Digitization and Development: Formalizing Property Rights and its Impact on Land and Labor Allocation

beg, Sabrin

Lerner College of Business and Economics, University of Delaware

29 January 2021

Digitization and Development: Property Rights Security, and Land and Labor Markets*

Sabrin Beg†

Last updated: January 29, 2021

Abstract

I test the land and labor market effects of a property rights reform that computerized rural land records in Pakistan, making digitized records and automated transactions accessible to agricultural landowners and cultivators. Using the staggered roll-out of the program, I find that while the reform does not shift land ownership, landowning households are more likely to rent out land and shift into non-agricultural occupations. At the same time, cultivating households have access to more land, as rented in land and overall farm size increase. I construct measures of farmer-level TFP and marginal product of land, and demonstrate evidence of improved allocative efficiency as land is redistributed towards more productive farmers. Aggregate district-level production data suggest a reduction in the dispersion of marginal products of land and an improvement in productivity. The results have implications for both the allocation of land across farmers and the selection of labor into farming, demonstrating that agricultural land market frictions present a constraint to scale farming and structural change in developing countries.

Keywords: Property Rights, Rural Mobility, Agricultural Land Markets, ICT in Development, Institutions, Misallocation

*I thank Imran Rasul (managing editor) and three anonymous referees for extensive comments. I am also grateful for useful feedback from Chris Udry, Marc Bellemare, Adrienne Lucas, Emily Oster, Laura Schechter, Shing-Yi Wang, James Fenske, Jeremy Tobacman, Ajay Shenoy, James Berry, Richard Hornbeck and seminar participants at various seminars and conferences. I am grateful to Ghazala Mansuri and the Punjab Land Records Management Information System for useful data. Attique Ur Rehman, Hana Zahir and Muhammad Zubair Rafique provided excellent research assistance. All errors are my own.

†Lerner College of Business and Economics, University of Delaware. Email: sabrin.beg@gmail.com

1 Introduction

Agricultural productivity growth is imperative for development and structural change. Developing countries, however, lag severely in aggregate agricultural productivity, despite the availability of modern and mechanized inputs.[1] Recent literature argues that misallocation of factors of production contributes to productivity differences across countries. An emerging agenda for development economists is thus to examine the causes of misallocation for unpacking agricultural productivity lags in the developing world.

Weak property rights and tenure insecurity lead to high transaction costs and market constraints that hinder the optimal allocation of productive inputs. Land market frictions thus impede the efficient trading of land and the occupational choices of individuals (de Janvry et al., 1991; Adamopoulos et al., 2017; Chen, 2017). Agricultural landowners facing restrictions in renting out or selling their land choose to farm when it might be optimal to practice a non-agricultural activity. Relatedly, barriers to purchasing or tenancy prevent productive famers from expanding the scale of operation and realizing returns to scale and mechanization. As a result, farms in lower middle income countries, including Pakistan, are small, unmechanized, and lag in productivity (Foster and Rosenzweig, 2017). Moreover, about a third to half of Pakistan's labor force works in agriculture, which constitutes only 15% of the total GDP (World Bank, 2013).

This paper establishes a causal link between tenure security and market activity, and the consequent implication for allocative efficiency, farm-scale, and productivity. Property rights and land market transactions are non-existent or excessively informal in the vast majority of developing countries. In the context of this study (Punjab, Pakistan), land records have been maintained under the same structure since the colonial period—paper-based records of millions of landowners were held by 8000 local officers or *patwaris*, who manually updated and managed these records.[2] The inefficient and dispersed land records system has led to tenure insecurity, with owners relying on the discretion of the *patwaris* for any transaction or proof of ownership and tenancy rights. These barriers to land transactions and security of property result in low mobility of land, affecting land use and labor market choices of rural landowners.

In 2009, the Punjab government launched the Land Record Management Information System to formalize and centralize land records in the province. Through this program, which was phased out in stages across all districts of the province, land records were obtained from the *patwaris*, computerized, and made available to the

[1] Gollin et al. (2002).

[2] A *Patwari* was a historically appointed officer during the British colonial government, and has persisted as an office in the present land management system.

public at a service center in each subdistrict. While no titles were given out as part of the program, an owner or tenant can go to the designated center and obtain a government-attested copy of his ownership or tenancy status, implying improved access to land records and security of rights due to the program. All land transactions and changes to ownership or tenancy are conducted digitally at this designated center. The program thus represents an overhaul of an informal system that is replaced with a more centralized and computerized system. I use the staggered rollout of the program between 2011 and 2015 to document effects of the program. Specifically, I exploit variation in the timing of program start in any district and the share of program subdistricts within a district to identify causal effects. I test the program's effect on rental market participation and labor choices of landowning households, on allocation of land across farmers and on farm operation, particularly, farm scale, input usage, and productivity. To validate the identification strategy I conduct tests to ensure early and late program districts do not demonstrate differential prior levels or trends in the main outcome variables, underlying soil quality and productivity, or macroeconomic indicators. These tests confirm that program timing or intensity are unlikely to be driven by preexisting differences across districts.

I find that the program increased rental market transactions, as landowners are more likely to rent out land. Consistent with higher rental activity, the rate of agricultural participation by landowners declines, supporting the significance of market frictions in affecting the selection of workers across sectors. Landowning households shift into non-agricultural occupations, particularly business ownership. This increase in renting out is driven by lower income households, who are more likely to face tenure insecurity and market constraints. I do not find any significant effects on land ownership or land sales and purchases, suggesting the market frictions for land sales are higher, or that renting and selling are possible substitutes.

While some landowners rent out land and exit agriculture, households that continue to cultivate increase the scale of farming as shown by more rented in land and higher average farm size in program districts. I rank farmers by productivity based on farm-level TFP calculated using detailed information on farm output and inputs. I find that higher TFP farmers in a district have greater farm land (and lower marginal product of land) after the program relative to low TFP farmers. Additionally, the dispersion in marginal products of land within a district is lower after the program. These findings support the hypothesis that market activity due to the program results in a more efficient allocation of land.

Suggestive evidence of improved input usage and investment supports the scale and allocation effects of the program. Cultivating households are marginally more likely to switch crop choice and use pesticides. The program has no effect on average farm-level yield, but a positive effect on two different measures of aggregate productivity. Remote sensing data on vegetation across subdistricts is used as a proxy for crop production; I find a significant increase in the vegetation index as the program is rolled out. Additionally, district-level data on aggregate output by crop show greater improvements in cereal yield due to the program (significant at the 10%). Taken together, the changes to land allocation, farming scale and inputs, and aggregate and remote sensing measures of agricultural output all suggest allocative efficiency and productivity improvement due to the program.

I test the robustness of the main findings in a number of ways. In additional specifications, I adjust the control variables and sample years, control for simultaneous macroeconomic trends and drop the early program districts that may be subject to selection bias. I complement the main findings using alternate identification strategies, including a 'stacked' difference-in-difference, a standard timing difference-in-difference using just timing variation, and an event study analysis. The results are highly stable across the various robustness specifications and strategies.

Misallocation in the industrial sector is documented by Hsieh and Klenow (2009) and in agriculture by Restuccia and Santaeulalia-Llopis (2017) and de Janvry et al. (1991). The documentation of misallocation of land and capital across farming entities is supported by parallel research noting the dramatic differences in the scale of farming across countries (Adamopoulos and Restuccia, 2014).[3] Adamopoulos et al. (2017) demonstrate that aggregate agricultural productivity depends not just on the allocation of land and capital across farmers, but also on allocation of workers across sectors—in particular, the type of farmers who operate in agriculture (selection). Much of the literature is focused on the extent and consequences of misallocation, and less on the sources. Existing papers highlight the role of markets in allocative efficiency by using theoretical arguments or by demonstrating a correlation between market activity and misallocation.[4]

As market activity is endogenous it is challenging to identify its role in factor allocation. Chen et al. (2017) use variation in the degree of land certification in Ethiopia to show that land rentals are associated with lower misallocation and higher agricultural productivity. Chari et al. (2017) demonstrate that legalizing land rentals in China improves the allocation of land across farmers and boosts aggregate productivity. In both these contexts land is communally or state owned, and therefore the status quo is characterized by a lack of any land market. By demonstrating improved allocative efficiency and productivity as a result of legalizing land rentals, these papers provide a justification for private property rights. However, even with private property rights, tenure insecurity can be high under informal or partially enforceable rights and

[3] There is a 34-fold difference in average farm size (land per farm) between rich and poor countries.

[4] For instance, the extent of inefficiency is larger on farms without marketed land in Malawi (Restuccia and Santaeulalia-Llopis, 2017).

contracts causing significant market frictions. This is apparent in the context Punjab, where only 20% of landowners lease their land and over 80% of farms are under 10 acres.

Theoretical work on property rights testifies to the role of tenure security on resource allocation. Besley and Ghatak (2010) identify two broad channels through which property rights affect allocation: first, limiting expropriation; and second, facilitating market transactions.[5] Empirically, the positive effects of land titling and certification programs on 'limiting expropriation' and incentivizing investment are well documented.[6] Less consistent evidence has been documented for the theoretical argument that tenure security facilitates market activity. Field and Torero (2006), Do and Iyer (2008) and Galiani and Schargrodsky (2010) do not find that titling significantly improves credit access, while Wang (2012), Carter and Olinto (1996) and López and Romano (2000) argue that that they do. Deininger and Goyal (2012) find that land registry computerization in India increases credit access, though the effects are modest and only in urban areas. The existing literature lacks comprehensive evidence of how tenure insecurity affects land rental and sales in particular.[7] Even fewer papers systematically identify the effect of property rights and security on labor choices, particularly in rural areas.[8] This paper fills this gap in the property rights literature by documenting the benefits of a land rights computerization program in progressing tenure security, and facilitating land rental and labor market allocation.

Specifically, I make two major contributions to the extensive body of empirical literature on property rights and misallocation. First, I provide direct evidence of the role of property rights insecurity in hindering agricultural land rental. I build on former work by demonstrating frictions in land rental activity even with privately owned property. The second contribution of my paper is the additional effects that I document on labor allocation of landowning households as rental transaction costs go down. These contributions depart from the focus of the existing property rights literature on their effect on investment incentives, and are complementary to the broader literature on property rights institutions and agricultural productivity (Bellemare, 2013; Newman et al., 2015; Gottlieb and Grobovšek, 2019).

These findings also contribute to understanding the process of structural change and urbanization in the context of South Asia (Binswanger-Mkhize, 2013). Agricultural

[5] Chen (2017) offers additional theoretical support by demonstrating that untitled land cannot be traded across farmers, creating land misallocation and distorting individuals' occupational choice between farming and working outside agriculture.

[6] See Field (2007); Do and Iyer (2008); Galiani and Schargrodsky (2010); Deininger et al. (2011); Ali et al. (2014); Feder (1988); Besley (1995); Goldstein and Udry (2008); Hornbeck (2010).

[7] Deininger et al. (2010) and Lunduka et al. (2010) provide evidence suggesting tenurial insecurity prevents the efficient functioning of the land rental market in Ethiopia and Malawi. Macours et al. (2010) find that tenurial insecurity constrains the matching of landlords and tenants in Nicaragua, affecting contractual outcomes.

[8] de Janvry et al. (2015) find that land certificates in Mexico increases the likelihood of households to have a migrant member.

participation is still considerably high in South Asian countries — approximately 50% of the total labor force in India, Pakistan and Bangladesh (compared to 24% for middle income countries) (World Bank, 2013). On the other hand agriculture accounts for just 18% of the GDP on average for South Asia.[9] Improving tenancy security and rights of land use can stimulate labor market allocation and structural transformation. The program further highlights that Information and Communication Technology (ICT) in governance and public service delivery holds substantial promise for lower income-nations with limited state capacity (Banerjee and Jain, 2003; Ghosh and Banerjee, 2006), and contribute to the literature on the positive impacts of digitization broadly on productivity and development (Bresnahan et al., 2002; Bloom et al., 2014).

The next section describes the background of land records in Punjab and the Land Record Management and Information Systems program. Section 3 describes the data and empirical specification for the main results, and Section 4 describes the results and mechanisms. Section 5 discusses the validity of findings and offers additional robustness checks. Finally, Section 6 concludes.

2 Background[10]

2.1 Agriculture and Land Records in Punjab

Punjab, the context of the study, is the most populated province of Pakistan with 80.5 million inhabitants (55.6% of the country's population), 70% of whom live in rural areas. The Board of Revenue bears responsibility for the administration of agricultural land, which is mostly privately owned. The history of the land revenue system in Pakistan dates to pre-colonial rulers who introduced a system of land administration, which was improved and formalized by the British colonial government and then underwent minimal changes over a 60-year period after Pakistan's independence.

Several levels of administration are involved in land record maintenance: the District, Subdistrict, Kanungo circle, and Patwar circle. *Patwaris*, or the local officers at the Patwar Circle level, are the custodians of land rights records—in Punjab, about 8,000 *Patwaris* maintain paper-based land records pertaining to 20 million land owners, at times holding them in cloth bags. Among various land record statements described in further detail in the Appendix, the most relevant is the 'Land Right Holders Register' that lists the owners of each land parcel demarcated in a corresponding cadastral map of each village (Figures A1-A2). Any changes to land rights are recorded in a separate register of mutations, which is used to update the register of right holders every four

[9] 17%, 25%, and 16% for India, Pakistan, and Bangladesh, respectively. Consistent with the high participation in agriculture, the average proportion of rural population in South Asia is 67% of the total, a decrease since 1960 but a much slower decline compared to Latin America.

[10] Background about land record documents is based on United Nations Human Settlements Programme (2012)

years. Tenants' and landowners' rights, as well as updates that arise due to rental or sale, are thus recorded at the discretion of the *patwari* and revenue officers above him in the bureaucratic chain.

The manual and decentralized system is potentially prone to corruption and mistakes, lowering tenure security for owners and cultivators. A survey conducted by Gallup Pakistan for the Board of Revenue found that 42% of a sample of land owners and cultivators from Punjab report higher dissatisfaction with the system of land records than with other government departments. Sixty-four percent of farm households describe the system as lacking transparency, and 82% report ever having to pay a bribe to obtain land record services. Seventy-six percent of respondents in the poll reported illegal occupation of land as the main form of land dispute, and 56% identified that the major source of all land disputes was incorrect land records.

Land transactions are uncommon. In the Pakistan Rural Household Survey (2001), 87% of landowners either inherited their land or obtained it from the government. Universal certification or titling is not prevalent, though official documents can be obtained based on verification by land revenue officials, albeit through a lengthy process. A request for obtaining a title begins at the *patwari* level and goes up the bureaucratic chain to the revenue office. Land ownership is verified by the revenue office through correspondence with the *patwari* who locates and confirms the rights of the landholder in his manual records. After verification, senior revenue officials issue a title to the landowner.

Among the rural landowning household sample from 2001, only 45% have a *'fard'* (title) or an ownership document on a registered stamp paper for their property. Of the owners with title documents, only 25% report not having to submit payment to a revenue official beyond the legal title registration fee.[11] Even for those with titles tenure security may be low as land records are dispersed and not easily accessible or verifiable. Eleven percent of households report they cannot sell their land if they wanted to. In 2010, with a similar land administration system, the Government of Punjab in India made attempts to abolish colonial posts like *patwaris* who were often accused of corruption and making 'fraudulent changes' in revenue records under their jurisdiction (Sural, 2013).

