding Remarks

This study provides large-scale evidence of the effects of accessing information via mobile phones on the adoption of modern agricultural technologies and crop yields in rural India, using detailed geo-referenced data on the construction of new mobile phone towers, farmers' calls for agricultural advice and the prevalence of local languages across fine geographical areas. The results indicate that mobile phones can have long-lasting effects on farmers' productivity by facilitating the adoption of modern technologies. The findings also suggest that the effects are larger for farmers with the lowest initial level of productivity, highlighting the potential of mobile phones to reduce the large productivity gap between farmers in India.

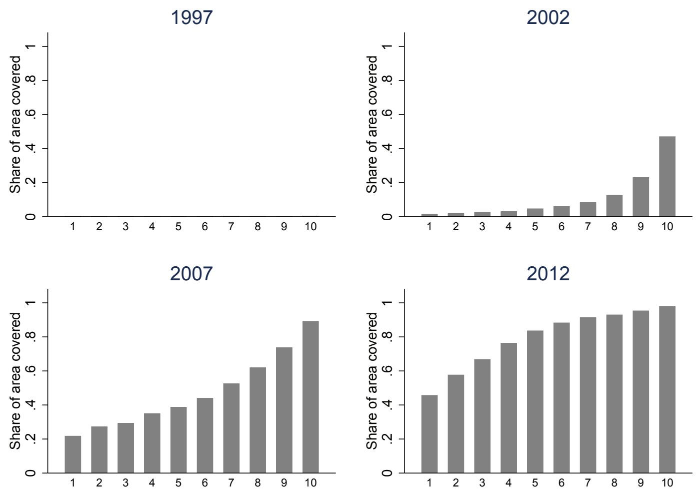
Access to mobile phones, however, is not in itself sufficient to foster this transition. A key element is the ability of farmers to access high-quality agricultural advice on their phones. The results show that the benefits of mobile phone coverage are much more limited, when not entirely absent, for farmers facing language barriers with agricultural advisors. The results imply that, in areas where 50 percent of more of the population cannot access agricultural advice, an increase in mobile phone coverage does not lead to a modernization of agricultural practices and technologies, nor to increased productivity for farmers.

As the number of mobile-based agricultural advisory services worldwide increases steadily (GSMA 2020), the results therefore provide a tale of cautious optimism about their effectiveness. On the one hand, the ability to connect with farmers in hard-to-reach rural areas and to provide continuous and personalized advice during the agricultural cycle can make mobile-based agricultural extensions such as the KCC an effective tool to lift farmers out of poverty. On the other hand, however, the design and specific features of these programs – such as the language in which agricultural advice is provided – may generate an uneven access to information and exacerbate disparities among farmers, thus creating winners and losers from their introduction. In the context of the study, this highlights the importance of expanding the KCC service to the 40 million Indian farmers who do not speak any officially-recognized language.

The results refer to a period when the only available technology in India was effectively 2G. In recent years, the country has made advancements towards the expansion of 3G/4G mobile services and the universal availability of broadband Internet. These improvements 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 but they can also further exacerbate the gap between those who can and cannot access agricultural advice. This will make it increasingly important to ensure universal access to mobile phones and information in the years to come.

Figures and Tables

Figure 3.1. 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 3.2. Timeline of tower construction under SMIS Phase I

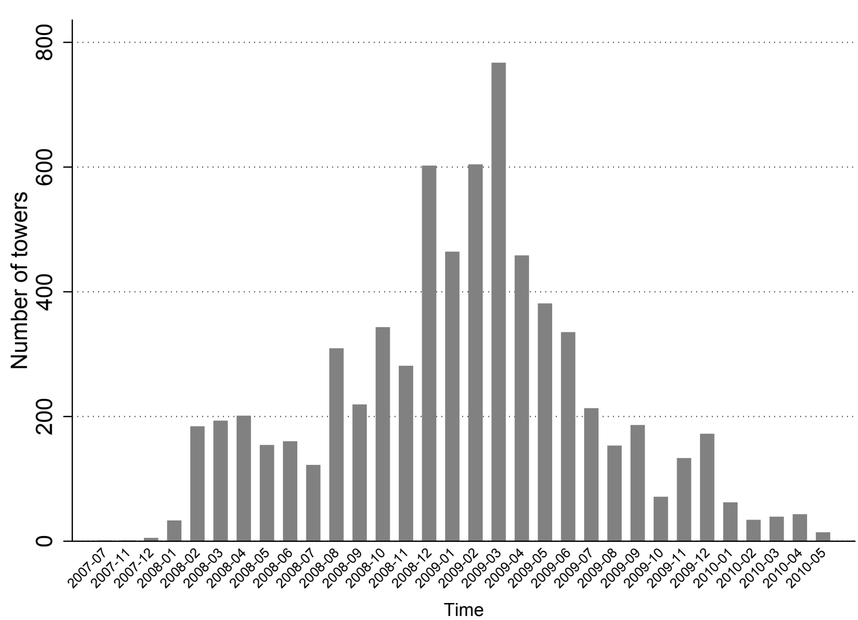
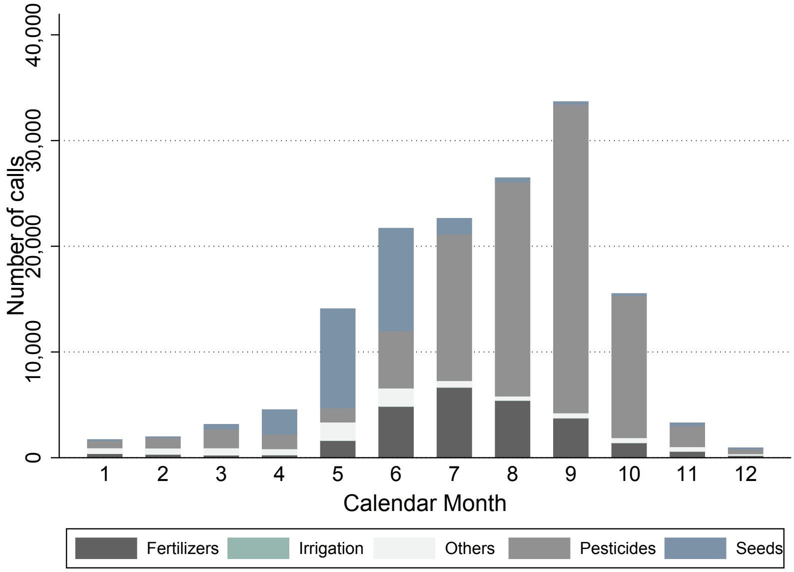
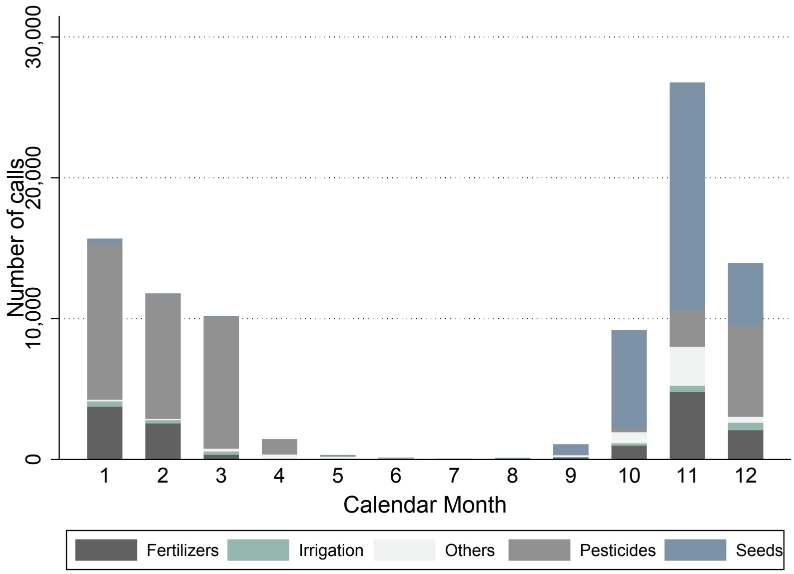