2.2 Land Record Management and Information Systems Program

Beginning in the years 2005–2009, the Government of Punjab received financial support from the World Bank to begin the computerization of land records to improve service delivery and enhance the perceived level of tenure security. The main objective of this endeavor was to facilitate increased access to land records at low costs, specifically for

[11] The remainder report paying an illegal fee or do not respond as this payment is illegal by design.

the poorest and least-connected households. The provincial government department noted that:

> Inequalities of land distribution, tenure insecurity and difficulties associated with the land administration and registration system are closely interrelated and continue to impose significant constraints on both rural and urban populations, particularly the poor. Land transactions are relatively high cost, and disputes about accuracy of land rights are caused, among others, by the inefficient and dispersed land records system. As a result land markets are thin and land prices are in excess of the discounted value of potential agricultural earnings from land. (World Bank – Project Information Document 2005)

The first objective of the program was to computerize all rural land records, including the list of land right holders (owners and cultivators), as maintained by the *patwaris*. The second objective was to establish a service center in each subdistrict of the province to host these records and to replace the lower-level land record officers for maintaining and updating these records, and providing citizens with land mutation, title issuance, and other land record related services directly. The computerized records establish both the identity of the owner and tenant, and can be located on the internet or obtained from the designated service centers.

The right holders (owners or tenants) can visit a service center where the staff can use their national ID number to search and verify their record, providing the client with a government-attested copy within minutes. Any mutation, due to to sale, transfer, or inheritance, is to be registered at the same service center. One hundred and fifty centers across the province now provide automated land records services, reducing the average time required to complete transactions from 2 months to 45 minutes (Gonzalez, 2016).

While the service centers increased access to digitized records of ownership and cultivation rights, they may have increased distance to land records. Initially, a *patwari* was available for each *patwar* circle, which comprises a few proximate villages, and was well-known to all village members in his jurisdiction—once all service centers are fully operational, only one center is available per subdistrict.[12] Though it may seem individuals face higher travel costs to access the centralized records, transacting parties were required to visit district revenue offices at several stages of any transaction and thus incurred high travel costs even prior to the reform. As the service centers provided all land transaction services at one location, the time and distance costs could effectively be lower after the program even with fewer service centers than *patwaris*. Changes to cultivation, for instance in the case of land rental, are still initially reported to the *patwari*, who then sends updated records to service centers at the beginning of

[12]Even though the *patwar's* role is not abolished, 150 service centers took on the tasks provided previously by approximately 8000 *patwaris*.

each agricultural season. Rental transactions thus do not entail the travel costs, but the records and rights are transitioned from being manual and disaggregated to digital, central, and verifiable.

The program thus resulted in two main changes to the pre-existing system: (1) centralized record keeping for ownership and tenancy rights, and (2) low cost and centralized land transactions including access to title documents. By making the computerized land record centrally available at a subdistrict level, the new system decreased the influence of the local officers and *patwaris* and can have potential effects on tenure security of owners and tenants, and consequently on the land market.

3 Empirical Strategy and Data

3.1 Data

Program Rollout: The program data are obtained from the Board of Revenue of Punjab, outlining the operational date for each subdistrict level service center in the province. Using the district boundaries from the pre-program period, there are 34 districts and 150 subdistricts in total. All 150 land records service centers opened between 2011 and 2015. Figure 1 shows the rollout of the service center openings and Figure 2 shows the number of subdistrict centers opening in each year. I construct $ProgramIntensity_{dt}$ in any year t and district d as the share of subdistricts in d that have received that program by year t. Table A1 shows the average values of $ProgramIntensity_{dt}$ over the sample period. The government sought to roll out the service centers in no specific order though not strictly randomly, and the proposed identification strategy will alleviate any selection bias.

Household Outcomes: The household data are obtained from Household Income and Expenditure surveys (HIES), conducted bi-yearly across the country—I use 5 HIES survey rounds from 2005 to 2015. These surveys, conducted in 2005–6, 2007–8, 2011–12, 2013–14 and 2015–16, collect demographics, employment, expenditure, and saving information from a repeated cross-section of approximately 6,600 (3,800 rural) households from Punjab in each round.[13] Thus, the data set has 19,067 rural households across 5 data rounds, and I focus on agricultural households. Specifically, I report outcomes for landowners (households that own agricultural land) and cultivators (households that operate a farm). There are 7,597 landowning households and 7,256 cultivating households.[14] Summary statistics from the household data used for

[13]In addition to the HIES, the social living standard measurement (PSLM) surveys interview 80,000 households (nationally) and collect information on demographics, employment, access to public services, and key social indicators. The HIES has a larger questionnaire and smaller sample, while the PSLM has a larger sample but does not contain key farm related data. For this reason, I use the HIES in the main regressions, but show additional analysis in the Appendix using data from the PSLM for outcomes measured in both surveys.

[14]There is significant overlap between the landowning and cultivating sample, as approximately 80% of the

the analyses are shown in Table 1.

To test for the program effect on allocative efficiency, I construct household level TFP for cultivating households in the sample using reported data on total farm area cultivated, farm output, and input expenditures. I first assume inputs can be converted to output through the following Cobb-Douglas function:

$$\log Output_i = \theta_l \log l_i + \theta_k \log k_i + \theta_h \log h_i$$

Land l is the number of acres of operational farm size (whether owned or otherwise). Labor h is the sum of hired and family labor in number of days. Hired labor is the household's expenditures on farm workers divided by the median daily wage in the district for any survey year. I count total number of days worked by household members who report being unpaid family laborers with primary occupations in farming to obtain the measure of family labor in days. Since family members only report days worked in the last month, I use the median number of months worked by agricultural farm workers to arrive at the total family labor days in the previous year.

Capital k is the sum of the value of rented capital and expenses, adjusting for owned capital (expenses include intermediate inputs, like seeds, pesticides and fertilizer). I do not have a measure of value of owned farmed machinery, so I use a proxy ψ calculated by regressing the log of total observed farm output on indicators for ownership of various assets.[15] I then assume that the value of expenses and owned capital are separable and have the same factor share in production.

To compute $\log Output$, I choose $\theta_k = 0.11$, $\theta_l = 0.25$, and $\theta_h = 0.31$ from the input elasticities for capital (k), land (l), and labor (h) calculated in Shenoy (2017) for rice farmers in Thailand.[16]

Weather variability and farmer-specific shocks are expected to be important factors contributing to realized output. To account for weather shocks common to all farmers in the same village, I regress the value of the difference between logged realized output Y_i and computed $\log Output_i$ from the above production function on village-year fixed effects. I use the residual from this regression as my measure of farm TFP (in logs). Figure 3 shows the distribution of log TFP from 2011. Given the individual TFPs for each farming household, I categorize households by quartiles of the TFP distribution in a specific district and year. I later use this ranking to test if

landowning sample also cultivates land.

[15] Assets include agricultural land, non-agricultural land, commercial land, residential land, cash savings, precious metals, and financial assets.

[16] Using these input shares assumes farming all other crops is similar to rice farming, and the TFP residual is not over- or under-estimated due to differences in input usage for farming other crops. To test for robustness, I also use the input elasticities for all crops calculated in Chari et al. (2017) and find that the TFPs from the two methods have a correlation of 0.97. In the later analysis, I rank farmers in a district by their productivity and examine how the program affects allocation of land across this ranking. The results are qualitatively identical when I use the alternate input shares for productivity calculation.

land is allocated towards farmers in higher TFP quartiles.[17]

Using the reported output Y_i and operational land input, the marginal product of land for each farmer is calculated as follows:

$$MPL_i = \theta_l \frac{Y_i}{l_i}$$

Agricultural Production: I obtain aggregate crop output data from the Agricultural Statistics of Pakistan, which record the overall production and area cultivated for each crop at the district and year level. I also obtain normalized difference vegetation index (NDVI), a disaggregated measure of greenness, for all the study regions and over the study time period. These remote sensing data are based on 24 images per year from an average of 29,000 pixels per subdistrict, distributed by NASA (Didan, 2015). I construct a subdistrict by year measure of the NDVI by using the following steps: I first get an average NDVI across all pixels within a subdistrict for each point the images are taken. Since the images are taken every 16 days, there are 24 such images per pixel in each year, or averaged NDVI for each subdistrict is available at 24 different points in any year. I allow for differences in agricultural seasons across regions and over time, and take the maximum NDVI across the 24 measurements over the year to obtain my measure of NDVI for each subdistrict and year. The NDVI measure is standardized to have mean 0 and standard deviation 1 across the subdistricts.

Soil Quality: For identification checks I obtain remote sensing data for soil quality characteristics published by the International Institute for Applied Systems Analysis (IIASA) and part of the Harmonized World Soil Database (HWSD). These evaluate soil quality according to the following criteria: nutrient availability; nutrient retention capacity; rooting conditions; oxygen availability to roots; excess salts; toxicity; and workability. For each of the seven dimensions, I compute the average value for all cells in each subdistrict's perimeter and construct an index for each subdistrict using a principal components analysis.

Administrative Data from Land Record Service Centers: I use two sources of data from the Land Record Management and Information System database. First, I have visit level records from all service centers for the year 2016. This data includes over 400,000 visits in the 12-month period, with a unique identity for each visitor, the center they patronized and the nature of services received. In addition, I have land parcel level list from the land records database for 18 of the 34 districts in the province.

Primary Data on farmers: I conduct a phone survey with a subsample of over a million farmers from Punjab province comprising farm households enrolled in various government programs for which they submitted contact numbers. In mid-2020, I called roughly 1800 randomly selected farmers and solicited information on access to titles and land record service centers in their localities. This data identifies

[17] Correlates of farmer TFP are discussed in Appendix Section C.

landowners, whether they have a title for their property and when they obtained the title. I construct a retrospective panel of households' title ownership for households who were untitled before the program to test if the timing of when they obtain the title is correlated with the timing of the program in their district. I construct a dummy 'Titled' at the household year level for 10 years spanning the program rollout in Punjab. 'Titled' takes on value 0 for untitled households until the year they obtain a title, after which it takes on value 1.

3.2 Theoretical Predictions

Appendix B illustrates a simple framework to provide the intuition for the effects of improving tenure security on land allocation, based on Restuccia and Santaeulalia-Llopis (2017). Market frictions are incorporated as transaction costs in land leasing, and the framework predicts that under high transaction costs, low productivity farmers operate larger than optimal land as they are restricted from renting out, while productive farmers operate smaller than optimal land and have high marginal product of land. Thus, transaction costs manifest as *wedges* in the marginal products across farmers. If transaction costs were infinite, everyone would cultivate their land endowment.

The reform is expected to lower transaction costs and improve market participation of agricultural households. In particular, high TFP (high MPL) households can have access to more land through the market, while low productivity farmers rent out land and reduce participation in agriculture. As a result of land mobility, farm scale and input choices may respond. Increasing market activity moves the allocation closer to the efficient allocation and lowers the *wedges* in the marginal products, which implies that aggregate level dispersion in marginal products goes down. Improved allocation of land and improved selection in farming both imply that aggregate production is higher.

Farm-level output and yield may be affected through two channels. On the one hand, the selection effect means more productive (high TFP) farmers are farming while the least productive exit, implying higher farm output and yield for those who stay in farming. On the other hand, if there is an inverse farmsize-yield relationship, land reallocation implies that farmers cultivate larger amount of land on average and may experience lower farm yield. Thus, the effect on average farm level yield is ambiguous, and may also differ across farmers.

3.3 Empirical Strategy

I exploit the staggered rollout of the program by using a difference-in-difference strategy to compare trends in districts that received the program earlier relative to those that received it later. The program proposed one service center for each 'tehsil' or subdistrict, but due to the nature of the data used, the subsequent analysis is at the district-year level. I use the fraction of subdistricts in a district d that have a functioning service center by year t to obtain program intensity at district level and run the following household level specification:

$$y_{idt} = \beta_0 + \beta_1 ProgramIntensity_{dt} + X'_{idt}\Psi + \mu_d + \eta_t + \mu_d \times t + \epsilon_{idt}, \qquad (1)$$

where $ProgramIntensity_{dt}$ is the percentage of subdistricts in a district with an active service center, y_{idt} is an outcome for household i in district d and year t. X_{idt} are household demographic controls, and μ_d and η_t are district and year fixed effects, respectively. Household level controls include household head age, age-squared, education and gender. To control for district specific trends, I include an interaction of district fixed effects with a linear yearly trend. Standard errors are clustered at the district level. To account for the number of clusters, I also present wild-bootstrapped p-values (Cameron and Miller, 2015). $ProgramIntensity_{dt} = 1$ indicates that all subdistricts in district d have the program. The coefficient β_1 thus estimates an average treatment effect and represents the change in the outcome (beyond the district-specific trend and aggregate year fixed effects) due to an increase in program intensity.

Identification is achieved from the variation in timing of program start as well as variation in the degree of program completion once it starts in any district. When subdistrict level data is available I run a standard timing difference-in-difference, where the primary independent variable is an indicator for the program at the subdistrict level.

Identifying Assumptions: The identification assumes that the timing of program start is quasi random; in particular, it is uncorrelated with district specific trends after accounting for district and year fixed effects. I test for the validity of this identifying assumption with two different balance tests. First, I test for balance in pre-program characteristics across the various timing groups. Specifically, I use data from the pre-2011 survey round to regress the district level outcomes of interest on fixed effects for each start year group. Standard errors are clustered at the district level. Columns (1)-(2) of Table A2 Panel A show the F-statistics from a test of the joint significance of the start year group fixed effects and the corresponding p-values. These regressions test if the timing of program start is correlated with the prior *levels* of the main outcomes and macroeconomic variables. Second, I test if the timing of program start is correlated with prior *changes* in these outcomes. To test the relationship between program timing and prior trends, I regress the district level change in the main outcomes of interest on indicators for the year of program start. Columns (3)-(4) of Table A2 Panel A show the F-statistics from a test of the joint significance of the start year group fixed effects and the corresponding p-values. In addition to the district level outcomes, two remote-sensing variables (Soil Quality Index and NDVI) are available at the subdistrict level. I thus conduct the balance test for these outcomes with respect to program start at the subdistrict level. For all different outcomes, we can reject that the program start year fixed effects are jointly significant. Thus, the timing of treatment across districts does not appear to be driven by the level or changes in the main outcomes of interest, key macroeconomic variables or underlying soil quality and productivity.

In addition to the variation in timing, the identification strategy uses variation in the number of centers opened as a share of total subdistricts in a district. The identification thus assumes that the opening of centers at the *subdistrict level* is quasi-random. In the absence of subdistrict level data, I test this assumption by conducting balance tests for program intensity similar to above. I regress $Program\_Intensity$ on the prior year levels and changes of district level outcomes, district and year fixed effects.[18] The results of these tests are presented in Panel B of Table A2. The balance tests confirm the validity of this assumption, as changes in $Program\_Intensity$ are not driven by the prior levels or changes in main outcomes of interest as the coefficients on various district level outcomes are small and statistically insignificant.[19] In similar spirit, I construct a placebo treatment variable by assuming program rollout prior to the actual launch of the program. If the differences in outcomes that are correlated with program intensity were pre-existing, we would see significant correlation between the outcomes and the placebo program intensity variable. I later confirm that the program effects are unlikely to be spurious as the placebo program variable has no significant effects.