Figure 3.3. Distribution of calls on rice and wheat across agricultural cycle

(a) Rice



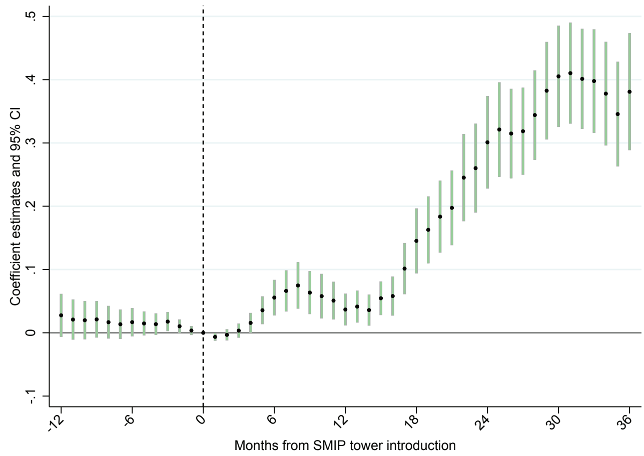
(b) Wheat



Notes: Source: Kisan Call Center, Ministry of Agriculture

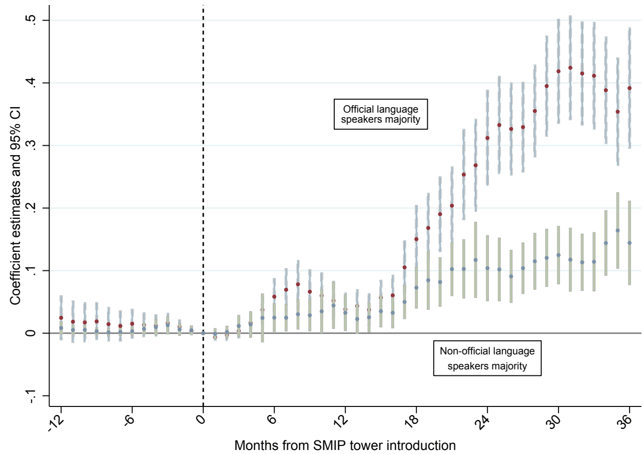
Figure 3.4. Farmers' Calls to KCC relative to Tower Construction - Event Study

(a) Average effects



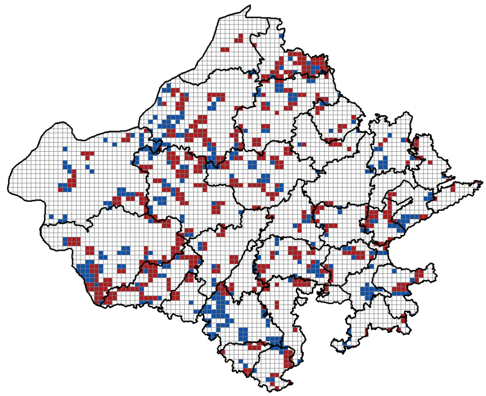
Notes: The figure plots the coefficients obtained with the following specification $Ln\,(1 + \text{calls})_{it} = \alpha_i + \alpha_t + \sum_{k=-12}^{+36} \beta_k D_{it}^k + \varepsilon_{it}$. Where i cell, t month, D_{it}^k dummy equal to 1 if month $t = k$ for cell i

(b) Heterogeneous Effects by Language



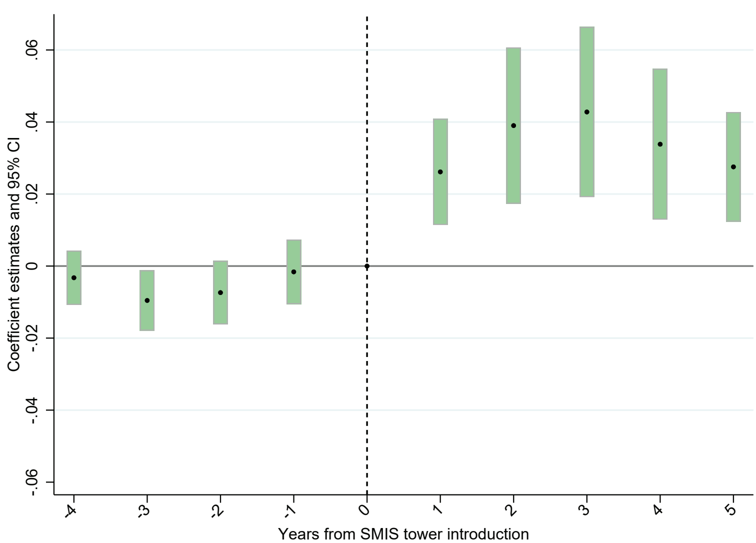
Notes: The figure plots the coefficients obtained with the following specification $Ln\,(1 + \text{calls})_{it} = \alpha_i + \alpha_t + \sum_{k=-12}^{+36} \beta_k D_{it}^k + \varepsilon_{it}$. Where i cell, t month, D_{it}^k dummy equal to 1 if month $t = k$ for cell i. We estimate this specification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

Figure 3.5. 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 SMIS Phase I. Control cells are those that are proposed *and not* covered by mobile tower under SMIS Phase I.

Figure 3.6. Agricultural productivity relative to Tower Construction: Event Study



Notes: The figure plots the coefficients β_k obtained with the following specification $\log(yield)_{it} = \alpha_i + \alpha_t + \sum_{k=-4}^{+5} \beta_k D_{it}^k + \varepsilon_{it}$. Where i denotes cell, t denotes year, D_{it}^k dummy equal to 1 if year t is k years after (or before) the construction of first SMIS tower in cell i.