I further account for factors that may compound the treatment effect by controlling for district-specific linear trends in all the regressions. There may be some concerns about district specific macro-economic cycles, as the study time-period represents a period of recovery from the 2008 global recession. In additional robustness checks I control for macro-economic variables at the district level, quadratic district-level trends, and account for district specific economic recovery by allowing for pre- and post-recession district-specific linear trends. I also conduct a placebo test with urban households to alleviate concerns that the program effects are capturing macroeconomic trends across districts. These robustness checks are discussed in additional details in Section 5 and the findings from the preferred specification are robust to

[18] I use the 2011 and 2013 survey rounds for these tests, since the variation in $Program\_Intensity$ is primarily during those rounds.

[19] For instance 6 percentage point higher level of land renal (the effect size from the program) is associated with a 0.001 percentage point higher program intensity in the following period. Thus the difference in land rental rate across districts has almost no relation with the program's rollout.

these checks.

Additionally, I estimate an event study specification to explicitly test for pre-existing differential trends and identify a post-program shift in trend. Last, I employ two additional identification strategies to compliment the findings from the main empirical analysis. First, I present the findings from the standard timing difference in difference strategy, replacing $ProgramIntensity_{dt}$ in (1) with a dummy $PostProgram_{dt}$ that switches from 0 to 1 when the program starts in district d.[20] Second, I use an alternate stacked difference-in-difference identification strategy following Deshpande and Li (2019). The main findings are confirmed by the results of the complementary identification strategies. To account for multiple hypotheses being tested, I adjust my p-values following Anderson (2008) and Benjamini and Hochberg (1995). The details of the additional empirical strategies and tests are presented in Section 5.

For land sales and rental market participation, I limit the regression sample to landowning households, while for farm input and output (farm size, crop and input choices, and yield), the sample includes cultivating households (including both landowners and landless cultivators). An additional concern arises if these samples are shifting over time, particularly in response to the program. I test the program effect on an indicator for inclusion in the rural, 'landowners' and 'cultivators' sample in Appendix Table A3; these tests provide assurance that the likelihood of being a rural, or a landowner or farming household among rural households does not respond to the program.

4 Results

4.1 Take-up of the Program

The argued mechanisms for the program effects are improved access to land records and tenure security as a result of land record digitization. In this section, I support these proposed 'first-stage' effects with both qualitative and empirical evidence on access and usage of the land record service centers.

First, landowner reported data prior to the program and qualitative reports from *patwaris* indicate that access to titles and property rights security was the dominant

[20] Districts belong to 4 possible timing groups depending on when the program starts in any district (2011, 2012, 2013 or 2014), and the timing difference-in-difference (DD) estimate is a weighted average of 2 types of difference-in-difference estimators (Goodman-Bacon, 2019). The first type compares the change in outcomes for early districts (treated) to late districts (control) before and after the start of the program. The second type compares the change in outcomes for late districts to early districts before and after the treatment starts for the late districts (while the early districts have already been treated for some time). Appendix Figure A3 illustrates two possible DD estimators when there are two timing groups. However, Goodman-Bacon (2019) points out concerns with the standard timing difference-in-difference if the treatment effects vary over time. In this context, the treatment effect may increase as the program expands to the other subdistricts after it is initiated at the district level. Thus, the preferred specification uses the intensity variable that accounts for subsequent openings after program start in a district.

constraint resolved by the program. After the roll-out of the reform, 72% of visitors to land record service centers perceive the new system to be more secure, and 60% also report service centers are expected to reduce land disputes.[21] Qualitative interviews of land revenue officers including *patwaris* also testament to the increased access to land titles and rights verification due to the program. These observations suggest that that the reform improved tenure security by obviating manual manipulation of records, diminishing the role of *patwaris* and other revenue officers, and facilitating improved access and verifiability of ownership and cultivation rights.

Second, I observe substantial usage of the service centers shortly after they are operational. The service centers received over 400,000 unique visitors in 2016, with an average ratio of visitors to land parcels of 18% over the 12 month period. I note that the same household or landowner can own multiple parcels so the ratio of visits to landowners is expected to be higher. The type of services obtained at the service centers show that 70% of all visits are for the purpose of obtaining a title, and another 10% are for confirmation of land rights.

Last, I document that service centers led to an increase in title ownership. In the phone-survey sample, 75% percent of the landowners have some title or *'fard'* for their property in 2020—56% have a computerized title from the land record center, while the rest have the old or 'manual' version of the title. Based on farmer responses, I infer that before computerized records were available 46% of farmers did not a have a title for their land. Of the untitled landowners, about 45% had obtained a title by the time of the survey, a substantial increase in title ownership over the 4-8 year period for which the program had been operating across different regions.

For the causal effect of the land records program on access to land titles, I regress *'Titled'* on *'Post'* or an indicator that switches to 1 when the program starts in the subdistrict of the landowning households' agricultural property, household fixed effects, and year fixed effects interacted with household's landholding. Consistent with the other analyses, I cluster standard errors at the district level and wild-bootstrap the standard errors. This clustering also accounts for the serial correlation in the outcome within households. The findings from this regression, shown in Appendix Table A4 show that the opening of the service center is associated with a 2 percentage point increase in acquiring a title. This amounts to 16% of the title ownership rate among the sample households by the end of the regression period. This effect may be an underestimate as we expect title ownership to continue to increase after the rollout of the program, whereas the difference in difference specification estimates the increase in title access in the year the program dummy switched from 0 to 1.

Self-reported 'take-up' is also high. In the entire phone survey sample (including some farmers who do not own land), 50% have used the land record service center

[21] Based on an independent report of the land records service centers usage.

and 75% of these respondents patronized the service centers for obtaining a title or record confirmation.[22] 'Record confirmation' is typically reported by cultivators who are not land owners, which suggests that property rights are more accessible and transparent for tenants as well. These data together provide an underpinning for the expected effects on tenure security and land market activity.

4.2 Program Effects on Land and Labor Market Participation by Landowners

Lack of ownership security restricts landowners from trading their land, i.e. renting out or selling it. Only 22 percent of landowning households report renting out their land, while only about 1 percent report having sold or purchased a portion of their agricultural land holding in the prior year. Among cultivating households, 15 percent are landless, and a third report renting in any land for cultivation.

To examine the land market effects of the program, I first test if land ownership shifts as the program is rolled out. The first outcome in Table 2 is an indicator for land ownership among all rural households. The program has no effect on the rate of land ownership, which could be consistent with no market activity or market transactions that caused land ownership to shift across households without changing the overall rate of ownership. To investigate this, I consider the change in recent land transactions by landowning households as the program is rolled out. I find that the rate of land purchase (Column 2) or sale (Column 3) do not respond to the program, suggesting that the program does not change the constraints on land ownership transactions.[23]

Improving tenure security can increase the likelihood of tenancy transactions even if land ownership stays stable. Column (4) in Table 2 shows that among landowning households, the likelihood of renting out increases by 6 percentage points when the program is completed in their district. This is a large effect, given the 22% rate of renting out on average across the districts prior to the program. Landowners renting out could be those who previously owned land or households that are able to purchase more land due to the program and then rent it out. Since there are no significant effects on the agricultural land ownership rate, ownership transactions, or the average size of land-owned, we can deduce that the change in tenancy is driven by previous landowners. Thus, the program resolved land market frictions that constrained existing landowners from renting out.

Relieving constraints on renting out for existing landowners can have spillover effects on the labor market. Specifically, agricultural participation is allegedly high due to insecure property rights on agricultural land that prevent households from participating in off-farm activities for better income, as vacating land bears the risk

[22]This usage rate measures any visit since the service center has been open (a 4-8 year period) while the administrative visit data above measures visits within one calendar year.

[23]There is also no significant effect on the size of owned holdings.

of losing it (Field, 2007). Increased rental activity by landowning households implies some landowning household members no longer need to practice cultivation if they have opportunities for participating in non-farm activities. The next set of results in Table 3 examine the effect of the program on participation in agricultural activities by landowning households. Consistent with high likelihood of renting out, I find that on average, these households are less likely to participate in agriculture as a result of the program. Three different outcome variables indicate this. Households are less likely to cultivate a farm (column 1), less likely to 'self-cultivate' or cultivate owned land (column 2) and less likely to to have members that participate in agriculture broadly, including wage work (column 3). In sum, landowners are, on average, 27% more likely to rent out their agricultural land and 12% more likely to quit agriculture due to the program. Column 4 of Table 3 shows the intensive margin measured by the share of households' income from agricultural activities. Consistent with the changes in the occupational choices of landowning households, the proportion of income from agricultural activities falls by 7 percentage points, when comparing wholly treated districts to wholly untreated ones. This corresponds to a 12% drop in income share of landowning households from agriculture.

I also test the changes in the alternate occupational choices of landowning households as they are able to rent out their agricultural land and exit agriculture. Table A5 in the Appendix shows an increase in the share of household members that participate in non-agricultural activities. These are statistically significant at the 10% level for participation in large business ownership, and positive (but statistically insignificant) for participation in small businesses, self-employment or as paid employees.[24]

In summary, the results above demonstrate that weak property rights constrain landowners from leaving agriculture, forcing them to cultivate their owned land instead of renting out land and engaging in other economic activities. Improved ownership security through the computerization of ownership and tenancy rights reduced market frictions, allowing landowners to rent out their land and increase participation in non-agricultural activities. I test the heterogenous effects of the program by income quartile in the Appendix Tables A7–A8. Heterogeneous effects demonstrate that land market frictions are particularly extreme for poorer, and plausibly less-connected, households. The program improves land rental probability and reduces agricultural participation for households in the lowest income quartile; the richest households experience significantly lower impact on both land rental and labor participation, relative to the poorest households. These effects are consistent with the motivation behind

[24]A natural outcome to test would be the rate of migration. The data does not allow us to test this explicitly, but the demonstrated effects suggest migration may have increased for landowning households. We can test if households are more likely to receive remittances or participate in the credit market. These outcomes are shown in Table A6, which demonstrates no significant effects on the likelihood of a loan or the likelihood of receiving remittance income. The table also shows that total income for landowning households improves by 22% due to improved land and labor markets with the roll-out of the program.

the design of this computerization program, which intended to increase accessibility of records for the marginalized sections of the rural population.

4.3 Program Effects on Farm Operation by Cultivators

The next set of regressions estimate the program effects for cultivating households, which includes landowning households that stay in cultivation as well as landless farm households. Table 4 shows the effect of the program on the intensive margin of renting in, measured by average quantity of rented in land among cultivating households. The program has a strong positive effect on land rented in on fixed cash rent, and no significant effect on land that is sharecropped. This is consistent with the view that land owners with less secure property rights may choose sharecropping, as it allows landlords to exert stricter property control by bearing a higher amount of production risk than in fixed rent contracts (Bellemare, 2012). Sharecropping is also typically arranged between landlords and tenants in the same village due to the sharing nature of this tenancy arrangement and for ease of monitoring; thus, the threat of weak property rights might be less binding for sharecropping.[25]

The program proves to relieve the constraints in the fixed rent lease market for agricultural land. Column (3) shows that owned cultivated area has no statistically significant change, which is consistent with the absence of any effects on land sales. Finally, Column (4) suggests that as more land is rented in, average farm size increases, indicating meaningful impacts of the program on scale of agriculture in Punjab. Average operational farm size is about 1 acre or 15% higher just after the program's completion in a district. In the Appendix, I explore heterogeneity across households in these outcomes; Table A11 presents these effects and shows that among cultivating households, landless households benefit from greater access to land due to improved rental markets (marginally more than landed farm households).

4.4 Program Effect on Land Allocation across Farmers

To test the program effect on the allocation of land across farmers, I re-run specification (1) interacting program intensity with TFP quartiles:

$$
\begin{aligned}
y_{idt} = &\phi_0 + \phi_{1,j} ProgramIntensity_{dt} \times TFPQuartile_{ij} + \phi_{2,j} TFPQuartile_{ij} \\
&+ X'_{idt}\Psi + \mu_d + \eta_t + \mu_d \times t + \epsilon_{idt}
\end{aligned} \tag{2}
$$

y measures total cultivated land, or MPL for any farmer. $TFPQuartile_{ij}$ is an

[25]Table A9 shows that the extensive margin of renting in among cultivating households does not change due to the program. As the composition of the sample of cultivating households may change due to the program (as landowners quit cultivation), I check for robustness by controlling for an indicator for land ownership by the household (Table A10).

indicator for a farmer *i* being in any quartile *j* of the TFP distribution in a district-year cell.[26]

The theoretical framework predicts that reducing transaction costs in the leasing market will induce lower TFP farmers to rent out land and exit agriculture, while higher TFP farmers have greater access to land and lower marginal product of land. Thus, in equation 2 we expect ϕ_1 to be positive for land and negative for MPL for farmers in higher TFP quartiles.[27]

Table 5 shows the effect of the program on land allocation across farmers at different parts of the productivity distribution. The table shows the log of total cultivated area (or farm size in acres) and the marginal products of land as a function of each farmers' TFP. The findings in Table 5 indicate that farmers in the higher TFP quartiles are more likely to have greater access to land, as indicated by high operational farm size and lower marginal product of land. In the highest TFP quartile, farmzise increases by 24% with the program.

A more efficient allocation implies that at a market level, the dispersion of marginal products would go down.[28] In order to test this predictions, I construct measures of MPL dispersion by calculating the standard deviation, coefficient of variation and interquartile range within a district in each survey year. The sample size restrictions do not allow the construction of these measures at a village level (rental market operation may be more likely to occur at the village level, but the sample size within the same village is too small). In Table 6, I find a statistically significant negative effect on MPL dispersion measured in three different ways. Additional effect on the dispersion of TFPs is also negative with moderate level of significance, and shown in Table A12. These outcomes are consistent with Table 5, showing a reallocation of land across farmers, and also corroborate the earlier evidence on reallocation of some households away from farming.

Since the construction of TFP requires a number of assumptions, I examine land allocation in response to the program using two alternate measures of productivity: agricultural yield per acre and profits per acre. These results are in Appendix Table A13 and show that land is higher for all households that stay in farming after the program, but significantly higher for farmers in the 3rd and 4th quartiles of the productivity distribution, whereas MPL is lower for them. These results are consistent with the program effects estimated across farmers ranked by TFP.

Together, the findings provide credible support for the hypothesis that an im-

[26]Allocative efficiency arises from moving land to high MPL farmers. In the pre-reform data TFP and MPL are positively correlated, as may be expected in markets with friction in land leasing or transfers.

[27]The lowest TFP farmers may choose to rent out all their land and exit agriculture entirely. Since the cultivation data is used to calculate farmer TFP, households that do not participate in cultivation are excluded from the above regressions and I cannot directly test if lowest TFP households exit cultivation.

[28]It can be noted that MPL is directly proportional to farm yield—MPL dispersion would simply capture the dispersion in yield.

proved allocation of land across farmers is underlying the average effects on rental market and farm size.

4.5 Program Effect on Farm-level and Aggregate Output and Yield

The effect of the program on farming scale has important implications for agricultural input choices, mechanization, and productivity improvement, as farm size is assumed to be a constraint to adoption of capital intensive technologies (Foster and Rosenzweig, 2011). Moreover, insecurity of tenure affects investment incentives, as Jacoby and Mansuri (2008) demonstrate that non-contractible investments are under-provided on leased land in Pakistan due to incomplete contracts. If optimal farm area and/or increased tenure security induces higher input, especially capital, usage on farms, then output should increase. Even if the capital margin is unaffected, improved allocative efficiency will result in higher aggregate productivity. I examine the effects of the program on farm output in Table 7 and aggregate output in Tables 8-9. The farm output, yield, and profits are expected to be measured with error, and winsorized values are used to test for program effects.