Table 3.1. SMIS coverage ($\mathbb{1}$ (Tower)) and cell characteristics (Balance Test)

Dependent variable:	1(**Tower**)				1(N.O. 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 SMIS 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 3.2. 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 SMIS 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 3.3. 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 SMIS 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 3.4. 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 $\times$ 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 SMIS 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 3.5. 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 SMIS 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 3.6. Mobile Coverage and long-run agricultural productivity

Outcome:	Δ log(yield) (2007-2017)					
			by baseline productivity (2007):			
			First Quartile	Second Quartile	Third Quartile	Fourth Quartile
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Coverage	0.053	0.068	0.055	0.004	-0.064	-0.059
	[0.024]	[0.025]	[0.029]	[0.048]	[0.060]	[0.454]
Δ Coverage $\times$ Non-official Languages (%)		-0.117	-0.030	-0.028	-0.199	0.769
		[0.043]	[0.033]	[0.028]	[0.594]	[6.127]
Non-official Languages (%)		-0.039	-0.007	-0.014	-0.077	0.180
		[0.017]	[0.009]	[0.012]	[0.213]	[1.426]
Observations	5,023	5,023	1,254	1,170	1,181	1,254
District f.e.	✓	✓	✓	✓	✓	✓
Baseline Controls	✓	✓	✓	✓	✓	✓
Other Controls	✓	✓	✓	✓	✓	✓

Notes: The table reports IV-2SLS estimates of the effect of mobile phone coverage on changes in long-run agricultural productivity between 2007-2017. Column (1) reports average effects. Column (2) reports heterogeneous effects depending on share of cell's population speaking non-official languages. Columns (3)-(6) report heterogeneous effects depending on the baseline productivity levels in 2007. Column (3) considers cells in the lowest quartile of baseline productivity and Column (6) 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 SMIS 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).

BIBLIOGRAPHY

Agarwal, S., S. Alok, P. Ghosh, S. Ghosh, T. Piskorski, and A. Seru (2017). Banking the unbanked: What do 255 million new bank accounts reveal about financial access? Technical report.

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

Allen, T. and D. Atkin (2016). Volatility and the gains from trade. *Working Paper*.

Atkeson, A. and A. Burstein (2008). Pricing-to-Market, Trade Costs, and International Relative Prices. *American Economic Review 98*(5), 1998–2031.

Atkin, D., A. Chaudhry, S. Chaudhry, A. K. Khandelwal, and E. Verhoogen (2015). Markup and Cost Dispersion across Firms: Direct Evidence from Producer Surveys in Pakistan. *American Economic Review 105*(5), 537–544.

Atkin, D. and D. Donaldson (2015). Who's Getting Globalized? The Size and Implications of Intranational Trade Costs. *Working Paper*.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? Technical Report 1.

Burdzy, K., D. M. Frankel, and A. Pauzner (2001). Fast equilibrium selection by rational players living in a changing world. *Econometrica 69*(1), 163–189.

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.

Carlton, D. W. and J. M. Klamer (1983). The need for coordination among firms, with special reference to network industries. *The University of Chicago Law Review 50*(2), 446–465.

Casaburi, L., M. Kremer, S. Mullainathan, and R. Ramrattan (2014). Harnessing ict to increase agricultural production: Evidence from kenya. *Harvard University*.

Cole, S. A. and A. Fernando (2016). Mobile-izing agricultural advice: Technology adoption, diffusion and sustainability.

Comin, D. A., M. Dmitriev, and E. Rossi-Hansberg (2012). The spatial diffusion of technology. Technical report, National Bureau of Economic Research.

Conley, T. G. and C. R. Udry (2010). Learning about a new technology: Pineapple in ghana. *American economic review 100*(1), 35–69.

Crouzet, N., A. Gupta, and F. Mezzanotti (2021). Shocks and Technology Adoption: Evidence from Electronic Payment Systems. *Working Paper*.

Dalrymple, D. G. et al. (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 Loecker, J. and P. K. Goldberg (2014). Firm Performance in a Global Market. *Annual Review of Economics 6*(1), 201–227.

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik (2016). Prices, Markups, and Trade Reform. *Econometrica 84*(2), 445–510.

De Loecker, J. and F. Warzynski (2012). Markups and Firm-Level Export Status. *American Economic Review 102*(6), 2437–71.

Domar, E. D. (1961). On the Measurement of Technological Change. *The Economic Journal 71*(284), 709–729.

Edmond, C., V. Midrigan, and D. Y. Xu (2015). Competition, Markups, and the Gains from International Trade. *American Economic Review 105*(10), 3183–3221.

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.

Faber, B. and T. Fally (2020). Firm Heterogeneity in Consumption Baskets: Evidence from Home and Store Scanner Data. *Working Paper*.

Fabregas, R., M. Kremer, M. Lowes, R. On, and G. Zane (2019). SMS-extension and Farmer Behavior: Lessons from Six RCTs in East Africa. Technical report, Working paper.

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. *Annu. Rev. Econ. 2*(1), 395–424.

Foster, L., J. Haltiwanger, and C. Syverson (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review 98*(1), 394–425.

Goldberg, P. K. and R. Hellerstein (2008). A Structural Approach to Explaining Incomplete Exchange-Rate Pass-Through and Pricing-to-Market. *American Economic Review 98*(2), 423–29.

Gollin, D., D. Lagakos, and M. E. Waugh (2014). Agricultural productivity differences across countries. *American Economic Review 104*(5), 165–70.

GSMA (2020). Digital Agriculture Maps: 2020 State of the Sector in Low and Middle-Income Countries. *London: GSMA*.

Gupta, A. (2021). Demand For Quality, Variable Markups and Misallocation: Evidence from India. *Working Paper*.

Haltiwanger, J., R. Kulick, and C. Syverson (2018). Misallocation Measures: The Distortion That Ate the Residual. *Working Paper*.

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). Market inefficiencies and the adoption of agricultural technologies in developing countries.

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.

Katz, M. and C. Shapiro (1994). Systems competition and network effects. *Journal of Economic Perspectives 77*(77), 93–115.

Katz, M. L. and C. Shapiro (1985). Network externalities, competition, and compatibility. *The American economic review 75*(3), 424–440.

Katz, M. L. and C. Shapiro (1986). Technology adoption in the presence of network externalities. *Journal of Political Economy 94*(4), 822–841.

Kugler, M. and E. Verhoogen (2011). Prices, Plant Size, and Product Quality. *The Review of Economic Studies 79*(1), 307–339.

Linder, S. B. (1961). An Essay on Trade and Transformation.

Mahajan, V. and R. K. Gupta (2011). Non Farm Opportunities for Smallholder Agriculture.

Matsuyama, K. (1995). Complementarities and cumulative processes in models of monopolistic competition. *Journal of Economic Literature 33*(2), 701–729.

Melitz, M. J. and G. I. P. Ottaviano (2008). Market Size, Trade, and Productivity. *The Review of Economic Studies 75*(1), 295–316.

Murphy, K. M., A. Shleifer, and R. W. Vishny (1989). Industrialization and the big push. *Journal of Political Economy 97*(5), 1003–1026.

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.

Peters, M. (2020). Heterogeneous Markups, Growth, and Endogenous Misallocation. *Econometrica 88*(5), 2037–2073.

Petrin, A. and J. Levinsohn (2012). Measuring aggregate productivity growth using plant-level data. *The RAND Journal of Economics 43*(4), 705–725.

Rauch, J. E. (1999). Networks versus markets in international trade. *Journal of International Economics 48*(1), 7 – 35.

Restuccia, D. and R. Rogerson (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics 11*(4), 707 – 720.

Rosenstein-Rodan, P. N. (1943). Problems of industrialisation of eastern and south-eastern europe. *The Economic Journal 53*(210/211), 202–211.

Rotemberg, M. (2019). Equilibrium Effects of Firm Subsidies. *American Economic Review 109*(10), 3475–3513.

Rysman, M. (2007). An empirical analysis of payment card usage. *The Journal of Industrial Economics 55*(1), 1–36.

Weyl, E. G. and M. Fabinger (2013). Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition. *Journal of Political Economy 121*(3), 528–583.

World Bank (2014). *World development indicators 2014*. World Bank Publications.

World Bank (2019). *Harvesting Prosperity: Technology and Productivity Growth in Agriculture*. The World Bank.