Table 7 shows positive effects on the total farm output, output per acre, and profits or value added, though the effect on output and profit per acre is not statistically significant.[29] The increase in output is consistent with an increase in farm scale, while the null effect on farm-level yield and profit is consistent with the combined effect of a negative farmsize productivity relationship (Foster and Rosenzweig, 2017) and a selection of more productive farmers into farming.

My measure of aggregate production, NDVI, is obtained from remote sensing data on vegetation, as explained in Section 3.1. Since the data is now at the subdistrict level, I run a subdistrict-year level regression, with an indicator $PostProgram$ as the dependent variable of interest, which is 1 for each year after the program has started in a subdistrict. I include subdistrict and year fixed effects and linear trends at the subdistrict and district level. The results, presented in Table 8, show the effect of the program on the NDVI in standard deviations, and show a robust positive effect on productivity as measured by the NDVI. Introducing the program increases the level of production at the subdistrict level by 9 percent of a standard deviation.

Table 9 employs alternate, aggregate district-level crop production data from administrative sources. The regressions are at district-crop-year level, where the outcome of interest is log of crop yield (ton/ha) for each district by year and for the major cereal and cash crops[30]. The district-level yield regressions include crop, district and year fixed effect as well as district specific linear trends, and span all years

[29] Value-added or profits are calculated as the difference in the value of output per acre and the total expenses per acre.

[30] Cereals: rice, wheat, and maize. Cash crops: cotton and sugarcane

from 2005 to 2015. I find that while the program does not affect total cultivated area under the major cereal or cash crops, aggregate cereal output and yield are 6% and 5% higher, respectively, when the program is completed in any district. These effects are significant at the 10% level. The major cash crops show an average 2% decline in yield, but this effect is statistically indistinguishable from zero. This district-level data provides supporting evidence of improved aggregate productivity, but suffers from caveats that are typical for government-collected administrative data.

Appendix Tables A14-A15 show the changes in inputs and crop choice underlying the farmsize effects. Farmers shift into rice and away from planting maize on their land, as shown in Table A14. Such a shift may or may not be expected, and there are not many consistent mechanisms that explain this shift. Secondly, the effect on input choices in Table A15 shows an increase in the usage of pesticides (marginally statistically significant). The usage of rented equipment is lower, though not statistically significantly, which could suggest an increased likelihood that farmers' use owned machinery and equipment. Since data on farmer ownership of agricultural machinery is not available, I cannot explicitly test this hypothesis. Farmers do report if they acquired (purchased or received) any agricultural machinery, including tube wells, tractors, ploughs, threshers, harvesters or trucks, in the previous year. I test the effect of the program on acquisition of agricultural machinery and find that while there is no overall effect, landless households are particularly more likely to have acquired agricultural equipment when the program is completed in their districts (Appendix Table A16). Altogether, greater farm size, higher usage of some inputs, increased owned equipment, and reduced equipment rental are consistent with increased incentives to invest in productive mechanized inputs.

4.6 Discussion of Findings and Mechanisms

Evidence on take-up shows that the program's effects are largely driven by an improvement of perceived property rights security through access to land records and titles. The survey data used for the analysis does not measure perceived property rights directly but asks landowners how much they expect to receive if they were to sell their land. I create a proxy for ability to sell one's land using an indicator that equals one if the respondent reports some positive expected payment from the sale of their property.[31] Using specification 1, I find a significant improvement in the proxy for perceived selling rights with the program. This is despite no change in land purchase/sale as shown earlier, or in the total area owned. The program has no impact on perceived rights for property types that were unaffected by the land records program, as measured by a similarly constructed proxy for nonagricultural land. This

[31] The proxy assumes that farmers who do not have rights to sell their property report no expected payment or an expected payment of zero.

suggestive evidence (presented in Table A17) in combination with other types of evidence provide a strong case for improved access to land rights and property rights security due to the program.

The lack of effects on land ownership and sales can be due to a number of reasons. First, land sale and rental markets may be substitutes. Second, landowners may rent out their land, but not sell it, as lack of complete insurance and credit markets induce them to hold onto land as a precautionary asset (Rosenzweig, 2001). Finally, the program effects are strongest for the lowest income quartiles among landowning households. These households may be less likely to participate in the sales market regardless of the program. If only large landowners participate in the land sales market, the sales margin is less likely to respond as large landowners are less likely to face tenure insecurity before the program and are therefore least affected by it.

The results, put together, demonstrate that a light-touch reform has promising prospects for the outlook of agricultural land markets and structural change in Punjab. Aside from facilitating fixed cash rent transactions and increasing farming scale, the program can induce a productive shift in the allocation of labor, paving the way for improved agricultural productivity and transition of rural labor into non-agricultural sectors, which are necessary predecessors of transition in the economy.

5 Robustness and Alternate Identification Strategies

I conduct a number of additional tests to validate the identification strategy and robustness of the findings.

5.1 Additional Time Periods

I test the primary outcomes using an extended dataset that combines the HIES with the Pakistan Social and Living Standards Measurement (PSLM) surveys. The outcomes that are measured in both surveys include land ownership, rental, and household occupation and are shown in Table A18. I find that the likelihood of renting out and leaving agriculture are higher for households in the districts with the program and no change in land ownership, confirming the primary results. Further, in Appendix Table A19 (Panel A) I ensure that the choice of controls do not drive the main results by showing the effects from the main specification without household controls.

5.2 Alternate macro-economic trends

I address the concern that the study period coincides with a period of global recession followed by a recovery and that some of the program effects may be driven by differential rates of recovery across districts. I rule out differential business cycle events across districts in four ways. I allow for non-linear district specific trend by controlling for quadratic trends by districts. I control directly for macroeconomic outcomes at the district level in the main specification. Particularly, I add district level unemployment and size of the labor force to the regression. In an additional robustness check, I allow for pre- and post-recovery trends for districts. Figure A4 shows the GDP per capita for Pakistan, which stagnates during global recession in 2007-09 and recovers in the post-2010 period. I interact the district specific trend with an indicator for the post recession period to allow for varying rates of recovery across districts. These results are presented in panels B-D of Table A19 and demonstrate that the treatment effects of the program are not sensitive to these additional controls. The coefficient estimates are qualitatively and quantitively unmoved when I account for the possibly confounding macroeconomic changes. Lastly, I conduct a placebo test using the income of urban households that is unlikely to be affected by the land records program and find a precisely estimated null effect (Table A20). This provides further reassurance that the program rollout is not capturing a differential recovery from the global recession across districts.

Similarly, the earliest program districts may have been selected endogenously as the pilot districts for a salient program and may have differential trends. I ensure that the treatment effects are robust to excluding the three districts where the program began in 2011 (Table A21).

5.3 Event Study Analysis

I test for pre-existing trends using an event study analysis using the expanded data set with the PSLM and HIES surveys that has data from consecutive years with the exception of 2009 when neither survey was conducted. Figure 4 shows the event study graphs, which plot the coefficients, γ_l, from the following district level regression.

$$y_{dt} = \gamma_0 + \sum \gamma_l YearsSinceProgram_{dt,l} + \mu_d + \eta_t + \varepsilon_{dt} \tag{3}$$

$YearsSinceProgram_{dt,l}$ is an indicator that equals one if it has been l years since the start of the program in district d and year t; the omitted category is $l = -1$, or the year just before the program starts in any district. Due to the limits on the time periods covered by the survey data and the staggered timing of program start, the lags and leads relative to start date represented in the survey data can vary for the program start timing groups. For instance, for a district in the 2012 timing group, the survey data represents the following lags and leads with respect to program start: -7, -6, -5, -4, -2, -1, 0, 1, 2, 3. Similarly, for the 2013 timing group, the lags range from -8 to 2 (-4 is missing). Thus, in specification (3), each *YearsSinceProgram* dummy coefficient would be driven by a different set of districts. To keep the sample of districts mostly stable, I show 6 lags and 2 leads. To account for the missing year, I follow McCrary (2007) and interpolate linearly between observation years correcting the standard errors for the induced serial correlation. With the two surveys combined, the data ranges from 2005-2015 with 9 time fixed effects ranging from -6 to +2 for a balanced sample of districts.[32] The graphs show that the program start is not driven by changes in land market activity, as the trend is flat in the pre-program period. In the post-program period land rental increases while agricultural participation declines. Land owned shows a flat pre- and post-program trend as reflected in the regression analysis earlier.

5.4 Placebo Program Rollout

As an alternate test to rule out that pre-existing trends in the main outcomes drive the rollout of the program, I construct a placebo variable to measure 'program intensity', assuming the program rollout began two survey years prior to the actual program date in each district.[33] Table A22 provides the outcomes from specification (1) replacing intensity with the placebo treatment and shows no effect on the main outcomes.

5.5 Standard timing Difference-in-difference estimation

I use two alternate identification strategies to measure the program effects, a standard timing difference in difference and a stacked difference in difference. First, I use a dummy variable, $PostProgram$, indicating the years after one of the subdistricts in a district has received the program in Table A23. This treatment classification would avoid any concerns about endogenous pace of program delivery after the first opening in any district. The $PostProgram$ indicator specifically captures the average effect of one subdistrict receiving the program (while the $ProgramIntensity$ coefficient in the primary specification captures the effect of all subdistricts receiving the program), thus the effects may be smaller. Table A23 shows a significant increase in land rental by landowners and a significant drop in agricultural participation, while cultivating households have significantly more area rented in for cultivation and higher farm output. The magnitude of the effects are smaller, and the effect on cultivated area is positive but loses statistical significance.

Goodman-Bacon (2019) cautions about an important feature of timing difference-in-difference (DD) strategies with early and late timing groups. In particular, the comparisons of late timing groups to early timing groups rely on a comparison of just treated groups to already treated groups. If the early treated units are set on a differential trend by the treatment, they are no longer good 'control' units. Goodman-Bacon (2019) proposes a decomposition to calculate the weight on each DD estimate.

[32] The districts that start the program in 2014 are not included in the estimation of the coefficient for $YearsSinceProgram = +2$.

[33] This robustness exercise yields similar results if I assume the placebo program started in the prior survey year.

In my context, a calculation of the weights shows that the DD estimates based on the comparisons of early to late units have the majority of the weight (70%) while the DD effects comparing newly treated units to already treated ones have a much lower weight (30%). Moreover, the average effects from the two types of comparisons are qualitatively similar for all the main outcomes. Thus, the timing DD estimates are a meaningful comparison for the effects from the main regressions.

5.6 Stacked Difference-in-difference estimation

Another strategy uses a stacked difference-in-difference to compare different timing groups to 'control' units that are treated in a future period but are untreated when they act as controls, as in Deshpande and Li (2019). In particular, I construct a data set as follows: for each center opening I label the district with the opening as treated and the districts that do not have any opening yet (but will receive the program 1-2 years after the treated district) as control. For each opening I also include outcomes for the treated and control districts from survey rounds in the year before and just after opening to capture the change in trend due to the opening. The time period just after the opening is indicated by a *Post* indicator. I construct the treated and control districts for each of the 150 openings and stack these datasets. A district is treated when the opening corresponds to a service center in one of its subdistricts, and a district could be treated for some openings and control for others. However, treated districts never switch to being controls because the set of control districts correspond to districts that have not yet been treated. The specification I run is as follows:

$$y_{idt,o} = \phi_0 + \phi_1 Treated_{d,o} \times Post_{t,o} + \phi_2 Post_{t,o} + \omega_{d,o} + X'_{idt}\Psi + \mu_d + \eta_t + \epsilon_{idt,o}, \quad (4)$$

where $y_{idt,o}$ is an outcome for household i in district d, year t for opening o. $Treated_{d,o}$ is 1 if d is the district where opening o occurred. $Post_{t,o}$ is 1 for the year after opening o occurs. District and year fixed effects are included, as well as the same household controls from the main specification. Additionally an opening by district fixed effect is included, which is collinear with $Treated_{d,o}$. The coefficient on $Treated_{d,o} \times Post_{t,o}$ captures the shift in outcome y just after the center opening in the treated district versus control districts.

Table A24 shows the outcomes from this empirical specification. In comparison to the main treatment effects that measure the effect of center openings in all subdistricts of a district, the coefficient in Table A24 captures the effect of an opening in one subdistrict. These effects are naturally smaller in magnitude, but are significant and consistently in line with the main treatment effects. The effects on farm area and output are less precisely estimated but have the expected magnitude. In summary, the two additional empirical strategies provide reassuring complementary evidence to the established treatment effects above.

5.7 Multiple hypotheses testing correction

I implement Anderson's (2008) methods to correct my standard errors for multiple hypothesis testing, controlling for the false discovery rate Benjamini and Hochberg (1995) following Banerjee et al. (2015) and Ksoll et al. (2016) (Table A25). The program effects on the primary outcomes, including land rental, agricultural participation by landowners, the total farm size, rental land size, and total output of cultivating households, survive this adjustment of p-values.

Together, the robustness checks validate the identification strategy and provide confirmation of the program effects documented in the primary empirical analysis.

6 Concluding Remarks

I focus on a 'light touch' property rights formalization program in a context with private ownership, without explicit titling or direct targeting of market transactions. Roth and McCarthy (2013) note a continuum of land rights formalization that extends from strengthening tenurial rights in law or formal titling and registration to better communicating those rights to land holders or strengthening informal land leasing arrangements and contracts. The Punjab land record computerization program resulted in a formalization of property rights through better clarity of and access to rights, and through automation of market transactions bypassing bureaucratic hurdles and corrupt officers. The formalization of property rights can have potentially large positive effects while obviating the financial and feasibility hurdles of titling programs.

The paper offers evidence that the program managed to significantly affect land markets, affecting *allocation* of land within agriculture and *selection* of cultivators into agriculture. Landowners who faced market constraints rent out land and exit agriculture after the program. On the flip side, households that stay in cultivation, rent in more land, effectively increasing average farm size, which has implications for modernization and aggregate agricultural productivity. Consistent with the increased rental activity and improved land allocation, aggregate yield improves in districts with the program, although the yield effects are not observed in farm-level data. I provide additional evidence that these changes in market activity are driven by improved security of tenure and verifiability of land rights.

The results thus illustrate that land and labor market constraints limit rural mobility in the South Asian context, shedding light on the rural-urban divide and the prospect of structural transformation. The paper further reinforces our understanding of development economics by exhibiting how ICT use is manifested in public service processes and can ease market frictions in lower income-countries. Effective use of ICT has been demonstrated for agricultural initiatives (Aker et al., 2016), delivering education and improving learning (Muralidharan et al., 2019; Beg et al., 2020), increasing service delivery staff accountability (Duflo et al., 2012), and reducing leakages in government welfare program payments (Banerjee et al., 2014; Muralidharan et al., 2016). The land record computerization program similarly improves access to property rights records through digitized and automated land record maintenance.

References

Adamopoulos, T., L. Brandt, J. Leight, and D. Restuccia (2017). Misallocation, selection and productivity: A quantitative analysis with panel data from China. *National Bureau of Economic Research working paper No. 23039*.

Adamopoulos, T. and D. Restuccia (2014). The size distribution of farms and international productivity differences. *American Economic Review 104*(6), 1667–97.

Aker, J. C., I. Ghosh, and J. Burrell (2016). The promise (and pitfalls) of ICT for agriculture initiatives. *Agricultural Economics 47*(S1), 35–48.

Ali, D. A., K. Deininger, and M. Goldstein (2014). Environmental and gender impacts of land tenure regularization in Africa: pilot evidence from Rwanda. *Journal of Development Economics 110*, 262–275.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Banerjee, A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Parienté, J. Shapiro, B. Thuysbaert, and C. Udry (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science 348*(6236), 1260799.