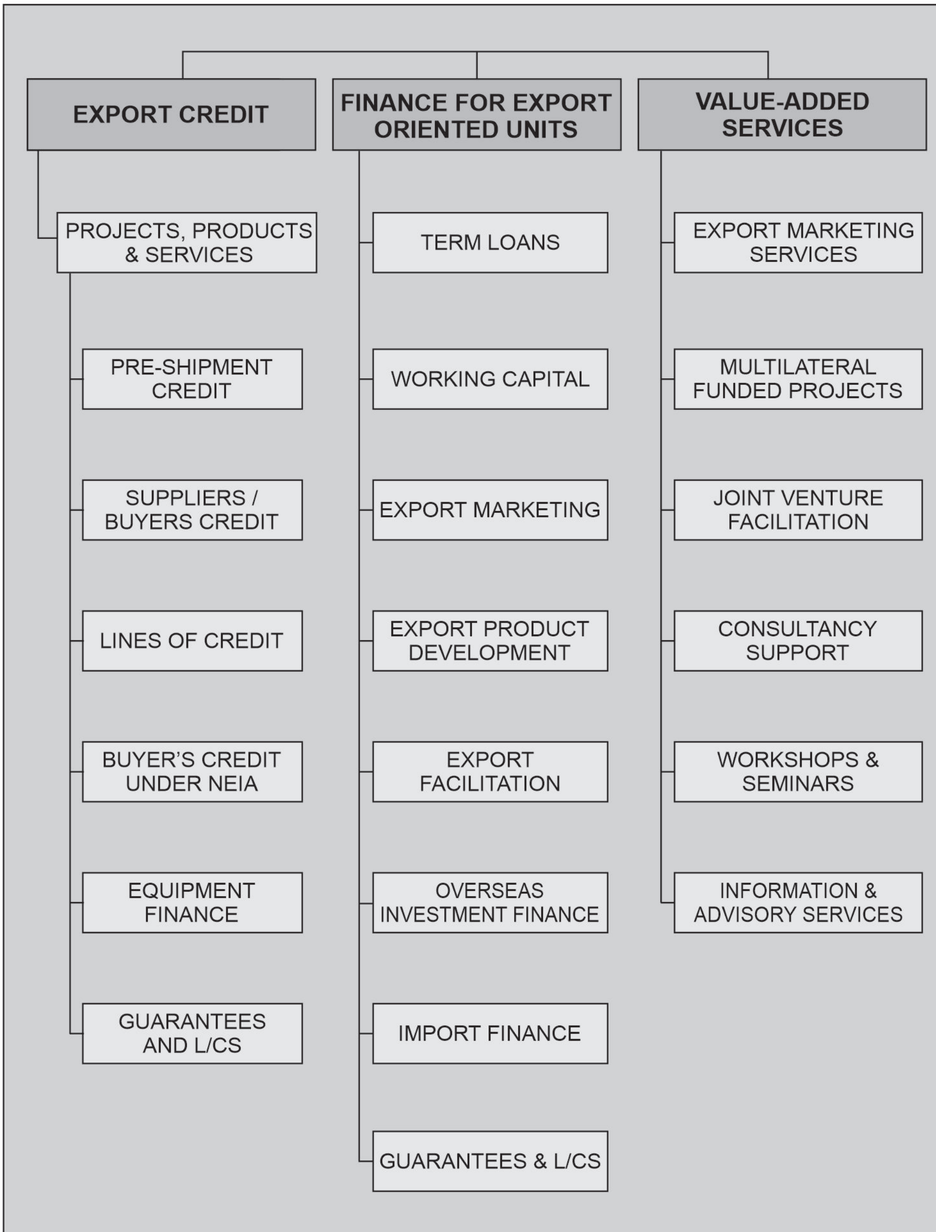
Zhelobodko, E., S. Kokovin, M. Parenti, and J.-F. Thisse (2012). Monopolistic Competition: Beyond the Constant Elasticity of Substitution. *Econometrica 80*(6), 2765–2784.

OP No.	Title
RECENT OCCASIONAL PAPERS	
141.	Comesa (Common Market for Eastern and Southern Africa): A Study of India's Trade and Investment Potential
142.	Indian Shipping Industry: A Catalyst for Growth
143.	New Renewable Energy in India: Harnessing the Potential
144.	Caribbean Community (Caricom): A Study of India's Trade and Investment Potential
145.	West African Region: A Study of India's Trade and Investment Potential
146.	India's Trade and Investment Relations with LDCs (Least Developed Countries): Harnessing Synergies
147.	Indian Electronic Industry : Perspectives and Strategies
148.	Export Potential of Indian Plantation Sector: Prospects and Challenges
149.	Mercosur: A Study of India's Trade and Investment Potential
150.	Openness and Growth of the Indian Economy: An Empirical Analysis
151.	The Commonwealth: Promoting a Shared Vision on Trade and Investment
152.	Southern African Development Community (SADC): A Study of India's Trade and Investment Potential
153.	Strategic Development of MSMEs: Comparison of Policy Framework and Institutional Support Systems in India and Select Countries
154.	Indian Chemical Industry : Exploring Global Demand
155.	Technological Interventions In Indian Agriculture for Enhancement of Crop Productivity
156.	Exports of Services and Offshore Outsourcing: An Empirical Investigation in the Indian Context
157.	Indian Ocean Rim Association for Regional Co-operation (IOR-ARC): A Study of India's Trade and Investment Potential
158.	West Africa: A Study of India's Trade and Investment Potential
159.	The Effects of Financial Openness: An Assessment of the Indian Experience
160.	Comparison of Labour Laws: Select Countries
161.	India's Trade and Investment Relations with Cambodia, Lao PDR, Myanmar, Vietnam (CLMV): Enhancing Economic Cooperation
162.	Indian Horticulture-Imperatives to Enhance Trade from India
163.	India's Trade and Investment Relations with Gulf Cooperation Council (GCC): Strengthening Economic Ties
164.	India's Hi-Tech Exports: Potential Markets and Key Policy Interventions
165.	Outward Direct Investment from India: Trends, Objectives and Policy Perspectives
166.	East African Community (EAC): A Study of India's Trade and Investment Potential
167.	Trade Liberalization, Product Variety and Growth
168.	Research & Development in BRICS: An Insight
169.	Indian Capital Goods Industry: A Sector Study
170.	Bangladesh: A Study of India's Trade and Investment Potential
171.	Indian Electronic Goods Industry: Neutralizing Trade Deficit with China