Banerjee, A., E. Duflo, C. Imbert, S. Mathew, and R. Pande (2014). Can e-governance reduce capture of public programs? Experimental evidence from a financial reform of India's employment guarantee. Technical report, Technical Report, MIT December 2014.

Banerjee, A. and S. Jain (2003). E-governance in India: Models that can be applied in other developing countries. In *International Conference on Electronic Government*, pp. 488–491. Springer.

Beg, S. (2019). Tenancy and clientelism. *Working Paper*.

Beg, S. A., A. M. Lucas, W. Halim, and U. Saif (2020). Engaging Teachers with Technology Increased Achievement, Bypassing Teachers Did Not. National Bureau of Economic Research Working Paper No.25704.

Bellemare, M. F. (2012). Insecure land rights and share tenancy: Evidence from Madagascar. *Land Economics 88*(1), 155–180.

Bellemare, M. F. (2013). The productivity impacts of formal and informal land rights: Evidence from madagascar. *Land Economics 89*(2), 272–290.

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological) 57*(1), 289–300.

Besley, T. (1995). Property rights and investment incentives: Theory and evidence from Ghana. *Journal of Political Economy*, 903–937.

Besley, T. and M. Ghatak (2010). Property rights and economic development. In *Handbook of Development Economics*, Volume 5, pp. 4525–4595. Elsevier.

Binswanger-Mkhize, H. P. (2013). The stunted structural transformation of the Indian economy. *Economic and Political weekly 48*(26-27), 5–13.

Bloom, N., L. Garicano, R. Sadun, and J. Van Reenen (2014). The distinct effects of information technology and communication technology on firm organization. *Management Science 60*(12), 2859–2885.

Bresnahan, T. F., E. Brynjolfsson, and L. M. Hitt (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics 117*(1), 339–376.

Cameron, A. C. and D. L. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317–372.

Carter, M. R. and P. Olinto (1996). *Decomposition and identification of the impact of land titling on agricultural productivity in Paraguay*.

Chari, A., E. M. Liu, S.-Y. Wang, and Y. Wang (2017). Property Rights, Land Misallocation and Agricultural Efficiency in China. National Bureau of Economic Research Working Paper No. w24099.

Chen, C. (2017). Untitled land, occupational choice, and agricultural productivity. *American Economic Journal: Macroeconomics 9*(4), 91–121.

Chen, C., D. Restuccia, and R. Santaeulàlia-Llopis (2017). The effects of land markets on resource allocation and agricultural productivity. Technical report, National Bureau of Economic Research.

de Janvry, A., K. Emerick, M. Gonzalez-Navarro, and E. Sadoulet (2015). Delinking land rights from land use: Certification and migration in Mexico. *American Economic Review 105*(10), 3125–49.

de Janvry, A., M. Fafchamps, and E. Sadoulet (1991). Peasant household behaviour with missing markets: some paradoxes explained. *The Economic Journal 101*(409), 1400–1417.

Deininger, K., D. A. Ali, and T. Alemu (2010). Land rental markets: transaction costs and tenure insecurity in rural Ethiopia. In *The Emergence of Land Markets in Africa*, pp. 70–86. Routledge.

Deininger, K., D. A. Ali, and T. Alemu (2011). Impacts of land certification on tenure security, investment, and land market participation: evidence from Ethiopia. *Land Economics 87*(2), 312–334.

Deininger, K. and A. Goyal (2012). Going digital: Credit effects of land registry computerization in India. *Journal of Development Economics 99*(2), 236–243.

Deshpande, M. and Y. Li (2019). Who is screened out? Application costs and the targeting of disability programs. *American Economic Journal: Economic Policy 11*(4), 213–48.

Didan, K. (2015). *MYD13Q1 MODIS/Aqua Vegetation Indices 16-Day L3 Global 250m SIN Grid V006.* Distributed by NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MYD13Q1.006. Accessed 2019-05-25.

Do, Q.-T. and L. Iyer (2008). Land titling and rural transition in Vietnam. *Economic Development and Cultural Change 56*(3), 531–579.

Duflo, E., R. Hanna, and S. P. Ryan (2012). Incentives work: Getting teachers to come to school. *American Economic Review 102*(4), 1241–78.

FAO/IIASA/ISRIC/ISSCAS/JRC (2012). Harmonized world soil database (version 1.2). *FAO, Rome, Italy and IIASA, Laxenburg, Austria*.

Feder, G. (1988). *Land policies and farm productivity in Thailand*. Johns Hopkins University Press.

Field, E. (2007). Entitled to Work: Urban Property Rights and Labor Supply in Peru. *The Quarterly Journal of Economics 122*(4), 1561–1602.

Field, E. and M. Torero (2006). Do property titles increase credit access among the urban poor? Evidence from a nationwide titling program. Technical report.

Foster, A. D. and M. R. Rosenzweig (2011). Are Indian farms too small? Mechanization, agency costs, and farm efficiency. *Unpublished Manuscript, Brown University and Yale University*.

Foster, A. D. and M. R. Rosenzweig (2017). Are there too many farms in the world? Labor-market transaction costs, machine capacities and optimal farm size. National Bureau of Economic Research Working Paper No. 23909.

Galiani, S. and E. Schargrodsky (2010). Property rights for the poor: Effects of land titling. *Journal of Public Economics 94*(9-10), 700–729.

Ghosh, A. and G. Banerjee (2006). A Study of E-governance in Rural India. *Delivering E-government*, 222.

Goldstein, M. and C. Udry (2008). The profits of power: Land rights and agricultural investment in Ghana. *Journal of Political Economy 116*(6), 981–1022.

Gollin, D., S. Parente, and R. Rogerson (2002). The role of agriculture in development. *American Economic Review 92*(2), 160–164.

Gonzalez, M. L. (2016). Land records go digital in Punjab, Pakistan. *End Poverty in South Asia (http://blogs.worldbank.org/endpovertyinsouthasia)*.

Goodman-Bacon, A. (2019). Difference-in-Differences with Variation in Treatment Timing. *National Bureau of Economic Research working paper No. 25018*.

Gottlieb, C. and J. Grobovšek (2019). Communal land and agricultural productivity. *Journal of Development Economics 138*, 135–152.

Hornbeck, R. (2010). Barbed wire: Property rights and agricultural development. *Quarterly Journal of Economics 125*(2).

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.

Hunter, W. W., J. S. Cotton, R. Burn, and W. S. Meyer. *Imperial Gazetteer of India.*

Jacoby, H. G. and G. Mansuri (2008). Land tenancy and non-contractible investment in rural Pakistan. *The Review of Economic Studies 75*(3), 763–788.

Ksoll, C., H. B. Lilleør, J. H. Lønborg, and O. D. Rasmussen (2016). Impact of Village Savings and Loan Associations: Evidence from a cluster randomized trial. *Journal of Development Economics 120*, 70–85.

López, R. and C. Romano (2000). Rural poverty in Honduras: Asset distribution and liquidity constraints. In *Rural Poverty in Latin America*, pp. 227–243. Springer.

Lunduka, R., S. T. Holden, and R. Øygard (2010). Land rental market participation and tenure security in Malawi. In *The Emergence of Land Markets in Africa*, pp. 125–143. Routledge.

Macours, K., A. de Janvry, and E. Sadoulet (2010). Insecurity of property rights and social matching in the tenancy market. *European Economic Review 54*(7), 880–899.

McCrary, J. (2007). The effect of court-ordered hiring quotas on the composition and quality of police. *American Economic Review 97*(1), 318–353.

Muralidharan, K., P. Niehaus, and S. Sukhtankar (2016). Building state capacity: Evidence from biometric smartcards in India. *American Economic Review 106*(10), 2895–2929.

Muralidharan, K., A. Singh, and A. J. Ganimian (2019). Disrupting education? Experimental evidence on technology-aided instruction in India. *American Economic Review 109*(4), 1426–60.

Newman, C., F. Tarp, and K. Van Den Broeck (2015). Property rights and productivity: The case of joint land titling in vietnam. *Land Economics 91*(1), 91–105.

Restuccia, D. and R. Santaeulalia-Llopis (2017). Land misallocation and productivity. National Bureau of Economic Research Working Paper No. 23128.

Rosenzweig, M. R. (2001). Savings behaviour in low-income countries. *Oxford Review of Economic Policy 17*(1), 40–54.

Roth, M. and N. McCarthy (2013). *Land and Tenure, Property Rights and Economic Growth in Rural Areas*. USAID Issue Brief.

Shenoy, A. (2017). Market failures and misallocation. *Journal of Development Economics 128*, 65–80.

Sural, A. (2013). Punjab may abolish patwari, kanungo systems. *Times of India*.

United Nations Human Settlements Programme (2012). Land and Property Rights in Pakistan. Islamabad, Pakistan.

Wang, S.-Y. (2012). Credit constraints, job mobility, and entrepreneurship: evidence from a property reform in China. *Review of Economics and Statistics 94*(2), 532–551.

World Bank (2013). *World Development Indicators 2013*. World Bank Publications.

7 Figures and Tables

Figure 1: *Program Rollout*

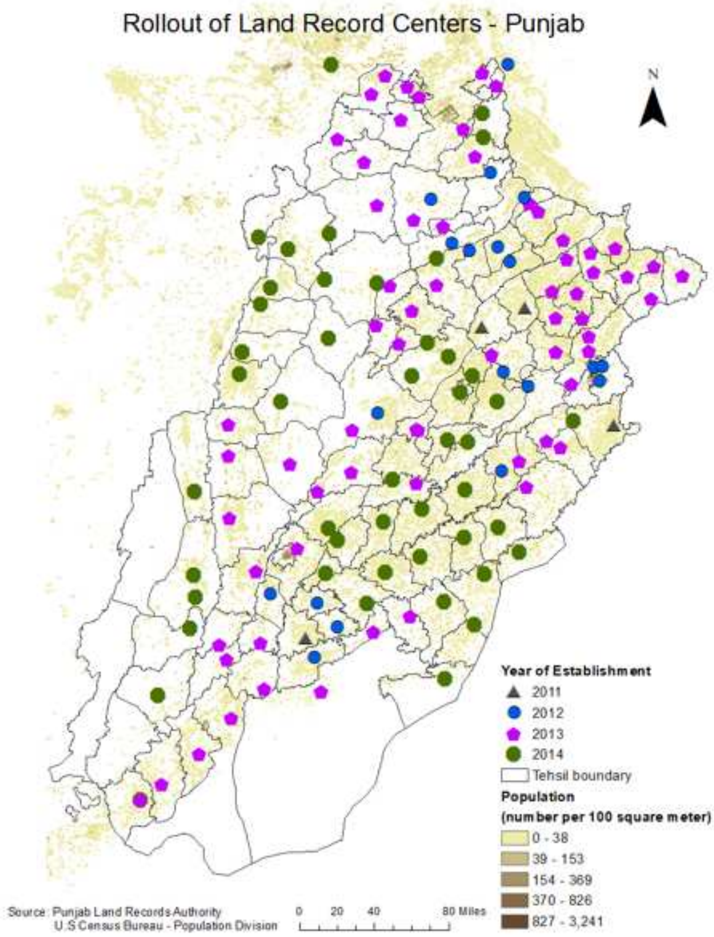


Figure 2: *Program Openings by Year*

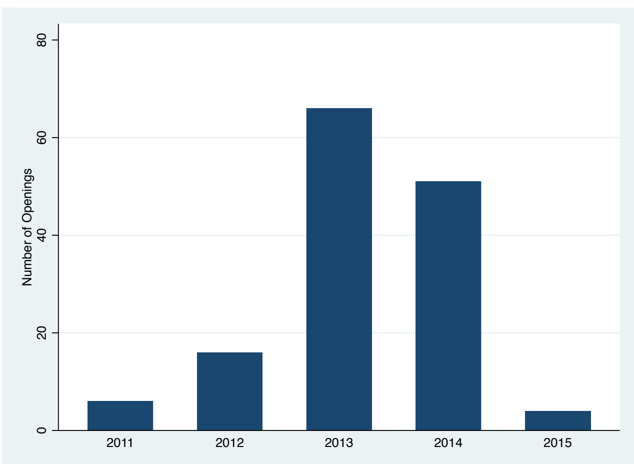


Figure 3: *Distribution of Farmer TFPs*

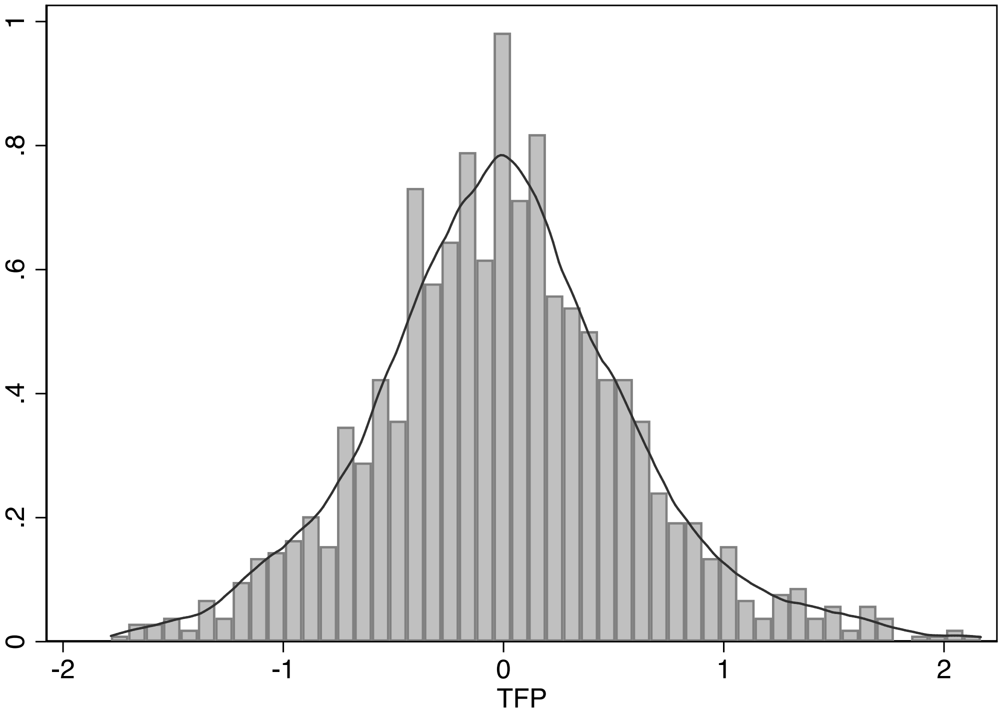
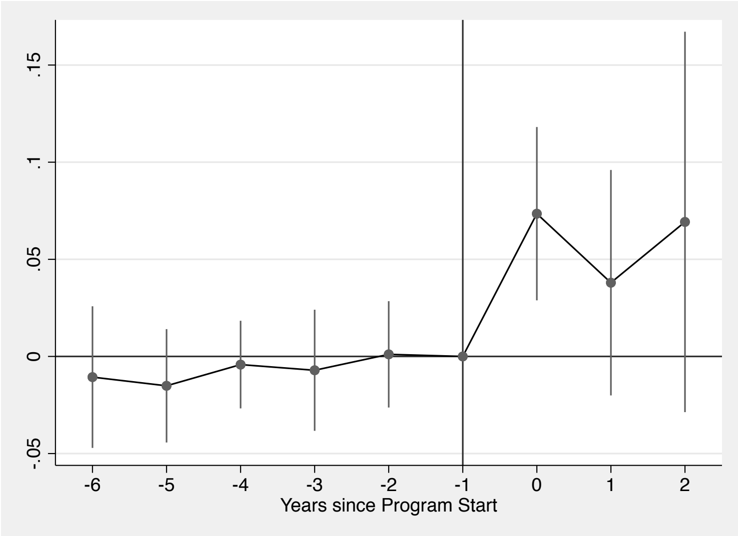
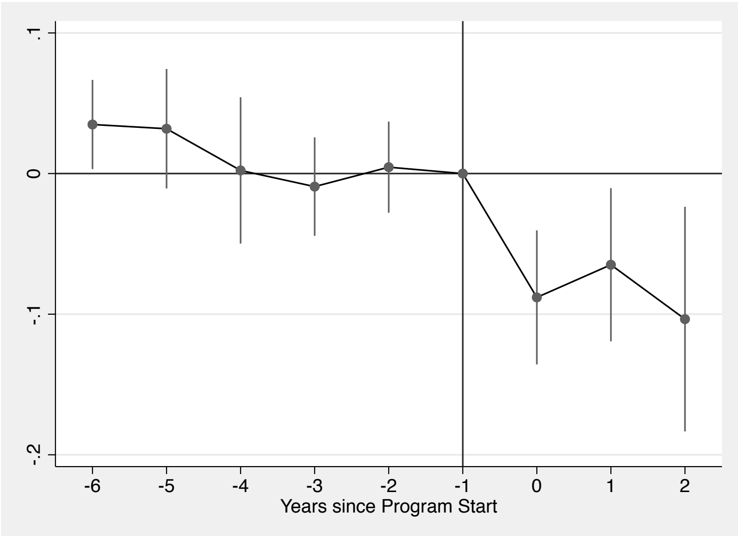