172. Indian Steel Industry: Export Prospects
173. Value Addition Chains and Trade in Manufactured Commodities in South-East Asia
174. Potential for Trade of Organic Products from India
175. Multilateral Development Bank- Funded Project: Trends and Opportunities for Indian Exports
176. Indian Pharmaceutical Industry: Challenges and Prospects
177. Migration and Remittances in India
178. Technical Progress and Structural Change: The Roles of Demand and Supply in Economic Growth
179. Inter-linkages between Exports and Employment in India
180. India's Engagements with CLMV: Gateway to ASEAN Markets
181. Export Promotion From Rajasthan: Key Insights And Policy Suggestions
182. Potential for Enhancing Exports from Andhra Pradesh
183. The Internationalisation of Indian firms through outbound Foreign Direct Investment: Nature, determinants and developmental consequences
184. Financialization and its Implications on the Determination of Exchange rate of Emerging Market Economies
185. Export from Uttar Pradesh: Trends, Opportunities and Policy Perspectives
186. Essays on International Trade, Welfare and Inequality
187. Essays in Indian Trade Policy
188. Exchange Rate Dynamics and its Impact on India's Exports to USA and EU: An Assessment
189. Liberalisation, Wages and Sector Growth : General Equilibrium Analysis for India
190. Manufacturing in SADC: Moving Up the Value Chain
191. North Africa: Unlocking India's Trade and Investment Potential
192. Essays on Education and Institutions in Developing Countries
193. Project Exports From India: Strategy for Reenergizing and Reorienting
194. Packaging Sector: Potential And Way Forward
195. Realizing India's Trade & Investment Potential with South Korea
196. India & Central Asia: Reinitalizing Trade & Investment Relations
197. AfCFTA: Opportunities for India in Africa's Economic Integration
198. India-CLMV: Building Supply Chains in Asia
199. Self-Reliant India- Approach and Strategic Sectors to Focus
200. Essays in Trade and Development Economics
201. The Political Origin and Firm-level Consequences of Bank Proliferation in China
202. Building Infrastructure in CLMV: Opportunities for India
203. Building Value Chain: Opportunities for India and ASEAN
204. Essays in International Trade in Post-Liberalization India
205. Indian Apparels - A Sector Study
206. India Engagements in FTAs"Looking Ahead
207. Indian Dairy Sector : Recent Trends and Prospects for Greater Export Orientation
208. Export Potential of India's Shrimps and Prawns
209. Empirical Studies in International Trade
210. Project Exports from India: Tapping Potential AMID Changing Dynamics

INDIA EXIM BANK'S MAJOR PROGRAMMES

Bank's Major Programmes



EXPORT-IMPORT BANK OF INDIA

HEAD OFFICE

Centre One Building, 21st Floor, World Trade Centre Complex, Cuffe Parade, Mumbai 400 005.
Phone: (91 22) 22172600 Fax : (91 22) 22182572
E-mail : ccg@eximbankindia.in Website: www.eximbankindia.in

LONDON BRANCH

5th Floor, 35 King Street, London EC2V 888 United Kingdom
Phone : (0044) 20 77969040 Fax : (0044) 20 76000936 E-Mail :eximlondon@eximbankindia.in

DOMESTIC OFFICES

Ahmedabad
Sakar II, 1st Floor,
Next to Ellisbridge Shopping Centre,
Ellisbridge P. 0., Ahmedabad 380 006
Phone : (91 79) 26576852
Fax : (91 79) 26577696
E-mail : eximahro@eximbankindia.in