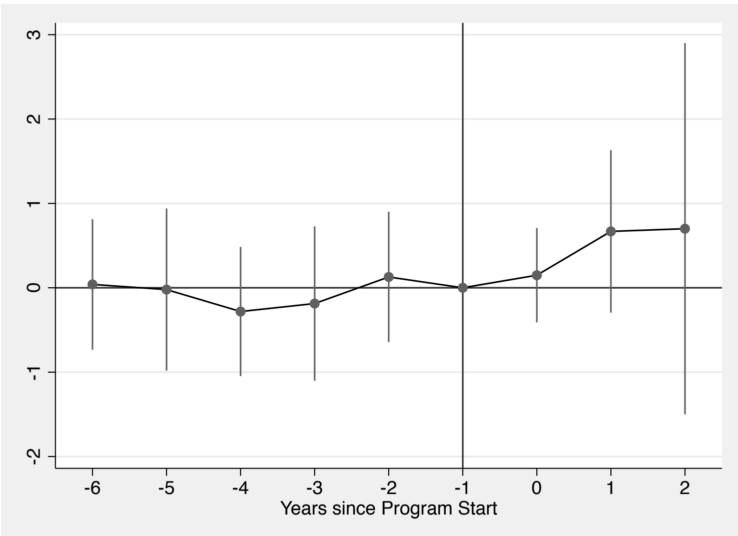
Figure 4: *Trend in Renting out, Agricultural participation and Agricultural Land ownership*



(a) *Renting out*



(b) *Agricultural Participation*



(c) *Agricultural Land ownership*

Table 1: *Summary Statistics for Households*

Variables	(1)
Panel A: Landowning Households (N=7,597)	
HH Rents out Ag Land	0.21
	(0.405)
HH Cultivates a Farm	0.81
	(0.395)
HH Cultivates Own Farm	0.62
	(0.484)
HH Member works in Ag	0.81
	(0.392)
HH Share Ag Income	0.65
	(0.391)
Panel B: Cultivating Households (N=7,256)	
Total Farm area cultivated (acres)	6.17
	(8.070)
Land rented in on cash rent (Y/N)	0.25
	(0.433)
Area rented in on cash rent (acres)	6.33
	(10.27)
Land rented in on sharecropping (Y/N)	0.07
	(0.253)
Area rented in on sharecropping (acres)	7.56
	(14.49)
Output (value) per acre	55.63
	(37.99)
Profit (output value - expenses) per acre	32.07
	(24.38)
Grows Wheat (Y/N)	0.91
	(0.281)
Grows Rice (Y/N)	0.31
	(0.461)
Grows Maize (Y/N)	0.07
	(0.257)
Grows Cotton (Y/N)	0.34
	(0.472)
Grows Sugarcane (Y/N)	0.17
	(0.373)

Notes: Data are from the HIES surveys. Landowning households report agricultural land ownership; cultivating households report farming agricultural land.

Table 2: *Program Effect on Market Activity for Land Owners*

	Own Agland (Y/N) (1)	Agland Purch. (Y/N) (2)	Agland Sold (Y/N) (3)	Agland Rentout (Y/N) (4)
Program Intensity	0.002	0.001	-0.002	0.061**
	(0.030)	(0.003)	(0.006)	(0.027)
	[0.954]	[0.821]	[0.707]	[0.0327]
Observations	19,067	7,584	7,584	7,597
Mean Dep., Pre-program	0.420	0.006	0.010	0.219
Sample Households	All Rural	All Landowning	All Landowning	All Landowning

Notes: All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3: *Program Effect on Agricultural Participation*

	HH Operates Any Farm	HH Operates Owned Land	HH Member Ag Worker	Share Income from Ag
	(1)	(2)	(3)	(4)
Program Intensity	-0.098***	-0.089***	-0.099***	-0.080**
	(0.030)	(0.032)	(0.035)	(0.034)
	[0.000900]	[0.00660]	[0.0114]	[0.0315]
Observations	7,597	7,597	7,597	7,597
Mean Dep., Pre-program	0.786	0.756	0.807	0.650

Notes: Sample includes all agricultural landowning households. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4: *Program Effect on Farm Size and Rented in Land*

	Rented	Sharecropped	Owned	Total Cultivated
	(1)	(2)	(3)	(4)
Program Intensity	0.925**	0.084	0.731	1.110**
	(0.433)	(0.255)	(0.697)	(0.452)
	[0.0351]	[0.797]	[0.320]	[0.0151]
Observations	7,256	7,256	7,256	7,256
Mean Dep., Pre-program	1.648	0.686	5.423	7.055

Notes: Sample includes all cultivating households. Rent area corresponds to area under fixed cash rent contracts and S/C refers to area under sharecropping contracts. Farm size is total operational farm area including owned land. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5: *Program Effect on Allocation across Farmers*

	Land (1)	MPL (2)
Program Intensity	0.050 (0.089) [0.579]	0.154 (0.099) [0.134]
TFP Quartile 2 x Program Intensity	0.081 (0.074) [0.273]	-0.081* (0.044) [0.0716]
TFP Quartile 3 x Program Intensity	0.096 (0.064) [0.139]	-0.130** (0.054) [0.0202]
TFP Quartile 4 x Program Intensity	0.244** (0.102) [0.0281]	-0.244*** (0.059) [0.000200]
Observations	7,256	7,256

Notes: Sample includes all cultivating households. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
Farmer TFP is calculated as demonstrated in Section **??**, and TFP quartiles are calculated within the district.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 6: *Dispersion of Marginal Product of Land*

	S.D.	C.V.	75 - 25
	(1)	(2)	(2)
Program Intensity	-0.134**	-0.013*	-0.110
	(0.065)	(0.007)	(0.072)
	[0.0112]	[0.0133]	[0.0525]
Observations	170	170	170

Notes: All regressions include district and year fixed effects, and district level linear trend. District-level controls include average education and rate of land ownership. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 7: *Program Effect on Farm Level Agricultural Production*

	Total Output	Output per acre	Profit per acre
	(1)	(2)	(3)
Program Intensity	91.215***	3.159	3.874
	(31.839)	(5.283)	(4.250)
	[0.00890]	[0.561]	[0.385]
Observations	7,256	7,256	7,256
Mean Dep., Pre-program	156.338	25.611	15.514

Notes: Output for each farm is calculated as the sum of the value of all crops grown on the farm. The value of each crop is calculated using the yield times median price for the crop across all farm households. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** p<0.01, ** p<0.05, * p<0.1.

Table 8: *Program Effect on Aggregate Agricultural Production using Remote Sensing NDVI*

	(1)	(2)
Post Program	0.093**	0.093**
	(0.040)	(0.041)
Observations	1,792	1,792
Linear Trend	District	Sub-district

Notes: Regressions are at subdistrict-year level, and the outcome is the NDVI, measured in number of standard deviations. All regressions include subdistrict and year fixed effects, in addition to the linear trends mentioned in the table. *PostProgram* is an indicator for all years after the program starts in any subdistrict. NDVI data are from (Didan, 2015). Standard errors clustered at the subdistrict level are presented in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.

Table 9: *Program Effect on Aggregate Agricultural Production using District-Level Data*

	Cereal Crops			Cash Crops		
	Log Area	Log Output	Log Yield	Log Area	Log Output	Log Yield
	(1)	(2)	(3)	(4)	(5)	(6)
Program Intensity	0.009	0.061*	0.053*	-0.023	-0.044	-0.021
	(0.024)	(0.031)	(0.028)	(0.093)	(0.125)	(0.042)
	[0.702]	[0.0572]	[0.0816]	[0.824]	[0.727]	[0.646]
Observations	792	792	792	455	455	455

Notes: Regressions are at district-crop-year level, and the outcomes are logged total area, total output, and yield for each crop in each district and year. Cereal crops include maize, rice and wheat, while cash crops include cotton and sugarcane. All regressions include district, crop and year fixed effects, and district specific linear trends. District-level controls include average education and rate of landownership.

Data are from the national Agricultural Statistics. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

A Details of Traditional Land Records

Land records are maintained through various statements; here, I describe the most commonly used and relevant to this study. The *Register Haqdaran Zamin* or Land Right Holders Register, is the document with which most citizens are primarily concerned. It lists the owners of the lands, including identifying details of the cultivator or tenant, soil and rent (Figure A1). Individual ownership documents, or *Fard Malkiyat/Fard*, can be prepared using the *Register Haqdaran Zamin*. The register of right holders corresponds to a *Mussavi* or a cadastral map of a village (see Figure A2). These maps were initially prepared by the British and specified each land parcel in a village with a unique parcel or (*Khasra*) number and dimensions (Hunter, Cotton, Burn, and Meyer, Hunter et al.). Any changes to land rights are recorded in a separate register of mutations *(Register Inteqalaat)*, which is used to update the register of right holders every four years. All of the above documents are held and maintained by the *patwari*. Additionally, the *patwari* maintains a *Khasra Girdawari*, which records the cultivator and crop information by *khasra* or land parcel.

Tenants' and landowners' rights, as well as updates that arise due to rental or sale, are thus recorded and updated using a combination of the above documents at the discretion of the *patwari*. Recording updates, or acquisition of the *Fard* or *Khasra Girdawari*, are the most common services offered to citizens through the land revenue records system. For instance, if a land-owner rents out land to a tenant, the *Khasra Girdawari* is updated for the relevant land parcel, and the *Register Haqdaran Zamin* is updated to reflect cultivating rights of the tenant.

With urbanization, land rights continue to be maintained by *patwaris* until an agency or urban development authority acquires the land. The urban land record system is similarly opaque. Overall there is no single agency maintaining updated urban land records for all of Punjab, and there is limited coordination in record keeping functions being carried out by the various agencies. The ambiguity of law regarding records of land rights is particularly harmful to the poor, who cannot afford protracted land disputes. The existing complexity of land rights in Pakistan, as well as the lack of information on the part of the citizens as well as authorities, and the discrepancies in the distribution of power in a rural context where land rights and power are connected Beg (2019), make it infeasible and costly to implement a universal land titling program. Numerous legal disputes result from limited enforcement of land rental contracts, e.g. illegal possession of land, eviction of tenants, and recovery of rent.[34]

B Theoretical Appendix

Farm production for farmer i, y_i, is taken to be a function of farmer level TFP, ω_i, randomly drawn from a uniform distribution over $[\underline{\omega}, \overline{\omega}]$. Each farmer has land endowment, n_i , which is independently drawn from a distribution such that $n_i \in [0, \infty]$ and the total land $\sum n_i = L$. Without loss of generality, the price of farm output is normalized to 1. With the operational land of each farmer given by, l_i, the farmer's production is as follows:

[34]Cases of land disputes are either handled by the Revenue Courts or Civil Courts, but cannot be resolved efficiently due to lack of decisive land rights records.

$$y_i = \omega_i l_i^{\alpha}, \quad \alpha < 1$$

I abstract from other inputs for focusing the theoretical discussion on the land market. In the construction of the TFP estimates, I account for other inputs reported by farmers for production.

Efficient Allocation: The efficient value of total output is obtained by choosing optimal land allocations to maximize the sum of individual productions:

$$Y^e = \max_{l_i} \sum \omega_i l_i^{\alpha}$$

$$\text{s.t. } \sum l_i = L$$

The efficient land allocation is given by setting the marginal product of land equal across farmers, which gives:

$$l_i^e = kz_i \text{ , where}$$

$$k = \left(\tfrac{\alpha}{r}\right)^{\frac{1}{1-\alpha}}$$

$$\text{and } z_i = \omega_i^{\frac{1}{1-\alpha}}$$

r is the implied marginal cost of land. Adding across the individual land allocations, gives:

$$l_i^e = \tfrac{z_i}{\sum z_i} L$$

Thus, regardless of endowment, the efficient land allocations are increasing in each farmer's individual productivity. These efficient allocations are identical to those derived in Restuccia and Santaeulalia-Llopis (2017) and Chari et al. (2017).

Actual Allocation with market frictions: I introduce transaction costs in the land market by assuming that farmers renting land in or out pay a proportional tax on agricultural land rent, i.e. rental cost is $(1+\mu)r$. Individuals maximize net income (farm production plus rental income). Thus, individual farmers will maximize total profit:

$$\max_{l_i} \omega_i l_i^{\alpha} - (l_i - n_i)(1+\mu)r$$

The first term is farm production, and the second term is rental cost if renting in land, or rental income if renting out. I assume, $\mu = \mu_2$ for renting in and $\mu = -\mu_1$ for renting out ($\mu_1, \mu_2 > 0$). The marginal product of land for farmer i with operational land l_i is given by $MPL_i = \frac{\alpha\omega_i}{l_i^{1-\alpha}}$

I show that the operational land with land market frictions differs from the efficient allocation for some farmers, depending on their land endowment and productivity. I make a simplifying assumption and consider low and high endowment cases, i.e. I assume $n_i \in \{\underline{n}, \overline{n}\}$. First consider farmers with $n = \underline{n}$, with ω uniformly distributed between $\underline{\omega}$ and $\overline{\omega}$ as for the entire population. Individuals will farm land until the marginal product of land equals its cost to them. As earlier, the operational farm size is greater for higher ωindividuals. There exists a cutoff such that if $\omega_i < a$, individuals will farm up to a point where marginal product of land equals the rental rate for renting out, and supply the rest of their land to the market. Thus, the total farmed land for these farmers is given by:

$$l_i = pz_i, \text{ where}$$

$$p = (\frac{1}{1-\mu_1})^{1/(1-\alpha)} k > k$$

Farmers with $a < \omega_i < b$ have marginal product higher than the rental rate but below r. Thus, these farmers would have rented out if market frictions did not exist, but do not rent out if $\mu_1 > 0$. Thus, for farmers with $\omega_i < b$, higher μ_1 results in lower rate of renting out and higher operational farm size relative to a scenario with no land market frictions.

Higher TFP farmers would choose to rent in. There exists another cutoff $c > b$, such that if TFP exceeds the cutoff, farmers will farm land until the marginal product equals the cost for renting in $(1 + \mu_2)r$. The total farmed land for these farmers is given by:

$$l_i = qz_i, \text{ where}$$

$$q = (\frac{1}{1+\mu_2})^{1/(1-\alpha)} k < k$$

These farmers rent in $(l_i - \underline{n})$ units of land; the operation land size and quantity of rented in land is lower than the scenario without market frictions, i.e. $\mu_2 = 0$. There will be some farmers who would have rented in but do not do so due to the transaction costs, i.e. farmers with $b < \omega_i < c$, operate all of their owned land.

Given this discussion, the actual land operated depends on ω as shown in the table below:

	ω	l_i	MPL	**Rental Market participation**	**Operational size vs. efficient allocation**
A	$\underline{\omega} < \omega < a$	$l_i = pz_i$	$MPL = (1 - \mu_1)r$	Rents out	$l_i > l_i^e$
B	$a < \omega < b$	$l_i = n_i$	$r > MPL > (1 - \mu_1)r$	None	$l_i > l_i^e$
C	$b < \omega < c$	$l_i = n_i$	$r < MPL < (1 + \mu_2)r$	None	$l_i < l_i^e$
D	$c < \omega < \overline{\omega}$	$l_i = qz_i$	$MPL = (1 + \mu_2)r$	Rents in	$l_i < l_i^e$

If $n = \overline{n}$, the farmers in category B would rent out more if $\mu_1 = 0$. Similarly, farmers in category C would rent in less if $\mu_2 = 0$.

Aggregate production under efficient and actual allocations:

This general form of the actual land allocation is given by:

$$l_i = \frac{\tau_i z_i}{\sum \tau_i z_i} L$$

where $\tau_i = (r/MPL_i)^{1/(1-\alpha)}$ demonstrates a wedge between the actual and efficient land allocation. Under no transaction cost scenario, $\tau_i = 1$. $\tau_i > 1$ for category A,B farmers and $\tau_i < 1$ for category C-D farmers. Under the actual allocation, the aggregate output is given by:

$$Y = \sum \omega_i l_i^\alpha = \sum \omega_i (s_i L)^\alpha = (\sum \omega_i s_i^\alpha) \cdot L^\alpha = \Omega L^\alpha$$

where s_i denotes each farmer's land allocation share, i.e. $\frac{\tau_i z_i}{\sum \tau_i z_i}$ and Ω is a measure of aggregate productivity. Removing transaction costs, lowers the land allocation share of farmers with $\omega < b$ and increases it for farmers with $\omega > b$. Thus, the overall productivity increases when the transaction costs are removed.

C Correlates of Farmer TFP

In Appendix Table A26, I regress the farmer-level TFP in the pre-program data on household characteristics, including household land ownership, education, and gender of head. I note that households with higher land ownership are less productive, likely because land ownership is the main determinant of selection into farming. Thus, households that own agricultural land operate a farm regardless of how productive they are in agriculture, particularly when barriers to renting out land are high. On the other hand, landless households that choose to farm are likely to do so if they have higher productivity in agriculture, implying that land ownership would be negatively correlated with individual TFP. Households with at least some educated members are more productive in farming, and female-headed households, as well as those with older heads are less productive.

D Appendix Tables and Figues

Figure A1: *A Land Record Register as maintained by Patwari (Adeel 2010)*

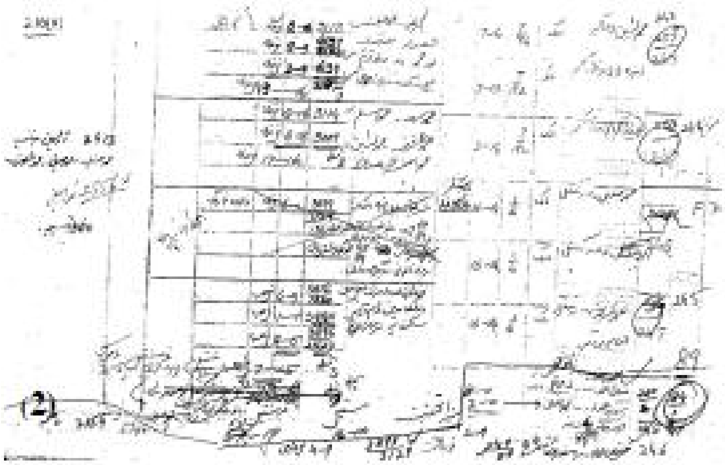


Figure A2: *A Cadastral Map for a village in Punjab (Adeel 2010)*

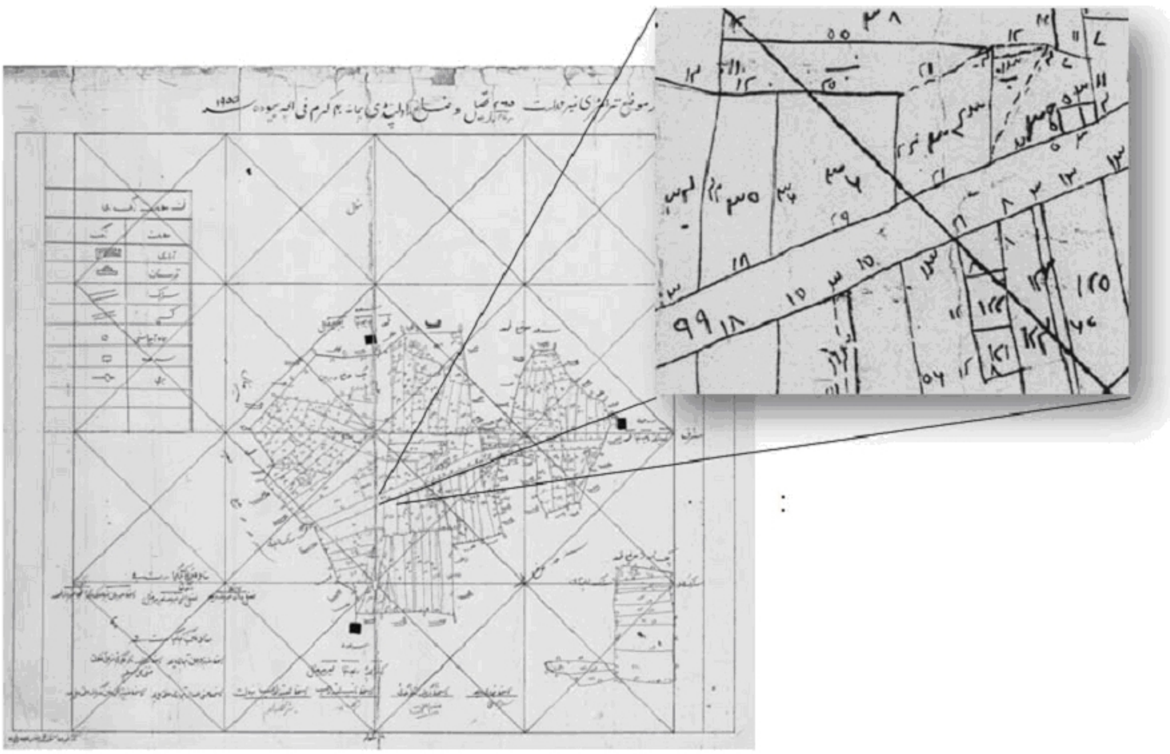
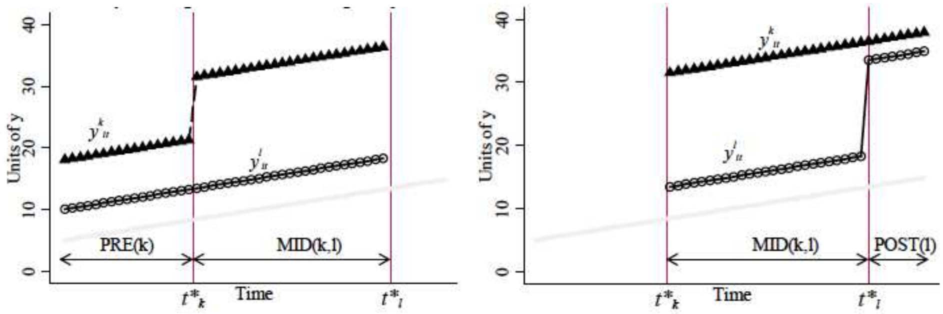


Figure A3: *Difference in difference estimators with two timing groups (Goodman-Bacon, 2019)*



Notes: The figure plots the groups and time periods that generate the difference-in-difference estimates in the case with an early treatment group and a late treatment group. The left panel compares early treated units to late treated units during the late group's pre-period and the right panel compares late treated units to early treated units during the early group's post-period.

Figure A4: *GDP per Capita by Year*

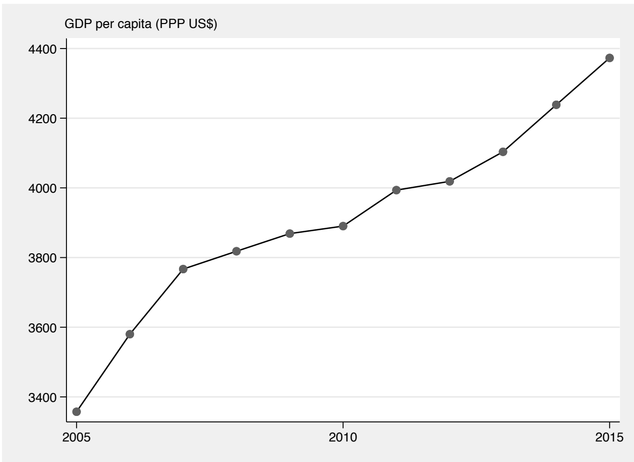


Table A1: *Summary of Program Progress*

	mean	sd
2011		
% of Subdistricts with a Center	0.042	0.201
2012		
% of Subdistricts with a Center	0.147	0.355
2013		
% of Subdistricts with a Center	0.608	0.490
2014		
% of Subdistricts with a Center	0.972	0.165

Table A2: *Balance Tests for Program Start and Progress*

A. Balance Test for Program Start			.	
	Prior Level		Prior Change	
	F-stat	p-value	F-stat	p-value
Outcome	(1)	(2)	(3)	(4)
Landowners Renting out	0.476	0.674	0.303	0.802
Landowners' Ag. participation	0.206	0.861	1.682	0.227
Farmsize (acres)	0.569	0.623	1.791	0.178
Acres Rented in	0.885	0.447	0.295	0.812
Population (mm)	0.431	0.698	0.259	0.850
Unemployment	0.381	0.759	1.518	0.198
NDVI	1.143	0.611	0.429	0.844
Soil Quality Index	2.468	0.101		
B. Balance Test for Program Progress after start				
	Prior Level		Prior Change	
Outcome	Coefficient	p-value	Coefficient	p-value
Landowners Renting out	0.019	0.779	0.025	0.896
Landowners' Ag. participation	-0.056	0.416	-0.735	0.262
Farmsize (acres)	0.019	0.830	-0.019	0.944
Acres Rented in	-0.081	0.411	-0.105	0.088
Population	0.000	0.848	0.006	0.980
Unemployment	0.147	0.032	0.037	0.805
F-stat of joint significance		**1.305**		**0.407**
p-value		**0.126**		**0.785**

Notes: Panel A shows the results from a regression of each outcome in the pre-program data rounds (level and change) on fixed effects for start year of the program. F-statistic and p-value of a joint test of significance of start year fixed effects are provided. Panel B shows the results from a regression of *Program_Intensity* on the lagged level and change for each outcome using data from the rounds after the program has begun. The F-statistic and p-value from joint test of significance of lagged levels and changes of all outcomes in reported in Panel B. Data are from the HIES surveys. NDVI data are from Didan (2015) and soil quality data are from the HWSD. Standard errors are clustered at district level and adjusted for wild-cluster bootstrapping in all regressions.

Table A3: *Sample Changes in Response to the Program*

	Rural	Landowning	Cultivating
	(1)	(2)	(3)
Program Intensity	0.024	0.002	-0.038
	(0.056)	(0.030)	(0.030)
	[0.696]	[0.954]	[0.224]
Observations	33,703	19,067	19,067
Mean Dep., Pre-program	0.686	0.420	0.395
Households	All HHs	Rural HHs	Rural HHs

Notes: Outcomes indicate if a household is rural, landowning, or cultivating. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** p<0.01, ** p<0.05, * p<0.1.

Table A4: *Access to Titles after Program*

	Land Title (1)
Post	0.016** (0.008) [0.0438]
Observations	2,322
Mean Title in 2016	0.097

Notes: Sample includes landowners without a formal title in 2006. Regressions are at farmer-year level and include farmer and year fixed effects. Data are from the farmer phone surveys. Standard errors clustered at the subdistrict level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** p<0.01, ** p<0.05, * p<0.1.

Table A5: *Program Effect on Non-Agricultural Participation*

	Small Business Owners	Large Business Owners	Self-Employed	Paid Employees
	(1)	(2)	(3)	(4)
Program Intensity	0.139	0.127*	0.438	1.117
	(0.124)	(0.066)	(0.461)	(0.830)
	[0.288]	[0.0765]	[0.365]	[0.212]
Observations	7,597	7,597	7,597	7,597
Mean Dep., Pre-program	0.082	0.021	2.160	5.791

Notes: Outcomes indicate if any member of the household participates in the specific activity. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A6: *Program Effect on Credit, Migration and Overall Income*

	Any Loan (Y/N)	Any Remittance (Y/N)	Log Total Income
	(1)	(2)	(3)
Program Intensity	0.047	0.006	0.217***
	(0.036)	(0.032)	(0.073)
	[0.216]	[0.862]	[0.00960]
Observations	7,597	7,593	7,223
Mean Dep., Pre-program	0.313	0.237	

Notes: The outcomes in columns (1) and (2), respectively, indicate if a household has any outstanding loan or any remittance income. Column (3) shows the natural log of total household income. Sample includes landowning households. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A7: *Program Effect on Market Activity by Income Quartile*

	Own Agland (Y/N) (1)	Agland Purch. (Y/N) (2)	Agland Sold (Y/N) (3)	Own Agland (Acres) (4)	Agland Rentout (Y/N) (5)
Program Intensity	0.003	0.001	-0.005	1.411*	0.086**
	(0.034)	(0.004)	(0.009)	(0.818)	(0.042)
	[0.933]	[0.731]	[0.568]	[0.0956]	[0.0566]
Program Intensity x Inc Quartile 2	-0.006	0.000	0.011	-0.693	0.009
	(0.022)	(0.001)	(0.009)	(0.706)	(0.037)
	[0.805]	[0.766]	[0.244]	[0.333]	[0.804]
Program Intensity x Inc Quartile 3	0.013	-0.002	0.006	-0.517	-0.016
	(0.026)	(0.002)	(0.011)	(0.759)	(0.038)
	[0.617]	[0.509]	[0.544]	[0.493]	[0.669]
Program Intensity x Inc Quartile 4	0.010	-0.001	-0.003	-1.389*	-0.040
	(0.024)	(0.003)	(0.010)	(0.759)	(0.041)
	[0.699]	[0.848]	[0.803]	[0.0773]	[0.342]
Observations	19,059	7,584	7,584	7,579	7,597
Mean Dep., Pre-program	0.997	0.006	0.009	6.359	0.219
Sample Households	All Rural	All Landowning	All Landowning	All Landowning	All Landowning