Bengaluru
Ramanashree Arcade, 4th Floor,
18, M. G. Road,
Bengaluru 560 001
Phone : (91 80) 25585755
Fax : (91 80) 25589107
E-mail : eximbro@eximbankindia.in

Chandigarh
C- 213, Elante offices, Plot No. 178-178A,
Industrial Area phase 1,
Chandigarh 160 002
Phone : (91 172) 4629171/73
Fax : (91 172) 4629175
E-mail : eximcro@eximbankindia.in

Chennai
Overseas Towers, 4th and 5th Floor,
756-L, Anna Salai, Chennai 600 002
Phone : (91 44) 28522830/31
Fax : (91 44) 28522832
E-mail : eximchro@eximbankindia.in

Guwahati
NEDFi House, 4th Floor, GS Road,
Dispur, Guwahati 781 006
Phone : (91 361) 2237607 /609
Fax : (91 361) 2237701
E-mail : eximgro@eximbankindia.in

Hyderabad
Golden Edifice, 2nd Floor,
6-3-639/640, Raj Bhavan Road,
Khairatabad Circle, Hyderabad 500 004
Phone : (91 40) 23307816
Fax : (91 40) 23317843
E-mail : eximhro@eximbankindia.in

Kolkata
Vanijya Bhawan, 4th Floor,
(International Trade Facilitation Centre),
1/1 Wood Street,
Kolkata 700 016
Phone : (91 33) 68261301
Fax : (91 33) 68261302
E-mail : eximkro@eximbankindia.in

Mumbai
8th Floor, Maker Chamber IV,
Nariman Point,
Mumbai 400 021
Phone : (91 22) 22861300
Fax : (91 22) 22182572
E-mail : eximmro@eximbankindia.in

New Delhi
Office Block, Tower 1, 7th Floor,
Adjacent Ring Road, Kidwai Nagar (E)
New Delhi - 110 023
Phone : (91 11) 61242600 / 24607700
Fax : (91 11) 20815029
E-mail : eximndo@eximbankindia.in

Pune
No. 402 & 402(B), 4th floor,
Signature Building, Bhamburda,
Bhandarkar Rd., Shivajinagar,
Pune - 411 004
Phone : (91 20) 26403000
Fax : (91 20) 25648846
E-mail : eximpro@eximbankindia.in

OVERSEAS OFFICES

Abidjan
5th Floor,
Azur Building,
18-Docteur Crozet Road,
Plateau,
Abidjan,
Côte d'Ivoi re
Phone : (225) 2720242951
Fax : (225) 2720242950
Email : eximabidjan@eximbankindia.in

Addis Ababa
House No. 46,
JakRose Estate Compound,
Woreda 07,
Bole Sub-city,
Addis Ababa,
Ethiopia.
Phone : (251) 118222296
Fax : (251) 116610170
Email : aaro@eximbankindia.in

Dhaka
Madhumita Plaza, 12th Floor,
Plot No. 11, Road No. 11, Block G,
Banani, Dhaka, Bangladesh - 1213.
Phone : (88) 01708520444
E-mail : eximdhaka@eximbankindia.in

Dubai
Level 5, Tenancy lB,
Gate Precinct Building No. 3,
Dubai International Financial Centre,
PO Box No. 506541, Dubai, UAE.
Phone : (971) 43637462
Fax : (971) 43637461
E-mail : eximdubai@eximbankindia.in

Johannesburg
2nd Floor, Sandton City Twin Towers East,
Sandhurst Ext. 3, Sandton 2196,
Johannesburg, South Africa.
Phone : (27) 113265103
Fax : (27) 117844511
E-mail : eximjro@eximbankindia.in

Singapore
20, Collyer Quay, #10-02,
Tung Centre, Singapore 049319.
Phone : (65) 65326464
Fax : (65) 65352131
E-mail : eximsingapore@eximbankindia.in

Washington D.C.
1750 Pennsylvania Avenue NW,
Suite 1202,
Washington D.C. 20006,
United States of America.
Phone : (1) 2022233238
Fax : (1) 2027858487
E-mail : eximwashington@eximbankindia.in

Yangon
House No. 54/A, Ground Floor,
Boyarnyunt Street, Dagon Township,
Yangon, Myanmar
Phone : (95) 1389520
E-mail : eximyangon@eximbankindia.in

India Exim Bank

Centre One Building, 21st Floor, World Trade Centre Complex, Cuffe Parade,
Mumbai - 400005 | Ph.: (91 22) 22172600 | Fax: (91 22) 22182572
E-mail: ccg@eximbankindia.in | Website: www.eximbankindia.in, www.eximmitra.in
Follow us on