Notes: Sample includes landowning households. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A8: *Program Effect on Agricultural Participation by Income Quartile*

	HH Operates Any Farm	HH Operates Owned Land	HH Member Ag Worker	Share Income from Ag
	(1)	(2)	(3)	(4)
Program Intensity	-0.115**	-0.051	-0.153***	-0.127***
	(0.045)	(0.056)	(0.045)	(0.045)
	[0.0176]	[0.383]	[0.00250]	[0.0131]
Program Intensity x Inc Quartile 2	-0.017	-0.058	0.020	0.033
	(0.035)	(0.041)	(0.037)	(0.047)
	[0.621]	[0.169]	[0.598]	[0.500]
Program Intensity x Inc Quartile 3	0.013	-0.030	0.100**	0.084*
	(0.037)	(0.049)	(0.041)	(0.049)
	[0.726]	[0.539]	[0.0260]	[0.101]
Program Intensity x Inc Quartile 4	0.026	-0.021	0.051	0.042
	(0.047)	(0.042)	(0.037)	(0.039)
	[0.596]	[0.625]	[0.182]	[0.282]
Observations	7,597	7,597	7,597	7,597
Mean Dep., Pre-program	0.786	0.620	0.807	0.650

Notes: Sample includes landowning households. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A9: *Program Effect on Land Rental for Cultivators*

	Rent in	Fixed Rent	Sharecrop
	(1)	(2)	(3)
Program Intensity	0.019	0.030	-0.007
	(0.043)	(0.038)	(0.020)
	[0.657]	[0.446]	[0.724]
Observations	7,256	7,256	7,256
Mean Dep., Pre-program	0.334	0.254	0.082

Notes: All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A10: *Program Effect on Farm Size and Rented in Land (with added controls)*

	Rented	Sharecropped	Owned	Total Cultivated
	(1)	(2)	(3)	(4)
Program Intensity	0.887**	0.065	0.802	1.116**
	(0.435)	(0.240)	(0.667)	(0.451)
	[0.0469]	[0.839]	[0.243]	[0.0140]
Observations	7,256	7,256	7,256	7,256
Mean Dep., Pre-program	1.648	0.686	5.423	7.055

Notes: Rent area corresponds to area under fixed cash rent contracts and S/C refers to area under sharecropping contracts. Farm size is total operational farm area including owned land. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include indicator for landownership, head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A11: *Program Effect on Land Rental for Landless Cultivators*

	Rent in (Y/N)	Rented	Sharecropped	Total Cultivated
	(1)	(2)	(3)	(4)
Program Intensity	0.005	0.775*	0.106	0.988**
	(0.038)	(0.431)	(0.234)	(0.452)
	[0.885]	[0.0847]	[0.724]	[0.0303]
Program Intensity x Landless	0.028	0.935	-0.343	1.073
	(0.021)	(0.686)	(0.211)	(0.781)
	[0.196]	[0.185]	[0.0965]	[0.175]
Observations	7,256	7,256	7,256	7,256
Mean Dep., Pre-program	0.334	1.648	0.686	7.055
p-value of sum	0.335	0.078	0.238	0.031

Notes: All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** p<0.01, ** p<0.05, * p<0.1.

Table A12: *Dispersion of TFP*

	S.D.	75 - 25
	(1)	(2)
Program Intensity	-0.102	-0.169
	(0.073)	(0.125)
	[0.0897]	[0.0977]
Observations	170	170

Notes: All regressions include district and year fixed effects, and district level linear trend. District-level controls include average education and rate of land ownership. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** p<0.01, ** p<0.05, * p<0.1.

Table A13: *Program Effect on Allocation across Farmers: Alternate Measures of Farmer Productivity*

	Land (1)	MPL (2)	Land (3)	MPL (4)
Program Intensity	0.148	0.185*	0.088	0.159
	(0.100)	(0.101)	(0.107)	(0.104)
	[0.156]	[0.0706]	[0.422]	[0.132]
Quartile 2 x Program Intensity	-0.014	-0.096**	0.101	-0.054
	(0.082)	(0.038)	(0.091)	(0.041)
	[0.864]	[0.0122]	[0.288]	[0.184]
Quartile 3 x Program Intensity	0.058	-0.149**	0.067	-0.158***
	(0.079)	(0.059)	(0.094)	(0.056)
	[0.470]	[0.0127]	[0.483]	[0.00670]
Quartile 4 x Program Intensity	0.024	-0.302***	0.128	-0.250***
	(0.106)	(0.057)	(0.112)	(0.060)
	[0.820]	[0]	[0.260]	[0.000100]
Observations	7,256	7,256	7,256	7,256
p-value for Q3	0.021	0.639	0.077	0.991
p-value for Q4	0.069	0.190	0.024	0.293
Productivity Measure	Yield	Yield	Profit	Profit

Notes: Sample includes all cultivating households. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A14: *Program Effect on Crop Choice*

	Cereal Crops			Cash Crops	
	Wheat	Rice	Maize	Cotton	Sugarcane
	(1)	(2)	(3)	(4)	(5)
Program Intensity	0.000	0.080*	-0.059*	0.044	0.006
	(0.026)	(0.041)	(0.031)	(0.045)	(0.029)
	[0.997]	[0.0766]	[0.0813]	[0.352]	[0.844]
Observations	7,256	7,256	7,256	7,256	7,256
Mean Dep., Pre-program	0.914	0.308	0.060	0.343	0.166

Notes: All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** p<0.01, ** p<0.05, * p<0.1.

Table A15: *Program Effect on Agricultural Inputs*

	Use Fertilizer	Use Pesticide	Use Hired Labor	Use Rented Equipment
	(1)	(2)	(3)	(4)
Program Intensity	0.003	0.052*	0.057	-0.024
	(0.018)	(0.031)	(0.049)	(0.021)
	[0.876]	[0.103]	[0.252]	[0.269]
Observations	7,256	7,256	7,256	7,256
Mean Dep., Pre-program	0.944	0.760	0.481	0.924

Notes: All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** p<0.01, ** p<0.05, * p<0.1.

Table A16: *Program Effect on Agricultural Equipment*

	(1)	(2)
Program Intensity	-0.019	-0.022
	(0.023)	(0.027)
	[0.430]	[0.409]
Program Intensity x Landless		0.042**
		(0.020)
		[0.0401]
Observations	6,789	6,789

Notes: Outcome variable is an indicator if a cultivator has acquired agricultural equipment in the last year (Tube well, Tractor, Plough, Thresher, Harvester, or Truck). All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A17: *Program Effect on Perceived Selling Rights*

	Ag. Land (knows selling price) (1)	Ag. Land (acres) (2)	Non Ag. Land (knows selling price)
Program Intensity	0.084**	0.683	-0.002
	(0.036)	(0.557)	(0.001)
	[0.0230]	[0.245]	[0.0883]
Observations	7,597	7,579	7,465
Mean Dep., Pre-program	0.225	6.359	0.999

Notes: Regressions are on the landowning sample. Column (3) includes landowners who also own residential, commercial or other non-agricultural property. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A18: *Program Effect on Using Data from Additional Years*

	Own Ag. Land (Y/N)	Owned Ag. Land Size (acres)	Rent out (Y/N)	Any Ag. Work	Share Ag. Income
	(1)	(2)	(3)	(4)	(5)
Program Intensity	0.000	0.478	0.042**	-0.070***	-0.048*
	(0.018)	(0.329)	(0.020)	(0.025)	(0.025)
	[0.999]	[0.157]	[0.0438]	[0.00890]	[0.0791]
Observations	123,574	53,984	54,007	54,007	54,007
Mean Dep., Pre-program	0.445	6.290	0.186	0.765	0.636
Households	All Rural	Landowning	Landowning	Landowning	Landowning

Notes: Regressions use outcome variables from HIES and PSLM data. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A19: *Robustness of main effects with Alternate Specifications*

	HH Rents Out (1)	HH Cultivates Farm (2)	Share Ag Income (3)	Rented in Area (4)	Total Cult. Area (5)	Total Output (6)
Panel A: No Household Level Controls						
Program Intensity	0.061*	-0.095**	-0.091**	0.921**	1.201**	94.622***
	(0.031)	(0.038)	(0.036)	(0.429)	(0.470)	(33.368)
	[0.0621]	[0.0164]	[0.0228]	[0.0328]	[0.00920]	[0.00760]
Panel B: District Specific Quadratic Trend						
Program Intensity	0.059*	-0.117***	-0.098***	1.032**	1.080**	86.414***
	(0.029)	(0.037)	(0.035)	(0.413)	(0.409)	(27.670)
	[0.0725]	[0.00130]	[0.0148]	[0.0108]	[0.0100]	[0.00400]
Panel C: Macroeconomic Controls						
Program Intensity	0.064**	-0.096***	-0.069**	0.894*	1.156**	94.217***
	(0.028)	(0.029)	(0.033)	(0.448)	(0.460)	(31.947)
	[0.0316]	[0.000700]	[0.0547]	[0.0537]	[0.0148]	[0.00770]
Panel D: Pre- & Post-Recession Trend						
Program Intensity	0.055*	-0.092***	-0.081**	1.037**	1.041***	95.523***
	(0.028)	(0.032)	(0.037)	(0.437)	(0.374)	(29.142)
	[0.0636]	[0.00350]	[0.0446]	[0.0178]	[0.00660]	[0.00280]
Observations	7,597	7,597	7,597	7,256	7,256	7,256
Sample HHs	Landowning	Landowning	Landowning	Cultivating	Cultivating	Cultivating

Notes: All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender (Panels B-D). Macroeconomic controls include unemployment and size of labor force at district level. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A20: *Placebo Test with Urban Households*

	Log Total Income
	(1)
Program Intensity	-0.008
	(0.078)
	[0.973]
Observations	13,700
Sample	Urban HHs

Notes: All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Data are from the HIES surveys. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** p<0.01, ** p<0.05, * p<0.1.

Table A21: *Robustness of Main Effects: Dropping Early Districts*

	HH Rents Out	HH Cultivates Farm	Share Ag Income	Rented in Area	Total Cult. Area	Total Output
	(1)	(2)	(3)	(4)	(5)	(6)
Program Intensity	0.063**	-0.108***	-0.085**	0.841*	0.991**	74.702**
	(0.029)	(0.032)	(0.036)	(0.460)	(0.385)	(34.352)
	[0.0478]	[0.00150]	[0.0334]	[0.0711]	[0.0186]	[0.0461]
Observations	7,084	7,084	7,084	6,758	6,758	6,758
Sample HHs	Landowning	Landowning	Landowning	Cultivating	Cultivating	Cultivating

Notes: All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A22: *Robustness Checks Using Placebo Program Rollout*

	HH Rents Out	HH Cultivates Farm	Share Ag Income	Rented in Area	Total Cult. Area	Total Output
	(1)	(2)	(3)	(4)	(5)	(6)
Placebo Program Intensity	0.004	0.022	0.071	-0.240	1.417	64.948
	(0.043)	(0.061)	(0.049)	(0.522)	(1.087)	(51.418)
	[0.923]	[0.732]	[0.225]	[0.660]	[0.260]	[0.308]
Observations	7,597	7,597	7,597	7,256	7,256	7,256
Sample Households	Landowning	Landowning	Landowning	Cultivating	Cultivating	Cultivating

Notes: The placebo program intensity assume the program rollout began 2 survey rounds before the actual program start date in a district. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A23: *Robustness of main effects with Alternate Identification Strategy: Standard Timing D-in-D*

	HH Rents Out	HH Cultivates Farm	Share Ag Income	Rented in Area	Total Cult. Area	Total Output
	(1)	(2)	(3)	(4)	(5)	(6)
Post Program	0.052**	-0.106***	-0.094***	0.752**	0.562	49.533*
	(0.022)	(0.026)	(0.028)	(0.345)	(0.385)	(24.843)
	[0.0708]	[0.000100]	[0.00450]	[0.0276]	[0.161]	[0.0623]
Observations	7,597	7,597	7,597	7,256	7,256	7,256
Sample Households	Landowning	Landowning	Landowning	Cultivating	Cultivating	Cultivating

Notes: Post Program indicates years where at least one subdistrict in a district has the program. All regressions include district and year fixed effects, and controls for linear district-level yearly trends. Additional household controls include head age, age squared, education and gender. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** p<0.01, ** p<0.05, * p<0.1.

Table A24: *Robustness of main effects with Alternate Identification Strategy: Stacked D-in-D*

	HH Rents Out	HH Cultivates Farm	Share Ag Income	Rented in Area	Total Cult. Area	Total Output
	(1)	(2)	(3)	(4)	(5)	(6)
Treated x Post	0.044**	-0.088***	-0.070**	0.481**	0.363	24.323
	(0.017)	(0.021)	(0.026)	(0.221)	(0.417)	(21.470)
	[0.0144]	[0]	[0.0176]	[0.0375]	[0.429]	[0.291]
Observations	185,889	185,889	185,889	182,263	182,263	182,263
Sample HHs	Landowning	Landowning	Landowning	Cultivating	Cultivating	Cultivating

Notes: Post Program indicates years where at least one subdistrict in a district has the program. All regressions include district and year fixed effects. Additional household controls include head age, age squared, education and gender. Standard errors clustered at the district level in parentheses. Wild cluster bootstrapped p-values in brackets.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A25: *Multiple Hypothesis Testing Correction*

	(1) Effect size	(2) Original pval	(3) FDR adj. q-val
Outcomes for Cultivating Households:			
Total Area Cultivated	1.110**	0.0183	0.055
Area Rented	0.925**	0.0354	0.067
Area Sharecropped	0.731	0.319	0.532
Area Owned	0.084	0.805	0.877
Total Output	90.439***	0.0144	0.054
Profit per acre	3.906	0.368	0.552
Output per acre	3.216	0.547	0.746
Outcomes for Landowning Households:			
Own agland	0.005	0.877	0.877
Household Sold Ag Land	-0.002	0.716	0.877
Household Purchased Ag Land	0.001	0.825	0.877
Rentout out agland	0.061**	0.0343	0.067
HH Cultivates a Farm	-0.098***	0.0015	0.023
HH Cultivates Owned Land	-0.089***	0.0073	0.044
HH in Ag. Work	-0.099***	0.0087	0.044
Share of Ag Income	-0.080**	0.028	0.067

Notes: Adjusted sharpened q-values calculated for all outcomes using the method suggested by Anderson (2008) based on Benjamini and Hochberg (1995).
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A26: *Correlates of TFP*

	(1)	(2)	(3)	(4)	(5)	(6)
Own Agland (Y/N)	-0.338***	-0.400***				-0.400***
	(0.026)	(0.026)				(0.026)
Own Agland (Acres)		0.010***				0.009***
		(0.001)				(0.001)
Head Schooling (Y/N)			0.101***			0.114***
			(0.022)			(0.022)
Female Head				-0.165***		-0.149***
				(0.047)		(0.045)
Age of Head					-0.009**	-0.006
					(0.004)	(0.004)
Observations	4,859	4,855	4,859	4,859	4,859	4,855

Notes: Sample includes cultivating households. All regressions include district and year fixed effects, controls for district level share of rural population and share of landowning households who participate in agriculture, and linear district-level yearly trend. Additional household controls include land ownership, head age, and education and household size. Standard errors clustered at the district-year level in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.